# The mechanisms of cell invasion in cnidarian-dinoflagellate symbiosis: learning from parasitic strategies

By

Emilie Fleur Neubauer

A thesis submitted to the Victoria University of Wellington in fulfilment of the requirements for the degree of Doctor of Philosophy in Science

Victoria University of Wellington

2016

# Abstract

Cnidarians such as corals, anemones and hydroids commonly form an intracellular symbiosis with photosynthetic dinoflagellates of the genus *Symbiodinium*. Dinoflagellate symbionts are most often obtained anew from the environment during larval development, and, once acquired reside inside host-derived vacuoles within the cnidarian gastrodermal cells. In order to gain entry to host cells, the symbionts likely interact with innate immune receptors in the extracellular matrix, the first line of defense against microbial attack. While several innate immune pathways have been described in cnidarians, little is known about the specific receptor-ligand interactions that allow the symbiont to gain entry to host cells. Furthermore, it is unclear how these pathways are involved in enabling friendly microbes to reside within host cells while maintaining an immune response to harmful pathogens.

The invasion strategies of vertebrate intracellular parasites are well studied, especially those used by members of the Apicomplexa. Apicomplexan parasites have evolved mechanisms to evade immune receptors in the extracellular matrix and exploit specific receptors to their own benefit, to gain entry to host cells. Apicomplexans are closely related to dinoflagellates, both belonging to the infrakingdom Alveolata. The malaria parasite *Plasmodium* spp. has evolved the thromobospondin-related anonymous protein, or TRAP, that uses a thrombospondin structural homology repeat (TSR) domain to bind to a scavenger receptor (SRB1) on the hepatocyte cell surface and gain entry to the cell. This is of particular interest, as class B scavenger receptors are upregulated in the symbiotic state of two anemone species. The aims of the research presented in this thesis were to: (1) characterize the scavenger receptor (SR) repertoire in cnidarians; (2) characterize the TSR-domain-containing protein repertoire of cnidarians and their symbiotic dinoflagellates; and (3) establish, through experimental manipulation, the potential role for SR-TSR domain interactions at the onset of symbiosis in the sea anemone *Aiptasia* sp, a model system for the study of the cnidarian-dinoflagellate symbiosis.

In Chapter 2, I characterized the large and diverse SR repertoire of six cnidarian species. Cnidarians lack the classic SR type-A collagen domain-containing proteins that are common in humans, however the cnidarian SR cysteine-rich domain-containing protein

repertoire is expanded and diverse. Phylogenetic analysis of SR type-B proteins defines two or three distinct groups. Functional experimental data presented here show that blocking SR binding sites with fucoidan significantly reduces dinoflagellate uptake by the anemone *Aiptasia* sp. These data provide further evidence that SRs are important to symbiont recognition and uptake, and may be an essential component of symbiont acquisition.

In Chapter 3, I investigated a SR ligand, the thrombospondin structural homology repeat, or TSR domain. In particular, I characterized the TSR-domain-containing protein repertoire of six cnidarian species and compared these proteins to vertebrate TSR proteins of known function. Searches revealed a large repertoire of TSR-domain-containing proteins. Of particular interest is the large number of Adams metalloprotease-like proteins, a group that is common in both humans and cnidarians, suggesting that this is an ancestral TSR protein group. Phylogenetic analysis of TSR domains shows that binding motifs and 3-D folding sites are highly conserved. These data suggest that TSR domains are ancient and have changed very little in amino acid sequence from lower metazoans to vertebrates.

In Chapter 4, I explored the role of TSR-domain-containing proteins at the onset of symbiosis in the model *Aiptasia* sp. system. In functional experiments, aposymbiotic anemones were challenged with proteins and antibodies to either block or stimulate TSR domain binding. Symbiont uptake was measured over several time-points to determine the effects on symbiont acquisition. Adding an excess of TSR-domain-containing protein or TSR synthetic peptide increased symbiont uptake, while blocking TSR domains prevented symbiont uptake. Finally, the addition of exogenous TGFβ to TSR antibody-challenged anemones, reversed the blocking effect. These data suggest that the immune-suppressive TGFβ pathway is involved in early onset of the symbiosis. Since the TSR domain is implicated in the TGFβ pathway, these results support previous findings of the involvement of TGFβ in promoting tolerance of symbionts within the host. Apicomplexan parasites exploit scavenger receptor-TSR domain-binding to gain entry, and also use immune modulation to persist inside host cells. Data presented here suggest that dinoflagellates are utilizing the same mechanisms to form a mutualistic relationship with the cnidarian host.

Overall, the work presented here provides new information about several cnidarian extracellular matrix proteins, with searches revealing large repertoires of both

scavenger receptors and TSR-domain-containing proteins. Functional data suggest that both protein families are involved in the cnidarian-dinoflagellate symbiosis. Searches of the dinoflagellate genome did not find a clear dinoflagellate homologue to the apicomplexan TRAP proteins. However, this research provides further evidence that similar receptor-ligand interactions are involved in the entry of both beneficial and pathogenic microbes to host cells. These results add to growing knowledge about the complex molecular pathways that enable and support cnidarian-dinoflagellate symbiosis. An understanding of the mechanisms that support healthy symbiosis is essential when trying to predict the vitality and productivity of reef ecosystems in the face of climate change.

# Acknowledgements

‘Sky above me, earth below me, fire within me’

Producing this thesis has been an epic journey, over three continents, two premature babies, and major surgery. Along the way, I have had the privilege of meeting and working with some incredibly wonderful scientists who shared and sparked the fire that has enabled me to keep going despite more than a few bumps in the road.

Thank you to my advisors:

Simon Davy at VUW, for being a calm place in the midst of the storm, despite all that has happened you have always remained constant and supportive, unflappable and open to deal with administrative issues. Thank you for enabling me, and for your editorial and grammatical efforts to shape this work.

Virginia Weis at Oregon State University, for inviting me to work in her lab, supporting my research investing both funds and time. For being an inspiring mentor, teacher and for keeping the faith, many a positive comment from you has kept me going especially this last year – thank you.

I would like to thank my Oregon family:

Weis lab members: Mauricio Rodriguez-Lanetty for inspiring the quest for SRB1; Santiago Perez for protocols and advice; Olivier Detournay for advice, inspiration, teaching, friendship and for being so much fun to work with; Christy Schnitzler for protocols, teaching, advice and friendship; Elisha Wood-Charleson for help with *Nematostella* data; Wendy Philips for teaching, friendship, support in thousand ways – you always know what, where or how, and for introducing me to the Wailing Jennies; Jessi Kershner for protocols, friendship and road-trip adventures; Jen Osbourne, Daryl Khaw, Brian Henrichs and Gabe Carrick for being awesome students and so much fun in and out of the lab; finally Team *Aiptasia* and Team Gym founding member Angela Poole, the best office and lab mate anyone could wish for, thank you for making such long hours so much fun.

Thank you to my Wellington family: Anne La Flamme for early advice and support; Mary Murray and Sandra Taylor for general support and watching Theo, Patricia Stein for endless administrative and emotional support; Shona De Sain for constant support and encouragement; Lizzie Towl for the thesis boot camp; Dan Logan for advice, microscopy knowhow and friendship; Tom Hawkins for enthusiasm; Dorota Starzak for pilates;

Joanne Davy for support and friendship and raw chocolate fudge making skills and Nishanie Pereira for constant text support, cheering me on, and being a thesis tigress!
Thank you to Team Umpqua members far and wide for believing in me, keeping me sane, feeding me and for so much fun – you are loved: Patricia Aron, Nia Haf Jones, Samantha Bryan, Stacie Mills Webber; Joshua Webber; and Anna Brantsch.
To Yoga for keeping me alive, sane and pain free
Finally thank you to Philipp Neubauer, for everything and so much more and to Theo James Neubauer and Jonah Kale Neubauer for being the most beautiful distractions – you are my everything.

# Table of Contents

**Abstract III**

**Acknowledgements VI**

**Chapter 1 General Introduction 19**

1.1 Symbiosis 19
1.2 Cnidarian-dinoflagellate symbiosis 22
1.3 Establishment of cnidarian-dinoflagellate symbiosis 26
*1.3.1 Symbiont recognition and uptake via phagocytosis 26*
*1.3.2 Selection (specificity) 28*
*1.3.3 Arrest of phago-lysosome maturation and persistence within host cells 29*
1.4 Innate immunity in the context of cnidarian algal-symbiosis 30
*1.4.1 Complement pathway PRRs 33*
*1.4.2 C-type lectins 33*
*1.4.3 Scavenger receptors 34*
*1.4.4 Toll-like receptors 36*
*1.4.5 TSR and TRAP proteins 37*
1.5 Parasitic behaviour – links to the Apicomplexa 38
1.6 *Aiptasia* sp. as a model organism 44
1.7 Available sequence resources for cnidarians 46
1.8 Aims and scope of this study 48

**Chapter 2 The scavenger receptor repertoire in six cnidarian species indicates a potential role in the establishment of cnidarian-dinoflagellate symbiosis 50**

2.1 Introduction 50
2.2 Materials and Methods 55
*2.2.1 Cnidarian genomic and transcriptomic resources 55*
*2.2.2 SR sequence searching and verification 55*
*2.2.3 Maintenance and preparation of anemone and dinoflagellate cultures 57*
*2.2.4 Addition of fucoidan to block SR binding function 57*
*2.2.5 Assessing infection success using confocal microscopy 59*
2.3 Results 60

*2.3.1 Searches of six anthozoan species reveal an expanded repertoire of SRCR-domain-containing proteins in cnidarians ..... 60*
*2.3.2 Phylogenetic analysis of Class B CD36-domain containing proteins ..... 63*
*2.3.3 Experimental blocking of SR proteins in the symbiotic anemone Aiptasia sp. ..... 66*
2.4 Discussion ..... 68
*2.4.1 The SRCR-domain-containing protein repertoire in cnidarians is expanded compared to vertebrate ..... 69*
*2.4.2 Putative CTLD-domain-containing SRs in cnidarians also contain the SRCR domain ..... 70*
*2.4.3 CD36-domain-containing proteins in cnidarians ..... 71*
*2.4.4 Functional experiments suggest that blocking SRs decreases symbiont acquisition and increases the stress response to immune challenge in the anemone Aiptasia sp. ..... 72*
2.5 Concluding remarks ..... 73

**Chapter 3 The large and diverse TSR-domain-containing protein repertoire of six cnidarian species contains highly conserved binding sites and motifs ..... 74**

3.1 Introduction ..... 74
3.2 Materials and methods ..... 79
*3.2.1 Cnidarian genomic and transcriptomic resources ..... 79*
*3.2.2 TSR sequence searching, verification and phylogenetic analysis ..... 80*
3.3 Results ..... 82
*3.3.1 TSR-domain-containing proteins in Symbiodinium minutum ..... 82*
*3.3.2 Analysis of potential binding sites and conserved motifs in cnidarian TSR domains ..... 84*
*3.3.3 Cnidarian TSR-domain-containing proteins ..... 85*
*3.3.4 Phylogenetic analysis of TSR domains reveals high conservation between vertebrates and invertebrates. ..... 86*
3.4 Discussion ..... 90
*3.4.1 Dinoflagellate TRAP-like proteins ..... 90*
*3.4.2 What can be inferred from known functions of vertebrate TSR domains? ..... 90*
*3.4.3 Which TSR proteins are present in cnidarians? ..... 91*

*3.4.4 Data constraints on protein analysis: data source, developmental stage and symbiotic state* ..... *93*
*3.4.5 Phylogenic analysis of TSR domains reveals high conservation* ..... *94*
3.5 Concluding remarks ..... 95

***Chapter 4*** **The thrombospondin structural homology repeat (TSR) domain is important in successful dinoflagellate infection of the anemone *Aiptasia* sp. ..... 96**

4.1 Introduction ..... 96
4.2 Materials and Methods ..... 101
*4.2.1 Maintenance and preparation of anemone and dinoflagellate cultures* ..... *101*
*4.2.2 Western blot analysis of TSP1 antibody protein targets* ..... *101*
*4.2.3 Cryosectioning and immunofluorescence microscopy to localise binding of the human TSP antibody* ..... *102*
*4.2.4 Experimental manipulation of anemones* ..... *104*
*4.2.5 Addition of anti-TSP to aposymbiotic anemones during the onset of symbiosis* ..... *104*
*4.2.6 Addition of TGFβ protein to anemones already treated with anti-TSP* ..... *105*
*4.2.7 Addition of human TSP1 protein and synthetic TSR peptides* ..... *105*
*4.2.8 Assessing infection success using confocal microscopy* ..... *106*
4.3 Results ..... 107
*4.3.1 Evidence of TSR domain proteins in host but not symbiont* ..... *107*
*Blocking TSR domains inhibits symbiont uptake* ..... *110*
*4.3.2 110*
*4.3.3 Blocking effect of TSP antibodies can be partially reversed by the addition of TGFβ protein* ..... *111*
*4.3.4 An excess of human Thrombospondin-1 induces 'super' infection* ..... *112*
*4.3.5 Addition of TSR peptide fragments also increases infection success at the onset of infection* ..... *113*
*4.3.6 TSR-domain-containing proteins involved in cnidarian-dinoflagellate symbiosis are likely of host origin* ..... *115*
*4.3.7 Re-infection experiments implicate the TSR domain in symbiont uptake* ..... *117*
4.4 Concluding remarks ..... 119

**Chapter 5 General Discussion ..... 121**

5.1 Summary of research findings ..... 121
5.2 How are scavenger receptors implicated at initial contact, host-symbiont signaling and host phagocytosis? ..... 123
5.3 What happens post phagocytosis? How does the symbiont modulate the host immune response and evade digestion? ..... 127
5.4 Future directions ..... 130

**Bibliography ..... 132**

**Appendices**

Appendix A1 ..... 158
Appendix A2 ..... 159
Appendix A3 ..... 168
Appendix A4 ..... 169
Appendix B1 ..... 175
Appendix B2 ..... 176
Appendix B3 ..... 184
Appendix C1: ..... 185

# Table of Figures

Figure 1.1: Simplified schematic diagram of a symbiotic anthozoan (phylum Cnidaria).....23

Figure 1.2: Cnidarians as a basal group in metazoan evolution. Within the Cnidaria, the Anthozoa are basal and the Hydrozoa are derived.....25

Figure 1.3: Schematic summary of nutritional interactions in the cnidarian-dinoflagellate symbiosis.....27

Figure 1.4: Host-microbe signalling during microbial invasion and phagocytosis in human phagocytic cells and macrophages.....34

Figure 1.5: Signalling pathway of toll-like receptors (TLRs).....36

Figure 1.6: A schema for the origin and evolution of plastids by primary and secondary endosymbiosis, following the fate of the red algal lineage..41

Figure 1.7: The model anemone *Aiptasia* sp.....47

Figure 1.8: Confocal microscope image of a tentacle squash of the symbiotic anemone *Aiptasia* sp..48

Figure 2.1: Domain architecture of vertebrate SRs. Class A mammalian SRs are characterized by having one or more collagen domain.....53

Figure 2.2: Domain architecture of cnidarian SR's compared to Human SR's of known function. SR's display eclectic domain architecture, multiple combinations and permutations create considerable diversity even within classes.....64

Figure 2.3: Maximum likelihood protein tree produced using SR class B protein CD36 domains only.....67

Figure 2.4: Experimental re-infection of aposymbiotic *Aiptasia* sp. treated with the known SR ligand, fucoidan, from the algae, *Fucus vesiculosus.*.....69

Figure 2.5: Effect of immune stimulation by LPS on NO production in symbiotic *A. pallida* treated with the known SR ligand fucoidan from the seaweed *Fucus vesiculosus*. ..... 70

Figure 3.1: What is a TSR?. ..... 77

Figure 3.2: What is a TRAP (thrombospondin-related anonymous protein)?.... 84

Figure 3.3: TSR domain alignment comparing apicomplexan TRAP TSR domains. ..... 85

Figure 3.4: The TSR domain is very well conserved from cnidarians up to humans with binding motifs for glycosaminoglycans (GAG's) and the type B scavenger receptors, CD36/SRB1. ..... 86

Figure 3.5: | Domain architecture of cnidarian TSR super-family proteins compared to known vertebrate TSR-domain-containing proteins. .. 88

Figure 3.6: Maximum-likelihood tree of TSR protein domains ..... 89

Figure 4.1: a) A simplified TGFβ signaling pathway. TGFβ is secreted in a latent form, and activation of latent TGFβ to its active form often involves the TSR domain of the thrombospondin 1 protein binding to the Class B scavenger receptor, CD36. ..... 100

Figure 4.2: A section of the antibody binding region of the human thrombospondin 1/2 antibody, aligned to a TSR protein fragment from *Aiptasia* sp. ..... 103

Figure 4.3: Alignment of the second TSR domains from human thrombospondin 1 and TSR proteins from the anemone *Aiptasia sp*. and the dinoflagellate *Symbiodinium minutum*. ..... 107

Figure 4.4: TSP antibody binding proteins are expressed in higher quantities in symbiotic *Aiptasia sp.* than in the aposymbiotic state of this anemone. ..... 109

Figure 4.5: Confocal microscope images of dinoflagellate cells taken from culture and freshly isolated cells taken from *Aiptasia sp*. anemone homogenate. ..... 110

Figure 4.6: Confocal microscope images of cryosections from whole symbiotic anemones (*Aiptasia* sp.). ..... 110

Figure 4.7: Infection rates in anemones treated with a rabbit polyclonal antibody with an epitope corresponding to amino acids 401-700 of thrombospondin 2 of human origin. ..... 111

Figure 4.8: The blocking effect of thrombospondin antibody can be reversed by the addition of TGFβ protein. ..... 112

Figure 4.9: The addition of exogenous (human platelet) thrombospondin (TSP -1) significantly increased the infection rate at the onset of infection. ..... 113

Figure 4.10: The effect of synthetic TSR peptides 1 and 2 on infection rates.. 114

Figure 4.11: Three possible hypotheses for how TSR proteins may be involved in the onset and maintenance of cnidarian/dinoflagellate symbiosis. ..... 116

Figure 5.1: Recognition and phagocytosis – testable cellular model. ..... 124

Figure 5.2: Post phagocytosis – in the host gastrodermal cell, the symbiont encapsulated is within the symbiosome membrane complex. ..... 126

Figure A.1: Protein alignment of SR class B proteins used to produce the ML protein tree in Figure 2.3. ..... 169

Figure B.1: Protein alignment of TSR protein domains used to produce the ML protein tree in Figure 3.6. ..... 176

# Table of Tables

Table 1.1: Comparison of molecules and mechanisms used by two apicomplexan intracellular parasites, a bacterium and *Symbiodinium* sp. ..... 43

Table 2.1: Anthozoan resources searched for SR proteins ..... 63

Table 3.1: Anthozoan resources searched for TSR-domain-containing proteins ..... 88

Table A.1: Resources from which invertebrate SR sequences were obtained. 158

Table A.2: Accession numbers for CD36-domain-containing proteins used for phylogenetic analysis in Figure 2.3 ..... 168

Table A.2: SR protein information ..... 168

Table B.1: Resources from which invertebrate TSR-domain-containing sequences were obtained ..... 167

Table B.2: TSR domain containing sequence information ..... 168

## List of Abbreviations

- BLAST: basic local alignment search tool
- BLASTn: search a nucleotide database using a nucleotide query
- BLASTp: search protein database using a protein query
- BLASTx: search protein database using a translated nucleotide query
- BSA: bovine serum albumen
- C3: complement protein
- CD14: cluster of differentiation 14
- CD163: (Cluster of Differentiation 163) is a protein that in humans is encoded by the CD163 gene. CD163 is the high affinity scavenger receptor for the hemoglobin-haptoglobin complex.
- CD36: cluster of differentiation 36, also known as FAT (fatty acid translocase)
- CD5: (Cluster of Differentiation 5) found in a subset of immunoglobulin M (IgM) cells called B1 cells also on T cells.
- CD6: (Cluster of Differentiation 6) is a human protein encoded by the CD6 gene. This gene encodes a protein found on the outer membrane of T-lymphocytes
- CD81: (Cluster of Differentiation 81), is a protein, which in humans is encoded by the CD81 gene. It is also known as 26kDa cell surface protein, TAPA-1 (Target of the Antiproliferative Antibody 1), and Tetraspanin-28 (Tspan-28).
- $CO_2$: carbon dioxide
- CTLD: C-type lectin domain
- CUB: The CUB domain
- DAPI: (4',6-diamidino-2-phenylindole) is a fluorescent stain that binds strongly to A-T rich regions in DNA
- DCMU: 3-(3,4-dichlorophenyl)-1,1-dimethylurea
- DIC: dissolved inorganic carbon
- DIN: dissolved inorganic nitrogen (ammonium or nitrate)
- DMBT: Deleted in malignant brain tumors 1 protein
- DNA: deoxyribonucleic acid
- DOM: dissolved organic matter
- ECM: extracellular matrix

- EGF: epidermal growth factor
- EST: expressed sequence tag
- FcR: receptor for the Fc portion of immunoglobulin
- FSW: filtered sea water
- GO: Gene Ontology
- GPIs: glycosylphosphatidylinositol anchors
- $HCO_{3^-}$: bicarbonate
- ITS2: internal transcribed spacer-2
- IκB: IκB kinase enzyme complex is part of the upstream NF-κB signal transduction cascade
- KEGG: Kyoto Encyclopedia of Genes and Genomes
- LMP2: latent membrane protein 2
- LOX: Lysyl oxidase (LOX), also known as protein-lysine 6-oxidase, is a protein that, in humans, is encoded by the LOX gene.
- LPS: lipopolysaccharide
- MAM: the MAM domain is an evolutionary conserved protein domain. It is an extracellular domain found in many receptors.
- MAMPs: microbe associated molecular patterns
- MARCO: macrophage receptor with collagenous structure (also known as SCARA2 AND SR-A2).
- MASP: mannan-binding lectin serine protease also known as mannose-associated serine protease
- MBL: mannose binding lectin
- mLDL: modified low density lipoprotein
- MR: mannose receptor
- MSR1: macrophage scavenger receptor 1
- MyD88: Myeloid differentiation primary response gene 88
- NFκB: nuclear factor kappa-light-chain-enhancer of activated B cells
- NGS: next generation sequencing
- NO: nitric oxide
- NOS: nitric oxide synthase
- PBS: phosphate-buffered saline

- PFA: paraformaldehyde
- PG: peptidoglycan
- POM: particulate organic matter
- PRRs: pattern recognition receptors
- PV: parasitophorous vacuole
- PVM: parasitophorous vacuole membrane
- SCARA5: scavenger receptor class A member 5
- SCARB1: Scavenger receptor class B member 1 (SRB1) also known as SR-BI is a protein that in humans is encoded by the SCARB1 gene.
- SR-A: Class A scavenger receptors
- SR-A1: Scavenger receptors type 1 also known as SCARA1 or MSR1: SR-A1 scavenger receptors have a cysteine-rich domain
- SR: scavenger receptor
- SRA: sequence read archive
- SRB1: scavenger receptor class B member 1, also known as SCARB1
- SRCL: scavenger receptor C-type lectin, (also known as SR-A2) have a C-type lectin domain (CTLD)
- SRCR: scavenger receptor cysteine rich domain
- tBLASTn: search translated nucleotide database using a protein query
- tBLASTx: search translated nucleotide database using a translated nucleotide query
- TGFβ: transforming growth factor beta
- TLR4: toll-like receptor 4
- TLRs: toll-like receptors
- TRAF6: TNF receptor associated factor 6
- TRAP: thrombospondin-related anonymous protein
- TSR: thrombospndin structural homology repeat
- UBOX: The U box is a domain of approximately 70 amino acids that is present in proteins from yeast to humans.

# Chapter 1

# General Introduction

## 1.1 Symbiosis

Symbiosis describes a close relationship between two or more phylogenetically different organisms. Symbiotic associations are prevalent in all environments and are a driving force in evolution: many species have evolved to use one another to achieve greater reproductive successes. The term symbiosis was formally described by the German mycologist Heinrich Anton de Bary, who stipulated that the relationship must be constant, intimate and between dissimilar species (De Bary 1878). Although mutualism, the symbiotic association where all partners benefit, is the type of association most commonly thought of in the context of symbiosis, the original definition implied no cooperation or mutualism between partners anymore than it implied exploitation; it simply described an intimate association. Mutualisms thus represent a fraction of symbiotic relationships: The term symbiosis also describes commensal associations (where one partner receives neither benefit nor harm) and parasitic associations (where one partner benefits at the expense of another) (Douglas 2009). Despite this, parasitic relationships are very rarely described as symbiotic and for this reason the term 'symbiont' is usually used to describe a microorganism that is beneficial to its host.

Symbiotic associations can also be either obligate or facultative, depending on the necessity of the association for one or more partner's survival. An example of obligate mutualism is the mycorrhizal fungi that associate with plant roots; the fungus increase water uptake by the plant and the plant provides essential nutrients to the fungus (Neuhauser 2004). Neither partner can survive without the other. By contrast, in a facultative mutualism, both organisms can survive independently, however both derive benefit from the association. In commensalism or parasitism, the relationship is usually obligate for the commensal or the parasite, since, by definition, they depend on the host. On the other hand, the host is in a facultative relationship with the commensal or

parasite, as it gains nothing from the association, and in the case of the parasite it would be better off without it (Sapp 1994, Leung and Poulin 2008).

Microorganisms living within a host are called endosymbionts or endoparasites in contrast to ectosymbionts and ectoparasites that reside on the external surface of a host. In the case of endosymbiosis, symbionts or parasites can be intracellular (inside host cells) or extracellular (outside host cells in the body cavity or in a specific organ). In humans and other vertebrates, the vast majority of microbial diversity is extracellular. The human gut, for example, hosts large microbial communities - in a healthy human body more than 90% of the cells are of microbial origin - and the only intracellular microorganisms are parasites (Douglas 2009). Endosymbiotic relationships are more common in invertebrates, perhaps due to the more complex adaptive immune system found in vertebrate hosts (Leung and Poulin 2008).

Symbiosis at the cellular level is highly complex and evolutionarily important, as highlighted by the endosymbiotic theory of evolution (Margulis 1975, Margulis and Bermudes 1985, Margulis 1988, Vesteg and Krajcovic 2008). Inside every eukaryotic cell there are organelles that have bacterial origin. Nucleated (or eukaryotic) cells are thought to be a collection of non-nucleated prokaryotic cells that have evolved into organelles that fulfil specialist roles within cells, such as photosynthetic chloroplasts and mitochondria. Endosymbiotic theory states that these organelles were once free-living specialists in their own right, but formed an alliance with nucleated cells, evolving through a series of multiple endosymbiosis (Schwartz and Dayhoff 1978, Vesteg and Krajcovic 2008). In its strongest form, symbiosis can lead to symbiogenesis – the evolution of a new species *via* the genetic integration of a symbiotic partner. Indeed, the photosynthetic dinoflagellate symbionts of many marine invertebrates, including cnidarians, evolved by an endosymbiosis between a red alga and a heterotrophic host (Archibald and Keeling 2002).

The cnidarian-dinoflagellate symbiosis is one of only very few examples of a mutualism formed between an animal and a unicellular eukaryote. The vast majority of eukaryotic microbes are parasitic. Indeed approximately half of all dinoflagellates themselves are heterotrophic and/or parasitic – some parasitic dinoflagellates even parasitize other parasitic dinoflagellates (Coats 1999). In reality, the lines between mutualism, commensalism and parasitism are often fluid on evolutionary and ecological timescales (Leung and Poulin 2008, Sachs et al. 2011a, 2011b): a single mutation can transform a

beneficial symbiont into a pathogen and *vice versa* (Douglas 2008), and a symbiotic relationship can shift from mutualism to parasitism due to a change in environmental conditions upsetting a complicated balance between partners (Neuhauser 2004, Sachs and Wilcox 2006, Sachs and Simms 2006, Douglas 2008). Indeed, even the cnidarian-dinoflagellate symbiosis can be thought of on a continuum of symbiotic states (Lesser et al. 2013): cnidarians associate with a range of *Symbiodinium* spp. phylotypes (see below for further discussion), and some of these may not be beneficial for their hosts. Furthermore, coral bleaching, the loss of coral symbionts due to high temperature or other environmental stresses, can be viewed as a state in which the cost of the symbiosis outweighs its benefits, and the host expels the symbionts, which act as parasites under such unfavourable conditions (Wooldridge 2010).

Due to their clinical, veterinary or economic importance, many parasites have been the focus of extensive research. In apicomplexans, a group of obligate intracellular parasites related to dinoflagellate coral symbionts, decades of research have revealed complex cellular mechanisms of host-cell entry and maintenance of the parasite. A famous apicomplexan, *Plasmodium* sp. - the malaria parasite - is reported as having approximately 50 different proteins involved in merozoite invasion of red blood cells and approximately 30 proteins in sporozoite invasion of liver cells (Patarroyo et al. 2015). This complexity of invasion mechanisms has developed over long-term co-evolution with the animal innate immune system (Sacks and Sher 2002). The animal innate immune system is designed to recognise, internalise and destroy pathogenic microbes, so parasitic microbes have evolved sophisticated mechanisms to gain entry to host cells and manipulate immune responses to their own benefit (Medzhitov and Janeway 2000, Janeway and Medzhitov 2002, Plüddemann et al. 2011). Our wealth of knowledge about parasitic mechanisms can both inform and benefit the study of mutualistic relationships. In particular, advances in apicomplexan parasitology may offer important clues of how the taxonomically related dinoflagellates gain entry to cnidarian hosts (Schwarz 2008). Many of the mechanisms of host entry and persistence within host cells are potentially shared among these systems (Sachs et al. 2011a), and cross-pollination between the two fields is a way to gain an understanding of the complexity of signalling mechanisms involved in intracellular symbioses.

## 1.2 Cnidarian-dinoflagellate symbiosis

Cnidarians are diploblastic: they have two-tissue layer organisation rather than the triploblastic (three-tissue layer) organisation found in the majority of metazoan phyla. A single opening acts as both mouth and anus to the gastrovascular cavity, and is essentially a blind sac filled with seawater (Figure 1.1a). The epidermal tissue layer houses specialised stinging cells called nematocysts, which can capture small zooplankton prey and enable the cnidarian to feed heterotrophically. Despite this heterotrophic ability, many cnidarians rely heavily on energy gained from a partnership with photosynthetic microalgae (Venn et al. 2008, Yellowlees et al. 2008). The cnidarian-dinoflagellate symbiosis is ecologically important as it is responsible for supporting the construction of coral reefs (Muscatine and Cernichiari 1969), which contain a quarter of global marine biodiversity (Wilson and Peter 1988, Reaka-Kudla et al. 1996).

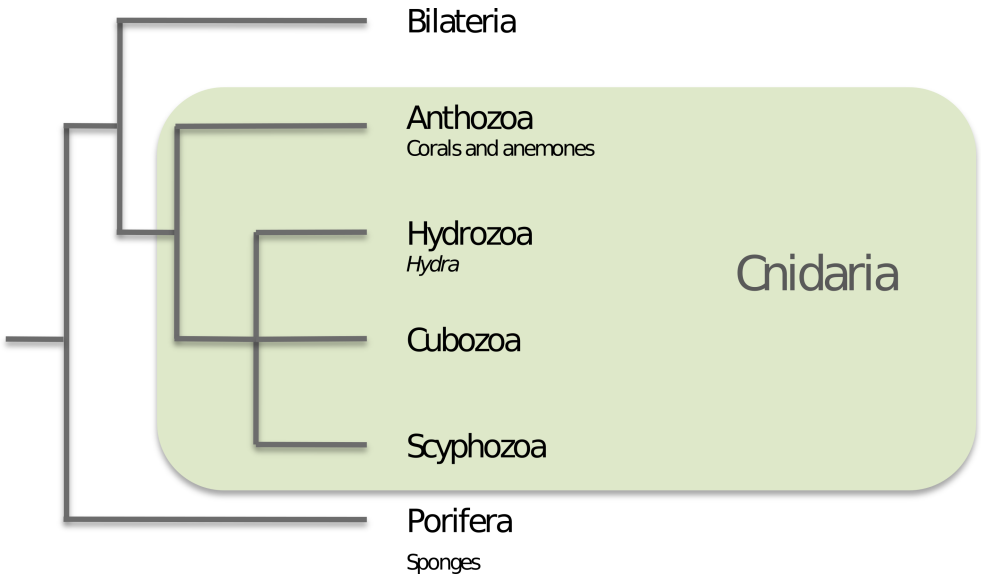



Figure 1.1: Cnidarians as a basal group in metazoan evolution. Within the Cnidaria, the Anthozoa are basal and the Hydrozoa are derived. Adapted from Miller et al. (2007).

Photosynthetic symbionts (microalgae and cyanobacteria) are common in the marine invertebrate phyla Porifera (sponges) and Cnidaria (hydroids, jellyfish, anemones and corals). The simple body organisation and large surface area-to-volume ratio of these animal groups makes them ideal algal hosts (Venn et al. 2008). Cnidarians form an important group of basal metazoans, and as a sister taxon to the bilaterians, they are important to the study of metazoan evolution (Dunn et al. 2008). Figure 1.2 shows the

phylogenetic relationships between members of the Cnidaria and other metazoans, with a basal group, the Anthozoa, and a derived group, the Hydrozoa (Miller et al. 2007). Corals and anemones belong to the cnidarian class Anthozoa, characterised by the lack of a medusoid stage in their life cycle. Many anthozoans, including hermatypic corals (reef-building corals), form a symbiotic association with dinoflagellate algae belonging to the genus *Symbiodinium.* This relationship is ancient and widespread in tropical and, to a lesser extent, temperate oceans (Stanley 2003).

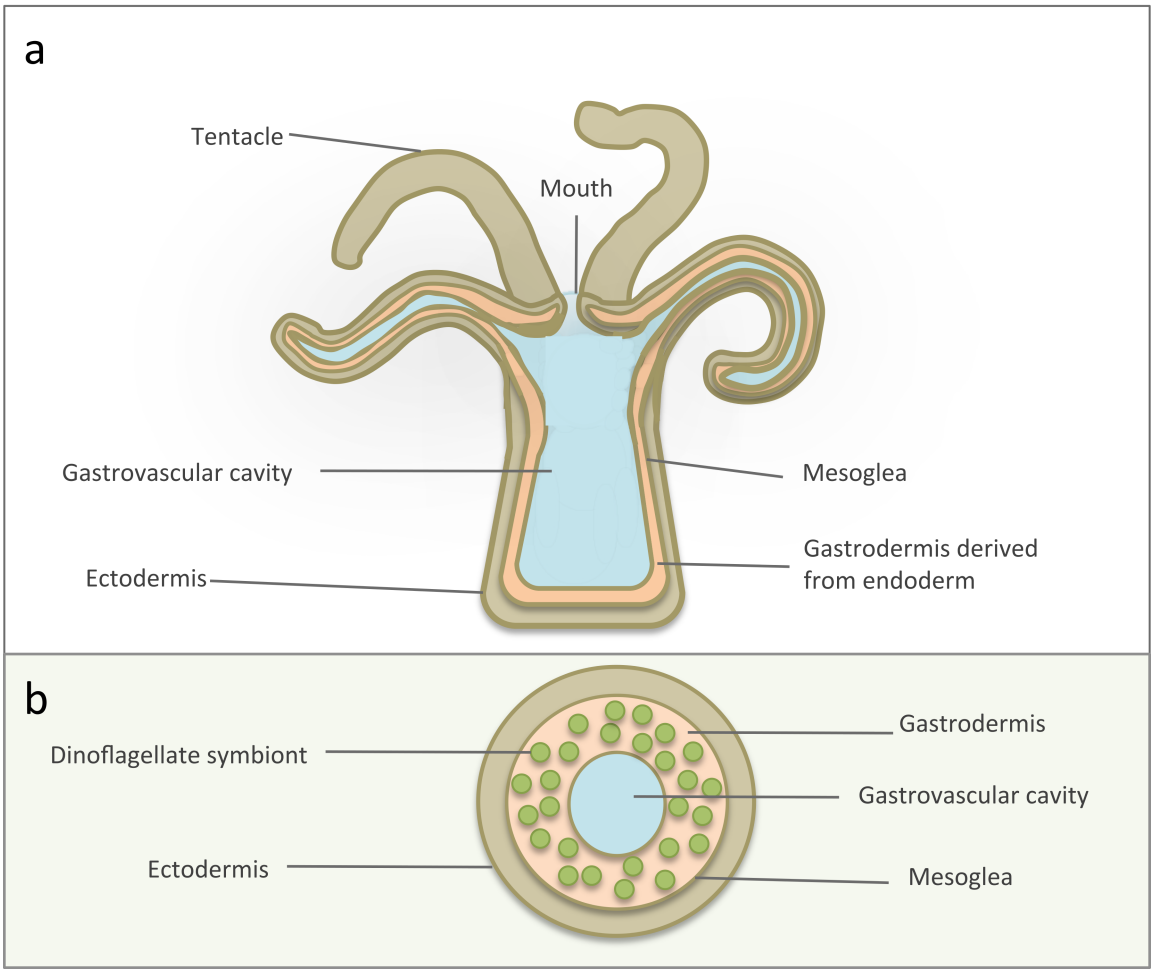



Figure 1.2: Simplified schematic diagram of a symbiotic anthozoan (phylum Cnidaria). (a) Polyp showing two layered (diploblastic) tissue organisation and gastrovascular cavity filled with seawater. (b) Transverse section through a tentacle showing dinoflagellate symbionts within the gastrodermal tissue layer.

The cnidarian-dinoflagellate relationship is mutualistic; both partners derive benefit. The partnership is based on nutritional exchange: dinoflagellates offer photosynthetically fixed carbon (e.g. glucose, glycerol, amino acids) to the host in exchange for inorganic nitrogen, phosphorus and carbon and a high-light, herbivore-

free home (Yellowlees et al. 2008). Dinoflagellate symbionts reside inside host gastrodermal cells (Figure 1.2b). In optimal sunlight, when the tentacles are fully expanded, the symbionts form a monolayer to optimise light capture (Glider et al. 1980). The host has adapted to recycle inorganic nutrients that would ordinarily be excreted and has developed mechanisms to deliver these to the symbiont (Yellowlees et al. 2008). The symbiont can be described as providing a 'waste disposal service' to the host by removing respiratory $CO_2$ and waste metabolites (Allemand et al. 1998). Figure 1.3 provides a detailed summary of the nutritional interactions between host and symbiont. Many of the key questions regarding the mechanisms behind the metabolic integration and transfer of metabolites in the alga-invertebrate symbiosis were raised in the 1970s and early 80s by Muscatine and Trench, but remain unanswered (Muscatine and Cernichiari 1969, Trench 1971, Muscatine 1974, Muscatine et al. 1975, Colley et al. 1983, Fitt and Trench 1983). Indeed, despite decades of research in this area, it is still not clear which organic compounds are available to the host and the regulation of their release is even less clear (Venn et al. 2008, Davy et al. 2012).

Much of the early work exploring the cell biology of cnidarian-algal symbiosis used the freshwater hydroid, *Hydra viridis* (or green hydra) that forms an intracellular association with a chlorophyte belonging to the genus *Chlorella* (Cernichiari et al. 1969, Muscatine et al. 1975, Jolley and Smith 1980, McNeil et al. 1982, McAuley and Smith 1982). This green *Hydra* research provided the basis for more recent studies of the cnidarian-dinoflagellate symbiosis. Despite the difference in symbiotic partners, this work is still relevant to the anthozoan-dinoflagellate symbioses, as very little research of this type has been done since. Several studies from these early years are especially relevant to the present study as they examine the mechanisms of symbiont recognition and uptake, and will be considered further below (Cernichiari et al. 1969, Jolley and Smith 1980, McNeil et al. 1982, McAuley and Smith 1982).

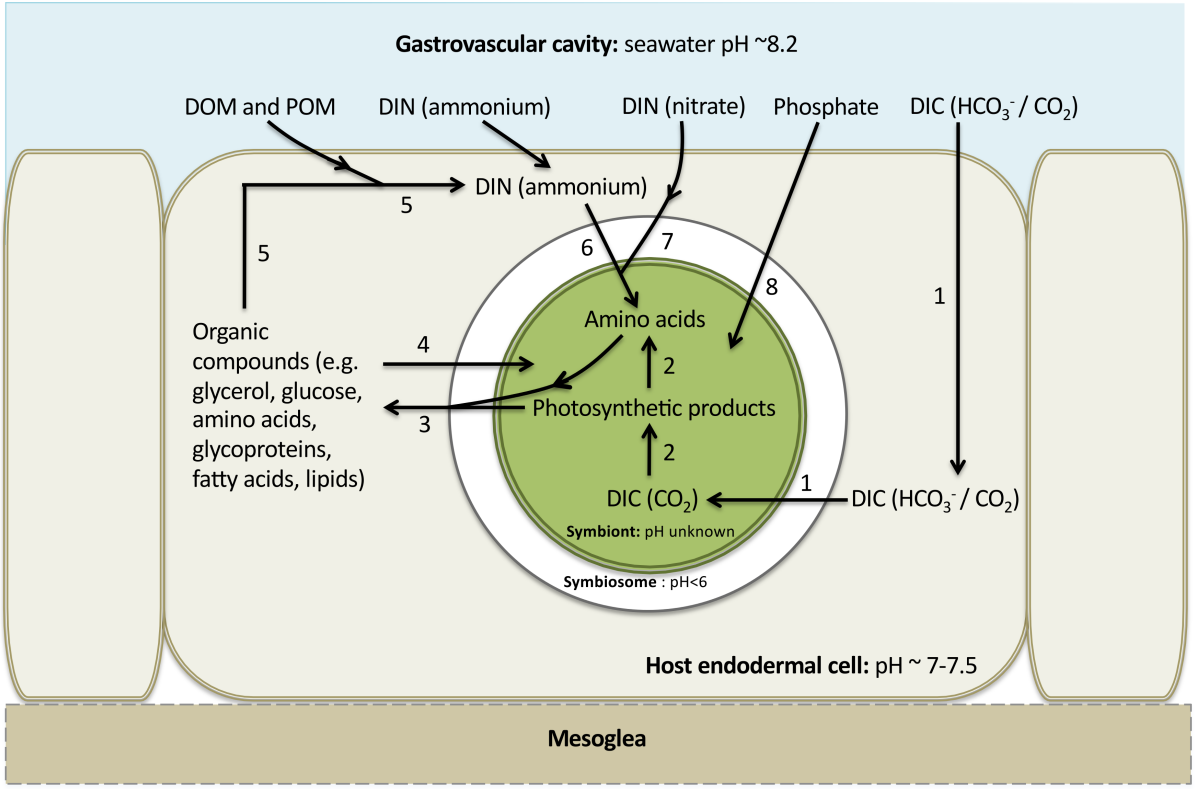



Figure 1.1: Schematic summary of nutritional interactions in the cnidarian-dinoflagellate symbiosis. 1. Dissolved inorganic carbon (DIC) uptake. DIC is acquired either as bicarbonate ($HCO_3^-$) from the surrounding seawater or as $CO_2$ from the seawater or host metabolism/calcification. In the case of $HCO_3^-$, it must be converted to $CO_2$ prior to photosynthesis by the dinoflagellate symbiont. 2. Photosynthesis. $CO_2$ is photosynthetically fixed through the Calvin-Benson cycle (i.e., the C3 pathway), with the dinoflagellate ultimately synthesizing a range of organic compounds, including amino acids. 3. Translocation. A portion of the photosynthetic products are translocated to the host cell. 4. Reverse translocation. Organic compounds are likely translocated from the host to the symbiont; these compounds could arise from host metabolism or be in the same forms as those originally translocated by the symbiont. 5. Host metabolism. Translocated compounds are used, alongside dissolved organic matter (DOM) and particulate organic matter (POM) taken up from seawater, to support host metabolism. The catabolism of nitrogenous compounds ultimately leads to the generation of ammonium waste that can be assimilated by the symbiont. 6. Ammonium assimilation. Excretory and seawater ammonium can be assimilated by both the host cell (pathway not shown) and the symbiont, with translocated organic compounds providing carbon skeletons necessary for host assimilation. The assimilation of excretory ammonium back into amino acids by the dinoflagellate symbiont completes the process of "nitrogen recycling" by the symbiosis. 7. Nitrate assimilation. Nitrate is taken up from the seawater, but only the symbiont can convert it to ammonium for subsequent assimilation into amino acids. 8. Phosphate assimilation. Phosphate is likewise taken up from seawater and can be assimilated by the dinoflagellate symbiont. Note that uptake of nutrients can also occur from the ambient seawater *via* the epidermis (not illustrated), but for simplicity these pathways are not shown. Re-drawn from Davy et al. (2012).

## 1.3 Establishment of cnidarian-dinoflagellate symbiosis

In relatively few cnidarians, dinoflagellate symbionts are transmitted through maternal inheritance (i.e., in asexual reproduction or budding, or *via* implantation into the egg) (Davy and Turner 2003), but in most cases symbionts must be acquired from the surrounding seawater by cnidarian larvae or settling polyps (Babcock et al. 1986). The mechanism by which the symbiont colonizes the host and is tolerated is a subject of recent interest (Schwarz 2008, Weis et al. 2008, Davy et al. 2012, Lehnert et al. 2014). The establishment of a symbiont population within the host can be described by three sequential steps: (1) recognition and phagocytosis; (2) selection (specificity); and (3) arrest of phago-lysosome maturation and symbiont persistence (Nyholm and Mcfall-Ngai 2004, Davy et al. 2012). Each step involves inter-partner signalling and a degree of cooperation between partners.

### *1.3.1 Symbiont recognition and uptake via phagocytosis*

Host-symbiont recognition involves molecular signalling between host and symbiont, in the form of complex partner crosstalk using a vast array of ligands, receptors and protein cascades that control and regulate the onset of symbiosis. This initial contact occurs between receptors within the extracellular matrix (ECM) of host gastrodermal cells and the algal cell surface proteins. The *Symbiodinium* cell secretes large molecular weight glycoproteins (Markell and Trench 1993). These secreted glycoconjugates provide candidates for signalling and binding to host ECM receptors (Markell and Wood-Charlson 2010). Lectin-glycan interactions are among some of the most studied inter-partner recognition mechanisms. Several studies have investigated the role of lectin-glycan signalling in the onset of cnidarian-dinoflagellate symbiosis (Lin et al. 2000, Wood-Charlson et al. 2006, Kvennefors et al. 2008, Wood-Charlson and Weis 2009, Logan et al. 2010, Markell and Wood-Charlson 2010), the green *Hydra-Chlorella* symbiosis (Meints and Pardy 1980), and in the squid-*Vibrio* symbiosis (Nyholm et al. 2003, Nyholm and Mcfall-Ngai 2004). Experimental evidence in several symbiotic systems implicates lectin-glycan interactions in recognition and phagocytosis of symbionts by host cells (Wood-Charlson et al. 2006, Schwarz et al. 2008, Kvennefors et al. 2008, Kerrigan and Brown 2009, Markell and Wood-Charlson 2010). The lectin-

glycan signalling mechanism is discussed in terms of the host's innate immune system in the relevant section below.

Photosynthetic products can also act as signalling molecules enabling symbiont uptake by the host. As previously mentioned, much of the early research in this area was performed using the green *Hydra-Chlorella* symbiosis (Jolley and Smith 1980, McNeil et al. 1982, McAuley and Smith 1982). Examination of the potential signalling molecules found that certain strains of *Chlorella* released large amounts of photosynthate in the form of the sugar maltose. These *Chlorella* strains were more successful at colonizing the host than those releasing little or no maltose (Hohman et al. 1982, McNeil et al. 1982). Low maltose-releasing symbionts were attacked by host lysosomes, whereas high maltose-releasing algae persisted inside the host cell. Interestingly, high maltose-releasing algal cells treated with the photosynthetic inhibitor 3-(3,4-dichlorophenyl)-1,1-dimethylurea (DCMU) were also destroyed by the host (McAuley and Smith 1982).

Symbionts are taken in through the host mouth and are subsequently taken up by gastrodermal cells *via* phagocytosis (Schwarz et al. 1999). Dinoflagellate cells present in the seawater can enter the gastrodermal cavity of polyps via the oral disc. Captured particles can be wiped from tentacles on the oral disc, transported via ciliated grooves on the tentacle and/or taken in with seawater by micro-currents generated by cilia around the oral disc; alternatively, motile forms of the symbiont may swim directly into the gastrodermal cavity (Fitt and Trench 1983). Once inside, the algae are ingested by phagocytosis (Muscatine et al. 1975, McNeil 1981, Colley et al. 1983). Phagocytosis is a special form of endocytosis, used to ingest large particles such as microorganisms and dead cells (Sansonetti 2000).

Three distinct methods of phagocytosis of live symbionts have been characterised in hydra: either a multiple membrane fold, a meshwork of microvilli or a funnel shaped extension of the plasmalemma (McNeil 1981, McNeil et al. 1981). Specifically, heat-killed cells were only ingested by a modified funnel mode and that *Artemia* (food particles) were ingested only by a multiple membrane-fold mode (McNeil 1981). By using plain latex beads and beads suspended in various solutions containing polyanions, it was shown that ingestion via the microvilli mesh mode can be induced (McNeil et al. 1981). These observations lead to the following questions: does the mode of phagocytosis determine the fate of the phagosome? Or does recognition or sorting into digestible and non-digestible phagosomes occur post-phagocytosis? Phagocytic

recognition by digestive cells is a separate mechanism to post phagocytic symbiont recognition and symbiont selection (McNeil et al. 1982). Phagocytosis can be both nutritive and act as an innate immune protective mechanism, whereby digestive cells take up dangerous microbes with intent to destroy them (Sansonetti 2000). Phagocytosis therefore be can be a non-specific uptake of a wide range of cells and subsequent discrimination between symbionts, food particles and pathogens occurs after entry to cells not before or during phagocytosis (McNeil et al. 1982).

### *1.3.2 Selection (specificity)*

The dinoflagellate symbiont enters the host *via* phagocytosis, surrounded by a phagosome composed of host-derived plasma membrane. Following successful uptake, phagosomes containing specific dinoflagellate strains selectively develop into a specialised vacuole referred to as the 'symbiosome', while other strains are expelled or destroyed (Kazandjian et al. 2008, Peng et al. 2010). Only live and photosynthetically active symbionts are retained by the host; dead and dysfunctional cells are rejected by this post-phagocytic selective process (Dunn and Weis 2009).

Since the *Symbiodinium* phylotype can determine cnidarian host fitness, selection of phylotypes by the host is important on both evolutionary and ecological timescales (Lesser et al. 2013). Indeed, cnidarian hosts have been observed to preferentially form associations with specific *Symbiodinium* phylotypes; this is termed 'specificity' (Baker 2003, LaJeunesse et al. 2004, Coffroth and Santos 2005, Van et al. 2005). Genetic diversity within the genus *Symbiodinium* has been categorised into nine clades (A-I) using the nucleotide sequence of the large ribosomal subunit to distinguish between clades (Rowan and Powers 1991, Pochon et al. 2004) and the internal transcribed spacer-2 (ITS2) to distinguish within clades. Clade C in particular shows considerable within-clade diversity (LaJeunesse 2001). Some of these clades may be more beneficial to their hosts than others: clade C, for example, may be interpreted as the optimal symbiont for coral, whereas clades A & D may lower host fitness and may thus act as parasites (Lesser et al. 2013). Most hosts are therefore equipped to preferentially retain or reject certain *Symbiodinium* phylotypes, and the majority (approx. 75%) of host species associate with one or two main *Symbiodinium* phylotypes, while maintaining a very small population of diverse phylotypes (Baker 2003, Silverstein et al. 2012). For

instance, in the coral *Fungia scutaria*, the larval preference for certain symbiont types appears stable over geographic and temporal scales (Weis et al. 2001). These larvae employ apoptosis-like pathways to prevent colonisation by unsuitable symbionts, and inhibition of these pathways results in persistent infection of the host by these previously incompatible *Symbiodinium* types (Dunn and Weis 2009).

### *1.3.3 Arrest of phago-lysosome maturation and persistence within host cells*

Early cell biology research into the cnidarian-algal association investigated the mechanisms of phago-lysosome maturation arrest and evasion of digestion. $^{14}$C-labelled and heat-killed *Chlorella* cells were readily and equally taken into hydra digestive cells by phagocytosis, however heat-killed cells were rapidly digested while healthy symbionts persisted (Hohman et al. 1982). The same study also found that healthy symbionts were selectively transported from the apical site of phagocytosis to the base of digestive cells. By labelling lysosomes with ferritin, it was shown that live symbiont-containing phagosomes failed to attach to lysosomes whereas ferritin was found in 50% of vacuoles containing heat-killed symbionts and all vacuoles containing *Artemia* food particles (Hohman et al. 1982). Similar experiments were repeated using a marine system, the jellyfish-dinoflagellate (*Cassiopeia xamachana*-*Symbiodinium*) symbiosis. Again, phago-lysosome fusion occurred in heat-treated symbiont and food-containing vacuoles, but there was no evidence of lysosome fusion with those containing live symbionts (Fitt and Trench 1983). Evasion of phago-lysosome fusion can be reversed by treatment with photosynthetic inhibitors such as polycationic polypeptides, DCMU and darkness (Perez et al. 2001, Belda-Baillie et al. 2002, Weis et al. 2008). Dead or photosynthetically-compromised dinoflagellate cells have been observed to accumulate surface proteins that lead to lysosome targeting and eventual destruction (Fitt and Trench 1983, McNeil and McAuley 1984). This would suggest that photosynthetic products are necessary signals to avoid lysosome development.

The phagosome membrane contains proteins that provide signals to progress phago-lysosome maturation. By manipulating the phagosome membrane, these essential signalling proteins can be excluded (Zerial and McBride 2001, Hong et al. 2009b). The symbiosome lacks characteristics of an ordinary phagosome (Chen et al. 2003b). Rab GTPases coordinate membrane traffic, including vesicle formation and vesicle

movement (Zerial and McBride 2001). Rab proteins can be used as markers to indicate stages in endosome development: Rab5, 4, 15, 18, 20 and 22, for example, indicate early endosome, while Rab11, 25 and 17 indicate recycling endosome, and Rab7, 9, 27 and 24 are associated with late endosome and lysosome fusion (Zerial and McBride 2001, Greenberg and Grinstein 2002). In the sea anemone, *Aiptasia-Symbiodinium* association, Rab proteins (Rab 7 and Rab 11) are excluded from the symbiosome membrane (Chen et al. 2003b). These proteins are involved in the recycling process essential for phagosome maturation, and their active exclusion, halts phagosome maturation and prevents phago-lysosome fusion (Schwarz 2008). Further study of Rab proteins and their association with the symbiosome in *Aiptasia* sp. showed that early endosomal Rab proteins (Rab 3, 4 and 5) are associated with healthy symbionts within symbiosome membranes (Chen et al. 2004, Hong et al. 2009b, 2009a).

## 1.4 Innate immunity in the context of cnidarian algal-symbiosis

The intracellular nature of the cnidarian-dinoflagellate mutualism raises many questions regarding how the symbiont manages a potential immune response from the host to gain entry and persist within host cells. All animals possess an innate immune defence system that responds to microbial attack by recognising microbial associated molecular patterns (MAMP's) in the form of tell-tale microbial surface proteins, such as glycans and lipopolysaccharide (LPS). The surface receptors of the innate immune system are known as pattern recognition receptors (PRRs) and are responsible for initiating various host defence pathways (Janeway and Medzhitov 2002). Several studies have investigated innate immune pathways in cnidarians (Miller et al. 2007, Schwarz et al. 2007, Bosch 2008, Wood-Charlson and Weis 2009, Kvennefors et al. 2010, López et al. 2011, Hamada et al. 2012, Poole and Weis 2014) and in particular potential tolerogenic signals between the two partners (Detournay et al. 2012). It is not known how the cnidarian host manages the balance between beneficial microbes, such as their symbiotic dinoflagellates, and the maintenance of an immune response to pathogenic attack.

In mammalian cells, the extracellular matrix (ECM) provides the first line of defense against invading microbes. PRRs of the animal innate immune system are responsible for managing the microbiome, maintaining a balance between tolerating beneficial

microflora and a response to pathogenic microbial attack (Schwarz 2008). PRRs in the ECM are both secreted and transmembrane multi-ligand proteins, which recognise MAMPs and trigger phagocytosis of alien microbes, leading to a cascade of phago-lysosome maturation and proteolytic degradation or apoptosis (Sacks and Sher 2002, Sotolongo et al. 2012). ECM proteins are therefore potential target ligands for microbial surface proteins to bind and gain entry to host cells. Figure 1.4 describes host-microbe signalling mechanisms during microbial invasion and phagocytosis in human phagocytic cells and macrophages. Many of these receptor-ligand signalling mechanisms have been identified in cnidarian systems through analysis of cnidarian genomic resources (Miller et al. 2007, Bosch 2008, Kvennefors et al. 2008, Wood-Charlson and Weis 2009, Kvennefors et al. 2010, Detournay et al. 2012). Phagocytic surface receptors of the ECM include non-opsonic receptors (e.g. C-type lectins and Scavenger Receptors), as well as opsonic receptors (e.g. complement receptor and Fc receptors).

