Integrin subunits: expression and function in early development of *Strongylocentrotus purpuratus*

by

Mary Elizabeth Brothers
BSc, St. Francis Xavier University, 2005

A Thesis Submitted in Partial Fulfillment
of the Requirements for the Degree of

MASTER OF SCIENCE

in the Department of Biology

Integrin subunits: expression and function in early development of *Strongylocentrotus purpuratus*

by

Mary Elizabeth Brothers
BSc (Honours), St. Francis Xavier University, 2005

Supervisory Committee

Dr. Robert D. Burke, Supervisor
(Department of Biology, Department of Biochemistry and Microbiology)

Dr. Robert J. Ingham, Departmental Member
(Department of Biology)

Dr. John S. Taylor, Departmental Member
(Department of Biology)

Supervisory Committee

Dr. Robert D. Burke, Supervisor
(Department of Biology, Department of Biochemistry and Microbiology)

Dr. Robert J. Ingham, Departmental Member
(Department of Biology)

Dr. John S. Taylor, Departmental Member
(Department of Biology)

Abstract

Integrins are heterodimeric transmembrane receptors composed of an α and a β subunit, that are expressed on the surface of all metazoan cells. These bidirectional signaling molecules are involved in many well-known aspects of cell function, although the role of integrins in early embryonic development remains a mystery. The purpose of this study was to characterize *S. purpuratus* integrins and determine if they are necessary for early embryonic development. Full length cDNA sequences for four incomplete gene predictions, αC, αD, αF, and βD, were determined by amplifying overlapping fragments and sequencing EST clones. Each cDNA has a single open reading frame predicting a protein with canonical integrin features. QPCR results show αC, αD, and βD are expressed in the embryo at relatively constant levels during the first 96 hours of development. αF is expressed in blastulae, during morphogenesis and tissue differentiation, at up to 35 times the levels of mRNA in the egg. Using a morpholino antisense oligonucleotide to block translation of αC results in a higher than normal mortality rate (57.1%) by 24 hours of development and 36.7% of embryos during this period have defects in aspects of cell division. These results indicate that αC is an essential gene for early development and that it may function in coordination of mitosis and cytokinesis. The expression of multiple subunits and the demonstration that αC has an essential role suggests that there are several non-overlapping functions for integrins in early embryonic development.

Table of Contents

Supervisory Committee....ii
Abstract....iii
Table of Contents....iv
List of Tables....vi
List of Figures....vii
List of Abbreviations....ix
Acknowledgments....xii

Chapter 1 – Introduction....1
1.1 Project overview and objectives....1
1.2 Integrins....3
1.2.1 Integrin structure and function....5
1.2.2 Integrin binding....6
1.2.3 Integrin signaling....8
1.2.4 Sea urchin integrins....12
1.3 *Strongylocentrotus purpuratus* as a model organism....14
1.4 Early embryonic development of *S. purpuratus*....15

Chapter 2 – Materials and Methods....18
2.1 Adult *S. purpuratus* culturing and collection of gametes....18
2.2 PCR amplification of cDNA....18
2.2.1 Primer design....18
2.2.2 RNA isolation and RT-PCR....19
2.2.3 PCR....20
2.2.4 Cloning and sequencing....21
2.2.5 5' RACE Amplification of αC....22
2.3 Phylogenetic analysis....23
2.4 Quantitative PCR....24
2.4.1 Primer Design....24
2.4.2 mRNA isolation and RT-PCR....24
2.4.3 QPCR....25
2.5 Microinjection of Sp-αC MASO....27
2.5.1 Injection of embryos....27
2.5.2 Immunofluorescence of injected embryos....28

Chapter 3 – Results....30
3.1 Sequencing results and protein domains....30
3.1.1 αC....30
3.1.2 αF....34
3.1.3 αD....37
3.1.4 βD....40
3.2 Phylogenetic Analysis....43

3.3 Temporal Expression 46
3.4 Anti-sense morpholino knockdown of αC 49
3.4.1 αC is necessary for early development 49
3.4.2 Abnormalities of cleavage-stage embryos 51
3.4.3 Morphogenetic defects 57
3.4.5 Polarity in early and blastula stage embryos 63

Chapter 4 – Discussion 67
4.1 Integrin subunit structure 67
4.2 Phylogeny 68
4.3 Temporal Expression 70
4.4 Function of αC in early development 71

Chapter 5 – Conclusions and Future Directions 76

Literature Cited 79
Appendix I Primer Sequences 88
Appendix II Protein alignments 91
Appendix III Temporal Expression Patterns 105

List of Tables

Table 2.1 Components of a standard PCR reaction 20

Table 2.2 Components of a QPCR reaction 25

Table 2.3 Nanograms of cDNA used in each QPCR reaction varied between stages 26

Table 3.1 Tabular notes comparing the regions of αC cDNA..... 31

Table 3.2 Tabular notes comparing the regions of αF cDNA..... 34

Table 3.3 Tabular notes comparing the regions of αD cDNA..... 37

Table 3.4 Tabular notes comparing the regions of βD cDNA..... 40

List of Figures

Figure 1.1 Cartoon depiction of an integrin heterodimer.....6

Figure 1.2 Light micrographs of developmental stages of *S. purpuratus* from unfertilized egg with sperm (S) to prism-stage larva with a tripartite gut (TG).....17

Figure 3.1 Schematic diagrams of the cDNA sequence and predicted protein domains of αC. 33

Figure 3.2 Schematic diagrams of the cDNA sequence and predicted protein domains of αF..... 36

Figure 3.3 Schematic diagrams of the cDNA sequence and predicted protein domains of αD..... 39

Figure 3.4 Schematic diagrams of the cDNA sequence and predicted protein domains of βD. 42

Figure 3.5 Neighbour-joining tree of the integrin_α2 domain protein sequences constructed using MEGA 4.0. 45

Figure 3.6 Neighbour-joining tree of the integrin_α2 domain in *S. purpuratus* α integrin protein sequences constructed using MEGA 4.0 46

Figure 3.7 Relative expression levels of αC, αD, βD, and βC obtained by QPCR data during the first 96 hours of development of *S. purpuratus* embryos 48

Figure 3.8 Relative expression levels of αF obtained by QPCR data during the first 96 hours of development of *S. purpuratus* embryos 49

Figure 3.9 Knockdown of αC expression results in increased mortality 51

Figure 3.10 Abnormal phenotypes during early cleavage stages. 54

Figure 3.11 Abnormal nuclear phenotypes in early cleavage stages 55

Figure 3.12 Phalloidin staining in *S. purpuratus* embryos injected with Sp-αC MASO.. 56

Figure 3.13 *S. purpuratus* embryos injected with Sp-αC MASO that survive to 24 hours show blastocoel defects. 59

Figure 3.14 Nuclear phenotypes through 24 hours of development 60

Figure 3.15 *S. purpuratus* embryos injected with Sp-αC MASO exhibit abnormal phenotypes through gastrulation. 62

Figure 3.16 Sp-Par6 localization in αC MASO injected embryos at early cleavage stages. 64

Figure 3.17 Sp-Par6 localization in Sp-αC MASO injected embryos. 65

Figure 3.18 A surface view of Sp-Par6 localization in *S. purpuratus* embryos injected with Sp-αC MASO..... 66

List of Abbreviations

~	approximately
%	percent
<	less than
>	greater than
ºC	degrees Celsius
α	alpha
β	beta
µg	microgram
µL	microlitre
µm	micrometre
µM	micromolar
aPKC	atypical protein kinase C
Arp2/3	actin-related protein 2/3
ATA	3-amino-1,2,4-triazole
AV	animal-vegetal
BCM	Baylor College of Medicine
BLAST	basic local alignment search tool
bp	base pairs
CAM	cell adhesion molecule
CDK	cyclin-dependent kinase
cDNA	complementary deoxyribonucleic acid
DIC	differential interference contrast
DNA	deoxyribonucleic acid
dNTP	deoxynucleotidetriphosphate
ECM	extracellular matrix
EGF	epidermal growth factor
EMT	epithelial to mesenchymal transition
ERK	extracellular signal-related kinase 1
EST	expressed sequence tag

F-actin	filamentous actin
FAK	focal adhesion kinase
FERM	band 4.1 ezrin/radixin/moesin
FSW	filtered sea water
hr	hour
G	gravity
G-actin	globular actin
GPCR	G-protein coupled receptors
GTP	guanosine triphosphate
IAP	integrin associated protein
IgSF	immunoglobulin superfamily
ILK	integrin-linked kinase
IPTG	isopropyl β-D-1-thiogalactopyranoside
JNK	Jun N-terminal kinase
KCl	potassium chloride
Kb	kilobase
kD	kiloDalton
L	litre
LB	Luria-Bertani broth
LGL	lethal giant larvae
LRR	leucine rich repeats
MAPK	mitogen activated protein kinase
MASO	morpholino antisense oligonucleotide
MIDAS	metal ion dependent adhesion site
mg	milligram
mL	millilitre
mM	millimolar
mRNA	messenger ribonucleic acid
MTOC	microtubule organizing centre
N	number
NAD	nicotinamide adenine dinucleotide

NCBI	National Centre of Biotechnology Information
ng	nanograms
nM	nanomolar
NMR	nuclear magnetic resonance
OA	oral-aboral
PBS	phosphate buffered saline
PBS-T	phosphate buffered saline with Tween20
PCR	polymerase chain reaction
Pfam	protein family
PI	phosphatidylinositol
PI 3-kinse	phosphoinositide 3-kinase
PIP2	phosphatidylinositol(4,5)-bisphosphate
PKC	protein kinase C
PMC	primary mesenchyme cells
PS	phosphatidylserine
QPCR	quantitative polymerase chain reaction
RACE	rapid amplification of complementary DNA ends
RGD	arginine-glycine-aspartic acid
RNA	ribonucleic acid
RT-PCR	reverse transcriptase polymerase chain reaction
SMART	simple modular architecture research tool
SOC	super optimal catabolite repression
WASP	Wiskott-Aldrich syndrome protein
X-gal	5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside
U	units
UTR	untranslated region
v/v	volume to volume ratio
w/v	weight to volume ratio
x	times

Acknowledgments

First and foremost, thank you to Dr. Robert Burke for giving me the opportunity to come and study in his lab at the University of Victoria, and for being a wonderful teacher and supervisor. I have enjoyed this experience and learned so much. Thank you also to my committee members, Dr. Rob Ingham and Dr. John Taylor, for their advice, words of encouragement and support.

I would like to thank Diana Wang for her extensive advice in the technical realm of this project. Without her help and support I could not have accomplished as much as I did. I would also like to extend a huge thank you to Navraj Chima, another person who I could not have completed this project without. Nav not only helped me with the sequencing portion of the project, but became an invaluable expert in the technique of microinjections. Also thank you to all of the other students that have been in the Burke lab with me over the course of my degree – Andrew Juurinen, Chantel Anderson, Kate MacDonald, Nick Church, Nathan West, and Stu Trenholm. You've all become close friends and I've enjoyed the time we spent working, chatting, and laughing. Another thank you goes out to Allison Churcher, not only for her help in the 5' RACE endeavour, but also for offering wonderful advice for my many questions.

I would like to acknowledge and thank Dr. Caroline Cameron's lab who let me invade regularly and take over their QPCR machine. I would also like to say a big thank you to Nik Veldhoen for teaching me everything I know about the process of QPCR.

Finally, thank you to Amy Dove for agreeing to be my roommate through yet another thesis and listening about the wonderful world of sea urchins for two whole years. You've made my time in Victoria so much fun and I could never have finished this without you.

Chapter 1 – Introduction

1.1 Project overview and objectives

Integrins are important heterodimeric cell adhesion molecules that have been the subject of thousands of research papers. Even with this vast knowledge about these receptors and how they function in the cell, there is very little known about the role they play in early embryonic development. The recent publication of the *S. purpuratus* genome revealed gene predictions for several integrin subunits, making sea urchins a good candidate as a model organism to investigate integrins in these early stages. Previous research has found integrins to be expressed in *S. purpuratus* embryos and several subunits, including αP, βC, βG, and βL, have been characterized (Burke et al., 2004; Marsden and Burke, 1997; Marsden and Burke, 1998; Murray et al., 2000a; Susan et al., 2000). A study by Burke et al., (2004) used an antisense technology to knock down expression of one of the integrin subunits, βC, and found it to be essential for formation of the cortical actin cytoskeleton. This observation led to questions about the expression, regulation, and function of the other integrin subunits. It is these questions that initiated this study.

Using sea urchin embryos as a developmental model, I investigated the expression and role of integrins in development. I hypothesized that the integrin subunits expressed during early embryogenesis of *S. purpuratus* contribute to functional receptors, that they are temporally and spatially regulated, and that they are necessary for development.

S. purpuratus has eight predicted α integrin subunits and four predicted β subunits: αP, αC, αK, αD, αJ, αF, αH, αG, βC, βD, βL, and βG (Whittaker et al., 2006).

Four of these have been previously cloned, sequenced, and characterized, but the gene models for all of the others were incomplete. My first objective was to test the genomic predictions for these genes by getting complete cDNA sequences for the remaining subunits. Several of these sequences were obtained using PCR amplification of *S. purpuratus* cDNA. Although the genome of *S. purpuratus* has been sequenced, the current assembly (Version 2.1 is incomplete as are many individual gene predictions. Translation initiation sites, intron-exon boundaries, and alternative splice forms can only be determined with certainty from cDNA sequences. Primers to amplify these sequences were designed using several gene predictions. Gene predictions are generated using various gene recognition algorithms. These algorithms assemble genomic sequences to predict mRNA. These assembled genomic sequences and mRNA predictions (Sea Urchin Genome Sequencing Consortium, 2006) are available in public databases, including gene annotations by the Human Genome Sequencing Centre at the Baylor College of Medicine (http://annotation.hgsc.bcm.tmc.edu/Urchin/); *Strongylocentrotus purpuratus* Genome Version 2 at Genboree (Baylor College of Medicine) (http://www.genboree.org/java-bin/login.jsp); and *Strongylocentrotus purpuratus* Genome Version 2.1 at NCBI (http://www.ncbi.nlm.nih.gov/mapview/map_search.cgi?taxid=7668).

The second objective of my research was to determine temporal expression patterns of the predicted integrin genes using quantitative PCR (QPCR) of integrin subunits whose cDNA sequences had been amplified. Determining the relative abundance of the message of each gene during early development offers insight into their function during this critical period. I expected the integrins to be temporally regulated

during embryogenesis and for these expression patterns to indicate potential functions of the individual genes. Temporal expression patterns of all predicted genes in the *S. purpuratus* genome have been obtained using microarray hybridization data, and they are available in a public database (http://urchin.nidcr.nih.gov/blast/index.html) (Wei et al. 2006). This provides a useful data set with which to compare my results.

To further understand if integrins function in early embryonic development and gain some insight into their role, the third objective of my research was to use a morpholino antisense oligonucleotide (MASO) to interfere with translation of the αC subunit. This subunit was chosen because QPCR data indicating it is expressed during early embryogenesis and the cDNA sequence for the start of translation were available early on in the project for design of a MASO. A long term goal is to determine the function of all the subunits by using gene knockdowns, but the one chosen for this study was αC. I expected αC to be a critical gene in this stage of development and that blocking expression of the protein would have a serious, possibly fatal, consequence on the developing embryo. However, as there are multiple subunits expressed in early development, it was equally likely that they have overlapping functions and interfering with translation would produce no defects or affect embryo viability.

1.2 Integrins

Cell signaling is a basic and necessary cellular process. Although single-celled organisms are able to signal, it is especially important in metazoans as cells must communicate for the organism to function. All cells have mechanisms for perceiving signals in the external environment, transmitting these to the interior of the cell, and

responding appropriately. The ability of cells to communicate is critical for basic cellular functions such as motility, growth, development, repair, and homeostasis. All multicellular life shares common mechanisms of cell signaling (Pires-daSilva and Sommer, 2003).

Although cell signaling is critical, it is an intricate and complex process which is not yet fully understood. An important group of receptors involved in cell signaling, which are found in all metazoans, are integrins. Integrins are cell surface receptors involved in many aspects of cell-cell contact and cell-extracellular matrix (ECM) adhesion, so named because they integrate the external and internal environments of the cell. Integrin receptors are heterodimers, each composed of one α and one β subunit, which are noncovalently bound. There are numerous subunits, and in mammals eight β and eighteen α subunits have been identified, although only 24 receptors are known (Burke, 1999; Hughes, 1992; Hynes, 1987; Hynes and Lander, 1992; Hynes, 2002). Integrins are capable of binding to many different ligands, including various components of the ECM (fibronectin, vitronectin, and laminin), adhesive proteins in blood, immunoglobulin proteins, and even other integrins (Hynes, 1987; Hynes, 2002; Luo et al., 2007; Schwartz, 1994; Smith, 1994). In addition to providing a mechanism of cell-cell interaction and adhesion, integrins are also responsible for activating many intracellular signaling pathways (Hynes, 2002; Katz and Yamada, 1997; Schwartz, 1994). There has been a lot of research done on integrins in the past 25 years, with novel insights into their functions and mechanisms continually being revealed. Although it is impossible to incorporate all of the research done on integrins, I will provide a brief

summary of background on the structure and function of integrins and what is known about integrins in sea urchins.

1.2.1 Integrin structure and function

Although most α subunits are a single polypeptide, some are proteolytically processed to one heavy and one light chain, linked by a disulfide bond. In these cleaved forms, the light chain contains the transmembrane domain and short cytoplasmic tail (Hughes, 1992; Smith, 1994) (Fig. 1.1). The N-terminal region of this subunit is folded into a seven-bladed β-propeller. In some vertebrate α subunits an inserted (I) domain consisting of α-helices surrounding a central β-sheet is a 200 residue insertion replacing the first divalent cation binding site. In these subunits the I-domain is necessary for ligand recognition and binding. Mg^{2+} binds to the N-terminus of the I-domain via a cation-binding domain (Shimaoka et al., 2002). The extracellular region C-terminal to the β-propeller is known as the leg of the α subunit. It contains the genu region, which is the region at which the integrin bends as it changes between active and inactive conformations (Luo et al., 2007; Xiong et al., 2001).

Each β subunit contains 56 extracellular cysteine residues consistent with a folded conformation stabilized by disulfide bonds. The amino terminal region contains extensive internal disulfide bonding (Hughes, 1992; Smith, 1994). The majority of this subunit is extracellular and there is a transmembrane domain, followed by a short cytoplasmic tail (Fig. 1.1). The β subunit also has a leg domain with a genu region, similar to the α subunit.

Integrins are not fixed structures and are mobile within the plane of the cell membrane. They are clustered in higher order structures, such as focal adhesions.

Integrin reorganization takes place during cell migration and may be an important function of that process. Clustering and movement of integrins occurs while in the presence of ligands, and this rearrangement appears to be important to signal transduction (Luo et al., 2007).

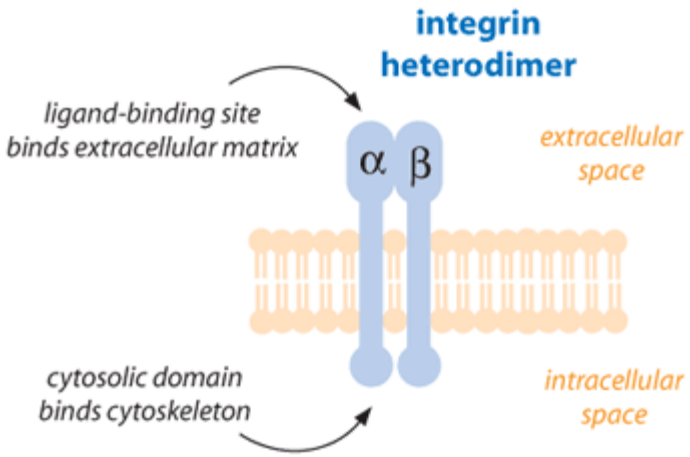


Figure 1.1 Cartoon depiction of an integrin heterodimer (Eslami and Philpot, 2005)**.** Ligand binding occurs in the extracellular N-terminal region of the integrin. Signal transduction and receptor activation involves the short cytoplasmic tails.

1.2.2 Integrin binding

Both α and β subunits contribute to the ligand binding region of the integrin receptor, although that relationship is complicated and it is often considered to be the α subunit primarily responsible for determining the ligand (Hynes, 1987; Hynes and Lander, 1992; Smith, 1994). β subunits are more promiscuous than α subunits and one β subunit can usually combine with several different α subunits to form a functional integrin (Hynes, 2002).

One well-studied integrin ligand is the arginine-glycine-aspartic acid tripeptide (RGD), found within an 11 kD fragment from fibronectin and other extracellular matrix

proteins (Pierschbacher and Ruoslahti, 1984; Smith, 1994). While this is a common integrin binding domain, not all integrins bind this motif. The one component of all integrin ligands is an aspartic acid residue (Smith, 1994) suggesting a similar mechanism of ligand engagement. Of the mammalian integrin receptors, α3, α6, and α7 bind to laminin, whereas α5, αV, α8, and αIIb bind to RGD-containing peptides. Of the *D. melanogaster* integrins, αPS1 binds to RGD and αPS2 binds to laminin (Hynes and Zhao, 2000; Hynes, 2002). *C. elegans* has two α subunits; INA-1 which binds laminin, and PAT-2 which binds RGD (Kramer, 2005).

Integrins have active and inactive conformations. The low-affinity, or inactive state, is a bent conformation and the active form is found in an upright, open conformation (Legate et al., 2006; Luo et al., 2007; Xiong et al., 2001). The conformational change between the two forms is accompanied by an increase in affinity for the specific ligand of the integrin. Integrin affinity for ligands is not completely lost in the inactive form (Luo et al., 2007).

Divalent cations, Mg^{2+} or Mn^{2+}, are required for ligand binding. The metal-coordinating residues and residues surrounding the metal-binding site contribute directly to formation of the ligand binding site. This site has been named the metal ion-dependent adhesion site (MIDAS) (Luo et al., 2007; Shimaoka et al., 2002; Xiong et al., 2001). Mn^{2+} induces a high-affinity or active state in some integrin receptors (Luo et al., 2007; Smith, 1994). Low levels of Ca^{2+} are required to maintain the association of α and β subunits and promote binding of ligands, while high levels can interfere with ligand binding (Luo et al., 2007).

While in the inactive state, the legs of the α and β domains are close together and weakly bound via a salt bridge (Hughes et al., 1996). Disruption of this bond results in a spatial separation of the two cytoplasmic domains via movement of the β domain, leading to a change in the leg domains. Mechanisms which induce integrin activation, such as ligand binding, extracellular addition of Mn^{2+}, and inside-out signaling also result in this movement (Luo et al., 2007; Xiong et al., 2001). Integrins appear to be in equilibrium between several forms and they readily switch conformation, although the default state of the integrin seems to be the bent, inactive form (Liddington and Ginsberg, 2002). Activated integrins result in ligand binding, integrin clustering, and recruitment of cytoplasmic proteins into focal adhesions (Grashoff et al., 2004; Luo et al., 2007).

1.2.3 Integrin signaling

Outside-in signaling

Integrins have the capacity for bidirectional signaling. Integrin mediated ligand binding can cause a downstream signaling cascade which initiates a response in the cell. Carrying this message into the cell is known as outside-in signaling. Signals to integrin bearing cells that cause integrins to change from an inactive conformation to an active conformation is known as inside-out signaling.

Integrins often localize to focal adhesions, large stable complexes that mediate cell adhesion to the substrate and anchor actin microfilaments. An initial step in the formation of focal adhesions is the clustering of activated integrins, a process involving talin, a cytoplasmic scaffolding protein critical to receptor activation. There are many proteins associated with focal adhesions including numerous cytoskeletal and signal transduction molecules, as well as adaptor proteins that bind to integrins including Src,

FAK (focal adhesion kinase), and talin. In addition, talin is a signal transduction molecule that mediates integrin signaling, along with filamin, paxillin, and integrin-linked kinase (ILK) (Burke et al., 2007; Campbell, 2008; Filipenko et al., 2005; Geiger et al., 2001; Humphries et al., 2007; Katz and Yamada, 1997; Ozaki et al., 2007).

Integrin cytoplasmic domains signal through the kinases FAK and Syk (Ferrell and Martin, 1989; Schwartz, 1994). In an outside-in signaling event, when platelets bind to integrins for example, the earliest reaction detectable within the cell and therefore one of the earliest downstream reactions is activation of Src and Syk protein tyrosine kinases. Src kinases are activated by the binding of fibrinogen. Syk is recruited to the complex and activated by Src, allowing either one to phosphorylate downstream components. As the signal propagates downstream, other factors that have the ability to influence actin dynamics and reorganization become involved (Shattil and Newman, 2004).

FAK is recruited to the membrane through its C-terminal domain as a result of integrin clustering into the adhesion complexes, as well as proteins associated with integrins including paxillin and talin (Sieg et al., 2000). FAK is accompanied by the activation of the MAPK (mitogen activated protein kinase) pathway, one of the pathways that links FAK to integrin-dependent cell survival. FAK also activates the c-JNK (Jun-N-terminal kinase) pathway, which is thought to be one way in which integrins are involved in regulation of the cell cycle (Jan et al., 2004; Katz and Yamada, 1997; Schwartz, 2001).

The involvement of integrins in the cell cycle through both integrin-dependent adhesion and signaling pathways is more complicated than just signaling through FAK. Coordinate signaling between integrins and receptor tyrosine kinases is important in regulation of the cell cycle, but the pathways involved are interwoven and complex. In

vertebrates, growth factor receptor signaling, in addition to phosphorylation of FAK by integrins, is responsible for induction of cyclin D1, a protein involved in regulation of the cell cycle. The reorganization of the cytoskeleton as a result of integrin signaling also stimulates translation of cyclin D1 (Assoian and Schwartz, 2001). In addition to being involved in pathways responsible for proliferation of the cell cycle, integrins have also been linked to apoptosis. This can occur through interruption of these pathways or cell-ECM contact, a process known as anoikis (Jan et al., 2004). Integrins are able to intrinsically, through stress, and extrinsically, through growth factors, regulate apoptosis showing how crucial they are to development (Assoian and Schwartz, 2001; Hulleman and Boonstra, 2001; Stupack and Cheresh, 2002).

Integrins are involved in cooperative signaling with growth factor receptors and actually regulate aspects of growth factor activation. During embryogenesis, growth factors are involved in cellular growth and differentiation. There are two hypotheses for growth factor activation by integrins, direct and collaborative. Direct activation happens without a growth factor ligand, but integrin binding results in tyrosine phosphorylation, which subsequently clusters and activates growth factors. Collaborative activation is the result of clustering of both integrins and growth factor receptors by an integrin ligand (Yamada and Even-Ram, 2002).

Integrins are involved in regulation of assembly of the actin cytoskeleton. There are many components that link integrins to the cytoskeleton, including parvin, Syk, tetraspanin proteins, talin, vinculin, and α-actinin. Actin-related protein 2/3 (Arp2/3) plays a major role in actin polymerization, which is recruited to the site of integrin binding by vinculin. Members of the Wiskott-Aldrich syndrome protein (WASP) are

regulators of the Arp2/3 complex and have also been known to associate with integrins (DeMali et al., 2003). The WASP family of proteins is also associated with the small GTPases Cdc42 and Rac. The effect of WASP on Arp2/3 induced actin polymerization is stimulated by PIP_2, which is synthesized by kinases recruited to integrin binding sites via talin (DeMali et al., 2003).

Inside-out signaling

The affinity of integrin receptors for their ligands can be regulated by signaling through other receptors, a phenomenon known as inside out signaling. Inside-out signaling regulates adhesion of the integrins, while outside-in signaling affects behaviour of the cell by passing messages into the cytoplasm (Lallier et al., 1994; Luo et al., 2007). Conformation of the integrin plays an important role in mediating inside-out signaling, while clustering in addition to conformational changes are necessary in outside-in signaling (Luo et al., 2007). The best known example of inside-out signaling is activation of the integrin to initiate ligand binding in leukocytes and platelets (Hynes, 2002).

The integrin cytoplasmic domain has an important role in signaling as it controls the transition from inactive to active form of the integrin. Although small, α cytoplasmic tails are 20 – 40 amino acids and β cytoplasmic tails are 45 – 60 amino acids, they are crucial for a functional integrin. Interference with the interaction that takes place between the cytoplasmic domains of the α and β subunits leads to activation and an increase in receptor affinity. There are several intracellular proteins that bind to the cytoplasmic tails of integrins, particularly the β subunit, that mediate integrin activation. An important one of these is talin, an antiparallel dimer of approximately 270-kD with a

50-kD N-terminal FERM domain and a 220-kD C-terminal rod domain. The N-terminal domain of talin binds to the cytoplasmic domain of some β subunits resulting in activation of the integrin molecule (Oxley et al., 2007; Shattil and Newman, 2004; Simon and Burridge, 1994; Tadokoro et al., 2003). This binding replaces the weak interaction between the α and β subunit causing the domains to separate and the integrin to be activated. This domain separation and subsequent activation is the basis for inside-out signaling (Travis et al., 2003).

1.2.4 Sea urchin integrins

The *S. purpuratus* genome has predictions for eight α subunits, only one of which had been cloned and sequenced before this study (Susan et al., 2000; Whittaker et al., 2006). At the outset of this project the predictions themselves were incomplete and of the eight subunits, only two encoded what appeared to be complete genes. Analysis of the genome predicted four β subunits, three of which have had full length cDNA sequences confirmed (Marsden and Burke, 1997; Murray et al., 2000a; Whittaker et al., 2006).

The one α subunit with a confirmed cDNA sequence, αP, predicts a protein that is 1038 amino acid residues with a molecular weight of 113 kDa. αP has conserved cysteine residues and motifs found in other integrins. Levels of αP mRNA are low in the unfertilized egg, completely disappear during early cleavage stages, and increase during gastrulation, reaching a peak at prism stage. Western blots have shown protein expression follows a similar pattern with a peak in prism stage, although the protein never disappears completely as mRNA does (Susan et al., 2000).