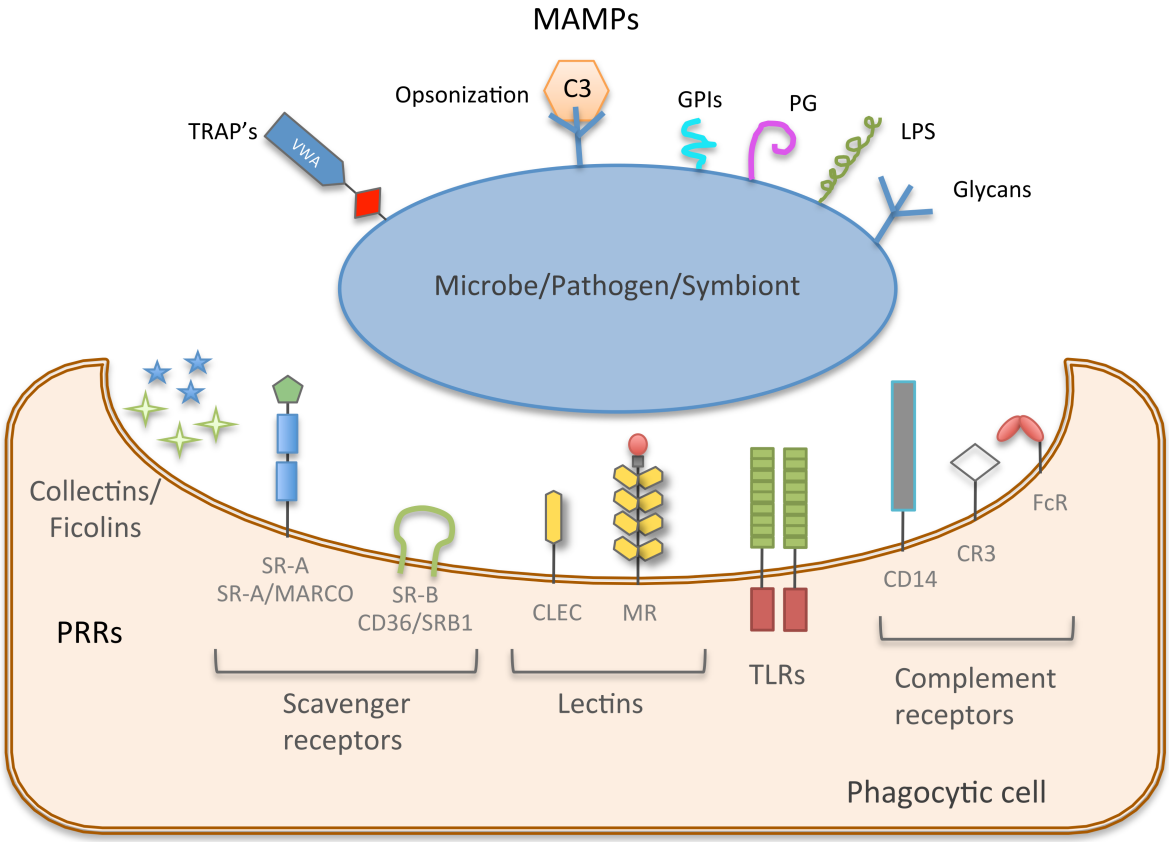



Figure 1.2: Host-microbe signalling during microbial invasion and phagocytosis in human phagocytic cells and macrophages. Multiple pattern recognition receptors (PRRs) are expressed in the extracellular matrix (ECM), either secreted or membrane bound. PRRs recognize pathogens and symbionts alike *via* telltale microbe associated molecular patterns (MAMPs) on the microbe surface, either through direct binding or by opsonization. Many pathogenic microbes utilize these innate immune PRRs to gain entry to host cells *via* phagocytosis, though certain apicomplexan parasites use active penetration after binding to the host cell surface receptors, rather than receptor-mediated phagocytosis to enter host cells. Phagocytic surface receptors include non-opsonic receptors (e.g. C-type lectins and scavenger receptors) as well as opsonic receptors (e.g. complement receptor and Fc receptors). The toll-like receptor, TLR4 is responsible for activating an immune response *via* NF-κB transcription. Collectins and ficolins are secreted pattern recognition molecules that bind microbial oligosaccharides, enabling complement activation and phagocytosis. MAMP signals can be amplified by the process of opsonization, where secreted host opsonins such as C3, collectins and ficolins, bind to pathogen cells to allow easier detection. MAMPs are sugar, protein, lipid, and nucleic acid compounds. Abbreviations: TRAPs: thrombospondin-related anonymous proteins; C3: complement protein; GPIs: glycosylphosphatidylinositol anchors; PG: peptidoglycan; LPS: lipopolysaccharide; SR: scavenger receptors Class A and B; CD: cluster of differentiation; CLEC: C type lectin receptor; MR: mannose receptor and FcR: receptor for the Fc portion of immunoglobulin. Figure compiled from McGuinness et al. (2003), Pluddemann et al. (2011) and Davy et al. (2012).

### *1.4.1 Complement pathway PRRs*

The complement system is an innate immune pathway that, in vertebrates, promotes phagocytosis and lysis of invading microbes. The complement protein C3, collectins and ficolins are secreted pattern recognition molecules or opsonins that bind microbial oligosaccharides, enabling complement activation and phagocytosis (Pinto et al. 2007, Kimura et al. 2009, Dunkelberger and Song 2010, Nonaka 2011). This process of labelling the surface of invading microbes in proteins refered to as opsonins such as C3, is referred to as opsonization. Opsonin-coated microbes are recognized by complement receptors such as CD14, complement receptor 3 (CR3) and Fc receptors, which in turn initiate phagocytosis by host cells and also activate an inflammatory response (Gros et al. 2008). The complement protein C3 has been identified in cnidarians and implicated in symbiosis. C3 has been characterized in the coral *Acropora millepora*, where the protein was shown to localize directly around symbionts in adult tissue (Kvennefors et al. 2010). In addition, in the anemone *Anemonia viridis*, one C3 isoform showed tissue-specific expression, while the other varied with symbiotic state (Ganot et al. 2011). Other complement system molecules, including Factor B and MASP, have also been characterized in cnidarians (Kimura et al. 2009, Shinzato et al. 2011). Three factor B and two MASP sequences were characterised in the anemone *Aiptasia* sp., where functional work determined that both proteins showed reduced expression in symbiotic animals, indicating potential immune suppression in the host (Poole 2014).

### *1.4.2 C-type lectins*

The C-type lectin defines a super family of proteins that are characterised by the presence of one or more C-type lectin domains or CTLDs. CTLDs are involved in the recognition of specific glycans, cell-cell adhesion and phagocytosis of potential pathogens (Cambi et al. 2005, Kerrigan and Brown 2009). Examples of CTLD proteins in mammalian systems are the mannose receptor (MR), which is implicated in opsonic recognition and activation of the complement system and Dectin 1, which can directly recognise MAMPs on the surface of microbes and mediate phagocytosis. In addition, some C-type lectins work in combination with the complement system (e.g. mannose binding lectin (MBL) promotes opsonisation of microbes by C3, and associated MASP

proteins are activated on binding to pathogens which in turn cleave complement components and activate the complement system).

Sixty-seven putative CTLD-domain-containing proteins, corresponding to a total of 92 putative CTLD domains have been identified in the genome of the non-symbiotic anemone *Nematostella vectensis* (Wood-Charlson and Weis 2009). The *Symbiodinium* cell surface glycome is recognised by host lectins, and glycan removal significantly affected infection success in functional experiments involving the enzymatic removal of surface glycans and the blocking of binding sites with various lectins prior to infection (Bay 2011). Furthermore, experiments using fluorescently-labelled lectin probes revealed very different binding patterns among *Symbiodinium* types, suggesting that the dinoflagellate *Symbiodinium* cell surface glycome differs among different *Symbiodinium* types (Wood-Charlson et al. 2006, Logan et al. 2010).

### *1.4.3 Scavenger receptors*

Scavenger receptors (SR) comprise a large family of transmembrane cell surface glycoproteins. They bind and initiate phagocytosis of microorganisms and their products, such as LPS. SRs can alter cell morphology and their expression is modulated by a wide range of cytokines (Areschoug and Gordon 2009). SRs are multi-domained, however no single domain is common to all (Gough and Gordon 2000). As the name 'scavenger' suggests, all these proteins have a high affinity for a wide range of ligands, and such is their plasticity that they have been described as "molecular fly paper" (Krieger 1992). This flexibility of binding to potential ligands suggests that SRs provide a wide range of cellular functions additional to host defence and innate immunity (Greaves and Gordon 2009).

Class A and I SRs contain the scavenger receptor cysteine rich (SRCR) domain, which consists of a 110 amino acid residue motif with conserved spacing of six to eight cysteines that are involved in intradomain disulphide bridges (Hohenester et al. 1999). The SRCR domain is an ancient and conserved protein domain that often occurs in repeats in membrane and soluble proteins. In vertebrate systems, proposed functions include acting as phagocytic receptors, epithelial homeostasis, cell aggregation and immunity (Yamada et al. 1998, Janeway and Medzhitov 2002, Mukhopadhyay et al. 2004, Sarrias et al. 2004, Martinez et al. 2011, Whelan et al. 2012). SRCR-domain-

containing proteins are important in pathogenesis and diseases such as atherosclerosis, autoimmune diseases and cancer (Martinez et al. 2011). There are six different subclasses of scavenger receptors (Classes A-F). Vertebrate SR Classes A and E contain C-type lectin domains (CTLDs), which are involved in lectin-glycan interactions. Vertebrate Class B SRs, such as CD36 and SRB1, consist of the CD36 domain, named after the CD36 protein. The CD36 domain, via its ability to recognise oxidized phospholipids and lipoproteins, is involved in lipid transport, recognition and instigating the phagocytosis of apoptotic cells, bacterial and fungal pathogens and modified low density lipoprotein (mLDL) (Silverstein and Febbraio 2009). In addition to this, the vertebrate CD36 protein has been implicated in both assisting the pathogenesis of various microbial pathogens (Gantt et al. 1997, Sacks and Sher 2002, Schäfer et al. 2009) and the control of angiogenesis (Silverstein and Febbraio 2007).

SRs have been identified in invertebrate systems. Studies of the purple sea urchin, *Strogylocentrotus purpuratus*, genome discovered a vastly expanded innate immune receptor repertoire, including 218 genes encoding SRCR-domain-containing proteins (Pancer et al. 1999, Pancer 2000, Rast et al. 2008). In cnidarians, SRs have been identified in searches of the genomic data for *Nematostella vectensis* (Putnam et al. 2007), *Acropora* sp. (Schwarz et al. 2008) and the reef building coral, *Pseudodiploria strigosa* (Ocampo et al. 2015). Two functional transcriptomic studies examined differential expression of genes between symbiotic and aposymbiotic individuals of the sea anemones *Anthopleura elegantissima* (Rodriguez-Lanetty 2006) and *Aiptasia* sp. (Lehnert et al. 2012). Both studies found an SRB1-like protein to be upregulated in the symbiotic state. However, to date, no comprehensive characterisation of the cnidarian SR repertoire has been completed. The cnidarian SR repertoire represents one of many potential surface receptors involved in the recognition and entry of symbionts into host gastrodermal cells.

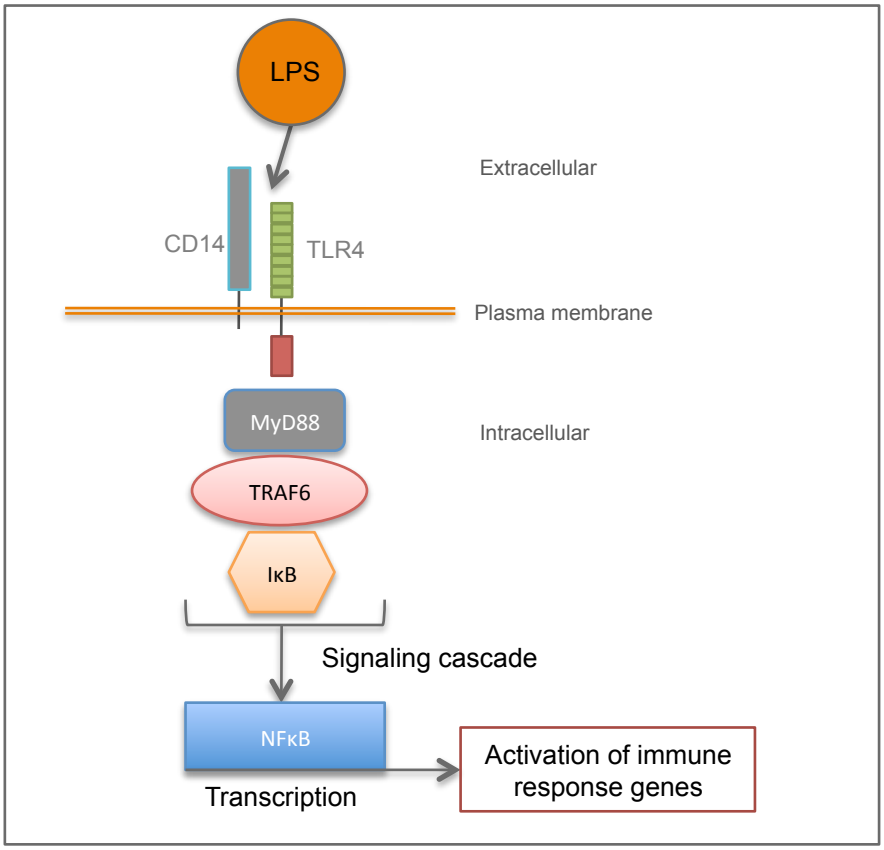



Figure 1.3: Signalling pathway of toll-like receptors (TLRs). The pattern recognition receptors (PRRs) such as the complement receptors, CD14 and TLR4, work in combination to respond to microbe associated molecular patterns (MAMPs) such as lipopolysaccharide (LPS). The TLR signalling cascade leads to the activation of NFκB and the activation of immune response genes. (Takeda et al. 2003).

### *1.4.4 Toll-like receptors*

Toll-like receptors (TLRs) are an evolutionarily conserved family of pattern recognition receptors. TLRs are sensing receptors, some of which are expressed on the surface (e.g. TLR4), while others are vacuolar (e.g. TLR9).. The TLR4 signalling cascade involves up to nine signalling molecules (main molecules shown in Figure 1.5), and leads to the activation of the major immunity gatekeeper transcription factor, NFκB. NFκB migrates into the cell nucleus and activates a wide variety of genes that instigate an immune response, such as the expression of nitric oxide synthase (NOS), which triggers the production of cytotoxic nitric oxide (NO) and pro-apoptotic genes (Chang et al. 2002, Takeda et al. 2003, Takeuchi and Akira 2010). TLR pathway components have been characterised in cnidarians: five putative TLRs and all other pathway components were identified in the non-symbiotic anemone *Nematostella vectensis*, and various

components have been identified in *Acropora* sp., *Hydra* sp., *Fungia scutaria* and *Aiptasia* sp. (Miller et al. 2007, Schnitzler 2010). The production of NO in response to multiple stressors, including elevated temperature, photosynthetic inhibitors and LPS has been demonstrated in functional experiments in the anemone *Aiptasia* sp. (Perez and Weis 2006, Detournay et al. 2012, Hawkins et al. 2013). The production of NO is associated with the immune response triggered during cnidarian (coral) bleaching, which is a collapse or dysfunction of the symbiosis at the cellular level as a response to environmental stressors such as increased temperature and solar radiation (Douglas 2003, Hughes et al. 2003, Hawkins 2015).

### *1.4.5 TSR and TRAP proteins*

The thrombospondin structural homology repeat, or TSR domain, was originally characterized in the human thrombospondin-1 protein (TSP1). Protein-protein interactions involving the TSR domain are central to TSP1 protein function, and the discovery of multiple invertebrate and vertebrate TSR-domain-containing proteins has raised questions about the importance and shared functions of this highly conserved domain (Adams and Tucker, 2000). All TSR-domain-containing proteins have one or more TSR domains, similar to the three repeats originally characterized in the human TSP1 protein (Tucker 2004). The TSR domain consists of approximately 60 amino acids, with several highly conserved motifs. TSR domains contain 6 conserved cysteine residues, forming three disulphide bridges with the exception of a few malaria thrombospondin-related anonymous proteins (TRAP's) and human complement factor, which contain five cysteines.

Several apicomplexan parasites use a TSR-domain-containing protein to assist in gliding motility, recognition and to gain entry to host cells (Kappe et al. 1999, Vaughan et al. 2008, Morahan et al. 2009). The involvement of a TSR-domain in recognition and entry of an apicomplexan parasite into host cells makes the TSR domain a target of interest in the study of cnidarian-dinoflagellate symbiosis. Dinoflagellates are a sister taxon to the apicomplexans within the infrakingdom Alveolata (Baldauf 2003, Saldarriaga et al. 2004, Janouskovec et al. 2010a), and might therefore share mechanisms of host invasion (Schwarz 2008).

A recent investigation into the evolution of thrombospondin proteins revealed that

cnidarians do not possess a classic TSP1 gene; the thrombospondin-like proteins in cnidarians lack TSR domains (Bentley and Adams 2010). However, TSR-domain-containing proteins were identified in all ten cnidarian species searched in a recent study of candidate symbiosis related genes in all available cnidarian sequence resources (Meyer and Weis 2012). A rhamnospondin gene with eight TSR domain repeats was identified in the colonial hydroid *Hydractinia symbiolongicarpus* (López et al. 2011), and several TSR domain-containing proteins were identified in two species of corals, *Acropora palmata* and *Montastraea faveolata.* (Schwarz et al. 2008).

## 1.5 Parasitic behaviour – links to the Apicomplexa

The exact role of innate immune receptors in the onset and maintenance of the cnidarian-dinoflagellate symbiosis is only just starting to be understood. However, the cellular process of invasion of several parasitic Protozoa is well studied and understood, particularly for members of the phylum Apicomplexa (Sibley 2004). This group is comprised entirely of obligate intracellular parasites. Parasitic protozoans are important causes of infectious disease globally, and are responsible for prolonged and chronic mammalian infections (Sacks and Sher 2002). Of these, three apicomplexan parasites of medical or agricultural importance have been investigated in detail: *Plasmodium* spp (the causative agent of human malaria), *Toxoplasma gondii* and *Cryptosporidium* sp.

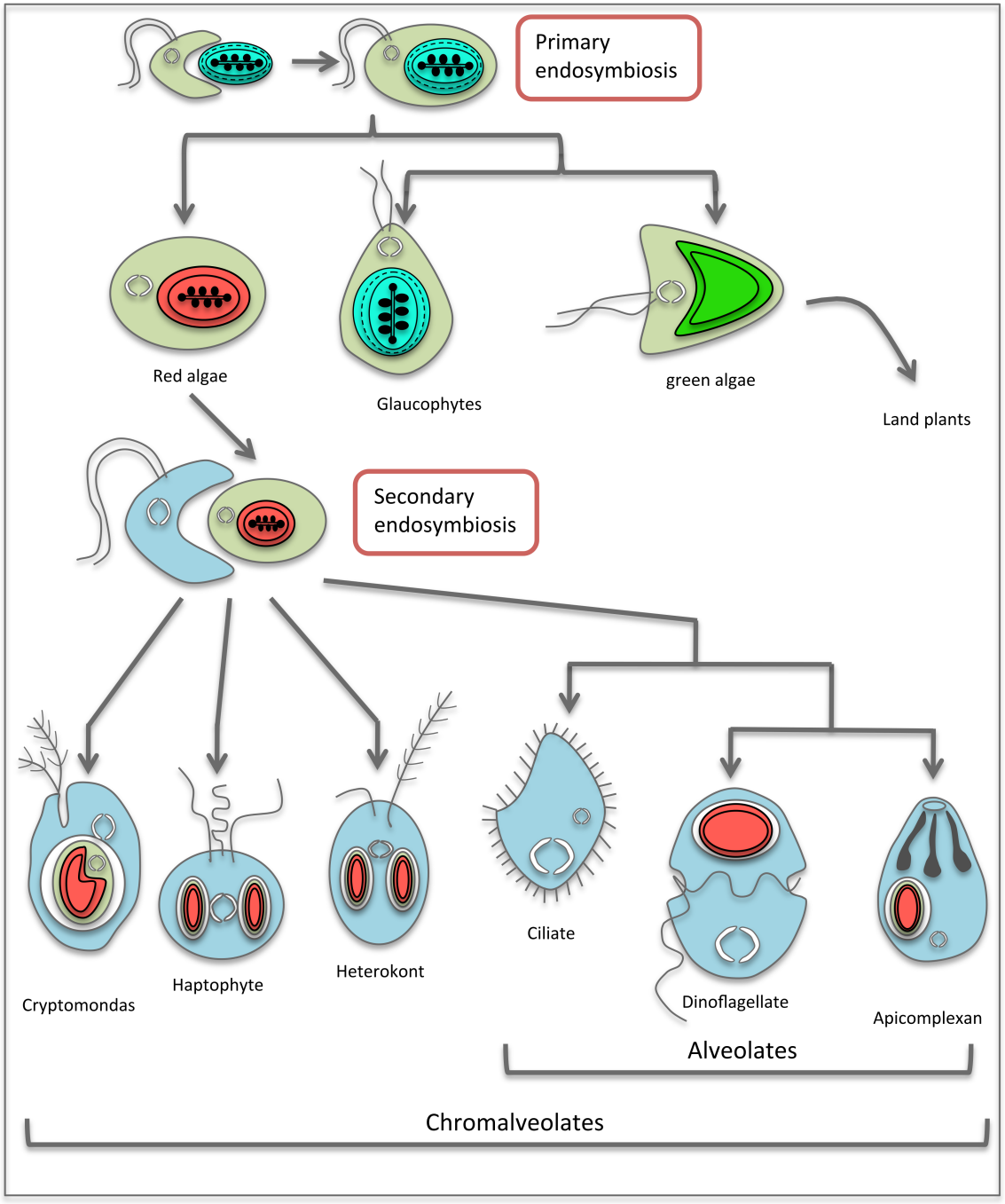



Figure 1.4: A schema for the origin and evolution of plastids by primary and secondary endosymbiosis, following the fate of the red algal lineage. A single primary endosymbiosis between an unknown heterotrophic eukaryote (green) and a cyanobacterium led to the three primary-plastid-bearing lineages (top). A secondary endosymbiosis between a red alga and a heterotrophic host led to all red plastid containing eukaryotic algae (blue). Loss of photosynthesis is pervasive in several of these lineages and, in the ciliates, the entire lineage is non-photosynthetic. Dinoflagellates are a sister taxon to the Apicomplexa. Adapted from Archibald et al. (2002), Keeling (2004) and Keeling (2013).

The apicpomplexans are a sister taxon to the dinoflagellates within the infrakingdom Alveolata (Baldauf 2003, Janouskovec et al. 2010a). Apicomplexans have a cryptic (non-

photosynthetic) plastid discovered in the malaria parasite, *Plasmodium* sp. (McFadden et al. 1996). Whole plastid genome phylogenies provide evidence that the plastid found in both apicomplexans and dinoflagellates originates from a common red algal endosymbiont ancestor to both (Janouskovec et al. 2010a). This secondary endosymbiosis can be described by the event of a phototrophic eukaryote containing a primary red plastid (a red algal primary endosymbiont) being phagocytosed but not digested by a heterotrophic eukaryote (Archibald and Keeling 2002, Keeling 2004, 2013)(Figure 1.6). In apicomplexans, the cryptic plastid is called an apicoplast and in *Plasmodium* sp. its prokaryotic cellular pathways have become important targets for drug therapy against the parasite (Foth and McFadden 2003).

A close photosynthetic relative to apicomplexan parasites has been described and provides a new model with which to study the evolution of parasitism within the alveolates (Moore et al. 2008, Okamoto and McFadden 2008). Interestingly, this new photosynthetic relative to the Apicomplexa, *Chromera velia*, is also a coral symbiont (Moore et al. 2008). Within the alveolates, symbiosis and parasitism are thus widespread, and indeed, these protists have a long history of interaction with animal immunity. Due to the taxonomic proximity of dinoflagellates and apicomplexan parasites, it may therefore be possible to draw comparisons between the mechanisms of host invasion and host microbe signalling with those operating at the onset and establishment of cnidarian-dinoflagellate symbiosis (Sacks and Sher 2002, Schwarz 2008, Areschoug and Gordon 2009).

Table 1.1: Comparison of molecules and mechanisms used by two apicomplexan intracellular parasites, a bacterium and *Symbiodinium* sp. Adapted from Schwarz et al. (2008).

| Microbe | *Toxoplasma* | *Plasmodium* | *Mycobacteria* | *Symbiodinium* |
|---|---|---|---|---|
| Site of host-microbe contact | Gastrointestinal tract | Blood and liver via blood or lymphatic system | Gastrointestinal tract and lung | Gastrovascular cavity |
| Active invasion | Active invasion | Active invasion | Host-mediated phagocytosis | Host-mediated phagocytosis |
| Host target cell type | All nucleated cells | Merozoites target erythrocytes<br>Sporozoites target Hepatocytes | Phagocytes within macrophage cells | Phagocytes within gastrodermal cells |
| Recognition/adhesion molecules | Glycosaminoglycans<br>Sialic acid | SRB1 + CD81 | Complement C3<br>SRB1 | Lectins<br>SRB1? |
| Intracellular niche | Non-phagosome vacuole | Non-phagosome vacuole | Modified phagosome | Modified phagosome |
| Manipulation of host response | Manipulation of host signalling pathways | Manipulation of host signalling pathways | Phagosome maturation arrest | Phagosome maturation arrest |

In the case of host invasion, two general strategies are observed, active penetration and host-mediated phagocytosis (Table 1.1). For example, both the apicomplexan parasites, *Toxoplasma gondii* and *Plasmodium* spp., enter the host by active penetration. *T. gondii* parasites attach to the outside of the host cell, and there is subsequent invagination of the host-cell membrane. This invagination of the membrane is selective, excluding specific host trans-membrane proteins by means of a moving junction, and creates a specialised compartment within the host cell referred to as the parasitophorus vacuole (PV) (Mordue et al. 1999). The PV is a non-phagosomal compartment, providing a safe intracellular environment in which the parasite can divide and grow, eventually killing the host cell by lysis and invading a new host cell. The malaria parasite, *Plasmodium falparcium*, uses a transmembrane protein on the parasite's cell surface, the thrombospondin-related anonymous protein (TRAP), which binds to the CD36 domain of SRB1 and to the tetraspanin CD81 on host hepatocyte cells (Ejigiri and Sinnis 2009). CD36 aids formation of a tight junction, whereby selective invagination of the host membrane forms the PV membrane (PVM), which is composed of both host- and parasite-derived proteins, and bears no resemblance to the membrane of host phagolysosomes or to that of digestive vacuoles (Lingelbach and Joiner 1998). The PVM

allows complex signalling between host and parasite, and manipulation of the host immune response (Musumeci et al. 2003).

*Symbiodinium* sp. uses the innate phagocytic mechanisms of cnidarian host digestive cells to gain entry (McNeil 1981, McNeil et al. 1981). In relatively simple organisms, such as cnidarians, digestive cells can also act like vertebrate macrophages and trigger innate immune pathways, via phagolysosome maturation (McNeil 1983, McNeil and McAuley 1984). Intracellular parasites, commonly manipulate host defence mechanisms to their own advantage, making use of innate immune phagocytic mechanisms to gain entry to host cells (Dietrich and Doherty 2009, Flannagan et al. 2009). For example, *Mycobacterium tuberculosis*, the bacterium causing tuberculosis, makes use of phagocytes within macrophages. More specifically, it arrests phagosome maturation by interfering with the Rab-mediated progression of endosomal interaction with lysosomes that deliver proteases and cause acidification within the phagolysosome (Dietrich and Doherty 2009). By halting phagosome maturation at the early endosome Rab5-mediated stage, the pathogen is able to persist (Schwarz 2008). *Symbiodinium* sp. enters the cnidarian host *via* ingestion and is phagocytosed into host gastrodermal cells. Invagination of the host membrane forms the host-derived symbiosome vacuole (Muscatine et al. 1975, Roth & Stacey 1989), and phago-lysosome maturation is manipulated *via* the same method of excluding specific Rab proteins from the symbiosome membrane (Chen et al. 2003b, Fransolet 2012).

Both the symbiosome and the PV avoid fusion with the host endolysomal system. By manipulating the 'alien'-containing vacuole-membrane components, both apicomplexans and dinoflagellates alter cell-signalling pathways and evade or halt innate immune responses from the host cell. The vertebrate immune response is undoubtedly more complex than the immune response provided by the cnidarian host. However, macrophage innate immune receptors (depicted in Figure 1.4) such as scavenger receptors (SRs), C-type lectins and toll-like receptors all have homologs identified within cnidarian models (Rodriguez-Lanetty et al. 2006, Miller et al. 2007, Wood-Charlson and Weis 2009, Lehnert et al. 2014).

The scavenger receptor class B type 1 (SRB1), which is upregulated in symbiotic individuals of the sea anemone *Anthopleura elegantissima* (Rodriguez-Lanetty et al.

2006), is known to facilitate infection of hepatocyte cells by *Plasmodium* (Rodrigues et al. 2008, Yalaoui et al. 2008). In particular, the inhibition of SRB1 function reduced *Plasmodium berghei* infection in mouse hepatocytes, and additional analyses revealed that SRB1 plays a dual role in *Plasmodium* infection, affecting both sporozoite invasion and intracellular parasite development (Rodrigues et al. 2008). Another study used SRB1-knockout, SRB1-hypomorphic and SRB1-transgenic primary hepatocytes, as well as specific SRB1-blocking antibodies, to show that SRB1 significantly boosts hepatocyte permissiveness to *Plasmodium* entry and promotes parasite development (Yalaoui et al. 2008). Over-expression of SRB1 produced a conformational change in the plasma membrane, making it easier for malaria sporozoites to attach and gain entry. Given the taxonomic relatedness of *Symbiodinium* and members of the Apicomplexa, it is conceivable that *Symbiodinium* makes use of similar pathways to gain entry and establish itself within cnidarian host cells. SRB1 is also implicated in mycobacterial recognition, however the pathogen employs multiple adhesion molecules and SRB1 is not essential for successful pathogen entry (Schäfer et al. 2009).

While many microbial pathogens avoid host recognition or dampen immune activation in the host via sophisticated signalling, some pathogens benefit from the stimulation of host immune pathways (Medzhitov 2007). A good example of this is the TGFβ immune pathway, which is largely responsible for inducing tolerance and containing inflammation (Li et al. 2006). In general, most macrophage pathogens have evolved mechanisms to induce the production of TGFβ protein production as a mechanism for supressing the macrophage inflammatory response and thus enhance pathogen proliferation (Ming et al. 1995). Experimental evidence suggests that there is a tolerogenic response to the symbiont *via* activation of the TGFβ immune suppressive pathway (Detournay et al. 2012). Blocking the putative cnidarian TGFβ protein reduced infection success in re-infection experiments using aposymbiotic *Aiptasia* sp. anemones. Blocking TGFβ ligands also induced an immune response in symbiotic anemones in the form of increased NO production (Detournay et al. 2012). In comparison to aposymbiotic *Aiptasia* sp., symbiotic anemones produce significantly less NO in response to immune stimulation by LPS (Detournay et al. 2012). This response is mirrored in vertebrate macrophages parasitized by *Leishmania major* and *Toxoplasma gondii*, where NO production is suppressed due to the tolerogenic effects of the TGFβ pathway (Li et al. 2006).

## 1.6 *Aiptasia* sp. as a model organism

The field of cnidarian research has historically lacked a clear single model organism that enables the study of all aspects of the cnidarian-dinoflagellate symbiosis. Corals are notoriously difficult to keep in aquaria, however anemones lack the skeleton of hermatypic (reef building) corals. The symbiotic Indo-Pacific branching coral *Acropora millepora* is the preferred model species for the study of the coral skeleton and tissue interface, as it is relatively easy to collect branch fragments and maintain them in laboratory aquaria (Weis et al. 2008). However, this species is not ideal for symbiosis studies as it cannot be rendered aposymbiotic (cleared of algal symbionts) and specimens re-infected under laboratory conditions do not survive. Using *A. millepora* also involves taking coral fragments from living reefs. For these reasons, the glass anemone *(Aiptasia* sp.) has been proposed as a model organism for studying the cnidarian-dinoflagellate symbiosis (Weis et al. 2008). The functional work presented in this thesis is all performed on this model organism.

*Aiptasia* sp. is a common pest species colonising marine tropical aquaria, producing large clonal populations that are difficult to remove (Figure 1.7a). Due to its robust nature, *Aiptasia* sp. is a good species for laboratory culture, where it is easily maintained in small aquaria. Symbionts can be removed completely by cold shock, and aposymbiotic anemones can be maintained and re-infected with a variety of *Symbiodinium* strains (Schoenberg and Trench 1980, Lin et al. 2000, Belda-Baillie et al. 2002, Perez and Weis 2008, Starzak et al. 2014) (Figure 1.7b). Access to both symbiotic and aposymbiotic animals allows for the study of symbiosis onset (Figure 1.7 c and d).

*Aiptasia* sp. is dioecious (of both sexes), and reproduction is both sexual and asexual. As with corals, gametes are released into the water column, where fertilisation occurs and planula larvae develop. Asexual reproduction occurs by basal or pedal laceration, whereby small pieces of tissue bud from the base of the anemone where it attaches to the substrate. These small buds metamorphose into tiny clones of the original anemone. During times of extreme stress, such as low light, disturbance, heat stress and low oxygen, *Aiptasia* sp. increases asexual reproduction by pedal laceration (Belda-Baillie et al. 2002).

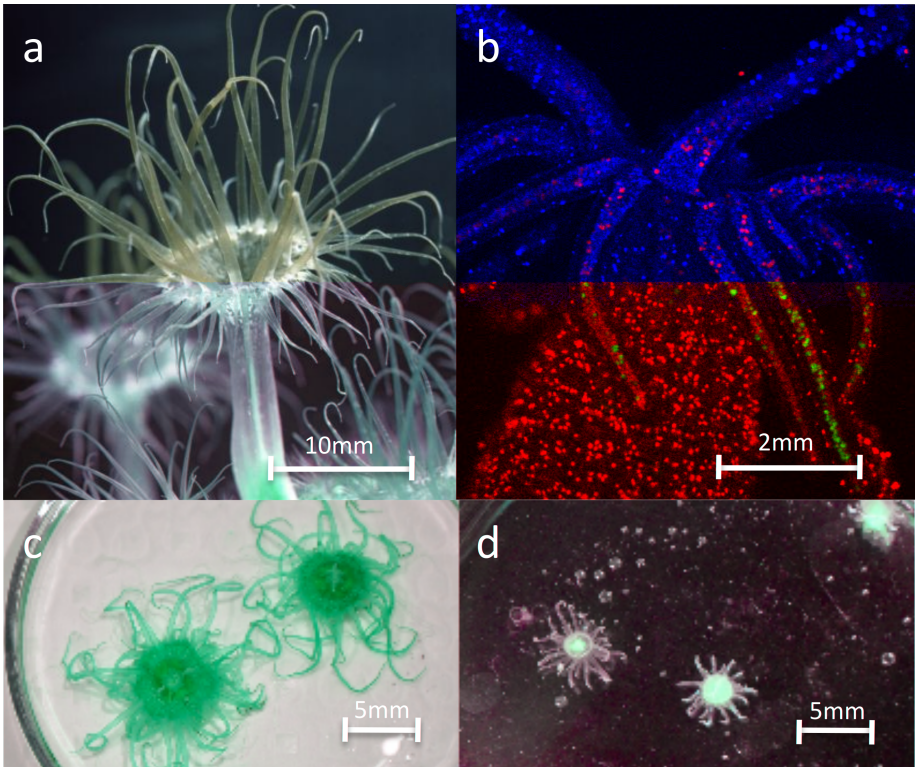



Figure 1.5: (a) The model anemone *Aiptasia* sp. is a considered a pest by tropical marine aquarium enthusiasts, as it grows quickly by asexual reproduction (Image: Dan Logan). (b) Confocal microscope image of whole *Aiptasia* sp. anemone. DAPI (blue) stains the nuclei of both anemone and *Symbiodinium* cells; red chloroplast autoflorescence labels healthy dinoflagellate cells within the tentacle gastrodermal layer during the early stages of infection (Image: Dan Logan). (c) *Aiptasia* sp. in the symbiotic (c) and aposymbiotic (d) states (Images: Tom Hawkins).

*Aiptasia* sp. has been used experimentally to examine: recognition and specificity (Schoenberg and Trench 1980, Lin et al. 2000, Belda-Baillie et al. 2002, Starzak et al. 2014), stress (Perez and Weis 2006, Dunn et al. 2007a), gene expression (Kuo et al. 2004, Dunn et al. 2006, 2007b, Lehnert et al. 2014), the biology of the symbiont (Glider et al. 1980, Steen 1986, Lesser and Shick 1989), symbiosis maintenance and cellular alteration (Chen et al. 2003b, 2003a, 2004, 2005, Detournay and Weis 2011, Detournay et al. 2012), and nutrition (Wang and Douglas 1997, Muller-Parker 2001). In addition, a transcriptome for *A. pallida* is available (Lehnert et al., 2012; Lehnert et al., 2014), and genome sequencing was completed during the writing of this thesis (Baumgarten et al. 2015). Figure 1.8 shows confocal microscope images (a and b) of symbionts within the tentacles of *Aiptasia* sp, and a light microscope image of dinoflagellate cells in laboratory culture (c).

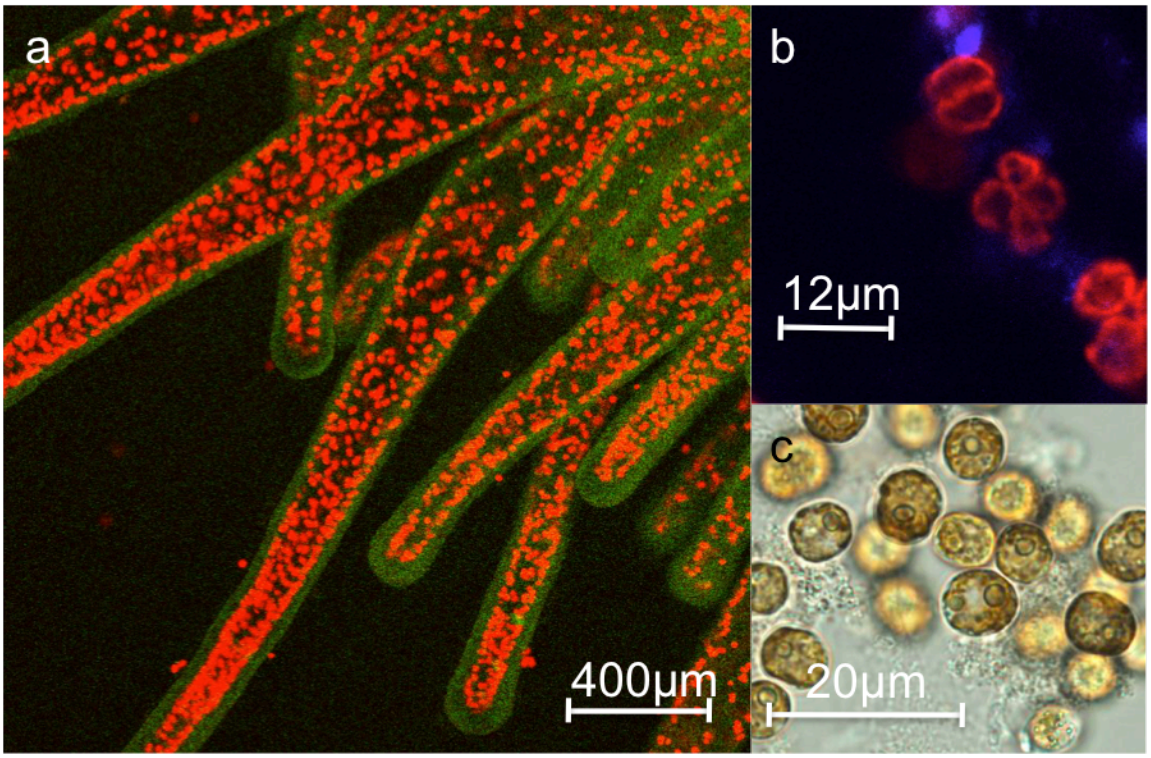



Figure 1.6: (a) Confocal microscope image of a tentacle squash of the symbiotic anemone *Aiptasia* sp. Dinoflagellate symbionts are clearly visible by the red chlorophyll autoflorescence (Image: Dan Logan). (b) Confocal microscope image of proliferation of dinoflagellate cells within an *Aiptasia* sp. tentacle (Image: Emilie Neubauer). A single host gastrodermal cell may contain up to four dinoflagellate cells. (c) Light microscope image of *Symbiodinium* cells in culture (Image: Tom Hawkins).

## 1.7 Available sequence resources for cnidarians

Due to the lack of a clear cnidarian model organism and the relatively small field of cnidarian researchers, cnidarian sequence resources have lagged behind those for all other taxa, including sponges. Until very recently there was no whole sequenced genome of a symbiotic cnidarian available. The combination of a general cooperative approach and next generation sequencing (NGS) technologies becoming both easier and cheaper has led to a recent sharp increase in available sequence data for cnidarian species. The newly available data have provided much needed comparative possibilities for the study of the cnidarian-dinoflagellate symbiosis. The first publically available, sequenced anthozoan genome was for the non-symbiotic brackish water anemone, *Nematostella vectensis* (Putnam et al. 2007). The *N. vectensis* genome provided insight into the genetic complexities of cnidarians, and several studies revealed vertebrate

signalling pathways to be conserved in cnidarians (Ryan et al. 2006, Sullivan et al. 2006, Putnam et al. 2007, Dunn et al. 2008, Wood-Charlson and Weis 2009). Recently a symbiotic anthozoan genome has become available, for the reef building coral *Acropora digitifera* (Shinzato et al. 2011). Comparison of *N. vectensis* and *A. digitifera* genomes has revealed putative symbiosis-associated genes that may function in the onset and maintenance of cnidarian-dinoflagellate symbiosis (Meyer and Weis 2012).

The majority of genetic data projects over the last decade have produced transcriptome data rather than full genomes. Transcriptomes generated using next-generation sequencing (expressed sequence tags (ESTs), 454 pyrosequencing and Illumina HiSeq technologies), have been published for 12 anthozoan species: *Montastraea faveolata* (Schwarz et al. 2008) *Acropora millepora* (Schwarz et al. 2008, Meyer et al. 2009, Moya et al. 2012); *Acropora palmata* (Polato et al. 2011); *Pocillopora damicornis* (Traylor-Knowles et al. 2011, Vidal-Dupiol et al. 2013); *Porites astreoides* (Kenkel et al. 2013); *Aiptasia pallida* (Lehnert et al. 2012); *Porites australiensis* (Shinzato et al. 2014a); *Corallium rubrum* (Pratlong et al. 2015); *Fungia scutaria*, *Montastraea cavernosa*, *Seriatopora hystrix* and *Anthopleura elegantissima* (Kitchen et al. 2015). These resources represent various developmental stages and symbiotic states, which limits the study of interplay between the two partners. Indeed, several studies have shown differential gene expression when comparing transcriptome data from symbiotic *versus* aposymbiotic anemone tissues (Rodriguez-Lanetty et al. 2006, Moya et al. 2012, Lehnert et al. 2014).

While these data are valuable, they can only provide a snapshot of the proteins expressed at the time the transcriptome was made. Comparison of specific proteins found between species for which only transcriptome data are available is problematic, as they simply may not be expressed at the time of sequencing. The amount of available cnidarian sequence data is increasing exponentially, thereby increasing the number and diversity of anthozoan resources for comparative, phylogenetic and functional analyses. During the completion of this thesis, the genome for the symbiotic anemone *Aiptasia* sp. has been published (Baumgarten et al. 2015), this allows a direct comparison of genome and transcriptome data.

## 1.8 Aims and scope of this study

This thesis aimed to unravel certain aspects of the cellular mechanisms underlying the establishment and maintenance of the cnidarian-dinoflagellate symbiosis, learning from the cellular mechanisms employed by various parasitic Protozoa. By taking a model-systems approach with *Aiptasia* sp, I aimed to contribute to the global effort to advance our understanding of the cell biology of this ecologically important symbiosis. My work combined cell biology and microscopy techniques with molecular techniques, as well as immunological manipulation and bioinformatics.

The three experimental chapters have the following objectives and hypotheses:

Chapter 2:

- Objective 1: Characterise the scavenger receptor repertoire in six cnidarian species using available sequence resources and bioinformatics techniques.
- Objective 2: Examine the role of SRs in symbiont uptake and recognition using functional experiments with *Aiptasia* sp.

  Hypotheses: a) SRs are involved in symbiont recognition and uptake; b) blocking SRs will induce an immune response in the host, indicating SR involvement in modulation of a host immune response to the symbiont.

Chapter 3:

- Objective 1: Characterise the TSR-domain-containing protein repertoire in six cnidarian species and the dinoflagellate *Symbiodinium minutum*, using available sequence resources and bioinformatics techniques.
- Objective 2: Compare and contrast vertebrate TSR-domain-containing proteins of known function with the cnidarian and *Symbiodinium minutum* TSR repertoires.

- Objective 3: Establish the presence of known binding motifs and their conservation within the cnidarian TSR-domains.

  Hypotheses: a) Cnidarians posses multiple TSR-domain-containing proteins; b) Cnidarian TSR domains contain conserved binding motifs for CD36/SRB1 and Glycosaminoglycans.

Chapter 4:

- Objective: Establish functional evidence that TSR-domain-containing proteins are involved in the regulation of the cnidarian–dinoflagellate symbiosis.

  Hypotheses: a) A TSR-domain-containing protein is essential for successful symbiont entry to host cells; b) a TSR-domain-containing protein provides a ligand for SRB1 proteins upregulated in the cnidarian symbiotic state; c) a TSR-domain-containing protein is involved in initiation of the TGFβ pathway in cnidarians.

This thesis draws on the knowledge-base of parasite research, which provides a focus and suggests potentially important receptor-ligand signalling pathways for host entry and intracellular persistence. By studying specific aspects of innate immune pathways I will draw parallels between known mechanisms of parasite invasion and symbiont acquisition, thereby providing further evidence for shared strategies among parasites and mutualistic symbionts.

# Chapter 2

# The scavenger receptor repertoire in six cnidarian species indicates a potential role in the establishment of cnidarian-dinoflagellate symbiosis.

## 2.1 Introduction

The scavenger receptor 'superfamily' comprises a large family of structurally diverse transmembrane cell surface glycoproteins. Scavenger receptors (SRs) are named for their role in scavenging and clearing of modified host molecules, microbial invaders and apoptotic cell debris, and play an important role in innate immune defense by acting as pattern recognition receptors (PRRs) against microbial pathogens (Yamada et al. 1998, Areschoug and Gordon 2009). The first SR was discovered in cholesterol experiments in 1979 (Goldstein et al. 1979). Vertebrate SRs were originally defined functionally by their ability to bind modified low-density lipoproteins (mLDL), but not native LDL. However, it is now known that all SRs have a high affinity for a wide range of ligands. This flexibility of binding to potential ligands has led SRs to be described as 'molecular fly paper' (Krieger 1992). The most commonly described function for SRs is to act as phagocytic receptors mediating direct non-opsonic phagocytosis of pathogenic microbes (Areschoug and Gordon 2009). Unlike complement pathway receptors, SRs recognise microbe associated molecular patterns (MAMPs) on the microbe surface without the need for opsins such as C3 to first bind the microbe and aid detection.

SRs exist in multiple Classes with overlapping specificities that allow for the recognition of many microbial- and pathogen-associated molecular patterns (MAMPs and PAMPs) (Krieger 1997). This extensive binding capability is due not only to the wide range of proteins included in the SR superfamily, but also to their capacity to partner with various co-receptors (Canton 2013). Similar to other innate immune receptors, it is likely the SRs arose early in evolution to recognize a multitude of endogenous and exogenous structures (Krieger 2001). Recent studies show that several pathogens have

evolved mechanisms to evade SR-mediated recognition (Areschoug et al. 2008, Faure and Rabourdin-Combe 2011). Indeed, several human pathogens exploit specific SRs for their own benefit, using their phagocytic abilities to gain entry to host cells. Both the Hepatitis C virus (HCV) (Catanese et al. 2007) and sporozoites from the malaria parasite (*Plasmodium falciparum)* (Ndungu et al. 2005, Rodrigues et al. 2008) have surface ligands that are recognized by the scavenger receptor SRB1, and use this method to gain entry to human hepatocyte cells.

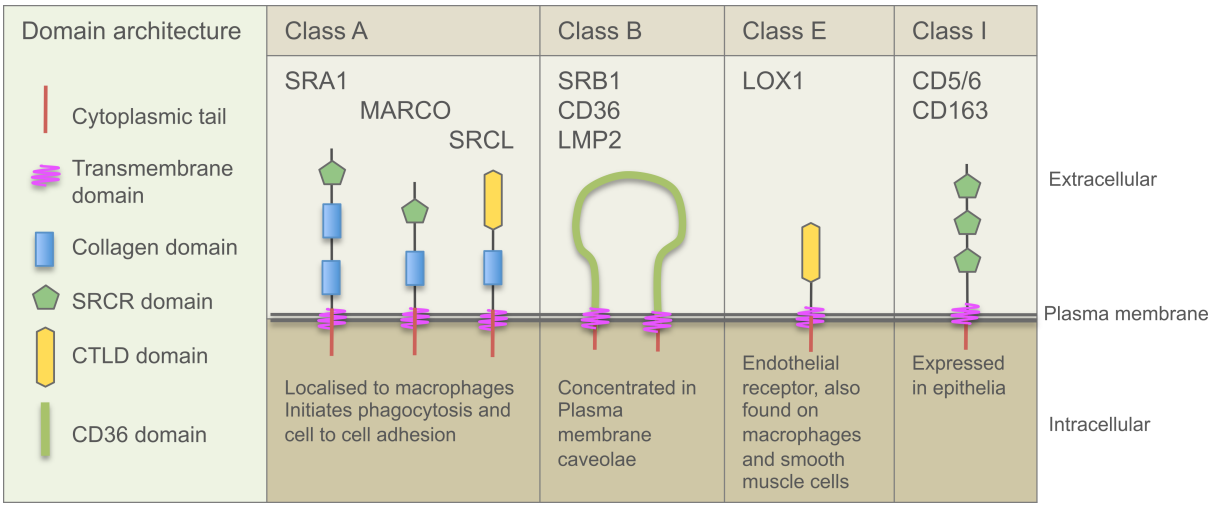



Figure 2.1: Domain architecture of vertebrate SRs. Class A mammalian SRs are characterized by having one or more collagen domains. Class B SRs are defined by having two cytoplasmic tails either side of the CD36 domain (there is an SRB protein also named CD36), Class E SRs have a C-type lectin domain (CTLD). Class I SRs have multiple copies of the SRCR and no other identifiable domains. SRCR, scavenger receptor cysteine-rich domain; CTLD, C type lectin domain; MARCO, macrophage receptor with collagenous structure (also known as SCARA2 and SR-A2); SRCL, scavenger receptor with C-type lectin (also known as SCARA4 and CLP1); LOX1, lectin-like oxidized LDL receptor 1; LMP2, lysosomal integral membrane protein.

Membrane associated SRs are divided into nine Classes, labeled SR-A to -I, although Class C scavengers are only found in the fruitfly, *Drosophila melanogaster* (Canton 2013). Figure 2.1 provides a diagrammatic representation of four of the nine SR classes. Members within a given Class share some sequence similarity, though among classes there is little similarity of domain structure; SRs are multi-domained, and no single domain is common to all (Gough and Gordon 2000, Gordon 2002). SRs are found in the extracellular matrix (ECM), anchored in the cell membrane with a cytoplasmic tail, which is short and does not contain any identifiable protein domains. The extracellular

domain contains a myriad of receptor domains. This study focuses on three common SR domains, the SRCR domain, the CTLD, and the CD36 domain.

Class A and I SRs contain the scavenger receptor cysteine rich (SRCR) domain, which consists of a 110 amino acid residue motif with conserved spacing of six to eight cysteines that are involved in intradomain disulphide bridges (Hohenester et al. 1999). The SRCR domain is an ancient and conserved protein domain that often occurs in repeats in membrane and soluble proteins. In vertebrate systems, proposed functions include acting as phagocytic receptors, epithelial homeostasis, cell aggregation and immunity (Yamada et al. 1998, Janeway and Medzhitov 2002, Mukhopadhyay et al. 2004, Sarrias et al. 2004, Martinez et al. 2011, Whelan et al. 2012). SRCR-domain-containing proteins are also important in pathogenesis and diseases such as atherosclerosis, autoimmune diseases and cancer (Martinez et al. 2011). Vertebrate SR Classes A and E contain C-type lectin domains (CTLDs), which are involved in lectin-glycan interactions. Vertebrate Class B SRs consist of the CD36 domain, named after the CD36 protein. Class B SRs have two cytoplasmic tails rooted in the membrane, forming a loop in the ECM. SR genes coding for all three of these domains - SRCR, CTLDs and CD36 - have also been described in invertebrates (Pancer et al. 1997, Hibino et al. 2006, Schwarz et al. 2007, Wood-Charlson and Weis 2009, Lehnert et al. 2014). However, little is known about the functions of these SR domains, and in particular the CD36 domain, in invertebrates, including cnidarians.

Cnidarians represent a basal metazoan phylum of invertebrates that includes sea anemones, corals, jellyfish and hydroids. Cnidarians regularly form an intracellular symbiosis with photosynthetic dinoflagellates of the genus *Symbiodinium*. The relationship between reef-building corals and dinoflagellate endosymbionts provides the trophic and structural foundation of the coral reef ecosystem, supporting immense biodiversity. The symbiont resides within a host-derived vacuole within host gastrodermal cells. The relationship is based on mutualistic nutrient exchange, where the dinoflagellate provides fixed carbon to the cnidarian host in return for inorganic nutrients and a high light environment safe from predation (Yellowlees et al. 2008). In contrast to our knowledge of the mechanisms employed by vertebrate parasites and

pathogens to invade host cells, there is limited research and understanding of how the cnidarian-dinoflagellate symbiosis is established and maintained.

There is strong evidence to suggest that the establishment of a steady state cnidarian-dinoflagellate symbiotic relationship involves a number of complex steps: recognition, phagocytosis, selection, arrest of phago-lysosome maturation, and control of the cell cycle and symbiont proliferation (Davy et al. 2012). Dinoflagellate symbionts are most commonly obtained from the environment during host early life history stages rather than directly via the gametes. Symbionts are taken in through the host mouth and are subsequently taken up by gastrodermal cells via phagocytosis (Schwarz et al. 1999). Initial recognition between cnidarian host and dinoflagellate symbiont, as well as symbiont uptake, are thought to involve a series of complex receptor-ligand interactions (Davy et al. 2012). Dinoflagellates somehow gain entry into host gastrodermal cells and avoid digestion once inside the host vacuole. It is not known how the cnidarian host manages the balance between beneficial microbes, such as their symbiotic dinoflagellates, and the maintenance of an immune response to pathogenic attack.

While the exact role of innate immune receptors in the onset and maintenance of the cnidarian-dinoflagellate symbiosis is only just starting to be understood, known invasion mechanisms from apicomplexan intracellular parasites may provide clues about mechanisms at play. The Apicomplexa comprise a well-studied group of unicellular eukaryote parasites such as *P. falciparum* and *Toxoplasma gondii*, which are the sister taxon to dinoflagellates within the kingdom Alveolata (Baldauf 2003, Janouskovec et al. 2010b). Due to the taxonomic proximity of dinoflagellates and apicomplexan parasites, it is possible to draw comparisons with the mechanisms operating at the onset and establishment of cnidarian-dinoflagellate symbiosis (Sacks and Sher 2002, Schwarz 2008, Areschoug and Gordon 2009). The host cell invasion strategies employed by apicomplexan parasites, such as the malaria parasite *P. falparcium*, are well studied and involve a unique process by which a transmembrane protein on the parasite's cell surface, the thrombospondin-related anonymous protein (TRAP), binds to the CD36 domain of SRB1 on host hepatocyte cells. SRB1 plays an important role in *Plasmodium* infection, as it promotes sporozoite invasion of hepatocyte cells and subsequent intracellular parasite development (Rodrigues et al.

2008). Thus, in malaria, the CD36 domain is targeted by the parasite and prevented from raising the alarm in host defense.

The cnidarian SR repertoire represents one of many potential surface receptors involved in the recognition and entry of symbionts into host gastrodermal cells. Two functional transcriptomic studies examined differential expression of genes in symbiotic compared to aposymbiotic individuals of the sea anemones *Anthopleura elegantissima* (Rodriguez-Lanetty 2006) and *Aiptasia* sp. (Lehnert et al. 2012). Both studies found an SRB1-like protein to be upregulated in the symbiotic state. In *Aiptasia* sp., the SR type B gene showed a 28-fold increase in expression between symbiotic and aposymbiotic anemone clones. This substantial increase in expression suggests that the SR Class B gene is involved in maintenance of the cnidarian-dinoflagellate symbiosis. The Class E SRs, which contain C-type lectin domains (CTLDs), are another potential SR ligand involved in recognition and uptake of dinoflagellate symbionts by cnidarian host cells. Studies of the make-up of the *Symbiodinium* cell surface glycome used florescent lectin probes to identify potential glycan-lectin ligands on freshly isolated *Symbiodinium* cells (Wood-Charlson et al. 2006, Logan et al. 2010). Another study identified 92 C-type lectins in the genome of the non-symbiotic sea anemone *Nematostella vectensis*, suggesting that a repertoire of lectins could be operating as MAMP-PRR interactions during establishment of a symbiosis. Two studies of coral species identified mannose-binding lectins with CTLDs: Millectin from *Acropora Millepora* (Kvennefors et al. 2008, 2010) and PdC lectin from *Pocillopora damicornis* (Vidal-Dupiol et al. 2009).

The aim of this chapter was to examine the repertoire of SRs in six cnidarian species using a variety of available genomic and transcriptomic resources. A complete genome is available for the symbiotic anthozoan *Acropora digitifera*, and transcriptome data are available for several more symbiotic cnidarian species (Kitchen et al. 2015). The description of the cnidarian SR repertoire, together with a comparison to vertebrate SRs of known function, will provide a platform for identifying potential roles of cnidarian SR proteins. These bioinformatic searches were paired with a set of simple functional experiments, which examined a role for SRs in symbiont recognition and uptake by host gastrodermal cells.

## 2.2 Materials and Methods

### *2.2.1 Cnidarian genomic and transcriptomic resources*

To characterize the SR protein repertoire in cnidarains, six anthozoan (coral, sea anemone) species with publically available genomic or transcriptomic resources were searched for SR proteins. These included three anemone species: *Anthopleura elegantissima* (Kitchen et al. 2015), *Aiptasia* sp. (Lehnert et al. 2012, Baumgarten et al. 2015) and *Nematostella vectensis* (Putnam et al. 2007), and three coral species: *Acropora digitifera* (Shinzato et al. 2011), *Acropora millepora* (Moya et al. 2012) and *Fungia scutaria* (Kitchen et al. 2015). These resources represent various developmental stages and symbiotic states (Table 2.1). All resources were used without manipulation, with the exception of the *Aiptasia* sp. transcriptome, for which raw Illumina sequence reads for accession SRR696721 were downloaded from the sequence read archive entry for the aposymbiotic CC7 transcriptome (www.ncbi.nlm.nih.gov/sra/SRX231866) and reassembled using Trinity (Grabherr et al., 2011).

### *2.2.2 SR sequence searching and verification*

Twenty-four non-cnidarian sequences were obtained, primarily from GenBank, for use in creating multiple sequence alignments and protein trees (accession numbers listed in Appendix A3). Eleven human SR genes were chosen to produce reference protein domain architecture diagrams, to compare predicted cnidarian proteins with human SR proteins of known function (Figure 2.2).

To search for cnidarian SR proteins, initial searches were performed in *N. vectensis*, through the Joint Genome Institute Genome portal (http://genome.jgi-psf.org/pages/blast.jsf?db=Nemve1). Keyword searches within the *N. vectensis* online genome portal were performed using the keywords: scavenger, SR, SRCR, CD36, and LMP2. In annotated genomes, keyword searches are comprehensive, and BLAST searches using mouse and human SR protein sequences indicated that all SR-like sequences had been identified through keyword searches. Consensus sequences (pfam01130: CD36, pfam00530: SRCR) from the conserved domain database (http://www.ncbi.nlm.nih.gov/cdd) (Marchler-Bauer et al., 2013) were used as queries in tBLASTn searches of each anthozoan resource. A high E-value cutoff (1x10-1) was

used in the tBLASTn searches to recover more divergent sequences. All BLAST searches were performed using Geneious pro version 5.4.3 (Drummond et al., 2011).

All databases were queried in the same way, using either BLASTp or tBLASTn searches using mouse and human SR protein sequences (SR-A1, MARCO, SRCL, CD36, SRB1/2, LMP2, and LOX1), combined with keyword searches (CD36, SRCR and scavenger) of the GO or KEGG annotation. To ensure that as many SR protein sequences as possible were recovered, representative *N. vectensis* sequences of each protein type (SRCR-domain-containing, CD36, SRB1, and LOX1) were also used as queries for tBLASTn searches of the other five anthozoan resources. To confirm that the sequences obtained contained SR domains, nucleotide sequences were translated using the program Geneious version 7.1.8 (http://www.geneious.com, (Kearse et al. 2012) and then annotated using the Geneious InterProScan plugin (Quevillon et al. 2005). InterProScan simultaneously searches nine different protein domain databases, only sequences where two or more databases found either SRCR, CD36 and CTLD domains and with an E- value of less than 1x10-4 were used. Where InterProScan was unable to resolve protein domains these sequences were used as query sequences for the online protein domain database PfamA (http://pfam.sanger.ac.uk) (Punta et al., 2012). Sequences for each species were aligned and those that were identical or almost identical (less than 5aa difference in the conserved domains) were omitted from the analysis as they likely represented artifacts of assembly issues or different isoforms of the same protein. Protein fragments missing a start or a stop codon were removed from the analysis.

To verify SR proteins, only proteins that showed significant pfam A matches to a SR domain or motif were included in the analysis. Proteins were annotated using the Geneious plugin InterProScan, using the Pfam A database, and then checked by eye. Diagrammatic representations of the various protein domain configurations were then produced using this information. Protein-domain architectures were grouped together according to common domains and compared to known human SR proteins (Figure 2.3).

CD36 domains were extracted and a multiple sequence alignment was performed with the MAFFT v 7.017 plug-in through Geneious, using the default settings (Katoh et al. 2002, Kearse et al. 2012). To choose the best-fit model of protein evolution, I used the program ProtTest v2.4 (Abascal et al. 2005) to apply AIC1, AIC2 and BIC2 model

selection criteria to a variety of possible substitution matrices and rate assumptions. The results from the overall comparison of these metrics indicated that the best-fit model for the full-length alignment was WAG+G+F (Whelan and Goldman 2001). A maximum likelihood tree was produced using FastTree v2.1.5. Bootstrap support values were generated using the online program SEQBOOT and values above 60-100% support were displayed in the tree as decimals at tree nodes.

### *2.2.3 Maintenance and preparation of anemone and dinoflagellate cultures*

Symbiotic *Aiptasia* sp. cultures were maintained in saltwater aquaria at 26°C with a 12/12 h light/dark photoperiod, and were fed twice weekly with live brine shrimp. Animals were rendered aposymbiotic with incubation for 8 h at 4 °C followed by maintenance in the dark. Anemones were fed daily with brine shrimp, and cleaned of expelled symbionts and food debris daily.
Cultured dinoflagellates, *Symbiodinium* sp., clade B1 (culture ID: CCMP830) were maintained in 50 ml flasks in sterile Guillard's f/2 enriched seawater culture medium (Sigma, St. Louis, MO, USA). Dinoflagellate cultures were maintained at 26°C on a 12/12 h light/dark photoperiod.

In preparation for experimental manipulations, individual anemones were placed in 24-cell well-plates in 2.5 ml of 1-µm filtered seawater (FSW) and acclimated to the well-plate for 3-4 days, with the water replaced daily. Well plates containing aposymbiotic anemones were exposed to as little light as possible and symbiotic anemones were maintained in an incubator at 26°C with a 12/12 h light/dark photoperiod. Animals were not fed during the experimental time period.

### *2.2.4 Addition of fucoidan to block SR binding function*

To explore a role for scavenger receptors in the onset of symbiosis, fucoidan, a known SR ligand, was added to anemones to block SR binding sites. Fucoidan is a protein derived from the brown alga *Fucus vesiculosus*; this polyanionic ligand is known to bind positively charged portions of Class A and B SR proteins, and block the binding of modified LDL (Hsu et al. 2001, Dinguirard and Yoshino 2006, Thelen et al. 2010).

To examine the effect of blocking SR binding capabilities on symbiont infection success aposymbiotic anemones (three anemones per treatment per time point) were pre-incubated in fucoidan (Sigma Eldrich #F563), at a concentration range of 0 (FSW control), 100, 200 and 400 µg/ml for 18 hours, according to Bowdish Lab protocols (online at McMaster University; www.bowdish.ca/lab/protocols). Fucoidan-treated aposymbiotic anemones were subsequently re-infected with *Symbiodinium* clade B1 (culture ID: CCMP830). *Symbiodinium* cell cultures were rinsed clean with FSW twice, re-suspended in FSW, and then added to well-plates (containing anemones) to a final concentration of 2 x $10^5$ symbionts per ml. After incubation for 12 h at 26°C in the light, anemones were rinsed twice with FSW and fucoidan treatments were refreshed. To test the effect of fucoidan on host health, a second control treatment (fucoidan-washed control), was prepared where aposymbiotic anemones were pre-incubated in 200 µg/ml fucoidan for 18 hours, and then washed clean with FSW prior to re-infection with symbionts as described above. Anemones for all treatments were sampled at 48 and 96 h post-infection (three tentacles per anemone, for n = 3 anemones per treatment per time point).

A second experiment examined the potential stress response to blocking SR binding capabilities in symbiotic anemones. The production of nitric oxide (NO) is known to indicate an immune or stress response in anemones (Perez and Weis 2006, Detournay and Weis 2011). If a symbiont is utilizing host SRs in order to initiate tolerogenic pathways (such as the TGFβ pathway) that dampen or prevent an immune or inflammatory response, blocking SR ligand binding capabilities should induce a stress response with the addition of lipopolysaccharide (LPS) to illicit an immune response. To induce a stress response, symbiotic anemones were incubated for 12 hours in FSW with 1 µg/ml LPS (Sigma, St. Louis, MO, USA). LPS has been shown to induce a stress response measured as increased NO production in aposymbiotic anemones but not in symbiotic anemones (Detournay et al. 2012). Anemones were incubated at increasing concentrations of Fucoidan: 0 (FSW control), 100, 200, 400 and 800 µg/ml, for 4 h, prior to the addition of 1 µg/ml of LPS for a further 12 h to elicit an immune response. The FSW control was also exposed to 1 µg/ml LPS for 12 hours.

### *2.2.5 Assessing infection success using confocal microscopy*

Infection success was assessed flourometrically by confocal microscopy, following methods detailed in Detournay et al. (2012). Following experimental manipulation, anemone treatments were replaced with 1 ml of relaxing solution (1:1 0.37 M $MgCl_2$ : FSW). Samples were observed under a Zeiss LSM 510 Meta microscope with a 40x/0.8 water objective lens and a working distance of 0.8–3.2 mm. Dinoflagellate cells present were visualised by detecting chlorophyll autofluorescence with excitation and emission wavelengths of 543 and 600-700 nm, respectively. Before image scanning, the focal plane of the optical section was adjusted to include the gastrodermal cells within the anemone tentacle. For each experiment, all images were obtained with the same software scanning settings, including detector gain and laser intensity. Quantification of fluorescence was achieved by first defining the gastrodermal tissue area within the anemone tentacles as a region of interest and then measuring the mean fluorescence intensity (MFI) for that region with the LSM 5 software (Zeiss). Symbiont numbers were quantified in host tentacles using confocal imaging. Chlorophyll autofluorescence intensity for each pixel was measured and a threshold value corresponding to the background was defined by measuring the MFI at 600 nm of a gastrodermal section without symbionts (threshold MFI = 20). Infection success was expressed as percent of pixels with autofluorescence intensity above the threshold. In reinfection experiments, each treatment represents a sample size of four anemones *per* treatment and time-point, with % infection taken as the mean over six tentacles *per* anemone. Three untreated symbiotic anemones (six tentacles *per* anemone) were examined to determine a baseline infection level for symbiotic anemones.

To measure and visualise production of NO, animals were transferred from the well-plate to a microfuge tube containing 500 µl of relaxing solution and 15 µM 4-amino-5-methylamino-2,7 difluorofluorescein diacetate (DAF-FM DA, Molecular Probes, Eugene, OR, USA) (Detournay and Weis 2011, Detournay et al. 2012). Animals were incubated for 30 min in the dark and then rinsed twice with relaxing solution. Confocal microscopy of anemone tentacles was used to visualize fluorescence of the DAF FM DA molecular probe following the methods detailed in Detournay et al (2012). The DAF FM DA molecular probe has excitation and emission wavelengths of 488 and 510– 530 nm,

respectively. Quantification of fluorescence was achieved by first defining the gastrodermal tissue area within the anemone tentacles as a region of interest and then measuring the mean fluorescence intensity (MFI) for that region with the LSM 5 software (Zeiss). For each experiment, all images were obtained with the same software scanning settings, including detector gain and laser intensity. Quantification of fluorescence was achieved by first defining the gastrodermis of the tentacles as a region of interest, and then measuring the mean fluorescence intensity (MFI) for that region with LSM 5 software (Zeiss, Germany).

The statistical significance of treatment effects was assessed using a Bayesian mixed-effects analysis of variance model (Gelman 2005). As multiple samples from a single anemone likely violate independence assumptions, a random effect was used for individual anemones in the experiment. Main effects included time and treatment, and their interaction was estimated to account for differences in slope between treatments. The model was estimated using Laplacian approximation methods implemented in the INLA package (Rue et al. 2009), for the statistical computing software R (www.R-project.org/).

## 2.3 Results

### *2.3.1 Searches of six anthozoan species reveal an expanded repertoire of SRCR-domain-containing proteins in cnidarians*

Annotated predicted cnidarian SR proteins were illustrated according to their domain architecture and compared with known human SR protein domain organisation (Figure 2.2). The SRCR domain is present in five of the six SR protein types described. Although Class-A SRs are poorly represented in the cnidarian study group, the SRCR domain is very well represented. The vertebrate Class-A SRs are defined by a collagen domain coupled with either an SRCR domain or a C-type lectin at the C terminus. Cnidarian SRCR-domain-containing proteins lack collagen domains, with the exception of two

putative cnidarian Class-A SRs identified in one of the six species searched, the symbiotic coral *A. digitifera*. Table 2.1 shows numbers of SRCR-domain-containing proteins in relation to developmental stage, data source and symbiotic state for the six cnidarian species searched. The overall numbers of SRCR-domain-containing proteins

within these six species can be explained by data source. In particular, genomic data are available for *Aiptasia* sp., *A. digitifera* and *N. vectensis*, and these have the three highest numbers of SRCR-domain-containing proteins, (66, 62 and 61) respectively. In constrast, the overall numbers of TSR-domain-containing proteins are lower in species for which only transcriptome data are available. The three remaining species had the following numbers of genes encoding SRCR-domain-containing proteins: *A. elegantissima* (12); *A. millepora* (15); and *F. scutaria* (11). *Aiptasia* sp. has both transcriptome and genome data available and a direct comparison shows that the transcriptome has less than a third (20) SRCR-domain containing proteins versus 66 identified in the genome.

Table 2.1 Anthozoan resources searched for SR proteins

| Organism | Family | Developmental stage | Data type | Symbiotic state | SRCR domain containing proteins |
|---|---|---|---|---|---|
| Anemones | | | | | |
| *Nematostella vectensis* | Edwardsiidae | Larvae | Genome | Non-symbiotic | 61 |
| *Anthopleura elelantissima* | Actiniidae | Adult | Transcriptome | Aposymbiotic | 12 |
| *Aiptasia* sp. | Aiptasiidae | Adult | Transcriptome | Aposymbiotic | 20 |
| *Aiptasia* sp. | Aiptasiidae | various | Genome | various | 66 |
| Corals | | | | | |
| *Acropora digitifera* | Acroporidae | Sperm | Genome | Symbiotic | 62 |
| *Acropora millepora* | Acroporidae | Adult and Larvae | Transcriptome | Symbiotic | 15 |
| *Fungia scutaria* | Fungiidae | Larvae | Transcriptome | Aposymbiotic | 11 |

Database searches identified a total of 18 full-length putative cnidarian Class-B SR protein sequences. These proteins all contained a CD36 domain with significant PfamA database matches. Full-length proteins were defined as those containing both transmembrane regions that form the Class-B SR extracellular loop protein configuration. Humans contain four distinct Class-B SR proteins - CD36, SRB 1 & 2, and LMP2 - while the six cnidarian species searched contained between two and four full-length proteins containing a CD36 domain.

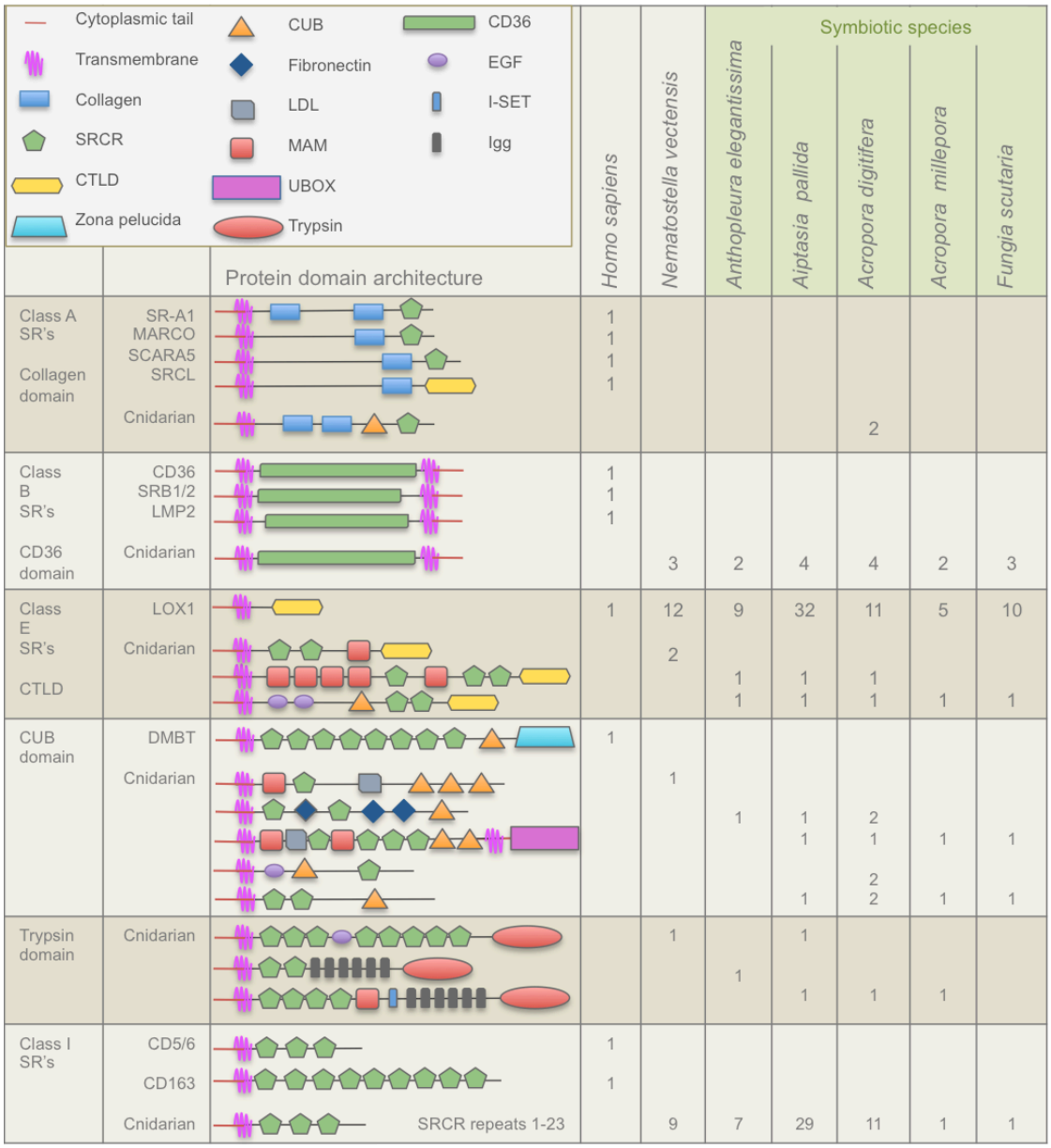



Figure 2.2 Domain architecture of cnidarian SR's compared to Human SR's of known function. SR's display eclectic domain architecture, multiple combinations and permutations create considerable diversity even within classes. In the above subset 15 different domains are listed. All proteins have one transmembrane domain and associated cytoplasmic tail, the three members of class B, SRB1, CD36 and lysosomal integral membrane protein, LMP2 have two transmembrane domains and two tails rooted in the cell forming an extracellular loop structure (see figure 4).Cytoplasmic tails do not display any identifiable protein domains, acting simply as a root for the extracellular protein portion. For full descriptions of domain abbreviations and known functions see PFAM or SMART websites. Human class A SR's are characterised by having one or more collagen domains, in this analysis very few cnidarian SR's were identified with collagen domains and SRCR or Lectin domains. CLEC, C-type lectin; CTLD, C type lectin domain; EGF, epidermal growth factor; LOX1, lectin-like oxidized LDL receptor 1; MARCO, macrophage receptor with collagenous structure (also known as SCARA2 and SR-A2); SCARA5, scavenger receptor class A member 5; SRCL, scavenger receptor with C-type lectin (also known as SCARA4 and CLP1); SRCR, scavenger receptor cysteine-rich domain; DMBT, deleted in malignant brain tumour. Human SR data taken from Canton et al. 2013.

Class-E SRs in humans are defined by the presence of a C-type lectin terminal domain (CTLD). The human LOX1 has a cytoplasmic tail, a transmembrane domain and a single CTLD. In contrast, several predicted cnidarian CTLD-containing proteins contain multiple other domains including SRCR repeats, MAM, EGF and CUB domains. The MAM domain likely has an adhesive function, it contains 4 conserved cysteine residues. The EGF domain is an evolutionary conserved protein domain, which derives its name from the epidermal growth factor where it was first described. It comprises about 30 to 40 amino-acid residues and is found in the extracellular domain of membrane-bound proteins or in secreted proteins. The CUB domain is a structural motif of approximately 110 residues found almost exclusively in extracellular and plasma membrane-associated proteins, many of which are developmentally regulated. The human CUB domain protein 'deleted in malignant brain tumour (DMBT) protein' contains eight SRCR repeats, a single CUB domain and the zona pelucida terminal domain. Predicted cnidarian proteins that resemble DMBT contain 1-3 CUB domains combined with a wide range of other protein domains, including multiple SRCR repeats and fibronectin domains. Another protein resembling DMBT, but with two transmembrane domains, was found in four out of the five symbiotic cnidarian species; this protein has a terminal UBOX domain, four SRCR repeat domains, two MAM domains and two CUB domains.

A potentially novel cnidarian SRCR-domain-containing protein group, found in all cnidarians (accept the *F. scutaria* transcriptome) in various forms, contains a trypsin terminal domain. Various configurations exist, all containing a trypsin terminal domain and 2-8 SRCR repeats. Other repeat domains include the immunoglobulin repeat domain.

Class I SRs are defined by containing only SRCR domains in various numbers of repeats. In humans there are three Class I SRs: CD5, CD6 and CD163. Class I SRs were abundant in all cnidarian study species, in identical configurations to the human Class I SRs, CD5/6 and CD163, and with variable numbers (3-23) of SRCR repeats.

### *2.3.2 Phylogenetic analysis of Class B CD36-domain containing proteins*

The Class-B scavenger receptor SRB1 is upregulated in the symbiotic state of two anemones, implicating the CD36 domain in symbiosis maintenance (Rodriguez-Lanetty

et al. 2006, Lehnert et al. 2014). Phylogenetic analysis was carried out on the CD36 domains identified in this study, with the aim of identifying potential binding motifs and a possible role for this protein domain in symbiosis (Figure 2.3). Protein sequence alignments of the predicted SRB-like proteins from cnidarians, combined with a subset of known vertebrate and invertebrate sequences, revealed that the CD36 domain is highly conserved from sponges to humans. Cnidarian sequences showed between 26% and 32% identity to human CD36 protein, between 28% and 37% identity to human LMP2 protein, and between 28% and 33% identity to human SRB1 protein. Percent identity among the cnidarian sample group was substantially higher, ranging from 39% to 95%. Proteins from the two *Acropora* species (*A. digitifera* and *A. millepora)* showed the highest percent identity with one another (95%). Cnidarian sequences showed between 21% and 27% identity to the predicted SRB-like protein sequence from the sponge, *Suberites domuncula.* Cnidarian sequences also showed between 31% and 34% identity to the LMP2-like sequence from the amphibian *Xenopus laevis.* The full alignment is available as supplemental data (Appendix A4).

Predicted cnidarian proteins used in this analysis lacked one of the three sets of cysteine residues known to form three disulphide bridges in the human CD36 protein. However, a pair of cysteine residues was found in all cnidarian study species at positions C107 and C117. Predicted cnidarian proteins had 8-10 N-linked glycosylation sites compared with 11 known sites in human SRB1 and eight known sites in human CD36 proteins.

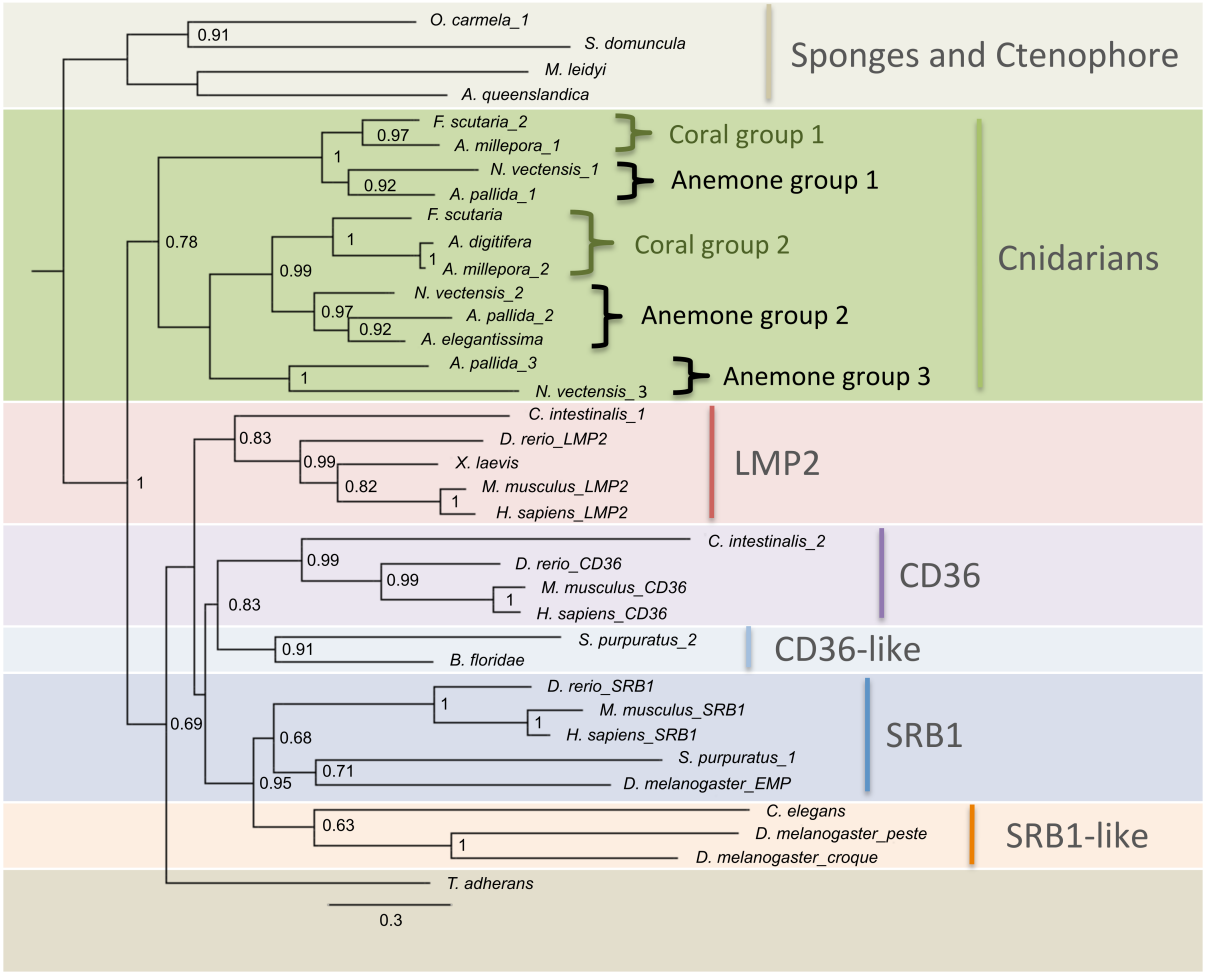



Figure 2.3| Maximum likelihood protein tree produced using SR class B protein, CD36 domains only within Geneious v7.1.2 created by Biomatters, available from http://www.geneious.com/. Proteins aligned using MAFT v7.017 with Blossum 80. Prottest v2.4 was used to suggested best model of evolution for this alignment. Tree produced using FastTree v2.1.5 with Whelan And Goldman, 2001. Bootstrap support values were generated using SEQBOOT, values above 60% (0.6) are shown as proportion at nodes. PhyML alternate tree produced identical topography. Cnidarians display at least 2 if not 3 distinct SR class B proteins, corals show two distinct groups as shown by the two labels in green. The anemones *A. pallida* and *N. vectensis* group to indicate 3 possible SRB variants. The SRB CD36 protein domain is highly conserved among all taxa.

There were two well-supported cnidarian clades, containing both coral and anemone species (Figure 2.3). Corals and anemones formed distinct groupings within each of the two clades. The putative cnidarian SRB proteins grouped together, forming a separate clade from other metazoans except for sponges and ctenophores, which formed a separate ancestral grouping. In contrast, other invertebrates such as the fruitfly *Drosophila melanogaster,* the sea urchin, *Strongylocentrotus purpuratus,* and the urochordate *Ciona intestinalis* grouped with mammalian sequences in several different sub-clades of SRBs: LMP2, CD36, CD36-like, SRB1, and SRB1-like proteins.

### *2.3.3 Experimental blocking of SR proteins in the symbiotic anemone Aiptasia sp.*

Fucoidan-treated anemones showed lower levels of infection (0-3% infection) than both the control and the fucoidan-washed control (7-8% infection), and infection success was significantly decreased with increasing fucoidan concentration (Figure 2.4). The Bayesian ANOVA suggested that the probability of these effects occurring under a control treatment was indistinguishable from zero at both time points and for all treatments. Lowered infection levels were reversed by washing fucoidan-treated anemones with FSW and resting anemones for 48 h in FSW before re-infection; this control suggests the treatment is not toxic to the anemone.

A second fucoidan experiment investigated the possible immune-regulation role of a SR in symbiosis maintenance. Symbiotic anemones were treated with increasing concentrations of fucoidan to block positively-charged binding sites of SRs and were subsequently immune-challenged by incubation with LPS. In the present experiment, the FSW/symbiotic anemone control treatment showed low levels of NO production in response to LPS. However, fucoidan treatments showed increasing NO production with increasing concentrations of fucoidan (Figure 2.5).

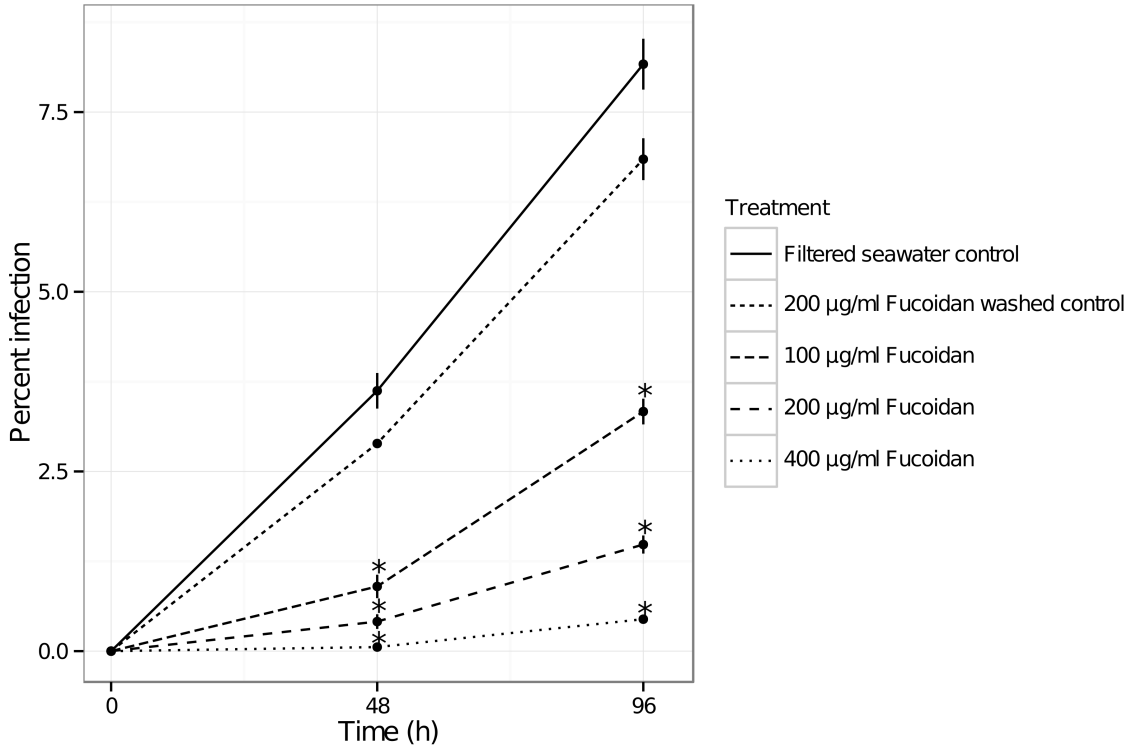



Figure 2.3: Experimental re-infection of aposymbiotic *Aiptasia* sp. treated with the known SR ligand, fucoidan, from the algae, *Fucus vesiculosus.* Graph shows percent infection success at two time points post re-infection (48 and 96h). Five treatments include two controls; filtered sea water (FSW) control and 200µg/ml fucoidan treatment washed after 18h incubation with fucoidan and rested 48h in FSW. Fucoidan is a known polyanionic ligand of positively charged binding sites on both SRCR and CD36 SR domains. Fucoidan treatments show significantly lower symbiont uptake than both controls, the highest concentration of fucoidan shows the lowest infection success. Bars represent means + SE, n = 3 anemones per treatment. Stars indicate high ($p>0.999$) posterior probability of treatment effects being different from controls under the Bayesian ANOVA model.

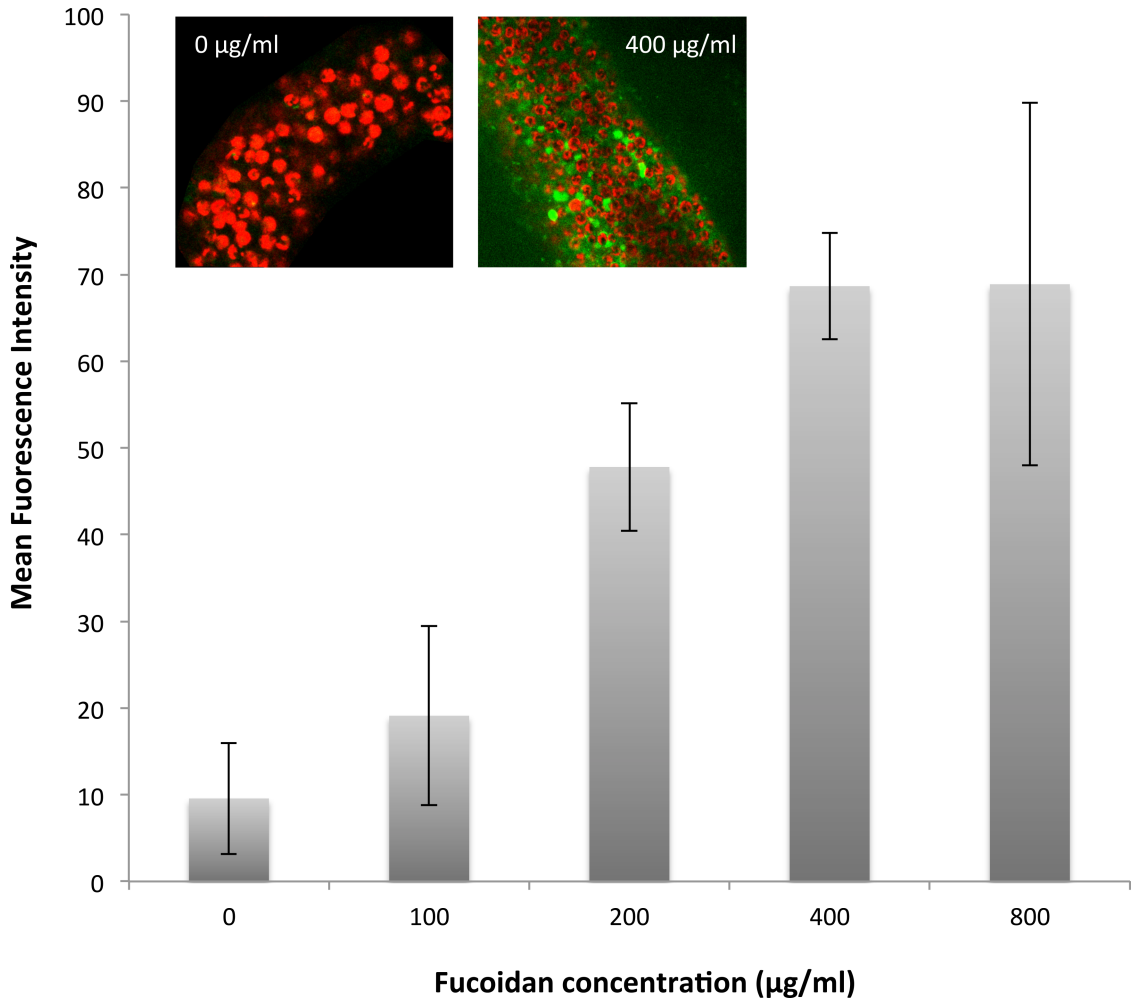



Figure 2.4: Effect of immune stimulation by LPS on NO production in symbiotic *A. pallida* treated with the known SR ligand fucoidan from the seaweed *Fucus vesiculosus*. Graph shows mean florescence intensity of the molecular probe DAF-FM used to label nitric oxide as an indicator of stress; plotted against treatments of increasing concentrations of the sulphated polysaccharide fucoidan. Anemones were incubated in four increasing concentrations of fucoidan from 100µg/ml to 800µg/ml, for 4 hours prior to the addition of 1 lg/ml LPS for a further 12 hours to challenge an immune response. Filtered sea water control was also exposed to 1 lg/ml LPS overnight. Fucoidan is a known polyanionic ligand of positively charged binding sites on both SR class A and class B proteins. Inset representative confocal images show the DAF-FM DA (510–530 nm) NO-dependent fluorescent signal appears in green and symbiont autofluorescence (600–700 nm) in red. Bars represent means + SD; n = 3 anemones.

## 2.4 Discussion

The aim of this study was to explore the SR protein repertoire in cnidarians, and to identify potential SR protein candidates for involvement in the establishment and maintenance of the cnidarian-dinoflagellate symbiosis. The six cnidarian species

searched have a large SR repertoire, including an expanded SRCR-domain-containing protein complement. Class-B CD36-domain-containing proteins were found in all six species in numbers comparable to the four found in humans. However five of the six study species lack the classic vertebrate macrophage SRA1/MARCO-like proteins that contain collagen domains, with the exception of *A. digitifera*, which has two proteins containing both collagen and SRCR domains. Searches targeted SRs containing three specific domains: the SRCR domain, the CTLD and the CD36 domain. These three domains have been previously identified as potentially important to cnidarian-dinoflagellate symbiosis; SRCR and CD36 are upregulated in the symbiotic compared to aposymbiotic state in two anemone species (Rodriguez-Lanetty et al. 2006, Schwarz et al. 2008, Lehnert et al. 2014), and CTLD's have been implicated in the recognition of symbiont cell-surface glycans (Wood-Charlson et al. 2006, Logan et al. 2010). Searches identified cnidarian SR proteins belonging to four out of nine possible sub-classes of human SRs (Classes A, B, E, and I).

### *2.4.1 The SRCR-domain-containing protein repertoire in cnidarians is expanded compared to vertebrates*

The purple sea urchin, *Strogolocentrotus purpuratus*, contains a vastly expanded innate immune receptor repertoire, including 218 genes encoding SRCR-domain-containing proteins (Pancer et al. 1999, Pancer 2000, Rast et al. 2008). Comparable expansion has been reported in the amphioxus, *Branchiostoma floridae*, genome with 270 genes encoding SRCR-domain-containing proteins (Huang et al. 2008). The SRCR-domain-containing protein repertoire in cnidarians is likewise expanded although not as dramatically as in the sea urchin and in amphioxius. This study found 62 and 61 genes encoding SRCR-domain-containing proteins in the genomes of *A. digitifera* and *N. vectensis*, respectively, compared with 16 genes in humans. In addition, cnidarian SRCR proteins contain a diverse complement of other protein domains, including CUB, CTLD and immunoglobulin domains.

Vertebrate Class-I SR proteins contain only SRCR domains. Examples of human Class I proteins include CD5/6 and CD163, which are described as cell adhesion molecules (Bowen 1995, Madsen et al. 2004, Martinez et al. 2011). Class-I SR proteins have been identified in the sponge *Geodia cydonium* and the starfish *Asterina pectinifera* (Pancer et

al. 1997, Furukawa et al. 2012). In *G. cydonium*, a protein composed of 14 SRCR domains is described as 'sponge aggregation factor', and mediates the re-aggregation of disassociated sponge cells (Blumbach et al. 1998). In the sea urchin, *Strongylocentrotus purpuratus*, another Class-I SR protein has been named 'speract receptor', and assists sperm/egg adhesion. All of the known functions for Class-I SR proteins involve cell adhesion.

In the context of cnidarian-dinoflagellate symbiosis, SRCR-domain-containing proteins could function in recognition of *Symbiodinium* cells and initial binding to the host cell surface. There is also some evidence to suggest a potential role in initiating phagocytosis. In the purple sea urchin, for example, SRCR proteins are expressed by coelomocytes, phagocytic cells that function in immunity. Sea urchin and amphioxus SRCR proteins share structural similarities to the mammalian DMBT protein. DMBT has eight SRCR repeats, a CUB domain and a zona pelucida terminal domain. In mammals, TLR4, NOD2 and DMBT1 function together in gut immunity, and DMBT has been shown to bind bacteria. CUB domains are found in many cell-surface receptors and are involved in binding target molecules (Martinez et al. 2011). Several cnidarian SRs containing 1-3 CUB domains were identified, however domain organization in these proteins shows few similarities with DMBT or sea urchin and amphioxus examples.

### *2.4.2 Putative CTLD-domain-containing SRs in cnidarians also contain the SRCR domain*

C-type lectin domain (CTLD)-containing proteins have been identified as having a potential role in the recognition of *Symbiodinium* cell surface glycans (Wood-Charlson et al. 2006, Logan et al. 2010). Experimental removal and blocking of glycans on the *Symbiodinium* cell surface significantly reduced infection success rates in aposymbiotic anemones and coral larvae infected with treated dinoflagellates (Wood-Charlson et al. 2006). Previous searches identified 92 CTLDs in the genome of the non-symbiotic anemone *N. vectensis* (Wood-Charlson and Weis 2009). Searches carried out in the current study identified several putative Class-E SRs that contain CTLD domains. All six cnidarian species have between 1-32 vertebrate LOX1 homologues and 1-4 other CTLD-domain-containing proteins combined with SRCR, MAM, CUB and EGF domains. In vertebrates, the LOX1 protein is involved in intracellular trafficking and signalling, activating apoptosis and mediating phagocytosis of apoptotic cells (Murphy et al. 2005).

Cnidarian CTLD-domain-containing proteins described here provide potential target proteins for experimental investigation of the lectin-glycan interactions between cnidarian gastrodermal cell-surface receptors and the *Symbiodinium* cell-surface glycome. Further investigation is needed to confirm whether these CTLD SRs are symbiosis-specific proteins that have evolved to detect specific symbiont strains in the selection ("winnowing") process during symbiont acquisition.

### *2.4.3 CD36-domain-containing proteins in cnidarians*

This study presents the first detailed analysis of multiple cnidarian CD36-domain-containing proteins. All six cnidarian species have between 1-4 genes encoding CD36-domain-containing proteins, belonging to the vertebrate Class-B SR group. Class B SRs have been implicated in the maintenance of steady state symbiosis, due to their upregulation in the symbiotic state of *A. elegantissima* and *A. pallida* (Rodriguez-Lanetty 2006, Lehnert et al. 2014). At the domain level of organization, predicted cnidarian Class-B SRs appear identical to all other metazoan Class-B SRs; with two transmembrane domains flanking a large CD36 domain. Further analysis of cnidarian CD36 domain amino acid sequences, in a pairwise protein alignment with a full complement of metazoan species, revealed differences in folding potential. While many amino acid motifs are conserved within all metazoan species, the six cysteine residues that are known to form three disulphide bridges in human CD36 domains, were not all present in cnidarians. Specifically, the CD36 domain has six cysteine residues that form bridges at $C^{243}$- $C^{311}$, $C^{272}$-$C^{333}$, and $C^{313}$-$C^{322}$, (Silverstein and Febbraio 2009), yet in cnidarians the $C^{243}$- $C^{311}$ bridge cysteine residues are both missing. However, cnidarians have two cysteine residues at $C^{84}$ and $C^{86}$. This cysteine pair was also found in the SRB1-like sequence from the nematode *Caenorhabditis elegans*. In contrast, the three sponges, *Oscarella carmella*, *Suberites domuncula*, and *Amphimedon queenslandica*, and the ctenophore *Mnemiopsis leidyi* do not contain any of the conserved cysteine residues found in vertebrate proteins. These differences may explain why human and mouse antibodies for SRB1 and CD36 failed to bind Class-B SRs in *Aiptasia* sp. in repeated western blot experiments (E.F. Neubauer; unpublished data). A difference in 3-D folding structure will alter the binding capabilities of these glycoproteins.

### *2.4.4 Functional experiments suggest that blocking SRs decreases symbiont acquisition and increases the stress response to immune challenge in the anemone Aiptasia sp.*

Fucoidan blocks the positively-charged binding sites on vertebrate Class-A and Class-B SRs (Hsu et al. 2001, Dinguirard and Yoshino 2006, Li et al. 2008), however this blocking effect is indiscriminate, blocking both SRCR and CD36 domains. Despite the lack of discrimination between SR Classes and binding domains, fucoidan is regularly used to block vertebrate macrophage SR activity, and in the absence of specific antibodies to SRs, it provides a rapid method for testing the effect of blocking SR-binding within the context of symbiont uptake. Experimental blocking of SRs in aposymbiotic anemones resulted in a marked decrease in symbiont infection rate. This result provides the first functional evidence that one or many SRs with SRCR or CD36 domains are essential for symbiont uptake by the cnidarian host.

My experimental results further support a role for SRs in immune suppression and symbiont tolerance. The symbiont must avoid digestion once inside the host vacuole; mechanisms for promoting tolerance may involve immune modulation or suppression. Previous experiments with *Aiptasia* sp. showed that symbiotic anemones produced significantly less nitric oxide (NO) as a stress response to LPS challenge than aposymbiotic anemones exposed to the same conditions (Detournay et al. 2012). In a second fucoidan-treatment experiment, I demonstrated that the stress response (NO production) in symbiotic *Aiptasia* sp. can be re-activated by blocking SR-binding domains with fucoidan. This suggests that SRs play an active role in regulating the stress response in symbiotic cnidarians. In *Aiptasia* sp., previous experimental evidence suggests that the activation of the host TGFβ innate immune pathway promotes tolerance of the dinoflagellate symbiont at the onset of symbiosis and during the stable symbiotic state (Detournay et al. 2012), while in humans, CD36 is required to turn latent TGFβ into its active form (Khalil 1999). I suggest that the dinoflagellate symbiont manipulates the cnidarian host SR's, initiating phagocytosis and subsequently initiating the TGFβ pathway. The dominant role of the TGFβ immune pathway is to promote tolerance and resolve inflamation (Li et al. 2006). Many intracellular parasites manipulate the host innate immune defence mechanisms to their own advantage (Medzhitov and Janeway 2002). Immune receptors have a dual role, they exist not only

to protect from pathogenic attack but also to maintain and control beneficial microbial communities within the host or holobiont (Mcfall-Ngai 2007).

## 2.5 Concluding remarks

This study provides the first detailed description of scavenger receptors in cnidarians and contributes to basic knowledge about the transmembrane protein repertoire for this basal metazoan group. The cell membrane and transmembrane protein receptors form the first line of defense for cells against invading microbial attack and is therefore important to research on coral disease and immunity, as well as in the context of mutualistic partners such as *Symbiodinium* spp. These genomic data enable and inform ongoing work investigating a potential role for SRs in the establishment of a healthy cnidarian-dinoflagellate symbiosis. The bioinformatic searches and analyses carried out in this study reveal a wide variety of cnidarian SR proteins with potential involvement in recognition and uptake of dinoflagellate symbionts. All three domains, the SRCR domain, the CTLDs, and the CD36 domain are implicated in recognition and phagocytosis in other systems, such as those involving apicomplexan parasites and vertebrates, and within the cnidarian-dinoflagellate symbiosis. In particular, The CD36 domain is implicated in immune modulation and maintenance of the symbiosis by potential involvement in the tolerogenic TGFβ pathway. The apicomplexan TRAP protein provides a suggested ligand for the CD36 domain; a TRAP homologue or TSR-domain-containing protein may therefore be involved in phagocytosis and modulation of the host immune response in the cnidarian-dinoflagellate symbiosis.

While results presented here clearly implicate the SR superfamily of proteins in the establishment of the cnidarian-dinoflagellate symbiosis, further investigation is necessary to establish the specific SR proteins involved. Blocking of specific SR domains and proteins using antibodies may reveal a clearer picture of events.

# Chapter 3

# The large and diverse TSR-domain-containing protein repertoire of six cnidarian species contains highly conserved binding sites and motifs

## 3.1 Introduction

In mammalian cells the extracellular matrix (ECM) is made up of modular secreted proteins that contain a number of repeat motifs and binding domains; many of these act as pattern recognition receptors (PRRs) and form the first line of defense against invading microbes. PRRs of the animal innate immune system recognize microbial molecular patterns from both parasitic and beneficial microbial invaders (Schwarz 2008). Extracellular and transmembrane proteins are therefore potential target ligands for microbial surface proteins to bind and gain entry to host cells. Human thrombospondin proteins (thrombospondin 1,2,3,4 and 5 or cartilage oligomeric matrix protein (COMP) are a well-studied family of ECM glycoproteins due to their multiple roles, including platelet aggregation, clot formation and cell attachment, anti-angiogenic capabilities (Adams and Lawler, 2004). The human thrombospondin protein family provides potential angiogenesis inhibitors, treatments for cancer and other diseases that involve a proliferation of blood vessels (Folkman 2004, Silverstein and Febbraio 2007, Kazerounian et al. 2008, Yee et al. 2009). The thrombospondin structural homology repeat, or TSR domain, was originally characterized in the human thrombospondin-1 protein (TSP1). Protein-protein interactions involving the TSR domain are central to TSP1 protein function, and the discovery of multiple invertebrate and vertebrate TSR-domain-containing proteins has raised questions about the importance and shared functions of this domain (Adams and Tucker, 2000).

All TSR-domain-containing proteins have one or more TSR domains, similar to the three repeats originally characterized in the human TSP1 protein (Figure 3.1) (Tucker 2004). TSR domains are found in secreted proteins or in the extracellular portions of

transmembrane proteins, with functions as varied as extracellular matrix remodelling, cell-cell interactions, and development (Lawler 1986, 2000). The TSR domain consists of approximately 60 amino acids, with several highly conserved motifs. TSR domains contain six conserved cysteine residues, forming three disulphide bridges with the exception of a few human complement factors and two malaria proteins (TRAP's), that have TSR domains that contain five cysteines (Tan et al. 2002). The $NH_2$-terminal portion of the TSR domain contains two or three tryptophan residues separated by two amino acids (i.e. WSXWSXW (Adams and Tucker 2000)). This latter motif binds glycosaminoglycans (GAGs), integrins and other proteins (see Chen et al. (2000) for a review).

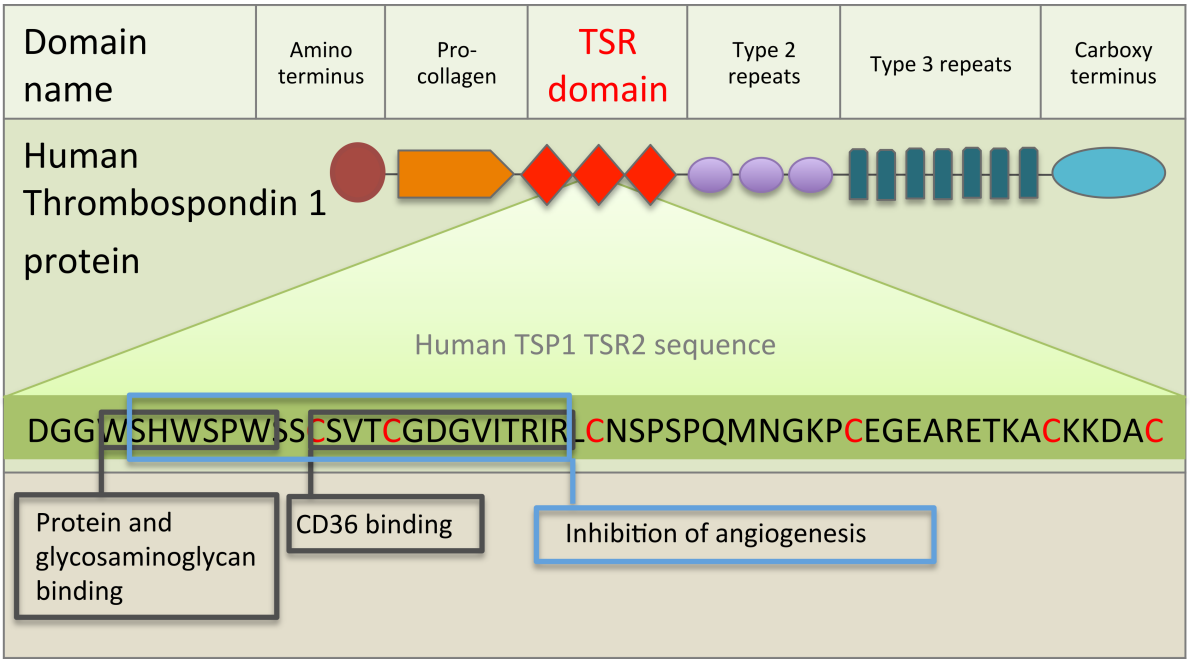



Figure 3.1: What is a TSR? Schematic representation of Human Thrombospondin 1 (TSP1) protein. The human TSP1 protein comprises multiple functional domains and each domain has multiple individual functions. The three TSR (Thrombospondin Structural homology Repeat) domains are depicted by three red diamonds. The amino acid sequence of the second TSR sequence is shown with six conserved cysteines in red. Known binding motifs and capabilities of the human TSP1 TSR2 domain are listed and depicted in boxes (Re-drawn from Zhang and Lawler 2007).

Several studies have used synthetic peptides designed from the TSR domains of human TSP1 to determine which motifs have anti-angiogenic activity. The type-B scavenger receptor, CD36, which is located on human endothelial cells, is responsible for the anti-

angiogenic activity of TSP1 (Dawson et al. 1997). Synthetic peptide experiments show that the conserved 'CSVTCG' motif within the TSR domain, are important to anti-angiogenic activity and binding of CD36 (Tolsma et al. 1993, Dawson et al. 1997). Areas flanking, but not including the CSVTCG motif, also have anti-angiogenic capabilities. These include the tryptophan GAG-binding motif 'WXXW' (Guo et al. 1992), and the 'GVITRIR' motif, which contains highly conserved arginine residues (Dawson et al. 1999, Anderson et al. 2007, Garside et al. 2010). However, whether these flanking motifs are specifically binding CD36 to give anti-angiogenic effects has not been tested (see (Kazerounian et al. 2008) for a review of thrombospondins and cancer therapeutics).

The TSR domain is evolutionarily ancient and has been duplicated and shuffled many times. For instance, 41 human genes contain one or more TSR domain copies (Silverstein 2002), while 27 genes in *C. elegans* and 14 in *Drosophila* contain TSR repeats (Tan et al. 2002). This large and diverse group of TSR-domain-containing proteins is collectively referred to as the TSR super-family (Adams and Tucker 2000). The key members of the TSR-domain-containing protein super family are considered in more detail here.

Several apicomplexan parasites use a TSR-domain-containing protein to assist in gliding motility, recognition and to gain entry to host cells (Kappe et al. 1999, Vaughan et al. 2008, Morahan et al. 2009). The host cell invasion strategies employed by apicomplexans, such as the malaria parasite *Plasmodium falparcium*, are well studied and involve a unique process by which a transmembrane protein on the parasite's cell surface, the thrombospondin-related anonymous protein (TRAP), links the parasite's cytoskeleton with the host cell membrane. The TRAP protein's cytoplasmic tail links to actin within the parasite, while extracellular domains bind host hepatocytes (Morahan et al. 2009). TRAP proteins have subsequently been discovered in all *Plasmodium* species, and in other members of the Apicomplexa, including *Toxoplasma gondii*, *Cryptosporidium*, *Theileria*, *Eimeria* and *Babesia* (Templeton and Kaslow 1997, Robson et al. 1998, Morahan et al. 2009).