The three β subunits for which there are cloned cDNAs are βC, βG, and βL. βG has an open reading frame that predicts a protein of 783 residues, 686 of which belong in

the extracellular domain. The 56 conserved cysteine residues are present in the extracellular domain of βG. The mRNA of βG increases throughout early development peaking at gastrulation before it decreases. Immunolocalization with a subunit specific antiserum found the protein localizes to the apical domains of blastomeres during cleavage (Marsden and Burke, 1997). The cDNA sequence of βL also has an open reading frame with a predicted protein consisting of 796 amino acids. Although βL has the typical cytsteine residues in the extracellular domain, some of these residues are in locations unique to this subunit. Expression of βL mRNA increases as development progresses with its peak expression times during the late gastrula and pluteus stages. The βL protein is localized to the basolateral domains and is necessary for gastrulation of the sea urchin embryo. Inhibition of βL also affects actin localization in all cells with the exception of vegetal plate cells, suggesting an important role for βL in embryonic development (Marsden and Burke, 1998).

βC has a single open reading frame with a protein prediction of 806 amino acids, 712 in the extracellular domain. Within this domain there are 10 potential N-linked glycosylation sites and 56 conserved cysteine residues. A MIDAS domain is also found. The βC protein is expressed in the unfertilized egg and although the protein is proteolytically removed at fertilization, it is re-expressed within 30 minutes of fertilization and localizes to the outer surface of the embryo (Burke et al., 2004; Burke et al., 2007; Murray et al., 2000a). A gene knockdown of βC using a morpholino antisense oligonucleotide results in loss of cortical actin suggesting that βC plays a role in cortex development. A βC protein lacking the cytoplasmic domain fails to rescue eggs injected with the βC MASO whereas full length cDNAs encoding βC or chicken β1 subunits

produced normal larvae. This study hypothesized that the actin rich cortex of the sea urchin egg may be anchored to a focal adhesion-like complex at the cell surface (Burke et al., 2004).

1.3 *Strongylocentrotus purpuratus* as a model organism

Strongylocentrotus purpuratus (S. purpuratus), the purple sea urchin, has been used as a model organism for well over a century (Briggs and Wessel, 2006; Hertwig, 1876). Although the adult sea urchin is radially symmetric, the sea urchin embryo has bilateral symmetry. The pattern of cleavage and aspects of gastrulation and mesoderm formation are similar to chordates. The embryonic features that ally the deuterostomes are strengthened by a clear molecular kinship; seventy percent (70%) of sea urchin genes have orthologues in the human genome, compared to 50% of *Drosophila melanogaster* genes and 35% of *Caenorhabditis elegans* genes (*C. elegans* Genome Sequencing Consortium, 1998; *Drosophil*a Genome Sequencing Consortium, 2000; Harada et al., 1995; Sea Urchin Genome Sequencing Consortium, 2006; Stewart et al., 2005). The sea urchin genome supported the long-held assertion that deuterostomes (echinoderms, hemichordates, and chordates) are monophyletic (Harada et al., 1995; Sea Urchin Genome Sequencing Consortium, 2006).

The *S. purpuratus* embryo is also useful as a developmental model because the adults are easy to keep and gametes are easy to obtain and handle. Once fertilization occurs, developing embryos require little special care and attention for the first week of development. During this time the embryos are easy to observe and the number of embryos that can be cultured at one time allows one to easily obtain a large number of

synchronous embryos (Poustka et al., 1999). Sea urchins have five gonads suspended in the coelomic cavity and each one is connected to a gonopore. Contractions of muscles in the gonads cause the release of eggs or sperm into the water where fertilization takes place. This muscular contraction can be induced with the injection of 0.55M KCl directly into the coelomic cavity.

1.4 Early embryonic development of *S. purpuratus*

Much is known about early sea urchin development (Carlson, 1996a; Gilbert, 2000; Wolpert et al., 2007). The *S. purpuratus* embryo undergoes first cleavage within two hours of fertilization. The first division is rapidly followed by a series of equal and synchronous cell division. Like most cleavage divisions, the cell cycle is modified to remove G1 and G2, so each mitotic division is followed by an S phase. The first two divisions are meridional, from animal pole to vegetal pole, followed by the third equatorial cleavage which is perpendicular to the first two. The fourth division is asymmetrical and four micromeres reveal the position of the vegetal pole. This is one of the first signs of polarity in the sea urchin embryo (Carlson, 1996a; Davidson et al., 1982; Gilbert, 2000). Subsequent cleavage divisions produce a set of blastomeres that by virtue of their adherence to an apical extracellular layer (the hyaline layer), form a hollow ball of cells, the blastula. During cleavage, the cells develop adherent junctions so that the blastula wall is an epithelium with apical basal polarity. By 18-20 hours after fertilization, the embryo has become a blastula and hatching of the embryo occurs around 24 hours when it becomes a free swimming larva. The mesenchyme blastula stage is characterized when the former micromeres undergo epithelial to mesenchymal transition,

the process in which epithelial cells are converted to mesenchymal cells, and migrate into the blastocoel to form the primary mesenchyme cells (PMCs) (Carlson, 1996a; Davidson et al., 1982; Gilbert, 2000). This transition is marked by a loss of adhesion of the PMCs to the surrounding epithelial cells through the down regulation of E-cadherin (Thiery and Sleeman, 2006).

Beginning at about 30 hours, the embryo begins morphogenesis and the three primary germ layers are formed. Initially mesoderm forms as loose mesenchyme cells and endoderm is a hollow tube derived from the cells surrounding the vegetal pole. During this morphogenesis, gastrulation begins. The initiation of this stage is marked by the separation of the PMCs from the cell wall to form a ring-like structure around the invaginating archenteron. Secondary mesenchyme cells form at the tip of the archenteron and extend filipodia to the opposite wall of the blastomere while the archenteron is elongating. Finally the archenteron comes in contact with the blastocoel wall near the animal pole and a full gut is formed. The anus forms at the original invagination while the mouth forms from the second opening. The embryo develops into a bilaterally symmetrical pluteus larva by five days after fertilization in which there is a plane of symmetry separating the right and left halves of the embryos. The larva undergoes metamorphosis to become a radially symmetric adult (Carlson, 1996b; Davidson et al., 1982; Gilbert, 2000; Wolpert et al., 2007). Figure 1.2 highlights the various stages of *S. purpuratus* development through the first 96 hours.

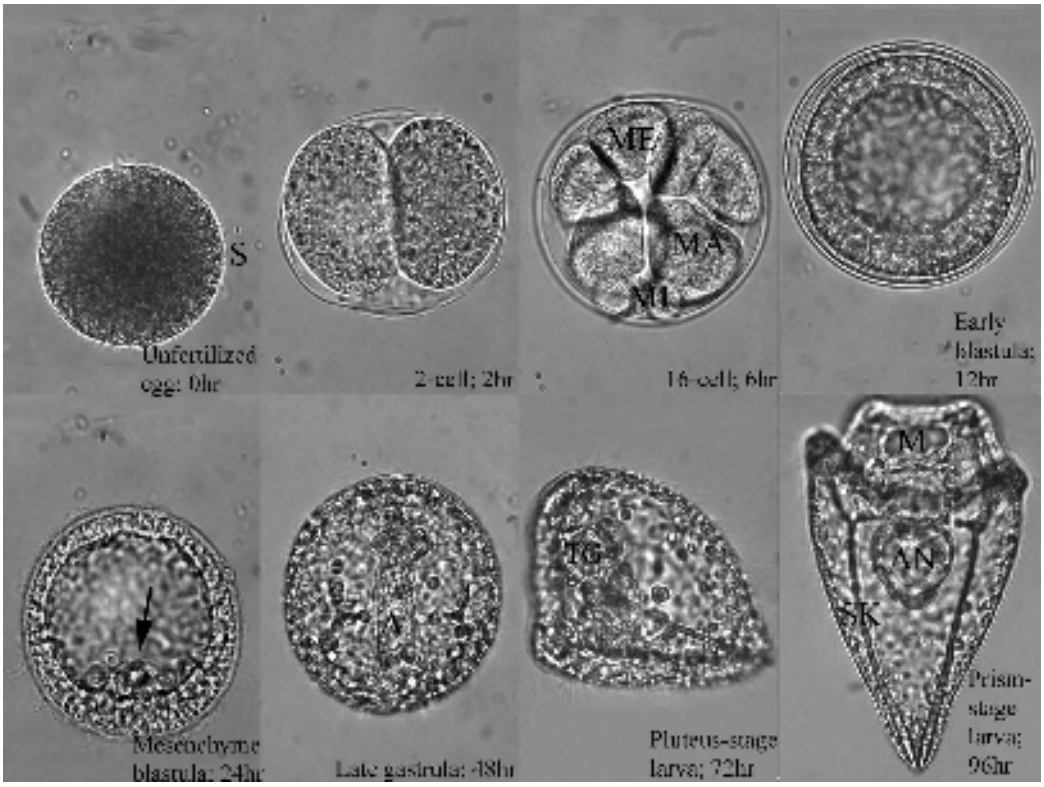


Figure 1.2 Light micrographs of developmental stages of *S. purpuratus* from unfertilized egg to prism-stage larva. Within the 16-cell embryo there are mesomeres (ME), macromeres (MA), and micromeres (MI). PMCs being to invaginate during the mesenchyme blastula stage (arrow) and an archenteron forms during gastrulation. The larva has a mouth, (M), stomach (AN), and a developing skeleton (SK).

Chapter 2 – Materials and Methods

2.1 Adult *S. purpuratus* culturing and collection of gametes

Adult *Strongylocentrotus purpuratus* were collected from Sooke, British Columbia and maintained in a photo-controlled sea water system. Spawning was initiated by intracoelomic injection of 0.55M KCl. Eggs were collected by inverting the spawning female onto a beaker of sea water. Sperm was collected from the spawning males and stored at 4 ºC until required for fertilization. Eggs were rinsed three times in filtered sea water (FSW) to remove the outer jelly coat and fertilized with sperm activated in sea water.

Embryos were cultured at an initial concentration of approximately 5000 eggs/mL of FSW. This concentration was reduced over time as the surviving embryos were removed from the original culture and diluted with more FSW. Cultures were grown in FSW with streptomycin sulphate (50 mg/L) to control bacterial growth (Vilela-Silva et al., 2001). Bacteria that grow in the cultures can infect the developing embryos, to the point that healthy growth of the cultures is inhibited.

2.2 PCR amplification of cDNA

2.2.1 Primer design

Primers were designed using gene models available for the *S. purpuratus* genome, as described in Section 1.1. There was no difficulty in designing gene-specific primers as *S. purpuratus* integrins exhibit relatively low sequence identity (~25%). Primers for all

PCR amplification were supplied by Alpha DNA (Montreal, QC). The complete list of primers used can be found in Appendix I.

2.2.2 RNA isolation and RT-PCR

Embryos for RNA extraction were collected by centrifuging embryos and lysing them immediately with the RNA extraction medium. For QPCR a total number of 5000 embryos were collected by taking average counts of small volumes of culture and collecting enough to make 5000 embryos. The number of embryos used for general PCR amplification was not standardized in order to obtain many embryos and extract a significant amount of RNA at each stage of development.

Eggs and embryos were collected for RNA isolation as unfertilized egg, 2 hours, 4 hours, 6 hours, 12 hours, 24 hours, 48 hours, 72 hours, and 96 hours after fertilization. Total RNA isolation followed the procedure of Ransick (2004), although TRIzol (Invitrogen, Catalogue No: 10296-028) was used instead of RNAzol-B. Embryos were resuspended in 200 µL of TriZol, to which 10% (v/v) chloroform was added. The samples were shaken vigorously for 15 seconds and left on ice for 5 minutes, before being centrifuged for 15 minutes at 4 ºC and at 12000 x G. The aqueous layer was removed and to it was added an equal volume of isopropanol and 5 µg glycogen. The RNA was left overnight at -20 ºC to precipitate.

The precipitated RNA was centrifuged for 15 minutes at 4 ºC and at 12000 x G. The pellet was rinsed twice in 75% ethanol and left to dry, before being resuspended in 15 µL nuclease-free H_2O. RNA concentrations were determined using an ND-1000 NanoDrop spectrophotometer.

RT-PCR used 1-5 µg of total RNA combined with 250 ng of random primers (Invitrogen, Catalogue No: 48190-011), 1 µL of 10mM (500 µM final concentration) dNTPs (Amersham Biosciences, Catalogue No: 27-2035-01), and nuclease free water to 12 µL. This solution was heated to 65 ºC for 5 minutes, followed by 2 minutes on ice. To this, 4 µL of 5X First Strand Buffer, 2 µL of 0.1M DTT, and 1 µL of RNase inhibitor (Invitrogen, Catalogue No: 15518-012) were added. The reaction was incubated at 25 ºC for 2 minutes, after which 200 units of SuperScript II (Invitrogen, Catalogue No: 18064-022) was added. The reaction was incubated for 10 minutes at 25 ºC, followed by a 60 minute incubation at 42 ºC, and a 15 minute inactivation at 70 ºC.

2.2.3 PCR

Four *S. purpuratus* integrin subunits were cloned and sequenced using cDNA from various stages of development as template. αC was cloned using 24 hour cDNA as a template, αD was completed using 2 hour cDNA, 96 hour cDNA was used for αF, and βD was amplified using 48 hour cDNA. These stages of development were used based on preliminary PCR experiments that showed each of these genes to be expressed during these stages. The PCR reactions included:

Table 2.1 Components of a standard PCR reaction

Component	Volume (in a 50 µL reaction)	Final Concentration
10X Ex Taq Buffer	5 µL	1X
25mM dNTP	4 µL	200 µM each
cDNA	2.5 µL	Varying
Forward primer	1 µL	200 – 400 nM
Reverse primer	1 µL	200 – 400 nM
5U/ µL Taq enzyme	0.25 µL	1.25 U

Takara *Ex Taq* DNA Polymerase (Fisher Scientific, Catalogue No: TAK RR001A) was used for all PCR amplifications, which were done using a MyCycler thermocycler (Bio-Rad, Catalogue No: 170-9703). The standard PCR program used was 94 ºC for 3 minutes; 94 ºC for 45 seconds, 60 ºC for 45 seconds, 72 ºC for 2 minutes (35X); 72 ºC for 8 minutes. The variability in this program included annealing temperature, which changed according to the primer set being used, and elongation time, which was set at approximately 1 minute elongation for 1 Kb of expected product. For very small products (< 300 bp), the program was 94 ºC for 3 minutes; 94 ºC for 15 seconds, 60 ºC for 15 seconds, 72 ºC for 30 seconds (35x); 72 ºC for 7 minutes. As before, the annealing temperature was changed according to the primer set being used. Primer sequences, annealing temperatures, and elongation times can be found in Appendix I.

2.2.4 Cloning and sequencing

The presence of PCR products was determined using gel electrophoresis on 1.25% (w/v) agarose gels. All isolated PCR products were ligated into pGEM-T Easy vector (Promega, Catalogue No: A1360) according to the manufacturer's protocol. Following the recovery step, cells were plated on LB agar plates containing 100 µg/mL ampicillin. Before plating, agar plates were spread with 40 µL of 5 mM stock X-gal and 4 µL of 100 mM stock IPTG for blue-white screening. LB agar plates were incubated overnight at 37 ºC. Positive white colonies were selected and grown up in LB media containing 0.1% (v/v) ampicillin at 37 ºC overnight, shaking. Cultured *E. coli* cells were PCR screened for the presence of the insert using the original PCR primers and program to amplify the insert with Taq DNA Polymerase (NEB, Catalogue No: M0273S).

Plasmids were isolated from clones that contained an insert using QIAprep Spin Miniprep Kit (Qiagen, Catalogue No: 27106) or GeneJET™ Plasmid Miniprep Kit (Fermentas, Catalogue No: K0503). A restriction digest was performed to verify the size of the insert. NotI (NEB, Catalogue No: R0189L) was used to digest the sample from the pGEM-T Easy vector, and approximately 1 µg of DNA was used in each reaction. The reactions were incubated for 2 hours at 37 ºC and correct digestion was confirmed by separation of products on a 1.25% (w/v) agarose gel. Plasmid samples were submitted for sequencing to the DNA Sequencing Facility at the Centre for Biomedical Research, University of Victoria.

2.2.5 5' RACE Amplification of αC

The 5' end of αC was amplified using 5' RACE. Tube feet from adult *S. purpuratus* were processed to isolate mRNA for the FirstChoice RLM RACE Kit (Ambion, Catalogue No: AM1700). Tube feet were processed with Aurum Total RNA Fatty and Fibrous Tissue Kit (Bio-Rad, Catalogue No: 732-6830) to isolate total RNA. The total RNA was processed using the MicroPoly(A) Purist Kit (Ambion, Catalogue No: AM1919) to obtain poly(A)RNA. The poly(A)RNA was processed according to the protocol of the FirstChoice RLM RACE Kit.

Primers designed to amplify the 5' end of αC are found in Appendix I. Reverse transcription was performed on the RNA tagged for 5' RACE according to the manufacturer's instructions. Random decamers were used, along with M-MLV Reverse Transcriptase (Ambion, Catalogue No: AM1700). The reaction was incubated for 60 minutes at 42 ºC.

Nested PCR was used to amplify the 5' RACE product with Takara Taq polymerase. The reaction mix included 1X PCR Buffer, 200 μM dNTP mix, 400 nM 5' RACE outer/inner primer, 400 nM gene specific outer/inner primer, 1 μL cDNA from the RT reaction, 1.25U of Taq, and nuclease free water to 50 μL. The program used was 94 ºC for 3 minutes; 94 ºC for 30 seconds, 60 ºC for 30 seconds, 72 ºC for 2 minutes (35x); 72 ºC for 7 minutes. Amplification of the inner 5' RACE product was identical, with the exception that the outer product was used as template. PCR amplification was done using a MyCycler thermocycler (Bio-Rad, Catalogue No: 170-9703).

2.3 Phylogenetic analysis

Protein sequences were used for phylogenetic analyses of α integrin subunits in four different phyla. Sequences similar to the confirmed *S. purpuratus* integrin subunits were identified from the NCBI database using the BLAST protein search tool (http://blast.ncbi.nlm.nih.gov/Blast.cgi). Integrin sequences were chosen to represent a variety of organisms and phyla and both protostomes and deuterostomes. These species included *S. purpuratus* (purple sea urchin), *L. variegatus* (green sea urchin), *M. musculus* (house mouse), *C. intestinalis* (sea squirt), *D. melanogaster* (fruit fly), *C. elegans* (roundworm), and *N. vectensis* (sea anemone).All sequences were run through Pfam (Finn et al., 2006) and all were found to have an integrin_α2 domain. This domain was aligned in MEGA 4.0 using ClustalW alignment and neighbour joining trees were generated using bootstrap values of 1000 (Tamura et al., 2007).

2.4 Quantitative PCR

2.4.1 Primer Design

QPCR primers were originally designed for five *S. purpuratus* integrin genes, with the final amplicon sizes of: αC – 136 bp; αD – 149 bp; αF – 150 bp; βD – 164 bp; βC – 148 bp. Ubiquitin had an amplicon size of 147 bp. Four of these genes (αC, αD, αF, and βD) were analyzed because they were novel genes whose full length cDNA sequences had been confirmed. The fifth, βC, was analyzed to confirm its expression during early cleavage stages that had been suggested by previous data (Murray et al., 2000a; Murray et al., 2000b). QPCR primers were designed so they, or the product, would span an intron/exon boundary, which reduces the possibility of amplifying a product from genomic DNA if there is any contamination. All QPCR amplicons were sequence confirmed. All primer pairs were checked against a no template control and a no enzyme control.

2.4.2 mRNA isolation and RT-PCR

For quantitative PCR studies, in each RT-PCR reaction, total RNA from 2500 embryos (half of the batch of 5000 that was originally collected) was used at stages of development including unfertilized egg, 2 hours, 12 hours, 24 hours, 48 hours, 72 hours, and 96 hours. In one set of experiments, total RNA was isolated using TriZol as described in section 2.2.2. For two subsequent experiments, mRNA was used in the RT-PCR reaction, and isolated from the embryos using the MicroPoly(A)Purist Kit (Ambion, Catalogue No: AM1919), according to the manufacturer's instructions.

Before being used for QPCR, cDNA was tested using standard PCR and TBPint primers (sequence in Appendix I). These were primers designed by Javier Tello

(University of Victoria) for the Tata binding protein in *S. purpuratus*. They were designed for the product to span an intron, giving a 283 bp product when amplified from cDNA, or an 800 bp product if the template had genomic contamination. PCR amplifications and gel electrophoresis analyses were done as described previously in sections 2.2.4 and 2.2.5. After verifying the presence of cDNA, the quality of cDNA was established using a QPCR reaction. Ubiquitin forward and reverse primers were used to test 10 fold dilutions of cDNA. Log cDNA concentration was plotted against the Ct value and the correlation of a best fit line was found. Correlations ≥0.95 were deemed to be acceptable cDNA samples.

2.4.3 QPCR

All QPCR reactions were set up in 96-well plates (Eppendorf, Catalogue No: 951022055) and run as 15 µL reactions. The fluorescent marker used was iQ Sybr Green Master Mix (Bio-Rad, Catalogue No: 170-8885) and all runs were done using an Eppendorf Mastercycler Realplex.

Table 2.2 Components of a QPCR reaction

Component	Volume (in 15 µL reaction)	Final Concentration
iQ Sybr Green Master Mix (2X)	7.5	1X
H_2O	3.2	
F1 primer (10 µM stock)	0.15	100 nM
R1 primer (10 µM stock)	0.15	100 nM
cDNA	4	Varying

Table 2.3 Nanograms of cDNA used in each QPCR reaction varied between stages

Amount of cDNA in 15 µL reaction (ng)				
	Round 1 (total RNA)	Round 2 (mRNA)	Round 2 (mRNA)	Round 3 (mRNA)
Egg	520.8	3.0	2.7	5.6
2 hours	182.8	----	4.0	5.4
12 hours	225.2	----	3.4	5.5
24 hours	534.5	7.8	----	9.8
48 hours	808.8	7.8	----	11.1
72 hours	1143.2	11.9	----	9.8
96 hours	1187.2	12.2	----	10.4

The program used for QPCR was 95 ºC for 2 minutes and 30 seconds; 95 ºC for 15 seconds, 60 ºC for 15 seconds, 68 ºC for 20 seconds (40x); followed by a 20 minute melting curve to establish the presence of a single amplicon. Data was analyzed using the software provided with the Eppendorf Mastercycler as well as Microsoft Office Excel 2007 and GraphPad Prism 4.03. Ubiquitin was used as a reference gene against the five *S. purpuratus* integrin genes. Ubiquitin is commonly used as an internal standard for QPCR analysis as it is known to be present in constant amounts during development (Howard-Ashby et al., 2006a; Howard-Ashby et al., 2006b; Nemer et al., 1991; Oliveri et al., 2002; Ransick et al., 2002).

The relative expression levels for each gene at each stage were calculated using the ΔCt method with ubiquitin as an internal standard. The first assumption for this method is that the primers for the target gene and reference gene are amplifying at the same efficiency, preferably 100%. This assumption was satisfied by amplifying each gene along with ubiquitin by QPCR using 2 fold cDNA dilutions and plotting the Ct values (y) against log of cDNA concentration (x) and determining the slope of the best fit line. The efficiency of the primer was calculated using the formula $10^{(-1/\text{slope})}$-1. The

combined primer efficiency of the gene in question and the normalizer gene (ubiquitin) was calculated to determine relative expression levels (discussed below).

The second assumption to use this method of relative quantification that was satisfied was that the primers were amplifying at similar efficiencies over various dilutions. ΔCt values were calculated at the various dilutions using the formula $\Delta Ct = Ct_{gene} - Ct_{ubiquitin}$. The ΔCt values (y) were plotted again the log of cDNA concentration (x) and the slope of the best fit line was determined. Ideally the slope of the line would be 0, although slopes under 0.1 were accepted.

Relative expression levels were determined using the formula $X^{-\Delta Ct}$ (where X = primer efficiency $_{gene}$ + primer efficiency $_{ubiquitin}$). The base values for each gene were: αC = 1.93, αF = 1.96, αD = 1.92, βD = 1.94, and βC = 1.92. The values for each gene were multiplied by a common factor so the relative expression level of the egg was 1.

2.5 Microinjection of Sp-αC MASO

The sequence of the Sp-αC MASO can be found in Appendix I. It comprises -25 to -1 of the αC sequence.

2.5.1 Injection of embryos

S. purpuratus gametes were obtained as described previously in section 2.1. Following the third rinse in filtered sea water, eggs to be used for microinjections were filtered 3–5 times through an 80 μm filter to completely remove the jelly coat and to allow the eggs to stick to the microinjection dishes.

Needles used for injection were made from thin-walled single filament glass capillaries (World Precision Instrument, Inc., Catalogue No: TW100F-4) using a micropipette puller (Sutter Instrument Co. Flaming/Brown Micropipette puller model P-97). Needles containing the Sp-αC morpholino antisense oligonucleotide (MASO) were loaded at a concentration of 200 or 300 μM in 6.76% (v/v) glycerol and 84 mM KCl, resulting in a final MASO concentration in the egg of 2 or 3 μM. The control morpholino, ZF-Chordin, was loaded in an identical concentration. All solutions were filtered through a 0.2 μm RNase-free microfuge filter before being loaded into needles. Microinjection dishes were prepared with a strip of 1% (w/v) protamine sulphate to allow the eggs to adhere to the dish. Unfertilized eggs were lined up in a row along the protamine sulphate and injected using a Picospritzer II (General Valve Corporation) injector and MMN-1 (Narishige) manipulator. Injections were done primarily by Navraj Chima. Injected eggs were fertilized in 1mM ATA in FSW. Approximately 30 minutes post-fertilization, the eggs were rinsed in FSW to dilute the ATA.

2.5.2 Immunofluorescence of injected embryos

Embryos were collected at 2 hours, 6 hours, 24 hours, and 48 hours post fertilization for staining with anti-Sp-Par6, a polyclonal rat antibody, or 2D2, a mouse monoclonal antibody. Embryos were fixed in ice cold 100% methanol for 20 minutes, followed by 3x 15 minute washes in 1X PBS. Blocking was done for 30 minutes in 5% lamb serum in PBS-T and the embryos were incubated in primary antibody (diluted in 5% lamb serum in PBS) overnight at 4 ºC. For anti-Sp-Par6, a 1:500 dilution of primary antibody was used. A 1:800 dilution was used for 2D2.

Following incubation in primary antibody, the embryos were rinsed 3 times in 1X PBS. Incubation in secondary antibody followed for 2 hours. Secondary antibody was either goat-anti-rat (Alexa Fluor 488, Invitrogen Molecular Probes, Catalogue No: A-11066 or Alexa Fluor 568, Invitrogen Molecular Probes, Catalogue No: A-11077) or goat-anti-mouse (Alexa Fluor 488, Invitrogen Molecular Probes, Catalogue No: A11029 or Alexa Fluor 568, Invitrogen Molecular Probes, Catalogue No: A11031). Alexa Fluor 488 antibodies were diluted 1:900, while Alexa Fluor 568 antibodies were diluted 1:1500. After incubation of the secondary antibody, embryos were incubated for 5 minutes with 1:3000 DAPI, followed by 3x 15 minute washes in 1X PBS. They were imaged on a Leica CTR6000 fluorescence microscope using OpenLab software. Images were cropped and adjusted for brightness/contrast using Adobe Photoshop 6.0.

Fixation was slightly different for embryos stained for F-actin localization. Embryos were collected at 6 hours and 24 hours after fertilization and fixed for 8 minutes at room temperature in 4% (w/v) Paraformaldehyde/Tris in FSW (50 mM Tris, pH 7.4) for actin staining. The embryos were washed twice with 1X PBS and blocked (PBS-T with 5% lamb serum) for 20 minutes. After another wash with 1X PBS, embryos were incubated for 20 minutes in a 1X solution of Alexa Fluor 594 Phalloidin (Invitrogen Molecular Probes, Catalogue No: A12381). One more wash in 1X PBS followed and imaging was done as described previously.

Statistical comparisons between injected and uninjected embryos were done with a 2 by 2 contingency table and Fisher's exact tests and chi-squared tests using Graphpad Prism (version 4.03).

Chapter 3 – Results

3.1 Sequencing results and protein domains

3.1.1 αC

Primers were designed for the αC sequence based on four GLEAN predictions (Baylor College of Medicine [BCM]) and one prediction from Scaffold_V2 GENSCAN (National Centre for Biotechnology Information [NCBI]). The full length cDNA sequence of αC was deduced from six overlapping clones. An EST clone obtained from Charles A. Ettensohn (Zhu et al., 2001) contained 1478 bp at the 3' end of the cDNA. Four clones were the product of PCR amplification and they extended the sequence to within 887 bp of the 5' end of the sequence. 5' RACE was used to amplify the 5' region of αC (Fig. 3.1b). The cDNA sequence consolidated six genomic predictions into a single cDNA indicating that the predictions are from incomplete fragments of a single gene. The cDNA sequence varied from the predicted sequence at several locations (Table 3.1).

Table 3.1 Tabular notes comparing the regions of αC cDNA with the predictions, detailing regions of consistency, dissimilarity, and exons that vary between the sequences (Fig 3.1a).

	Base pair location of each feature within αC cDNA sequence			
Prediction	Exact matches	Inconsistent sequences	Exons not predicted but found in αC cDNA sequence	Predicted exons not found in confirmed sequence
Scaffold_v2_32336_1	893-1648 2098-2900 2901-3315	840-893 1649-2097	None	2900-2901
Scaffold_v2_32336_2	198-800	801-857	None	None
GLEAN3_15378	1650-1980	1255-1649 1981-2037	None	None
GLEAN3_00547	2213-3515	2059-2212	None	None
GLEAN3_15377	387-801 893-1992 2047-3315	802-892 1993-2046	None	None
GLEAN3_15379	2046-3051		None	None

The cDNA sequence has a single open reading frame that predicts a 120.1 kD protein 1105 amino acids in length (Appendix II). The prediction contains typical features of an integrin: a signal peptide, beta-propeller repeats, a cation binding site, a transmembrane domain, and a short cytoplasmic domain. SignalP 3.0 (http://www.cbs.dtu.dk/services/SignalP/) (Bendtsen et al., 2004) predicts a signal peptide in the protein sequence (P = 0.989) that cleaves between amino acids 24 and 25 (P = 0.971). Pfam and SMART analyses identify five beta-propeller repeats, an integrin_α2 domain, which is the leg region of the integrin found from 485 – 961 of the predicted amino acid sequence (Fig. 3.1c – sequence is underlined in Appendix II), a transmembrane domain, and a cytoplasmic domain that is 37 amino acids long (Finn et al., 2006; Letunic et al., 2006; Schultz et al., 1998). The cytoplasmic tail of αC contains the conserved amino acid sequence K/R-R-E/D from 1068 – 1095 of the predicted

protein sequence. BLAST searches of the non-redundant protein data base suggest that αC is most similar to α6 in *M. musculus* with 35% sequence identity.