The extracellular TRAP TSR domain is responsible for binding proteins on the host cell membrane. The amino acid motif 'WSPCSVTCG' is specifically responsible for the ability

to bind sulphated glyconjugates and hepatocytes, in the *P. falciparum* TRAP protein and portion two of the circumsporozoite protein (CS) (Müller et al. 1993). This suggests that TRAP proteins are involved in recognition and entry of hepatocytes, as sporozoite entry was inhibited by antisera raised against WSPCSVTCG.

The specific involvement of a TSR-domain motif in recognition and entry of an apicomplexan parasite into host liver cells makes the TSR domain a target of interest in the study of cnidarian-dinoflagellate symbiosis. Dinoflagellates are a sister taxon to the apicomplexans within the infrakingdom Alveolata (Baldauf 2003, Saldarriaga et al. 2004, Janouskovec et al. 2010a), and might therefore share mechanisms of host invasion (Schwarz 2008). Initial recognition between a cnidarian host and its dinoflagellate symbionts involves several complex ligand/receptor interactions (Davy et al. 2012). The TSR domain is a candidate for motifs involved in the recognition and entry of the dinoflagellate into cnidarian gastrodermal cells.

Cnidarian species regularly form an intracellular symbiosis with photosynthetic dinoflagellates of the genus *Symbiodinium*. The dinoflagellate is housed inside a vacuole in the host gastrodermal cell, and symbionts proliferate within the host, maintaining an optimal symbiont population for light capture. The relationship is based on nutritional exchange: the symbiont translocates fixed carbon to the host in return for inorganic nutrients and a high light environment, the exact detail of the exchange are not fully understood (Yellowlees et al. 2008). This symbiotic partnership allows corals to form the basis for the entire reef ecosystems in nutrient-poor shallow tropical seas. It is unknown how cnidarians manage the balance of beneficial microbes such as dinoflagellates, while maintaining an immune response to microbial invasion. Several studies have investigated innate immune pathways in cnidarians (Miller et al. 2005, Schwarz et al. 2007, Bosch 2008, Wood-Charlson and Weis 2009, Shinzato et al. 2014b, Poole and Weis 2014), and in particular potential tolerogenic signals between the two partners (Detournay et al. 2012).

Aside from their taxonomic proximity to apicomplexan obligate intracellular parasites, other factors point to the involvement of TSR proteins in the establishment of the cnidarian-dinoflagellate symbiosis: Firstly, the TSR domain binds to the Class-B SRs,

CD36/SRB1 (Asch et al. 1987, Li et al. 1993, Frieda et al. 1995, Dawson et al. 1997). A homolog to the CD36/SRB1 gene was found to be upregulated in the symbiotic state of the anemones *Anthopleura elegantissima* (Rodriguez-Lanetty et al. 2006) and *Aiptasia* sp. (Lehnert et al. 2014). Secondly, the human TSP1 protein functions in transformation of latent TGFβ to the active form, and the subsequent activation of the immunosuppressive TGFβ pathway responsible for blocking a response to immune challenge in humans. Previous work in cnidarians identified many proteins within the TGFβ immune suppressive pathway and suggesting a tolerogenic balance between host and symbiont (Detournay et al. 2012).

Cnidarians do not possess a classic TSP1 gene; the thrombospondin-like proteins in cnidarians lack TSR domains (Bentley and Adams 2010), however other studies have identified TSR domains within cnidarians. A rhamnospondin gene with eight TSR domain repeats was identified in the colonial hydroid *Hydractinia symbiolongicarpus* (López et al. 2011), and several TSR domain-containing proteins were identified in two species of corals, *Acropora palmata* and *Montastraea faveolata.* TSR domains were found in four life-history stages (spawned eggs, early stage planulae, late stage planulae, and adult corals) (Schwarz et al. 2008). Furthermore, TSR-domain-containing proteins were identified in all ten cnidarian species searched in a recent study of candidate genes in all available cnidarian sequence resources (Meyer and Weis 2012).

The aim of this chapter was to identify and compare the TSR-domain-containing protein repertoire of six cnidarian species (five symbiotic, one non-symbiotic) and one symbiotic dinoflagellate, to investigate potential ligands for scavenger receptor class-B members (CD36/SRB1) and potential TRAP-like proteins in the host and symbiont respectively. Using a variety of genomic and transcriptomic resources, I compared and contrasted vertebrate TSR proteins of known function with the cnidarian TSR repertoire. I investigated the presence of known binding motifs and their conservation within the cnidarian TSR-domains. In addition, I identified potential TSR proteins and TRAP homologues in the *Symbiodinium minutum* genome (Shoguchi et al. 2013). The current study characterizes a large repertoire of TSR-domain-containing proteins within all six cnidarian species investigated. Detailed description of TSR-domain-containing protein sequences will inform the design of further functional experiments

to investigate the possibility that one or more of these proteins is involved in the initiation of the cnidarian-dinoflagellate symbiosis.

## 3.2 Materials and methods

### *3.2.1 Cnidarian genomic and transcriptomic resources*

To characterize the TSR-domain-containing protein repertoire of cnidarians, the publically available genomic or transcriptomic resources for six Anthozoan species were searched. These included three sea anemone species: *Anthopleura elegantissima* (Kitchen et al., submitted: http://people.oregonstate.edu/~meyere/ data.html), *Aiptasia* sp. (Lehnert et al. 2012) and *Nematostella vectensis* (Putnam et al. 2007), and three coral species: *Acropora digitifera* (Shinzato et al. 2011), *Acropora millepora* (Moya et al. 2012) and *Fungia scutaria* (Kitchen et al., submitted). In addition the genome of the symbiotic dinoflagellate *Symbiodinium minutum* (Shoguchi et al. 2013) was searched for TSR-domain-containing proteins, to investigate the presence of a potential TRAP-like protein. These resources represent various developmental stages and symbiotic states (Table 3.1). All resources were used without manipulation, with the exception of *Aiptasia* sp, for which raw Illumina sequence reads for accession SRR696721 were downloaded from the sequence read archive (SRA) entry for the aposymbiotic CC7 transcriptome (http:// www.ncbi.nlm.nih.gov/sra/SRX231866) and reassembled using Trinity (Grabherr et al., 2011).

Table 3.1: Anthozoan resources searched for TSR-domain-containing proteins

| Organism | Family | Developmental stage | Symbiotic state | Data type | Publication |
|---|---|---|---|---|---|
| *Nematostella vectensis* | Edwardsiidae | Larvae | Non-symbiotic | Genome | Putnam et al. 2007<br>Kimura et al. 2009 |
| *Anthopleura elegantissima* | Actiniidae | Adult | Aposymbiotic | Transcriptome | Kitchener et al. in prep |
| *Aiptasia* sp. | Aiptasiidae | Adult | Aposymbiotic | Transcriptome | Lehnert et al. 2012 |
| *Acropora digitifera* | Acroporidae | Sperm | Symbiotic | Genome | Shinzato et al. 2011 |
| *Acropora millepora* | Acroporidae | Adult and Larvae | Symbiotic | Transcriptome | Moya et al. 2012 |
| *Fungia scutaria* | Fungiidae | Larvae | Aposymbiotic | Transcriptome | Kitchener et al. in prep |

### *3.2.2 TSR sequence searching, verification and phylogenetic analysis*

A variety of invertebrate genomic and transcriptomic resources were searched for TSR-domain-containing proteins (Appendix B1). Searches of the N. vectensis genome protein models were performed using a combination of BLASTp and tBLASTn searches using protein fragments from the TSR domains in human and mouse TSP1, and the keywords: thrombospondin (162 genes found), tsp1 (125 genes found), adam (28 genes found), and sema (two genes found). All databases were queried in the same way, using either BLASTp or tBLASTn searches with the most closely related annotated protein sequences available, combined with keyword searches of the GO or KEGG annotations. Transcriptome data was loaded into Geneious version 7.1.8 (http://www.geneious.com, (Kearse et al. 2012) and BLAST searches were carried out within this program. To ensure as many TSR-domain-containing protein sequences as possible were recovered, representative N. vectensis sequences of each protein type (ADAMTS, Astacin, Trypsin, VWA, SEMA, Igg, TSR domains only) were also used as queries for tBLASTn searches of the other five anthozoan resources. To confirm that the sequences obtained contained TSR domains, nucleotide sequences were translated using the program Geneious version 7.1.8 (http://www.geneious.com, (Kearse et al. 2012) and then annotated using the Geneious InterProScan plugin (Quevillon et al. 2005). InterProScan simultaneously searches nine different protein domain databases, only sequences where two or more

databases found TSR domains and with an E- value of less than 1x10-4 were used. Where InterProScan was unable to resolve protein domains these sequences were used as query sequences for the online protein domain database PfamA (http://pfam.sanger.ac.uk) (Punta et al., 2012). Sequences for each species were aligned and those that were identical or almost identical (less than 5aa different in the conserved domains) were thrown out of the analysis due to the fact that they likely represented artifacts of assembly issues or different isoforms of the same protein. Protein fragments were identified as missing a start or a stop codon and therefore left out of the analysis.

To verify that proteins contained TSR domains, only proteins that showed significant pfam A matches to a TSR domain were included in the analysis. Proteins were annotated using the Geneious plugin InterProScan, using the Pfam A database, and then checked by eye. Diagrammatic representations of the protein domain configurations were then generated using this information. Consensus sequences (cl15278: TSP_1 Superfamily) from the conserved domain database (http://www.ncbi.nlm.nih.gov/cdd) (Marchler-Bauer et al., 2013) were used as queries in tBLASTn searches of each anthozoan resource. A high E-value cutoff (1x10-1) was used in the tBLASTn searches to recover more divergent sequences. All BLAST searches were performed using Geneious pro version 5.4.3 (Drummond et al., 2011). Full details and accession numbers for all cnidarian scavenger receptors identified in this study are listed in Appendix A2.

A multiple sequence alignment was performed on the protein sequences with the MAFFT v 7.017 plug-in through Geneious, using the default settings (Drummond et al., 2011; Katoh et al., 2002). To standardize the data used, the second TSR domain from proteins with multiple repeats was always chosen for alignment. This method is consistent with previous studies involving human TSP1 protein (Zhang and Lawler 2007). Accession numbers for all proteins used are listed in Appendix B2. To determine the best model of protein evolution for each alignment, ProtTest v 2.4 (Abascal et al., 2005) was run and results were compared using statistical model selection for all possible substitution matrices and improvements. The best model was LG+I+G. Maximum likelihood phylogenetic analysis was performed using the PhyML 3.0 web server (http://www.atgc-montpellier.fr/phyml) (Guindon and Gascuel, 2003) using the

appropriate model of protein evolution and 500 bootstrap replicates. FigTree v1.4 (http://tree.bio.ed.ac.uk/software/figtree/) was used to visualize and annotate the resulting tree.

## 3.3 Results

### *3.3.1 TSR-domain-containing proteins in Symbiodinium minutum*

Searches of the *Symbiodinium minutum* genome identified 175 contigs containing TSR domains, however predicted proteins contained no other identifiable domains. TSR domains were alone or in repeats of up to 15 (Figure 3.2). In contrast, most other apicomplexan sequences possessed other domains, including a Von Willebrand factor (VWA) domain and all had a C-terminal transmembrane domain. Figure 3.3 shows a protein domain alignment of TSR domains, including TRAP-protein TSR domains from several apicomplexans, *S. minutum* TSR domains, human TSP1 TSR2, and TSR domains from two cnidarian TSR-domain-containing proteins. TSR domains from *S. minutum* TSR-domain-containing proteins have five or six cysteines, a variation that is consistent with two apicomplexan TRAP proteins (Tan et al. 2002). The CD36/SRB1 binding sites but not the glycosaminoglycan-binding sites are well-conserved in *S. minutum*.

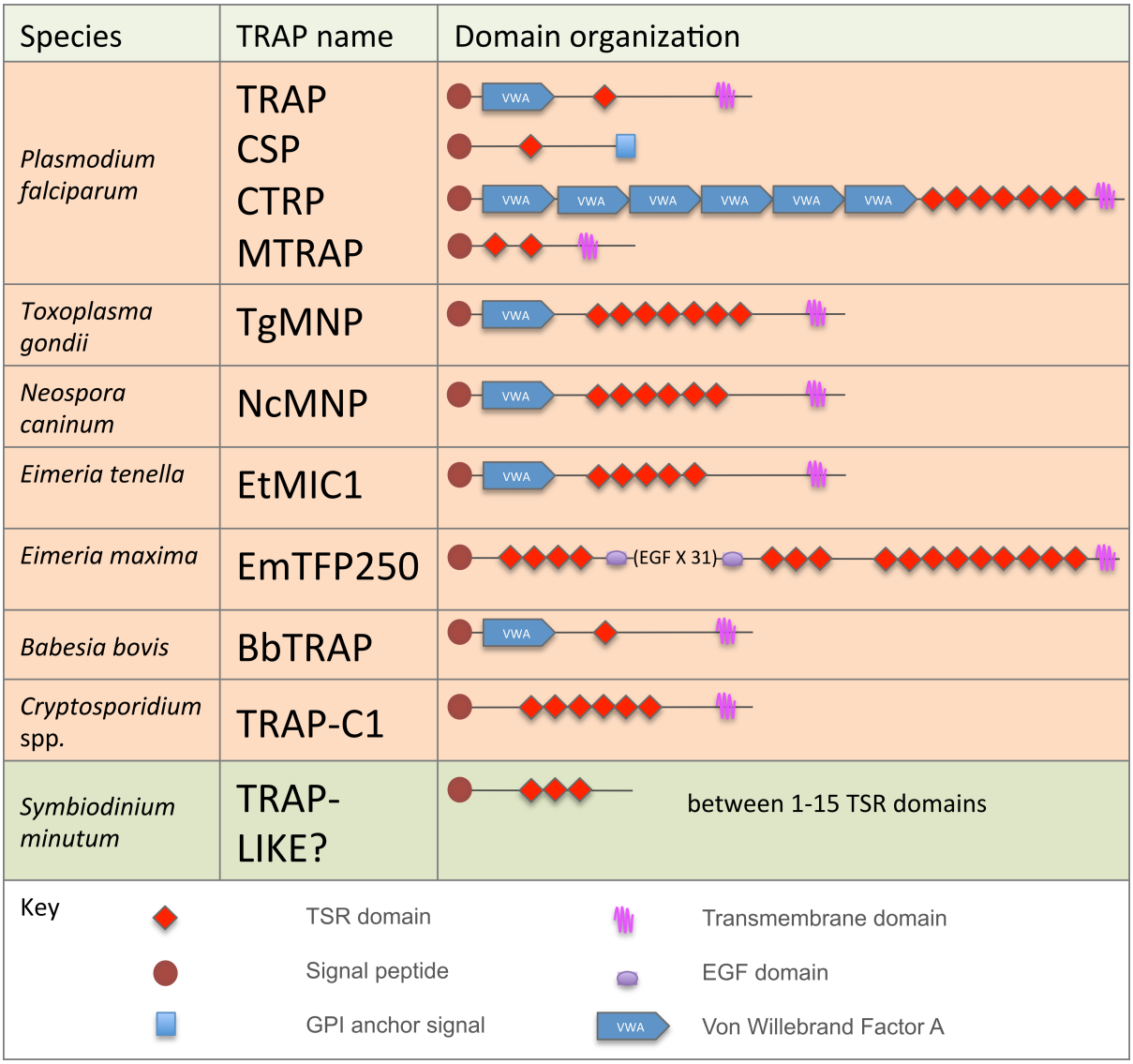



Figure 3.2: What is a TRAP (thrombospondin-related anonymous protein)? Schematic representation of different members of the thrombospondin gene family in apicomplexan parasites. Apicomplexan TRAP proteins (orange background) and TSR-domain-containing proteins from the dinoflagellate *Symbiodinium minutum* (green background).

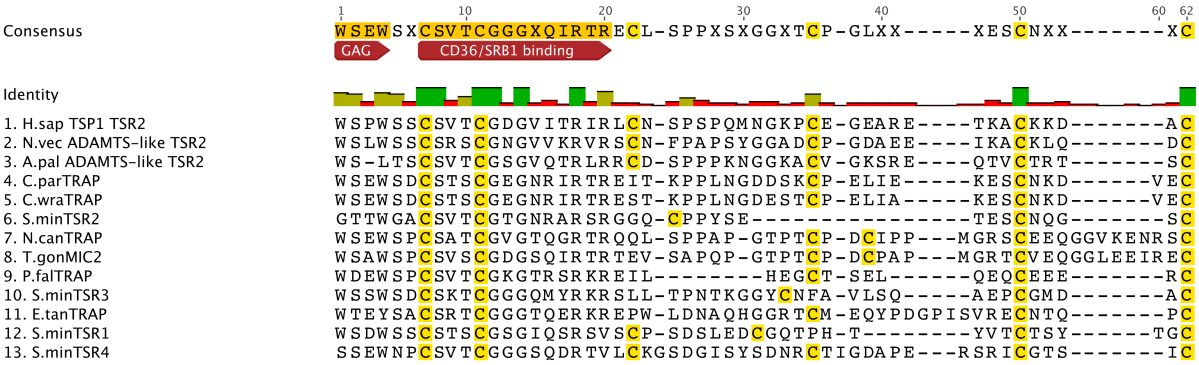



Figure 3.3: TSR domain alignment comparing apicomplexan TRAP TSR domains with TSR domains from the dinoflagellate *Symbiodinium minutum,* TSR 2 from human TSP1, and Adamts-like TSR domains from the anemones *Nematostella vectensis* and *Aiptasia sp*. Positioning and absence of specific cysteine residues (coloured yellow) in TRAP and *S. minutum* TSRs will result in different patterns of disulphide bonds and three-dimentional folding. Binding sites for glycosaminoglycans (GAGs) and the scavenger receptors CD36/SRB1 (annotated in red) are somewhat conserved. Consensus sequence is generated from the cnidarian, apicomplexan and human sequences in the above alignment.

### *3.3.2 Analysis of potential binding sites and conserved motifs in cnidarian TSR domains*

TSR domains taken from a selection of TSR-domain-containing proteins in the cnidarian species searched, show strong amino acid sequence homology to the second TSR repeat in the human TSP1 protein (Figure 3.4). Three-dimensional folding sites described for the TSR domains of human TSP1 (Tan et al. 2002) are present in the cnidarian TSRs: all six cysteine residues are present, forming three disulphide bridges. Arginine residues are present in the motif RXRXR, form salt bridges with other polar residues to form further folding these are referred to by Tan et al. (2002) as the R layers. Three tryptophan residues form the 'WXXWXXW' motif provides protein and glycosaminoglycan-binding sites. The 'CSVTCG' and 'GVQTRXR' motifs, which bind CD36/SRB1, are also very well conserved in cnidarian TSRs.

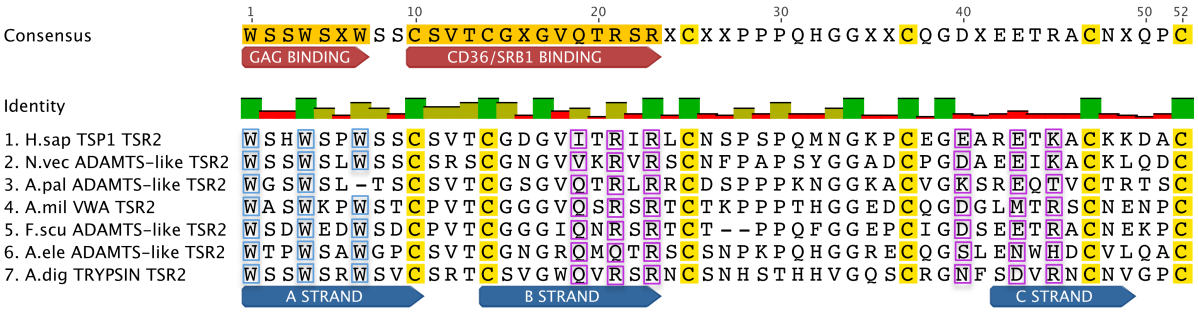



Figure 3.4: The TSR domain is very well conserved from cnidarians to humans with binding motifs for glycosaminoglycans (GAG's) and the type B scavenger receptors, CD36/SRB1. All the three dimensional folding sites are present as described by Tan et al (2002) for the crystal structure of human TSP1 TSR2. Six conserved cysteine residues are highlighted in yellow and form three disulphide bridges (C1-C5, C2-C6 and C3-C4). Three conserved tryptophan residues are shown in blue boxes and mark the 'WXXW' protein binding motif. Amino acids that form the R layers are marked with purple boxes, pairings forming three R layers are as follows R3-R4, R2-R5 and R1-R6. The B strands are annotated at the bottom in blue strands A,B and C. Please refer to Tan et al. (2002) for a more detailed explanation of the three-dimensional folding. Consensus sequence is generated from the cnidarian and human sequences in the above alignment.

### *3.3.3 Cnidarian TSR-domain-containing proteins*

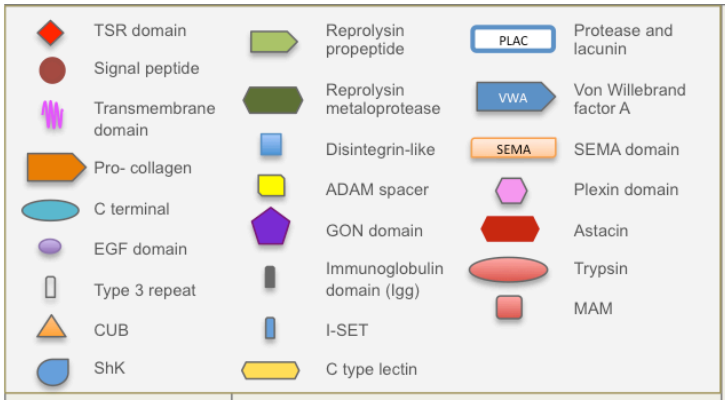
TSR-domain-containing proteins identified from the six cnidarian reources searched were compared to mammalian TSR-domain-containing proteins of known function (Figure 3.5). Sequences obtained from searches of transcriptomes represent proteins identified within the available data, not necessarily the total number of these proteins present in these species. The overall numbers of TSR-domain-containing proteins within the six cnidarian species can be explained by data source: *N. vectensis* and *A. digitifera*, the two species with a full genome sequence, have the highest number of TSR-domain-containing proteins (Figure 3.5).

None of the cnidarians resources searched contained a TSR-domain-containing thrombospondin protein. Putative thrombospondins lacking TSR repeats and a pro-collagen domain and similar to human TSP 3,4 and 5 were identified in all species. The ADAMTS metalloproteases are the largest group of TSR proteins with cnidarian homologues. Searches of the *N. vectensis* genome database recovered 134 contigs encoding predicted proteins that contain one or more TSR domains. 59% of the 134 contigs encode short protein fragments containing only 1-6 TSR domains, while 23% encode ADAMTS metalloprotease-like proteins. Searches of the *Acropora digitifera*

genome database identified 109 TSR-domain-containing genes; 28% of the 109 genes contain only TSR domains, while 17% encode ADAMTS metalloprotease-like proteins. The majority of cnidarian ADAMTS-like proteins lack the disintegrin domain. TSR-domain-containing proteins containing the astacin metallopeptidase domain are found in five of the six species searched. A novel TSR-domain-containing protein, with a trypsin domain and ShK domains, was found in all five of the symbiotic cnidarians searched. Stichodactyla toxin is a potassium channel toxin first discovered in the sea anemone *Stichodactyla helianthus* (Castañeda et al. 1995). VWA-domain-containing TSR proteins were also found in all five of the symbiotic cnidarian species searched. Homologues of the human semaphorin proteins were found in all species, except for *F. scutaria*. Immunoglobulin-domain-containing TSR proteins were only found in three out of the six cnidarian species searched. Large numbers of proteins with TSR repeats alone, and no other identifiable domains, were also found. In comparison, in humans, only the complement factor properdin contains six TSR repeats and no other domains.

### *3.3.4 Phylogenetic analysis of TSR domains reveals high conservation between vertebrates and invertebrates.*

Phylogenetic analysis of the TSR domains in cnidarians, other invertebrates, humans, *S. minutum* and a variety of apicomplexans shows that the TSR domain is very highly conserved (Figure 3.6). Human domains group with invertebrate domains, and TSR domains from very different protein types group closely; for example, the human TSP1 TSR domain clusters with the TSR domain from *Aiptasia* sp. trypsin-containing TSR protein. TSR domains from apicomplexan TRAP proteins group together with domains from *S. minutum* in a single clade.




| Protein name | Protein domain architecture | *Homo sapiens* | *Nematostella vectensis* | *Anthopleura elegantissima* | Symbiotic species: *Aiptasia* sp. | Symbiotic species: *Acropora digitifera* | Symbiotic species: *Acropora millepora* | *Fungia scutaria* |
|---|---|---|---|---|---|---|---|---|
| **Thrombospondins** | | | | | | | | |
| TSP1/2 | | 2 | 0 | 0 | 0 | 0 | 0 | 0 |
| TSP 3/4/5 and COMP | | 4 | 4 | 2 | 3 | 3 | 3 | 2 |
| **ADAMTS Metalloproteases** | | | | | | | | |
| ADAMTS 1-19 Basic structure | | 5 | 0 | 1 | 0 | 1 | 0 | 0 |
| ADAMTS 9 | TSR X14 | 2 | 0 | 0 | 0 | 1 | 0 | 0 |
| ADAMTS 6,7,10,12,16-19 | PLAC | 8 | 0 | 0 | 1 | 1 | 1 | 0 |
| ADAMTS-like 1,2,3,and papilin | PLAC | 4 | 2 | 0 | 1 | 1 | 0 | 0 |
| Cnidarian ADAMTS-like | (1-5 TSR) | 0 | 21 | 1 | 9 | 12 | 15 | 0 |
| | | 0 | 3 | 0 | 2 | 2 | 1 | 0 |
| **Astacin metallopeptidases** | | | | | | | | |
| Cnidarian astacin domain TSR proteins | (2-15 TSR) | 0 | 1 | 2 | 1 | 1 | 1 | 0 |
| | | 0 | 1 | 3 | 2 | 1 | 1 | 0 |
| **Trypsin** | | | | | | | | |
| Cnidarian trypsin domain TSR proteins | | 0 | 0 | 1 | 1 | 1 | 3 | 1 |
| **VWA** | | | | | | | | |
| Cnidarian VWA domain TSR proteins | (2-17 TSR) VWA (1-5 VWA) | 0 | 0 | 2 | 1 | 3 | 3 | 1 |
| **Semaphorins** | | | | | | | | |
| SEMA 5A and B | SEMA | 2 | 1 | 1 | 1 | 1 | 1 | 0 |
| **Immunoglobulin containing TSR proteins** | | | | | | | | |
| Cnidarian immunoglobulin domain proteins | | 0 | 1 | 0 | 0 | 1 | 0 | 0 |
| | | 0 | 1 | 1 | 0 | 2 | 0 | 0 |
| **TSR domains only** | | | | | | | | |
| Properdin | | 1 | 0 | 0 | 0 | 0 | 0 | 0 |
| Cnidarian TSR domain | (1-15 TSR) | 0 | 9 | 7 | 6 | 4 | 2 | 6 |

Figure 3.5: Domain architecture of cnidarian TSR super-family proteins compared to known vertebrate TSR-domain-containing proteins.

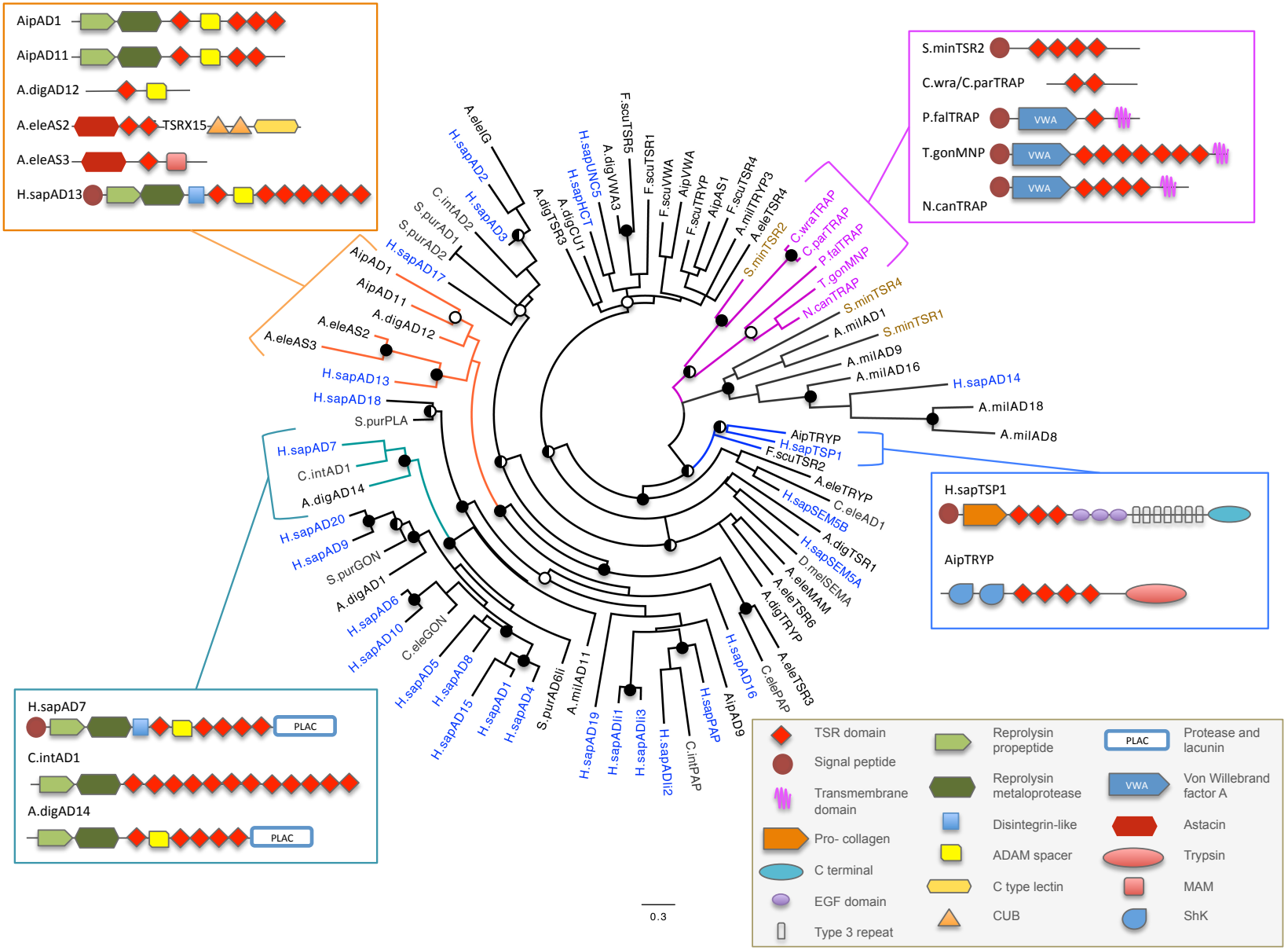

AipAD1
AipAD11
A.digAD12
A.eleAS2
TSRX15
A.eleAS3
H.sapAD13
S.minTSR2
C.wra/C.parTRAP
P.falTRAP
VWA
T.gonMNP
VWA
VWA
N.canTRAP
H.sapAD7
PLAC
C.intAD1
A.digAD14
PLAC
H.sapTSP1
AipTRYP
TSR domain
Signal peptide
Transmembrane domain
Pro- collagen
C terminal
EGF domain
Type 3 repeat
Reprolysin propeptide
Reprolysin metaloprotease
Disintegrin-like
ADAM spacer
C type lectin
CUB
PLAC
Protease and lacunin
VWA
Von Willebrand factor A
Astacin
Trypsin
MAM
ShK
0.3

Figure 3.6: Maximum-likelihood tree of TSR protein domains, constructed using PhyML with 200 bootstraps. Circles at nodes indicate bootstrap support values: full circles 85-100%, half circles 70-85% and open circles 50-70%. Colour coded inlays show the full protein domain organisations and for parts of the tree. TSR domains are depicted by a red diamond. The apicomplexan TRAP protein TSR domains all group together (pink), with one protein from *Symbiodinium minutum* (brown)(S.min). Human (H.sap) reference TSR domain sequences are shown in blue. A.dig, *A. digitifera;* A.mil*, A. millepora;* Aip, *Aiptasia* sp.; A.ele, *A. elegantissima*; N.vec, *N. vectensis*; F.scu, *F. scutaria*; S.pur, *S. purpuratus*; C.ele, *C. elegans*; C.int, *C. intestinalis*; D.mel, *D. melanogaster*. The alignment used to produce this tree is given in Appendix B3.

## 3.4 Discussion

The purpose of this study was to characterise the TSR-domain-containing protein repertoire of both partners involved in the cnidarian-dinoflagellate symbiosis. The TSR protein motif is one of many potential ligands involved in the establishment and regulation of a stable symbiotic relationship. Data presented here reveal a large and diverse TSR-domain-containing protein repertoire in all six cnidarian species searched. The TSR domain is very well conserved from cnidarians to humans, with binding domains for GAGs and CD36/SRB1. The 3-D folding sites, as described by Tan et al (2002) for the crystal structure of the TSR domains in human TSP 1, are present in all six cnidarian species searched.

### *3.4.1 Dinoflagellate TRAP-like proteins*

TSR domain sequences obtained from the *Symbiodinium minutum* genome, in contrast to cnidarian TSR domains, have variable numbers of cysteine residues (4-6), and known binding motifs and 3-D folding sites are not well conserved. Phylogenetic analysis grouped all three *S. minutum* TSR sequences within the same clade as the apicomplexan TRAP TSRs. However, it is not clear from the data presented here whether *Symbiodinium* sp. has a TRAP-like protein that it could use to gain entry into cnidarian host cells. There are similarities between well-studied apicomplexan TRAP proteins and the TSR-domain-containing protein repertoire of *S. minutum*, but further investigation is required to ascertain whether the dinoflagellate is using a TRAP-like protein. Experiments investigating *Symbiodinium* cell surface proteins, using techniques previously used to identify surface lectins (Wood-Charlson et al. 2006, Logan et al. 2010) and techniques used to investigate the release of photosynthetic products (e.g. including protein-containing glycoconjugates) in response to the presence of host tissue (Davy 2001), may be necessary to answer this question.

### *3.4.2 What can be inferred from known functions of vertebrate TSR domains?*

The TSR domains in human TSP 1/2 bind a wide range of ligands and are responsible for many of the functions of the thrombospondin protein. Tan et al. (2002) crystallized

the structure of TSR repeats 2 and 3 in the human thrombospondin-1 protein, revealing many of the essential binding domains. The human TSP1 protein has a large repertoire of cellular functions, many of which appear to be antagonistic. TSP1 promotes cell adhesion and tumour progression, disrupts of cellular adhesion and inhibits of tumour growth (Lopez-Dee et al. 2011). This duality likely stems from the ability to bind multiple receptors presented by a variety of cell types: Receptors include integrins, CD36 and cellular glycosaminoglycans (GAGs) (Lawler 2000).

Human TSP1 TSR domains bind the Class-B scavenger receptors CD36, SRB1 and LIMP in humans; this receptor/ligand interaction has been studied in some detail due to important roles in human disease and immunity (Leung et al. 1992, Li et al. 1993, Silverstein and Febbraio 2007). A CD36-like SRB1 protein is upregulated in the symbiotic state of two different anemones (Rodriguez-Lanetty et al. 2006, Lehnert et al. 2014), however the function of this scavenger receptor is still unknown. The TSR protein domain provides a potential ligand for the upregulated SRB1. Possible roles for scavenger receptors in relation to the establishment and maintenance of a healthy cnidarian-dinoflagellate symbiosis are discussed in detail in chapter 2. Briefly, firstly SRB1 could be used by the symbiont to gain entry to the host cell as in an immune receptor capacity SRB1 initiates phagocytosis of microbes (Schäfer et al. 2009, Silverstein and Febbraio 2009). Secondly, in humans TSP1 protein binds CD36 in order to activate latent TGFβ and initiate the TGFβ pathway, which has an immune suppressive effect that enables tumour progression (Khalil 1999, Murphy-Ullrich and Poczatek 2000, Koli et al. 2001, Uchida et al. 2008), this tolerogenic response could be beneficial to an invading dinoflagellate cell.

### *3.4.3 Which TSR proteins are present in cnidarians?*

The six cnidarian species searched in this study have genes that encode proteins with only TSR domains, in repeats from 1-15. In humans, the complement factor properdin is a protein that contains only six repeat TSR domains and no other protein domains. Properdin is involved in tissue inflammation and pathogen engulfment by phagocytes (Nolan et al. 1992); initiating phagocytosis is an important mechanism for establishing an endosymbiotic relationship. The TSR domain could thus represent a mechanism by which the dinoflagellate uses the host cell's immune receptors to initiate phagocytosis

and gain entry to the host gastrodermis. This mechanism is used by *Plasmodium* spp. and the Hepatitis C virus to enter human hepatocyte cells (Areschoug and Gordon 2009). In both cases, the host receptor is the scavenger receptor SRB1, and in the case of *Plasmodium* the ligand used is the TSR domain of a TRAP protein (Müller et al. 1993, Yalaoui et al. 2008).

ADAMTS metalloprotease-like proteins form the second largest group of TSR-domain-containing proteins identified in five of the six cnidarian species searched. The ADAMTS metalloproteases are a group of disintegrin metalloproteases with thrombospondin motifs. In humans there are 19 ADAMTS genes that encode for secreted proteases, some of which bind to the ECM (Tang and Tang 2001). The TSR domains within these ADAMTS proteins anchor the protease to the extracellular matrix (Kuno et al. 1998). The TSR domain is involved in proteolysis and can cleave the proteoglycans, aggrecan, versican, brevican and neurocan (Porter et al. 2005). ADAMTS metalloproteases have important roles in connective tissue organization, inflammation, angiogenesis and cell migration (Kuno et al. 1998, Tang and Tang 2001, Davis et al. 2009).

The ADAMTS proteases are much less common in non-mammalian metazoans. There are five ADAMTS genes in *C. elegans* and three ADAMTS-like protein sequences in *Drosophila melanogaster*. The ascidian, *Ciona intestinalis*, has six ADAMTS proteases and the purple sea urchin, *Strongylocentrotus purpuratus*, has six ADAMTS-like proteins (Nicholson et al. 2005, Huxley-Jones et al. 2005, Fernando et al. 2011). Four of the six cnidarian species searched show expanded numbers of ADAMTS genes compared to the numbers of ADAMTS genes identified in other invertebrate model organisms. *N. vectensis* has 26 ADAMTS-like genes, while *A. digitifera* and *A. millepora* both have 17 and *Aiptasia* sp. has 13. The number of ADAMTS genes found here is closer to that of vertebrates than to other invertebrate species. In humans, ADAMTS 13 binds CD36, localizing the protease on the endothelial cell surface where it regulates the cleavage of VWA. The ADAMTS-like proteins thus have potential roles in anchoring dinoflagellates to cnidarian gastrodermal cells, dinoflagellate cell migration within the gastrodermis, and initiation of an immune response.

A novel member of the TSR-domain-containing superfamily was found in all five symbiotic cnidarians, but not in the non-symbiotic anemone *N. vectensis*. This protein contains 3-5 TSR repeats, the toxin ShK domain and a terminal trypsin domain. The presence of the trypsin domain suggests that this protein is involved in protein

digestion, but it is not clear why this protein would be found in symbiotic cnidarians and not in the non-symbiotic species. TSR proteins containing 1-5 VWA domains and between 2-15 TSR domain repeats were also found in the five symbiotic cnidarians but not in the non-symbiotic anemone *N. vectensis*. These VWA TSR proteins are unlike any known vertebrate TSR proteins, but they are very similar to the CTRP protein of *P. falciparum* and the other TRAP proteins. The TRAP protein binds to hepatocytes through a dual ligand system that uses the TSR and VWA domains; while invasion is possible when one or other domain bears a mutation, invasion is more successful when both domains are present (Morahan et al. 2009). Due to their exclusive presence in symbiotic species, both of these TSR proteins are potential targets for further study in the context of cnidarian-dinoflagellate symbiosis.

A homologue to the human SEMA5 protein was found in five of the six cnidarian species searched. The semaphorins are a family of growth-guidance proteins that guide axons during neural development, initiate bone growth, and bind to B1 plexin, while SEMA4 and -7 are involved in immune functions (Kruger et al. 2005). SEMA5 is highly evolutionarily conserved and contains between five and seven TSR repeats. Homologues to the human SEMA5 have been found in mice, flies, zebrafish and now cnidarians. Very little is known about the specific functions of SEMA5 and its TSR domains. It may be involved in cell morphology, cytoskeletal organisation, neural connectivity and vasculature patterning in humans (Yazdani and Terman 2006). A homologue to the human B1 plexin transmembrane receptors has also been identified in *N. vectensis* (Putnam et al. 2007).

### *3.4.4 Data constraints on protein analysis: data source, developmental stage and symbiotic state*

The genetic data used in this study represent various developmental stages and symbiotic states. Fully annotated genomes are only available for the non-symbiotic anemone *N. vectensis* and the symbiotic coral *A. digitifera*, while a draft genome assembly is available for the dinoflagellate *S. minutum*. The four remaining cnidarian species used in this study have only transcriptome data available. While these data are valuable, they can only provide a snapshot of the proteins expressed at the time the transcriptome was made. Therefore, it is not possible to compare numbers of specific

proteins found between species for which transcriptome data are available. With the exception of the coral *A. millepora*, the transcriptomes were all produced using aposymbiotic tissues, and hence the lack of symbionts could affect the types of proteins expressed and could exclude symbiosis-specific proteins. Indeed, several studies have shown differential gene expression when comparing transcriptome data from symbiotic *versus* aposymbiotic anemone tissue (Rodriguez-Lanetty et al. 2006, Moya et al. 2012, Lehnert et al. 2014). These data constraints can explain several of the discrepancies in the numbers of proteins. For example, both *A. elegantissima* and *F. scutaria*, for which only transcriptome data are available, have low numbers of ADAMTS-like proteins, whereas all other species searched have high numbers of these proteins. Indeed, the transcriptome for *F. scutaria* has low numbers of all TSR proteins compared to the five other species searched. Further comparisons between genomic and transcriptomic data at various developmental stages and symbiotic states within the same species could provide a more detailed picture of the expression patterns of these genes within cnidarians.

### *3.4.5 Phylogenic analysis of TSR domains reveals high conservation*

The patterns observed in the protein domain phylogeny presented here indicate that TSR domain sequences are very similar, with the exception of the apicomplexan TRAP TSR domains. TSR domains for all taxa sampled group together and the TSR domains from very different proteins also group together. For example, the TSR domain from a trypsin-domain-containing protein from the anemone *Aiptasia* sp. groups with the TSR domain from the human TSP1 protein. The TSR domains from ADAMTS proteins of all species searched here group together; these groups include Anthozoan TSR proteins of all types, while some dinoflagellate TSR domains group with human ADAMTS14 and *A. millepora* ADAMTS-like protein domains. The phylogeny indicates that the TSR domains across taxa and different TSR-domain-containing proteins are highly similar. Conservation of 3-D folding and binding motifs would suggest that the TSR domains and proteins are interchangeable between systems. This hypothesis was tested in the next chapter, where treatment of *Aiptasia* sp. anemones with human TSP1 protein produced the same effect as peptides derived from native *Aiptasia* sp. TSR proteins.

## 3.5 Concluding remarks

The bioinformatic searches and analyses carried out in this study suggest that the TSR domain, its functions and binding sites are ancient and highly conserved. Anthozoan cnidarians contain a large and diverse TSR protein repertoire, exhibiting similarities to the human TSR protein repertoire and also some novel proteins, which could be potential targets for further investigation. More work is needed to determine the function and evolutionary history of these uncharacterized proteins. This study provides the first detailed description of TSR-domain-containing proteins in cnidarians and contributes to basic knowledge about the secreted and ECM protein repertoire for this basal metazoan group. The ECM is the first line of defense for cells against invading microbial attack and is therefore important to research on coral disease and immunity, as well as in the context of mutualistic partners such as *Symbiodinium* spp. These genomic data enable and inform ongoing work investigating a potential role for TSR-domain-containing proteins in the establishment of a healthy cnidarian-dinoflagellate symbiosis.

# Chapter 4

# The thrombospondin structural homology repeat (TSR) domain is important in successful dinoflagellate infection of the anemone *Aiptasia* sp.

## 4.1 Introduction

The innate immune system is responsible for managing the microbiome, maintaining a balance between tolerating beneficial microbes and a response to pathogenic microbial attack. It responds to microbial attack by recognising microbe associated molecular patterns (MAMPs) in the form of tell-tale microbial surface molecules such as glycans and lipopolysaccharides (LPS). Host pattern recognition receptors (PRRs) – located on host membranes or in the extracellular matrix (ECM) – recognise MAMPs and can trigger phagocytosis of invading microbes, quickly leading to a cascade of phagolysosome maturation and ending in proteolytic degradation.

Many intracellular parasites manipulate host defence mechanisms to their own advantage. Both the malaria parasite *Plasmodium* spp. and the hepatitis C virus have ligands recognised by the scavenger receptor (SR) SRB1 and exploit the phagocytic mechanisms of SRB1 to gain entry to host hepatocytes (Rodrigues et al. 2008, Catanese et al. 2010). Many microbial pathogens avoid host recognition or dampen immune activation through sophisticated signalling, however some pathogens benefit from the stimulation of host defence pathways. The parasite *Trypanosoma cruzii* requires the presence of thrombospondin 1 (TSP1) protein in the ECM of host cells and the activation of the transforming growth factor-β (TGFβ) signalling pathway for entry into mammalian cells (Ming et al. 1995, Waghabi et al. 2005, Simmons et al. 2006, Nde et al. 2012). TGFβ is a regulatory cytokine; its main role in the immune system is to maintain tolerance *via* the regulation of lymphocyte proliferation, and to resolve inflammation (Li et al. 2006). This tolerance-promoting immune pathway is exploited by pathogens to facilitate entry, replication, and persistence within the host. Both the intracellular parasites *Leishmania major* and *Toxoplasma gondii* have developed mechanisms to

induce macrophages to produce high levels of active TGFβ, which in turn suppresses nitric oxide (NO) production and reduces an inflammatory response (Li et al. 2006).

The TGFβ pathway has the potential to be both beneficial and detrimental to the pathogen. In cancer, TGFβ has been described as a molecular 'Jekyll and Hyde', acting as both a tumour suppressor by initiating apoptosis, and as a tumour promoter by enhancing growth, invasion and metastasis (Bierie and Moses 2006, Massagué 2008, Ikushima and Miyazono 2010). Activation of the TGFβ pathway in vertebrates is not fully understood, however in many cases the activation of latent TGFβ into its active form requires the interaction of the TSP1 protein with CD36, a Class B scavenger receptor found in the ECM (Khalil 1999). Within TSP1, the TSR domain specifically binds the CD36 domain on the CD36 protein (Murphy-Ullrich and Poczatek 2000) (see Figure 4.1a for an illustration of a simplified TGFβ signalling pathway). The mechanisms employed by obligate intracellular parasites to gain entry to host cells and manipulate immune responses to their own benefit are of particular interest to the study of mutualistic relationships such as the cnidarian-dinoflagellate symbiosis.

The cnidarian-dinoflagellate symbiosis is an intracellular relationship where dinoflagellate symbionts of the genus *Symbiodinium* reside within the host gastrodermis, the innermost cell layer lining the gastrovascular cavity. The relationship is based on mutual nutritional exchange: the symbiont translocates fixed carbon to the host in return for inorganic nutrients and a high light environment (Yellowlees et al. 2008). In the majority of cnidarian species, symbionts must be acquired from the surrounding seawater by cnidarian larvae. The mechanism by which the symbiont colonizes the host and is tolerated is a subject of recent interest (Dietrich et al. 2006, Schwarz 2008, Davy et al. 2012, Lehnert et al. 2014). The symbiont enters the host *via* phagocytosis, and becomes surrounded by a phagosome composed of host-derived plasma membrane that develops into a specialised vacuole referred to as the 'symbiosome' (Kazandjian et al. 2008, Peng et al. 2010).

The establishment of a healthy cnidarian-dinoflagellate symbiotic relationship involves a series of complex steps: symbiont recognition and uptake *via* phagocytosis, selection, arrest of phago-lysosome maturation and cell cycle control, followed by symbiont proliferation within the host cells (Davy et al. 2012). Each step involves inter-partner

signalling and a degree of cooperation between partners. Several recent genomic studies have uncovered a complex innate immune repertoire in cnidarians (Miller et al. 2007, Bosch 2008, Kvennefors et al. 2010, Lehnert et al. 2014, Baumgarten et al. 2015). Partner crosstalk is complex and involves a vast array of ligands, receptors and protein cascades. Individual studies are contributing to our understanding of the complexity of the molecular mechanisms that support cnidarian-dinoflagellate symbiosis (Chen et al. 2004, Wood-Charlson et al. 2006, Wood-Charlson and Weis 2009, Detournay and Weis 2011, Meyer and Weis 2012, Detournay et al. 2012).

In figure 4.1b, I introduce a testable cellular model for how the TGFβ pathway may be involved in cnidarian-dinoflagellate symbiosis. A previous study in *Aiptasia* identified several proteins within the TGFβ immune pathway and proposed a tolerogenic balance between host and symbiont (Detournay et al. 2012). A TGFβ *sensu stricto* has been characterized in *Aiptasia* sp., along with the Smad transcription factors that function downstream of TGFβ activation (Technau et al. 2005, Detournay et al. 2012). Phosphorylated Smad proteins 2/3 were present in higher quantities in symbiotic *versus* aposymbiotic *Aiptasia* sp., *via* immunoblot analysis (Detournay et al. 2012). In functional genomic studies examining genes differentially expressed in aposymbiotic *versus* symbiotic states, the class B SR, SRB1, is upregulated in the symbiotic state of both the anemones *A. elegantissima* and *Aiptasia* sp. (Rodriguez-Lanetty et al. 2006, Lehnert et al. 2014). The up-regulation of this protein post-phagocytosis indicates a potential role in symbiosis maintenance. Both CD36 and TSR-domain-containing proteins identified in this thesis (Chapters 2 and 3) provide new putative TGFβ pathway components.

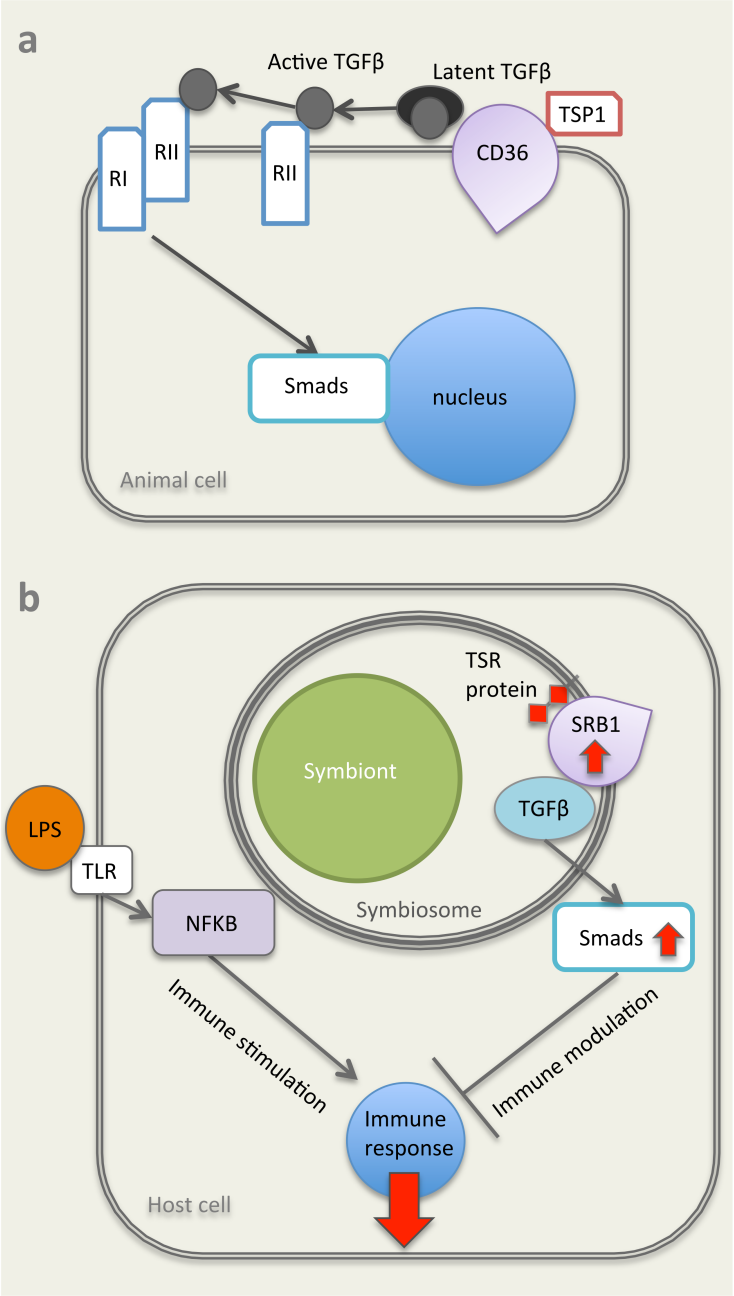



Figure 4.1: a) A simplified TGFβ signaling pathway. TGFβ is secreted in a latent form, and activation of latent TGFβ to its active form often involves the TSR domain of the thrombospondin 1 protein binding to the Class B scavenger receptor, CD36 (Schultz-Cherry et al. 1994; Murphy-Ullrich et al. 2000). TGFβ activation initiates a signalling cascade *via* transmembrane TGFβ Receptors I and II, activating the phosphorylation of the Smad transcription factors. Smads migrate to the nucleus and promote or inhibit transcription of a large variety of genes (Li et al. 2006). b) Proteins and potential elements of the TGFβ pathway identified in cnidarians. Previous work supports the involvement of TGFβ in promoting tolerance of symbionts within the cnidarian host (Detournay et al. 2012). Upward pointing red arrows indicate proteins identified as upregulated in the symbiotic state: SRB1 (Rodriguez-Lanetty et al. 2006; Lehnert et al. 2014); Smad 2/3 (Detournay et al. 2012). Downward pointing red arrow indicates a down-regulation of the host immune response. In comparison to aposymbiotic *Aiptasia* sp., symbiotic anemones produce significantly less NO in response to immune stimulation by LPS (Detournay et al. 2012).

Experimental evidence suggests that there is a tolerogenic response to the symbiont *via* activation of the TGFβ immune suppressive pathway (Detournay et al. 2012). Blocking the putative cnidarian TGFβ protein reduced infection success in re-infection experiments using aposymbiotic *Aiptasia* sp. anemones. Blocking TGFβ ligands also induced an immune response in symbiotic anemones in the form of increased NO production (Detournay et al. 2012). In comparison to aposymbiotic *Aiptasia* sp, symbiotic anemones produce significantly less NO in response to immune stimulation by LPS (Detournay et al. 2012). This response is mirrored in vertebrate macrophages parasitized by *Leishmania major* and *Toxoplasma gondii*, where NO production is suppressed due to the tolerogenic effects of the TGFβ pathway (Li et al. 2006).

Searches of the available genomic and transcriptomic resources for six cnidarian species revealed a large and complex repertoire of TSR-domain-containing proteins in cnidarians (Chapter 3). Several apicomplexan parasites use a TSR domain-containing protein referred to as the "thrombospondin related anonymous protein" (TRAP) to bind to and gain entry to host cells (Morahan et al. 2009). The TSR domains in cnidarian TSR-domain-containing proteins show strong amino acid sequence conservation when compared to vertebrate TSR domains: The 3-D folding sites and significant binding motifs are highly conserved, indicating that this domain is ancient and highly conserved from vertebrates to lower metazoans (Chapter 3). The characterisation of cnidarian and dinoflagellate TSR domains and the TSR-domain-containing repertoire thus provides targets for the functional investigation of these proteins in the onset of the cnidarian-dinoflagellate symbiosis.

The aim of this chapter was to investigate the role of TSR-domain-containing proteins in the establishment of the cnidarian-dinoflagellate symbiosis within the anemone *Aiptasia* sp. This anemone represents a model system, for the study of cnidarian-dinoflagellate symbiosis, as it can be maintained in both symbiotic and aposymbiotic states in the laboratory (Dietrich et al. 2006). The central hypothesis tested is that a TSR-domain-containing protein is involved in symbiont uptake. Specifically I focussed on: (1) whether TSR-domain-containing proteins involved are of symbiont or host origin; and (2) if the dinoflagellate uses a TRAP-like TSR-domain-containing surface protein to gain entry to host cells or alternatively if the symbiont utilises host TSR-domain-containing protein(s) to gain entry to host cells. Western blots showed that TSR-domain-containing proteins are present in higher quantities in symbiotic *versus*

aposymbiotic anemones. Functional experiments measured symbiont uptake in aposymbiotic anemones, after challenge with proteins and antibodies to either stimulate or block TSR-domain-binding. Overall, results suggest that host derived TSR-domain-containing proteins are important to symbiont acquisition in the anemone *Aiptasia* sp. Immunofluorescence microscopy of symbiotic anemones showed that anti-TSP localises in gastrodermal cells containing symbionts. These images suggest a potential role for TSR-domain-containing proteins in symbiosis maintenance. This study provides information on the role of a novel group of proteins in symbiosis onset and also potential mechanisms of inter-partner signalling and innate immune modulation.