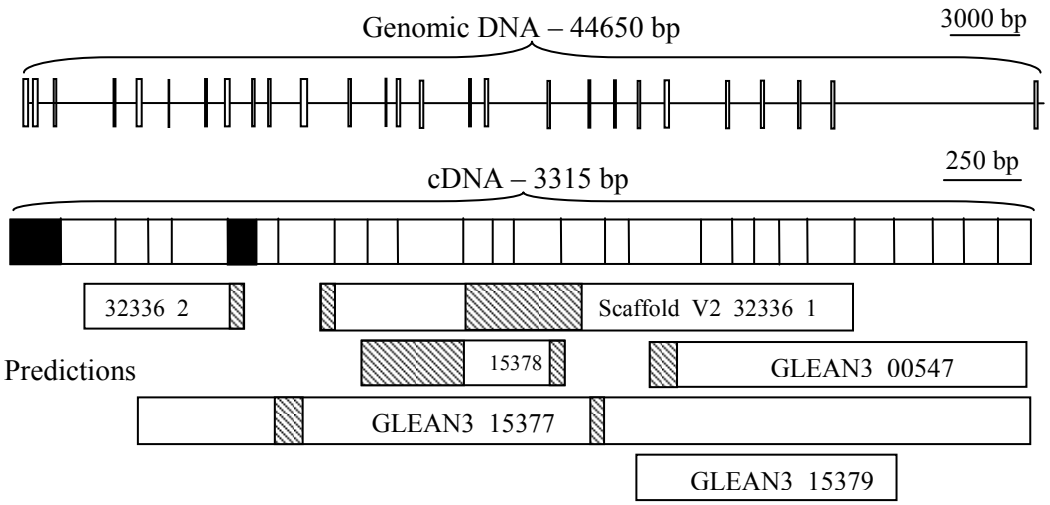


Figure 3.1a Schematic of genomic DNA, cDNA, and predictions used in obtaining the full length sequence of αC. Black regions represent unconfirmed exons. Shaded regions represent discrepancies between predicted and confirmed cDNA sequences.

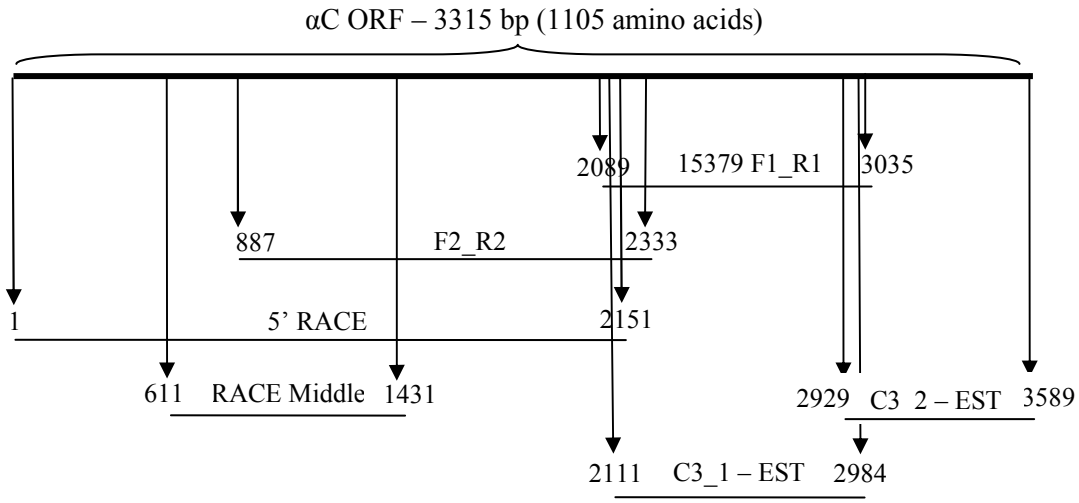


Figure 3.1b Schematic diagram of regions amplified from cDNA to obtain the full length sequence of αC. C3_1-EST and C3_2-EST refer to an expressed sequence tag clone from a previously existing cDNA library.

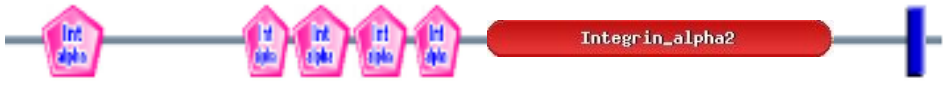


Figure 3.1c Hybridized SMART and Pfam domain predictions for the αC protein sequence (Finn et al., 2006; Letunic et al., 2006). The Int_alpha represent beta-propeller repeats. The integrin_alpha2 domain is one recognized within all α subunits and constitutes the leg region of the integrin (Xiong et al., 2001).

Figure 3.1 Schematic diagrams of the cDNA sequence and predicted protein domains of αC.

3.1.2 αF

Cloning and sequencing of αF was completed by Kate MacDonald. Four incomplete gene predictions were used to deduce the full length sequence of αF; a GLEAN prediction (BCM), 2 models from Scaffold_V2 (NCBI), GENSCAN and BCM Ensemble: CDS, and 1 model from Scaffold_V2.1 (NCBI). The full length cDNA sequence of αF was deduced from PCR amplified clones and two EST clones, (obtained from Dr. James A. Coffman, Mount Desert Island Biological Laboratory) (Fig. 3.2b). The cDNA sequence consolidated four genomic predictions into a single cDNA indicating that the predictions are from incomplete fragments of a single gene. The cDNA sequence varied from the predicted sequence at several locations (Table 3.2).

Table 3.2 Tabular notes comparing the regions of αF cDNA with the predictions, detailing regions of consistency, dissimilarity, and exons that vary between the sequences (Fig 3.2a).

	Base pair location of each feature within αF cDNA sequence			
Prediction	Exact matches	Inconsistent sequences	Exons not predicted but found in αF cDNA sequence	Predicted exons not found in αF cDNA sequence
Scaffold_v2_ GENSCAN_16313_11	1898-2253	None	None	None
Scaffold_v2_ BCM_26279	74-464 884-1196 1299-1540 1595-2173 2338-2884	465-883 1197-1298 1541-1594 2174-2337 2885-2984	None	None
GLEAN3_15920	1367-3232	None	None	None
Scaffold_v2.1_ LOC581907	231-601 884-1207 1281-3232	137-231	602-883 1208-1280	None

The αF sequence has a single open reading frame that encodes a 117.3 kD protein 1076 amino acids in length (Appendix II) that contains the typical components of an

integrin: a signal peptide, beta-propeller repeats, a cation binding site, a transmembrane domain, and a short cytoplasmic domain. SignalP 3.0 (Bendtsen et al., 2004) predicts a signal peptide (P=0.997) with a cleavage site between amino acids 24 and 25 (P=0.995). The protein domains recognized by Pfam and SMART include 5 beta-propeller repeats, an integrin_α2 domain from 460 to 911 of the predicted sequence (Fig. 3.2c – sequence is underlined in Appendix II), a transmembrane domain, and a 37 amino acid cytoplasmic domain (Finn et al., 2006; Letunic et al., 2006; Schultz et al., 1998). The highly conserved amino acid sequence K/R-R-E/D is found in the cytoplasmic domain of the αF protein sequence from 1020 - 1033. BLAST searches of the non-redundant protein data base indicated that αF is most similar to α8 in *M. musculus* with 31% sequence identity and it has sequence identity with other *S. purpuratu*s α integrins ranging from 28% to 38%.

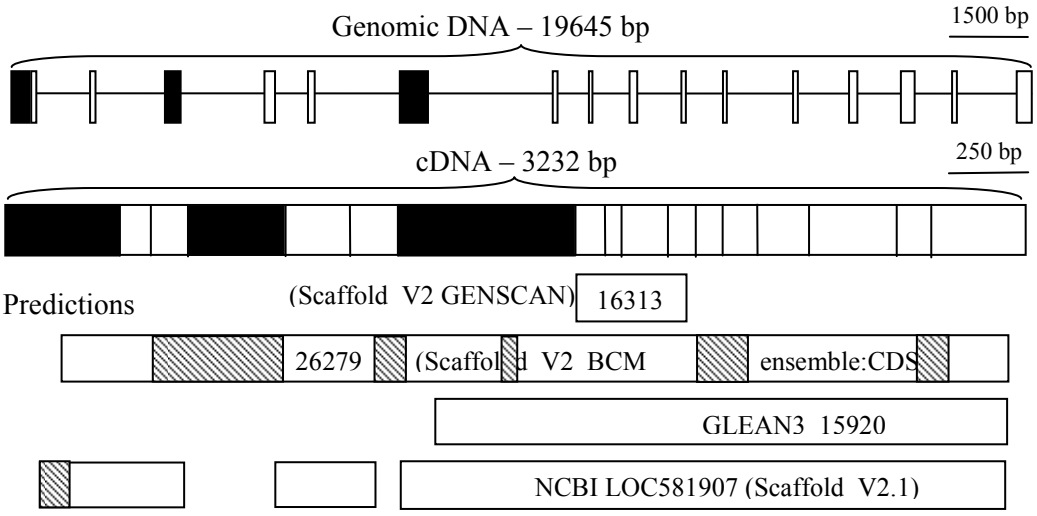


Figure 3.2a Schematic of genomic DNA, cDNA, and predictions used in obtaining a full length sequence of αF. Black regions represent unconfirmed exons. Shaded regions represent discrepancies between predicted and confirmed cDNA sequences.

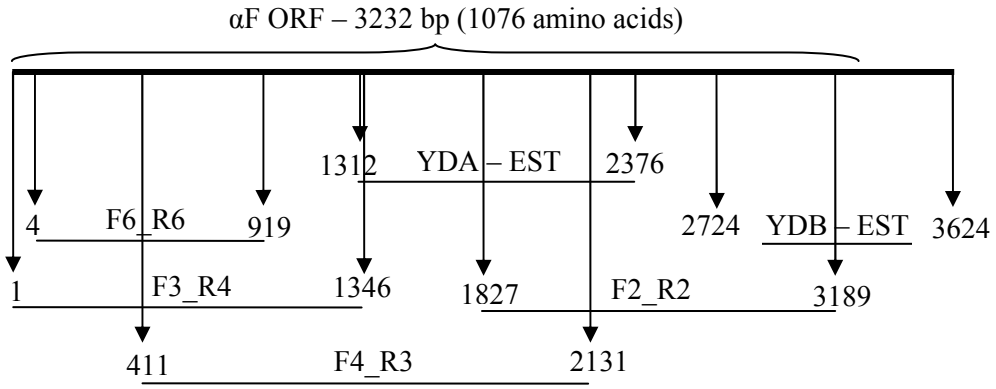


Figure 3.2b Schematic diagram of regions amplified from cDNA to obtain the full length sequence of αF. YDA-EST and YDB-EST refer to an expressed sequence tag clone from a previously existing cDNA library.

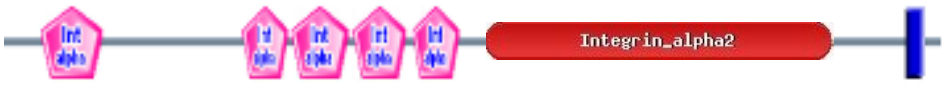


Figure 3.1c Hybridized SMART and Pfam domain predictions for the αF protein sequence (Finn et al., 2006; Letunic et al., 2006). The Int_alpha represent beta-propeller repeats. The integrin_alpha2 domain is one recognized within all α subunits and constitutes the leg region of the integrin (Xiong et al., 2001).

Figure 3.2 Schematic diagrams of the cDNA sequence and predicted protein domains of αF.

3.1.3 αD

Cloning and sequencing of αD was completed by Navraj Chima. Four gene predictions were used for design of primers for αD: a GLEAN prediction (BCM), two from Scaffold_V2 GENSCAN (NCBI), and one from Scaffold_V2 GNOMON (NCBI). The full length cDNA sequence of αD was assembled from six overlapping PCR amplified clones (Fig. 3.3b). The assembled cDNA sequence varied from the sequence predictions in having three unpredicted exons and three regions in which the sequence varied from the predicted sequence (Table 6).

Table 3.3 Tabular notes comparing the regions of αD cDNA with the predictions, detailing regions of consistency, dissimilarity, and exons that vary between the sequences (Fig 3.3a).

	Base pair location of each feature within αD cDNA sequence			
Prediction	Exact matches	Inconsistent sequences	Exons not predicted but found in αD cDNA sequence	Predicted exons not found in αD cDNA sequence
Scaffold_v2_ 16313_6	473-694 829-3113	1-472 695-828	None	None
Scaffold_v2_ 16313_7	1-365	366-498	None	None
GLEAN3_15921	4-733 833-3011	None	734-832	None
Scaffold_v2_ GNOMON_235439	4-418 623-733 833-3113	None	419-622 734-832	None

The αD cDNA sequence consists of a single open reading frame that encodes a 112.5 kD protein 1028 amino acids long (Appendix II) and contains the typical properties of an integrin: a signal peptide, beta-propeller repeats, a cation binding site, a transmembrane domain, and a short cytoplasmic domain. SignalP 3.0 (Bendtsen et al., 2004) predicts a signal peptide for αD (P = 0.989) with a predicted cleavage site between

amino acids 25 and 26 (P = 0.824). SMART and Pfam predict the domains found in the resulting protein to include five beta-propeller repeats, an integrin_α2 domain from 486 – 903 of the predicted amino acid sequence (Fig. 3.3c – underlined in Appendix II), a transmembrane domain, and a cytoplasmic domain that is 14 amino acids long (Finn et al., 2006; Letunic et al., 2006; Schultz et al., 1998). This is quite small compared to other α cytoplasmic domains and with no stop codon found in the cytoplasmic domain, there are approximately 26 amino acids, or 78 nucleotides, that remain to be confirmed at the 3' end of the sequence. Scaffold_v2_16313_6 is a prediction that may allow for complete sequence confirmation as it aligns with αD at the 3' end and extends beyond the cDNA that has been sequence confirmed. The highly conserved amino acid sequence K/R-R-E/D is found in the cytoplasmic domain of the αD protein sequence from 1014 – 1026. BLAST searches of the non-redundant protein database suggest that αD is most similar to α8 in *M. musculus* with 29% sequence identity. αD has sequence identity with other *S. purpuratus* α integrins ranging from 25% to 34%.

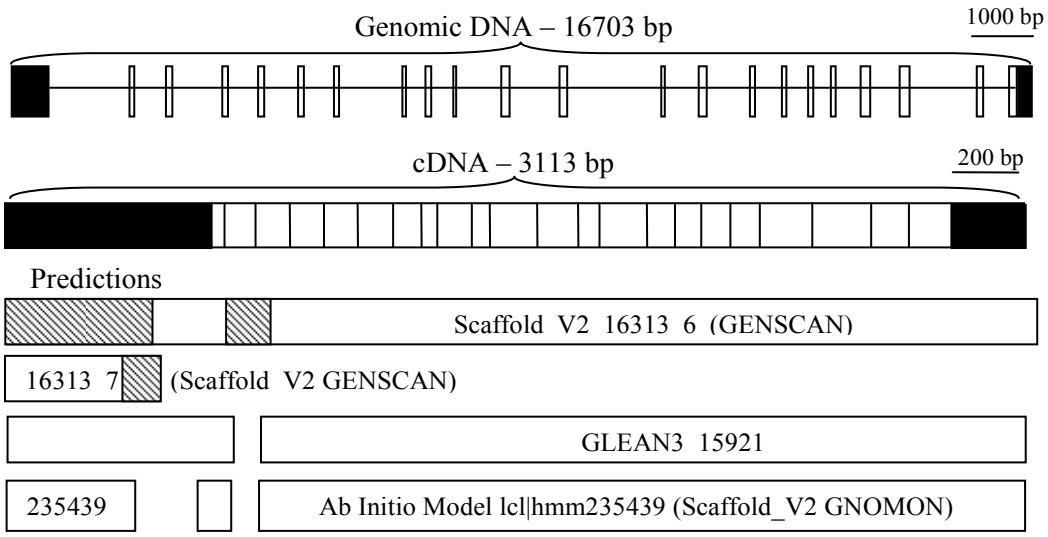


Figure 3.3a Schematic of genomic DNA, cDNA, and predictions used in obtaining a full length sequence of αD. Black regions represent unconfirmed exons. Shaded regions represent discrepancies between predicted and confirmed cDNA sequences.

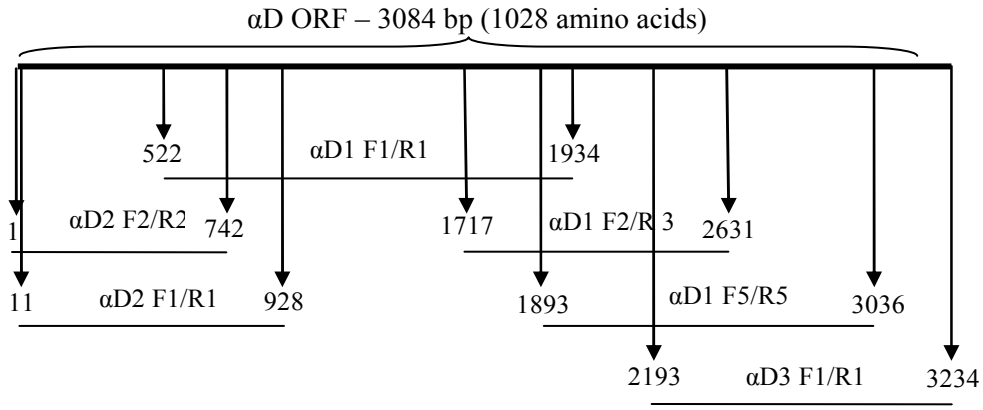


Figure 3.3b Schematic diagram of regions amplified from cDNA to obtain the full length sequence of αD.

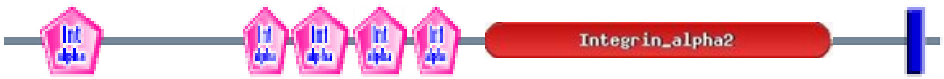


Figure 3.1c Hybridized SMART and Pfam domain predictions for the αD protein sequence (Finn et al., 2006; Letunic et al., 2006). The Int_alpha represent beta-propeller repeats. The integrin_alpha2 domain is one recognized within all α subunits and constitutes the leg region of the integrin (Xiong et al., 2001).

Figure 3.3 Schematic diagrams of the cDNA sequence and predicted protein domains of αD.

3.1.4 βD

Cloning and sequencing of βD was completed by Navraj Chima. Primers to amplify the cDNA sequence were designed from four predictions: two GLEAN predictions (BCM), one prediction from Scaffold_V2 GENESCAN (NCBI), and one prediction from Scaffold_V2.1 (NCBI). The cDNA sequence of βD was deduced from five overlapping PCR clones (Fig. 3.4b) and varies from the predicted sequence in having seven unpredicted exons (Table 7).

Table 3.4 Tabular notes comparing the regions of βD cDNA with the predictions, detailing regions of consistency, dissimilarity, and exons that vary between the sequences (Fig 3.4a).

	Base pair location of each feature within βD cDNA sequence			
Prediction	Exact matches	Inconsistent sequences	Exons not predicted but found in βD cDNA sequence	Predicted exons not found in βD cDNA sequence
Scaffold_v2_15_8	151-566 690-828 900-1489 1592-2433	1-150	567-689 829-899 1490-1591	None
GLEAN3_12985	964-1800 2020-2433	None	1801-2019	None
GLEAN3_25755	1-504 658-828	829-1457	505-657	None
Scaffold_v2.1_ LOC587915	1-566 690-828 910-2436	None	567-689 829-909	None

The sequence consists of a single open reading frame of 2427 bp that encodes an 88.2 kD, 809 amino acid polypeptide (Appendix II). The predicted protein has integrin beta domain, a transmembrane domain, and a cytoplasmic tail of 44 amino acids (Fig. 3.4c) (Finn et al., 2006). SignalP 3.0 (Bendtsen et al., 2004) predicts a signal peptide (P = 0.998) that cleaves between amino acids 22 and 23 (P = 0.987). There is a conserved

metal ion binding domain (DXSXS) and 56 cysteine residues in the extracellular domain that align with the cysteine residues in the other *S. purpuratus* β subunits (underlined in Appendix II). BLAST searches of the non-redundant protein data base suggest that βD is most similar to β1 in *M. musculus* with 41% sequence identity. The deduced βD protein has sequence identity with other *S. purpuratus* β integrins ranging from 40% to 47%.

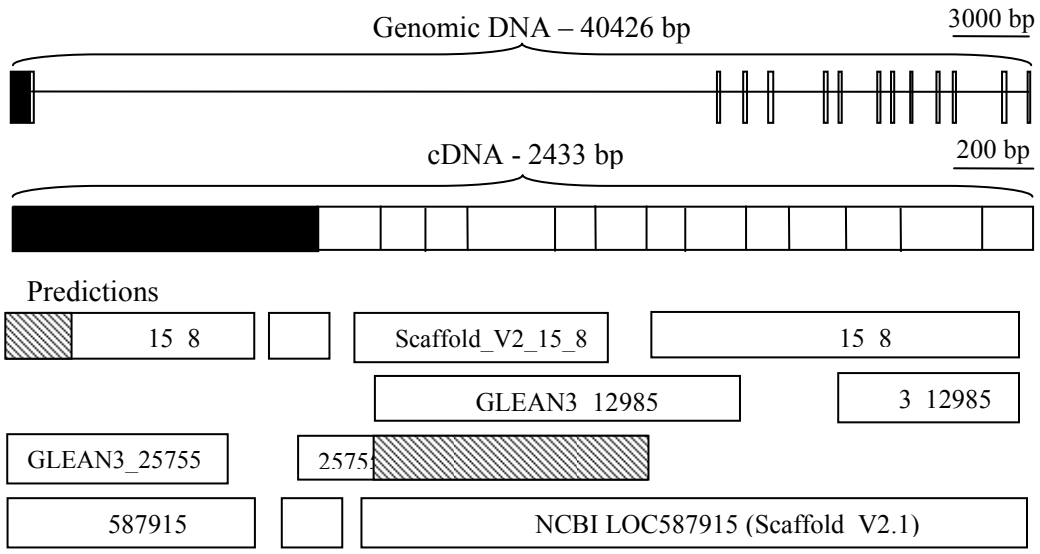


Figure 3.4a Schematic of genomic DNA, cDNA, and predictions used in obtaining a full length sequence of βD. Black regions represent unconfirmed exons. Shaded regions represent discrepancies between predicted and confirmed cDNA sequences.

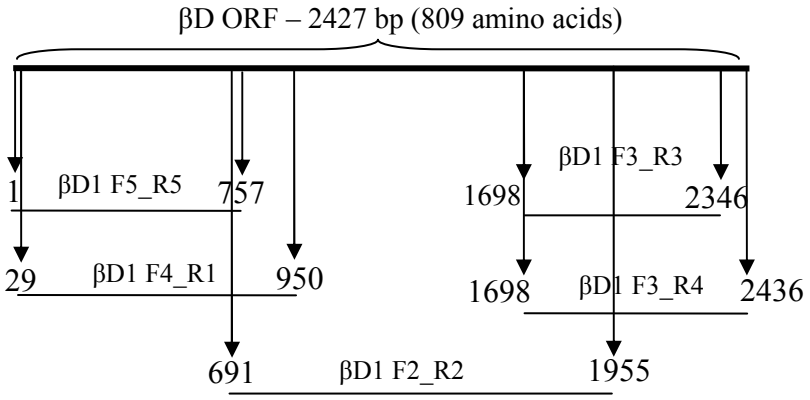


Figure 3.4b Schematic diagram of regions amplified from cDNA to obtain the full length sequence of βD.

Figure 3.4c Pfam domain prediction for the βD protein sequence. The integrin_beta domain represents a cysteine-containing domain. The red region is the transmembrane region and the yellow is the cytoplasmic domain (Finn et al., 2006).

Figure 3.4 Schematic diagrams of the cDNA sequence and predicted protein domains of βD.

3.2 Phylogenetic Analysis

Five *S. purpuratus* α integrin subunit sequences were analyzed, in addition to one sequence from *L. variegatus*, seven sequences from *M. musculus*, two sequences from *C. intestinalis*, two sequences from *C. elegans*, two sequences from *D. melanogaster*, and one sequence from *N. vectensis*. Several sequences from the *S. purpuratus* genome were not used, αG, αH, and αK, as there was not enough confirmed cDNA sequence to predict an integrin_α2 domain. A full length cDNA sequence for αJ does not yet exist, but the predicted protein from the αJ cDNA that has been confirmed contains the integrin_α2 domain. Eleven more integrin subunit sequences from the *M. musculus* genome exist, but the ones used were chosen because they had BLAST hits with some similarity to the *S. purpuratus* sequences.

Pfam software (version 22.0) identified an integrin_α2 domain within each α subunit (http://pfam.sanger.ac.uk/). This domain ranges from 404 amino acids to 528 amino acids in size and is found in the extracellular domain within the leg region of the integrin, C-terminal to the beta-propeller repeats (Xiong et al., 2001). A ClustalW alignment (Appendix II) shows the percent identity of this conserved domain between all of the sequences to be 25%, which could not be increased with minor adjustments. The *S. purpuratus* sequences of the integrin_α2 domain were aligned in ClustalW and found to have 25% amino acid identity, although the sequence was adjusted by cropping until identity was 32% (Appendix II). Whether or not the integrin_α2 domain, or even the α integrin sequences are homologous cannot be confirmed as the sequence identities are low. However, as this domain is found in all α subunits, it has the potential to be a homologous region and is the most useful for a phylogenetic study.

The α integrin phylogenetic tree shows lineage-specific diversifications that cluster as two major groups. However, the low bootstrap values prevent the relationship between these groups from being resolved (Fig. 3.5). There is representation within each group for all of the taxa included in the analysis. Organisms with only two α subunits, *C. elegans* for example, have one subunit in each of the groups. The majority of *S. purpuratus* α subunits are found with the RGD binding subunits, and αC groups with the laminin binding subunits (Fig. 3.5). When the *S. purpuratus* α integrin protein sequences are compared to each other, αC diverges from the other subunits while αD and αJ appear to be the most closely related (Fig. 3.6).

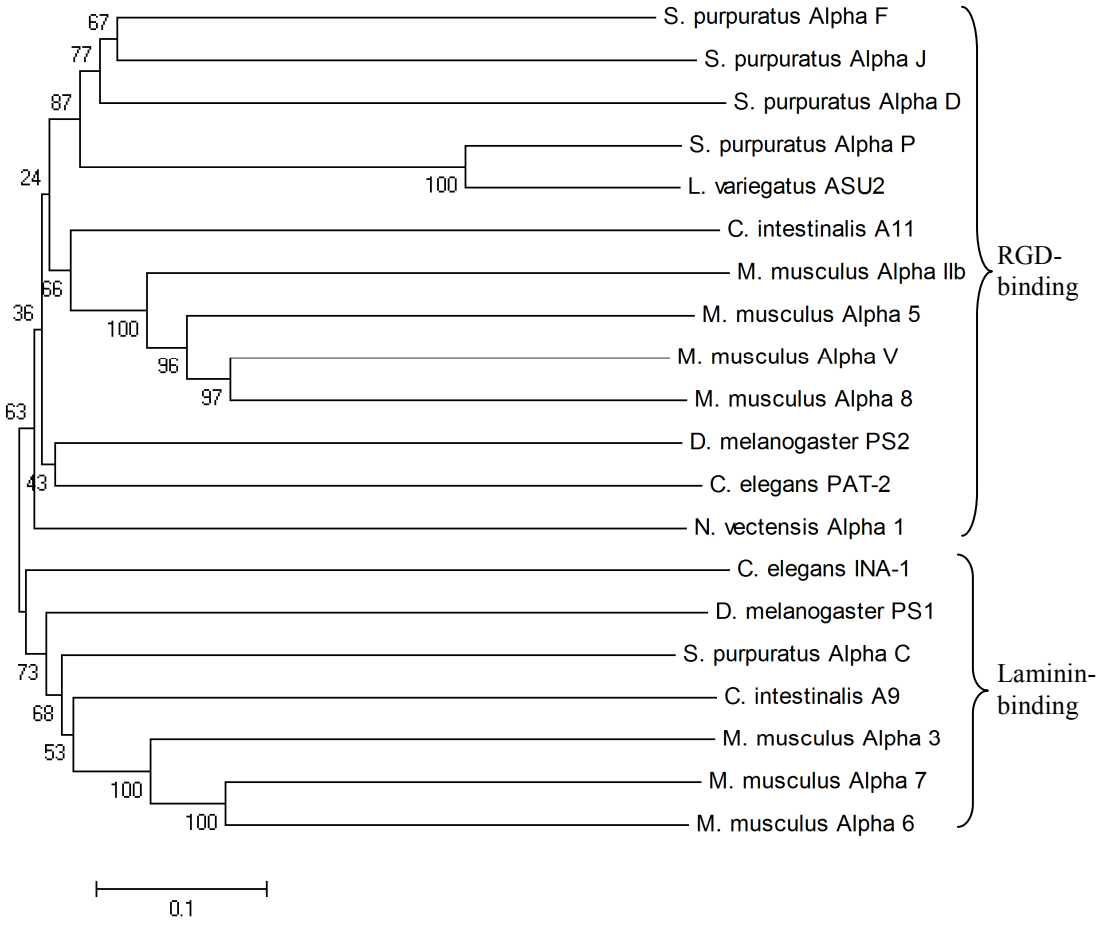


Figure 3.5 Neighbour-joining tree of the integrin_α2 domain protein sequences constructed using MEGA 4.0. Bootstrap values were obtained by 1000 replicates and are reported at the nodes as percentages. The branch lengths are proportional to the divergence between two nodes, and the scale bar corresponds to 0.1 amino acid substitution per site. The α subunits separate into two functional groups, those that bind RGD and those that bind laminin. Most of the *S. purpuratus* α subunits group with those that bind RGD, with the exception of αC which is more similar to the laminin binding subunits.

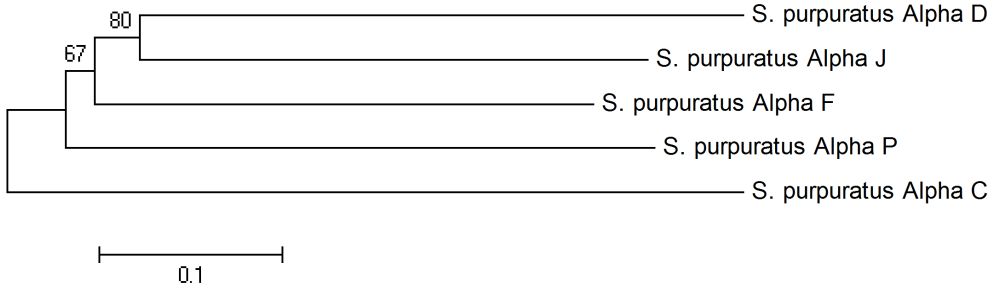


Figure 3.6 Neighbour-joining tree of the integrin_α2 domain in *S. purpuratus* α integrin protein sequences constructed using MEGA 4.0. Bootstrap values were obtained by 1000 replicates and are reported at the nodes as percentages. The branch lengths measure the amount of divergence between two nodes, and the scale bar corresponds to 0.1 nucleotide substitutions per site. αC has diverged from the other *S. purpuratus* α subunits, although the relationship cannot be determined.

3.3 Temporal Expression

QPCR data indicate variation in αC message abundance during cleavage and throughout the first 24 hours, or until the mesenchyme blastula stage. Relative levels of mRNA decrease to about 25% of that found in the egg throughout the rest of embryonic development and the overall fold change is never more than twice what is expressed in the unfertilized egg (Fig. 3.7).

The αF mRNA abundance increases by successive steps as the embryo develops. By 72 hours post-fertilization, when the embryo reaches the prism stage, the expression of αF increases by approximately 35 times compared to the unfertilized egg (Figure 3.8). Between 72 and 96 hours of development there is an apparent decrease in the relative abundance of αF.

mRNA abundance of αD remains almost constant during the first 12 hours of development. There appears to be a decrease during gastrulation (24 – 48 hours) and a return to levels similar to the unfertilized egg during formation of the larva, although there is never more than a two-fold change in expression (Fig. 3.7).

QPCR data shows relative βD mRNA abundance is almost constant during the first 96 hours of development. There is a decrease in expression 2 hours after fertilization, but within the next 10 hours this recovers to levels similar to the unfertilized egg. There appears to be another decrease by the time the embryo reaches the pluteus larva stage at 96 hours (Fig. 3.7). While it fluctuates minimally, overall expression stays close to that of the unfertilized egg and never experiences more than a two-fold change.

Expression of βC mRNA does not change during the first 12 hours of development and expression decreases to approximately 25% of that in the unfertilized egg beginning at 24 hours and for the remainder of development (Fig. 3.7).

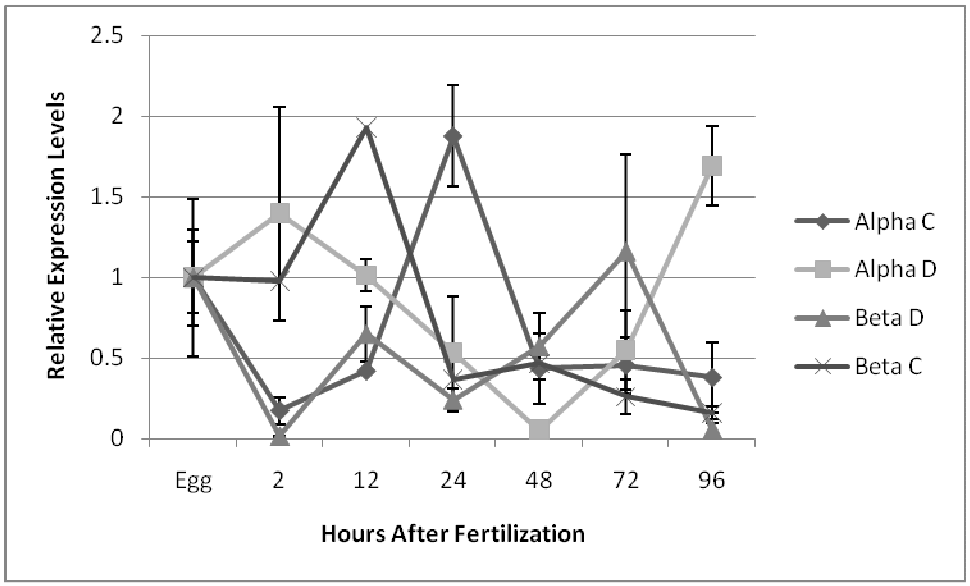


Figure 3.7 Relative expression levels of αC, αD, βD, and βC obtained by QPCR data during the first 96 hours of development of *S. purpuratus* embryos. All values are standardized against a base value of 1 for expression in the unfertilized egg. Although minor fluctuations of the genes occur during embryogenesis, the abundance of message never increases by more than two-fold, resulting in no significant change in expression. Error bars are generated from three separate experiments, with the exception of several points of βC that only had one experiment and therefore no error bars.

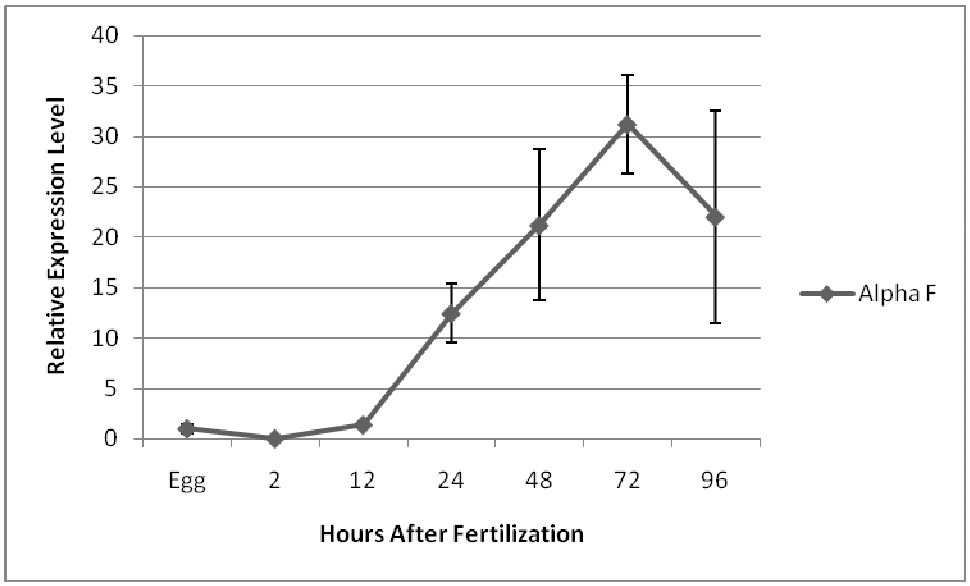


Figure 3.8 Relative expression levels of αF obtained by QPCR data during the first 96 hours of development of *S. purpuratus* embryos. All values are standardized against a base value of 1 for expression in the unfertilized egg. The abundance of αF mRNA increases dramatically during embryogenesis. Error bars are generated from three separate experiments. For the first three stages the error bars are negligible.

3.4 Anti-sense morpholino knockdown of αC

3.4.1 αC is necessary for early development

It is important to note that one significant control is missing from this morpholino knockdown experiment. Uninjected embryos and embryos injected with an irrelevant morpholino were used to make direct comparisons with embryos injected with Sp-αC MASO. However, a control to determine if the Sp-αC MASO targets αC specifically was not included. This control could be done by co-injecting the Sp-αC MASO and the full length αC sequence that is lacking the morpholino target sequence. As well, an

antibody against the αC protein could be used to determine if Sp-αC MASO injected embryos have reduced expression of the αC protein. This control should be completed before any conclusions of the function of αC can be reached.

Between 0 and 6 hours, there were no differences in the number of dead embryos injected with Sp-αC MASO or the control injected embryos. By 24 hours, 57.1% (N=373) of the Sp-αC MASO injected embryos are dead compared to 22.6% (N=548) for uninjected controls, and 12.8% (N=336) for control injected embryos. By 48 hours, 71.7% (N=265) of the Sp-αC MASO injected embryos are dead and 20.1% (N=309) of the uninjected embryos or 17.5% (N=286) of the control injected embryos are dead. Embryos injected with Sp-αC MASO experience a significantly higher mortality rate than the control groups during the first 24 hours of development (Fig. 3.8).

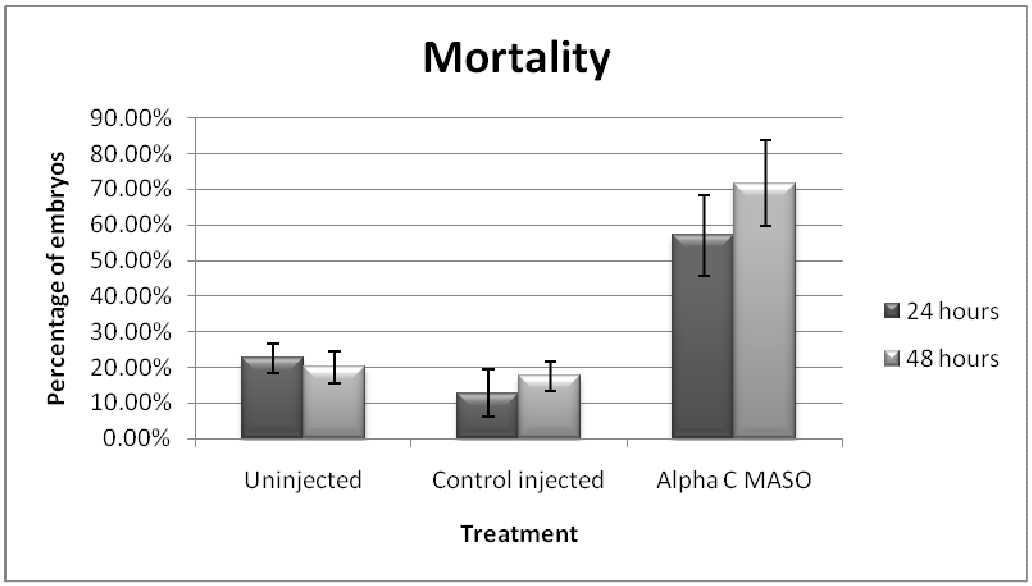


Figure 3.9 Knockdown of αC expression results in increased mortality. *S. purpuratus* embryos injected with a morpholino to the αC integrin subunits have a significantly higher proportion of dead embryos (57.1%, N=373) within the first 24 hours of development compared to the injected control (12.8%, N=336), which has a similar mortality to uninjected embryos (22.6%, N=548). Mortality in the three populations do not significantly change between 24 and 48 hours.

3.4.2 Abnormalities of cleavage-stage embryos

In some MASO injected embryos, blastomeres are not uniform in size and shape (Fig. 3.9). Six hours post-fertilization, normal embryos are typically at the 8- to 16-cell stage and cells within a tier are uniform in size and shape. Overall, 5% (N=154) of embryos injected with Sp-αC MASO had a defect in cell size and shape compared to 3% (N=96) of the injected controls and 0.3% (N=271) of the uninjected controls. During the first few hours of development cells adhere to each other and to the surrounding hyaline layer. Of the embryos injected with the Sp-αC MASO, 2% (N=154) have non-adherent cells or cells that are loose within the hyaline layer (Fig. 3.18). However, 2% (N=96) of

the control injected and none of the uninjected controls (N=271) had similar defects. The proportion of control injected and experimental injected embryos with defects in blastomere size or adhesion are very similar, suggesting that this may be an artefact of injection rather than a decrease in translation of αC.

During the first cleavages, embryos should have 2, 4, or 8 cells, but occasionally embryos with 3 or 5 cells are found (Fig. 3.9). Approximately 9.7% (N=154) of the Sp-αC MASO embryos have this defect, whereas 4.2% (N=96) of the control MASO injected embryos and none of the uninjected controls have abnormal cell numbers (N=271).

Some of the embryos with reduced translation of αC have unequal and incomplete cytokinesis in which the cleavage furrow forms from one pole and invaginates only part way (Fig. 3.9). There are also phenotypes in which cytokinesis does not occur at all as there is no distinction between individual cells in the embryo. Mitosis does occur in these embryos as DAPI staining indicates the presence of several nuclei (Fig. 3.10). Of MASO injected embryos, 4.5% (N=154) show these cytokinesis defects while none of the control injected (N=96) or uninjected (N=271) have the same defect. Overall, 14.3% (N=154) of Sp-αC MASO have an unusual number of cells, cytokinetic defects, or both, compared to 4.2% (N=96) of the control injected and none (N=271) of the uninjected controls. Using Fisher's exact test of the data in a 2 X 2 contingency table, the difference between the two treatements is not statistically significant (p=0.1415). Embryos, injected and uninjected, were fixed at 6 hours and stained with Alexa 568 phalloidin to localize filamentous actin (F-actin). The results for this experiment are inconclusive as the embryos have very high background and preservation of F-actin is inconsistent between

embryos. There are also very few control injected embryos that remain intact through processing. Of those that can be distinguished, there appear to be abnormal expression patterns of F-actin in Sp-αC MASO injected embryos. In one case, filamentous actin localizes to only one end of the cleavage furrow in an embryo exhibiting incomplete cytokinesis and the overall signal in the cleavage furrow is less pronounced than in the uninjected control (Fig. 3.11). Further investigation into the localization of F-actin in embryos with reduced αC translation is necessary to determine if these phenotypes are accurate due to the low number of controls that successfully stained.

Embryos prepared with the nuclear stain DAPI revealed that those injected with Sp-αC MASO had mitotic defects (Fig. 3.10). Abnormal numbers of nuclei occur in 19.5% (N=41) of embryos injected with Sp-αC MASO whereas one embryo injected with ZF-Chordin had the same phenotype (4.2%, N=24) and none of the uninjected controls had this phenotype (N=19).

Some embryos injected with Sp-αC MASO had problems with localization of nuclei; they were not centrally located in each blastomere (Fig. 3.10). Problems with localization occur in four experimental embryos (9.8%, N=41), in one (4.2%, N=24) ZF-Chordin MASO embryo, and in none of the uninjected controls (N=19).

The nuclear defects (6 hours after fertilization) for Sp-αC MASO embryos total 29.3% (N=41) with abnormalities in two ZF-Chordin MASO embryos (8.3%, N=24). Statistical tests of a 2 by 2 contingency table using Fisher's exact test results in $p=0.063$. Thus, the null hypothesis, that there is no difference in the treatment groups, cannot be rejected.

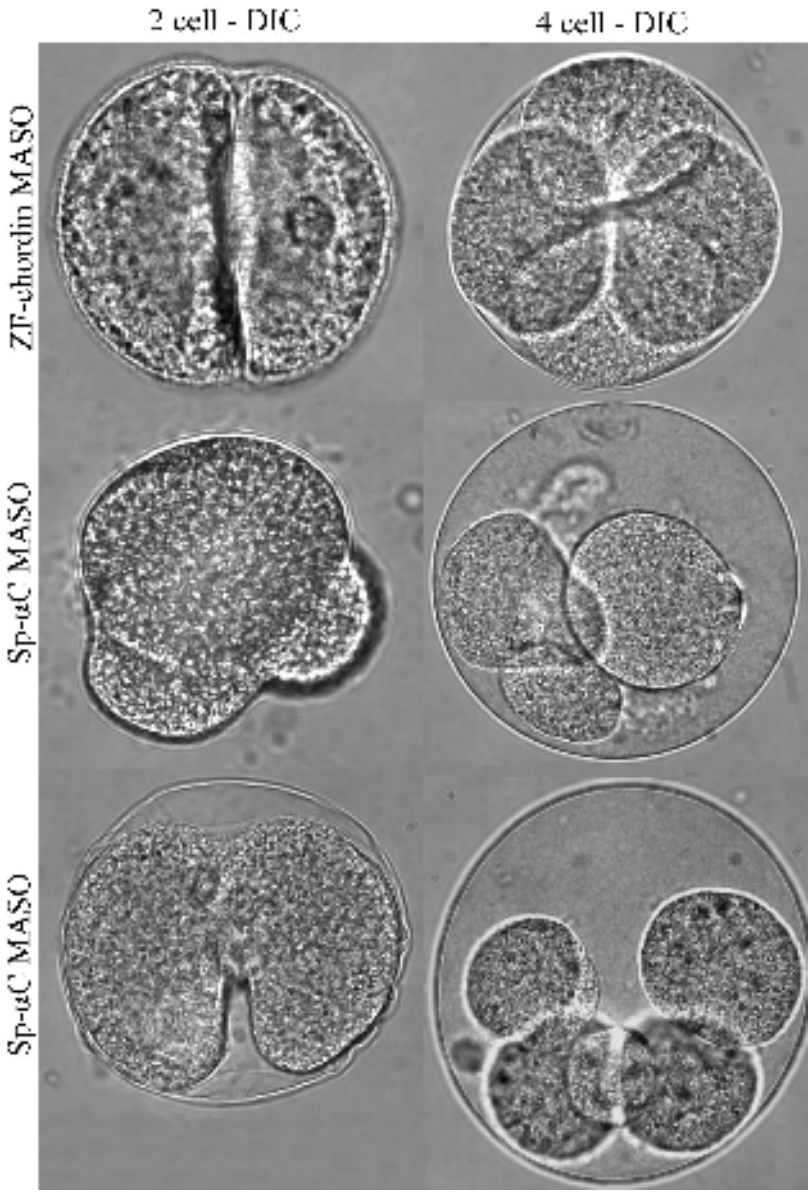


Figure 3.10 Abnormal phenotypes during early cleavage stages. *S. purpuratus* embryos injected with Sp-αC MASO have unusual cell numbers, cells of unusual size and shape, cytokinetic defects, and adhesion defects (21.2%, N=154) by the 2- and 4-cell stage compared to normal development, as seen in the injected control embryos.

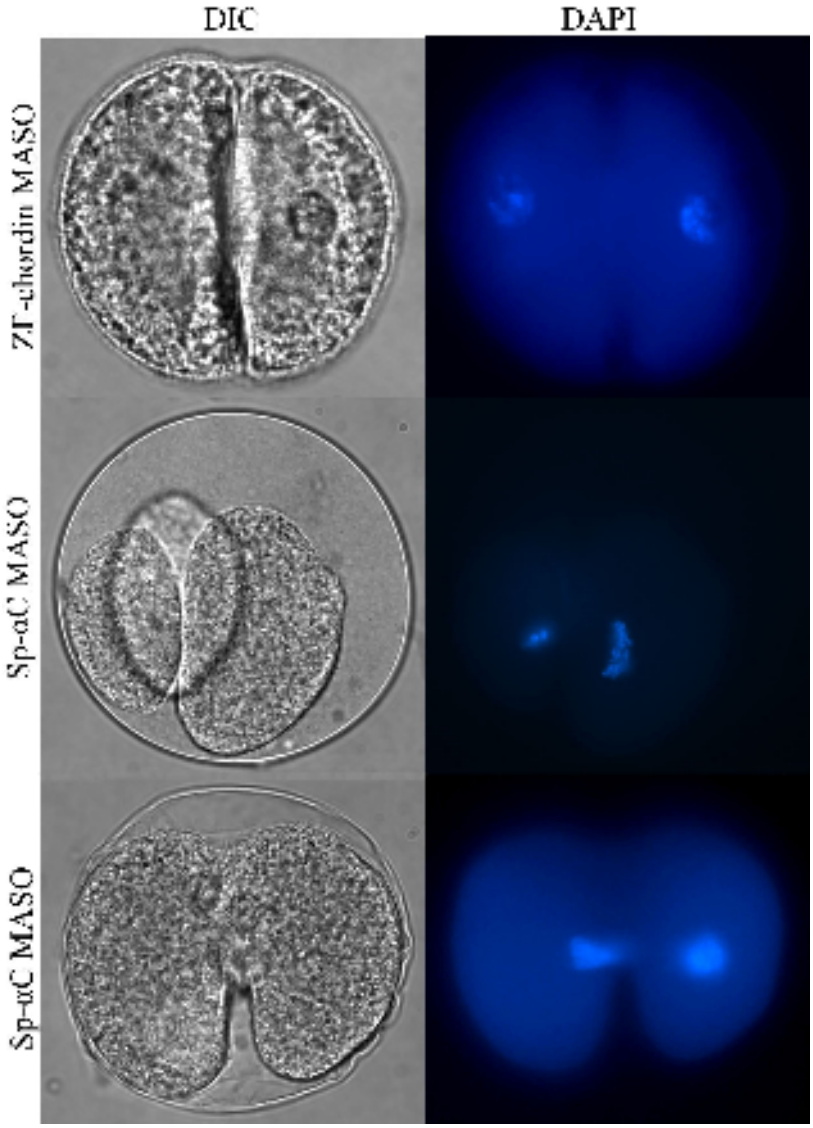


Figure 3.11 Abnormal nuclear phenotypes in early cleavage stages. *S. purpuratus* embryos injected with Sp-αC MASO have an odd number of nuclei, nuclei inconsistent in size and shape, and localization defects within the first 1-2 cleavages (43.9%, N=41). Control injected embryos have nuclei consistent in size and shape localized to the same region in each cell.

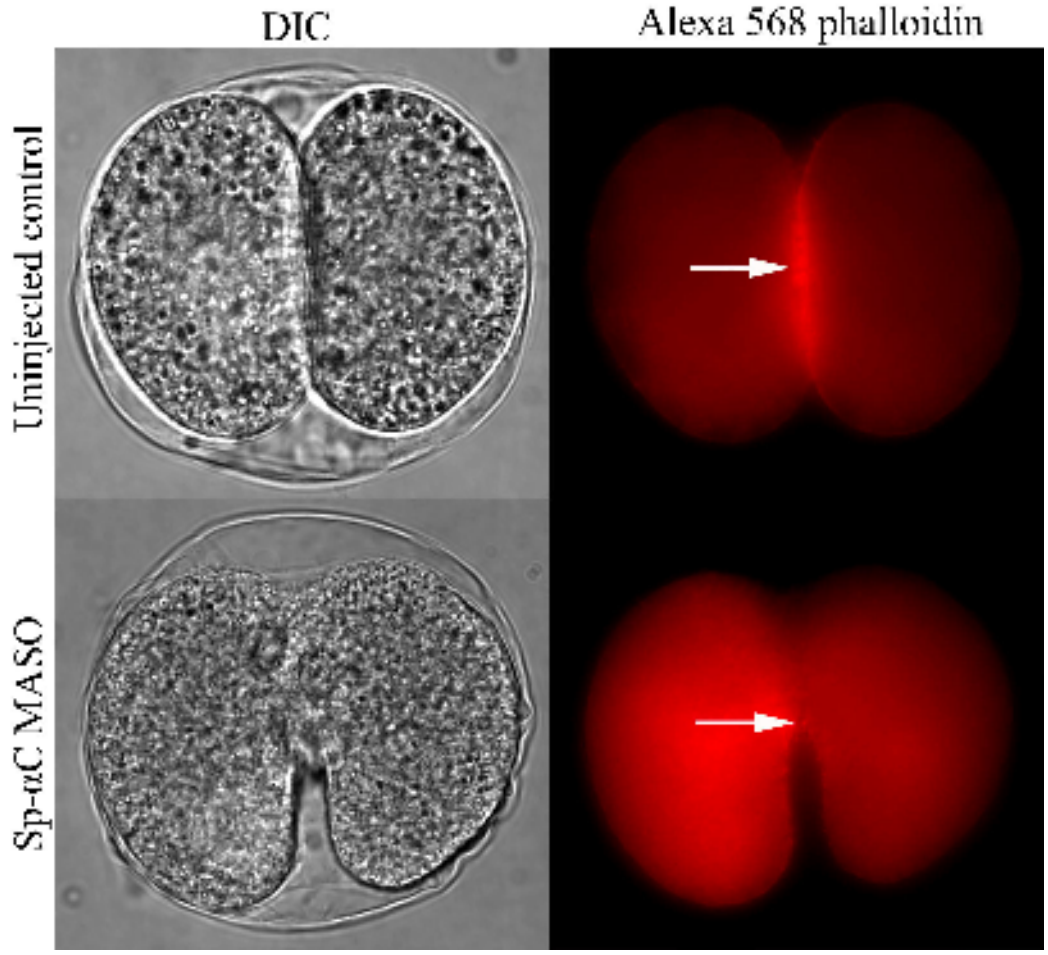


Figure 3.12 Phalloidin staining in *S. purpuratus* embryos injected with Sp-αC MASO. Uninjected embryos have F-actin localize to the cleavage furrow (arrow) during cytokinesis. Some embryos with a knockdown of αC that have cytokinesis defects do not appear to have the same intensity of F-actin in the cleavage furrow (arrow) but rather fluorescence is spread throughout the cell. No control injected embryos were stained for F-actin.

3.4.3 Morphogenetic defects

Normal embryos at 24 hours of development have reached the mesenchyme blastula stage when primary mesenchyme cells (PMCs) begin to ingress. Injected embryos, both control and experimental, suffer a delay in timing and at 24 hours of development are more like the hatched blastula stage which consist of an epithelial blastoderm surrounding an ECM filled space, the blastocoel. Of the 42.9% (N=373) of embryos that are still alive, 23.1% (N=104) of those injected with Sp-αC MASO have an abnormal blastocoel in which more than half of the blastocoel is filled with cells. The control injected embryos experience this phenotype in 3.7% (N=268) and it is found in 0.8% (N=253) of the uninjected control embryos (Figs. 3.12, 3.13). It is unlikely the cells in the abnormal blastocoels are PMCs as they do not form a regular pattern nor do they attach to the wall of the blastocoel (Fig. 3.12).

Sp-αC MASO embryos also fail to form a blastocoel entirely, which occurs in 28.9% (N=104) of those observed. Control injected embryos have the same defect in 2.6% (N=268), whereas uninjected embryos do not have this phenotype (N= 253) (Fig. 3.12). In total, 52.0% (N=104) of the Sp-αC MASO embryos had blastocoel defects (for a total of 79.4% abnormal or dead) while 6.3% (N=268) of the control injected and 0.8% (N=253) of the uninjected controls had blastocoel defects. The null hypothesis, that blastocoel abnormalities occur in the same number of embryos between Sp-αC MASO and control injected, can be rejected ($p<0.0001$).

Embryos lacking normal αC translation also have defects in their nuclei 24 hours post-fertilization. Normal nuclei are round and are equal in size and shape, whereas experimental injected embryos have nuclei that are not round and vary in size and shape

(Figs. 3.12, 3.13). The nuclei in the Sp-αC MASO embyros appear pyknotic with clumped, condensed chromatin (Figs. 3.12, 3.13). Of the embryos injected with SP-αC MASO, abnormal nuclei occurred in 36.7% (N=30) of them, whereas the uninjected control (N=10) and control injected (N=15) do not exhibit this phenotype. Using Fisher's exact test, the null hypothesis, that abnormal nuclear phenotypes occur at 24 hours of development in the same frequency between Sp-αC MASO and control injected embryos, can be rejected ($p<0.01$).

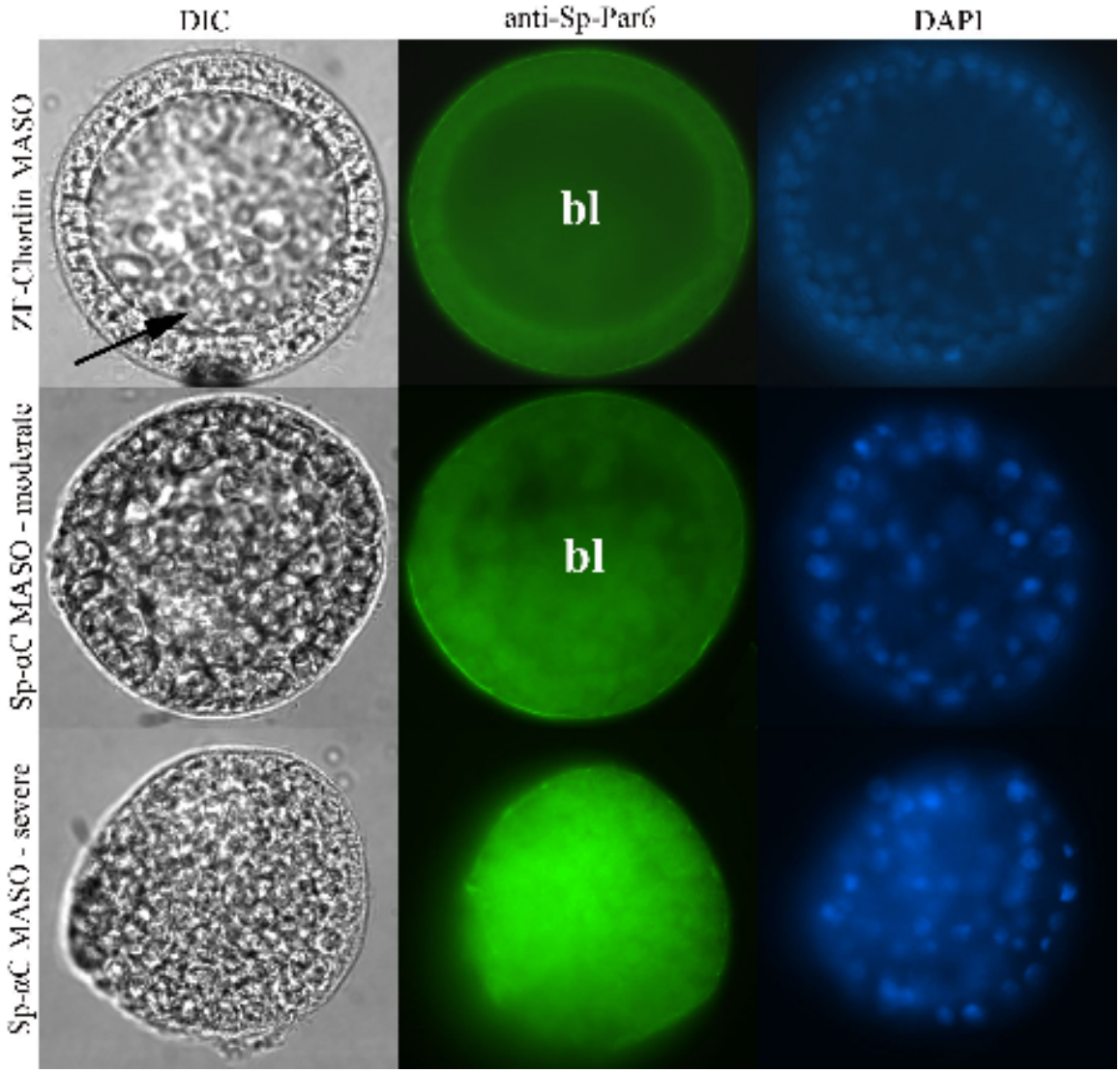


Figure 3.13 *S. purpuratus* embryos injected with Sp-αC MASO that survive to 24 hours show blastocoel defects. Of the 42.9% (N=373) that are still alive, 23.1% (N=104) exhibit a moderate phenotype which forms a blastocoel (bl), over 50% of which is filled with cells that are not PMCs. A severe phenotype in which no blastocoel is formed is seen in 28.9% (N=104) of the embryos, as opposed to the injected controls which form a normal blastocoel (bl) that will soon contain ingressing PMCs.