## 4.2 Materials and Methods

### *4.2.1 Maintenance and preparation of anemone and dinoflagellate cultures*

Symbiotic *Aiptasia* cultures were maintained in saltwater aquaria at 26°C with a 12/12 h light/dark photoperiod, and were fed twice weekly with live brine shrimp. Animals were rendered aposymbiotic by incubation for 8 h at 4 °C, after a few days recovery this cold shock was repeated 4-5 times, followed by maintenance in the dark. Anemones were fed daily with brine shrimp, and cleaned of expelled symbionts and food debris daily. Aposymbiotic animals were verified fusing a florescence microscope checking for dinoflagellate chloroplast auto-fluorescence.

Cultured dinoflagellates, *Symbiodinium* sp., ITS2 type B1 (culture ID: CCMP830) were maintained in 50 ml flasks in sterile Guillard's f/2 enriched seawater culture medium (Sigma, St. Louis, MO, USA). Dinoflagellate cultures were maintained at 26°C on a 12/12 h light/dark photoperiod.

### *4.2.2 Western blot analysis of TSP1 antibody protein targets*

Western blots were performed to assess the expression levels of TSR proteins binding to the rabbit polyclonal antibody with an epitope corresponding to the three TSR domains of human thrombospondin proteins 1 and 2 (H-300, Santa Cruz Biotechnology sc-14013). The epitope used to make the antibody showed significant similarity with a TSR-domain-containing protein identified in the experimental anemone *Aiptasia* sp. (Figure 4.2). Sets of eight aposymbiotic or symbiotic anemones were homogenized on

ice in 1 ml of homogenization buffer (50 mM Tris–HCl, pH 7.4, 300 mM NaCl, 5 mM EDTA) with a protease inhibitor cocktail (BD Biosciences, San Jose, CA, USA). Homogenates were centrifuged at 4°C for 15 min at 14,000 x g, supernatants were decanted and protein concentrations were determined using the Bradford assay. Protein concentrations were adjusted or diluted to a standard (50 µg/ml) and boiled for 5 min in loading dye. Proteins were resolved on a 7% SDS–PAGE gel. Proteins were electrophoretically transferred overnight onto nitrocellulose membrane. After blocking with 5% non-fat dry milk in TBS-Tween 20 (0.1%) for 1 h at 37 °C, membranes were incubated with anti-TSP or an IgG isotype control, both at a dilution of 1:200, for 2 h at room temperature. The blots were washed three times in TBS-Tween 20 followed by incubation in a secondary HRP-conjugate goat anti-rabbit IgG (0.2 lg/ml, Sigma, St. Louis, MO, USA) for 1 h. Each sample contains eight anemones; protein concentrations standardised to contain 50 µg/ml of protein per well. Bands were detected by enhanced chemiluminescence (Millipore, Temecula, CA, USA). Relative band intensities were quantified with IMAGE J software.


Figure 4.2: A section of the antibody binding region of the human thrombospondin 1/2 antibody (H-300, sc-14013 from Santa Cruz Biotechnology), aligned to a TSR protein fragment from *Aiptasia* sp.

### *4.2.3 Cryosectioning and immunofluorescence microscopy to localise binding of the human TSP antibody*

Immuno-fluorescence was used to investigate the presence of TRAP-like proteins on the surface of dinoflagellate cells. I compared anti-TSP binding in *Symbiodinium* cells from

cultured algal strain CCMP830 to algal cells freshly isolated from *Aiptasia* sp. To obtain freshly isolated algal cells with intact symbiosome membranes, anemones were homogenised in a microfuge tube with a micro-pestle and the resulting homogenate was centrifuged to produce an algal pellet. The pellet was washed several times in FSW and re-pelleted. Algal cells were re-suspended to a concentration of 25,000 cells per ml. Both cultured and freshly isolated *Symbiodinium* cells were incubated with the anti-TSP conjugated to a fluorescent probe. The lipophilic membrane stain, 'DiI' (1,1'-Dioctadecyl-3,3,3',3'-Tetramethylindocarbocyanine Perchlorate; DiIC18(3); Molecular Probes), was used to test for the presence of host derived plasma symbiosome membrane surrounding freshly isolated algal cells.

To investigate where anti-TSP binds within both symbiotic and aposymbiotic anemones, 20um thick cryosections of both whole symbiotic and aposymbiotic anemones were made using methods modified from (Dunn et al. 2007b) (see: http://people.oregonstate.edu/~weisv/assets/cryosectioning.pdf for the full protocol pdf). Whole anemone sections were frozen to slides and stored at -20°C. Slide-mounted sections were then thawed to room temperature for further processing. The sections were washed twice in phosphate-buffered saline (PBS) and fixed with 4% paraformaldehyde (PFA) for 10 min, and then washed twice in PBS. Sections were then made permeable with 0.2% Triton-X-100 in PBS for 5 min and blocked in 3% bovine serum albumen (BSA), 0.2% Triton-X-100 in PBS for 30 min, before being incubated in the anti-TSP rabbit polyclonal antibody at a 1:200 dilution (in blocking buffer) for 4 h at 4°C. Slides were subsequently washed three times for 5 min each with 0.2% Triton-X-100 in PBS at room temperature. Alexa Fluor 488 (ThermoFisher Scientific) secondary fluorescent antibody was diluted in blocking buffer (1:150 dilution) and incubated on the slides at room temperature in the dark for 1 h. Slides were washed in the dark three times for 5 min with 0.2% PBS/Triton-X-100. A drop of Vectashield DAPI hard set mounting medium was then used to stain nuclei and mount cover slips onto slides. Immunofluorescence was visualised using the Zeiss LSM 510 Meta microscope through a Plan-APOCHROMAT 63x/1.4 Oil DIC objective lens. The fluorescence excitation/emission was 488/556-573 nm for Alexa Fluor 488 secondary antibody and 688/679-754 nm for *Symbiodinium* chlorophyll autofluorescence.

### *4.2.4 Experimental manipulation of anemones*

In preparation for experimental manipulations, individual anemones were placed in 24-cell well plates in 2.5 ml of 1-µm filtered seawater (FSW) and acclimated to the well plate for 4 days, with FSW replaced daily. Well plates containing aposymbiotic anemones were exposed to as little light as possible and symbiotic anemones were maintained in an incubator at 26°C with a 12/12 h light/dark photoperiod. Animals were not fed during the experimental time period.

Aposymbiotic *Aiptasia* sp. anemones were subjected to experimental re-infection with *Symbiodinium* cells; infection success was determined by counting the number of symbionts present in host gastrodermal cells (see 4,2,7 for full details). Experimental treatments (see below) were initiated 2 hours prior to infection with *Symbiodinium*. For infection experiments, cultured *Symbiodinium* sp. cells of ITS2 type B1 (culture ID: CCMP830) were added to each well to give a final concentration of 2 x $10^5$ cells/ml. After incubation with dinoflagellate cells for 4 h in an incubator at 26°C under 40 µmol quanta/$m^2$/s of light, anemones were washed twice in FSW and experimental treatments were refreshed. Well-plates were then placed in an incubator at 26°C under 40 µmol quanta/$m^2$/s of light and a 12/12 h light/dark cycle.

### *4.2.5 Addition of anti-TSP to aposymbiotic anemones during the onset of symbiosis*

To investigate the effects of blocking TSR domains at the onset of the symbiosis, anemones were incubated in a rabbit polyclonal antibody with an epitope corresponding to the three TSR domains of human thrombospondin proteins 1 and 2 (H-300, Santa Cruz Biotechnology sc-14013). Anemones were incubated for 2 h prior to infection with dinoflagellate cells (described above) in anti-TSP at a concentration of 0.5 µg/ml. Control animals were given fresh FSW at the same time. A second IgG control to control the use of an antibody used normal rabbit IgG isotype/pre-immune serum at 1000 µg/ml also from Santa Cruz (sc:2027). Anemones were sampled at 48, 72, 96 and 120 h post-infection, with a sample size of four anemones per treatment per time-point. Treatment conditions of these animals were refreshed once every 24 h.

### *4.2.6 Addition of TGFβ protein to anemones already treated with anti-TSP*

To investigate if the inhibitory effect on infection success observed with the addition of anti-TSP could be reversed, TGFβ protein was added to anemones pre-treated with anti-TSP. The hypothesis supporting this experiment suggests that TSR-domain-containing proteins are involved in initiating the tolerogenic TGFβ pathway. For an overview of the TGFβ pathway components see Figure 4.1. Recombinant TGFβ protein (Sigma, St.Louis, MO, USA) at a concentration of 100 µg/ml was added at the time of infection to well-plates containing aposymbiotic anemones pre-treated (2 hours prior) with human thrombospondin antibody at a concentration of 0.5 µg/ml. Control animals were given fresh FSW at the same time. Anemones were sampled at 65 and 95 h post-infection; sampling time-points were limited by the available quantity of TGFβ protein. Treatment conditions of these animals were refreshed once every 24 h.

### *4.2.7 Addition of human TSP1 protein and synthetic TSR peptides*

To investigate the effect of TSR-containing proteins on dinoflagellate infection success, soluble human thrombospondin-1 protein (TSP-1) (thrombospondin human platelet, Athens research and technology, #:16-20-201319) was added to aposymbiotic anemones. Anemones were incubated for 2 h prior to infection in 25 µg/ml human TSP-1 in FSW. Control animals were given fresh FSW at the same time. Anemones were sampled at 48, 72, 96 and 120 h post-infection, with a sample size of four anemones *per* treatment per time-point. Treatment conditions of these animals were refreshed once every 24 h.

To investigate whether native *Aiptasia* sp. TSR domains would produce a similar effect to human TSP1 protein, anemones were incubated in synthetic TSR peptides. Several studies have used TSR peptide fragments to investigate the binding sites of specific receptors such as the scavenger receptor type B CD36 (Tolsma et al. 1993, Li et al. 1993, Karagiannis and Popel 2007, Cano et al. 2009). The putative TSR domain from *Aiptasia* sp. contains multiple binding motifs, WXXWXXW, CSVTCG and GVQTRLR, which are all known to bind glycosaminoglycans and class B scavenger receptors in humans (Figure 4.3). Two separate peptides were designed: Peptide 1 was identical to TSR domain 2 from the predicted protein *Aiptasia* sp. comp25690 (taken from an *Aiptasia* sp.

transcriptome (Lehnert et al. 2012)). Peptide 2 cysteine residues in peptide 1 were substituted with alanine residues to avoid peptide self-adhesion and resultant loss of adhesion to target molecules. Peptides were designed according to the peptide design guidelines (version 3) available at www.biomatik.com. Anemones were incubated for 2 h prior to infection in either of the two peptides at a concentration of 150 µg/ml of FSW. Control animals were given fresh FSW at the same time. Anemones were sampled at 48, 72, 96 and 120 h post-infection, with a sample size of four anemones *per* treatment *per* time-point. Treatment conditions of these animals were refreshed once every 24 h.

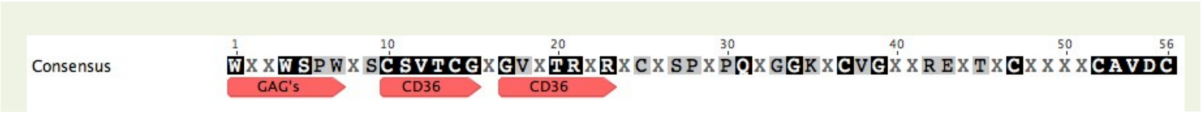



Figure 4.3: Alignment of the second TSR domains from human thrombospondin 1 and TSR proteins from the anemone *Aiptasia sp*. and the dinoflagellate *Symbiodinium minutum.* In red are the binding sites for glycosaminoglycans (GAGs) and CD36; greyscale indicates the % identity of the three sequences. Pink annotation indicates the TSR peptide sequence covering all three binding domains; inset are the synthetic peptide sequences for experimental peptides. In peptide 2, the cysteine residues were replaced with alanine residues as shown in red.

### *4.2.8 Assessing infection success using confocal microscopy*

Symbiont numbers were quantified in host tentacles using confocal imaging. Infection success was assessed flourometrically by confocal microscopy, following methods detailed in Detournay et al. (2012). Following experimental manipulation, anemone treatments were replaced with 1 ml of relaxing solution (1:1 0.37 M $MgCl_2$ : FSW). Relaxing solution is used to keep the live anemone still and prevent it from contracting. DAPI (4',6-Diamidino-2-Phenylindole, Dihydrochloride) ThermoFisher scientific, (catalog number: D1306) was added to the relaxing solution to stain nuclei. Samples were observed under a Zeiss LSM 510 Meta confocal microscope with a 40x/0.8 water

objective lens and a working distance of 0.8–3.2 mm. Dinoflagellate cells present were visualised by detecting chlorophyll autofluorescence with excitation and emission wavelengths of 543 and 600-700 nm, respectively. Before image scanning, the focal plane of the optical section was adjusted to include the gastrodermal cells within the anemone tentacle. For each experiment, all images were obtained with the same software scanning settings, including detector gain and laser intensity. Quantification of fluorescence was achieved by first defining the gastrodermal tissue area within the anemone tentacles as a region of interest and then measuring the mean fluorescence intensity (MFI) for that region with the LSM 5 software (Zeiss). Chlorophyll autofluorescence intensity for each pixel was measured and a threshold value corresponding to the background was defined by measuring the mean fluorescence intensity (MFI) at 600 nm of a gastrodermal section without symbionts (threshold MFI = 20). Infection success was expressed as percent of pixels with autofluorescence intensity above the threshold. In reinfection experiments, each treatment represents a sample size of four anemones per treatment and time-point, with % infection taken as the mean over six tentacles per anemone. Three untreated symbiotic anemones (six tentacles per anemone) were examined to determine a baseline infection level for symbiotic anemones.

## 4.3 Results

### *4.3.1 Evidence of TSR domain proteins in host but not symbiont*

Anti-TSP labelled the same two bands between 72 and 43 kDa in Western blot analysis of homogenate from both symbiotic and aposymbiotic *Aiptasia* sp. (Figure 4.4). However band intensities were higher in symbiotic, compared to aposymbiotic samples. The exact nature of these target proteins remains unknown, as immuno-precipitation and mass spectrometry failed to resolve this matter (see Appendix C1).

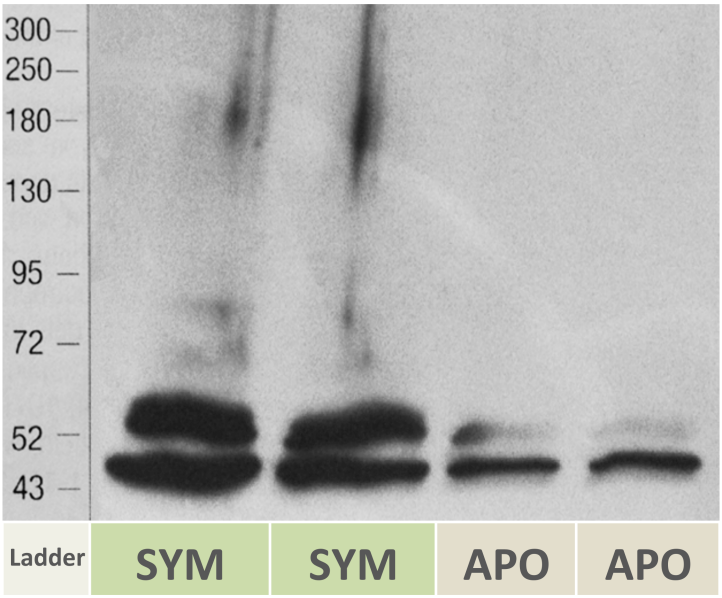



Figure 4.4: TSP antibody binding proteins are expressed in higher quantities in symbiotic *Aiptasia sp.* than in the aposymbiotic state of this anemone. Western blot image shows protein from symbiotic (labelled SYM) and aposymbiotic (labelled APO) anemones. All four wells showed bands at approx. 47 kD and approx. 53-55 kD. Stronger intensities of bands in symbiotic samples indicate larger quantities of TSR domain-containing proteins binding to the TSP antibody.

Anti-TSP labelled freshly isolated, but not cultured *Symbiodinium* cells (Figure 4.5). Label appeared around the outside of cells, suggesting that it was labelling symbiosome membrane or host material associated with freshly isolated algae. Freshly isolated algae were surrounded by a host-derived symbiosome membrane and/or gastrodermal cell membranes, as indicated by the presence of DIL membrane stain fluorescence (Figure 4.5).

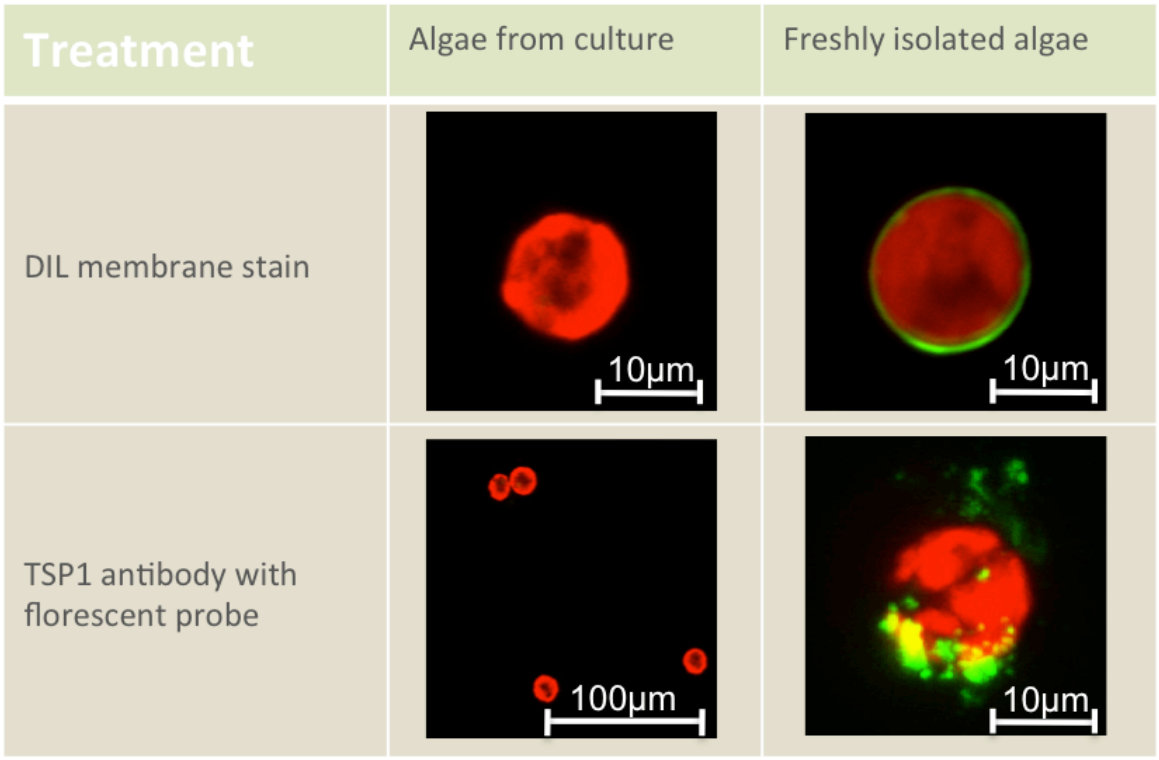



Figure 4.5: Confocal microscope images of dinoflagellate cells taken from culture (left) and freshly isolated cells taken from *Aiptasia sp*. anemone homogenate (right). The DIL membrane stain treatment indicated anemone plasma membrane (symbiosome membrane) present around the freshly isolated cells and absent around the cultured cells. A fluorescent probe conjugated to the TSP1 antibody showed that the antibody bound to a target protein within the anemone-derived plasma membrane present on the freshly isolated cells. The lower resolution image (bottom left) shows no binding in the cell culture.

Anti-TSP labelled anemone tissue closely associated with symbionts in whole mount specimens (Figure 4.6). Labelling was localised within host gastrodermal cells: both host and symbiont nuclei labelled blue with DAPI indicating symbionts tightly enclosed within host gastrodermal cells (Figure 4.6 c & d). Green anti-TSP labelling is visible surrounding the red autofluorescence of the dinoflagellate chloroplast.

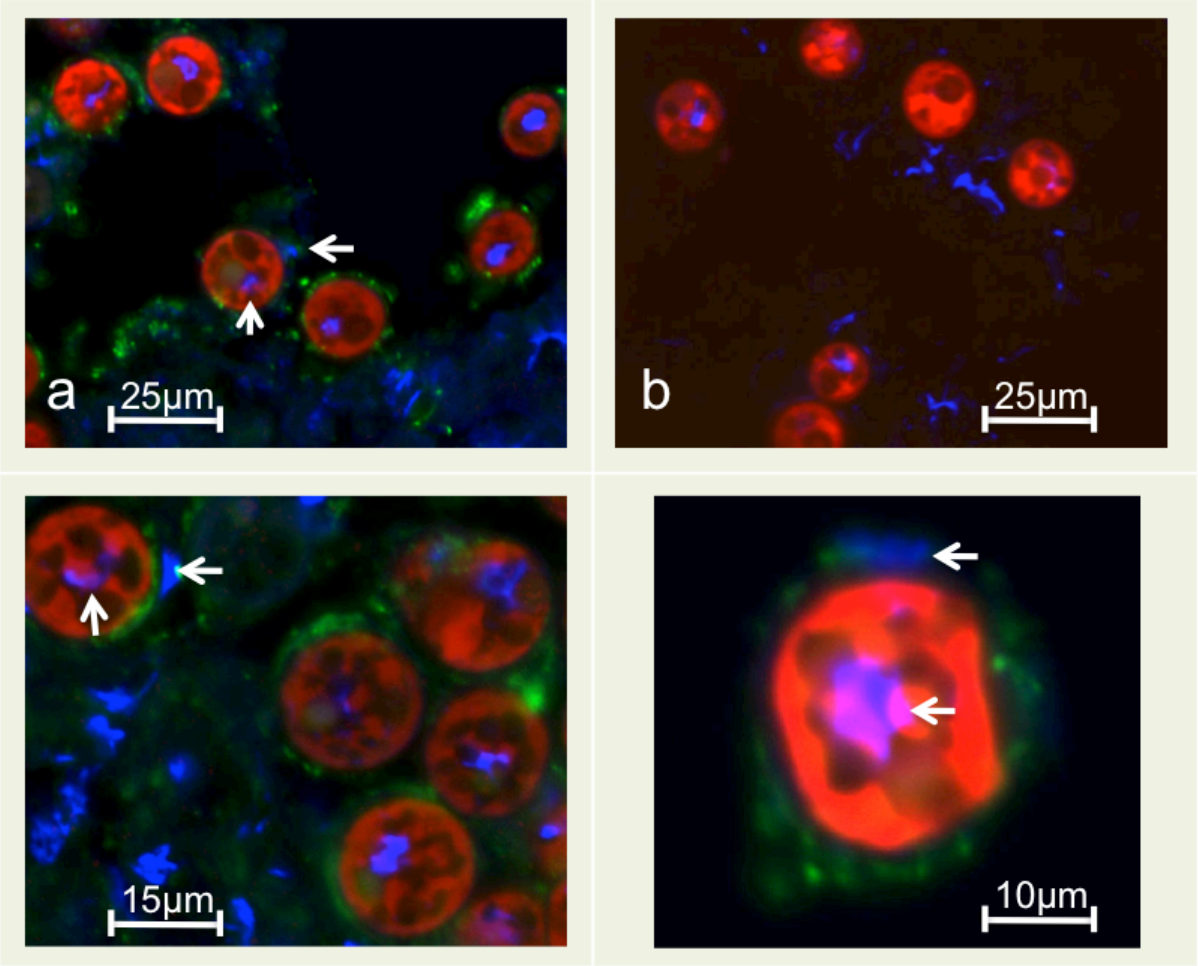



Figure 4.2: Confocal microscope images of cryosections from whole symbiotic anemones (*Aiptasia* sp.). (a) A section of anemone gastrodermis stained with anti-TSP conjugated with a florescent probe (green) and the dinoflagellate chloroplast autoflorescence (red). DAPI (blue) stains the nuclei of both anemone and *Symbiodinium* cells; arrows indicate where the anemone cell nuclei are visible near the dinoflagellate cell nuclei, indicating symbiont cells tight inside host gastrodermal cells. (b) The secondary antibody-only control, with no green fluorescent staining visible. (c) and (d) Thrombospondin antibody binding within the host gastrodermal cells surrounding the symbiont.

### *4.3.2 Blocking TSR domains inhibits symbiont uptake*

Incubation of aposymbiotic anemones with anti-TSP inhibited uptake of dinoflagellate cells (Figure 4.7; see inset confocal images showing representative tentacle slices at 72 h post-infection.). In comparison, anemones incubated in the FSW control showed relatively slow re-infection rates for the first 72 h and increased infection rates thereafter, and infection levels in the IgG antibody control followed a similar pattern. At 48 h post-infection, anemones in the anti-TSP, FSW and IgG treatments had a mean

percent infection of 0.6% (SE: 0.15), 2.4% (SE: 0.55) and 2.9% (SE: 1.04), respectively, while at 120 h, the percent infection was 1.26% (SE: 0.86%), 18.1% (SE: 2.65%) and 17.8 % (SE: 2.56%) respectively.

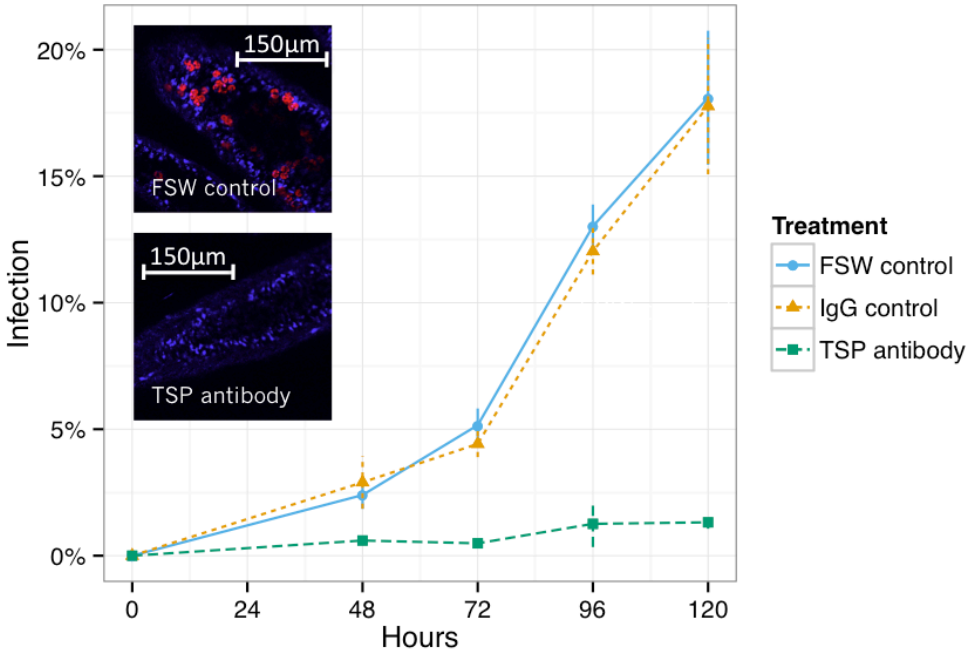



Figure 4.7: Infection rates in anemones treated with a rabbit polyclonal antibody with an epitope corresponding to amino acids 401-700 of thrombospondin 2 of human origin. Thrombospondin antibody (green dashed line) had a substantial effect on infection success compared to the filtered seawater (solid blue line) and IgG (orange dashed line) controlanemones, which showed expected infection rates. Infection success determined using autofluorescent area with the confocal microscope; each time-point represents four individual anemones per treatment, and three tentacles per anemone we sampled at random. Inset confocal images show representative tentacle slices at 72 h post-infection.

### *4.3.3 Blocking effect of TSP antibodies can be partially reversed by the addition of TGFβ protein*

The inhibitory effect on infection success of anti-TSP incubation was partially reversed by addition of exogenous TGFβ protein. Adding exogenous TGFβ protein to anti-TSP treated anemones restored infection to levels close to those found in FSW controls (Figure 4.8). At 65 h post-infection, anemones treated with both TSP antibody and TGFβ protein showed mean infection levels of 2.4% (SE: 0.3%), *versus* 0.02% (SE: 0.02%) in the antibody treatment, 3.14% (SE: 1.2%) in the FSW control, 3.1% (SE: 0.9%) in the

IgG control, and 3.3% (SE: 1.1%) in the IgG / TGFβ control. This suggests that tolerogenic TGFβ activity is linked to TSP activity in the recognition process. Experimental time points were limited by available TGFβ protein. Overall infection levels were low in these data compared with those measured in naturally symbiotic anemones (mean: 41.96%, SE: 2.31), suggesting that uptake is slow and proliferation of symbionts takes time to reach an optimal symbiont load within a host anemone.

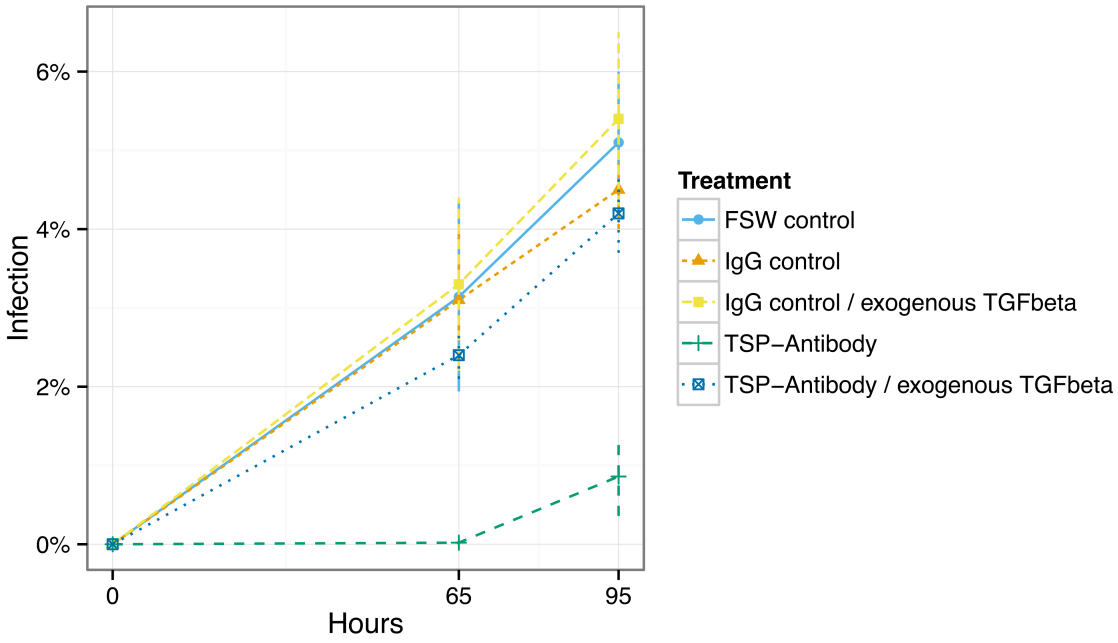



Figure 4.8: The blocking effect of thrombospondin antibody can be reversed by the addition of TGFβ protein. The addition of TSP1 antibody (green dashed line) had a substantial blocking effect on infection success, while adding exogenous TGFβ protein combined with TSP1 antibody (blue dotted line) restored infection success to those rates seen in the seawater control (solid blue line), and IgG (orange dashed line) and IgG combined with TGFβ protein treatments.

### *4.3.4 An excess of human Thrombospondin-1 induces 'super' infection*

Addition of exogenous human thrombospondin-1 (TSP1) protein accelerated the rate of infection by symbionts. Infection experiments using anemones pre-treated with soluble TSP1 protein showed markedly increased infection success when compared with FSW controls (Figure 4.9). Infection success after 48 h was 8.05% (SE: 0.98) in the TSP1 treatment *versus* 1.18% (SE: 0.28) in the FSW treatment, and after 96 h infection success was 25.1% (SE: 2.6) in the TSP1 treatment *versus* 9.87% (SE: 2.4) in the FSW control. This suggests that exogenous TSP is promoting recognition events between

host and symbiont. Confocal images (Figure 4.9 inset) show representative tentacle sections through the gastrodermis at 96 h post-infection and confirm a difference in infection levels between treated and untreated aposymbiotic anemones. Ultimately (120 h post-infection), infection levels in the TSP1 treatment and FSW control converged, at 31.35% (SE: 1.6) and 26.22% (SE: 2.49), respectively, highlighting that the effect of human TSP1 was most apparent at the onset of symbiont acquisition.

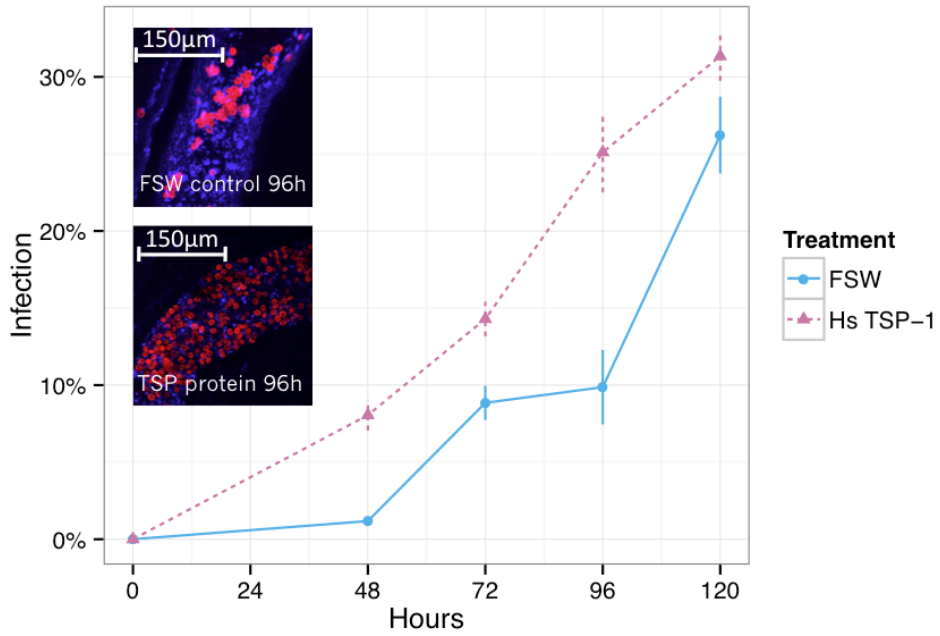



Figure 4.9: The addition of exogenous (human platelet) thrombospondin (TSP -1) (red dashed line) substantially increased the infection rate at the onset of infection, when compared with control anemones in filtered seawater (FSW) (blue line). Infection success determined using autofluorescent area with the confocal microscope; each time-point represents four individual anemones *per* treatment, and three tentacles *per* anemone were sampled at random. Inset confocal images show representative tentacle slices at 96 h post infection.

### *4.3.5 Addition of TSR peptide fragments also increases infection success at the onset of infection*

As with human TSP1, pre-treating anemones with short synthetic TSR peptides produced increased infection success (Figure 4.10), with higher infection levels in treated anemones across all time-points. In particular, at the onset of symbiosis, during the first 48 h post-infection, symbiont acquisition was faster in peptide-treated anemones (Peptide 1: 11.14%, SE: 1.1% and Peptide 2: 11.78%, SE: 0.9%) compared to the FSW control anemones (2.08% infection, SE: 0.29%). After 48 h, infection levels in

the Peptide 2 treatment were consistently higher than in the Peptide 1 treatment. This difference was particularly apparent at 72 h, where infection levels in anemones in the Peptide 2 treatment were 5% higher than in the Peptide 1 treatment (Peptide 2: 20.24% infection [SE: 1.4%] *versus* Peptide 1: 15.11% infection [SE: 1.98%]). The peptide treatments showed the largest increase relative to the FSW control at 96 h, with 18.8% (SE: 1.3%) and 20.9% (SE: 1.68) infection for Peptides 1 and 2, respectively, compared to only 6.15% (SE: 0.75%) for the FSW control. However, as in the TSP1 treatment, the extent of this difference declined by the end of the experiment (120 h), with infection in the FSW control increasing nearly 3-fold from the value at 96 h, to 17.01% (SE: 1.22%), and infection in the peptide treatments increasing only slightly over this same period (Peptide 1: 22.79% [SE: 2.66] and Peptide 2: 24.82% [SE: 1.87]).

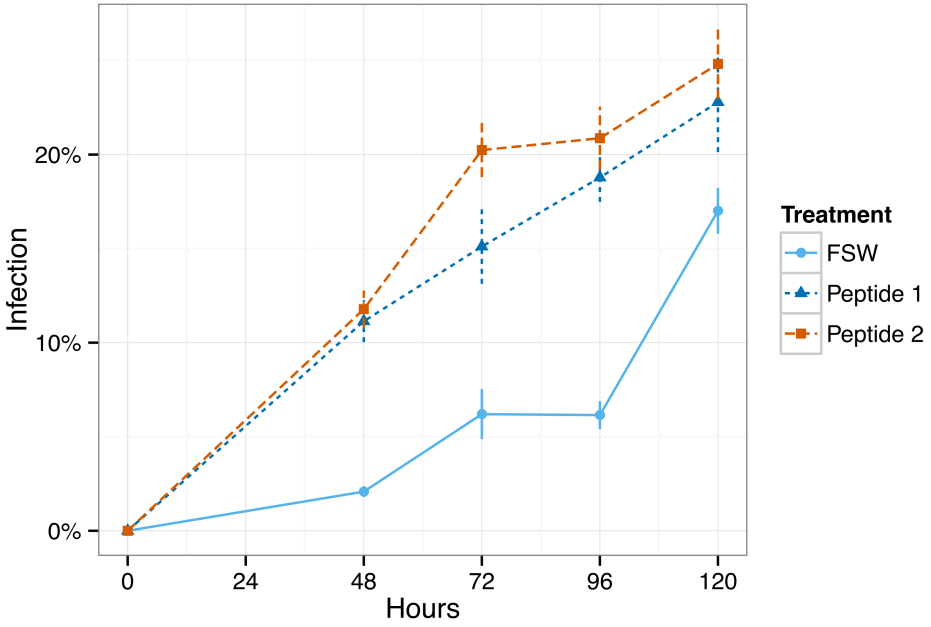



Figure 4.10: The effect of synthetic TSR peptides 1 (blue dashed line) and 2 (orange dashed line) on infection rates compared to the control anemones in filtered seawater (blue solid line). Anemones treated with both peptides 1 and 2 showed increased uptake of algae at the onset of infection.

### 4.4 Discussion

This functional experimental work aimed to investigate the possible involvement of TSR-domain-containing proteins in symbiont uptake by the model anemone *Aiptasia* sp. All TSR-domain-containing proteins have one or more TSR domains, similar to the three repeats originally characterized in the human TSP1 protein (Tucker 2004). The TSR

protein domain is one of many potential ligands involved in the establishment of a stable cnidarian-dinoflagellate symbiotic relationship. Using an anti-TSP antibody (directed against three TSR domains), I found that TSR-containing proteins are present in higher quantities in symbiotic *versus* aposymbiotic anemones, and confocal microscopy suggests that these proteins are localised within anemone gastrodermal cells containing dinoflagellates. Functional experiments further showed an inhibitory effect of anti-TSP, and a stimulating effect of human TSP and synthetic TSR-domain-containing peptides on symbiont uptake. Data presented here therefore suggest that TSR-domain-containing proteins play a key role in host-symbiont recognition and uptake.

### *4.3.6 TSR-domain-containing proteins involved in cnidarian-dinoflagellate symbiosis are likely of host origin*

One hypothesis for how TSR-domain-containing proteins are involved in initiating symbiont uptake is the use of a TRAP-like protein on the symbiont surface (Figure 4.11, Hypothesis 1). The TRAP protein is used by several apicomplexan parasites to gain entry to mammalian host cells. For example, the malaria parasite, *Plasmodium falciparum*, uses a TRAP to bind to the hepatocyte cell surface (Müller et al. 1993). A second hypothesis suggests that the symbiont utilises host TSR proteins, by initiating TSR protein expression within host gastrodermal cells (Figure 4.11, hypothesis 2). The symbiont may actively sort membrane proteins at the point of phagocytosis, excluding unwanted proteins and seeking out host TSR-domain-containing proteins to be part of the host-derived symbiosome membrane. This strategy is used by the apicomplexan parasite *Toxoplasma gondii.*, which initiates selective invagination of the membrane during phagocytosis by host cells, excluding specific host trans-membrane proteins by means of a moving junction (Mordue et al. 1999).

Immuno-fluorescent labelling of anti-TSP antibodies localised the TSR-domain containing proteins in the symbiosis. Whereas anti-TSP failed to label the surface of cultured *Symbiodinium* cells*.* (phylotype B1; culture ID: CCMP830) (Figure 4.5), it did label anemone tissue associated with *Symbiodinium* spp. freshly isolated from *Aiptasia* sp. (Figure 4.5). Whole mount sections of symbiotic anemone tissue confirmed that anti-TSP binding was localised within anemone gastrodermal tissue surrounding the

symbionts (Figure 4.6). Furthermore, anti-TSP labelled the same two bands in both symbiotic and aposymbiotic anemones in Western blot analysis, indicating that no new TSR proteins are present in the symbiotic state. The CD36 domain binding sites identified in human TSP1 are present in both cnidarian and dinoflagellate TSR domains. 3-D folding sites and glycosaminoglycan binding motifs are not as well conserved in sequences obtained from the dinoflagellate, *Symbiodinium minutum* draft genome (figure 4.2).

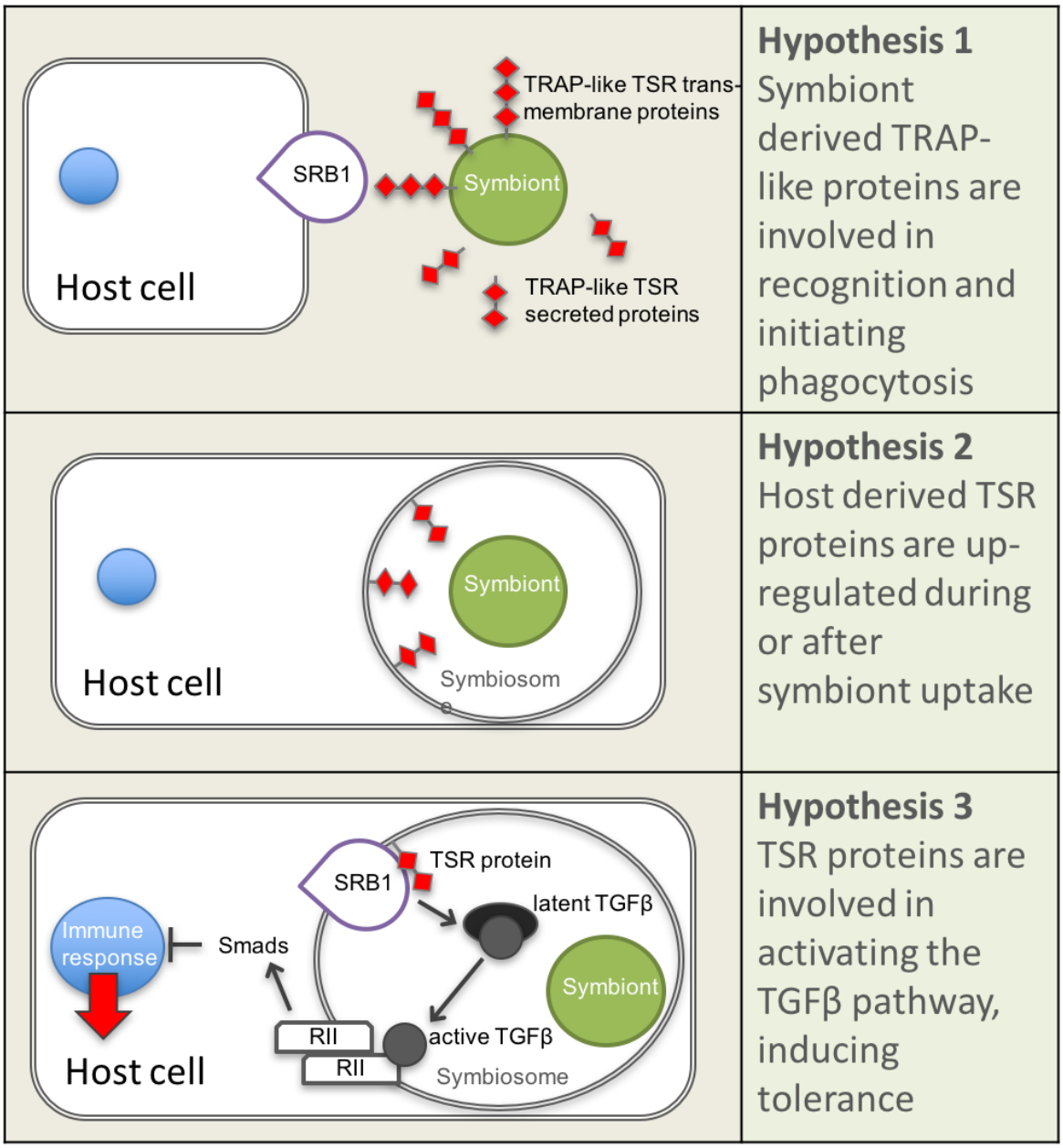



Figure 4.11: Three possible hypotheses for how TSR proteins may be involved in the onset and maintenance of cnidarian/dinoflagellate symbiosis.

Together, my experiments provide support for Hypothesis 2 (i.e., a host origin of TSR-domain-containing proteins involved in symbiosis), and no evidence in support of Hypothesis 1 (the use of symbiont-derived TRAP-like proteins). However, gene searches revealed multiple TSR domain-containing proteins within the *Symbiodinium minutum* genome (Chapter 3), and dinoflagellate TSR-domain-containing proteins could be released only in response to signals provided by the anemone host during the onset of symbiosis. Similar mechanisms have been found in *P. falciparum*, for which TRAP expression increases during sporozoite ontogeny, suggesting that the TRAP proteins are only expressed in response to specific cues (Robson et al. 1995). Indeed, the cnidarian host is known to stimulate the release of various metabolites from the dinoflagellate symbionts (e.g. (Gates et al. 1995, Davy 2001, Markell and Wood-Charlson 2010)). This study does not rule out the potential involvement of TRAP-like proteins in the onset of the cnidarian-dinoflagellate symbiosis. However my experiments suggest that TSR-domain containing proteins are upregulated within the cnidarian host in response to cues from the symbiont rather than the host stimulating the secretion of TRAP-like proteins by the symbiont.

### *4.3.7 Re-infection experiments implicate the TSR domain in symbiont uptake*

Functional experiments in this study involved the re-introduction of dinoflagellate cells to aposymbiotic anemones pre-treated to either block, stimulate or mimic TSR-domain-containing proteins. These experiments focussed on the involvement of a TSR domain in the early onset of symbiosis and symbiont uptake by the host. Initial interest in the TSR domain was prompted by the search for a binding target for the scavenger receptor SRB1. A cnidarian SRB1 gene was upregulated in the symbiotic state of the anemones *Anthopleura elegantissima* (Rodriguez-Lanetty et al. 2006) and *Aiptasia* sp. (Lehnert et al. 2014). The TSR domain also binds the Class B scavenger receptor CD36/SRB1 (Asch et al. 1987, Li et al. 1993, Frieda et al. 1995, Dawson et al. 1997). The binding of a TSR protein to SRB1 can initiate the TGFβ immune pathway within the host by activating latent TGFβ protein (Khalil 1999, Murphy-Ullrich and Poczatek 2000, Koli et al. 2001). This pathway provides a hypothesis for symbiosis maintenance, persistence and proliferation within the anemone host (Figure 4.11, Hypothesis 3).

Many intracellular parasites manipulate the host innate immune defence mechanisms to their own advantage (Medzhitov and Janeway 2002). The most commonly described function of scavenger receptors (SRs) is to act as phagocytic receptors (Areschoug and Gordon 2009). A common mechanism of entry by microbes to host cells is to alter and gain control over the host phagosome which engulfs them with the intent to destroy the pathogen (Mcfall-Ngai 2007, Schwarz 2008). The malaria parasite *P. falciparum* uses the motif 'WSPCSVTCG' to bind to sulphated glycoconjugates and SRB1 on hepatocytes, as the first step to gaining entry to host cells (Müller et al. 1993, Areschoug et al. 2008). SRB1 is thus a key element in *Plasmodium* infection, as it promotes sporozoite invasion of hepatocyte cells and subsequent intracellular parasite development (Rodrigues et al. 2008). TSR-domain-containing proteins provide a potential ligand for the upregulated scavenger receptor protein SRB1 (Rodriguez-Lanetty et al. 2006, Lehnert et al. 2014), thereby initiating phagocytosis of dinoflagellate symbionts.

In this study, an antibody raised against the three TSR domains in human TSP1 protein inhibited dinoflagellate entry to host gastrodermal cells. This inhibitory effect on symbiont uptake suggests that anti-TSP is binding to TSR-domain-containing proteins that play a role in host-symbiont recognition. It is not clear exactly which TSR-domain-containing proteins are actively involved in symbiont acquisition and are therefore blocked by the anti-TSP in these experiments. Previous work revealed a large cnidarian repertoire of TSR domain-containing proteins such as ADAMTS metalloprotease-like proteins (Chapter 3). The cnidarian TSR domains within these cnidarian TSR proteins is highly conserved and functional motifs are intact, including the tryptophan GAG-binding motif 'WXXW' and scavenger receptor binding motifs 'CSVTCG' and 'GVITRIR' (Adams and Tucker 2000, Silverstein 2002). In humans, the metalloprotease ADAMTS 13 binds to CD36 (Davis et al. 2009) and in *Caenorhabiditis elegans*, an ADAMTS protein (AD-2) is responsible for initiating the TGFβ pathway, regulating body growth and maintaining cuticle formation (Fernando et al. 2011). It is therefore conceivable that an ADAMTS-like TSR protein is involved in TGFβ initiation in the cnidarian-dinoflagellate symbiosis. In Chapter 3, a number of potential TSR proteins were presented that could form the basis of further investigation, using specific antibodies designed for known targets.

Binding motifs within the TSR domain are so highly conserved across taxa that the synthetic peptides designed from TSR domains have been employed by a number of studies, including determining which motifs have anti-angiogenic activity in human

TSP1 (Tolsma et al. 1993, Anderson et al. 2007, Cano et al. 2009, Garside et al. 2010), which motifs bind to CD36 (Li et al. 1993) and which *Plasmodium* sp. TSR peptides bind to red blood cells (Calderón et al. 2008). In the present study, a TSR peptide was designed from the TSR domain in an *Aiptasia* sp. TSR-domain-containing protein. This synthetic peptide contained both the tryptophan GAG-binding motif 'WXXW' and scavenger receptor binding motifs 'CSVTCG' and 'GVXTRXR'. Both soluble TSP1 protein and the *Aiptasia* sp. synthetic TSR peptides increased the rate of symbiont uptake in treated *versus* non-treated aposymbiotic anemones. This result suggests that one or multiples of these binding motifs are involved in successful entry to host cells by the dinoflagellate.

## 4.4 Concluding remarks

This study provides the first functional evidence for involvement of TSR-domain-containing proteins in both the onset and maintenance of cnidarian-dinoflagellate symbiosis. Western blots showed that anti-TSP labelled bands had stronger intensities in symbiotic compared to aposymbiotic anemone protein homogenates. The larger quantities of TSR-domain-containing protein in symbiotic anemones suggests that a TSR protein is involved in the long-term maintenance of the symbiosis. Pre-treatment of aposymbiotic anemones with anti-TSP inhibited symbiont uptake compared with untreated controls. This blocking effect was partially reversed by the addition of exogenous TGFβ protein; symbiont uptake levels were restored, supporting the hypothesis that TSR-domain-containing proteins may be involved in triggering a tolerogenic response in the host that occurs post-phagocytosis via the TGFβ pathway. In this case, host-derived TSR proteins would initiate the TGFβ immune pathway, resulting in persistence of symbionts within the host gastrodermal cells (Figure 4.11, Hypothesis 3). In mammalian systems, the TGFβ pathway is responsible for maintaining immune homeostasis; the dominant role of this pathway is to induce tolerance, and to contain and resolve inflammation(Li et al. 2006).

Proteins of the TGFβ pathway have been identified in cnidarians (Samuel et al. 2001, Technau et al. 2005), and are expressed in the symbiotic state of the anemone *Aiptasia* sp. (Detournay et al. 2012). Data presented here support both Hypotheses 2 and 3, which may well be sequential (Figure 4.11); host derived TSR proteins are upregulated

and important to both the process of symbiont acquisition via SRB1, as well as maintenance of a healthy symbiotic relationship between host and symbiont.

# Chapter 5

# General Discussion

## 5.1 Summary of research findings

The research detailed in this thesis can be divided into two main questions areas: (1) what are the potential receptor-ligand signaling mechanisms utilized during initial contact and symbiont invasion of the cnidarian host and (2) how may these ECM proteins be involved in immune modulation of the host and persistence inside the host phagosome. The motivation for investigating scavenger receptor (SR) proteins came from the results of a transcriptome study in the anemone *Anthopleura elegantissima*, where the SR SRB1 was upregulated in the symbiotic relative to the aposymbiotic state (Rodriguez-Lanetty et al. 2006). A review of the apicomplexan parasite literature, where TSR-domain-containing proteins are used to gain entry to host cells, motivated characterization of the TSR-domain-containing protein repertoire in cnidarians (Morahan et al. 2009). The TSR domain is also implicated in the tolerogenic TGFβ pathway. Most vertebrates macrophage pathogens have evolved mechanisms to induce TGFβ production, which, in turn, suppresses apoptosis, enhances intracellular proliferation of the pathogen, and thus favours parasite virulence (Li et al. 2006). The TSR domain provides a potential ligand for the upregulated SRB1 protein that is implicated in symbiosis maintenance.

The work presented here provides new information about several cnidarian extracellular matrix (ECM) proteins that are potentially important during symbiosis onset. In chapter 2, characterisation of the scavenger receptor protein repertoire of 6 cnidarian species suggested that cnidarians have an expanded SRCR-domain-containing protein repertoire compared to vertebrates. This expansion has been observed in other invertebrate models such as the purple sea urchin, *Strongylocentrotus purpuratus* and the amphioxus, *Branchiostoma floridae*. Although the reason for this expansion is unknown, in invertebrates, innate immune receptors maintain control over both pathogens and beneficial microbial communities. The latter may provide clues about the expansion of the SRCR-protein repertoire, as regulation of the holobiont requires a

greater complexity of signaling mechanisms (Mcfall-Ngai 2007, Messier-Solek et al. 2010). Simple functional experiments blocked positively charged binding sites on both SRCR and CD36 protein domains, resulting in decreased symbiont invasion success in aposymbiotic *Aiptasia* sp. This blocking experiment also produced an increase in the inflammatory response to immune stimulation by lipopolysaccharide in symbiotic *Aiptasia* sp. These data support the hypothesis that SRs are involved in both symbiont uptake and persistence within host phagosomes.

In chapter 3, a large and diverse TSR-domain-containing protein repertoire of 6 cnidarian species was characterized and compared to vertebrate TSR proteins of known function. Of particular interest is the large number of ADAMTS metalloprotease-like proteins. This group is large in both humans and cnidarians and my findings suggest this is an ancestral TSR protein group. The *Symbiodinium minutum* genome in comparison contains proteins with TSR domains in repeats without any other identifiable domains. The TSR domain is of interest in terms of symbiosis maintenance, as in humans it is instrumental in initiating the TGFβ pathway, which among many homeostatic functions is responsible for inducing tolerance, as well as containing and resolving inflammation (Li et al. 2006). Phylogenetic analysis of TSR domains revealed strong conservation of binding motifs and 3-D folding sites, suggesting this domain is ancient and highly conserved from lower metazoans to vertebrates. Conservation of 3-D folding and binding motifs would suggest that the TSR domains and proteins are interchangeable between vertebrate and cnidarian systems.