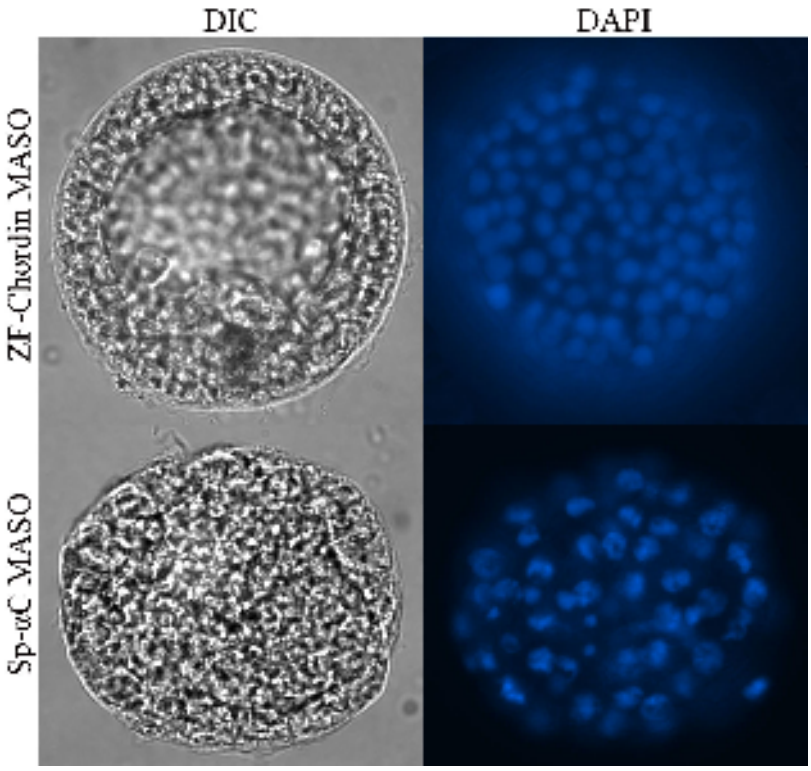


Figure 3.14 Nuclear phenotypes through 24 hours of development. Of the *S. purpuratus* embryos injected with Sp-αC MASO, 36.7% (N=30) exhibit nuclei that are irregular in size and shape and appear to be pyknotic. Control injected embryos at the same stage have round nuclei that are consistent in size, shape, and texture.

At 48 hours post-fertilization normal embryos have formed an archenteron and have distinctly patterned skeletongenic mesoderm. Only 28.3% (N=265) of Sp-αC MASO injected embryos are still alive at this stage and gastrulation has failed in 44.1% (N=179) of these embryos. Typically there is no archenteron formed and scattered clusters of cells are in the blastocoel (Fig. 3.14). Problems with gastrulation are seen in 8.9% (N=425) of control injected embryos, whereas no uninjected embryos have gastrulation defects (N=317). Using Fisher's exact test in a 2 by 2 contingency table, the null hypothesis that gastrulation defects occur to the same extent in control and experimental injected embryos can be rejected ($p<0.0001$).

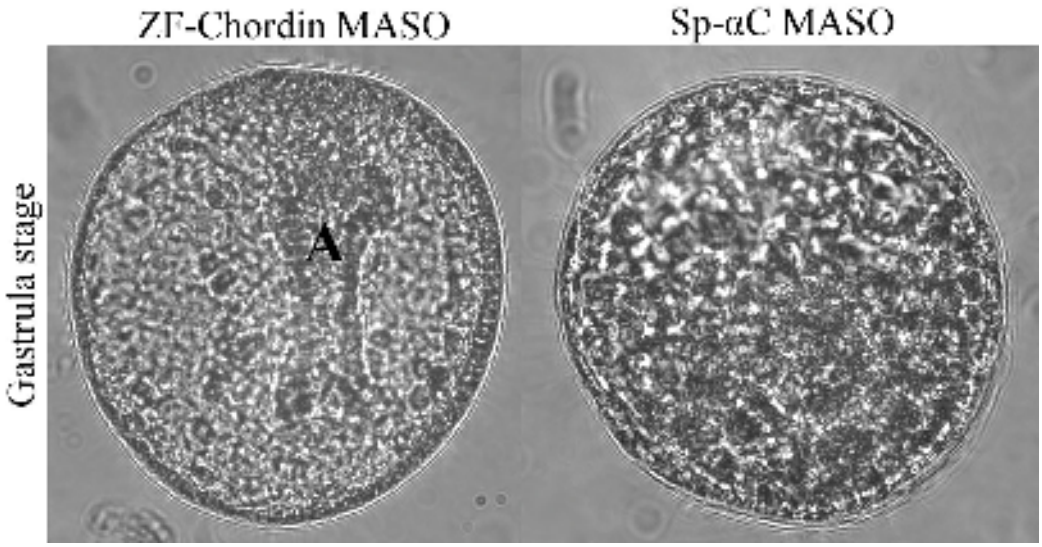


Figure 3.15 *S. purpuratus* embryos injected with Sp-αC MASO exhibit abnormal phenotypes through gastrulation. By 48 hours of development, control embryos have gastrulated and formed complete archenterons. Of the 28.3% (N=265) of embryos still alive, 43.8% (N=112) of those with a knockdown of αC, do not begin to gastrulate or form an archenteron although they are still alive.

3.4.5 Polarity in early and blastula stage embryos

In normal embryos, Sp-Par6 localizes to the apical domain of the cell and between the cells at the apical junctions by 6 hours and remains there throughout development (Figs. 3.15, 3.16). In the surface view of a control blastula, Sp-Par6 can be seen outlining each cell and marking the basal body of the embryo (Fig. 3.17). In a cross section of an embryo with reduced αC expression Sp-Par6 is still visible, but does not consistently localize to the apical surface in both early cleavage and blastula stage embryos (Figs. 3.15, 3.16). Sp-Par6 can be seen in granular form on the surface of 24 hour embryos injected with Sp-αC MASO that exhibit a severe phenotype, but it is not localized around the cell and the basal body cannot be distinguished (Fig. 3.17). Of those embryos injected with Sp-αC MASO, 20% (N=25) have problems with Sp-Par6 localization while none of the control injected (N=16) or uninjected (N=12) show an abnormal phenotype.

Embryos were also stained with 2D2, an antibody against the *S. purpuratus* integrin subunit βC, which has previously been used as a marker for epithelial polarity (Burgess, 2008). Sp-αC MASO injected embryos exhibit normal βC localization to the apical surface of the cells with the exception of one embryo (7%, N=14) in which βC was expressed throughout the embryo but not localized to the apical region. Statistical comparisons of these treatments indicate that the differences are not significant (Fisher's exact test, $p=0.17$).

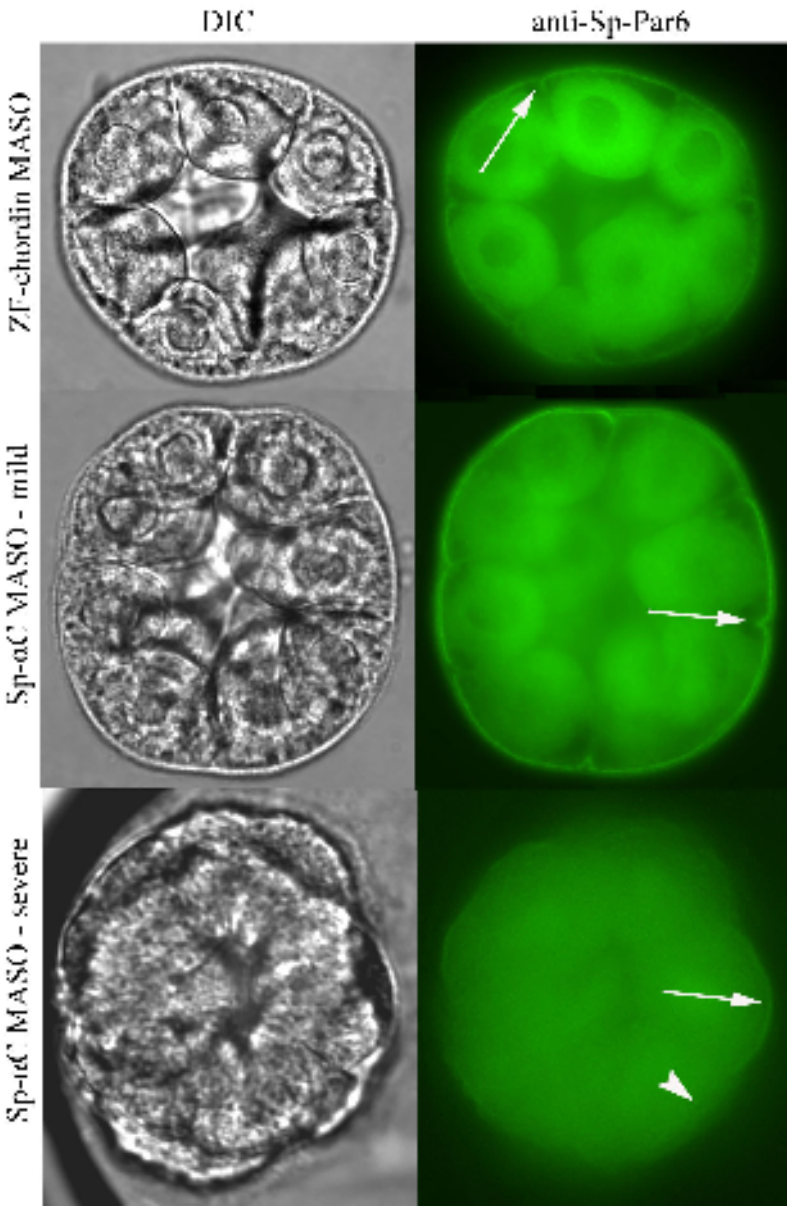


Figure 3.16 Sp-Par6 localization in αC MASO injected embryos at early cleavage stages. The majority of Sp-αC MASO *S. purpuratus* embryos have a mild phenotype very similar to the injected control, whereas Sp-Par6 localizes to the apical domain and junctions of the cells (arrow). Those with a severe phenotype that do exhibit the abnormal localization have Sp-Par6 in some regions (arrow) but not uniformly distributed along the apical domain (arrowhead).

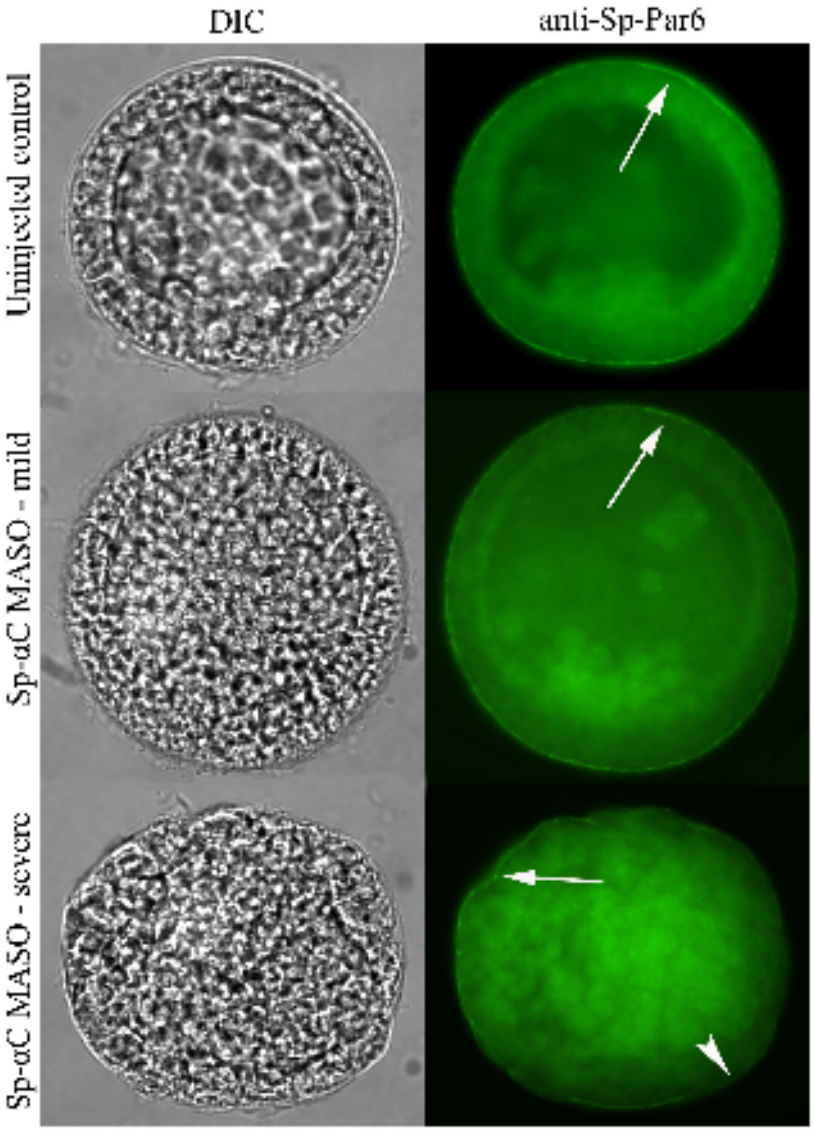


Figure 3.17 Sp-Par6 localization in Sp-αC MASO injected embryos. The majority of Sp-αC MASO *S. purpuratus* embryos have a mild phenotype very similar to the injected control, where Sp-Par6 localizes to the apical domain and junctions of the cells (arrow). Those with a severe phenotype that do exhibit the abnormal localization have Sp-Par6 in some regions (arrow) but not uniformly distributed along the apical domain (arrowhead).

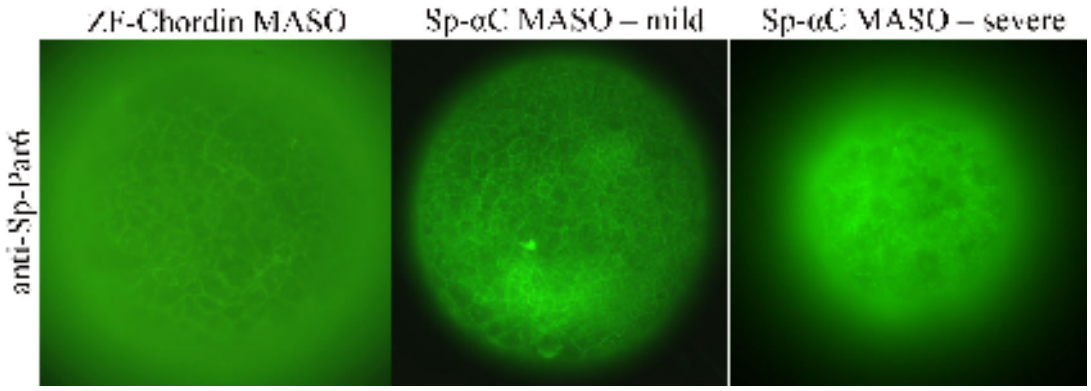


Figure 3.18 A surface view of Sp-Par6 localization in *S. purpuratus* embryos injected with Sp-αC MASO. Control embryos and those that exhibit a mild phenotype show Sp-Par6 localization surrounding each cell and staining at the basal body. Embryos with a severe phenotype lack any patterning at all and single cells cannot be distinguished.

Chapter 4 – Discussion

4.1 Integrin subunit structure

The sequencing and assembly of the sea urchin genome produced a number of gene predictions for integrin subunits; most of which were incomplete (Whittaker et al., 2006). The first objective of my research was to use partially sequenced cDNA clones (ESTs) and RT-PCR amplification of mRNA to determine full length cDNA sequences of integrins expressed in early development. There were at least four predictions of each of the four integrin subunits that I focused on and all of these turned out to be fragments and to contain errors in the prediction of exons. Although there were two assemblies of the genome produced (Sea Urchin Genome Sequencing Consortium, 2006), even when using the same assembly, different gene prediction programs recognized different translation start sites and ORFs within each genomic DNA sequence. To use one gene as an example, the predictions for αC vary, but there are often places where one prediction matches a region in the cDNA sequence where another prediction does not. For example, Scaffold_v2_32336_1 does not match the confirmed cDNA from 1649-2097 of the sequence, but GLEAN3_15377 does match in that same region. Many of the irregularities occur at intron/exon boundaries suggesting this is an area of uncertainty for gene prediction programs. Alternative splice forms are a possible explanation for variations between the predictions and cDNA sequences. Alternative splicing has been found to occur in vertebrate integrins, in extracellular and intracellular domains (de Melker and Sonnenberg, 1999). To determine if there are alternative splice forms, more sequencing of αC, αF, αD, and βD will need to be done. However, there were no

situations identified throughout my studies in which PCR amplification produced multiple products that would result from exon skipping or alternative splice sites.

One region lacking in the *S. purpuratus* α integrins is the inserted (I) domain. This domain consists of α-helices surrounding a central β-sheet and is a 200 residue insertion replacing the first divalent cation binding site. This I domain is necessary for ligand recognition and binding in the integrins that contain it. Mg^{2+} binds to the N-terminus of this region via a cation-binding domain. In the integrins that do not contain the I domain, including *S. purpuratus* subunits, the ligand binding domain is located directly within the α subunit (Shimaoka et al., 2002). This modification has only been found in vertebrate α subunits to date, so the fact that it is not present in the *S. purpuratus* subunits is not surprising.

With the addition of my data on the cDNA structure for integrin subunits, it is now clear that there are four β subunits (βC, βD, βG, βL) and at least seven α subunits (αD, αC, αF, αP, αH, αJ, and αK). The annotation process identified an additional subunit, αG, however it appears to be a gene fragment that cannot be amplified from embryonic or adult cDNA. The gene models for αJ, αK, and αH are also incomplete, although expression has been confirmed with PCR and EST sequencing. However, the 5' ends of all three genes have yet to be determined.

4.2 Phylogeny

The *S. purpuratus* α integrin subunits have low sequence identity ranging from 25% identity between αC and αD to 40% identity between αP and αF. Of the four

subunits, αC has the lowest sequence identity with the other three. In vertebrates α integrin identity ranges from 55-59%, although identity can be as high as 88-92% between several human and guinea pig sequences within the same region of the subunit (Susan et al., 2000). Of the newly sequenced *S. purpuratus* α integrins, they share identity between human α integrins ranging from 27% (αD with α8) to 31% (αC with α6). βD shares more identity with other *S. purpuratus* β subunits than the α subunits have with each other, ranging from 40% to 47%, although these values are still low.

It has been proposed that the α integrin subunits are monophyletic, with a pre-metazoan history (Hughes, 1992; Hughes, 2001; Hynes and Zhao, 2000; Whittaker et al., 2006). Phylogenies based on aligned α subunit amino acid sequences suggest that there are two main groups. These two groups appear to relate to the function of the receptors: those that bind laminin and those that bind RGD (Hynes and Zhao, 2000). The phylogenetic analysis I have done of the aligned integrin_α2 domains of the sea urchin α subunits with representative α subunits from vertebrates and other invertebrates produces a tree in which there are two main groups of α integrins. However, the bootstrap values for the division of the two major clades are low, indicating the relationship is not clearly resolved by the data. Within the two clades, relationships are more distinct and there appears to be one clade of sea urchin α subunits (αF, αJ, αD, αP) that represent an echinoid diversification of RGD binding subunits. The αC subunit clusters with other deuterostome laminin binding subunits. My data expands the phylogenetic analysis reported by Whittaker et al. (2006) on three of the subunits identified from analysis of the genome and confirms that urchins, like other organisms, have at least one α subunit from each functional group.

When the integrin_α2 domains are aligned using ClustalW, there is sequence identity in only 25% of the sequence. Anything below 20% identity can result in an incorrect alignment of up to 50% of the amino acid residues while 20-30% identity results in approximately 80% of the alignment being correct (Thompson et al., 1999), so this phylogenetic analysis of α integrins is in the threshold region of accuracy. The *S. purpuratus* integrin_α2 domains by themselves have 32% sequence identity, which is just above the threshold value allowing for greater than 80% of the alignment to be accurate (Thompson et al., 1999). This sequence alignment of α subunits may not be possible to improve on greatly. It seems the α integrin subunits have diverged substantially and have little sequence similarity, even within the integrin_α2 domain.

4.3 Temporal Expression

QPCR data of *S. purpuratus* integrin subunits identifies two main patterns of mRNA expression, which supports the initial hypothesis that expression of these subunits is regulated. αC, αD, βD, and βC are expressed throughout the first 96 hours of development at relatively constant levels. Although there are fluctuations in mRNA abundance it does not vary by more than a two-fold change from the egg.

The second pattern of expression currently found in only one integrin subunit, αF, is an increase in message abundance during late cleavage and morphogenesis. The upregulation of αF suggests that it is involved subsequent to cleavage, possibly during morphogenesis.

The patterns of expression obtained by QPCR support the genome-wide analysis of embryonic expression reported by Wei et al. (2006). However, a difference between the QPCR and the chip hybridization results is the fold change in relative levels of gene expression. Although the QPCR data indicates the expression levels of αC, αD, βD, and βC do not change significantly between stages, the chip hybridization data finds changes greater than two-fold do occur. Notably, αC, αD, and βC have significantly higher messages during early cleavage stages than they do during later stages, up to fifteen times higher (Appendix III). The chip data reports message abundance relative to expression at 48 hours, whereas I have reported expression relative to message abundance in the egg. There are some minor discrepancies, but overall it is possible to conclude that seven α subunits and all four β subunits are expressed in embryonic development (Marsden and Burke, 1997; Marsden and Burke, 1998; Susan et al., 2000; Wei et al., 2006). The subunits of αC, αH, αD, αJ, and αK have relatively constant levels of expression, whereas αP and αF increase in abundance during cleavage and later development. The subunits βD and βC are expressed throughout development at relatively constant levels and βG and βL increase in abundance after cleavage. Major questions remain as to what the precise roles are for all of these subunits and what subunits combine to form functional receptors.

4.4 Function of αC in early development

The αC MASO is lethal in over 70% of the embryos within 48 hours of fertilization. This supports the original hypothesis regarding the function of αC that

indicates that interfering with the translation of this subunit is lethal and that the αC subunit is necessary for early development. This evidence also supports the hypothesis that integrin subunits contribute to form a functional receptor. To determine what aspect of development requires the αC subunit I have taken into account two critical aspects of how the experiment was done. It is intended that eggs are injected with three picolitres, however it is impossible to ensure that precisely the same amount of MASO is injected into each egg. In any experiment there are some embryos strongly affected by the treatment and others less strongly affected. Thus, the speculation I have made is that embryos that are strongly affected received a higher dose of MASO than those that are less affected. The second aspect of the experimental design that I have applied to my speculation of these results is that development progresses with time and that a defect in a process that occurs early in development may not be lethal, but that defect may alter the outcome of a subsequent process that is dependent on the first. Over time, a defect in an earlier developmental process will slowly have consequences on subsequent processes. The embryo may then succumb to defects that are only indirectly related to processes first affected by the treatment. Thus, the embryos that fail to form a blastula or to gastrulate do not necessarily indicate that the αC subunit has a role in morphogenesis. When I take these two aspects of experimental design into account, the earliest defects that I observe in αC MASO injected embryos are likely to be manifestations of a high dose of the MASO on the aspect of development that first requires the αC subunit.

The necessary control to determine the effect of the MASO is to demonstrate that the αC protein is not produced, or is less abundant. As well, it is necessary to demonstrate that the MASO has not produced off target defects usually by use of multiple

morpholinos designed to block translation or to alter intron-exon splicing (Eisen and Smith, 2008). The control that I have done is to inject some eggs with an irrelevant MASO, which provides a reference and controls for injection artefacts, but it is not adequate to conclude that the defects are due to loss of the αC subunit alone. However, the distinctive nature of the defects produced by the αC MASO and the consistent results indicate that these additional controls should be done to confirm the specificity of the effects of the αC MASO.

The first defects that occurred in a proportion of embryos were early defects in cytokinesis and mitosis. These occurred in a larger proportion of the αC MASO injected embryos than in control injected embryos, although the proportion affected was not statistically significant. However, by 24 hours, a significant proportion of the embryos injected with the αC MASO had aberrations of cytokinesis and mitosis. The limited data I have on formation of the contractile actin ring during cytokinesis suggests that the αC MASO may interfere with its formation. Twenty-four hours after fertilization there were also a significant number of αC MASO injected embryos that had what appeared to be dead cells in the blastocoel or that failed to form a blastocoel.

What aspect of integrin function could produce a defect that alters cytokinesis and mitosis? In mammalian cells in culture it has been noted that as mitosis begins, cells loosen attachments to the substrate and disassemble focal adhesions, stress fibers, and microtubules (Maddox and Burridge, 2003). Microtubules then reassemble into the bipolar spindle to direct mitosis, and actin filaments from the contractile ring to separate daughter cells during cytokinesis (Glotzer, 2001; Vandre et al., 1984). As cell division nears completion, daughter cells re-spread and focal adhesions, stress fibres, and the

radial microtubule network re-form. This dynamic regulation of adhesion during cell division suggests a mechanism that links the function of integrins with cell division. Reverte et al. (2006) demonstrated that cells containing an integrin with a point mutation in the NPXY motif of the β1 cytoplasmic tail fail to form spindles and complete cytokinesis. These effects of expression of the mutant subunit could be rescued with an antibody that caused activation of integrins. Reverte et al. (2006) concluded that integrins can regulate the assembly of the microtubule cytoskeleton during interphase and the bipolar spindle during mitosis.

In urchin eggs, blastomere division involves a sequence of events that are similar to division of mammalian cells in culture; cells release their apical attachment to the hyaline layer and their basal attachment to the basal lamina before rounding up during mitosis. Following cytokinesis they again attach apically and basally to the extracellular matrix and adopt their former shape. An effect of interfering with the translation of the integrin αC subunit appears to be defects in cytokinesis and mitosis. There are precedents for integrins having a role in coordinating mitosis and cytokinesis, but the mechanisms are unknown. Urchin eggs are a well established model for studies of cytokinesis and mitosis, and many of the founding principles for coordination of the two processes were initially worked out in sea urchin eggs (Rappaport, 1996). The hypothesis that the αC subunit functions in mediating coordination of mitosis and cytokinesis merits a full examination.

Some of the defects observed in later stages include altered nuclear morphology and what appear to be dead cells accumulating in the blastocoels. These defects suggest the embryos lacking the αC integrin may be apoptotic. It has been suggested that

embryos injected with a morpholino against the βC integrin subunit, may also be apoptotic (Burke et al., 2007). Apoptosis is an event that integrins are intricately involved in. It can occur when the cell is no longer attached to the ECM (via integrins, for example), a process called anoikis, or integrins can interfere with signaling pathways that are responsible for the continuation of the cell cycle, as mentioned above (Assoian and Schwartz, 2001; Frisch and Screaton, 2001; Hulleman and Boonstra, 2001; Jan et al., 2004; Schwartz, 2001; Stupack et al., 2001; Stupack and Cheresh, 2002). Bcl-2 is an antiapoptotic protein that is upregulated by integrins, possibly via the PI3K pathway (Jan et al., 2004). Integrins may also initiate survival-signaling via FAK which can send signals for the cell cycle to continue even in conditions that might normally be apoptotic, such as a reduction in growth factor signaling (Stupack et al., 2001).

Although several of the later effects of αC MASO injection are consistent with apoptosis, this may be an indirect effect. The early defects in cytokinesis and mitosis are not clearly linked to activation of an apoptotic pathway. However, defects in mitotic mechanisms may well lead to cell death. It is possible that there are several direct effects of interfering with the translation of the integrin αC subunit. However, the expression data indicates there are several integrins expressed in early development and their function would not be clearly altered by elimination of αC. Thus, it seems more likely that the suite of defects I see are a combination of direct and indirect effects. There are a number of cytological and biochemical methods available to assess apoptosis and it is a priority for future research to determine if apoptosis occurs and if it is a direct effect of interfering with the function of αC containing integrin receptors or an indirect effect of altered mitosis and cytokinesis.

Chapter 5 – Conclusions and Future Directions

This study has described four novel integrin subunits, all of which are expressed in sea urchin embryos. Three α subunits and one β subunit each have a single open reading frame in their confirmed cDNA sequences which predicts a protein containing canonical integrin features. Knowing the primary structure of these subunits expands the phylogenetic analysis reported by Whittaker et al. (2006) and opens the way to functional studies. The temporal data on integrin expression reveals two major patterns of mRNA abundance; constant levels throughout early embryogenesis and later expression during morphogenesis. At least one of these subunits, αC, is necessary for embryonic development as interference with translation of αC results in embryonic lethality. This is in agreement with other studies which have shown some integrins to be absolutely required for early events in embryogenesis (Hynes, 2002; Katz and Yamada, 1997).

Defining the precise roles of the αC integrin subunit is challenging. The evidence I report indicates that αC is involved in cytokinesis and mitosis. One crucial piece of information missing is the localization of the αC protein. Knowing the temporal and spatial dynamics of the protein in the developing embryo offers insight into its functions. Determining the β subunit or subunits that interact with αC will also allow a more precise hypothesis on integrins functions in *S. purpuratus* development (Burke et al., 2004; Marsden and Burke, 1997; Marsden and Burke, 1998; Murray et al., 2000a). Localization of the protein can be determined using an antibody against αC or by expressing an αC:GFP fusion protein from injected RNA. Co-localization with

antibodies against β subunits or pull down experiments may reveal which subunits combine to form functional receptors.

The expression data indicates that all four β integrin subunits and at least six α subunits are expressed within the first 48 hours of development. Indeed, in the fertilized egg there may be as many as two β subunits and four α subunits expressed. This suggests that there are at least four and as many as eight distinct receptors expressed during the first cell cycle. It is not known whether integrins function independently or if different integrins have overlapping roles. However, interfering with the translation of αC and βC (Burke et al., 2004) both produced distinctive phenotypes, suggesting that these subunits have unique roles that are not compensated for by the other subunits. Clearly, data on the expression of the subunit proteins and loss of function experiments with individual subunits and combinations of subunits are necessary to determine the role of this complex array of integrin receptors.

Integrins are known to be expressed on the surface of eggs and early embryos in a number of species (Lallier et al., 1994). The function of integrins on mouse eggs has been thoroughly studied, yet remains enigmatic. Their putative role in fertilization has not been substantiated in mice with targeted deletions of subunits and little is known about their post-fertilization functions (He et al., 2003). Integrins clearly have a role in implantation and morphogenesis, but what their post-fertilization role is remains unknown (Wang et al., 2006).

This study provides insight into the involvement of integrins in early embryogenesis and provides a more refined hypothesis of integrin function during cleavage; that αC is necessary for coordination of mitosis and cytokinesis during early

embryonic cell division. Determining a specific role for αC will allow for a greater understanding of the role of integrins in development, embryonic development in *S. purpuratus*, but also development in general.

Literature Cited

Assoian, R. K. and Schwartz, M. A. (2001). Coordinate signaling by integrins and receptor tyrosine kinases in the regulation of G_1 phase cell-cycle progression. *Current Opinions in Genetics and Development* **11,** 48-53.

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

Briggs, E. and Wessel, G. M. (2006). In the beginning... Animal fertilization and sea urchin development. *Developmental Biology* **300,** 15-26.

Burgess, D. R. (2008). Cytokinesis and the establishment of early embryonic cell polarity. *Biochemical Society Transactions* **36,** 384-386.

Burke, R. D. (1999). Invertebrate integrins: structure, function, and evolution. *International Review of Cytology* **191,** 257-284.

Burke, R. D., Brothers, E. and Matviw, A. (2007). Integrin signaling in early sea urchin development. *Signal Transduction* **7,** 207-215.

Burke, R. D., Murray, G., Rise, M. and Wang, D. (2004). Integrins on eggs: the βC subunit is essential for formation of the cortical actin cytoskeleton in sea urchin eggs. *Developmental biology* **265,** 53-60.

***C. elegans* Sequencing Consortium.** (1998). Genome sequence of the nematode *C. elegans*: a platform for investigating biology. *Science* **282,** 2012-2018.

Campbell, I. D. (2008). Studies of focal adhesion assembly. *Biochemical Society Transactions* **36,** 263-266.

Carlson, B. M. (1996a). Chapter 5: Cleavage and formation of the blastula. In *Patten's Foundations of Embryology*, pp. 151-188. New York, NY: McGraw-Hill, Inc.

Carlson, B. M. (1996b). Chapter 6: Gastrulation and the formation of germ layers. In *Patten's Foundations of Embryology*, pp. 189-194. New York, NY: McGraw-Hill, Inc.

Chalmers, A. D., Pambos, M., Mason, J., Lang, S., Wylie, C. and Papalopulu, N. (2005). aPKC, Crumbs3, and Lgl2 control apicobasal polarity in early vertebrate development. *Development* **132,** 977-986.

Davidson, E. H., Cameron, A. R. and Ransick, A. (1998). Specification of cell fate in the sea urchin embryo: summary and some proposed mechanisms. *Development* **125,** 3269-3290.

Davidson, E. H., Hough-Evans, B. R. and Britten, R. J. (1982). Molecular biology of the sea urchin embryo. *Science* **217,** 17-26.

DeMali, K. A., Wennerberg, K. and Burridge, K. (2003). Integrin signaling to the actin cytoskeleton. *Current Opinion in Cell Biology* **15,** 572-582.

de Melker, A. A. and Sonnenberg, A. (1999). Integrins: alternative splicing as a mechanism to regulate ligand binding and integrin signaling events. *BioEssays* **21,** 499-509.

***Drosophila* Genome Sequencing Consortium.** (2000). The genome sequence of *Drosophila melanogaster*. *Science* **287,** 2185-2195.

Eisen, J. S., and Smith, J. C. (2008). Controlling morpholino experiments: don't stop making antisense. *Development* **135,** 1735-1743.

Eslami, A. and Philpot, J. (2005). The role of integrins in wound healing. *The Science Creative Quarterly*.

Estes, J. E., Selden, L. A. and Gershman, L. C. (1981). Mechanism of action of phalloidin on the polymerization of muscle actin. *Biochemistry* **20,** 708-712.

Ferrell, J. E. and Martin, G. S. (1989). Tyrosine-specific protein phosphorylation is regulated by glycoprotein IIb-IIIa in platelets. *Proceedings of the National Academy of Sciences of the United States of America* **86,** 2234-2238.

Filipenko, N. R., Attwell, S., Roskelley, C. and Dedhar, S. (2005). Integrin-linked kinase activity regulates Rac- and Cdc42-mediated actin cytoskeleton reorganization via alpha-PIX. *Oncogene* **24,** 5837-5849.

Finn, R. D., Mistry, J., Schuster-Böckler, B., Griffiths-Jones, S., Hollich, V., Lassmann, T., Moxon, S., Marshall, M., Khanna, A., Durbin, R. et al. (2006). Pfam: clans, web tools, and services. *Nucleic Acids Research* **34,** D247-D251.

Frisch, S. M. and Screaton, R. A. (2001). Anoikis mechanisms. *Current Opinion in Cell Biology* **13,** 555-562.

Geiger, B., Bershadsky, A., Pankov, R. and Yamada, K. M. (2001). Transmembrane extracellular-matrix cytoskeleton crosstalk. *Nature Reviews Molecular Cell Biology* **2,** 793-805.

Gilbert, S. F. (2000). The early development of sea urchins. In *Developmental Biology*: Sinauer Associates, Inc.

Glotzer, M. (2001). Animal cell cytokinesis. *Annual Review of Cell and Developmental Biology* **17,** 351-386.

Grashoff, C., Thievessen, I., Lorenz, K., Ussar, S. and Fassler, R. (2004). Integrin-linked kinase: integrin's mysterious partner. *Current Opinions in Cell Biology* **16,** 565-571.

Harada, Y., Yasuo, H. and Satoh, N. (1995). A sea urchin homologue of the chordate *Brachyury (T)* gene is expressed in the secondary mesenchyme founder cells. *Development* **121,** 2747-2754.

He, Z. Y., Brakebusch, C., Fässler, R., Kreidberg, J. A., Primakoff, P., and Myles, D. G. (2003). None of the integrins known to be present on the mouse egg or to be ADAM receptors are essential for sperm-egg binding and fusion. *Developmental Biology* 254, 226-237.

Heil-Chapdelaine, R. A. and Otto, J. J. (1996). Relative changes in F-actin during the first cell cycle: evidence for two distinct pools of F-actin in the sea urchin egg. *Cell Motility and the Cytoskeleton* **34,** 26-35.

Hertwig, O. (1876). Beiträge zur Kenntniss der Bildung, Befruchtung und Theilung des thierischen Eies. *Morphol. Jahr.* **1,** 452.

Hertzler, P. L. and McClay, D. R. (1999). AlphaSU2, an epithelial integrin that binds laminin in the sea urchin embryo. *Developmental Biology* **207,** 1-13.

Howard-Ashby, M., Materna, S. C., Brown, C. T., Chen, L., Cameron, R. A. and Davidson, E. H. (2006a). Identification and characterization of homeobox transcription factor genes in *Strongylocentrotus purpuratus*, and their expression in embryonic development. *Developmental Biology* **300,** 74-89.

Howard-Ashby, M., Materna, S. C., Brown, C. T., Tu, Q., Oliveri, P., Cameron, R. A. and Davidson, E. H. (2006b). High regulatory gene use in sea urchin embryogenesis: Implications for bilaterian development and evolution. *Developmental Biology* **300,** 27-34.

Hughes, A. L. (1992). Coevolution of the vertebrate integrin α- and β-chain genes. *Molecular Biology and Evolution* **9,** 216-234.

Hughes, A. L. (2001). Evolution of the Integrin α and β protein families. *Journal of Molecular Evolution* **52,** 63-72.

Hughes, P. E., Diaz-Gonzalez, F., Leong, L., Wu, C., McDonald, J. A., Shattil, S. J. and Ginsberg, M. H. (1996). Breaking the integrin hinge. *Journal of Biological Chemistry* **271,** 6571-6574.

Hulleman, E. and Boonstra, J. (2001). Regulation of G1 phase progression by growth factors and the extracellular matrix. *Cellular and Molecular Life Science* **58,** 80-93.

Humphries, J. D., Wang, P., Streuli, C., Geiger, B., Humphries, M. J. and Ballestrem, C. (2007). Vinculin controls focal adhesion formation by direct interactions with talin and actin. *Journal of Cell Biology* **179,** 1043-1057.

Hynes, R. O. (1987). Integrins: a family of cell surface receptors. *Cell* **48,** 549-554.

Hynes, R. O. (2002). Integrins: bidirectional, allosteric signaling machines. *Cell* **110,** 673-687.

Hynes, R. O. and Lander, A. D. (1992). Contact and adhesive specificities in the associations, migrations, and targeting of cells and axons. *Cell* **68,** 303-322.

Hynes, R. O. and Zhao, Q. (2000). The evolution of cell adhesion. *Journal of Cell Biology* **150,** F89-F96.

Jan, Y., Matter, M., Pai, J., Chen, Y., Pilch, J., Komatsu, M., Ong, E., Fukuda, M. and Ruoslahti, E. (2004). A mitochondrial protein, Bit1, mediates apoptosis regulated by integrins and Groucho/TLE corepressors. *Cell* **116,** 751-762.

Katz, B. Z. and Yamada, K. M. (1997). Integrins in morphogenesis and signaling. *Biochimie* **79,** 467-476.

Komura, H., Ogita, H., Ikeda, W., Mizoguchi, A., Miyoshi, J. and Takai, Y. (2008). Establishment of cell polarity by afadin during the formation of embryoid bodies. *Genes to Cells* **1,** 79-90.

Kramer, J. M. (2005). Basement membranes. In *Wormbook, Ed. the C. Elegans Research Community*.

Krysko, D. V., Berghe, T. V., D'Herde, K. and Vandenabeele, P. (2008). Apoptosis and necrosis: detection, discrimination, and phagocytosis. *Methods* **44,** 205-221.

Lallier, T., Hens, M. D. and Desimone, D. W. (1994). Integrins in Development. In *Integrins: Molecular and Biological Responses to the Extracellular Matrix* (ed. D. A. Cheresh and R. P. Mecham), pp. 111-133. San Diego, California: Academic Press Inc.

Legate, K. R., Montañez, E., Kudlacek, O. and Fässler, R. (2006). ILK, PINCH and parvin: the tIPP of integrin signaling. *Nature Reviews Molecular Biology* **7,** 20-31.

Letunic, I., Copley, R. R., Schmidt, S., Ciccarelli, F. D., Doerks, T., Schultz, J., Ponting, C. P. and Bork, P. (2004). SMART 4.0: towards genomic data integration. *Nucleic Acids Research* **32,** D142-D144.

Letunic, I., Copley, R. R., Pils, B., Pinkert, S., Schultz, J. and Bork, P. (2006). SMART 5: domains in the context of genomes and networks. *Nucleic Acids Research* **34,** D257-D260.

Liddington, R. C. and Ginsberg, M. H. (2002). Integrin activation takes shape. *Journal of Cell Biology* **158,** 833-839.

Luo, B., Carman, C. V. and Springer, T. A. (2007). Structural basis of integrin regulation and signaling. *Annual Review of Immunology* **25,** 619-647.

Mabuchi, I. (1994). Cleavage furrow: timing of emergence of contractile ring actin filaments and establishment of the contractile ring by filament bundling in sea urchin eggs. *Journal of Cell Science* **107,** 1853-1862.

Maddox, A. S. and Burridge, K. (2003). RhoA is required for cortical retraction and rigidity during mitotic cell rounding. *Journal of Cell Biology* **160,** 255-265.

Marsden, M. and Burke, R. D. (1997). Cloning and characterization of novel [beta] integrin subunits from a sea urchin. *Developmental Biology* **181,** 234-245.

Marsden, M. and Burke, R. D. (1998). The [beta]L integrin subunit is necessary for gastrulation in sea urchin embryos. *Developmental Biology* **203,** 134-148.

Marsden, M. and de Simone, D. W. (2003). Integrin-ECM interactions regulate cadherin-dependent cell adhesion and are required for convergent extension in *Xenopus*. *Current Biology* **13,** 1182-1191.

Miller, J. R. and McClay, D. R. (1997a). Changes in the pattern of adherens junction-associated beta-catenin accompany morphogenesis in the sea urchin embryo. *Developmental Biology* **192,** 310-322.

Miller, J. R. and McClay, D. R. (1997b). Characterization of the role of cadherin in regulating cell adhesion during sea urchin development. *Developmental Biology* **192,** 323-339.

Montero, J. and Helsenberg, C. (2003). Adhesive crosstalk in gastrulation. *Developmental Cell* **5,** 190-191.

Morey, J. S., Ryan, J. C. and Van Dolah, Frances M. (2006). Microarray validation: factors influencing correlation between oligonucleotide microarrays and real-time PCR. *Biological Procedures Online* **8,** 175-193.

Moro, L., Venturino, M., Bozzo, C., Silengo, L., Altruda, F., Beguinot, L., Tarone, G. and Defilippi, P. (1998). Integrins induce activation of EGF receptor: role in MAP kinase induction and adhesion-dependent cell survival. *The EMBO Journal* **17,** 6622-6632.

Murray, G., Reed, C., Marsden, M., Rise, M., Wang, D. and Burke, R. D. (2000a). The αBβC integrin is expressed on the surface of the sea urchin egg and removed at fertilization. *Developmental Biology* **227,** 633-647.

Murray, G., University of Victoria . Dept. of Biology and University of Victoria. (2000b). Expression of the beta-c integrin subunit in early development of the sea urchin. xiv, 108.

Nemer, M., Rondinelli, E., Infante, D. and Infante, A. A. (1991). Polyubiquitin RNA characteristics and conditional induction in sea urchin embryos. *Developmental Biology* **145,** 255-265.

Oliveri, P., Carrick, D. M. and Davidson, E. H. (2002). A regulatory gene network that directs micromere specification in the sea urchin embryo. *Developmental Biology* **246,** 209-228.

Oxley, C. L., Anthis, N. J., Lowe, E. D., Vakonakis, I., Campbell, I. D. and Wegener, K. L. (2007). An integrin phosphorylation switch: the effect of beta3 integrin tail phosphorylation on DOK1 and talin binding. *Journal of Biological Chemistry* **Epub,**.

Ozaki, M., Ogita, H. and Taki, Y. (2007). Involvement of integrin-induced activation of protein kinase C in the formation of adherens junctions. *Genes to Cells* **12,** 651-662.

Pierschbacher, M. D. and Ruoslahti, E. (1984). The cell attachment activity of fibronectin can be duplicated by small fragments of the molecule. *Nature* **309,** 30-33.

Pires-daSilva, A. and Sommer, R. J. (2003). The evolution of signaling pathways in animal development. *Nature Reviews Genetics* **4,** 39-49.

Poustka, A. J., Herwig, R., Krause, A., Hennig, S., Meier-Ewert, S. and Lehrach, H. (1999). Toward the gene catalogue of sea urchin development: the construction and analysis of an unfertilized egg cDNA library highly normalized by oligonucleotide fingerprinting. *Genomics* **59,** 122-133.

Rajeswari, J. and Pande, G. (2006). Direct association between caspase 3 and α5β1 integrin and its role during anoikis of rat fibroblasts. *Cell Biology International* **30,** 963-969.

Ramos, J. W., Whittaker, C. A. and DeSimone, D. W. (1996). Integrin-dependent adhesive activity is spatially controlled by inductive signals at gastrulation. *Development* **122,** 2873-2883.

Ransick, A. (2004). Detection of mRNA by *in situ* hybridization and RT-PCR. *Methods in Cell Biology* **74,** 601-620.

Ransick, A., Rast, J. P., Minokawa, T., Calestani, C. and Davidson, E. H. (2002). New early zygotic regulators expressed in endomesoderm of sea urchin embryos discovered by differential array hybridization. *Developmental Biology* **246,** 132-147.

Rappaport, R. (1996). *Cytokinesis in Animal Cells*, pp. 386. Cambridge, UK: Cambridge University Press.

Reverte, C. G., Benware, A., Jones, C. W. and LaFlamme, S. E. (2006). Perturbing integrin function inhibits microtubule growth from centrosomes, spindle assembly, and cytokinesis. *Journal of Cell Biology* **174,** 491-497.

Roovers, K., Davey, G., Zhu, X., Bottazzi, M. E. and Assoian, R. K. (1999). α5β1 integrin controls cyclin D1 expression by sustaining mitogen-activated protein kinase activity in growth factor-treated cells. *Molecular Biology of the Cell* **10,** 3197-3204.

Schultz, J., Milpetz, F., Bork, P. and Ponting, C. P. (1998). SMART, a simple modular architecture research tool: Identification of signaling domains. *Proceedings of the National Academy of Sciences of the United States of America* **95,** 5857-5864.

Schwartz, M. A. (1994). Integrins as Signal Transducing Receptors. In *Integrins: Molecular and Biological Responses to the Extracellular Matrix* (ed. D. A. Cheresh and R. P. Mecham), pp. 33-44. San Diego, California: Academic Press Inc.

Schwartz, M. A. (2001). Integrin signaling revisited. *Trends in Cell Biology* **11,** 466-470.

Sea Urchin Genome Sequencing Consortium. (2006). The genome of the sea urchin *Strongylocentrotus purpuratus*. *Science* **314,** 941-952.

Shattil, S. J. and Newman, P. J. (2004). Integrins: dynamic scaffolds for adhesion and signaling in platelets. *Blood* **104,** 1606-1615.

Shimaoka, M., Takagi, J. and Springer, T. A. (2002). Conformational regulation of integrin structure and function. *Annual Review of Biophysics and Biomolecular Structure* **31,** 485-516.

Sieg, D. J., Hauck, C., R., Ilic, D., Klingbeil, C. K., Schaefer, E., Damsky, C. H. and Schlaepfer, D. D. (2000). FAK integrates growth-factor and integrin signals to promote cell migration. *Nature Cell Biology* **2,** 249-257.

Simon, K. O. and Burridge, K. (1994). Interaction between Integrins and the Cytoskeleton: Structure and Regulation. In *Integrins: Molecular and Biological Responses to the Extracellular Matrix* (ed. D. A. Cheresh and R. P. Mecham), pp. 49-69. San Diego, California: Academic Press Inc.

Smith, J. C., Symes, K., Hynes, R. O. and DeSimone, D. (1990). Mesoderm induction and the control of gastrulation in *Xenopus laevis*: the roles of fibronectin and integrins. *Development* **108,** 229-238.

Smith, J. W. (1994). The Structural Basis of Integrin-Ligand (RGD) Interaction. In *Integrins: Molecular and Biological Responses to the Extracellular Matrix* (ed. D. A. Cheresh and R. P. Mecham), pp. 1-31. San Diego California: Academic Press Inc.

Sousa, A. D. and Cheney, R. E. (2005). Myosin-X: a molecular motor at the cell's fingertips. *Trends in Cell Biology* **15,** 533-539.

Stewart, P. L., Makabi, M., Lang, J., Dickey-Sims, C., Robertson, A. J., Coffman, J. A. and Suprenant, K. A. (2005). Sea urchin vault structure, composition, and differential localization during development. *BMC Developmental Biology* **5,** 3.

Stupack, D. G. and Cheresh, D. A. (2002). Get a ligand, get a life: integrins, signaling, and cell survival. *Journal of Cell Science* **115,** 3729-3738.

Stupack, D. G., Puente, X. S., Boutsaboualoy, S., Storgard, C. M. and Cheresh, D. A. (2001). Apoptosis of adherent cells by recruitment of caspase-8 to unligated integrins. *Journal of Cell Biology* **155,** 459-470.

Susan, J. M., Just, M. L. and Lennarz, W. J. (2000). Cloning and characterization of alpha P integrin in embryos of the sea urchin *Strongylocentrotus purpuratus*. *Biochemical and biophysical research communications* **272,** 929-935.

Tadokoro, S., Shattil, S., J., Eto, K., Tai, V., Liddington, R. C., de Pereda, Jose M., Ginsberg, M. H. and Calderwood, D. A. (2003). Talin binding to integrin β tails: a final common step in integrin activation. *Science* **302,** 103-106.

Tamura, K., Dudley, J., Nei, M. and Kumar, S. (2007). MEGA4: molecular evolutionary genetics analysis (MEGA) software version 4.0. *Molecular Biology and Evolution* **24,** 1596-1599.

Terasaki, M. (1996). Actin filament translocations in sea urchin eggs. *Cell Motility and the Cytoskeleton* **34,** 48-56.

Thiery, J. P. and Sleeman, J. P. (2006). Complex networks orchestrate epithelial-mesenchymal transitions. *Nature Reviews Molecular Cell Biology* **7,** 143.

Thompson, J. D., Plewniak, F. and Poch, O. (1999). A comprehensive comparison of multiple sequence alignment programs. *Nucleic Acids Research* **27,** 2682-2690.

Travis, M. A., Humphries, J. D. and Humphries, M. J. (2003). An unraveling tale of how integrins are activated from within. *Trends in Pharmacological Science* **24,** 192-197.

Vandre, D. D., Davis, F. M., Rao, P. N., and Borisy, G.G. (1984). Phosphoproteins are components of mitotic microtubule organizing centres. . *Proceedings of the National Academy of Sciences of the United States of America* **81,** 4439-4443.

Vilela-Silva, A. E. S., Werneck, C. C., Valente, A. P., Vacquier, V. D. and Mourao, P. A. S. (2001). Embryos of the sea urchin *Strongylocentrotus purpuratus* synthesize a dermatan sulfate enriched in 4-*O*- and 6-*O*-disulfated galactosamine units. *Glycobiology* **11,** 433-440.

Wang, J., Mayernik, L., Armant, D. R. (2006). Trophoblast adhesion of the peri-implantation mouse blastocyst is regulated by integrin signaling that targets phospholipase C. *Developmental Biology* **302,** 143-153.

Wei, Z., Angerer, R. C. and Angerer, L. M. (2006). A database of mRNA expression patterns for the sea urchin embryo. *Developmental Biology* **300,** 476-484.

Whittaker, C. A., Bergeron, K., Whittle, J., Brandhorst, B. P., Burke, R. D. and Hynes, R. O. (2006). The ecinoderm adhesome. *Developmental biology* **300,** 252-266.

Whittaker, C. A. and DeSimone, D. W. (1993). Integrin alpha subunit mRNA's are differentially expressed in early *Xenopus* embryos. *Development* **117,** 1239-1249.

Wolpert, L., Jessell, T., Lawrence, P., Meyerowitz, E., Robertson, E. and Smith, J. (2007). Development of nematodes, sea urchins, ascidians, and slime molds. In *Principles of Development*, pp. 202-212. New York: Oxford University Press.

Wong, G. K., Allen, P. G. and Begg, D. A. (1997). Dynamics of filamentous actin organization in the sea urchin egg cortex during early cleavage divisions: implications for the mechanism of cytokinesis. *Cell Motility and the Cytoskeleton* **36,** 30-42.

Wong, M. L., and Medrano, J. F. (2005). Real-time PCR for mRNA quantification. *BioTechniques* **39,** 1-11.

Xiong, J., Stehle, T., Diefenbach, B., Zhang, R., Dunker, R., Scott, D. L., Joachimiak, A., Goodman, S. L. and Arnaout, M. A. (2001). Crystal structure of the extracellular segment of integrin alpha5beta3. *Science* **294,** 339-345.

Yamada, K. M. and Even-Ram, S. (2002). Integrin regulation of growth factor receptors. *Nature Cell Biology* **4,** E75-E76.

Zhu, X., Mahairas, G., Illies, M., Cameron, R. A., Davidson, E. H. and Ettensohn, C. A. (2001). A large-scale analysis of mRNAs expressed by primary mesenchyme cells of the sea urchin embryo. *Development* **128,** 2615-2627.

Appendix I
Primer Sequences

Primer Name and Sequence (5' – 3')	Annealing Temperature	Extension Time/Product Size
Alpha C 15379F1 CCATTCTCCAGGAGAACACCCAAAC **Alpha C 15379R1** TGTGATGTCCTTTCCTGCTGTGAGT	56 ºC	1 minute 946 bp
Alpha C F2 AAGGAAACACCTGACTCTATC **Alpha C R2** AAGTAGGCAGGATAGGTGAC	60 ºC	1:30 1446 bp
Alpha C RACE Middle F1 TATATACCAGCCCGCCAGCGTG **Alpha C RACE Middle R1** CCTCGCCATTGTTCTCGTACGG	58 ºC	1 minute 820 bp
Alpha C 5' RACE Inner CGCATTCTTTGCTGAAGATGACC **Alpha C 5' RACE Outer** CCGATGAGAGGGAAATGAATCC	60 ºC 60 ºC	2:30 2:30 2150 bp
Sp-AC F1B GGATCCATGACGAAGATACCCGTCGCTTT **Sp_AC R1C** CATGTCGTCGTAGTAAGCGG	68 ºC	2:30 ~3100 bp
Alpha D1 F1 CGACCGATATTTACGAGGTGGCTG **Alpha D1 R1** CTGGCAAAGACAAGCGAGTTGAGTG	60 ºC	1:30 1412 bp
Alpha D1 F2 TCGCTACTCTTCGCCATCAAATCAG **Alpha D1 R3** TCGGACTGTCTCTGTTGATCCTTGG	62 ºC	1 minute 914 bp
Alpha D1 F5 CGCCGTCAATAACCAGACACTCAAC **Alpha D1 R5** GAGAAGACCGCCTAGTGAGGAGACG	60 ºC	1:30 1143 bp
Alpha D2 F1 GCCCGACATCGCTCTCAATG **Alpha D2 R1** ACAAGAACACCCGCCCAAATC	62 ºC	1 minute 917 bp

Alpha D2 F2 ACGATGTCCGGCCCGACA **Alpha D2 R2** AACAGCGTGACCATTCCACAAAC	59 ºC	1 minute 742 bp
Alpha D3 F1 CAATCCTGTACCAGCCGGTGATC **Alpha D3 R1** TTACTCTTGTATGAGGCCTGCCAC	65 ºC	1:30 1041 bp
Alpha F F2 GAGATACTGCCCATTCTGGACG **Alpha F R2** ACGCCATTACTGCCCTCTAGC	60 ºC	1:30 1362 bp
Alpha F F3 CTTTACTCGCATTCTGTCATTG **Alpha F R4** TGATCCAAATGTGTTGAGGC	58 ºC	1:30 1346 bp
Alpha F F4 GGCAATGGTTAGGAGCAAC **Alpha F R3** CTCTGATGACTGGACCATGC	55 ºC	2 minutes 1720 bp
Alpha F F6 CTGGAATAGGAAGAAAACCATG **Alpha F R6** CAAATAGGCGTTCATTGTCG	56 ºC	1 minute 905 bp
Beta D1 F5 ATGGGGGTGCCATGGAGAGTATG **Beta D1 R5** CATCCAACAACCTCGTCACAAACAG	60 ºC	1 minute 757 bp
Beta D1 F4 CAGTGGTCTTCTTGACATTGGTAGG **Beta D1 R1** TCCCTAACCAGACTCGACAAGTGACC	60 ºC	1:30 921 bp
Beta D1 F2 TTGGATAAACCGGAAGCTGGTCTTG **Beta D1 R2** CCACTGTTGAAAGCCTTGCATTGG	55 ºC	1:30 1264 bp
Beta D1 F3 TGACCGAACCACTCAGCAAAGCAC **Beta D1 R3** TGGCTTGTACAGCGGAGGCTCGTTC	55 ºC	2:30 2177 bp
Beta D1 F3 TGACCGAACCACTCAGCAAAGCAC **Beta D1 R4** AGGTTATGATTTTTCTTGCAGG	58 ºC	1 minute 738 bp

Sequencing Primers for Alpha C EST Clone		
Spu_AlphaC C3 M13F1 ATGAAATACAAGACCCCTCC **Spu_AlphaC C3 M13R1** AGGACTAAAGCCCAAGAACG		
Primers for Quantitative PCR		
Ubiquitin F1 (Sea Urchin Genome Project, Caltech) CACAGGCAAGACCATCACAC **Ubiquitin R1** GAGAGAGTGCGACCATCCTC	60 ºC	30 seconds 147 bp
Spu_AlphaC QPCR F1 TCGGTGAACAGCAGGAAATC **Spu_AlphaC QPCR R1** AGTGAATGCCTGTGAAGCCG	60 ºC	30 seconds 136 bp
Spu_AlphaD QPCR F2 CGCCAATGTTTACCGATGAGG **Spu_AlphaD QPCR R2** AATAGCCAAAGCGACAGCCAG	60 ºC	30 seconds 149 bp
Spu_AlphaF QPCR F4 ACGCCTATTTGAACCTCACGG **Spu_AlphaF QPCR R4** ATCTCCCATCCGTCAGTCGTC	60 ºC	30 seconds 150 bp
Spu_BetaD QPCR F1 ATGCTCATGGTACCTGCAC **Spu_BetaD QPCR R1** ATTGCTCCTCTGTGAGTGG	60 ºC	30 seconds 164 bp
Sp BetaC QPCR F1 AGCTGGTGAAATGAAGAACATC **Sp BetaC QPCR R1** GGAAGGAGAATGCTGGACC	60 ºC	30 seconds 148 bp
Control Primers		
Spu_TBPint F1 CAGGATGGAGGGCAACAGAGTC **Spu_TBPint R1** GCATGGAGGGCAATTTTCTTCAGATC	60 ºC	30 seconds ~300 bp (cDNA) ~800 bp (genomic)
Morpholino Antisense Oligonucleotide		
Spu_αC MASO TATTGATGATGATTCTGTCCACTAC		