In chapter 4 the role of TSR proteins at the onset of symbiosis was examined in the model anemone *Aiptasia* sp. system. Various functional experiments measured symbiont uptake in aposymbiotic anemones, after challenge with proteins and antibodies to either stimulate or block TSR-domain-binding. Adding soluble human TSP1 protein increased the rate of symbiont uptake; dinoflagellate cells were taken up faster and in larger numbers compared to the control. It is not clear exactly why the human protein has such an effect on the early stages of symbiont acquisition. One hypothesis is that the TSR domain is a limiting factor in symbiont acquisition. Parasites stimulate the overexpression of TGFβ protein to modulate the host immune response, ADAMTS and TRAP are involved in activation of TGFβ (Omer et al. 2003, Li et al. 2006). In the absence of a clear TRAP protein homologue found in the symbiont, a second hypothesis suggests that the symbiont stimulates production of one or more host TSR-

domain-containing proteins, such as ADAMTS. The addition of synthetic peptides designed from *Aiptasia* TSR-domain-containing proteins also increased symbiont uptake, while the addition of an antibody with an epitope corresponding to three TSR repeats shut down symbiont uptake. A final functional experiment added exogenous TGFβ to antibody challenged anemones and the blocking effect was reversed suggesting that the TGFβ pathway is involved in early onset of the symbiosis. Fluorescence microscopy of symbiotic *Aiptasia* sp. found the antibody localized within gastrodermal cells containing symbionts, showing that TSR protein expression may be localized and contained within the symbiosome membrane. Previous work identified proteins of the TGFβ pathway within cnidarians (Detournay et al. 2012). These results provide strong for support the involvement of TGFβ in promoting tolerance of symbionts within the host. These data further support the hypothesis that a TSR domain containing protein is important to symbiont acquisition and the maintenance of a healthy symbiosis.

## 5.2 How are scavenger receptors implicated at initial contact, host-symbiont signaling and host phagocytosis?

The mechanisms of symbiosis establishment have been broken down into a series of steps to allow each process to be studied in detail. These steps are: (1) recognition and phagocytosis, (2) selection (specificity), (3) arrest of phago-lysosome maturation and cell cycle control (Nyholm and Mcfall-Ngai 2004, Davy et al. 2012). Recognition involves molecular signaling between host and symbiont. This initial contact occurs between receptors within the ECM of host gastrodermal cells and the algal cell surface proteins. The cocoid phase of the *Symbiodinium* cell is surrounded by a cellulosic cell wall that secretes large molecular weight glycoproteins (Markell and Trench 1993). These secreted glycoconjugates provide candidates for signaling and binding to host ECM receptors (Markell and Wood-Charlson 2010). The lectin-glycan interactions are among some of the most studied inter-partner recognition mechanisms. In the ECM, lectins act as pattern recognition receptors (PRRs), binding microbial glycans and initiating a innate immune response, subsequent phagocytosis and destruction of the invading microbe (Vasta 2009). Several studies have investigated the role of lectin-glycan signaling in the onset of cnidarian-dinoflagelate symbiosis (Lin et al. 2000, Wood-Charlson et al. 2006, Kvennefors et al. 2008, Wood-Charlson and Weis 2009, Logan et al.

2010, Markell and Wood-Charlson 2010), the green *Hydra-Chlorella* symbiosis (Meints and Pardy 1980), and in the squid-vibrio symbiosis (Nyholm et al. 2003, Nyholm and Mcfall-Ngai 2004). Experimental evidence in several symbiotic systems implicates lectin-glycan interactions in recognition and phagocytosis of symbionts by host cells.

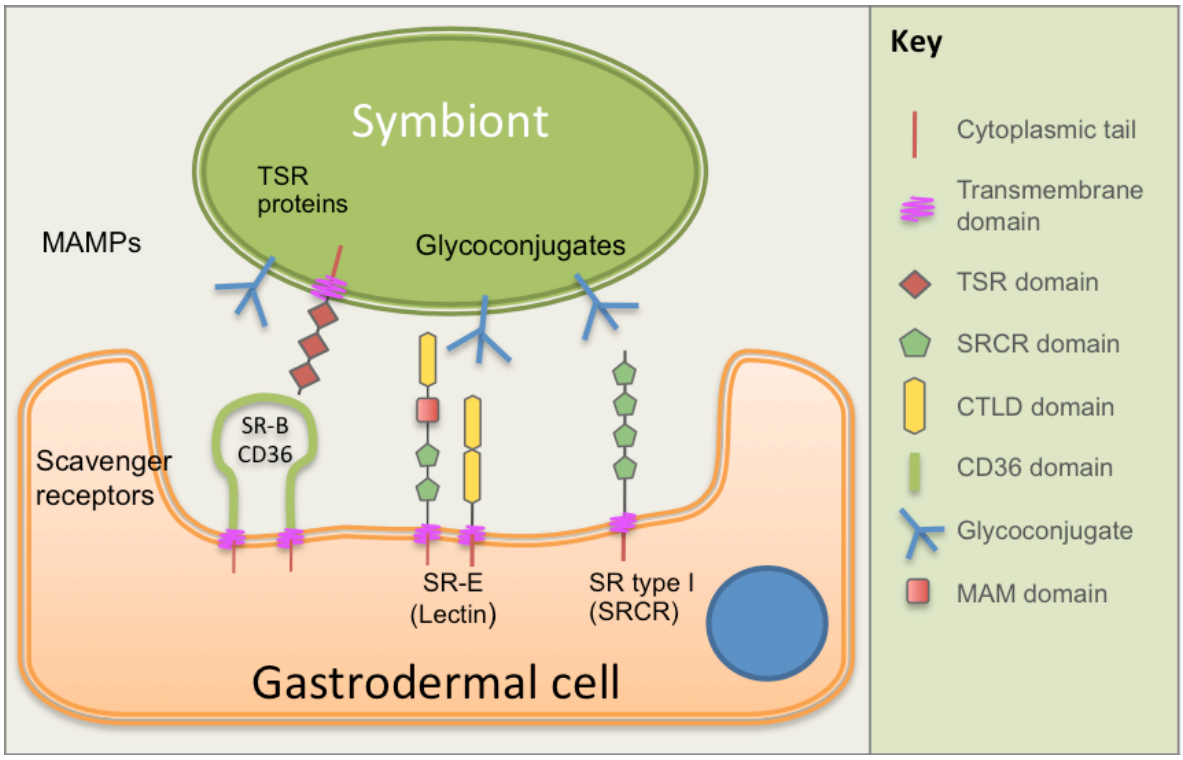



Figure 5.1: Recognition and phagocytosis – testable cellular model. Suggested symbiont cell surface ligands for the three groups of SRs identified. TSR-domain-containing proteins or a possible *Symbiodinium* TRAP homolologue binds to the CD36 domain of class B scavenger receptors. Gycoconjugates, such as glycoproteins, glycopeptides, peptidoglycans and glycolipids are common microbe associated molecular patterns (MAMPs) recognized by CTLD and SRCR domain containing SRs. All three SR groups have the ability to bind symbionts and initiate phagocytosis.

In the present study, several C-type lectin domain (CTLD) containing SR proteins were identified and described in the 6 cnidarian species searched. Figure 5.1 provides a testable cellular model for how three SR types may be involved in symbiont recognition and subsequent phagocytosis. Both CTLD and SRCR-domain containing SR proteins described in chapter 2 are potential receptors for secreted, symbiont derived glycoconjugates. In vertebrates, the LOX1, a CTLD containing SR protein, is involved in intracellular trafficking and signaling, it activates apoptosis and mediates the phagocytosis of apoptotic cells (Murphy et al. 2005). Cnidarians possess a LOX1

homologue and CTLD containing SRs with multiple other domains including SRCR domains. Much less is known about the ligands of the SRCR domain, however, they are implicated in aggregation and cell to cell signaling (Aruffo et al. 1997, Blumbach et al. 1998), phagocytosis and initiation of an innate immune response (Hohenester et al. 1999, Sarrias et al. 2004). I propose here that SRCR domains act in a similar capacity to the CTLD, and bind glycoconjugates in the exuded mucous layer associated with the algal cell surface and initiate phagocytosis of symbionts. The CD36 domain in class B SRs has the capacity to bind a wide range of ligands from lipoproteins to bacteria (Silverstein and Febbraio 2009). The CD36 domain has the potential to bind algal cell surface glycoconjugates in the same way it has been shown to bind the E2 glycoprotein of the hepatitis C virus (Areschoug et al. 2008, Catanese et al. 2010).

It is unclear from the data presented in this study if dinoflagellates posses a homologue to the TSR-domain-containing TRAP proteins used by apicomplexan parasites, *Plasmodium* spp, *Toxoplasmosa* spp, and *Cryptosporidium* spp. to invade host hepatocytes. Several *Symbiodinium minutum* TSR-domain-containing proteins were described in this study, however there is no evidence to suggest these are TRAP-like proteins or that they are expressed on the cell surface as they are on malaria sporozoites. The TSR domain amino acid motif 'WSPCSVTCG' of *Plasmodium falciparum* is required for binding sulphated glycoconjugates and to hepatocytes (Müller et al. 1993). The TSR motifs of human TSP1 protein bind to the CD36 domain of class B SRs using the amino acid motifs 'CSVTCG' and 'GVITRIR' (Zhang and Lawler 2007). The 'CSVTCG' motif is shared and highly conserved in axpicomplexan, dinoflagellate and cnidarian sequences analyzed in multiple sequence alignments in chapter 3. Based on this binding motif overlap, I suggest a symbiont derived TSR-domain-containing protein could be involved in binding the CD36 domain of cnidarian class B SRs. However data presented in this thesis do not support a dinoflagellate TRAP protein homologue. In chapter 4, functional experiments suggest that a TSR-domain-containing protein is necessary to symbiont uptake, however it is unclear whether these proteins are of host origin, symbiont origin or indeed TSR proteins from both are involved in enabling symbiont uptake. *Plasmodium* utilises both host ADAMTS proteins and the parasite TRAP protein to activate the tolerogenic TGFβ pathway (Omer et al. 2003). The present study supports the involvement of host derived TSR-domain-containing protein such as one of the many cnidarian ADAMTS proteins described in chapter 3.

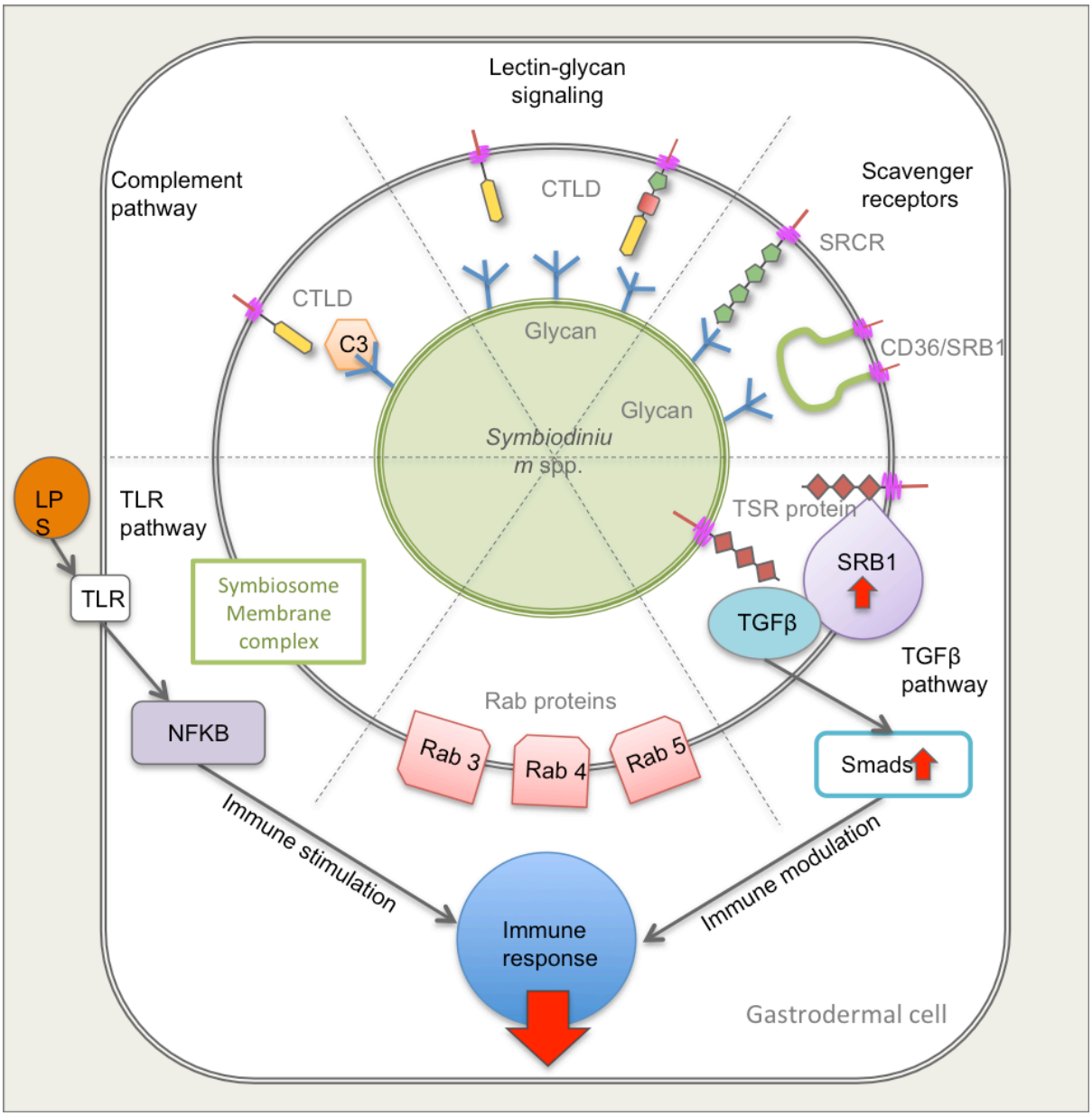



Figure 5.2: Post phagocytosis – in the host gastrodermal cell, the symbiont encapsulated is within the symbiosome membrane complex. The top three pie slices show ECM proteins involved in symbiont recognition; the lectin complement pathway, lectin/glycan signalling and scavenger receptors. The bottom three pie slices show modulation of host immune pathways, the TLR pathway, halting of phagosome maturation by selection of early endosome Rab proteins and tolerance promoted via the TGFβ pathway. Upward pointing red arrows show proteins that are upregulated in the symbiotic state, the large down arrow in the host nucleus indicates a down regulation of the immune response induced via LPS in symbiotic anemones.

## 5.3 What happens post phagocytosis? How does the symbiont modulate the host immune response and evade digestion?

Post phagocytosis, the symbiont is contained in the host membrane, and the host phagosome matures into the symbiosome membrane complex (Muscatine et al. 1975). The phagosome is composed of host plasma membrane, internalized at the point of host- symbiont contact. However, it is not clear if ECM receptors involved in recognition and initial binding are also internalized and remain part of the symbiosome membrane complex. Figure 5.2 shows the dinoflagellate symbiont within the host gastrodermal cell, combining the data presented here on SRs and TSR-domain-containing proteins in the context of current knowledge. This model proposes that the ECM receptors involved in recognition and uptake remain attached to the symbiont post-phagocytosis. The involvement of CTLD-containing receptors is well established within the recent literature, and the lectin complement pathway protein C3 has been characterized in the coral *A. millepora*. C3 could be opsonizing the symbiont, enabling interpartner communication and recognition (Kvennefors et al. 2010). Research into the functional mechanisms of how innate immune PRRs may be implicated in dinoflagellate recognition and uptake is in its infancy (Davy et al. 2012). Details of specific SR proteins described here will enable further functional experiments to investigate the role SRs in interpartner signaling and recognition. The CTLD-glycan signaling mechanism investigated using confocal imaging of infection, suggests selective uptake of homologous algae glycan signatures by *Fungia scutaria* larvae (Wood-Charlson et al. 2006). I suggest that all three SR domains; CTLD, SRCR and CD36 have the potential to be involved in symbiont recognition and uptake, involvement in post phagocytosis mechanisms is unclear. These mechanisms of inter-partner signaling may be essential to symbiosis maintenance as well as successful invasion.

The class B SR, SRB1, is upregulated in the symbiotic state of both the anemones *A. elegantissima* and *Aiptasia* sp, in functional genomics studies examining genes differentially expressed in aposymbiotic versus symbiotic states (Rodriguez-Lanetty et al. 2006, Lehnert et al. 2014). The up regulation of this protein post-phagocytosis indicates a potential role in symbiosis maintenance. Data presented in chapters 3 and 4 support a role for a cnidarian SRB1 protein in TSR-domain-containing protein binding and subsequent activation of the TGFβ pathway. Figure 5.2 shows cnidarian TGFβ

pathway proteins identified in cnidarians: a TGFβ ligand was characterized in *Aiptasia* sp., as have the Smad proteins that function downstream post TGFβ activation (Technau et al. 2005, Detournay et al. 2012). Phosphorylated Smad proteins 2/3 were present in higher quantities in symbiotic versus aposymbiotic *Aiptasia* sp., detected using immunoblots (Detournay et al. 2012). Both CD36 and TSR-domain-containing proteins identified in the current research; provide new potential TGFβ pathway components. In vertebrates, the activation of latent TGFβ into its active form requires the interaction of TSP1 protein with CD36 (Khalil 1999), and within TSP1, the TSR domain binds the CD36 domain on the CD36 protein. In chapter 3, multiple sequence alignments of TSR domains show conservation of the known CD36 domain binding sites within cnidarian TSR domains, suggesting that cnidarian TSR-domain-containing proteins can bind to CD36-domain-containing SRs.

The TGFβ pathway controls many interactions between host and pathogen. The majority of vertebrate macrophage pathogens have evolved mechanisms that promote TGFβ production to induce tolerance and enhance intracellular proliferation within the host (Li et al. 2006). The TGFβ pathway has the potential to be both beneficial and detrimental to the invading parasite. The parasite *Trypanosoma cruzii* requires the presence of TSP1 protein in the ECM of host cells and the activation of the TGFβ signaling pathway for entry into mammalian cells (Ming et al. 1995, Waghabi et al. 2005, Simmons et al. 2006, Nde et al. 2012). Both *Leishmania major* and *Toxoplasma gondii* have developed mechanisms to induce macrophages to produce high levels of active TGFβ, the TGFβ produced by infected macrophages suppresses NO production and reduces an inflammatory response (Li et al. 2006). Interestingly, in red blood cells infected with *Plasmodium* spp., high levels of active TGFβ are associated with slowing parasite replication and protection of the host. Activation of latent TGFβ by malaria parasites is a two-step process involving ADAMTS metalloproteases and the malaria TSR-domain-containing TRAP protein (Omer et al. 2003). The mechanism of TGFβ activation in *Plasmodium* spp, is of particular interest to cnidarian-dinoflagellate symbiosis as it provides evidence of both ADAMTS metalloprotease and TRAP TSR-domain-containing proteins activating TGFβ in a similar mechanism as is described for mammalian systems. Several ADAMTS metalloprotease TSR-domain-containing proteins were identified in the six cnidarian species searched in chapter 3. These and

other cnidarian TSR-domain-containing proteins could be involved in TGFβ activation in cnidarian cells.

Several studies in vertebrates show that high glucose levels increase TSP1 production (Wang et al. 2004, Yung et al. 2006, Dabir et al. 2008) and this effect subsequently increases levels of active TGFβ (Tada et al. 2001, Shalev et al. 2002). Within the cnidarian-dinoflagellate symbiosis, the TSR domain provides a potential ligand for the upregulated SRB1 proteins post symbiont phagocytosis. Confocal microscopy evidence presented in chapter 4 suggests that TSR-domain-containing proteins are localized in gastrodermal cells containing symbionts. I propose that photosynthetic glucose released by the symbiont could increase production of TSP-domain-containing proteins locally, within the host gastrodermal cell. This mechanism could explain the localization of TSR-domain-containing proteins within cells containing symbionts.

In comparison to aposymbiotic *Aiptsia* sp., symbiotic anemones produce significantly less NO in response to immune stimulation by LPS (Detournay et al. 2012). This response is mirrored in vertebrate macrophages parasitized by *Leishmania major* and *Toxoplasma gondii*, where NO production is suppressed due to the tolerogenic effects of the TGFβ pathway (Li et al. 2006). In figure 5.2, the TLR pathway is activated by addition of LPS, cnidarian homologues of the TLR receptor and NFKB have been identified (Miller et al. 2007, Palmer and Traylor-Knowles 2012). Data presented in this thesis supports the hypothesis presented by Detournay et al. (2012), the cnidarian host immune response is used and modulated by the symbiont via activation of the tolerance promoting TGFβ pathway.

The process of symbiosome membrane formation is unclear; it has been suggested that a process of sorting membrane proteins may occur (Peng et al. 2010). The apicomplexan parasite *Toxoplasmosa gondii*, actively penetrates its host cell by the creation of a moving junction in the host cell plasma membrane (Mordue et al. 1999). This junction allows selective control of host cell plasma membrane proteins, enabling the parasite to design its own specialized vacuole referred to as the parasitophorous vacuole (PV). In cnidarian-dinoflagellate symbiosis, the symbiosome membrane is also modified. In the *Aiptasia-Symbiodinium* association, Rab proteins (Rab 7 and Rab 11) are excluded from the symbiosome membrane (Chen et al. 2003b). These proteins are involved in the recycling process essential for phagosome maturation and, by their active exclusion, the symbiont halts phagosome maturation and prevents phago-

lysosome fusion. Further study of Rab proteins and their association with the symbiosome in *Aiptasia* sp. shows that early endosomal Rab proteins (Rab 3, 4 and 5) are associated with healthy symbionts within symbiosome membranes (Chen et al. 2004, Hong et al. 2009b, 2009a).

## 5.4 Future directions

The work presented here has generated a myriad of hypotheses to be explored by future functional investigations into the roles of ECM proteins in establishment and maintenance of cnidarian-dinoflagellate symbiosis. Blocking experiments performed in chapter 2 used fucoidan, which binds multiple SR classes. Commercially available antibodies have epitopes that are designed for mammalian proteins and are often problematic for use within invertebrate systems. SR protein sequences identified and described in chapter 2 provide the basis for designing specific *Aiptasia* sp. antibodies to block specific SRs in functional re-infection experiments, to identify the SRs involved in symbiont uptake. Western blot analysis of SRs at different stages of symbiosis will also determine which SRs are present at different points. So far we are only aware that SRB1 is present at higher levels in the symbiotic state as opposed to the aposymbiotic state, and its presence at the onset of symbiosis should be investigated.

TSR-domain-containing proteins characterized in chapter 3 are of interest for future research as their functions in invertebrates are largely unexplored. Furthermore, the question of a *Symbiodinium* spp. TRAP-like protein remains unanswered. Due to the nature of the dinoflagellate cell wall, I am skeptical that the TRAP proteins identified in apicomplexan parasites are found in dinoflagellates. It is possible that a TRAP-like protein is exuded by the dinoflagellate in response to host stimuli or that the symbiont could increase production of TSP-domain-containing proteins by the release of photosynthate in the form of glucose. These hypotheses require further investigation. It is possible that the addition of glucose to symbiotic anemones will increase infection success in the same way the addition of TSR protein did in functional re-infection experiments of aposymbiotic *Aiptasia* sp. Western blot analysis using a specific *Aiptasia* sp. TSR antibody could then be used to investigate increased expression of TSR proteins in glucose treated anemones.

Further work is needed to explore the potential role of the TGFβ pathway in regulating stable symbiosis and a possible role in symbiosis dysfunction. This work identifies several potential new components of a cnidarian TGFβ pathway. Future functional work is needed to identify whether one of the TSR-domain-containing proteins described here is responsible for activating the TGFβ pathway during symbiosis onset. Research into TGFβ pathway driven innate immune modulation by parasitic species such as *Trypanosoma* spp. suggests that the tolerogenic signal can easily be switched from encouraging pathogenesis to host protection and pathogen destruction. This suggests the possibility that the TGFβ pathway is involved in symbiosis dysfunction and coral bleaching. Heat stress experiments in *Aiptaisia* sp. indicated that TGFβ treated anemones showed significantly less symbiont loss than animals without TGFβ (Detournay et al. 2012). The dominant role of TGFβ in the immune system is to induce tolerance, contain and resolve inflammation (Li et al. 2006). This role of maintaining immune homeostasis is essential to a successful symbiotic relationship between host and symbiont.

Lastly the sequencing of more cnidarian genomes, combined with transcriptomes taken at various symbiotic states for comparison will enable further comparative genomic studies and lead to a better understanding of ECM protein cross-talk and innate immune complexity in cnidarians.

# Bibliography


Abascal, F., Abascal, F., R. Zardoya, R. Zardoya, D. Posada, and D. Posada. 2005. ProtTest: selection of best-fit models of protein evolution. Bioinformatics 21:2104–2105.

Adams, J. C., and R. P. Tucker. 2000. The thrombospondin type 1 repeat (TSR) superfamily: diverse proteins with related roles in neuronal development. Developmental Dynamics 218:280–299.

Allemand, D., P. Furla, and S. Bénazet-Tambutté. 1998. Mechanisms of carbon acquisition for endosymbiont photosynthesis in Anthozoa. Canadian Journal of Botany 76:925–941.

Anderson, J. C., J. R. Grammer, W. Wang, L. B. Nabors, J. Henkin, J. E. Stewart, and C. L. Gladson. 2007. ABT-510, a modified type 1 repeat peptide of thrombospondin, inhibits malignant glioma growth in vivo by inhibiting angiogenesis. Cancer Biology & Therapy 6:454–462.

Archibald, J. M., and P. J. Keeling. 2002. Recycled plastids: a “green movement” in eukaryotic evolution. Trends in Genetics : TIG 18:577–584.

Areschoug, T., and S. Gordon. 2009. Scavenger receptors: role in innate immunity and microbial pathogenesis. Cellular Microbiology 11:1160–1169.

Areschoug, T., J. Waldemarsson, and S. Gordon. 2008. Evasion of macrophage scavenger receptor A-mediated recognition by pathogenic streptococci. European Journal of Immunology 38:3068–3079.

Aruffo, A., M. A. Bowen, D. D. Patel, B. F. Haynes, G. C. Starling, J. A. Gebe, and J. Bajorath. 1997. CD6-ligand interactions: a paradigm for SRCR domain function? Immunology Today 18:498–504.

Asch, A. S., J. Barnwell, R. L. Silverstein, and R. L. Nachman. 1987. Isolation of the thrombospondin membrane receptor. Journal of Clinical Investigation 79:1054–1061.

Babcock, R. C., G. D. Bull, P. L. Harrison, A. J. Heyward, J. K. Oliver, C. C. Wallace, and B. L. Willis. 1986. Synchronous spawnings of 105 scleractinian coral species on the Great Barrier Reef. Marine Biology 90:379–394.

Baker, A. C. 2003. Flexibility and specificity in coral-algal symbiosis: diversity, ecology, and biogeography of Symbiodinium. Annual Review of Ecology 34:661–689.

Baldauf, S. L. 2003. The deep roots of eukaryotes. Science 300:1703–1706.

Baumgarten, S., O. Simakov, L. Y. Esherick, Y. J. Liew, E. M. Lehnert, C. T. Michell, Y. Li, E. A. Hambleton, A. Guse, M. E. Oates, J. Gough, V. M. Weis, M. Aranda, J. R. Pringle, and C. R. Voolstra. 2015. The genome of Aiptasia, a sea anemone model for coral symbiosis. Proceedings of the National Academy of Sciences of the United States of America 112:11893–11898.

Bay, L. 2011. Infection dynamics vary between Symbiodinium Types and cell surface treatments during establishment of endosymbiosis with coral larvae. Diversity 3:356–374.

Belda-Baillie, C. A., B. K. Baillie, and T. Maruyama. 2002. Specificity of a model cnidarian-dinoflagellate symbiosis. The Biological Bulletin 202:74–85.

Bentley, A. A., and J. C. Adams. 2010. The evolution of thrombospondins and their ligand-binding activities. Molecular Biology and Evolution 27:2187–2197.

Bierie, B., and H. L. Moses. 2006. Tumour microenvironment: TGFbeta: the molecular Jekyll and Hyde of cancer. Nature reviews. Cancer 6:506–520.

Blumbach, B., Z. Pancer, B. Diehl-Seifert, R. Steffen, J. Munkner, I. Muller, and W. Muller. 1998. The putative sponge aggregation receptor. Isolation and characterization of a molecule composed of scavenger receptor cysteine-rich domains and short consensus repeats. Journal of Cell Science 111:2635.

Bosch, T. 2008. *Innate Immunity of Plants, Animals, and Humans. 27 Nucleic Acids and Molecular Biology 21.* (H. Heine, Ed.). Springer-Verlag Berlin Heidelberg 2008.

Bowen, M. A. 1995. Cloning, mapping, and characterization of activated leukocyte-cell adhesion molecule (ALCAM), a CD6 ligand. Journal of Experimental Medicine 181:2213–2220.

Calderón, J. C., Calderón, J. C., H. Curtidor, O. González, O. González, G. Cifuentes, G. Cifuentes, C. Reyes, C. Reyes, and M. E. Patarroyo. 2008. High affinity interactions between red blood cell receptors and synthetic Plasmodium thrombospondin-related apical merozoite protein (PTRAMP) peptides. Biochimie 90:802–810.

Cambi, A., M. Koopman, and C. G. Figdor. 2005. How C-type lectins detect pathogens. Cellular Microbiology 7:481–488.

Cano, M. D. V., E. D. Karagiannis, M. Soliman, B. Bakir, W. Zhuang, A. S. Popel, and P. L. Gehlbach. 2009. A peptide derived from type 1 thrombospondin repeat-containing protein WISP-1 inhibits corneal and choroidal neovascularization. Investigative Ophthalmology & Visual Science 50:3840–3845.

Canton, J. 2013. Scavenger receptors in homeostasis and immunity. Nature Reviews Immunology 13:621–634.

Castañeda, O., V. Sotolongo, A. M. Amor, R. Stöcklin, A. J. Anderson, A. L. Harvey, Å. Engström, C. Wernstedt, and E. Karlsson. 1995. Characterization of a potassium channel toxin from the Caribbean Sea anemone *Stichodactyla helianthus*. Toxicon : official journal of the International Society on Toxinology 33:603–613.

Catanese, M. T., H. Ansuini, R. Graziani, T. Huby, M. Moreau, J. K. Ball, G. Paonessa, C. M. Rice, R. Cortese, A. Vitelli, and A. Nicosia. 2010. Role of scavenger receptor class B type I in hepatitis C virus entry: kinetics and molecular determinants. Journal of Virology 84:34–43.

Catanese, M. T., R. Graziani, T. von Hahn, M. Moreau, T. Huby, G. Paonessa, C. Santini, A. Luzzago, C. M. Rice, R. Cortese, A. Vitelli, and A. Nicosia. 2007. High-Avidity Monoclonal Antibodies against the Human Scavenger Class B Type I Receptor Efficiently Block Hepatitis C Virus Infection in the Presence of High-Density Lipoprotein. Journal of Virology 81:8063–8071.

Cernichiari, E., L. Muscatine, and D. C. Smith. 1969. Maltose Excretion by the Symbiotic Algae of Hydra viridis. Proceedings of the Royal Society of London B: Biological Sciences 173:557–576.

Chang, H., C. W. Brown, and M. M. Matzuk. 2002. Genetic analysis of the mammalian transforming growth factor-beta superfamily. Endocrine Reviews 23:787–823.

Chen, C., K. K. Lam, Y. Nakano, and W. S. Tsai. 2003a. A Stable Association of the Stress-Tolerant Zooxanthellae, Symbiodinium Clade D, with the Low-Temperature-Tolerant Coral, Oulastrea crispata (Scleractinia: Faviidae) in Subtropical Non-Reefal Coral Communities. Zoological Studies 42:540–550.

Chen, M.-C., M.-C. Hong, Y.-S. Huang, M.-C. Liu, Y.-M. Cheng, and L.-S. Fang. 2005. ApRab11, a cnidarian homologue of the recycling regulatory protein Rab11, is involved in the establishment and maintenance of the Aiptasia-Symbiodinium endosymbiosis. Biochemical and Biophysical Research Communications 338:1607–1616.

Chen, M.-C., Y.-M. Cheng, M.-C. Hong, and L.-S. Fang. 2004. Molecular cloning of Rab5 (ApRab5) in *Aiptasia pulchella* and its retention in phagosomes harboring live zooxanthellae. Biochemical and Biophysical Research Communications 324:1024–1033.

Chen, M.-C., Y.-M. Cheng, P.-J. Sung, C.-E. Kuo, and L.-S. Fang. 2003b. Molecular identification of Rab7 (ApRab7) in Aiptasia pulchella and its exclusion from phagosomes harboring zooxanthellae. Biochemical and Biophysical Research Communications 308:586–595.

Coats, D. 1999. Parasitic life styles of marine dinoflagellates. Pages 402–409 *in*.

Coffroth, M. A., and S. R. Santos. 2005, June. Genetic diversity of symbiotic dinoflagellates in the genus *Symbiodinium.* Protist 156:19–34.

Colley, N., Colley, N., R. Trench, and R. Trench. 1983. Selectivity in phagocytosis and persistence of symbiotic algae by the scyphistoma stage of the jellyfish Cassiopeia xamachana. Proceedings of the Royal Society of London. Series B, Biological Sciences:61–82.

Dabir, P., T. E. Marinic, I. Krukovets, and O. I. Stenina. 2008. Aryl hydrocarbon receptor is activated by glucose and regulates the thrombospondin-1 gene promoter in endothelial cells. Circulation Research 102:1558–1565.

Davis, A., R. Makar, R. Makar, C. Stowell, C. Stowell, D. Kuter, D. Kuter, W. Dzik, and W. Dzik. 2009. ADAMTS13 binds to CD36: a potential mechanism for platelet and endothelial localization of ADAMTS13. Transfusion 49:206–213.

Davy, S. 2001. The influence of “host release factor” on carbon release by zooxanthellae isolated from fed and starved *Aiptasia pallida* (Verrill). Comparative Biochemistry and Physiology - A Molecular and Integrative Physiology 129:487–494.

Davy, S. K., and J. R. Turner. 2003. Early development and acquisition of Zooxanthellae in the temperate symbiotic sea anemone Anthopleura ballii (Cocks). The Biological Bulletin 205:66–72.

Davy, S. K., S. K. Davy, D. Allemand, D. Allemand, and V. M. Weis. 2012. Cell Biology of Cnidarian-Dinoflagellate Symbiosis. Microbiology and Molecular Biology Reviews 76:229–261.

Dawson, D. W., Dawson, D. W., S. F. Pearce, S. F. Pearce, R. Zhong, R. Zhong, R. L. Silverstein, R. L. Silverstein, W. A. Frazier, W. A. Frazier, N. P. Bouck, and N. P. Bouck. 1997. CD36 mediates the in vitro inhibitory effects of thrombospondin-1 on endothelial cells. The Journal of Cell Biology 138:707–717.

Dawson, D. W., O. V. Volpert, S. F. Pearce, A. J. Schneider, R. L. Silverstein, J. Henkin, and N. P. Bouck. 1999. Three distinct D-amino acid substitutions confer potent antiangiogenic activity on an inactive peptide derived from a thrombospondin-1

type 1 repeat. Molecular pharmacology 55:332–338.

De Bary, A. H. 1878. Die Erscheinung der Symbiose. (K. J. Thubner, Ed.). Strassburg : Verlag von Karl J. Trübner, 1879.

Detournay, O., and V. M. Weis. 2011. Role of the sphingosine rheostat in the regulation of cnidarian-dinoflagellate symbioses. The Biological Bulletin 221:261–269.

Detournay, O., C. E. Schnitzler, A. Poole, and V. M. Weis. 2012. Regulation of cnidarian-dinoflagellate mutualisms: Evidence that activation of a host TGFβ innate immune pathway promotes tolerance of the symbiont. Developmental and Comparative Immunology 38:525–537.

Dietrich, J., and T. M. Doherty. 2009. Interaction of Mycobacterium tuberculosiswith the host: consequences for vaccine development. APMIS 117:440–457.

Dietrich, J., V. M. Weis, A. Robertson, T. M. Doherty, S. K. Davy, J. Croce, O. Hoegh-Guldberg, S. Carbonneau, M. Rodriguez-Lanetty, E. Voronina, J. R. Pringle, E. Miranda, D. R. Mcclay, and J. A. Coffman. 2006. The genomic underpinnings of apoptosis in Strongylocentrotus purpuratus. Developmental Biology 300:321–334.

Dinguirard, N., and T. P. Yoshino. 2006. Potential role of a CD36-like class B scavenger receptor in the binding of modified low-density lipoprotein (acLDL) to the tegumental surface of *Schistosoma mansoni* sporocysts. Molecular and Biochemical Parasitology 146:219–230.

Douglas, A. E. 2003. Coral bleaching--how and why? Marine Pollution Bulletin 46:385–392.

Douglas, A. E. 2008. Conflict, cheats and the persistence of symbioses. The New Phycologist 177:849–858.

Douglas, A. E. 2009. Endosymbionts and Intracellular Parasites. Pages 128–141 Encyclopedia of Microbiology. Third Edition.

Dunkelberger, J. R., and W. C. Song. 2010. Complement and its role in innate and adaptive immune responses. Cell Research 20:34-50.

Dunn, C. W., A. Hejnol, D. Q. Matus, K. Pang, W. E. Browne, S. A. Smith, E. Seaver, G. W. Rouse, M. Obst, G. D. Edgecombe, M. V. Sørensen, S. H. D. Haddock, A. Schmidt-Rhaesa, A. Okusu, R. M. Kristensen, W. C. Wheeler, M. Q. Martindale, and G. Giribet. 2008. Broad phylogenomic sampling improves resolution of the animal tree of life. Nature 452:745–749.

Dunn, S. R., and V. M. Weis. 2009. Apoptosis as a post-phagocytic winnowing mechanism

in a coral-dinoflagellate mutualism. Environmental Microbiology 11:268–276.

Dunn, S. R., C. E. Schnitzler, and V. M. Weis. 2007a. Apoptosis and autophagy as mechanisms of dinoflagellate symbiont release during cnidarian bleaching: every which way you lose. Proceedings of the Royal Society B: Biological Sciences 274:3079–3085.

Dunn, S. R., W. S. Phillips, D. R. Green, and V. M. Weis. 2007b. Knockdown of actin and caspase gene expression by RNA interference in the symbiotic anemone Aiptasia pallida. The Biological Bulletin 212:250–258.

Dunn, S. R., W. S. Phillips, J. W. Spatafora, D. R. Green, and V. M. Weis. 2006. Highly conserved caspase and Bcl-2 homologues from the sea anemone Aiptasia pallida: lower metazoans as models for the study of apoptosis evolution. Journal of Molecular Evolution 63:95–107.

Ejigiri, I., and P. Sinnis. 2009. Plasmodium sporozoite–host interactions from the dermis to the hepatocyte. Current Opinion in Microbiology 12:401–407.

Faure, M., and C. Rabourdin-Combe. 2011. Innate immunity modulation in virus entry. Current Opinion in Virology 1:6–12.

Fernando, T., Fernando, T., S. Flibotte, S. Flibotte, S. Xiong, S. Xiong, J. Yin, J. Yin, E. Yzeiraj, E. Yzeiraj, D. G. Moerman, D. G. Moerman, A. Meléndez, A. Meléndez, C. Savage-Dunn, and C. Savage-Dunn. 2011. C. elegans ADAMTS ADT-2 regulates body size by modulating TGFβ signaling and cuticle collagen organization. Developmental Biology 352:92–103.

Fitt, W. K., and R. K. Trench. 1983. Endocytosis of the symbiotic dinoflagellate Symbiodinium microadriaticum Freudenthal by endodermal cells of the scyphistomae of Cassiopeia xamachana and resistance of the algae to host digestion. Journal of Cell Science 64:195–212.

Flannagan, R. S., G. Cosío, and S. Grinstein. 2009. Antimicrobial mechanisms of phagocytes and bacterial evasion strategies. Nature Reviews Microbiology 7:355–366.

Folkman, J. 2004. Endogenous angiogenesis inhibitors. APMIS.

Foth, B. J., and G. I. McFadden. 2003. The apicoplast: a plastid in Plasmodium falciparum and other Apicomplexan parasites. International review of cytology 224:57–110.

Fransolet, D. 2012. Establishment of endosymbiosis: The case of cnidarians and Symbiodinium. Journal of Experimental Marine Biology and Ecology 420-421:1–7.

Frieda, S., A. Pearce, J. Wu, and R. L. Silverstein. 1995. Recombinant GST/CD36 fusion proteins define a thrombospondin binding domain. Evidence for a single calcium-dependent binding site on CD36. The Journal of Biological Chemistry 270:2981–2986.

Furukawa, R., M. Matsumoto, and H. Kaneko. 2012. Characterization of a scavenger receptor cysteine-rich-domain-containing protein of the starfish, Asterina pectinifera: ApSRCR1 acts as an opsonin in the larval and adult innate immune systems. Developmental and Comparative Immunology 36:51–61.

Ganot, P., A. Moya, V. Magnone, D. Allemand, P. Furla, and C. Sabourault. 2011. Adaptations to endosymbiosis in a cnidarian-dinoflagellate association: differential gene expression and specific gene duplications. PLoS Genetics 7:e1002187.

Gantt, S. M., P. Clavijo, X. Bai, J. D. Esko, and P. Sinnis. 1997. Cell adhesion to a motif shared by the malaria circumsporozoite protein and thrombospondin is mediated by its glycosaminoglycan-binding region and not by CSVTCG. The Journal of Biological Chemistry 272:19205–19213.

Garside, S. A., J. Henkin, K. D. Morris, S. M. Norvell, F. H. Thomas, and H. M. Fraser. 2010. A Thrombospondin-Mimetic Peptide, ABT-898, Suppresses Angiogenesis and Promotes Follicular Atresia in Pre- and Early-Antral Follicles in Vivo. Endocrinology 151:5905–5915.

Gates, R., Gates, R., O. Hoegh-Guldberg, M. McFall-Ngai, M. McFall-Ngai, K. Bil, K. Bil, and L. Muscatine. 1995. Free amino acids exhibit anthozoan “host factor” activity: they induce the release of photosynthate from symbiotic dinoflagellates in vitro. Proceedings of the National Academy of Sciences of the United States of America 92:7430–7434.

Gelman, A. 2005. Analysis of variance—why it is more important than ever. The Annals of Statistics 33:1–53.

Glider, W. V., D. W. Phipps join(' ', and R. L. Pardy. 1980. Localization of Symbiotic Dinoflagellate Cells within Tentacle Tissue of Aiptasia pallida (Coelenterata, Anthozoa). Transactions of the American Microscopical Society 99:426.

Goldstein, J. L., Y. K. Ho, S. K. Basu, and M. S. Brown. 1979. Binding site on macrophages that mediates uptake and degradation of acetylated low density lipoprotein, producing massive cholesterol deposition. Proceedings of the National Academy of Sciences of the United States of America 76:333–337.

Gordon, S. 2002. Pattern recognition receptors doubling up for the innate immune response. Cell 111:927–930.

Gough, P. J., and S. Gordon. 2000. The role of scavenger receptors in the innate immune system. Microbes and infection / Institut Pasteur 2:305–311.

Greaves, D. R., and S. Gordon. 2009. The macrophage scavenger receptor at 30 years of age: current knowledge and future challenges. Journal of Lipid Research 50 Suppl:S282–6.

Greenberg, S., and S. Grinstein. 2002. Phagocytosis and innate immunity. Current opinion in immunology 14:136–145.

Gros, P., F. J. Milder, and B. J. C. Janssen. 2008. Complement driven by conformational changes. Nature Reviews Immunology 8:48–58.

Guo, N. H., H. C. Krutzsch, E. Negre, V. S. Zabrenetzky, and D. D. Roberts. 1992. Heparin-binding peptides from the type I repeats of thrombospondin. Structural requirements for heparin binding and promotion of melanoma cell adhesion and chemotaxis. The Journal of Biological Chemistry 267:19349–19355.

Hamada, M., E. Shoguchi, C. Shinzato, T. Kawashima, D. J. Miller, and N. Satoh. 2012. The Complex NOD-Like Receptor Repertoire of the Coral Acropora digitifera Includes Novel Domain Combinations. Molecular Biology and Evolution 30:167–176.

Hawkins, T. 2015. Antioxidant responses to heat and light stress differ with habitat in a common reef coral. Coral Reefs 34:1229–1241.

Hawkins, T. D., B. J. Bradley, and S. K. Davy. 2013. Nitric oxide mediates coral bleaching through an apoptotic-like cell death pathway: evidence from a model sea anemone-dinoflagellate symbiosis. FASEB journal : official publication of the Federation of American Societies for Experimental Biology 27:4790–4798.

Hibino, T., M. Loza-Coll, C. Messier, A. J. Majeske, A. H. Cohen, D. P. Terwilliger, K. M. Buckley, V. Brockton, S. V. Nair, K. Berney, S. D. Fugmann, M. K. Anderson, Z. Pancer, R. A. Cameron, L. C. Smith, and J. P. Rast. 2006. The immune gene repertoire encoded in the purple sea urchin genome. Developmental Biology 300:349–365.

Hohenester, E., T. Sasaki, and R. Timpl. 1999. Crystal structure of a scavenger receptor cysteine-rich domain sheds light on an ancient superfamily. Nature Structural Biology 6:228–232.

Hohman, T., P. McNeil, and L. Muscatine. 1982. Phagosome-lysosome fusion inhibited by algal symbionts of *Hydra viridis.* The Journal of Cell Biology 94:56.

Hong, M.-C., Y.-S. Huang, P.-C. Song, W.-W. Lin, L.-S. Fang, and M.-C. Chen. 2009a. Cloning and characterization of ApRab4, a recycling Rab protein of Aiptasia pulchella, and its implication in the symbiosome biogenesis. Marine Biotechnology 11:771–785.

Hong, M.-C., Y.-S. Huang, W.-W. Lin, L.-S. Fang, and M.-C. Chen. 2009b. ApRab3, a biosynthetic Rab protein, accumulates on the maturing phagosomes and symbiosomes in the tropical sea anemone, *Aiptasia pulchella*. Comparative Biochemistry and Physiology 152:249–259.

Hsu, H. Y., S. L. Chiu, M. H. Wen, K. Y. Chen, and K. F. Hua. 2001. Ligands of macrophage scavenger receptor induce cytokine expression via differential modulation of protein kinase signaling pathways. The Journal of Biological Chemistry 276:28719–28730.

Huang, S., S. Yuan, L. Guo, Y. Yu, J. Li, T. Wu, T. Liu, M. Yang, K. Wu, H. Liu, J. Ge, Y. Yu, H. Huang, M. Dong, C. Yu, S. Chen, and A. Xu. 2008. Genomic analysis of the immune gene repertoire of amphioxus reveals extraordinary innate complexity and diversity. Genome Research 18:1112–1126.

Hughes, T. P., A. H. Baird, D. R. Bellwood, M. Card, S. R. Connolly, C. Folke, R. Grosberg, O. Hoegh-Guldberg, J. B. C. Jackson, J. Kleypas, J. M. Lough, P. Marshall, M. Nyström, S. R. Palumbi, J. M. Pandolfi, B. Rosen, and J. Roughgarden. 2003. Climate change, human impacts, and the resilience of coral reefs. Science 301:929–933.

Huxley-Jones, J., S. S. Apte, S. S. Apte, D. L. Robertson, and R. P. Boot-Handford. 2005. The characterisation of six ADAMTS proteases in the basal chordate *Ciona intestinalis* provides new insights into the vertebrate ADAMTS family. The international Journal of Biochemistry & Cell Biology 37:1838–1845.

Ikushima, H., and K. Miyazono. 2010. TGFbeta signalling: a complex web in cancer progression. Nature reviews. Cancer 10:415–424.

Janeway, C. A., Jr, and R. Medzhitov. 2002. Innate immune recognition. Annual Review of Immunology 20:197–216.

Janouskovec, J., A. Horák, M. Oborník, J. Lukes, and P. J. Keeling. 2010a. A common red algal origin of the apicomplexan, dinoflagellate, and heterokont plastids. Proceedings of the National Academy of Sciences of the United States of America 107:10949–10954.

Janouskovec, J., Armstrong, L. C., A. Horák, P. Bornstein, M. Oborník, J. Lukes, and P. J. Keeling. 2010b. A common red algal origin of the apicomplexan, dinoflagellate, and

heterokont plastids. Proceedings of the National Academy of Sciences 107:10949–10954.

Jolley, E., and D. Smith. 1980. The green hydra symbiosis. II. The biology of the establishment of the association. Proceedings of the Royal Society of London. Series B, Biological Sciences:311–333.

Kappe, S., T. Bruderer, S. Gantt, H. Fujioka, V. Nussenzweig, and R. Menard. 1999. Conservation of a Gliding Motility and Cell Invasion Machinery in Apicomplexan Parasites. The Journal of Cell Biology 147:937–944.

Karagiannis, E. D., and A. S. Popel. 2007. Peptides derived from type I thrombospondin repeat-containing proteins of the CCN family inhibit proliferation and migration of endothelial cells. The International Journal of Biochemistry & Cell Biology 39:2314–2323.

Katoh, K., K. Misawa, K.-I. Kuma, and T. Miyata. 2002. MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. Nucleic Acids Research 30:3059–3066.

Kazandjian, A., V. Shepherd, M. Rodriguez-Lanetty, W. Nordemeier, A. Larkum, and R. Quinnell. 2008. Isolation of symbiosomes and the symbiosome membrane complex from the zoanthid *Zoanthus robustus*. Phycologia 47:294–306.

Kazerounian, S., K. O. Yee, and J. Lawler. 2008. Thrombospondins in cancer. Cellular and Molecular Life Sciences (CMLS) 65:700–712.

Kearse, M., R. Moir, A. Wilson, S. Stones-Havas, M. Cheung, S. Sturrock, S. Buxton, A. Cooper, S. Markowitz, C. Duran, T. Thierer, B. Ashton, P. Meintjes, and A. Drummond. 2012. Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. Bioinformatics 28:1647–1649.

Keeling, P. J. 2004. Diversity and evolutionary history of plastids and their hosts. American Journal of Botany 91:1481–1493.

Keeling, P. J. 2013. The Number, Speed, and Impact of Plastid Endosymbioses in Eukaryotic Evolution. Annual Review of Plant Biology 64:583–607.

Kenkel, C. D., E. Meyer, and M. V. Matz. 2013. Gene expression under chronic heat stress in populations of the mustard hill coral (Porites astreoides) from different thermal environments. Molecular Ecology 22:4322–4334.

Kerrigan, A. M., and G. D. Brown. 2009. C-type lectins and phagocytosis. Immunobiology 214:562–575.

Khalil, N. 1999. TGF-beta: from latent to active. Microbes and infection / Institut Pasteur 1:1255–1263.

Kimura, A., E. Sakaguchi, and M. Nonaka. 2009. Multi-component complement system of Cnidaria: C3, Bf, and MASP genes expressed in the endodermal tissues of a sea anemone, Nematostella vectensis. Immunobiology 214:165–178.

Kitchen, S. A., C. M. Crowder, A. Z. Poole, V. M. Weis, and E. Meyer. 2015. De Novo Assembly and Characterization of Four Anthozoan (Phylum Cnidaria) Transcriptomes. G3 (Bethesda, Md.) 5:2441–2452.

Koli, K., J. Saharinen, M. Hyytiäinen, C. Penttinen, and J. Keski-Oja. 2001. Latency, activation, and binding proteins of TGF-beta. Microscopy Research and Technique 52:354–362.

Krieger, M. 1992. Molecular flypaper and atherosclerosis: structure of the macrophage scavenger receptor. Trends in Biochemical Sciences 17:141–146.

Krieger, M. 1997. The other side of scavenger receptors: pattern recognition for host defense. Current Opinion in Lipidology 8:275.

Krieger, M. 2001. Scavenger receptor class B type I is a multiligand HDL receptor that influences diverse physiologic systems. Journal of Clinical Investigation 108:793–797.

Kruger, R. P., J. Aurandt, and K.-L. Guan. 2005. Semaphorins command cells to move. Nature Reviews Molecular Cell Biology 6:789–800.

Kuno, K., Kuno, K., K. Matsushima, and K. Matsushima. 1998. ADAMTS-1 protein anchors at the extracellular matrix through the thrombospondin type I motifs and its spacing region. The Journal of Biological Chemistry 273:13912–13917.

Kuo, J., M.-C. Chen, C.-H. Lin, and L.-S. Fang. 2004. Comparative gene expression in the symbiotic and aposymbiotic Aiptasia pulchella by expressed sequence tag analysis. Biochemical and Biophysical Research Communications 318:176–186.

Kvennefors, E. C. E., W. Leggat, C. C. Kerr, T. D. Ainsworth, O. Hoegh-Guldberg, and A. C. Barnes. 2010. Analysis of evolutionarily conserved innate immune components in coral links immunity and symbiosis. Developmental and Comparative Immunology 34:1219–1229.

Kvennefors, E. C. E., W. Leggat, O. Hoegh-Guldberg, B. M. Degnan, and A. C. Barnes. 2008. An ancient and variable mannose-binding lectin from the coral Acropora millepora binds both pathogens and symbionts. Developmental and Comparative Immunology

32:1582–1592.

L. Muscatine, J. E. Christensen, R. R. Pool, A. R. Thomsen, and R. K. Trench. 1975. Symbiosis of algae and invertebrates: aspects of the symbiont surface and the host-symbiont interface. Transactions of the American Microscopical Society 94:450–469.

LaJeunesse, T. C. 2001. Investigating The Biodiversity, Ecology, And Phylogeny Of Endosymbiotic Dinoflagellates In The Genus Symbiodinium Using The Its Region: In Search Of A "Species" Level Marker. Journal of Phycology 37:866–880.

LaJeunesse, T., D. J. Thornhill, E. F. Cox, F. G. Stanton, W. K. Fitt, and G. W. Schmidt. 2004. High diversity and host specificity observed among symbiotic dinoflagellates in reef coral communities from Hawaii. Coral Reefs 23:596–603.

Lawler, J. 1986. The structural and functional properties of thrombospondin. Blood 67:1197–1209.

Lawler, J. 2000. The functions of thrombospondin-1 and-2. Current Opinion in Cell Biology 12:634–640.

Lehnert, E. M., M. E. Mouchka, M. S. Burriesci, N. D. Gallo, J. A. Schwarz, and J. R. Pringle. 2014. Extensive differences in gene expression between symbiotic and aposymbiotic cnidarians. G3 (Bethesda, Md.) 4:277–295.

Lehnert, E. M., M. S. Burriesci, and J. R. Pringle. 2012. Developing the anemone Aiptasia as a tractable model for cnidarian-dinoflagellate symbiosis: the transcriptome of aposymbiotic *A. pallida*. BMC Genomics 13:271.

Lesser, M. P., M. Stat, and R. D. Gates. 2013. The endosymbiotic dinoflagellates (Symbiodinium sp.) of corals are parasites and mutualists. Coral Reefs:603–611.

Lesser, M., and J. M. Shick. 1989. Effects of irradiance and ultraviolet radiation on photoadaptation in the zooxanthellae of Aiptasia pallida: primary production, photoinhibition, and enzymic defenses against oxygen toxicity. Marine Biology 102:243–255.

Leung, L. L., W. X. Li, J. L. McGregor, G. Albrecht, and R. J. Howard. 1992. CD36 peptides enhance or inhibit CD36-thrombospondin binding. A two-step process of ligand-receptor interaction. The Journal of Biological Chemistry 267:18244–18250.

Leung, T., and R. Poulin. 2008. Parasitism, commensalism, and mutualism: exploring the many shades of symbioses. Vie et Milieu - Life and Environment 58:107–115.

Li, B., F. Lu, X. Wei, and R. Zhao. 2008. Fucoidan: structure and bioactivity. Molecules

(Basel, Switzerland) 13:1671–1695.

Li, M. O., Y. Y. Wan, S. Sanjabi, A.-K. L. Robertson, and R. A. Flavell. 2006. Transforming growth factor-beta regulation of immune responses. Annual Review of Immunology 24:99–146.

Li, W. X., R. J. Howard, and L. L. Leung. 1993. Identification of SVTCG in thrombospondin as the conformation-dependent, high affinity binding site for its receptor, CD36. The Journal of Biological Chemistry 268:16179–16184.

Lin, K., J. T. Wang, and L. S. Fang. 2000. Participation of glycoproteins on zooxanthellal cell walls in the establishment of a symbiotic relationship with the sea anemone, *Aiptasia pulchella*. Zoological Studies 39:172–178.

Lingelbach, K., and K. A. Joiner. 1998. The parasitophorous vacuole membrane surrounding Plasmodium and Toxoplasma: an unusual compartment in infected cells. Journal of Cell Science 111 ( Pt 11):1467–1475.

Logan, D. D. K., A. C. LaFlamme, V. M. Weis, and S. K. Davy. 2010. Flow-Cytometric Characterization Of The Cell-Surface Glycans Of Symbiotic Dinoflagellates (*Symbiodinium* Spp.). Journal of Phycology 46:525–533.

Lopez-Dee, Z., K. Pidcock, and L. S. Gutierrez. 2011. Thrombospondin-1: multiple paths to inflammation. Mediators of Inflammation 2011:296069–10.

López, J. A., M. G. Fain, and L. F. Cadavid. 2011. The evolution of the immune-type gene family Rhamnospondin in cnidarians. Gene 473:119–124.

Madsen, M., H. J. Møller, M. J. Nielsen, C. Jacobsen, J. H. Graversen, T. van den Berg, and S. K. Moestrup. 2004. Molecular characterization of the haptoglobin.hemoglobin receptor CD163. Ligand binding properties of the scavenger receptor cysteine-rich domain region. The Journal of Biological Chemistry 279:51561–51567.

Margulis, L. 1975. Symbiotic theory of the origin of eukaryotic organelles; criteria for proof. Symposia of the Society for Experimental Biology:21–38.