Appendix II
Protein alignments

>*S. purpuratus* full length α integrins

```
                    10         20         30         40         50
Sp Alpha C   -MTKIPVALA TILCVLVAVV ERTLCFNFET RLPVIKEGPE NS-YFGFSVA
Sp Alpha F   -MAAAGIRFF TRILSLWTIL TVVQGFNMDL NAPVLQNISD TSSYFGFSVT
Sp Alpha D   MTTGYIFSIA LCITICSRWA LVVHGFNVDT SFPLYFTHDD TDSQFGLSLA
Sp Alpha P   ----MARFLL LSFCTTLVLL DSTVAFNFDL RAPVKFDGPQ GS-LFGFSVA

                   60         70         80         90        100
Sp Alpha C   EHILTKTKPT TPFAESLLIV GAPIYQTN-- ---IDAIKPG GVFTCPFTSN
Sp Alpha F   MHRHNGE--- -----NMVIV GAPRFQTS-- --EPDVTRGG ALFRCPLTPL
Sp Alpha D   LHSEGKESEG K---ENMLLV GAPRAMSSGY SADQSIVRSG ALYKCPIDTS
Sp Alpha P   QHRDQNT--- -----DWVLI GAPEAPTT-- --QPGVTNGG AVYKCPVTPL

                  110        120        130        140        150
Sp Alpha C   ---YSDCTVL ELDENGN--- ------TESD IKDNQWLGVA VVSQGPGGPV
Sp Alpha F   TNPGECEELS AFDALGTDDN EVG---AKIA DKSGQWLGAT VKSTGENGKI
Sp Alpha D   IGTRADCQEI ILDTTGNEYL DTNNPDFQST NASGQWLGST LQSGGPDGSV
Sp Alpha P   SGSGPCEQVP -FDTTGN--- ------TEVL DKSNQWFGAT LASSGPDGRI

                  160        170        180        190        200
Sp Alpha C   AACAHRHIYQ PAS-----VD EIQGLGKCYF MSADLQKDVF IRDGYQPCLS
Sp Alpha F   LACRPLYTWF IKP-ADATKT AREPIGGCFL ANSDFTN--- --IEVYEPCR
Sp Alpha D   VVCRPLYSYQ TIN--QNQDI DRYLRGGCFV INGGLRSPDP GLLNTAVNFN
Sp Alpha P   LACAPRLVWL QTSTISPTDK EREPTGTCFV GHSDFTN--- --FVNYSPCQ


                  210        220        230        240        250
Sp Alpha C   QTRPNNYDGI LYKYCQAGTD VDVTRLKYTD STSVSEYIMG VPGSDNWLGG
Sp Alpha F   TSRESEQSVF GITHCEAGIA ADITSD---- -----EILIG APGSFFWQGQ
Sp Alpha D   PCAANTVGYR TKYRCEFGGS AAVRDT---- -----KIVYG ATGGCLWNGR
Sp Alpha P   STDRDLFGFD KITHCQAGFS AQIPSDNS-- -----TLVMG APGSYYLQGQ

                  260        270        280        290        300
Sp Alpha C   IALANVPDSI RSINLITPLD FENNKVGFNS YLGFSVASAK LMDGNVQTYI
Sp Alpha F   IWVTGT-DLQ SERR-TNEGP VRFD----DS YRGYAVAFGD FKGDSRPEYA
Sp Alpha D   IWYTND-GIT DASALFYFPQ TCRFNLDFEY YFGHAVAFGN FRDGPTLEVA
Sp Alpha P   IFAQSL-STL SDVSNTPEQA VAFD----NS YRGYSLALGD FNGDGLEDYV

                  310        320        330        340        350
Sp Alpha C   AGAPRSESMG AVLTLKSEAG QSPSPIYRID AEKPSSSFGY DVAAADVNGD
Sp Alpha F   VGIPRADNLV GIVSIFDST- -MNAYLN-LT GSQTGAFFGH TLVVSDFNDD
Sp Alpha D   IAAPRWDHGF GRVFLYDDA- -FMEYDI-IT DGDLATYFGH ALATSDLNND
Sp Alpha P   VGTPRGESLR GLVAIFDQS- -LVEIITPVV GEQIVSYFGY SVASVDVNGD

                  360        370        380        390        400
Sp Alpha C   GYDDLIVGAP QYFRRDPTNY QGGEGGRVYI YLNEAHDGTF ENVEPIKLTG
Sp Alpha F   GFTDLVVSAP FYIDESKTTD G-WEIGRVRI FYNDENG--G FSEGTTIQGY
Sp Alpha D   GFDDLVIGAP MFTDEERSID AGWEVGKIFI YYNDQQGSFY FDDNDVIIGT
Sp Alpha P   GLDDLLVGAP MFTNREPATE K-WEAGRVYV YLQNADH--S LGAPQMLTGK
```

```
                    410        420        430        440        450
Sp Alpha C   KIDSLFGQVV TNIGDINLDG FEDIAIGAPY EN-NGEGAVY IYLGH-GQDG
Sp Alpha F   KVRARFGFSM AALSDVNQDG YNDLAVSAPY GGKTGEGIVY IYHSQ-GRLG
Sp Alpha D   GAGCRFGYSI AALGDINQDG FNDLAVGAPF CNGGNDGKVF IYHGSGINLP
Sp Alpha P   KIRARFGFPI TSIGDSNQDG FNDVAIGAPY DGEDNSGVVY IYHG--SAEG

                    460        470        480        490        500
Sp Alpha C   IRQPAAQKIT PSDLPNGING YPFPNTSFGY SISGGVDLDG NGFPEVAIGA
Sp Alpha F   LSTTFAQSLE PSDFGLS--- ----LNTFGS SLSAGMDMDN NMYPDLVVGA
Sp Alpha D   LNLSPQQTLS PSVLGRP--- ----LRFFGF VMSAGLDVDQ NTYPDFAVGA
Sp Alpha P   LRLTESQVLT PSELGFSD-- ----ITTFGF SVDGGQDMDQ NDYPDLVVGA

                    510        520        530        540        550
Sp Alpha C   YEVQQIAVLR GRPVINIEAQ LTLSVTELDP NTTNCEYS-G KDSLCFTVQL
Sp Alpha F   YKSSTVVLVR SRPVVHVTGS LVSKTESIDL KNTNHNTASG TAVSSFDFEV
Sp Alpha D   HESQTAIIFR SRAVVWTVAE IYPEENAIDL DVKDETTSRG VLVTGFNVVV
Sp Alpha P   ESADAAILVR TRPVVLLEAE LTIEPIGINL DNKTYELPDG TMVTSFVAMA


                    560        570        580        590        600
Sp Alpha C   CMSYYCLAPA FNDPITINFD IEAEGVRRSK VLNSRVVFED SGIYSLTKQT
Sp Alpha F   CIRYSGTN-- VPPNLDFDYS LSLDYNRQST -RRANFDLGG NDASSVSLTV
Sp Alpha D   CIHYMGQG-- LPANLDFSYE IVLDSLRQVI NRRAAFITN- -NIATLRHQI
Sp Alpha P   CFTYTGNH-- LPARIGISYT LTVDSSITSG -RRALLEVN- -ELSQVTKNR

                    610        620        630        640        650
Sp Alpha C   LVITNDDEAC TPEYNVILKA GFTDIFRPIP INLTYYIPEV EAVMPQPGDP
Sp Alpha F   TLTSNVEYCN P--FTAYAKP TIVDKLSPIP VSLSFGLPDS -SPAPG----
Sp Alpha D   SAPIEQRTCM T--HAAYIQP TIRDKQTPMS LRLIHSVAQT -SQSDS----
Sp Alpha P   NLDVGMKFCD P--LRAYVVN TIQDKLTPIA VDLQYELTDE SILLPY----

                    660        670        680        690        700
Sp Alpha C   LPSMTPFPIL QENTQTIMLE EVIFSKECAK DDGLCITDLD VRASVDLEGT
Sp Alpha F   ----EILPIL DESSDNFRAL SVPIERNCMN --ETCVPDLR VRASTETA--
Sp Alpha D   ----AVQPIM SNAVNNQTLN SLVFARDCAG --ETCYPDLA VQTIVSTT--
Sp Alpha P   ----EILPII NKEAVSSQTK QVSIQNNCVN --NICIPEIG ITVTPNLP--


                    710        720        730        740        750
Sp Alpha C   PPILKVGEQQ EISMSAHIAN REEDAYNAKL VVTYPAYLGF ISLSS-ASQA
Sp Alpha F   --SLLVGSSE SITIEVTIDN NGEDAYLSTL TVEVPQELQY AGFTR-VQTD
Sp Alpha D   --ELMIGKAE TFLLTVEVMN SGEDAFLSVL DIVEPPGLFF VNVLR-SNTG
Sp Alpha P   --NIVIGQAQ ELTLVVSINN RGEDAFQSTL AVYYPEGLQY VRLERRANMD

                    760        770        780        790        800
Sp Alpha C   FTADCNPEPF EEGANETSII CELGNPYSEG SVDSFVVLYD ASSVPPDAEV
Sp Alpha F   QIVTCN---- -SMVQSSEIS CDVGNPFTAG TTVILKMQFS STNLPGNGDN
Sp Alpha D   IIVSCS---- -VSTTTRIRS CNIGNPVPAG DQITVGLQYQ TATYSTFTGP
Sp Alpha P   FSVTCT---- -EDSALRMIT CDTGNPLVGK YNLEFGLTLS TLQVSGDKDN

                    810        820        830        840        850
Sp Alpha C   FNITLVASST NDAD-SNPGN NQYIITIEVE SITDILISGK -GPEQLYFSG
Sp Alpha F   VHFTFGVSSV ESE--QMIAD NEFNITVPLA IAADLSFHGN VIPETVIYSE
Sp Alpha D   AVLNISTRSI DAERAGRDVD NEAMVSIPVY ATSSLALSGN SIPDTLVITE
Sp Alpha P   IEFYLVAGSE NNEDPNTLDN NELNVTVAVI VDATLKLLSA SYPEIVTYRV
```

```
                    860        870        880        890        900
Sp Alpha C   QVIG-----E SAMNYFEDIG LAVNQRWTIF NEGPGAVNTA RVTIDFPYEV
Sp Alpha F   EN----YLLT RAIKLEKDVG PDITHVFSLY NNGPSSVART EITIPWPMRY
Sp Alpha D   DS-------- ---NATDDKG PLMTHVIRLQ NNGPSFIGPS TIEILWPIRL
Sp Alpha P   PEDNIVPEFP TKNASEADIG MEVVHLYEVR NTGSSNAAEV TLNIRWPEKD

                    910        920        930        940        950
Sp Alpha C   ANGKWLLYMT EMPFVEDNKG SCNVTPAVYV NELGLKPKNG GGAYNPVAPG
Sp Alpha F   NGVQRNYLLY LLSATMDTGE ECTIDGPWNP EGLGSVDNST GTPESKRRRR
Sp Alpha D   Q--DGKKLMN VTEVTMDTGM PCRLT----- ---------- ---------P
Sp Alpha P   E--NGDYLFY LLGIMTDEGV TCQIS----- ---QGQANPL GVKLEASTKE

                    960        970        980        990       1000
Sp Alpha C   TGGTTQSRKR RQAEAEAVTS TRELTAGKDI TLDCKTG-TA QCLTIICDLD
Sp Alpha F   QAATTDSPAV GTTDIPMSTT ESPVKV---R DIKCSGPQSQ FCYMITCVVD
Sp Alpha D   RINRDSPKQN NVTTFVDDVS NSTIYT---L WADCDS--LP ECVAIRCPLA
Sp Alpha P   QLSNSTTQVS GRRKREGEYA EALAQA---E PIFCTP---E SCVLINCTID


                   1010       1020       1030       1040       1050
Sp Alpha C   PLTKETDMNE VNIVVRSRLW NSTFLEDYIN ARKVKIITEG YVEIESAPYI
Sp Alpha F   SLGAG-SNAA RDNVIVTVRS RVWTNTLTVN LLELTVFVQA HAQVLEVPYT
Sp Alpha D   SLATGADSAV LISVQSRLWH DTVFAEDSSF LHFIEILSSV SVTVNGTVYP
Sp Alpha P   EIKATKSKVV RILGRFWERT FQKAVSEAVP VIQVTLASTA TATVRSIPYN

                   1060       1070       1080       1090       1100
Sp Alpha C   ----TQTRED NDKFRLALTI TPDIKTLPPS KPLQWWIIAL AVIGGIILLI
Sp Alpha F   -TTFTDPIEI SSNEVSTKIV PQFLIEEVEP EKVALWVYLV SILCGCLILV
Sp Alpha D   GIPYSPHPTA TPLEIQTEVR VQYEKQRTII KTTPIWIYIV SSLGGLLILI
Sp Alpha P   -IPLPMEFTD STKASTLITA EELVLPPVS- --IAWWIIVV SVLGGIILLL

                   1110       1120       1130       1140       1150
Sp Alpha C   LLILLLWKCG FFERKTGYKY ATVTQQQASG ---------- ----------
Sp Alpha F   LLIIVLWRVG FFKRKKKLTA DQRKMLEGSN ---------- ----------
Sp Alpha D   IFIAILYKVK FFQRKQISQE QREILSEIKR GTRSYDDAIR MSRFMNPLFD
Sp Alpha P   IIILGLWKCG FFERKK--PG EDQKEYEPVA ---------- ----------

                   1160       1170       1180       1190
Sp Alpha C   -KAKTEKTAY YDDMYY---- ---------- ---------- ---
Sp Alpha F   GVENEGLEGV DKTEE----- ---------- ---------- ---
Sp Alpha D   GDQRPNHALA SQDKDRLSAN AKVAFLSLLS KHCRKRALRL RGT
Sp Alpha P   VTEKDGPPEV YDAPDRSSAN SNTFLVCSY- ---------- ---
```

>*S. purpuratus* full length β integrins

```
                     10         20         30         40         50
Sp Beta D    ---------M GVPWRVCMSV VFLTLVGYCS GQDS---EEC RLAVNCADCI
Sp Beta C    MKLHRRPQ-L LFQWTISAVI FFLSLHVFLD AQSTGTNSAC LQAKNCGDCI
Sp Beta L    MPSVRLPHRT TRPGSVVVFF LTFVLAVFTV HANEELSCDL SRAQNCGECI
Sp Beta G    ---------M AWKCLLVLTI AFSAGQLYRT YAQT-SSSAC SDAKTCGECI
```

```
                    60         70         80         90        100
Sp Beta D    SVGPACGWCA QRDFFN---- ASCDLSSTLL VSNCS--DIQ NPEAYAVNIT
Sp Beta C    ITDPSCAWCA QPDFQDAENF PRCDNPDTLR ERGCQELHIE NPQSTMN-IT
Sp Beta L    SVNPECTWCK EDVFEG---- RRCDLEILLQ EAGCG--NIT NPLPSAV-PI
Sp Beta G    SLDSSCGWCT LLNYTDDTGN PQCDLASSLS QRGCS--QIV DPDSTM--VL

                   110        120        130        140        150
Sp Beta D    QNEELSNAGD APLGQAVQVQ PQEIHLTIRR GETATLRMNV RQAEDYPVDM
Sp Beta C    GKTPLSKAG- DPVENLVQVY PQVVDLTMRP GEKISIDLQI RQAEDFPVDL
Sp Beta L    EDKPLSEAN- ADLDAIVQVK PQMMRIKVRP REPINIKLYV RQAEDYPVDL
Sp Beta G    GWISLSNAGS APQGQAVQVR PQQVDLKLRR GKPVVMTLPG PTAEDYPVDL

                   160        170        180        190        200
Sp Beta D    YYLMDLSQSM KEDLETIRQL AGDLVDKMKE LTRNFQVGFG SFVDKNIIPF
Sp Beta C    YYLMDLSDSM SDDLVQLRSL GGILAGEMKN ITNNFRLGYG AFIDKTVMPY
Sp Beta L    YYAMDLSHSM KDDLENLKGL GTTLSEELNS ITRDFRLGFG SFVDKTVLPY
Sp Beta G    YYVMDLSKSM EDDLSKLMDL GDILASEMKN ITSNFRLGFG SFVDKTVMPY

                   210        220        230        240        250
Sp Beta D    ASLKENGEVL LDCAP-ECED PYLFKNGQTL TNSSELFGQT LGETGYSTSL
Sp Beta C    VDIYP-AKLE NPCLNKRCGP AFSFH-ILPL TLETDRFTEE ISKVNSSGNL
Sp Beta L    VSTVP-AQLI SPCTG--CAS PHGFHNALPL NQDPSLFANR ITNTTVSGNL
Sp Beta G    VSTVP-EKLI APCTG--CEA PYGFKNVLPL NENTNLFSET VMNQRASGNL

                   260        270        280        290        300
Sp Beta D    DKPEAGLDAL LQVAVCDEVV GWRDSARHLV VFLTDAPYHA AGDGRLGGIV
Sp Beta C    DSPEGGMDAL MQATVCTDEI GWRERARHLL VYTTDASFHI AGDGKLGGIV
Sp Beta L    DTPEGGFSAL MQIAVCGEVI GWRPKARHLV IFTTDASFHF AGDGRLGGIV
Sp Beta G    DAPEGGMDAL MQITVCGNQI GWRENARHLV IYTTDSS-SI MPETKLGGII

                   310        320        330        340        350
Sp Beta D    TPNDGLCHVD PVTGEYMIDK LQDYPSLGHL SSLVREKDII MMFVVG----
Sp Beta C    KPNDGKCHMD STGFEYTMAN EMDYPSISKL SQKMETLSIL PIFAIGKAEV
Sp Beta L    EPNDGQCHMD PNTNMYDFST LQDYPSIGHL SAKLRENNVI PIFAV-----
Sp Beta G    TPNDGQCYLD PISQNYTMSH YLDYPSIRHL NAKMRENSVI PIFAV-----

                   360        370        380        390        400
Sp Beta D    --AEVQQTYR DILTFFPGSQ IGTLAANSDN ILFWIENLYQ KITSVVTLSV
Sp Beta C    DKQDPFVFYE DLPQYFHESK AARLSADSSN IVDLIKNIYL NITSEVTVET
Sp Beta L    -TRDQTPLYM SLEKDIEGAT VGTLDEDSGN VVQLIRSNCD RITSQVRLTS
Sp Beta G    -IQKEFEIYN NLTQYIEGAT AGILAQDSNN IVQLVKDNYS KITSRVEVVD

                   410        420        430        440        450
Sp Beta D    QA-PEELLLE FTSHCDVDPT APGVPGEMTC EGLAIGDTVW FEIDVTAEE-
Sp Beta C    RLGADLFEVD YVAHCLDGSI TKDKQ---TC MGLKLGDQIS FDVGITMKNL
Sp Beta L    TA-PDDVTLS YRANCKD-QT YQDTN---EC SGLSLGDTVS FDITLTAER-
Sp Beta G    DA-PENVTID YGPHCPGGQV TPGSQ---VC EGLQLGDTVN FTLTITAT--

                   460        470        480        490        500
Sp Beta D    -CIEGGTT-S FSVSPVGFSE ELLVMVDVIC GCDCE-KEGI SNSPECNYNG
Sp Beta C    SCTVQQWNAH DSVGRPVFTE NLVLNVKALC ECDCSSSEQE PNSTKCNFHG
Sp Beta L    -CVEGGMT-S FNIGPVGFNE ELQIELEVTC ECDCQGLEEA NSTVCSSGNG
Sp Beta G    GCPPNKY--Q QFTVRPLIQR ELKVNVEFAC DCDCEAQKVE NSQVCSGGNG

                   510        520        530        540        550
Sp Beta D    TNQCGMCACN MGRYGDECQC SGESSSVGD- DRLACIASGS LSICSGRGEC
Sp Beta C    TFTCGACMCN EGRSGRTCEC --DTVEAEGL DPS-CVQPNS TVECSSRGTC
```

```
Sp Beta L    TLVCGECDCN PGRYGVKCEC SGNEINMEST DPSPCRTDNT TRTCSGRGEC
Sp Beta G    TLECGSCICN PGHYGRYCEC SSDDPTLEDN DAP-CDITKH IHRCSGRGSC


                      560        570        580        590        600
Sp Beta D    ICGECVCNKQ KDPTQVVSGE FCQCDNFNCP LFEGQLCGGP SRGTCVCDRT
Sp Beta C    VCGECECDTR GDPNRIITGE YCQCDNYLCP RSGGEVCGGS DKGTCLCDED
Sp Beta L    ICGKCVCDNT GNPGEVISGQ FCECDNFNCP YSRGLRCGGP DQGMCVCDVA
Sp Beta G    VCGNCICFPR PNPSEVVSGT FCECDNFNCD RYLGELCGGS DRGQCVCDEY

                      610        620        630        640        650
Sp Beta D    TQQSTCQCTA EYEGDSCNCL RSQEACIANN GLICNAHGTC TCGACRCDTD
Sp Beta C    VG-NFCGCLE GYQGSACECP TSNDTCRAPN GEICNGVRTC DCGKCQCND-
Sp Beta L    TRQPKCQCNP GFEGDSCDCP TRFDMCAASN GLECNAHGTC VCGCLRVFAD
Sp Beta G    TRRSQCRCRS GYTGDACECS TRVDTCMTG- DTICNGEGVC ICGECKCNAG

                       660        670        680        690        700
Sp Beta D    SPFKGPTCFD CATCT-GQCE VYQECVQCKA FNSGPLTEEQ CDMCTVEVII
Sp Beta C    PKYSGATCQI CPDCA-GECD INQPCVQCRA FHTGAYNKSQ CKNCPHPIYV
Sp Beta L    SQFQGKTCEK CPTCAFGICH IHRDCVECTV FGTGRLTPEQ CDMCTVNIIN
Sp Beta G    SSYRGALCQD CPTCS-GQCS RNEECVQCKA FGTG-LSKAD CDKCPFPVIM


                      710        720        730        740        750
Sp Beta D    VERFPEVDYP YN-CSVNHID NCTIVFIPYF MDGEDAVPVI YVLEDTVCPR
Sp Beta C    VKELRR-EEG HQECRFTDDD ECHVIFT--Y EILPNGTYYI EVQEDKICPG
Sp Beta L    VTSIDEYTED NPKCNFPLSD DCTFQFV--V VS-ENETVTV YVEGRETCIE
Sp Beta G    VDNLEI-PTG SERCLAEDED DCSIIFT--Y AKSANLALIL YVQKEKVCFE

                      760        770        780        790        800
Sp Beta D    GPVAPG---V DARWVAVGLA IAIVLIGILL IILYKCYITW LDRKEYNEWE
Sp Beta C    AP-------- EMLWVILGII IGVFLVGLAL LLIWRLLTYI HDRREFQNFE
Sp Beta L    PVGKPTLLGG RIRWIVIGII LGIVLIGMIL ARAWRLYTYV QDKREYAQWE
Sp Beta G    PV-------- DIMHVIIGIV VGIIIVGLAL LLVWRLLVYV QDSREFASFE

                      810        820        830        840
Sp Beta D    KEKANAQFSK NENPLYKPAS TVHKNPAYGA QFSDTGEQAE ALQEKS
Sp Beta C    KERANATWEG GENPIYKPST SVFKNPTYNI K--------- ------
Sp Beta L    NDCKKAQWDQ SDNPIYKSST TTFKNPTYGK ---------- ------
Sp Beta G    KERAGTHWGQ NENPIYKPST STFKNPTYQK ---------- ------
```