Margulis, L. 1988. Serial endosymbiotic theory (set)-undulipodia, mitosis and their microtubule systems preceded mitochondria. Endocytobiosis and Cell Research 5:133–162.

Margulis, L., and D. Bermudes. 1985. Symbiosis as a mechanism of evolution: status of cell symbiosis theory. Symbiosis 1:101–124.

Markell, D. A., and R. K. Trench. 1993. Macromolecules Exuded By Symbiotic Dinoflagellates In Culture: Amino Acid And Sugar Composition. Journal of Phycology

29:64–68.

Markell, D., and E. M. Wood-Charlson. 2010. Immunocytochemical evidence that symbiotic algae secrete potential recognition signal molecules in hospite. Marine Biology 157:1105–1111.

Martinez, V. G., S. K. Moestrup, U. Holmskov, J. Mollenhauer, and F. Lozano. 2011. The conserved scavenger receptor cysteine-rich superfamily in therapy and diagnosis. Pharmacological Reviews 63:967–1000.

Massagué, J. 2008. TGFbeta in Cancer. Cell 134:215–230.

McAuley, P. J., and D. C. Smith. 1982. The Green Hydra Symbiosis. V. Stages in the Intracellular Recognition of Algal Symbionts by Digestive Cells. Proceedings of the Royal Society of London. Series B, Biological Sciences 216:7–23.

McFadden, G. I., M. E. Reith, J. Munholland, and N. Lang-Unnasch. 1996. Plastid in human parasites. Nature 381:482.

Mcfall-Ngai, M. 2007. Adaptive Immunity: Care for the community. Nature 445:153–153.

McNeil, P. L. 1981. Mechanisms of nutritive endocytosis. I. Phagocytic versatility and cellular recognition in Chlorohydra digestive cells, a scanning electron microscope study. Journal of Cell Science 49:311–339.

McNeil, P. L. 1983. Acidification of phagosomes is initiated before lysosomal enzyme activity is detected. The Journal of Cell Biology 97:692–702.

McNeil, P. L., and P. J. McAuley. 1984. Lysosomes fuse with one half of alga-bearing phagosomes during the reestablishment of the European green hydra symbiosis. Journal of Experimental Zoology 230:377–385.

McNeil, P. L., T. C. Hohman, and L. Muscatine. 1981. Mechanisms of nutritive endocytosis. II. The effect of charged agents on phagocytic recognition by digestive cells. Journal of Cell Science 52:243–269.

McNeil, P., McNeil, P., and D. Smith. 1982. The Green Hydra Symbiosis. IV: Entry of Symbionts into Digestive Cells. Proceedings of the Royal Society of London. Series B, Biological Sciences:1–6.

Medzhitov, R. 2007. Recognition of microorganisms and activation of the immune response. Nature 449:819–826.

Medzhitov, R., and C. J. Janeway. 2000. Innate immune recognition: mechanisms and pathways. Immunological Reviews 173:89–97.

Medzhitov, R., and C. Janeway. 2002. Decoding the patterns of self and nonself by the

innate immune system. Science 296:298–300.

Meints, R. H., and R. L. Pardy. 1980. Quantitative demonstration of cell surface involvement in a plant-animal symbiosis: lectin inhibition of reassociation. Journal of Cell Science 43:239–251.

Messier-Solek, C., K. M. Buckley, and J. P. Rast. 2010. Highly diversified innate receptor systems and new forms of animal immunity. Seminars in Immunology 22:39–47.

Meyer, E., and V. M. Weis. 2012. Study of cnidarian-algal symbiosis in the "omics" age. The Biological Bulletin 223:44–65.

Meyer, E., Meyer, E., G. V. Aglyamova, G. V. Aglyamova, S. Wang, S. Wang, J. Buchanan-Carter, J. Buchanan-Carter, D. Abrego, D. Abrego, J. K. Colbourne, J. K. Colbourne, B. L. Willis, B. L. Willis, M. V. Matz, and M. V. Matz. 2009. Sequencing and de novo analysis of a coral larval transcriptome using 454 GSFlx. BMC Genomics 10:219.

Miller, D. J., E. E. Ball, and U. Technau. 2005. Cnidarians and ancestral genetic complexity in the animal kingdom. Trends in genetics : TIG 21:536–539.

Miller, D. J., G. Hemmrich, E. E. Ball, D. C. Hayward, K. Khalturin, N. Funayama, K. Agata, and T. C. Bosch. 2007. The innate immune repertoire in Cnidaria - ancestral complexity and stochastic gene loss. Genome Biology 8:R59.

Ming, M., M. E. Ewen, and M. E. Pereira. 1995. Trypanosome invasion of mammalian cells requires activation of the TGF beta signaling pathway. Cell 82:287–296.

Moore, R. B., M. Oborník, J. Janouskovec, T. Chrudimský, M. Vancová, D. H. Green, S. W. Wright, N. W. Davies, C. J. S. Bolch, K. Heimann, J. Slapeta, O. Hoegh-Guldberg, J. M. Logsdon, and D. A. Carter. 2008. A photosynthetic alveolate closely related to apicomplexan parasites. Nature 451:959–963.

Morahan, B. J., L. Wang, and R. L. Coppel. 2009. No TRAP, no invasion. Trends in Parasitology 25:77–84.

Mordue, D. G., N. Desai, M. Dustin, and L. D. Sibley. 1999. Invasion by *Toxoplasma gondii* establishes a moving junction that selectively excludes host cell plasma membrane proteins on the basis of their membrane anchoring. The Journal of Experimental Medicine 190:1783–1792.

Moya, A., L. Huisman, E. E. Ball, D. C. Hayward, L. C. Grasso, C. M. Chua, H. N. Woo, J.-P. Gattuso, S. Forêt, and D. J. Miller. 2012. Whole transcriptome analysis of the coral Acropora millepora reveals complex responses to $CO_2$-driven acidification during the initiation of calcification. Molecular Ecology 21:2440–2454.

Mukhopadhyay, S., S. Mukhopadhyay, S. Gordon, and S. Gordon. 2004. The role of scavenger receptors in pathogen recognition and innate immunity. Immunobiology 209:39–49.

Muller-Parker, G. 2001. Temperate and tropical algal-sea anemone symbioses. Invertebrate Biology 120:104–123.

Murphy, J. E., P. R. Tedbury, S. Homer-Vanniasinkam, J. H. Walker, and S. Ponnambalam. 2005. Biochemistry and cell biology of mammalian scavenger receptors. Atherosclerosis 182:1–15.

Murphy-Ullrich, J. E., and M. Poczatek. 2000. Activation of latent TGF-β by thrombospondin-1: mechanisms and physiology. Cytokine & Growth Factor Reviews 11:59–69.

Muscatine, L. 1974. Coelenterate Biology - Reviews and New Perspectives. Pages 359–395 Coelenterate Biology Reviews and New Perspectives. Elsevier.

Muscatine, L., and E. Cernichiari. 1969. Assimilation of photosynthetic products of zooxanthellae by a reef coral. Biological Bulletin 137:506–523.

Musumeci, M., L. Malaguarnera, J. Simpore, A. Messina, and S. Musumeci. 2003. Modulation of immune response in *Plasmodium falciparum* malaria: role of IL-12, IL-18 and TGF-beta. Cytokine 21:172–178.

Müller, H. M., I. Reckmann, M. R. Hollingdale, H. Bujard, K. J. Robson, and A. Crisanti. 1993. Thrombospondin related anonymous protein (TRAP) of Plasmodium falciparum binds specifically to sulfated glycoconjugates and to HepG2 hepatoma cells suggesting a role for this molecule in sporozoite invasion of hepatocytes. The EMBO journal 12:2881–2889.

Nde, P. N., M. F. Lima, C. A. Johnson, S. Pratap, and F. Villalta. 2012. Regulation and use of the extracellular matrix by *Trypanosoma cruzi* during early infection. Frontiers in Immunology 3:337.

Ndungu, F. M., B. C. Urban, K. Marsh, and J. Langhorne. 2005. Regulation of immune response by Plasmodium-infected red blood cells. Parasite immunology 27:373–384.

Neuhauser, C. 2004. A mutualism-parasitism continuum model and its application to plant-mycorrhizae interactions. Ecological Modelling 177:337–352.

Nicholson, A. C., S.-B. Malik, J. M. Logsdon, and E. G. Van Meir. 2005. Functional evolution of ADAMTS genes: Evidence from analyses of phylogeny and gene organization.

BMC Evolutionary Biology 5:11.

Nolan, K. F., S. Kaluz, J. M. Higgins, D. Goundis, and K. B. Reid. 1992. Characterization of the human properdin gene. The Biochemical Journal 287 ( Pt 1):291–297.

Nonaka, M. 2011. The complement C3 protein family in invertebrates. Invertebrate Survival Journal 8:21–32.

Nyholm, S. V., and M. Mcfall-Ngai. 2004. The winnowing: establishing the squid–vibrio symbiosis. Nature Reviews Microbiology 2:632–642.

Nyholm, S. V., Nyholm, S. V., M. J. McFall-Ngai, and M. J. McFall-Ngai. 2003. Dominance of Vibrio fischeri in secreted mucus outside the light organ of Euprymna scolopes: the first site of symbiont specificity. Applied and Environmental Microbiology 69:3932–3937.

Ocampo, I. D., A. Zárate-Potes, V. Pizarro, C. A. Rojas, N. E. Vera, and L. F. Cadavid. 2015. The immunotranscriptome of the Caribbean reef-building coral Pseudodiploria strigosa. Immunogenetics 67:515–530.

Okamoto, N., and G. I. McFadden. 2008. The mother of all parasites. Future Microbiology 3:391–395.

Omer, F. M., Omer, F. M., J. B. de Souza, J. B. de Souza, P. H. Corran, P. H. Corran, A. A. Sultan, A. A. Sultan, E. M. Riley, and E. M. Riley. 2003. Activation of Transforming Growth Factor {beta} by Malaria Parasite-derived Metalloproteinases and a Thrombospondin-like Molecule. Journal of Experimental Medicine 198:1817–1827.

Palmer, C. V., and N. Traylor-Knowles. 2012. Towards an integrated network of coral immune mechanisms. Proceedings of the Royal Society of London B: Biological Sciences 279:4106–4114.

Pancer, Z. 2000. Dynamic expression of multiple scavenger receptor cysteine-rich genes in coelomocytes of the purple sea urchin. Proceedings of the National Academy of Sciences of the United States of America 97:13156.

Pancer, Z., J. Munkner, I. Muller, and W. Muller. 1997. A novel member of an ancient superfamily: sponge (*Geodia cydonium*, Porifera) putative protein that features scavenger receptor cysteine-rich repeats1. Gene 193:211–218.

Pancer, Z., J. P. Rast, and E. H. Davidson. 1999. Origins of immunity: transcription factors and homologues of effector genes of the vertebrate immune system expressed in sea urchin coelomocytes. Immunogenetics 49:773–786.

Patarroyo, M. E., G. Arévalo-Pinzón, C. Reyes, A. Moreno-Vranich, and M. A. Patarroyo.

2015. Malaria Parasite Survival Depends on Conserved Binding Peptides' Critical Biological Functions. Current Issues in Molecular Biology 18:57–78.

Peng, S.-E., Y.-B. Wang, L.-H. Wang, W.-N. U. Chen, C.-Y. Lu, L.-S. Fang, and C.-S. Chen. 2010. Proteomic analysis of symbiosome membranes in Cnidaria-dinoflagellate endosymbiosis. Proteomics 10:1002–1016.

Perez, S., and V. Weis. 2006. Nitric oxide and cnidarian bleaching: an eviction notice mediates breakdown of a symbiosis. The Journal of Experimental Biology 209:2804–2810.

Perez, S., and V. Weis. 2008. Cyclophilin and the regulation of symbiosis in *Aiptasia pallida*. The Biological Bulletin 215:63–72.

Perez, S., C. Cook, and W. Brooks. 2001. The role of symbiotic dinoflagellates in the temperature-induced bleaching response of the subtropical sea anemone Aiptasia pallida. Journal of Experimental Marine Biology and Ecology 256:1–14.

Pinto, M. R., D. Melillo, S. Giacomelli, G. Sfyroera, and J. D. Lambris. 2007. Ancient origin of the complement system: emerging invertebrate models. Advances in Experimental Medicine and Biology 598:372–388.

Plüddemann, A., S. Mukhopadhyay, and S. Gordon. 2011. Innate immunity to intracellular pathogens: macrophage receptors and responses to microbial entry. Immunological Reviews 240:11–24.

Pochon, X., T. C. LaJeunesse, and J. Pawlowski. 2004. Biogeographic partitioning and host specialization among foraminiferan dinoflagellate symbionts (Symbiodinium; Dinophyta). Marine Biology 146:17–27.

Polato, N. R., J. C. Vera, and I. B. Baums. 2011. Gene discovery in the threatened elkhorn coral: 454 sequencing of the *Acropora palmata* transcriptome. PLoS ONE 6:e28634.

Poole, A. 2014, September 22. The Role of Innate Immunity in Cnidarian-dinoflagellate Symbiosis. Oregon State University.

Poole, A. Z., and V. M. Weis. 2014. TIR-domain-containing protein repertoire of nine anthozoan species reveals coral-specific expansions and uncharacterized proteins. Developmental and Comparative Immunology 46:480–488.

Porter, S., I. M. Clark, L. Kevorkian, and D. R. Edwards. 2005. The ADAMTS metalloproteinases. The Biochemical Journal 386:15–27.

Pratlong, M., A. Haguenauer, O. Chabrol, C. Klopp, P. Pontarotti, and D. Aurelle. 2015. The red coral (Corallium rubrum) transcriptome: a new resource for population

genetics and local adaptation studies. Molecular ecology resources 15:1205–1215.

Putnam, N. H., Putnam, N. H., M. Srivastava, M. Srivastava, U. Hellsten, U. Hellsten, B. Dirks, B. Dirks, J. Chapman, J. Chapman, A. Salamov, A. Salamov, A. Terry, A. Terry, H. Shapiro, H. Shapiro, E. Lindquist, E. Lindquist, V. V. Kapitonov, V. V. Kapitonov, J. Jurka, J. Jurka, G. Genikhovich, G. Genikhovich, I. V. Grigoriev, I. V. Grigoriev, S. M. Lucas, S. M. Lucas, R. E. Steele, R. E. Steele, J. R. Finnerty, J. R. Finnerty, U. Technau, M. Q. Martindale, D. S. Rokhsar, and D. S. Rokhsar. 2007. Sea Anemone Genome Reveals Ancestral Eumetazoan Gene Repertoire and Genomic Organization. Science 317:86–94.

Quevillon, E., V. Silventoinen, S. Pillai, N. Harte, N. Mulder, R. Apweiler, and R. Lopez. 2005. InterProScan: protein domains identifier. Nucleic Acids Research 33:W116–20.

Rast, J. P., Rast, J. P., C. Messier-Solek, and C. Messier-Solek. 2008. Marine invertebrate genome sequences and our evolving understanding of animal immunity. The Biological Bulletin 214:274–283.

Reaka-Kudla, M. L., D. E. Wilson, and E. O. Wilson. 1996. Biodiversity II: understanding and protecting our biological resources.

Robson, K. J., A. Dolo, I. R. Hackford, O. Doumbo, M. B. Richards, M. M. Keita, T. Sidibe, A. Bosman, D. Modiano, and A. Crisanti. 1998. Natural polymorphism in the thrombospondin-related adhesive protein of Plasmodium falciparum. The American journal of tropical medicine and hygiene 58:81–89.

Robson, K. J., U. Frevert, I. Reckmann, G. Cowan, J. Beier, I. G. Scragg, K. Takehara, D. H. Bishop, G. Pradel, and R. Sinden. 1995. Thrombospondin-related adhesive protein (TRAP) of Plasmodium falciparum: expression during sporozoite ontogeny and binding to human hepatocytes. The EMBO journal 14:3883–3894.

Rodrigues, C. D., M. Hannus, M. Prudêncio, C. Martin, L. A. Gonçalves, S. Portugal, S. Epiphanio, A. Akinc, P. Hadwiger, K. Jahn-Hofmann, I. Röhl, G.-J. van Gemert, J.-F. Franetich, A. J. F. Luty, R. Sauerwein, D. Mazier, V. Koteliansky, H.-P. Vornlocher, C. J. Echeverri, and M. M. Mota. 2008. Host scavenger receptor SR-BI plays a dual role in the establishment of malaria parasite liver infection. Cell Host & Microbe 4:271–282.

Rodriguez-Lanetty, M. 2006. Temporal and spatial infection dynamics indicate recognition events in the early hours of a dinoflagellate/coral symbiosis. Marine

Biology 149:713–719.

Rodriguez-Lanetty, M., W. S. Phillips, and V. M. Weis. 2006. Transcriptome analysis of a cnidarian-dinoflagellate mutualism reveals complex modulation of host gene expression. BMC Genomics 7:23.

Rowan, R., and D. A. Powers. 1991. A Molecular Genetic Classification of Zooxanthellae and the Evolution of Animal-Algal Symbioses. Science 251:1348–1351.

Rue, H., S. Martino, and N. Chopin. 2009. Approximate Bayesian inference for latent Gaussian models by using integrated nested Laplace approximations. Journal of the Royal Statistical Society B 71:319–392.

Ryan, J. F., P. M. Burton, M. E. Mazza, and G. K. Kwong. 2006. The cnidarian-bilaterian ancestor possessed at least 56 homeoboxes: evidence from the starlet sea anemone, *Nematostella vectensis*. Genome Biology 7:R64.

Sachs, J. L., and E. L. Simms. 2006. Pathways to mutualism breakdown. Trends in Ecology & Evolution 21:585–592.

Sachs, J. L., and T. P. Wilcox. 2006. A shift to parasitism in the jellyfish symbiont Symbiodinium microadriaticum. Proceedings of the Royal Society of London B: Biological Sciences 273:425–429.

Sachs, J. L., C. J. Essenberg, and M. M. Turcotte. 2011a. New paradigms for the evolution of beneficial infections. Trends in Ecology & Evolution 26:202–209.

Sachs, J. L., R. G. Skophammer, and J. U. Regus. 2011b. Evolutionary transitions in bacterial symbiosis. Proceedings of the National Academy of Sciences of the United States of America 108 Suppl 2:10800–10807.

Sacks, D., and A. Sher. 2002. Evasion of innate immunity by parasitic protozoa. Nature Immunology 3:1041–1047.

Saldarriaga, J. F., F. J. R. Max Taylor, T. Cavalier-Smith, S. Menden-Deuer, and P. J. Keeling. 2004. Molecular data and the evolutionary history of dinoflagellates. European Journal of Protistology 40:85–111.

Samuel, G., D. Miller, and R. Saint. 2001. Conservation of a DPP/BMP signaling pathway in the nonbilateral cnidarian *Acropora millepora.* Evolution & Development 3:241–250.

Sansonetti, P. 2000. Phagocytosis, a cell biology view. Journal of Cell Science 113 (Pt 19):3355–3356.

Sapp, J. 1994. Evolution by Association : A History of Symbiosis. Oxford University Press.

Sarrias, M., J. Grønlund, O. Padilla, M. Madsen, U. Holmskov, and F. Lozano. 2004. The Scavenger Receptor Cysteine-Rich (SRCR) domain: an ancient and highly conserved protein module of the innate immune system. Critical Reviews in Immunology 24:1–37.

Schäfer, G., R. Guler, G. Murray, F. Brombacher, and G. D. Brown. 2009. The role of scavenger receptor B1 in infection with Mycobacterium tuberculosis in a murine model. PLoS ONE 4:e8448.

Schnitzler, C. E. 2010, June 9. Temperature Stress, Gene Expression, and Innate Immunity at the Onset of Cnidarian-dinoflagellate Symbiosis. Oregon State University.

Schoenberg, D., and R. Trench. 1980. Genetic Variation in Symbiodinium (=Gymnodinium) microadriaticum Freudenthal, and Specificity in its Symbiosis with Marine Invertebrates. I. Isoenzyme and Soluble Protein Patterns of Axenic Cultures of Symbiodinium microadriaticum. Proceedings of the Royal Society of London. Series B, Biological Sciences 207:405–427.

Schwartz, R., and M. Dayhoff. 1978. Origins of prokaryotes, eukaryotes, mitochondria, and chloroplasts. Science 199:395–403.

Schwarz, J. 2008. Understanding the intracellular niche in cnidarian-Symbiodinium symbioses: parasites lead the way. Vie et Milieu - Life and Environment 58 :141–151.

Schwarz, J. A., D. A. Krupp, and V. M. Weis. 1999. Late Larval Development and Onset of Symbiosis in the Scleractinian Coral Fungia scutaria. The Biological Bulletin 196:70–79.

Schwarz, J. A., P. B. Brokstein, C. Voolstra, A. Y. Terry, C. F. Manohar, D. J. Miller, A. M. Szmant, M. A. Coffroth, and M. Medina. 2008. Coral life history and symbiosis: functional genomic resources for two reef building Caribbean corals, Acropora palmata and Montastraea faveolata. BMC Genomics 9:97.

Schwarz, R. S., L. Hodes-Villamar, K. A. Fitzpatrick, M. G. Fain, A. L. Hughes, and L. F. Cadavid. 2007. A gene family of putative immune recognition molecules in the hydroid Hydractinia. Immunogenetics 59:233–246.

Shalev, A., C. A. Pise-Masison, M. Radonovich, S. C. Hoffmann, B. Hirshberg, J. N. Brady, and D. M. Harlan. 2002. Oligonucleotide microarray analysis of intact human pancreatic islets: identification of glucose-responsive genes and a highly regulated

TGFbeta signaling pathway. Endocrinology 143:3695–3698.

Shinzato, C., E. Shoguchi, T. Kawashima, M. Hamada, K. Hisata, M. Tanaka, M. Fujie, M. Fujiwara, R. Koyanagi, T. Ikuta, A. Fujiyama, D. J. Miller, and N. Satoh. 2011. Using the *Acropora digitifera* genome to understand coral responses to environmental change. Nature 476:320–323.

Shinzato, C., M. Inoue, and M. Kusakabe. 2014a. A Snapshot of a Coral "Holobiont": A Transcriptome Assembly of the Scleractinian Coral, Porites , Captures a Wide Variety of Genes from Both the Host and Symbiotic Zooxanthellae. PLoS ONE 9:e85182.

Shinzato, C., S. Mungpakdee, N. Satoh, and E. Shoguchi. 2014b. A genomic approach to coral-dinoflagellate symbiosis: studies of Acropora digitifera and Symbiodinium minutum. Frontiers in Microbiology 5:336.

Shoguchi, E., C. Shinzato, T. Kawashima, F. Gyoja, S. Mungpakdee, R. Koyanagi, T. Takeuchi, K. Hisata, M. Tanaka, M. Fujiwara, M. Hamada, A. Seidi, M. Fujie, T. Usami, H. Goto, S. Yamasaki, N. Arakaki, Y. Suzuki, S. Sugano, A. Toyoda, Y. Kuroki, A. Fujiyama, M. Medina, M. A. Coffroth, D. Bhattacharya, and N. Satoh. 2013. Draft Assembly of the Symbiodinium minutum Nuclear Genome Reveals Dinoflagellate Gene Structure. Current Biology 23:1399–1408.

Sibley, L. D. 2004. Intracellular parasite invasion strategies. Science 304:248–253.

Silverstein, R. L. 2002. The face of TSR revealed: an extracellular signaling domain is exposed. The Journal of Cell Biology 159:203–206.

Silverstein, R. L., and M. Febbraio. 2007. CD36-TSP-HRGP interactions in the regulation of angiogenesis. Current Pharmaceutical Design 13:3559–3567.

Silverstein, R. L., and M. Febbraio. 2009. CD36, a Scavenger Receptor Involved in Immunity, Metabolism, Angiogenesis, and Behavior. Science Signaling 2:review 3.

Silverstein, R. N., A. M. S. Correa, and A. C. Baker. 2012. Specificity is rarely absolute in coral-algal symbiosis: implications for coral response to climate change. Proceedings of the Royal Society of London B: Biological Sciences 279:2609–2618.

Simmons, K. J., Simmons, K. J., P. N. Nde, P. N. Nde, Y. Y. Kleshchenko, Y. Y. Kleshchenko, M. F. Lima, M. F. Lima, F. Villalta, and F. Villalta. 2006. Stable RNA interference of host thrombospondin-1 blocks *Trypanosoma cruzi* infection. FEBS letters 580:2365–2370.

Sotolongo, J., J. Ruiz, and M. Fukata. 2012. The role of innate immunity in the host

defense against intestinal bacterial pathogens. Current infectious disease reports 14:15–23.

Stanley, G. 2003. The evolution of modern corals and their early history. Earth-Science Reviews 60:195–225.

Starzak, D. E., R. G. Quinnell, M. R. Nitschke, and S. K. Davy. 2014. The influence of symbiont type on photosynthetic carbon flux in a model cnidarian–dinoflagellate symbiosis. Marine Biology:711–724.

Steen, R. G. 1986. Evidence for heterotrophy by zooxanthellae in symbiosis with *Aiptasia pulchella*. The Biological Bulletin:267–278.

Sullivan, J. C., A. M. Reitzel, and J. R. Finnerty. 2006. A high percentage of introns in human genes were present early in animal evolution: evidence from the basal metazoan *Nematostella vectensis.* Genome informatics. International Conference on Genome Informatics 17:219–229.

Tada, H., K. Kuboki, K. Nomura, and T. Inokuchi. 2001. High glucose levels enhance TGF-β1–thrombospondin-1 pathway in cultured human mesangial cells via mechanisms dependent on glucose-induced PKC activation. Journal of Diabetes and its Complications 15:193–197.

Takeda, K., T. Kaisho, and S. Akira. 2003. Toll-like receptors. Annual Review of Immunology 21:335–376.

Takeuchi, O., and S. Akira. 2010. Pattern recognition receptors and inflammation. Cell 140:805–820.

Tan, K., M. Duquette, J.-H. Liu, Y. Dong, R. Zhang, A. Joachimiak, J. Lawler, and J.-H. Wang. 2002. Crystal structure of the TSP-1 type 1 repeats: a novel layered fold and its biological implication. The Journal of Cell Biology 159:373–382.

Tang, B. L., and Tang, B. L. 2001. ADAMTS: a novel family of extracellular matrix proteases. The International Journal of Biochemistry & Cell Biology 33:33–44.

Technau, U., Technau, U., S. Rudd, S. Rudd, P. Maxwell, P. Maxwell, P. GORDON, P. GORDON, M. Saina, M. Saina, L. GRASSO, L. GRASSO, D. HAYWARD, D. HAYWARD, C. SENSEN, C. SENSEN, R. Saint, R. Saint, T. HOLSTEIN, and T. HOLSTEIN. 2005. Maintenance of ancestral complexity and non-metazoan genes in two basal cnidarians. TRENDS in Genetics 21:633–639.

Templeton, T. J., and D. C. Kaslow. 1997. Cloning and cross-species comparison of the thrombospondin-related anonymous protein (TRAP) gene from Plasmodium

knowlesi, Plasmodium vivax and Plasmodium gallinaceum. Molecular and Biochemical Parasitology 84:13–24.

Thelen, T., Y. Hao, A. I. Medeiros, J. L. Curtis, C. H. Serezani, L. Kobzik, L. H. Harris, and D. M. Aronoff. 2010. The class A scavenger receptor, macrophage receptor with collagenous structure, is the major phagocytic receptor for *Clostridium sordellii* expressed by human decidual macrophages. The Journal of Immunology 185:4328–4335.

Tolsma, S. S., O. V. Volpert, D. J. Good, W. A. Frazier, P. J. Polverini, and N. Bouck. 1993. Peptides derived from two separate domains of the matrix protein thrombospondin-1 have anti-angiogenic activity. The Journal of Cell Biology 122:497–511.

Traylor-Knowles, N., B. R. Granger, T. J. Lubinski, J. R. Parikh, S. Garamszegi, Y. Xia, J. A. Marto, L. Kaufman, and J. R. Finnerty. 2011. Production of a reference transcriptome and transcriptomic database (PocilloporaBase) for the cauliflower coral, *Pocillopora damicornis*. BMC Genomics 12:585.

Trench, R. K. 1971. The The Physiology and Biochemistry of Zooxanthellae Symbiotic with Marine Coelenterates. II. Pages 237–250 *in*.

Tucker, R. P. 2004. The thrombospondin type 1 repeat superfamily. The international journal of Biochemistry & Cell Biology 36:969–974.

Uchida, H., Uchida, H., M. Kuroki, M. Kuroki, T. Shitama, T. Shitama, H. Hayashi, H. Hayashi, M. Kuroki, and M. Kuroki. 2008. Activation of TGF-β1 Through Up-Regulation of TSP-1 by Retinoic Acid in Retinal Pigment Epithelial Cells. Current Eye Research 33:199–203.

Van, O. M., A. J. Mahiny, and T. J. Done. 2005. Geographic distribution of zooxanthella types in three coral species on the Great Barrier Reef sampled after the 2002 bleaching event. Coral Reefs 24:482–487.

Vasta, G. R. 2009. Roles of galectins in infection. Nature Reviews Microbiology 7:424–438.

Vaughan, A. M., A. S. I. Aly, and S. H. I. Kappe. 2008. Malaria parasite pre-erythrocytic stage infection: gliding and hiding. Cell Host & Microbe 4:209–218.

Venn, A. A., J. E. Loram, and A. E. Douglas. 2008. Photosynthetic symbioses in animals. Journal of Experimental Botany 59:1069–1080.

Vesteg, M., and J. Krajcovic. 2008. Origin of eukaryotic cells as a symbiosis of parasitic

alpha-proteobacteria in the periplasm of two-membrane-bounded sexual prekaryotes. Communicative & Integrative Biology 1:104–113.

Vidal-Dupiol, J., D. Zoccola, E. Tambutté, C. Grunau, C. Cosseau, K. M. Smith, M. Freitag, N. M. Dheilly, D. Allemand, and S. Tambutté. 2013. Genes Related to Ion-Transport and Energy Production Are Upregulated in Response to CO 2 -Driven pH Decrease in Corals: New Insights from Transcriptome Analysis. PLoS ONE 8:e58652.

Vidal-Dupiol, J., M. Adjeroud, E. Roger, L. Foure, D. Duval, Y. Mone, C. Ferrier-Pages, E. Tambutté, S. Tambutté, D. Zoccola, D. Allemand, and G. Mitta. 2009. Coral bleaching under thermal stress: putative involvement of host/symbiont recognition mechanisms. BMC Physiology 9:14.

Waghabi, M. C., M. Keramidas, J.-J. Feige, T. C. Araujo-Jorge, and S. Bailly. 2005. Activation of transforming growth factor beta by *Trypanosoma cruzi*. Cellular Microbiology 7:511–517.

Wang, J., and A. E. Douglas. 1997. Nutrients, signals, and photosynthate release by symbiotic algae: The impact of taurine on the dinoflagellate alga Symbiodinium from the sea anemone *Aiptasia pulchella*. Plant Physiology 114:631–636.

Wang, S., J. Skorczewski, X. Feng, L. Mei, and J. E. Murphy-Ullrich. 2004. Glucose up-regulates thrombospondin 1 gene transcription and transforming growth factor-beta activity through antagonism of cGMP-dependent protein kinase repression via upstream stimulatory factor 2. The Journal of Biological Chemistry 279:34311–34322.

Weis, V. M., S. K. Davy, O. Hoegh-Guldberg, M. Rodriguez-Lanetty, and J. R. Pringle. 2008. Cell biology in model systems as the key to understanding corals. Trends in Ecology & Evolution 23:369–376.

Weis, V., W. Reynolds, M. deBoer, and D. Krupp. 2001. Host-symbiont specificity during onset of symbiosis between the dinoflagellates *Symbiodinium* spp. and planula larvae of the scleractinian coral *Fungia scutaria*. Coral Reefs 20:301–308.

Whelan, F. J., C. J. Meehan, G. B. Golding, B. J. McConkey, and D. M. E. Bowdish. 2012. The evolution of the class A scavenger receptors. BMC Evolutionary Biology 12:227.

Wilson, E. O., and F. M. Peter. 1988. Biodiversity. National Academies.

Wood-Charlson, E. M., and V. M. Weis. 2009. The diversity of C-type lectins in the genome of a basal metazoan, Nematostella vectensis. Developmental and Comparative Immunology 33:881–889.

Wood-Charlson, E. M., L. L. Hollingsworth, D. A. Krupp, and V. M. Weis. 2006. Lectin/glycan interactions play a role in recognition in a coral/dinoflagellate symbiosis. Cellular Microbiology 8:1985–1993.

Wooldridge, S. A. 2010. Is the coral-algae symbiosis really “mutually beneficial” for the partners? Bioessays 32:615–625.

Yalaoui, S., T. Huby, J.-F. Franetich, A. Gego, A. Rametti, M. Moreau, X. Collet, A. Siau, G.-J. van Gemert, R. W. Sauerwein, A. J. F. Luty, J.-C. Vaillant, L. Hannoun, J. Chapman, D. Mazier, and P. Froissard. 2008. Scavenger receptor BI boosts hepatocyte permissiveness to Plasmodium infection. Cell Host & Microbe 4:283–292.

Yamada, Y., T. Doi, T. Doi, T. Hamakubo, T. Hamakubo, T. Kodama, and T. Kodama. 1998. Scavenger receptor family proteins: roles for atherosclerosis, host defence and disorders of the central nervous system. Cellular and Molecular Life Sciences (CMLS) 54:628–640.

Yazdani, U., and J. R. Terman. 2006. Protein Family Review: The semaphorins. Genome Biology 7:211.

Yee, K. O., C. M. Connolly, M. Duquette, S. Kazerounian, R. Washington, and J. Lawler. 2009. The effect of thrombospondin-1 on breast cancer metastasis. Breast Cancer Research and Treatment 114:85–96.

Yellowlees, D., T. A. V. Rees, and W. Leggat. 2008. Metabolic interactions between algal symbionts and invertebrate hosts. Plant, Cell & Environment 31:679–694.

Yung, S., C. Y. Y. Lee, Q. Zhang, S. K. Lau, R. C. W. Tsang, and T. M. Chan. 2006. Elevated glucose induction of thrombospondin-1 up-regulates fibronectin synthesis in proximal renal tubular epithelial cells through TGF-beta1 dependent and TGF-beta1 independent pathways. Nephrology, dialysis, transplantation : official publication of the European Dialysis and Transplant Association - European Renal Association 21:1504–1513.

Zerial, M., and H. McBride. 2001. Rab proteins as membrane organizers. Nature Reviews Molecular Cell Biology 2:107–117.

Zhang, X., and J. Lawler. 2007. Thrombospondin-based antiangiogenic therapy. Microvascular Research 74:90–99.

# Appendices

## Appendix A1

Table A.1: Resources providing invertebrate SR sequences.

| Phylum | Class | Organism | Resource | Publication |
|---|---|---|---|---|
| Porifera | Homoscleromorpha | *Oscarella carmella* | http://www.compagen.org/index.html | Nichols et al. 2012 |
| | Domospongiae | *Suberites domuncula* | NCBI: http://www.ncbi.nlm.nih.gov/ | Harcet et al. 2010 |
| | Domospongiae | *Amphimedon queenslandica* | http://metazoa.ensembl.org/Amphimedon_queenslandica | Srivastava et al. 2010 |
| Ctenophora | Tentaculata | *Mnemiopsis leidyi* | http://research.nhgri.nih.gov/mnemiopsis/blast/ | Ryan et al. 2013 |
| Cnidaria | Anthozoa | *Nematostella vectensis* | http://genome.jgi-psf.org/Nemve1/Nemve1.home.html | Putnam et al. 2007, Kimura et al. 2009 |
| | | *Anthopleura elegantissima* | http://people.oregonstate.edu/~meyere/data.html | Kitchener et al. in prep |
| | | *Aiptasia pallida* | http://pringlelab.stanford.edu/projects.html | Lehnert et al. 2012 |
| | | *Acropora digitifera* | http://marinegenomics.oist.jp/genomes/gallery | Shinzato et al. 2011 |
| | | *Acropora millepora* | http://www.bio.utexas.edu/research/matz_lab/matzlab/Data.html | Moya et al. 2012 |
| | | *Fungia scutaria* | http://people.oregonstate.edu/~meyere/data.html | Kitchener et al. in prep |
| Placozoa | n/a | *Trichoplax adhaerens* | http://genome.jgi-psf.org/Triad1/Triad1.home.html | Srivastava et al. 2008 |
| Echinodermata | Echinoidea | *Strongylocentrotus purpuratus* | http://www.spbase.org/SpBase/ | Cameron et al. 2009 |
| Nemotoda | Chromadorea | *Caenorhabditis elegans* | http://www.wormbase.org/ | Chen et al. 2005 |
| Arthropoda | Insecta | *Drosophila melanogaster* | NCBI: http://www.ncbi.nlm.nih.gov/ | Smith et al. 2007 |
| Chordata | Acidiacea | *Ciona intestinalis* | http://genome.jgi-psf.org/Cioin2/Cioin2.home.html | Dehal et al. 2002 |
| Chordata | Leptocardii | *Branchiostoma floridae* | http://genome.jgi- psf.org/Brafl1/Brafl1.home.html | Putnam et al. 2008 |
| Chordata | Amphibia | *Xenopus laevis* | NCBI: http://www.ncbi.nlm.nih.gov/ | Klein et al.2002 |
| Chordata | Actinopterygii | *Danio rerio* | NCBI: http://www.ncbi.nlm.nih.gov/ | Strausberg et al. 2002 |
| Chordata | Mammalia | *Mus musculus* | NCBI: http://www.ncbi.nlm.nih.gov/ | Acton et al. 1996 |
| Chordata | Mammalia | *Homo sapiens* | NCBI: http://www.ncbi.nlm.nih.gov/ | Calvo et al. 1993 |

## Appendix A2

Table A.2: SR protein information

| Protein Type | Sequence | Organism | Original identifier/ Accession Number | Domain structure (Pfam) | Length of protein |
|---|---|---|---|---|---|
| Class A SR's | A.digitifera SRA-like1 | *Acropora digitifera* | aug_v2a.13777 | SIG PEP-TM-CollagenX2-CUB-SRCR | 555 |
| | A.digitifera SRA-like2 | *Acropora digitifera* | aug_v2a.04186 | SIG PEP-TM-Collagen-ISETX2-EGFCAX2-CUB-SRCR | 1690 |
| Class B SR's | N.vectensis SRB1 | *Nematostella vectensis* | jgi\|Nemve1\|12228\|gw.216.12.1 | CD36 | 379 |
| | N.vectensis SRB2 | *Nematostella vectensis* | jgi\|Nemve1\|93030\|e_gw.32.86.1 | CD36 | 397 |
| | N.vectensis SRB3 | *Nematostella vectensis* | jgi\|Nemve1\|115992\|e_gw.135.24.1 | CD36 | 394 |
| | A.elegantissima SRB1 | *Anthopleura elegantissima* | comp27898 | TM-CD36-TM | 408 |
| | A.elegantissima SRB2 | *Anthopleura elegantissima* | comp105988 | CD36-TM | 247 |
| | Aiptasia SRB1 | *Aiptasia* sp. | AIPGENE27709, COMP18601 | TM-CD36-TM | 517 |
| | Aiptasia SRB2 | *Aiptasia* sp. | AIPGENE13459, COMP 9241 | TM-CD36-TM | 518 |
| | Aiptasia SRB3 | *Aiptasia* sp. | AIPGENE25617, COMP18959 | CD36-TM | 495 |
| | Aiptasia SRB4 | *Aiptasia* sp. | AIPGENE9495, COMP18212 | TM-CD36 | 561 |
| | A.digitifera SRB1 | *Acropora digitifera* | aug_v2a.07404 | TM-CD36-TM | 553 |
| | A.digitifera SRB2 | *Acropora digitifera* | aug_v2a.02372 | CD36-TM | 254 |
| | A.digitifera SRB3 | *Acropora digitifera* | aug_v2a.04988 | CD36-TM | 237 |
| | A.digitifera SRB4 | *Acropora digitifera* | aug_v2a.16675 | CD36-TM | 323 |
| | A.millepora SRB1 | *Acropora millepora* | c004851 | CD36-TM | 424 |
| | A.millepora SRB2 | *Acropora millepora* | c007468 | CD36-TM | 524 |
| | A.millepora SRB3 | *Acropora millepora* | c011462 | CD36-TM | 281 |
| | F.scutaria SRB1 | *Fungia scutaria* | comp23243 | TM-CD36-TM | 520 |

| Protein Type | Sequence | Organism | Original identifier/ Accession Number | Domain structure (Pfam) | Length of protein |
|---|---|---|---|---|---|
| | F.scutaria SRB2 | *Fungia scutaria* | comp8783 | CD36-TM | 472 |
| | F.scutaria SRB3 | *Fungia scutaria* | comp207143 | TM-CD36 | 238 |
| Class E SR's | N.vectensis LOX1-like1 | *Nematostella vectensis* | jgi\|Nemve1\|239025\|estExt_fgenesh1_pg.C_110046 | CTLD | 115 |
| | N.vectensis LOX1-like2 | *Nematostella vectensis* | jgi\|Nemve1\|210164\|fgenesh1_pg.scaffold_117000056 | CTLD | 281 |
| | N.vectensis LOX1-like3 | *Nematostella vectensis* | jgi\|Nemve1\|154303\|e_gw.7191.2.1 | CTLD | 137 |
| | N.vectensis LOX1-like4 | *Nematostella vectensis* | jgi\|Nemve1\|138199\|e_gw.425.7.1 | CTLD | 92 |
| | N.vectensis LOX1-like5 | *Nematostella vectensis* | jgi\|Nemve1\|118816\|e_gw.156.96.1 | CTLD | 80 |
| | N.vectensis LOX1-like6 | *Nematostella vectensis* | jgi\|Nemve1\|103399\|e_gw.68.85.1 | CTLD | 129 |
| | N.vectensis LOX1-like7 | *Nematostella vectensis* | jgi\|Nemve1\|102786\|e_gw.66.179.1 | CTLD | 138 |
| | N.vectensis LOX1-like8 | *Nematostella vectensis* | jgi\|Nemve1\|88150\|e_gw.17.229.1 | CTLD | 125 |
| | N.vectensis LOX1-like9 | *Nematostella vectensis* | jgi\|Nemve1\|87968\|e_gw.17.235.1 | CTLD | 121 |
| | N.vectensis LOX1-like10 | *Nematostella vectensis* | jgi\|Nemve1\|83478\|e_gw.7.7.1 | CTLD | 121 |
| | N.vectensis LOX1-like11 | *Nematostella vectensis* | jgi\|Nemve1\|83039\|e_gw.7.243.1 | CTLD | 146 |
| | N.vectensis LOX1-like12 | *Nematostella vectensis* | jgi\|Nemve1\|66344\|gw.3.677.1 | CTLD | 99 |
| | A.elegantissima LOX1-like1 | *Anthopleura elegantissima* | comp29206 | TM-CTLD | 177 |
| | A.elegantissima LOX1-like2 | *Anthopleura elegantissima* | comp206797 | SIG-TM-CTLD | 241 |
| | A.elegantissima LOX1-like3 | *Anthopleura elegantissima* | comp147141 | SIG-TM-CTLD | 115 |
| | A.elegantissima LOX1-like4 | *Anthopleura elegantissima* | comp991 | SIG-TM-CTLD | 179 |
| | A.elegantissima LOX1-like5 | *Anthopleura elegantissima* | comp165958 | SIG-CTLD | 267 |
| | A.elegantissima LOX1- | *Anthopleura elegantissima* | comp160005 | CTLD | 115 |

| Protein Type | Sequence | Organism | Original identifier/ Accession Number | Domain structure (Pfam) | Length of protein |
|---|---|---|---|---|---|
| | like6 | | | | |
| | A.elegantissima LOX1-like7 | *Anthopleura elegantissima* | comp16589 | CTLD | 140 |
| | A.elegantissima LOX1-like8 | *Anthopleura elegantissima* | comp12131 | CTLD | 330 |
| | A.elegantissima LOX1-like9 | *Anthopleura elegantissima* | comp8014 | CTLD | 252 |
| | Aiptasia LOX1-like1 | *Aiptasia sp.* | AIPGENE25771 | TM-CTLD-TM | 378 |
| | Aiptasia LOX1-like2 | *Aiptasia sp.* | AIPGENE25770 | TM-CTLD-TM | 509 |
| | Aiptasia LOX1-like3 | *Aiptasia sp.* | AIPGENE25914 | SIG-TMX2-CTLD-TM | 524 |
| | Aiptasia LOX1-like4 | *Aiptasia sp.* | AIPGENE25775 | SIG-CTLD | 404 |
| | Aiptasia LOX1-like5 | *Aiptasia sp.* | AIPGENE24234 | SIG-CTLD | 748 |
| | Aiptasia LOX1-like6 | *Aiptasia sp.* | AIPGENE22689 | SIG-CTLD | 197 |
| | Aiptasia LOX1-like7 | *Aiptasia sp.* | AIPGENE17171 | SIG-CTLD | 175 |
| | Aiptasia LOX1-like8 | *Aiptasia sp.* | AIPGENE17156 | SIG-CTLD | 193 |
| | Aiptasia LOX1-like9 | *Aiptasia sp.* | AIPGENE17147 | SIG-CTLD | 120 |
| | Aiptasia LOX1-like10 | *Aiptasia sp.* | AIPGENE4175 | SIG-CTLD | 302 |
| | Aiptasia LOX1-like11 | *Aiptasia sp.* | AIPGENE2158 | SIG-CTLD | 385 |
| | Aiptasia LOX1-like12 | *Aiptasia sp.* | AIPGENE26260 | CTLD-TM | 379 |
| | Aiptasia LOX1-like13 | *Aiptasia sp.* | AIPGENE23661 | CTLD-TM | 589 |
| | Aiptasia LOX1-like14 | *Aiptasia sp.* | AIPGENE8793 | CTLD | 143 |
| | Aiptasia LOX1-like15 | *Aiptasia sp.* | AIPGENE25745 | CTLD | 300 |
| | Aiptasia LOX1-like16 | *Aiptasia sp.* | AIPGENE22161 | CTLD | 127 |
| | Aiptasia LOX1-like17 | *Aiptasia sp.* | AIPGENE20196 | CTLD | 132 |
| | Aiptasia LOX1-like18 | *Aiptasia sp.* | AIPGENE18051 | CTLD | 188 |
| | Aiptasia LOX1-like19 | *Aiptasia sp.* | AIPGENE18018 | CTLD | 224 |

| Protein Type | Sequence | Organism | Original identifier/ Accession Number | Domain structure (Pfam) | Length of protein |
|---|---|---|---|---|---|
| | Aiptasia LOX1-like20 | *Aiptasia sp.* | AIPGENE18000 | CTLD | 241 |
| | Aiptasia LOX1-like21 | *Aiptasia sp.* | AIPGENE17873 | CTLD | 214 |
| | Aiptasia LOX1-like22 | *Aiptasia sp.* | AIPGENE17207 | CTLD | 297 |
| | Aiptasia LOX1-like23 | *Aiptasia sp.* | AIPGENE17153 | CTLD | 173 |
| | Aiptasia LOX1-like24 | *Aiptasia sp.* | AIPGENE16190 | CTLD | 131 |
| | Aiptasia LOX1-like25 | *Aiptasia sp.* | AIPGENE15934 | CTLD | 112 |
| | Aiptasia LOX1-like26 | *Aiptasia sp.* | AIPGENE13331 | CTLD | 250 |
| | Aiptasia LOX1-like27 | *Aiptasia sp.* | AIPGENE11521 | CTLD | 668 |
| | Aiptasia LOX1-like28 | *Aiptasia sp.* | AIPGENE10805 | CTLD | 137 |
| | Aiptasia LOX1-like29 | *Aiptasia sp.* | AIPGENE9848 | CTLD | 209 |
| | Aiptasia LOX1-like30 | *Aiptasia sp.* | AIPGENE8532 | CTLD | 206 |
| | Aiptasia LOX1-like31 | *Aiptasia sp.* | AIPGENE2088 | CTLD | 364 |
| | Aiptasia LOX1-like32 | *Aiptasia sp.* | AIPGENE6834 | CTLD | 243 |
| | A.digitifera LOX1-like1 | *Acropora digitifera* | aug_v2a.15486 | TM-CTLD-TM | 564 |
| | A.digitifera LOX1-like2 | *Acropora digitifera* | aug_v2a.03583 | SIG-TM-CTLD | 188 |
| | A.digitifera LOX1-like3 | *Acropora digitifera* | aug_v2a.09977 | SIG-CTLD | 222 |
| | A.digitifera LOX1-like4 | *Acropora digitifera* | aug_v2a.20063 | CTLD-TM | 446 |
| | A.digitifera LOX1-like5 | *Acropora digitifera* | aug_v2a.12238 | CTLD-TM | 425 |
| | A.digitifera LOX1-like6 | *Acropora digitifera* | aug_v2a.06178 | CTLD-TM | 633 |
| | A.digitifera LOX1-like7 | *Acropora digitifera* | aug_v2a.23858 | CTLD | 140 |
| | A.digitifera LOX1-like8 | *Acropora digitifera* | aug_v2a.12298 | CTLD | 329 |
| | A.digitifera LOX1-like9 | *Acropora digitifera* | aug_v2a.10868 | CTLD | 121 |
| | A.digitifera LOX1-like10 | *Acropora digitifera* | aug_v2a.07446 | CTLD | 130 |
| | A.digitifera LOX1-like11 | *Acropora digitifera* | aug_v2a.06416 | CTLD | 180 |
| | A.millepora LOX1-like1 | *Acropora millepora* | c030119 | TM-CTLD | 220 |
| | A.millepora LOX1-like2 | *Acropora millepora* | c008721 | SIG-TM-CTLD | 639 |

| Protein Type | Sequence | Organism | Original identifier/ Accession Number | Domain structure (Pfam) | Length of protein |
|---|---|---|---|---|---|
| | A.millepora LOX1-like3 | *Acropora millepora* | c026988 | CTLD | 129 |
| | A.millepora LOX1-like4 | *Acropora millepora* | c024702 | CTLD | 156 |
| | A.millepora LOX1-like5 | *Acropora millepora* | c022421 | CTLD | 197 |
| | F.scutaria LOX1-like1 | *Fungia scutaria* | comp33746 | SIG-TM-CTLD | 578 |
| | F.scutaria LOX1-like2 | *Fungia scutaria* | comp30359 | SIG-CTLD | 167 |
| | F.scutaria LOX1-like3 | *Fungia scutaria* | comp22746 | SIG-CTLD | 633 |
| | F.scutaria LOX1-like4 | *Fungia scutaria* | comp246601 | CTLD | 137 |
| | F.scutaria LOX1-like5 | *Fungia scutaria* | comp111893 | CTLD | 169 |
| | F.scutaria LOX1-like6 | *Fungia scutaria* | comp96085 | CTLD | 161 |
| | F.scutaria LOX1-like7 | *Fungia scutaria* | comp42537 | CTLD | 72 |
| | F.scutaria LOX1-like8 | *Fungia scutaria* | comp30359 | CTLD | 152 |
| | F.scutaria LOX1-like8 | *Fungia scutaria* | comp11329 | CTLD | 174 |
| | F.scutaria LOX1-like10 | *Fungia scutaria* | comp280 | CTLD | 355 |
| CUB domain | N.vectensis CUB | *Nematostella vectensis* | jgi\|Nemve1\|242231\|estExt fgenesh1 pg.C 590046 | MAM SRCR MAM LDL CUB X3 | 1310 |
| | A.elegantissima CUB | *Anthopleura elegantissima* | comp13519 | SRCR FN3 SRCR FN3 FN3 CUB | 1997 |
| | Aiptasia CUB1 | *Aiptasia sp.* | AIPGENE22865 | SRCR FN3X2 SRCR FN3X2 CUB VWB EGFX3 | 3718 |
| | Aiptasia CUB2 | *Aiptasia sp.* | AIPGENE22864 | SRCR fn3 SRCR FN3X2 CUB VWD EGFX3 | 4280 |
| | Aiptasia CUB3 | *Aiptasia sp.* | AIPGENE2358 | CUB SRCRX2 | 345 |
| | A.digitifera CUB1 | *Acropora digitifera* | aug_v2a.02958.t1 | EGF CUB SRCR 2 | 375 |
| | A.digitifera CUB2 | *Acropora digitifera* | aug_v2a.03894.t2 | MAM LDL MAM SRCR2 CUB TM UBOX | 1126 |
| | A.digitifera CUB3 | *Acropora digitifera* | aug_v2a.10056.t1 | PAN1 EGFCA CUB SRCR | 453 |
| | A.digitifera CUB4 | *Acropora digitifera* | aug_v2a.13776.t1 | SIG PEP EGF CA CUB SRCR | 506 |

| Protein Type | Sequence | Organism | Original identifier/ Accession Number | Domain structure (Pfam) | Length of protein |
|---|---|---|---|---|---|
| | A.digitifera CUB5 | *Acropora digitifera* | aug_v2a.18885.t1 | SRCR FN3 CUB EGF | 3001 |
| | A.digitifera CUB6 | *Acropora digitifera* | aug_v2a.19537.t1 | EGF CA CUB SRCR2 | 382 |
| | A.digitifera CUB7 | *Acropora digitifera* | aug_v2a.23712.t1 | LECTIN C CUB SRCR2 | 1087 |
| | A.digitifera CUB8 | *Acropora digitifera* | aug_v2a.24394.t1 | EGF 2 CUB 3 SRCR | 376 |
| | A.digitifera CUB9 | *Acropora digitifera* | aug_v2a.18043.t1 | CUB SRCR2 | 305 |
| | A.millepora CUB1 | *Acropora millepora* | c000424 | SIG TM MAM SRCR MAM LDL MAM SRCRX3 WSC CUBX2 UBOX | 1948 |
| | A.millepora CUB2 | *Acropora millepora* | c009570 | EGF-CAX3 CUB SRCRX2 | 558 |
| | F.scutaria CUB1 | *Fungia scutaria* | comp2166 | SRCRX2 WSC CUB | 341 |
| | F.scutaria CUB2 | *Fungia scutaria* | comp2699 | SRCR FN3 CUB FN3 | 767 |
| Trypsin domain | N.vectensis TRY1 | *Nematostella vectensis* | jgi\|Nemve1\|199744\|fgenesh1 pg.scaffold 17000033 | SRCRX3 EGF SRCRX6 TRYPSIN TM | 1823 |
| | A.elegantissima TRY1 | *Anthopleura elegantissima* | comp713 | TRYPSIN SRCR | 621 |
| | A.elegantissima TRY2 | *Anthopleura elegantissima* | comp8236 c0 seq2 translation | SRCRX2 IG X6 TRYPSIN | 1096 |
| | Aiptasia TRY1 | *Aiptasia sp.* | AIPGENE6890 | SRCRX3 TRYPSIN | 530 |
| | Aiptasia TRY2 | *Aiptasia sp.* | AIPGENE15640 | SRCRX2 MAM ISET IG2 IG3 IG2X2 IG3 TRYPSIN | 1337 |
| | A.digitifera TRY1 | *Acropora digitifera* | aug_v2a.02016.t1 | SRCRX2 MAMX2 IG2X2 IG3X2 TRYPSIN | 1216 |
| | A.millepora TRY1 | *Acropora millepora* | c000899 | sig SRCRX4 MAM IGX7 TRYPSIN | 1647 |
| Class I SRs | N.vectensis SRCR1 | *Nematostella vectensis* | jgi\|Nemve1\|118215\|e_gw.152.94.1 | SRCRX2 | 242 |
| | N.vectensis SRCR2 | *Nematostella vectensis* | jgi\|Nemve1\|130928\|e gw.287.21.1 | SRCRX3 | 327 |
| | N.vectensis SRCR3 | *Nematostella vectensis* | jgi\|Nemve1\|212232\|fgenesh1_pg.scaffold_1 52000003 | SRCRX3 | 1282 |
| | N.vectensis SRCR4 | *Nematostella vectensis* | jgi\|Nemve1\|217833\|fgenesh1 pg.scaffold 314000022 | SRCRX3 | 1003 |

| Protein Type | Sequence | Organism | Original identifier/ Accession Number | Domain structure (Pfam) | Length of protein |
|---|---|---|---|---|---|
| | N.vectensis SRCR5 | *Nematostella vectensis* | jgi\|Nemve1\|224889\|fgenesh1 pg.scaffold 6484000001 | SRCRX3 | 328 |
| | N.vectensis SRCR6 | *Nematostella vectensis* | jgi\|Nemve1\|230058\|fgenesh1 pm.scaffold 134000005 | SRCRX3 | 322 |
| | N.vectensis SRCR7 | *Nematostella vectensis* | jgi\|Nemve1\|141947\|e gw.622.10.1 | SRCRX5 | 409 |
| | N.vectensis SRCR8 | *Nematostella vectensis* | jgi\|Nemve1\|208577\|fgenesh1 pg.scaffold 93000088 | SRCRX6 | 681 |
| | N.vectensis SRCR9 | *Nematostella vectensis* | jgi\|Nemve1\|133966\|e gw.325.30.1 | SRCRX9 | 951 |
| | A.elegantissima SRCR1 | *Anthopleura elegantissima* | comp8605 | SRCRX2 | 431 |
| | A.elegantissima SRCR2 | *Anthopleura elegantissima* | comp29440 | SRCRX2 | 222 |
| | A.elegantissima SRCR3 | *Anthopleura elegantissima* | comp754 | SRCRX11 | 1307 |
| | A.elegantissima SRCR4 | *Anthopleura elegantissima* | comp11108 | SRCRX5 | 801 |
| | A.elegantissima SRCR5 | *Anthopleura elegantissima* | comp30702 | SRCRX3 | 324 |
| | A.elegantissima SRCR6 | *Anthopleura elegantissima* | comp43117 | SRCRX3 | 477 |
| | A.elegantissima SRCR7 | *Anthopleura elegantissima* | comp885 | SRCRX3 | 315 |
| | Aiptasia SRCR1 | *Aiptasia sp.* | AIPGENE2115 | SRCR | 662 |
| | Aiptasia SRCR2 | *Aiptasia sp.* | AIPGENE5740 | SRCR | 142 |
| | Aiptasia SRCR3 | *Aiptasia sp.* | AIPGENE13496 | SRCR | 180 |
| | Aiptasia SRCR4 | *Aiptasia sp.* | AIPGENE14367 | SRCR | 162 |
| | Aiptasia SRCR5 | *Aiptasia sp.* | AIPGENE17877 | SRCR | 463 |
| | Aiptasia SRCR6 | *Aiptasia sp.* | AIPGENE19256 | SRCR | 166 |
| | Aiptasia SRCR7 | *Aiptasia sp.* | AIPGENE20918 | SRCR | 117 |
| | Aiptasia SRCR8 | *Aiptasia sp.* | AIPGENE10213 | SRCRX2 | 230 |
| | Aiptasia SRCR9 | *Aiptasia sp.* | AIPGENE14004 | SRCRX2 | 276 |
| | Aiptasia SRCR10 | *Aiptasia sp.* | AIPGENE15247 | SRCRX2 | 260 |
| | Aiptasia SRCR11 | *Aiptasia sp.* | AIPGENE28925 | SRCRX2 | 578 |

| Protein Type | Sequence | Organism | Original identifier/ Accession Number | Domain structure (Pfam) | Length of protein |
|---|---|---|---|---|---|
| | Aiptasia SRCR12 | *Aiptasia sp.* | AIPGENE5733 | SRCRX2 | 530 |
| | Aiptasia SRCR13 | *Aiptasia sp.* | AIPGENE7410 | SRCRX2 | 261 |
| | Aiptasia SRCR14 | *Aiptasia sp.* | AIPGENE15393 | SRCRX2 | 264 |
| | Aiptasia SRCR15 | *Aiptasia sp.* | AIPGENE19188 | SRCRX2 | 268 |
| | Aiptasia SRCR16 | *Aiptasia sp.* | AIPGENE19250 | SRCRX2 | 236 |
| | Aiptasia SRCR17 | *Aiptasia sp.* | AIPGENE19488 | SRCRX2 | 509 |
| | Aiptasia SRCR18 | *Aiptasia sp.* | AIPGENE28950 | SRCRX2 | 264 |
| | Aiptasia SRCR19 | *Aiptasia sp.* | AIPGENE18935 | SRCRX3 | 807 |
| | Aiptasia SRCR20 | *Aiptasia sp.* | AIPGENE10883 | SRCRX3 | 544 |
| | Aiptasia SRCR21 | *Aiptasia sp.* | AIPGENE18484 | SRCRX3 | 828 |
| | Aiptasia SRCR22 | *Aiptasia sp.* | AIPGENE28939 | SRCRX3 | 567 |
| | Aiptasia SRCR23 | *Aiptasia sp.* | AIPGENE20912 | SRCRX4 | 475 |
| | Aiptasia SRCR24 | *Aiptasia sp.* | AIPGENE28215 | SRCRX4 | 383 |
| | Aiptasia SRCR25 | *Aiptasia sp.* | AIPGENE7405 | SRCRX6 | 998 |
| | Aiptasia SRCR26 | *Aiptasia sp.* | AIPGENE13495 | SRCRX6 | 956 |
| | Aiptasia SRCR27 | *Aiptasia sp.* | AIPGENE13497 | SRCRX8 | 1269 |
| | Aiptasia SRCR28 | *Aiptasia sp.* | AIPGENE13500 | SRCRX12 | 1778 |
| | Aiptasia SRCR29 | *Aiptasia sp.* | AIPGENE4659 | SRCRX23 | 3592 |
| | A.digitifera SRCR1 | *Acropora digitifera* | aug_v2a.00627.t1 | SRCR | 160 |
| | *A.digitifera SRCR2* | *Acropora digitifera* | aug_v2a.06512.t1 | SRCR | 194 |
| | *A.digitifera SRCR3* | *Acropora digitifera* | aug_v2a.08252.t1 | SRCR | 112 |
| | *A.digitifera SRCR4* | *Acropora digitifera* | aug_v2a.12043.t1 | SRCR | 676 |
| | *A.digitifera SRCR5* | *Acropora digitifera* | adi_v1.05932 | SRCRX2 | 266 |
| | *A.digitifera SRCR6* | *Acropora digitifera* | aug_v2a.21107.t1 | SRCRX2 | 218 |
| | *A.digitifera SRCR7* | *Acropora digitifera* | aug_v2a.05932.t1 | SRCRX2 | 266 |
| | *A.digitifera SRCR8* | *Acropora digitifera* | aug_v2a.24114.t1 | SRCRX2 | 218 |

| Protein Type | Sequence | Organism | Original identifier/ Accession Number | Domain structure (Pfam) | Length of protein |
|---|---|---|---|---|---|
| | *A.digitifera SRCR9* | *Acropora digitifera* | adi_v1.02987 | SRCRX4 | 806 |
| | *A.digitifera SRCR10* | *Acropora digitifera* | adi_v1.20497 | SRCRX6 | 1014 |
| | *A.digitifera SRCR11* | *Acropora digitifera* | adi_v1.18340 | SRCRX8 | 1265 |
| | A.millepora SRCR1 | *Acropora millepora* | c001468 | SRCRX10 | 1120 |
| | F.scutaria SRCR1 | *Fungia scutaria* | comp217598 | SRCRX2 | 232 |

## Appendix A3

Table A.3: Accession numbers for CD36-domain-containing proteins used for phylogenetic analysis in Figure 2.3

| Phylum | Organism | Sequence identifier | Original identifier/ Accession Number |
|---|---|---|---|
| Porifera | *Oscarella carmela* | O. carmela_1 | comp37448 |
| | *Suberites domuncula* | S. domuncula | gi\|46981146 |
| | *Amphimedon queenslandica* | A.queenslandica | gi\|340380125 |
| Ctenophora | *Mnemiopsis leidyi* | M. leidyi | ML01096 |
| Cnidaria | *Nematostella vectensis* | N. vectensis_1 | jgi\|Nemve1\|12228\|gw.216.12.1 |
| | *Nematostella vectensis* | N. vectensis_2 | jgi\|Nemve1\|93030\|e_gw. 32.86.1 |
| | *Nematostella vectensis* | N. vectensis_3 | jgi\|Nemve1\|115992\|e_gw. 135.24.1 |
| | *Anthopleura ellegantissima* | A. elegantissima | comp27898 |
| | *Aiptasia pallida* | A. pallida_1 | comp18601 |
| | *Aiptasia pallida* | A. pallida_2 | comp9241 |
| | *Aiptasia pallida* | A. pallida_3 | comp18959 |
| | *Acropora digitifera* | A. digitifera | adi_v1.07404 |
| | *Acropora millepora* | A. millepora_1 | c004851 |
| | *Acropora millepora* | A. millepora_2 | c007468 |
| | *Fungia scutaria* | F. scutaria | comp23243 |
| | *Fungia scutaria* | F. scutaria_2 | comp8783 |
| Placozoa | *Tricoplax adherans* | T. adherans | XP_002112871.1 |
| Nematoda | *Caenorhabditis elegans* | C. elegans | gi\|25005153\|emb\|CAB11566.3 |
| Echinodermata | *Strongolocentrotus pupuratus* | S. pupuratus_1 | XP_003729136.1 |
| | *Strongolocentrotus pupuratus* | S. pupuratus_2 | XP_003723576.1 |
| Arthropoda | *Drosophila melanogaster* | D. melanogaster_peste | NP_609168.1 |
| | *Drosophila melanogaster* | D. melanogaster_croque | NP_787957.1 |
| | *Drosophila melanogaster* | D. melanogaster_EMP | NP_523859.2 |
| Chordata | *Ciona intestinalis* | C. intestinalis_1 | XP_002127015.1 |
| | *Ciona intestinalis* | C. intestinalis_2 | XP_002123211.2 |
| | *Branchiostoma floridae* | B. floridae | XP_002229898.1 |
| | *Xenopus laevis* | X. laevis | NP_001080106.1 |
| | *Danio rerio* | D. rerio_SRB1 | AAH44516.1 |
| | *Danio rerio* | D. rerio_CD36 | AAH76048.1 |
| | *Danio rerio* | D. rerio_LMP2 | NP_775366.1 |
| | *Mus musculus* | M. musculus_SRB1 | Q61009 |
| | *Mus musculus* | M. musculus_CD36 | Q08857.2 |
| | *Mus musculus* | M. musculus_LMP2 | NP_031670 |
| | *Homo sapiens* | H. sapiens_SRB1 | Q8WTV0.1 |
| | *Homo sapiens* | H. sapiens_CD36 | NP_001001547 |
| | *Homo sapiens* | H. sapiens_LMP2 | NP_005497 |

## Appendix A4

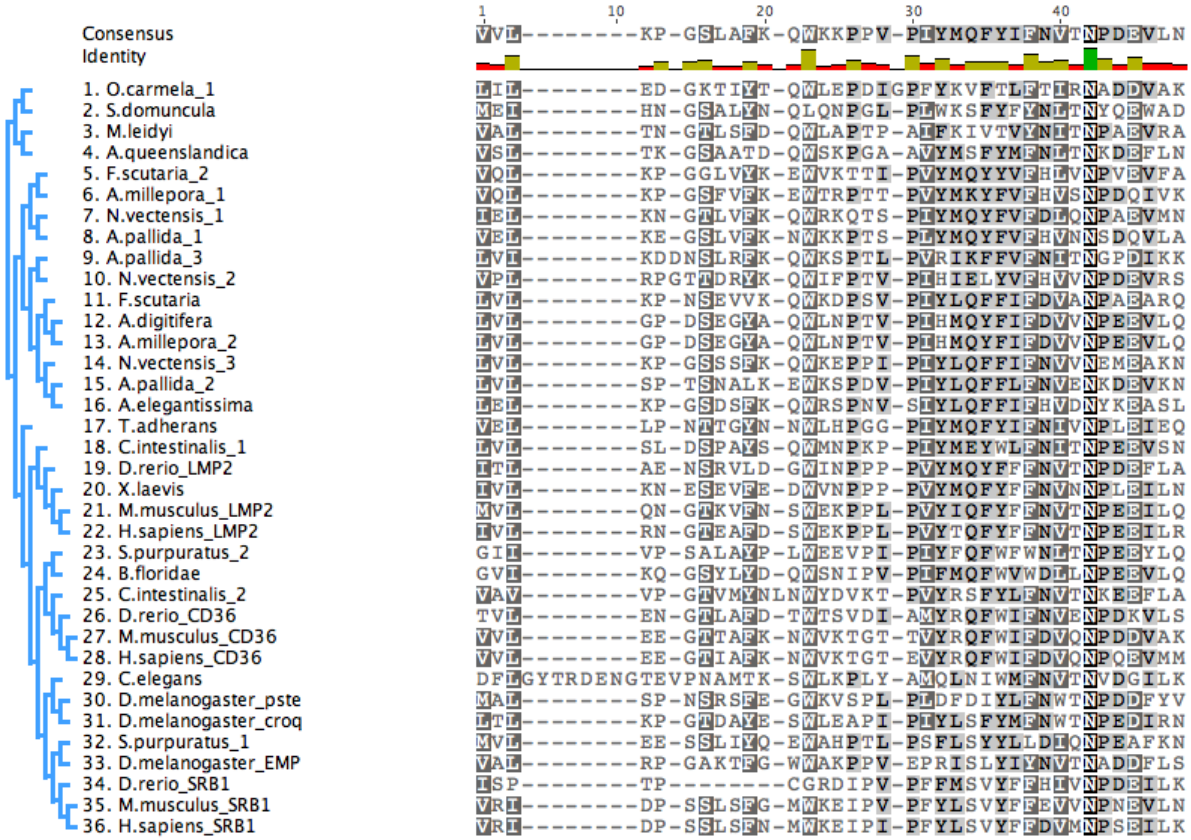



Figure A.1: Protein alignment of SR class B proteins used to produce the ML protein tree in Figure 2.3. Alignment shows similarity according to Blossum matrix 80, amino acids shaded according to similarity (black = fully conserved and white = variable).

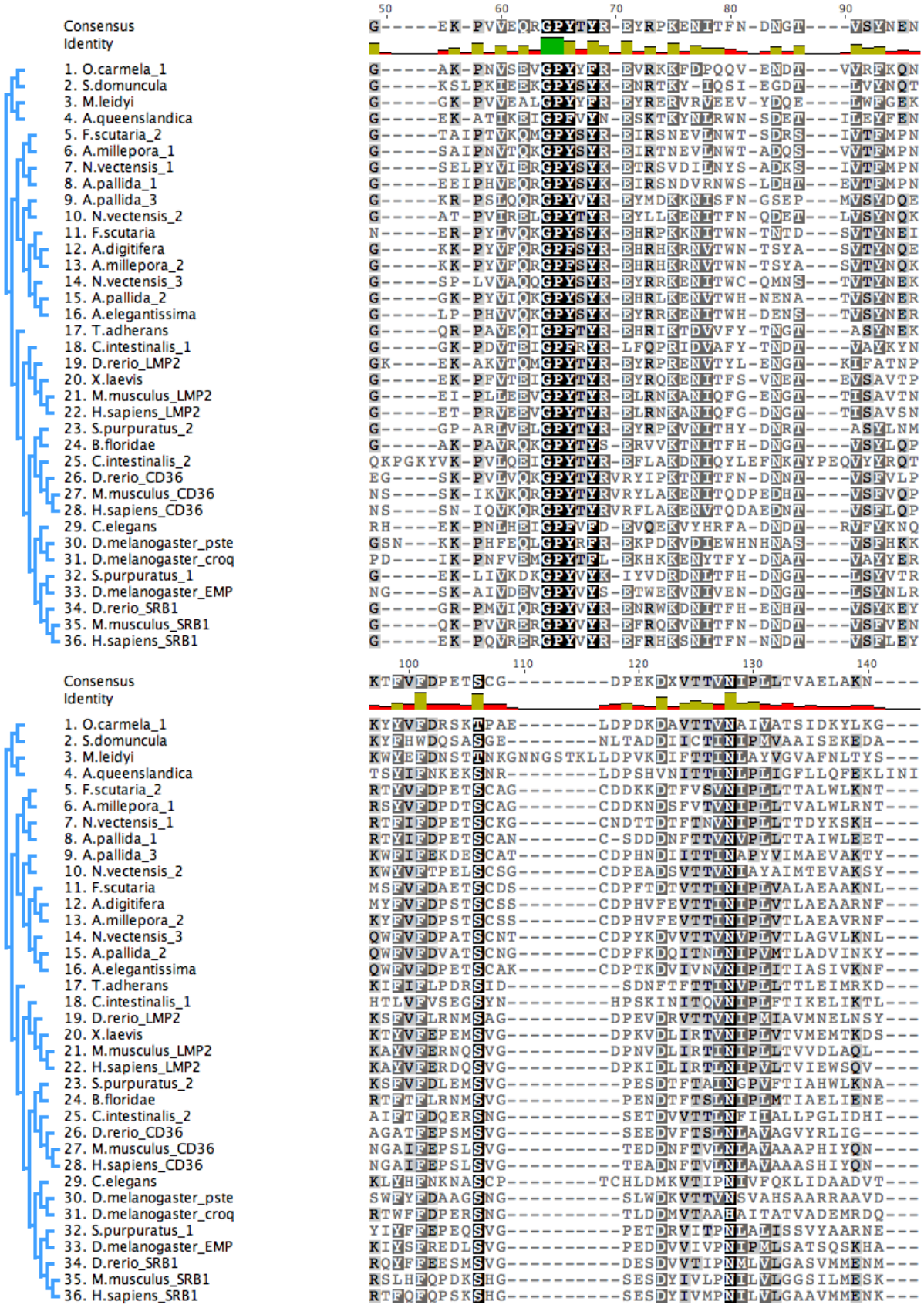



Figure A.1 continued

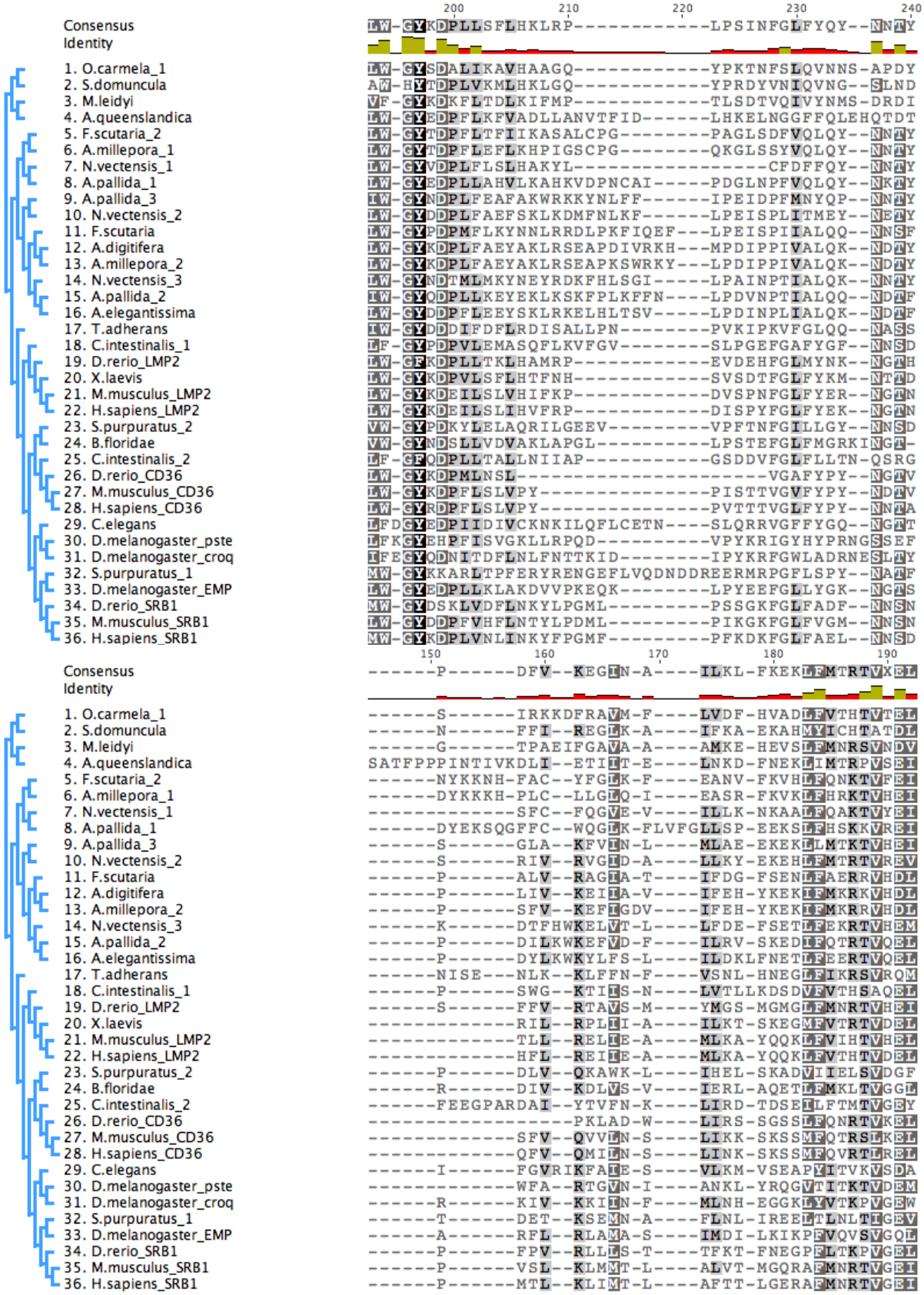



Figure A.1 continued

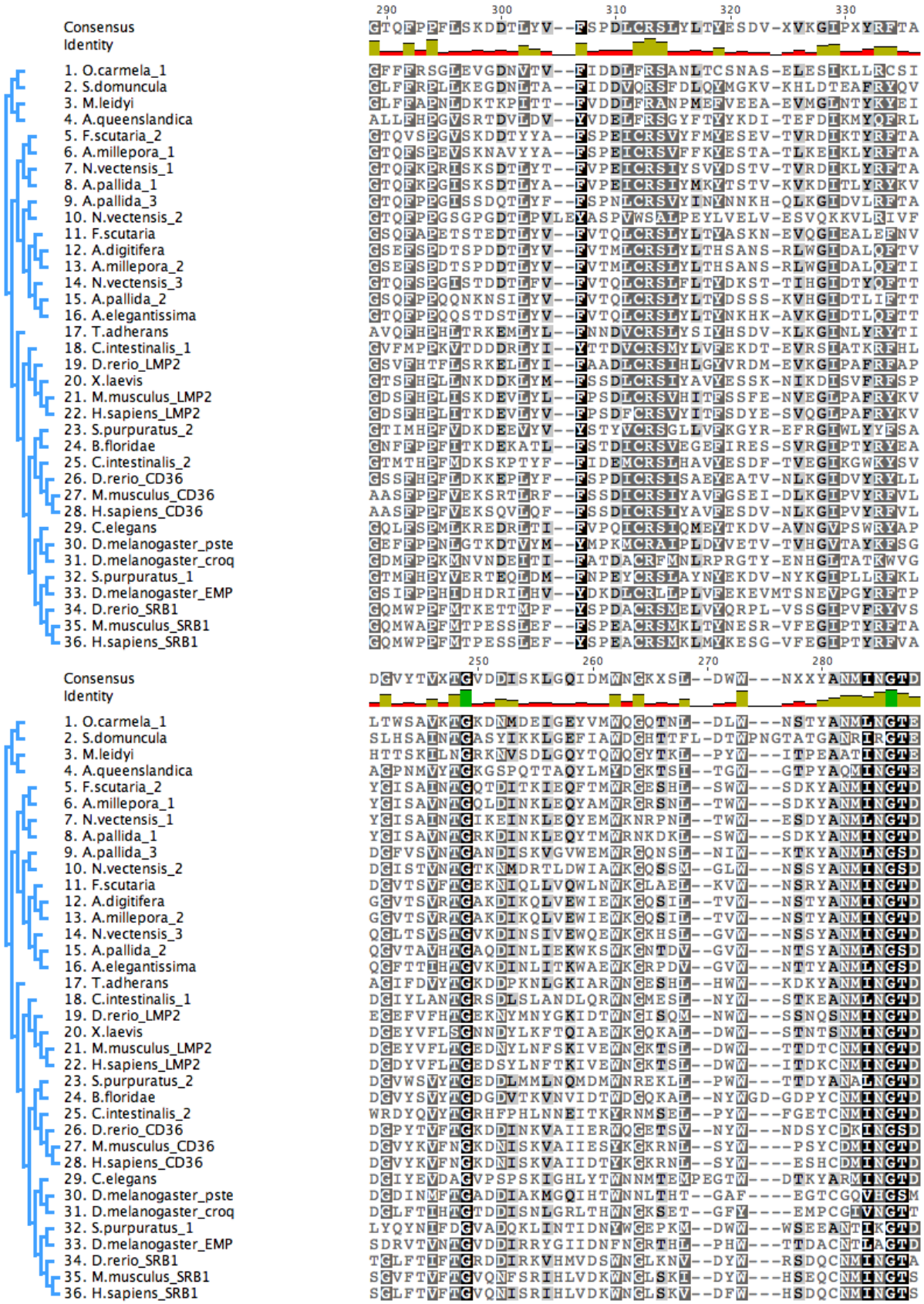



Figure A.1 continued

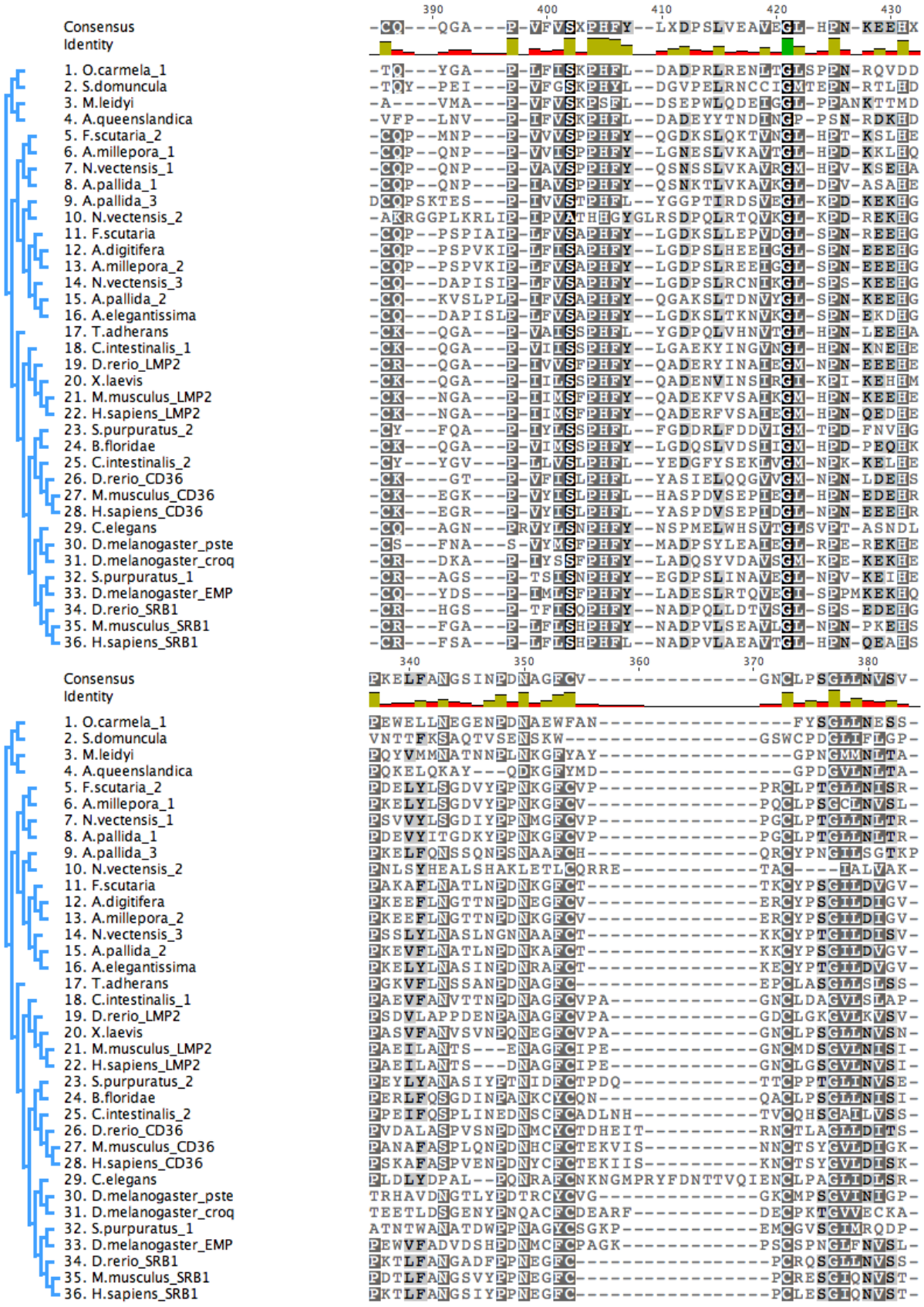



Figure A.1 continued

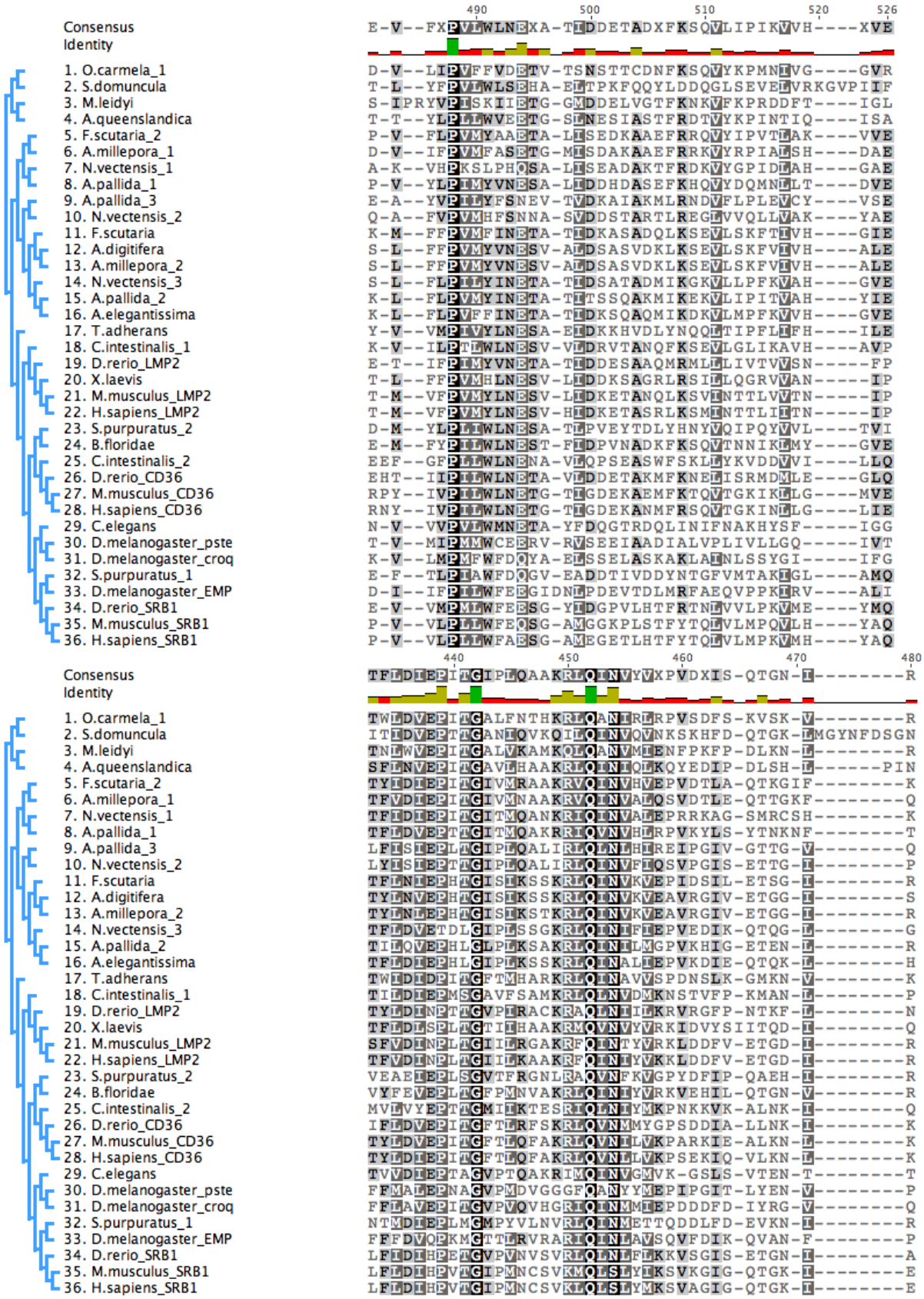



Figure A.1 continued

## Appendix B1

Table B.1: Resources from which invertebrate TSR-domain-containing sequences were obtained

| Phylum | Class | Organism | Resource | Publication |
|---|---|---|---|---|
| Cnidaria | Anthozoa | Nematostella vectensis | http://genome.jgi-psf.org/Nemve1/Nemve1.home.html | Putnam et al. 2007, Kimura et al. 2009 |
| | | Anthopleura elegantissima | http://people.oregonstate.edu/~meyere/data.html | Kitchener et al. in prep |
| | | Aiptasia pallida | http://pringlelab.stanford.edu/projects.html | Lehnert et al. 2012 |
| | | Acropora digitifera | http://marinegenomics.oist.jp/genomes/gallery | Shinzato et al. 2011 |
| | | Acropora millepora | http://www.bio.utexas.edu/research/matz_lab/matzlab/Data.html | Moya et al. 2012 |
| | | Fungia scutaria | http://people.oregonstate.edu/~meyere/data.html | Kitchener et al. in prep |
| Dinoflagellata | Dinophyceae | Symbiodinium minutum | http://marinegenomics.oist.jp/genomes/gallery | Shoguchi et al. 2013 |
| Apicomplexa | Aconoidasida | Plasmodium falciparum | NCBI: http://www.ncbi.nlm.nih.gov/ | Morahan et al.2008 |
| | Conoidasida | Toxoplasma gondii | | |
| | | Neospora caninum | | |
| | | Eimeria tanella | | |
| | | Eimeria maxima | | |
| | | *Cryptosporidium* sp. | | |
| | Piroplasmea | *Babesia bovis* | | |
| Echinodermata | Echinoidea | *Strongylocentrotus purpuratus* | Sp Base: http://www.spbase.org/SpBase/ | Cameron et al. 2009 |
| Nemotoda | Chromadorea | *Caenorhabditis elegans* | http://www.wormbase.org/ | Chen et al. 2005 |
| Chordata | Acidiacea | *Ciona intestinalis* | http://genome.jgi-psf.org/Cioin2/Cioin2.home.html | Dehal et al. 2002 |

## Appendix B2

Table B.2: TSR-domain-containing protein sequence information

| Sequence | Protein type | Organism | Original identifier/ Accession Number |
|---|---|---|---|
| A.digAD1 | ADAMTS-like | *Acropora digitifera* | adi_v1.06910 |
| A.digAD2 | | | adi_v1.22282 |
| A.digAD3 | | | adi_v1.01169 |
| A.digAD4 | | | adi_v1.05790 |
| A.digAD5 | | | adi_v1.02358 |
| A.digAD6 | | | adi_v1.16603 |
| A.digAD7 | | | adi_v1.22282 |
| A.digAD8 | | | adi_v1.10799 |
| A.digAD9 | | | adi_v1.14055 |
| A.digAD10 | | | adi_v1.14580 |
| A.digAD11 | | | adi_v1.14366 |
| A.digAD12 | | | adi_v1.13631 |
| A.digAD13 | | | adi_v1.13245 |
| A.digAD14 | | | adi_v1.01169 |
| A.digIG1 | Igg domain | *Acropora digitifera* | adi_v1.11759 |
| A.digIG2 | | *Acropora digitifera* | adi_v1.14581 |
| A.digIG3 | | *Acropora digitifera* | adi_v1.07946 |
| A.digVWA1 | VWA domain | *Acropora digitifera* | adi_v1.24510 |
| A.digVWA2 | | *Acropora digitifera* | adi_v1.20375 |
| A.digVWA3 | | *Acropora digitifera* | Adi_v1.00391 |
| A.digCU1 | CUB domain | *Acropora digitifera* | adi_v1.07445 |
| A.digCU2 | CUB domain | *Acropora digitifera* | adi_v1.06393 |
| A.digTRYP | Trypsin | *Acropora digitifera* | Adi_v1.01708 |
| A.digLAM | Laminin | *Acropora digitifera* | Adi_v1.07180 |
| A.digF5F8 | F5 F8 typeC | *Acropora digitifera* | Adi_v1.05945 |
| A.digTSR1 | TSR domain only | *Acropora digitifera* | adi_v1.01708 |
| A.digTSR2 | | *Acropora digitifera* | Adi_v1.19596 |
| A.digTSR3 | | *Acropora digitifera* | Adi_v1.05200 |
| A.digTSR4 | | *Acropora digitifera* | Adi_v1.13412 |

| Sequence | Protein type | Organism | Original identifier/ Accession Number |
|---|---|---|---|
| A.milAD1 | ADAMTS-like | *Acropora millepora* | C000480 |
| A.milAD2 | | *Acropora millepora* | C017471 |
| A.milAD3 | | *Acropora millepora* | C023777 |
| A.milAD4 | | *Acropora millepora* | C009153 |
| A.milAD5 | | *Acropora millepora* | C001892 |
| A.milAD6 | | *Acropora millepora* | C002682 |
| A.milAD7 | | *Acropora millepora* | C002928 |
| A.milAD8 | | *Acropora millepora* | C003233 |
| A.milAD9 | | *Acropora millepora* | C002240 |
| A.milAD10 | | *Acropora millepora* | C002566 |
| A.milAD11 | | *Acropora millepora* | C001019 |
| A.milAD12 | | *Acropora millepora* | C008221 |
| A.milAD13 | | *Acropora millepora* | C004467 |
| A.milAD14 | | *Acropora millepora* | C018737 |
| A.milAD15 | | *Acropora millepora* | C000549 |
| A.milAD16 | | *Acropora millepora* | C004934 |
| A.milAD17 | | *Acropora millepora* | C002834 |
| A.milAD18 | | *Acropora millepora* | C010562 |
| A.milVWA1 | VWA domain | *Acropora millepora* | c001933 |
| A.milVWA2 | | *Acropora millepora* | c001272 |
| A.milVWA3 | | *Acropora millepora* | c001380 |
| A.milVWA4 | | *Acropora millepora* | c002257 |
| A.milTRYP1 | TRYPSIN domain | *Acropora millepora* | c026763 |
| A.milTRYP2 | | *Acropora millepora* | c006349 |
| A.milTRYP3 | | *Acropora millepora* | c011355 |
| A.milSEMA | Sema | *Acropora millepora* | c001041 |
| A.milTYRO | Tyrosinase | *Acropora millepora* | c007877 |
| A.milTSR1 | TSR ONLY | *Acropora millepora* | c025507 |
| A.milTSR2 | | *Acropora millepora* | c016853 |

| Sequence | Protein type | Organism | Original identifier/ Accession Number |
|---|---|---|---|
| AipAD1 | ADAMTS-like | *Aiptasia sp.* | comp15249 |
| AipAD2 | | *Aiptasia sp.* | comp22854 |
| AipAD3 | | *Aiptasia sp.* | comp19879 |
| AipAD4 | | *Aiptasia sp.* | comp17790 |
| AipAD5 | | *Aiptasia sp.* | comp14764 |
| AipAD6 | | *Aiptasia sp.* | comp19492 |
| AipAD7 | | *Aiptasia sp.* | comp24280 |
| AipAD8 | | *Aiptasia sp.* | comp13560 |
| AipAD9 | | *Aiptasia sp.* | comp11171 |
| AipAD10 | | *Aiptasia sp.* | comp9866 |
| AipAD11 | | *Aiptasia sp.* | comp31581 |
| AipAD12 | | *Aiptasia sp.* | comp19879 |
| AipAD13 | | *Aiptasia sp.* | comp10621 |
| AipAD14 | | *Aiptasia sp.* | comp10097 |
| AipAD15 | | *Aiptasia sp.* | comp14944 |
| AipAD16 | | *Aiptasia sp.* | comp3980 |
| AipAS1 | Astacin | *Aiptasia sp.* | comp17259 |
| AipAS2 | | *Aiptasia sp.* | comp17991 |
| AipAS3 | | *Aiptasia sp.* | comp28344 |
| AipWVA | WVA | *Aiptasia sp.* | comp16584 |
| AipTRYP | TRYPSIN | *Aiptasia sp.* | comp25690 |
| AipLAM | Laminin_G3 | *Aiptasia sp.* | comp18801 |
| A.eleAD | ADAMTS-like | *Anthopleura elegantissima* | comp13383 |
| A.eleAS1 | Astacin | *Anthopleura elegantissima* | comp625 |
| A.eleAS2 | | *Anthopleura elegantissima* | comp23921 |
| A.eleAS3 | | *Anthopleura elegantissima* | comp27202 |
| A.eleAS4 | | *Anthopleura elegantissima* | comp12431 |
| A.eleAS5 | | *Anthopleura elegantissima* | comp6424 |
| A.eleVWA1 | VWA | *Anthopleura elegantissima* | comp7794 |
| A.eleVWA2 | | *Anthopleura elegantissima* | comp40534 |

| Sequence | Protein type | Organism | Original identifier/ Accession Number |
|---|---|---|---|
| A.eleIG | Igg domain | *Anthopleura elegantissima* | comp146376 |
| A.eleTRYP | TRYPSIN | *Anthopleura elegantissima* | comp6788 |
| A.eleCUB | CUB domain | *Anthopleura elegantissima* | comp27202 |
| A.eleMAM | MAM domain | *Anthopleura elegantissima* | comp37953 |
| A.eleSEA | SEA domain | *Anthopleura elegantissima* | comp199145 |
| A.eleLAM | Laminin G3 | *Anthopleura elegantissima* | comp111612 |
| A.eleCALX | Calx-beta | *Anthopleura elegantissima* | comp530 |
| A.eleCBM | CBM 14 | *Anthopleura elegantissima* | comp11405 |
| A.eleTSR1 | TSR only | *Anthopleura elegantissima* | comp2495 |
| A.eleTSR2 | | *Anthopleura elegantissima* | comp44203 |
| A.eleTSR3 | | *Anthopleura elegantissima* | comp52752 |
| A.eleTSR4 | | *Anthopleura elegantissima* | comp6899 |
| A.eleTSR5 | | *Anthopleura elegantissima* | comp137837 |
| A.eleTSR6 | | *Anthopleura elegantissima* | comp123466 |
| A.eleTSR7 | | *Anthopleura elegantissima* | comp109002 |
| F.scuTRYP | TRYPSIN | *Fungia scutaria* | comp3626 |
| F.scuVWA | VWA | *Fungia scutaria* | comp4904 |
| F.scuLECB | Lectin B | *Fungia scutaria* | comp1435 |
| F.scuTSR1 | TSR only | *Fungia scutaria* | comp153497 |
| F.scuTSR2 | | *Fungia scutaria* | comp54007 |
| F.scuTSR3 | | *Fungia scutaria* | comp134907 |
| F.scuTSR4 | | *Fungia scutaria* | comp35136 |
| F.scuTSR5 | | *Fungia scutaria* | comp15630 |
| F.scuTSR6 | | *Fungia scutaria* | comp108427 |
| N.vecAD1 | ADAMTS- like | *Nematostella vectensis* | 95408\|e_gw.39.117.1 |
| N.vecAD2 | | *Nematostella vectensis* | 202989\|fgenesh1_pg.scaffold_39000072 |
| N.vecAD3 | | *Nematostella vectensis* | 208839\|fgenesh1_pg.scaffold_98000003 |
| N.vecAD4 | | *Nematostella vectensis* | 212944\|fgenesh1_pg.scaffold_166000015 |
| N.vecAD5 | | *Nematostella vectensis* | 87270\|e_gw.15.9.1 |

| Sequence | Protein type | Organism | Original identifier/ Accession Number |
|---|---|---|---|
| N.vecAD6 | ADAMTS - like | *Nematostella vectensis* | 246133 |
| N.vecAD7 | | *Nematostella vectensis* | 201879 |
| N.vecAD8 | | *Nematostella vectensis* | 247539 |
| N.vecAD9 | | *Nematostella vectensis* | 193804 |
| N.vecAD10 | | *Nematostella vectensis* | 101427 |
| N.vecAD11 | | *Nematostella vectensis* | 30346 |
| N.vecAD12 | | *Nematostella vectensis* | 246133 |
| N.vecAD13 | | *Nematostella vectensis* | 246132 |
| N.vecAD14 | | *Nematostella vectensis* | 30019 |
| N.vecAD15 | | *Nematostella vectensis* | 217812 |
| N.vecAD16 | | *Nematostella vectensis* | 215900 |
| N.vecAD17 | | *Nematostella vectensis* | *215897* |
| N.vecAD18 | | *Nematostella vectensis* | 133001 |
| N.vecAD19 | | *Nematostella vectensis* | 246793 |
| N.vecAD20 | | *Nematostella vectensis* | 212948 |
| N.vecAD21 | | *Nematostella vectensis* | 30434 |
| N.vecAD22 | | *Nematostella vectensis* | 174601 |
| N.vecAD23 | | *Nematostella vectensis* | 95477 |
| N.vecAD24 | | *Nematostella vectensis* | 222694 |
| N.vecAD25 | | *Nematostella vectensis* | 212950 |
| N.vecAD26 | | *Nematostella vectensis* | 126434 |
| N.vecAD27 | | *Nematostella vectensis* | 139926 |
| N.vecAD28 | | *Nematostella vectensis* | 99266 |
| N.vecIG1 | Igg domain | *Nematostella vectensis* | 20471 |
| N.vecIG2 | | *Nematostella vectensis* | 12396 |
| N.vecSEMA | SEMA domain | *Nematostella vectensis* | 197331 |
| n/a | TSR only | *Nematostella vectensis* | 33857, 138675, 2608, 86049, 34322 |
| n/a | | *Nematostella vectensis* | 110159, 116985 |
| n/a | | *Nematostella vectensis* | 919 |
| n/a | | *Nematostella vectensis* | 101342, 241642 |

| Sequence | Protein type | Organism | Original identifier/ Accession Number |
|---|---|---|---|
| n/a | TSR only | *Nematostella vectensis* | 112027, 39497, 84016, 27059, 101677, 110056, 110059, 112221, 112324, 117067, 120076, 121678, 122783, 126293, 128245, 129869, 130270, 138181, 155648, 185165, 201458, 28108, 39429, 39467, 57118, 81752, 8468, 85411, 85746, 86478, 88028, 88031, 9083, 91405, 92306, 92318, 9250, 9631 |
| n/a | | *Nematostella vectensis* | 218033, 39112, 121670, 127215, 146328, 148689, 5633, 19215, 126298, 147698, 148395, 79287, 94713, 97881, 6402, 16630, 6589, 106372, 110258, 110590, 116940, 119602, 79262, 86063, 9468, 9859, 119266, 245175, 203983, 214340, 218031 |
| N.vecAS1 | ASTACIN peptidase | *Nematostella vectensis* | 232545 |
| N.vecAS2 | | *Nematostella vectensis* | 211915 |
| N.vecFIB1 | Fibronectin domain | *Nematostella vectensis* | 217643, 217642 |
| N.vecFIB2 | | *Nematostella vectensis* | 213652 |
| N.vecCAD | Cadherin | *Nematostella vectensis* | 247587 |
| N.vecSOMA | Somatomedin | *Nematostella vectensis* | 226328 |
| N.vecWSC | WSC domain | *Nematostella vectensis* | 208719 |
| S.minTSR1 | TSR only | *Symbiodinium minutum* | symb1.aug_v1.008489.t1 |
| S.minTSR2 | | *Symbiodinium minutum* | symb1.aug_v1.031565.t1 |
| S.minTSR3 | | *Symbiodinium minutum* | symb1.aug_v1.039705.t1 |
| S.minTSR4 | | *Symbiodinium minutum* | symb1.aug_v1.041934.t1 |
| P.falTRAP | TRAP | *Plasmodium falciparum* | XP_001350088 |
| N.canTRAP | | *Neospora caninum* | AAF01565 |
| C.wraTRAP | | *Cryptosporidium wrairi* | AAC48313 |
| C.parTRAP | | *Cryptosporidium parvum* | XP_628162 |
| T.gonMNP | | *Toxoplasma gondii* | XP_002367474 |

| Sequence | Protein type | Organism | Original identifier/ Accession Number |
|---|---|---|---|
| S.purAD1 | ADAMTS | *Strongylocentrotus purpuratus* | XP_793479 |
| S.purAD2 | | | XP_001180853 |
| S.purAD6li | | | XP_001180742 |
| S.purGON | | | XP_001189419 |
| S.purPLA | | | XP_001187744 |
| D.melSEMA | Semaphorin | *Drosophila melanogaster* | NP_001163430 |
| C.eleAD1 | ADAMTS | *Caenorhabditis elegans* | NP_510116 |
| C.eleAD2 | | *Caenorhabditis elegans* | NP_001024534 |
| C.eleGON | | *Caenorhabditis elegans* | NP_501792 |
| C.elePAP | | *Caenorhabditis elegans* | NP_505017 |
| C.intAD1 | ADAMTS | *Ciona intestinalis* | XP_002121766 |
| C.intAD2 | | *Ciona intestinalis* | XP_002120522 |
| C.intPAP | | *Ciona intestinalis* | XP_002121111 |
| H.sapAD1 | ADAMTS | *Homo sapiens* | NP_008919 |
| H.sapAD2 | | *Homo sapiens* | NP_055059 |
| H.sapAD3 | | *Homo sapiens* | NP_055058 |
| H.sapAD4 | | *Homo sapiens* | NP_005090 |
| H.sapAD5 | | *Homo sapiens* | NP_008969 |
| H.sapAD6 | | *Homo sapiens* | NP_055088 |
| H.sapAD7 | | *Homo sapiens* | NP_055087 |
| H.sapAD8 | | *Homo sapiens* | NP_008968 |
| H.sapAD9 | | *Homo sapiens* | NP_891550 |
| H.sapAD10 | | *Homo sapiens* | NP_112219 |
| H.sapAD11 | | *Homo sapiens* | NP_620594 |
| H.sapAD12 | | *Homo sapiens* | NP_542453 |
| H.sapAD13 | | *Homo sapiens* | NP_620686 |
| H.sapAD14 | | *Homo sapiens* | NP_620687 |
| H.sapAD15 | | *Homo sapiens* | NP_620686 |
| H.sapAD16 | | *Homo sapiens* | NP_620687 |
| H.sapAD17 | | *Homo sapiens* | NP_620688 |

| Sequence | Protein type | Organism | Original identifier/ Accession Number |
|---|---|---|---|
| H.sapAD18 | ADAMTS | *Homo sapiens* | NP_955387 |
| H.sapAD19 | | *Homo sapiens* | NP_598377 |
| H.sapAD20 | | *Homo sapiens* | NP_079279 |
| H.sapADli1 | | *Homo sapiens* | ATL1_HUMAN |
| H.sapADli2 | | *Homo sapiens* | NP_055509 |
| H.sapADli3 | | *Homo sapiens* | NP_997400 |
| H.sapHCT | Hemicentin | *Homo sapiens* | HMCN1_HUMAN Hemicentin-1 Q96RW7 |
| H.sapPAP | papilin | *Homo sapiens* | NP_775733 |
| H.sapSEMA5A | Semaphorin | *Homo sapiens* | NP_003957 |
| H.sapSEMA5B | Semaphorin | *Homo sapiens* | Q9P283 |
| H.sapTSP1 | Thrombospondin | *Homo sapiens* | |
| H.sapUNC5 | UNC-5 | *Homo sapiens* | UNC5A_HUMAN Q6ZN44 |

## Appendix B3

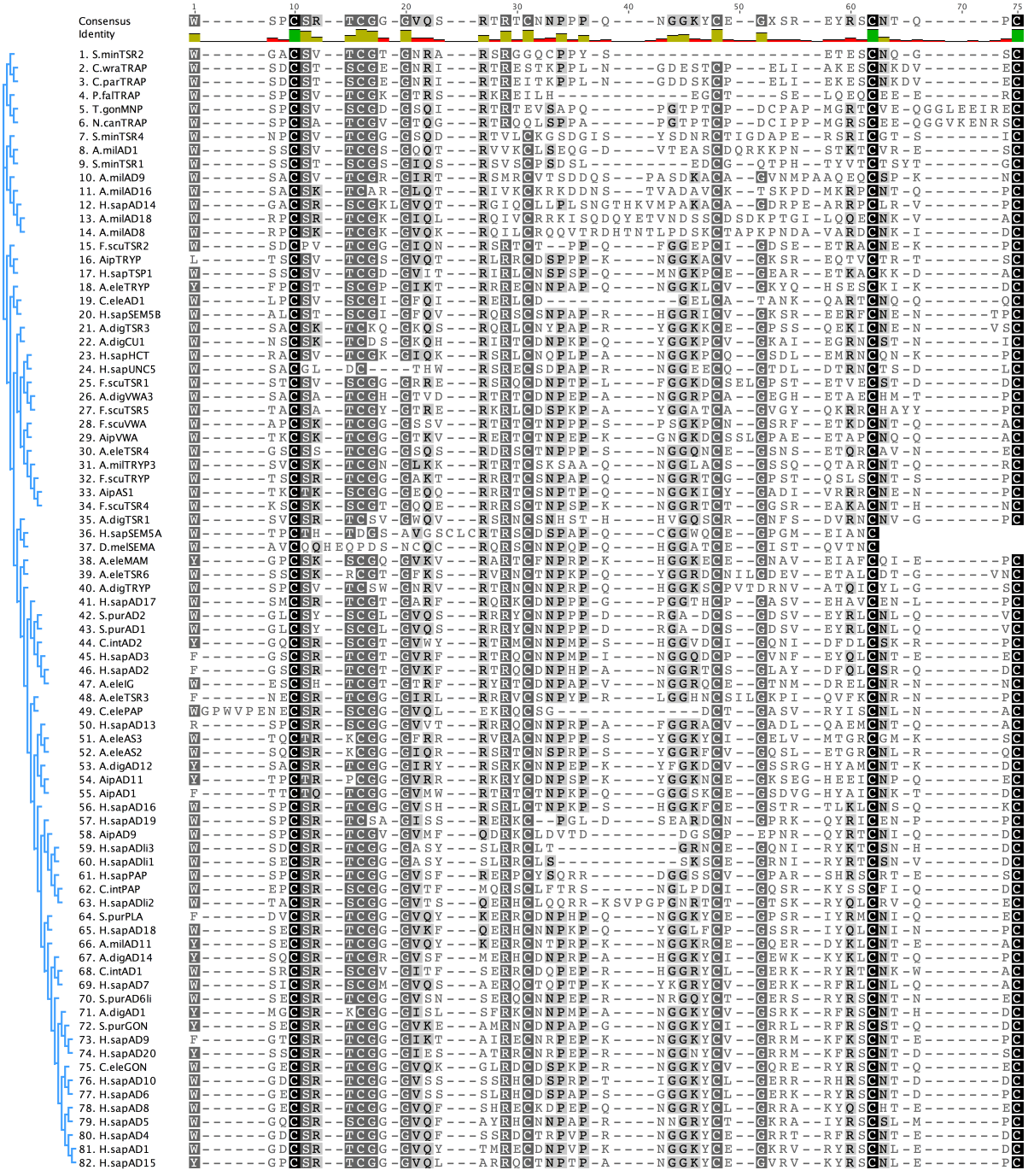



Figure B.1: Protein alignment of TSR protein domains used to produce the ML protein tree in Figure 3.6. Alignment shows similarity according to Blossum matrix 80, amino acids shaded according to similarity (black = fully conserved and white = variable).

**Appendix C1:**

## Immuno-precipitation and mass spectrometry failed to resolve anemone proteins binding to anti-TSP

*Methods*

Anti-TSP labeled two bands in western blot analysis of protein homogenate from both symbiotic and aposymbiotic *Aiptasia sp.* (Figure 4.4). Immuno-precipitation for mass spectrometry followed a combination of the Santa Cruz (http://www.scbt.com/protocols.html?protocol=immunoprecipitation) protocol. Sets of eight aposymbiotic or symbiotic anemones were ground in 1 ml of ice-cold homogenization buffer (50 mM Tris–HCl, pH 7.4, 300 mM NaCl, 5 mM EDTA) with protease inhibitor cocktail (BD Biosciences, San Jose, CA, USA). Homogenates were centrifuged at 4°C for 15 min at 14,000g, supernatants were decanted and pre-cleared by adding 100 µl of protein G beads according to the manufacturer's instructions (Sigma, St. Louis, MO, USA). After centrifugation at 14,000 µg at 4 °C for 10 min, supernatants were decanted and protein concentrations were determined using a Bradford assay. One mg of pre-cleared homogenate was immuno-precipitated overnight at 4 °C with 20 µg of rabbit polyclonal anti-TSP with an epitope corresponding to the three TSR domains of human thrombospondin proteins 1 and 2 (H-300, Santa Cruz Biotechnology sc-14013), followed by capture of the immunocomplex with protein G beads for 1 h at 4 °C. After a pulse centrifugation, the immunocomplexes were harvested and washed 3 times with 800 µl PBS. The immunocomplexes were then resuspended in Laemmli sample buffer (Biorad, Hercules, CA, USA) and boiled for 5 min, followed by SDS-PAGE analysis of the supernatent. The resulting bands were excised and digested using an in-gel digestion with ProteaseMax (Promega), and the final product was prepared for mass spectrometry analysis.

*Results*

Mass spectrometry analysis yielded no conclusive results as no relevant protein matches were obtained from database searches of the data obtained. Two possible explanations are suggested for the failure of this technique in this instance. Firstly it is

very difficult to get enough protein from small *Aiptaisia sp.* anemones to produce a strong enough sequence for mass spectrometry analysis. After the pre-clearing step I had approx. 7000 µg/ml of total anemone protein for the IP. The amount of protein within a band digestion from a 1-D gel was likely to be very low. Secondly, the available searchable database was limited to the *Nematostella vectensis* genome at the time of this analysis, as the *Aiptaisia* sp. transcriptome was in the early stages of development. Due to these two limitations, I decided not to repeat this experiment a second time.