>Integrin_α2 domains

```
                       10         20         30         40         50
                                                        **
Sp Alpha C   ---------- ---------- ---------- ---------- GRPVINIE--
Sp Alpha F   FGLSLNTFGS SLSAGMDMDN NMYPDLVVGA YKSSTVVLVR SRPVVHVT--
Sp Alpha D   ---------- ---------- ---------- ---------- SRAVVWTV--
Sp Alpha P   ---------- ---------- ---------- ---------- TRPVVLLE--
Sp Alpha J   ---------- ---------- ---------- ---------- SRPVVDVS--
Lv ASU2      ---------- ---------- ---------- ---------- TRPVVSLD--
Mm Alpha I   ---------- ---------- ---------- ---------- AQPGVMAT--
Mm Alpha V   ---------- ---------- ---------- ---------- ARPVVTVN--
Mm Alpha 8   ---------- ---------- ---------- ---------- ARPVVTVD--
Mm Alpha 7   ---------- ---------- ---------- ---------- ARPVLHVS--
```

```
Mm Alpha 6   ---------- ---------- ---------- ---------- SRPVINIL--
Mm Alpha 5   ---------- ---------- ---------- ---------- GRPIISAS--
Mm Alpha 3   ---------- ---------- ---------- ---------- ARPVINILH-
Dm PS1       ---------- ---------- ---------- ---------- ARPIISIQT-
Dm PS2       ---------- ---------- ---------- ---------- SRPVAAVNAE
Ce INA-1     ---------- ---------- ---------- ---------- SRPVISIE--
Ce PAT-2     ---------- ---------- ---------- ---------- TKPVVTVT--
Nv Alpha 1   ---------- ---------- ---------- ---------- SRPIVDVV--
Ci A9        ---------- ---------- ---------- ---------- SRPVITIE--
Ci A11       ---------- ---------- ---------- ---------- ARPVVKLL--

                      60         70         80         90        100
                                                     *
Sp Alpha C   AQLTLSVTEL DPNTTNCEYS GKD---SLCF TVQLCMSYYC LAP--AFNDP
Sp Alpha F   GSLVSKTESI DLKNTNHNTA SGT--AVSSF DFEVCIRYSG T----NVPPN
Sp Alpha D   AEIYPEENAI DLDVKDETTS RGV--LVTGF NVVVCIHYMG Q----GLPAN
Sp Alpha P   AELTIEPIGI NLDNKTYELP DGT--MVTSF VAMACFTYTG N----HLPAR
Sp Alpha J   AHVSVDLKGI RLKERNYRLP GGA--LVLGF NVTLCFSFDG I----NIPQE
Lv ASU2      ATLNTEPIGI NLENKTYELA DGT--MVTSF IAMTCFTYTG N----YLPDH
Mm Alpha I   VQLMVQDS-L NPTLKNCVLD QTKT-PVSCF NIQMCVGATG H----NIPQK
Mm Alpha V   AGLEVYPSIL NQDNKICPLP GTAL-KVSCF NVRFCLKADG KG---TLPRK
Mm Alpha 8   AQLLLHPMII NLENKTCQIP EFPT-PVACF SVRVCASIAG Q----SISNT
Mm Alpha 7   QEIFIDPRAI DLEQ---PNC ADG--RLVCV DIKICFSYVA VPS--SYSPS
Mm Alpha 6   KTITVTPNRI DLRQK--SMC GSP--SGICL KVKACFEYTA KPS--GYNPP
Mm Alpha 5   ASLTIFPSMF NPEERSCSLE GNP---VSCI NLSFCLNASG K----HVPNS
Mm Alpha 3   RTLVARPAVL DP-----ALC TAT----SCV QVELCFAYNQ SAGNPNYRRN
Dm PS1       SVQRKELHNM DPNTPGCLDD PAS--NLTCF TFRACCSIEP YD---EKNKE
Dm PS2       TSFASNSKLI SLDDRSCQLV RDHK-KVPCM LLTTCWSYTG R----YLPEQ
Ce INA-1     TKHKMEKRMV DIDKG-VNCP RGA---KTCF PLDMVIYVDE ETK--RGAEL
Ce PAT-2     GQTEPESALI SVEDKNCDVD GKLG-KQACK HINTCFKYEG KG---DTPND
Nv Alpha 1   GKIELSTNRI ILESNQTNLC TLQGNQHKCL TARVCLKYKH K---VPVSQG
Ci A9        SSLSASPTSL DLRNPNCTLA SKT---YSCF ELQFCFRYNA RHK--EYTEI
Ci A11       VIQTITPNKI DLETLSCDLN PTT--KSACF YVETCFSYSG K----TLPDT

                    110        120        130        140        150
Sp Alpha C   ITINFDIEAE GVR------- -RSKVLNSRV VFED--SGIY ----SLTKQT
Sp Alpha F   LDFDYSLSLD YNR------- -QST---RRA NFDLGGNDAS ----SVSLTV
Sp Alpha D   LDFSYEIVLD SLRQ------ -VIN---RRA AFIT--NNIA ----TLRHQI
Sp Alpha P   IGISYTLTVD SSI------- -TSG---RRA LLEV--NELS ----QVTKNR
Sp Alpha J   IFLDYDLIAD AKLP------ -VS----QRC YFELLDQPDT ---AGIRDRV
Lv ASU2      IDISYTVTVD SGI------- -IAN---RRA MFVD--NDMS ----EITKTR
Mm Alpha I   LHLKAELQLD LQKP------ -RQG---RRV LLLA--SQQA ----SLTLSL
Mm Alpha V   LHFQVDVLLD KLK------- -QKGA-IRRA LFLH--NRSP ----VHSKTM
Mm Alpha 8   IALLAEVQLD FLK------- -QKGA-IKRT LFLH--NHQS ----HFTFPF
Mm Alpha 7   VALDYMLDGD TDRR------ -LRGQ-VPRV TFLS--RGLD DLRHQSSGTV
Mm Alpha 6   ISILGILEAE KERR------ -KSGL-SSRV QFRN--QGSE ---PKYTQEL
Mm Alpha 5   IGFEVELQLD WQK------- -QKGG-VRRA LFLT--SKQA ----TLTQTL
Mm Alpha 3   ITLAYTLEAD RDR------- -RP----PRL RFAR--SQSS ----VFHGFF
Dm PS1       LRLAYSVEAE TFDH------ -LKKF--SRV FFFDRENKRT N---VLSRVV
Dm PS2       LDFDVSWLLD AKKL------ -LN----PRM FFLR--DEGK ---NIRNQTI
Ce INA-1     VDFSSDVFMC NLEAIP---- FRADT-TARG FIEG--SHSH ----NYSWPC
Ce PAT-2     LEFDLRFNLD DHSPEPRAYF LQKDVKSDRS IKVAQGSKTR DHPSSIEQRV
Nv Alpha 1   VELDFSLALD QDRR------ -TIEL--RRM FFDP--KNTT DKVFTKRGSA
Ci A9        LPIHYSITLD SELQ------ -KERK-AARV AFASP-QGPS ----LMSDVA
Ci A11       VNMSYSYDVD SGKE------ -ERE---KRS YLLD---DSP -------RNL
```

```
                    160        170        180        190        200
                                          *
Sp Alpha C    LVITNDDEAC TPEYNVILKA --GFTDIFRP IPIN-LTYYI PEVEAVMPQP
Sp Alpha F    TLTSNVEY-C NPFT-AYAKP --TIVDKLSP IPVS-LSFGL P-DSSPAPG-
Sp Alpha D    SAPIEQRT-C MTHA-AYIQP --TIRDKQTP MSLR-LIHSV A-QTSQSDS-
Sp Alpha P    NLDVGMKF-C DPLR-AYVVN --TIQDKLTP IAVD-LQYEL T-DESILLPY
Sp Alpha J    NISTGSEI-C LSHV-AYIPN --SIQDKLTP ITMR-ADYSL VIADQVNSTL
Lv ASU2       RLAVSTQF-C DPLR-AYVGN --SIEDKLTP IKVT-LQYDL NNDESRLQPH
Mm Alpha I    DLGGRDKPIC HTTG-AFLRD EADFRDKLSP IVLS-LNVSL P----PEET-
Mm Alpha V    TVFRGGQMQC EELV-AYLRD ESEFRDKLTP ITIF-MEYRL D-QRTAADAT
Mm Alpha 8    VMKQQKSLHC QDFM-VYLRD ETEFRDKLSP INIS-LNYSL D-DSTFKDSL
Mm Alpha 7    WLKHQHDRVC GDTV-FQLQE --NVKDKLRA IVVT-LSYGL RTPPLGRQAP
Mm Alpha 6    TLNRQKQRAC MEET-LWLQE --NIRDKLRP IPIT-ASVEI QEPSSRRRVN
Mm Alpha 5    LIQNGAREDC REMK-IYLRN ESEFRDKLSP IHIA-LNFSL D-PKAPMDSH
Mm Alpha 3    SMP---ETHC QTLE-LLLMD --NVRDKLRP IVIA-MNYSL PLRMPDRLKL
Dm PS1        RVHTNGRTEC QAVT-GYIKA --NTRDIQTP VRFR-LKYSL VEPPLADSAL
Dm PS2        RLNYGQKY-C LNET-VYLLD --KVQDKLTP LEVE-ARYNL RSSRPLDPMV
Ce INA-1      GSNSHVQKRT YRQL-IYLPV Q-ESKDWITP LKFR-FTVSI RNEKKPVQPP
Ce PAT-2      RLEKGRQK-C FRHR-FFASS --TMKDKLSP IHWS-VNYTY VESKTGKLRG
Nv Alpha 1    VLTKAGDWHC LPEYKIFLRL REEIGDLLTA LTFD-LQYDL K---KNAIC-
Ci A9         NVPRAGNYFR NSQY-IYSMK G-SIWDTLST SYLNVLIVEI NGVKRERPDQ
Ci A11        TLRAGDQQQC VHET-VFMKK --DVRDKQTE IAVT-VNYWL Q--ENHLPT-

                    210        220        230        240        250
                        *                     *       *   *
Sp Alpha C    GDP-LPSMTP FPILQENTQT IML--EEVIF S-KECAKDDG LCITDLDVR-
Sp Alpha F    --------EI LPILDESSDN FRA--LSVPI E-RNC-MNE- TCVPDLRVR-
Sp Alpha D    --------AV QPIMS-NAVN NQT--LNSLV FARDC-AGE- TCYPDLAVQ-
Sp Alpha P    --------EI LPIINKEAVS SQT--KQVSI Q-NNC-VNN- ICIPEIGIT-
Sp Alpha J    FR----PTDI QPILNKNIPA STS--LTVPI L-RNC-QNL- TCVPDLRLA-
Lv ASU2       --------EI LPIIDMATMS TQT--KQVSI Q-NNC-VNN- ICIPDLDVT-
Mm Alpha I    --------GG APAVVLHGET HVQ--EQTRI I-LDCGEDD- LCVPQLRLT-
Mm Alpha V    --------GL QPILNQFTPA NVS--RQAHI L-LDCGEDN- VCKPKLEVS-
Mm Alpha 8    --------EV KPILNHYRDN VVT--EQAHI L-VDCGEDN- LCVPDLKLS-
Mm Alpha 7    GQE---LPTV APILNAHQPS TQR--TEIHF LKQGCGEDK- ICQSNLQLER
Mm Alpha 6    --S---LPEV LPILNSNEAK TVQ--TDVHF LKEGCGDDN- VCNSNLKLE-
Mm Alpha 5    --------GL RPVLHYQSKS RIE--DKAQI L-LDCGEDN- ICVPDLQLD-
Mm Alpha 3    GLR---SLDA YPVLNQAQAM ENH--TEVHF Q-KECGPDN- KCDSNLQMR-
Dm PS1        -------VRL NPILDQTQAH VDF---EGTF Q-KDCGDDD- LCESNLIIRV
Dm PS2        RHR---RSIL EPVIDQNREI VLR--DAINI Q-KNCGPDN- ICEPDLKLK-
Ce INA-1      QGSQLVDLKH YPVLNKYGAS YEF---DVPF N-TLCGEDH- TCQTDLSLKA
Ce PAT-2      -------DKL EPAIDTTVPL SFQ--NKINI A-NNCGKDD- LCVPDLKVT-
Nv Alpha 1    --------EV CPILNDYDDA RQRTVRAEAI FQKHC-KDK- VCVPDLSVSG
Ci A9         PVFN---MIN DPIMDADFPH ART--SVVNL A-NNCGDDG- -CQSNLKLR-
Ci A11        ---------- EPVLDVLAGT SST--TRADI F-KDCGPDE- ICIPDLVVN-
```

```
                      260        270        280        290        300
                                                              *
Sp Alpha C    ----AS---- -------VDL EGTPPILKVG -EQQEISMSA HIAN------
Sp Alpha F    ----A----- ---------- STETASLLVG -SSESITIEV TIDN------
Sp Alpha D    ----T----- ---------- IVSTTELMIG -KAETFLLTV EVMN------
Sp Alpha P    ----V----- ---------- TPNLPNIVIG -QAQELTLVV SINN------
Sp Alpha J    ----V----- ---------- TKNVEKLFVG -DRKNIIMII TITN------
Lv ASU2       ----V----- ---------- TPNLPNIVIG -QTQELTLDV SLNN------
Mm Alpha I    ----A----- ---------- TAGDSPLLIG -ADNVLELKI EAAN------
Mm Alpha V    ----V----- ---------- NSDQKKIYIG -DDNPLTLTV KAQN------
Mm Alpha 8    ----A----- ---------- RPDKHQIIIG -DENHLMLII NARN------
Mm Alpha 7    YQFCSRISDT -EFQALPMDL DGRTALFALS -GQPFIGLEL TVTNLPSDPA
Mm Alpha 6    YKFGTREGNQ DKFSYLPIQ- KG-IPELVLK -DQKDIALEI TVTNSPSDPR
Mm Alpha 5    ----V----- ---------- YGEKKHVYLG -DKNALNLTF HAQN------
Mm Alpha 3    AAFLS----- -------EQL QPLSRLQYSR -DTKKLFLSI NVTNSPSS--
Dm PS1        EPNIT----- ---------- ESSGNEYTLI LDETELEVRI NVSN------
Dm PS2        ---------- ---------- VSTVDKYLFG -SPEPLVIEV FISN------
Ce INA-1      AFKDIP---- ---------L TSNGYVSNVG -EKDYLDLTF TVEN------
Ce PAT-2      ----A----- ---------- VADREKFLLG TQDNTMLINV TVQN------
Nv Alpha 1    ---SAS---- ---------F SGGFKDLRIG -ISQQMLVTI TVEN------
Ci A9         ---GSIP--- ---------- PEVVVGKLVG -RDSKMLLSL NVTN------
Ci A11        ----A----- ---------- KLSPDTVQVG -KYSEINVKA SVWN------

                      310        320        330        340        350
                      *
Sp Alpha C    --REED--AY NAKLVVTY-P AYLGFISLSS ASQA---FTA DCNPEPFEEG
Sp Alpha F    --NGED--AY LSTLTVEV-P QELQYAGFTR -VQT--DQIV TCN--SMVQS
Sp Alpha D    --SGED--AF LSVLDIVE-P PGLFFVNVLR SNTG---IIV SCS--VSTTT
Sp Alpha P    --RGED--AF QSTLAVYY-P EGLQYVRLER RANM--DFSV TCT--EDSAL
Sp Alpha J    --QGED--AY QSRLKIHT-P PELLFKKAEQ LRLIH-DFIA SCS--NDVQL
Lv ASU2       --RGED--AF QSSLSVYY-P LGLQFVRLER KANM--DFSV TCS--EDSDL
Mm Alpha I    --DGEG--AY EAELAVHL-P PGAHYMRALS NIEG--FERL VCTQKKENES
Mm Alpha V    --QGEG--AY EAELIVSI-P PQADFIGVVR NNEA--LARL SCAFKTENQT
Mm Alpha 8    --EGEG--AY EAELFVII-P EEADYVGIER NNKG--LRPL SCEYKMENVT
Mm Alpha 7    RPQADGDDAH EAQLLVTL-P ASLRYSGVRA LDSV--EKPL -CLS-NDSAS
Mm Alpha 6    NPRKDGDDAH EAKLIATF-P DTLTYSAYRE LRAFP-EKQL SCVA-NQNGS
Mm Alpha 5    --LGEGG-AY EAELRVTA-P LEAEYSGLVR HPGN--FSSL SCDYFAVNQS
Mm Alpha 3    --QRAGEDAH EALLTLEV-P SALLLSSVRP SGT------- -CQ----ANN
Dm PS1        --LADS--AY EAQLFIAH-Q AGVSYVATKK PTN------A TCN---SYNT
Dm PS2        --TNED--AF EAAFYMVT-P PDLQFRKLQQ LGEKK-DTPI TCSPPTPENN
Ce INA-1      --KKEK--AY QANFYLEYNE EELELPQVQG SKR------M IAE--TIGKN
Ce PAT-2      --GGED--SY ETKLYFDV-P QGFEYGGIES VGGDGSKSAP ACSPTSDEPD
Nv Alpha 1    --KAED-FAY PGKVLVTY-P SFLGYVGVAA KQD------V QCKRLITSDD
Ci A9         --DGEE--AH QAVVSAKL-P AWVYYDNYTV TRSS--GAEI VCTTSDIGGS
Ci A11        --NGEN--AY LTTMVVNY-P QYVTFIGLQE DKVD--GRVI SCV----DLN

                      360        370        380        390        400
                      *   ***                                    ?
Sp Alpha C    AN-ETSIICE LGNPYSEGSV DSFVVLYDAS S--VPPDAEV FNITLVASST
Sp Alpha F    SE----ISCD VGNPFTAGTT VILKMQFSST N--LPGNGDN VHFTFGVSSV
Sp Alpha D    RI----RSCN IGNPVPAGDQ ITVGLQYQTA T--YSTFTGP AVLNISTRSI
Sp Alpha P    RM----ITCD TGNPLVGKYN LEFGLTLSTL Q--VSGDKDN IEFYLVAGSE
Sp Alpha J    SL----VTCT VGNPLPAGAE LKLRVEFENV N--LSGQVDL LTFSMQVQSV
Lv ASU2       RI----ITCD TGNPMVGKNI LEFGLTLSTF Q--VSGDKDS IEFYFKAESE
Mm Alpha I    RV----ALCE LGNPMKKDTR IGITMLVSVE N--LEEAGES VSFQLQVRSK
Mm Alpha V    RQ----VVCD LGNPMKAGTQ LLAGLRFSVH Q--QSEMDTS VKFDLKIQSS
Mm Alpha 8    RM----VVCD LGNPMVTGTN FSLGLRFAVP R--LEKTNMS INFDLQIRSS
```

```
Mm Alpha 7   H-----VECE LGNPMKRGAQ VTFYLILSTS G--ITIETTE LEVKLLLATI
Mm Alpha 6   Q-----ADCE LGNPFKRNSS VTFYLILSTT E--VTFDTTD LDINLKLETT
Mm Alpha 5   RQ----LVCD LGNPMKAGTS LWGGLRFTVP H--LQDTKKT IQFDFQILSK
Mm Alpha 3   ET----ILCE LGNPFKRNQR MELLIAFEVI G--VTLHTRD LPVLLQLSTS
Dm PS1       TL----VACS LGNPMLRDTT TFVTIRFQPK G--LEPSEKI MLFHIFANTT
Dm PS2       HT----LKCD IGNPLESGKI AHFKISLVPE E--KYGSSSS YDFYWEANST
Ce INA-1     I-----VHLP LGNPMNGASK HQFTIQFKLT RGRTEGIGKA LKFMAHVNST
Ce PAT-2     SDGKWTFACD LGNPLPANKV VSSVVRVTAS S--DKPPLAP ISINAHVNSS
Nv Alpha 1   FEQ---AVCL VGMPFQGKTK KTFDMLFATA G--VKGNISM FKVDLEATSP
Ci A9        TF----VVCT LGNPYVEGSA DTIDILLDVQ Q--LTADTKQ IVIEVTGSTT
Ci A11       PF----LLCE VANPLKVDTR IDLVIQFGVN E--LQGDVDV LPLLLYANCT

                    410        420        430        440        450
                                                     *
Sp Alpha C   NDAD-SNPGN NQYIITIEVE SITDILISGK G-PEQLYFSG Q-----VIGE
Sp Alpha F   ESE--QMIAD NEFNITVPLA IAADLSFHGN VIPETVIYSE E---NYLLTR
Sp Alpha D   DAERAGRDVD NEAMVSIPVY ATSSLALSGN SIPDTLVITE D---------
Sp Alpha P   NNEDPNTLDN NELNVTVAVI VDATLKLLSA SYPEIVTYRV PE--DNIVPE
Sp Alpha J   EAERESTRRD NLMDVLIPVE IVASMELIGV SSPQEVVFRP E---HYPLDR
Lv ASU2      NSEDPNTLEN NELNMTVPVT VDCTLKLLSA SYPEIVMYST QE--DYVVPP
Mm Alpha I   NS---QNPNS KVVMLPVAIQ AEATVELRGN SFPASLVVAA E-----EGDR
Mm Alpha V   NSF--DNV-S PVVSYKVDLA EKAAVEIRGV SSPDHIFLPI P---NWEYKE
Mm Alpha 8   NK---DNPDS NFERVQINIT AIAQVEIRGV SHPPQIVLPI H---NWEPEK
Mm Alpha 7   S----EQE-L DPVSVRAHVF IELPLSISGV ATPQQLFFSG E-----VKGE
Mm Alpha 6   S----NQDNL APITAKAKVV IELLLSVSGV AKPSQVYFGG T-----VVGE
Mm Alpha 5   NL---NNSQS NVVSFPLSVE AQAQVSLNGV SKPEAVIFPV S---DWNPQD
Mm Alpha 3   S----HQDNL QPVLLTLQVD YTLQASLSLM NHRLQSFFGG T-----VMGE
Dm PS1       SKL--VGPER PERDLRVNVV RRAKLNFRGW AIPEQSFYSG SSVANSVANT
Dm PS2       NLEKPGSEYD NKIRQSVGIW VDTDLDIKGT SLPDYQLYKA D---DYKELE
Ce INA-1     SQETEEELKD NKWEAEVQII KKAELEIYGI SDPDRVFFGG K-----ARAE
Ce PAT-2     NDEEAHTVAD NKVTFTIPVD FKNQLSLNGR SNPEQVDFSM T----NKTRV
Nv Alpha 1   GAD--ANPSD NKVSLQLAVK FEADLSISGA SLPDELVFSS K----GKTSE
Ci A9        S----INPEI PPTPLYSDVI IQMMLSLSGY GKPQQVRYVN A----PIKGE
Ci A11       NE---QNNQS PLFHTMIHVE VVANVKFYNV STPSLIRLDK E------SEN

                    460        470        480        490        500
                             *          * *                   *
Sp Alpha C   SAMN------ -YFEDIGLAV NQRWTIFNEG P----GAVNT ARVTIDFPYE
Sp Alpha F   -AIK------ -LEKDVGPDI THVFSLYNNG P----SSVAR TEITIPWPMR
Sp Alpha D   -SNA------ --TDDKGPLM THVIRLQNNG P----SFIGP STIEILWPIR
Sp Alpha P   FPTKNA---- -SEADIGMEV VHLYEVRNTG S----SNAAE VTLNIRWPEK
Sp Alpha J   TINS------ --YDDVGPEI SHCFSLINKG P----SHIGP TIVHIEWPLF
Lv ASU2      FPAKNA---- -SEADIGMEV MHLYEVRNTG S----SNAGE VSLNIQWPQK
Mm Alpha I   EQED------ --LDSWVSRL EHTYELHNIG P----GTVNG LRLLIHIPGQ
Mm Alpha V   NPET------ --EEDVGPIV QHIYELRNNG P----SSFSK AILNLQWPYK
Mm Alpha 8   KPHK------ --EEEVGPLV EHIYELHNIG P----STISD SILDVGWPFS
Mm Alpha 7   SAMR------ -SERDVGSKV KYEVTVSNQG QSL--NTLGS AFLNIMWPHE
Mm Alpha 6   QAMK------ -SEDEVGSLI EYEFRVINLG KPL--KNLGT ATLNIQWPKE
Mm Alpha 5   QPQK------ --EEDLGPAV HHVYELINQG P----SSISQ GVLELSCPQA
Mm Alpha 3   AAMK------ -TAEDVGSPL KYEFQVSPVG DGL--AALGT LVLGLEWPYE
Dm PS1       AATDIEGHGP MGMDDVGSQV HHMFTIFNEG P----STAPK VQMVIHWPYS
Dm PS2       NATK------ --EDDIGPQV VHIYEIRNNR P----SIIEE AEVFIHLPYE
Ce INA-1     SELE------ -LEEDIGTMV RHNYTIINHG P----WTVRN VEAHISWPYQ
Ce PAT-2     DAFD------ --DNEIGPVV SHLYQISNRG P----SEVDS ATLDIFWPSF
Nv Alpha 1   EVRT------ --ERDIGPLV QQTIMVRNNG P----GAVDG SEVTVSLPWK
Ci A9        SAIH------ -YTNEIGPFV EYTFTVKNDG KYHXLQLANG VRLIVDFPIE
Ci A11       AALN------ ---QSASKPL THTYEITNAG P----AVISK AEISLLWPLS
```

```
                      510        520        530        540        550
                            *                     *
Sp Alpha C    VANG------ -KWLLYMTEM PFVED----- --NKGSCNVT P-AVYVNELG
Sp Alpha F    YNGVQR---- -NYLLYLLSA TMDTG----- ----EECTID G---PWNPEG
Sp Alpha D    LQDG------ -KKLMNVTEV TMDTG----- ----MPCRLT P---RINRDS
Sp Alpha P    DENG------ -DYLFYLLGI MTDEG----- ----VTCQIS QG--QANPLG
Sp Alpha J    TNNG------ -SLLLLLTRA SMNNG----- ----QMCTVN A---YVNPAK
Lv ASU2       NEDG------ -EYLFYLLGI MTEEG----- ----VTCQLT QG--KANPEG
Mm Alpha I    SQP------- -SDLLYILDV QPQGG----- ----LLCSTQ P---SPK-VD
Mm Alpha V    YNN------- -NTLLYILHY DIDGP----- ----MNCTAD T---EINPLR
Mm Alpha 8    ARD------- -EFLLYIFHL QTLGP----- ----LQCQTN P---EINPQD
Mm Alpha 7    IANG------ -KWLLYPMRV ELEGGQGPG- --KRGICSPR P---NI--LQ
Mm Alpha 6    ISNG------ -KWLLYLMKV ESKG----L- --EQIVCEPH N---EINYLK
Mm Alpha 5    LEG------- -QQLLYVT-- KVTGL----- ----SNCTSS Y---TPNSQG
Mm Alpha 3    VTNG------ -KWLLYPTEI TIHSN----- --GSWPCQPS GN--LVNPLN
Dm PS1        LYSDPQSGRP VQYLLYLEQV PTVEVS---- ---QGECHVA K--EYVNPLN
Dm PS2        TIVG------ -DPLMYLLNQ PETGG----- ---KIQCDDV A----FNEYN
Ce INA-1      LRSRFGRG-- -KNALYLLDV PTITTEFTDG TSEVRKCFIK QQYEYVNPAE
Ce PAT-2      STEG------ -GHLLYIITE PVVNP---P- --NKGRCRVK QL-QNVNPLN
Nv Alpha 1    KSEEDL---- -SYLLYLLSV QVHGS----- ---AATCDVK T-----NPLN
Ci A9         INNG------ -KWLMYLVSA HVRSGG---- LQIVGECEQQ Y----QNTLH
Ci A11        VNGDSK---- -DLLLPLLEV QHSGP----- ----VICHYS H---------

                      560        570        580        590        600
Sp Alpha C    LKPKNG---- --GGAYNPVA PGTGGTTQSR KRRQAEAEAV TSTRELTAGK
Sp Alpha F    LG-------- ---SVDNSTG TPE------- -SKRRRRQAA ----------
Sp Alpha D    PK-------- ----QNNVTT FVD------- --DVSNSTIY TLWA------
Sp Alpha P    VKL------- ---EASTKEQ LSN---STTQ VSGRRKREGE YAEA--LAQA
Sp Alpha J    LP-------- ----MESGVM YED------- -TGTRPDRSY CKDTNTAMTS
Lv ASU2       VKL------- ---EPSTKAK LSN---STTQ VSGRKRREPE VAEA--LAQT
Mm Alpha I    WKLS------ ---TPSPS-S IRP------V HHQRERRQAF LQGPKPGQQD
Mm Alpha V    IK-------- ---TPEKN-D TGA-AGQGER SHLITKRG-- -LTL--REGD
Mm Alpha 8    IKPAA----- ---SPEDTPE LSAFLRNATI PHLVRKRDVP VVQLH-RQSP
Mm Alpha 7    LDVDS----- ---RDRRRRE LGQ-PEPQEP PEKVEPSTSW WPVSS-AEKR
Mm Alpha 6    LKESH----- ---NSRKKRE L---PEKQ-- ---IDDSRKF SLFP---ERK
Mm Alpha 5    LEL------- ---DPETS-- ---------- PHHLQKREAP GRSS--TASG
Mm Alpha 3    LTLS------ ---DPGVTPL SPQRRRRQLD PGGDQSSPPV TLAAAKKAKS
Dm PS1        LASGSRENPA YLSAPAQMRM FPSQSRHSFN KSLIHSQRSY YSSSHRDDHS
Dm PS2        LL-------- ----LDEKLV KKS------- --YLQAQGAI WNSA---QVS
Ce INA-1      IKLN------ ---TKYSTQE TAPHRVEHRM KREIDEDEEE QSDDLGAVEE
Ce PAT-2      LR-------- ---ITNEHVP TEP------- --PVAKTPNE YSRE---EDD
Nv Alpha 1    IKMSN----- ---NTDKPAV LES------- TSRVGANKRT RKEV----KK
Ci A9         FKLP------ --VTASVSAA RRT---KREV TSSLYEPRRL SSTQ--SGSN
Ci A11        ---------- ----IADA-- ---------- ---ICNKNTF SSFS------

                      610        620        630        640        650
Sp Alpha C    DITLDCKTGT AQCLTIIC-- ---------- ---------- ----------
Sp Alpha F    ---------- ---------- ---------- ---------- ----------
Sp Alpha D    ----DCDSLP -ECVAIR--- ---------- ---------- ----------
Sp Alpha P    EP-IFCTPE- -SCVLIN--- ---------- ---------- ----------
Sp Alpha J    HHMQCCSLG- -NCATITC-- ---------- ---------- ----------
Lv ASU2       DNVIYCASD- -SCVLINC-- ---------- ---------- ----------
Mm Alpha I    PVLVSCDGS- ASCTVVEC-- ---------- ---------- ----------
Mm Alpha V    VHTLGCGIA- -KCLQITC-- ---------- ---------- ----------
Mm Alpha 8    ARILNCTNI- -DCLQISC-- ---------- ---------- ----------
```

```
Mm Alpha 7   NVTLDCAQGT AKCVVFSC-- ---------- ---------- ----------
Mm Alpha 6   YQTLNCSVN- VRCVNIRC-- ---------- ---------- ----------
Mm Alpha 5   TQVLKCPEA- -KCFRLRC-- ---------- ---------- ----------
Mm Alpha 3   ETVLTCSNGR ARCVWLEC-- ---------- ---------- ----------
Dm PS1       DDTQSNRNRV RRSFLERVTR LERLMYDPES SNAANGKKQD IVELDCNKGA
Dm PS2       GQSSSSSSSG GASVHIEK-- ---------- ---------- ----------
Ce INA-1     NIPWFSTANF WNLFAIKGGD GRPREVKHLS CQDNTANCFT VIC-------
Ce PAT-2     ESYEDETTT- -QSQSTRH-- ---------- ---------- ----------
Nv Alpha 1   GDQLSCDNV- -KCLLFTC-- ---------- ---------- ----------
Ci A9        FVTLDCHGDT ARCVKIGC-- ---------- ---------- ----------
Ci A11       -SLQNCHTDP QNCYEMTC-- ---------- ---------- ----------

                      660        670        680
Sp Alpha C   --------.. .......... .......... .
Sp Alpha F   --------.. .......... .......... .
Sp Alpha D   --------.. .......... .......... .
Sp Alpha P   --------.. .......... .......... .
Sp Alpha J   --------.. .......... .......... .
Lv ASU2      --------.. .......... .......... .
Mm Alpha I   --------.. .......... .......... .
Mm Alpha V   --------.. .......... .......... .
Mm Alpha 8   --------.. .......... .......... .
Mm Alpha 7   --------.. .......... .......... .
Mm Alpha 6   --------.. .......... .......... .
Mm Alpha 5   --------.. .......... .......... .
Mm Alpha 3   --------.. .......... .......... .
Dm PS1       TNCVRIEC.. .......... .......... .
Dm PS2       --------.. .......... .......... .
Ce INA-1     --------.. .......... .......... .
Ce PAT-2     --------.. .......... .......... .
Nv Alpha 1   --------.. .......... .......... .
Ci A9        --------.. .......... .......... .
Ci A11       --------.. .......... .......... .
```

>*S. purpuratus* integrin_α2 domains

```
                     10         20         30         40         50
                                                       ****
Sp Alpha C   ---------- ---------- ---------- ---------- GRPVINIEAQ
Sp Alpha F   FGLSLNTFGS SLSAGMDMDN NMYPDLVVGA YKSSTVVLVR SRPVVHVTGS
Sp Alpha D   ---------- ---------- ---------- ---------- SRAVVWTVAE
Sp Alpha P   ---------- ---------- ---------- ---------- TRPVVLLEAE
Sp Alpha J   ---------- ---------- ---------- ---------- SRPVVDVSAH

                     60         70         80         90        100
                     *          *   *  *     *  * * *      *  *
Sp Alpha C   LTLSVTELDP NTTNCEYS-G KDSLCFTVQL CMSYYCLAPA FNDPITINFD
Sp Alpha F   LVSKTESIDL KNTNHNTASG TAVSSFDFEV CIRYSGTN-- VPPNLDFDYS
Sp Alpha D   IYPEENAIDL DVKDETTSRG VLVTGFNVVV CIHYMGQG-- LPANLDFSYE
Sp Alpha P   LTIEPIGINL DNKTYELPDG TMVTSFVAMA CFTYTGNH-- LPARIGISYT
Sp Alpha J   VSVDLKGIRL KERNYRLPGG ALVLGFNVTL CFSFDGIN-- IPQEIFLDYD
```

```
                    110        120        130        140        150
                    *    *                          *     *
Sp Alpha C   IEAEGVRRSK VLNSRVVFED SGIYSLTKQT LVITNDDEAC TPEYNVILKA
Sp Alpha F   LSLDYNRQS- TRRANFDLGG N-DASSVSLT VTLTSNVEYC NPFTAYAKPT
Sp Alpha D   IVLDSLRQV- INRRAAFITN N--IATLRHQ ISAPIEQRTC MTHAAYIQPT
Sp Alpha P   LTVDSSITS- GRRALLEVNE L---SQVTKN RNLDVGMKFC DPLRAYVVNT
Sp Alpha J   LIADAKLPV- SQRCYFELLD QPDTAGIRDR VNISTGSEIC LSHVAYIPNS

                    160        170        180        190        200
                 * *   **    *                        **
Sp Alpha C   GFTDIFRPIP INLTYYIPEV EAVMPQPGDP LPSMTPFPIL QENTQTIMLE
Sp Alpha F   -IVDKLSPIP VSLSFGLPDS SPA------- --PGEILPIL DESSDNFRAL
Sp Alpha D   -IRDKQTPMS LRLIHSVAQT SQS------- --DSAVQPIM SNAVNNQTLN
Sp Alpha P   -IQDKLTPIA VDLQYELTDE SIL------- -LPYEILPII NKEAVSSQTK
Sp Alpha J   -IQDKLTPIT MRADYSLVIA DQVNSTL--- FRPTDIQPIL NKNIPASTSL

                    210        220        230        240        250
                 * *   *       * ***                * **          * *
Sp Alpha C   EVIFSKECAK DDGLCITDLD VRASVDLEGT PPILKVGEQQ EISMSAHIAN
Sp Alpha F   SVPIERNCMN --ETCVPDLR VRASTETA-- --SLLVGSSE SITIEVTIDN
Sp Alpha D   SLVFARDCAG --ETCYPDLA VQTIVSTT-- --ELMIGKAE TFLLTVEVMN
Sp Alpha P   QVSIQNNCVN --NICIPEIG ITVTPNLP-- --NIVIGQAQ ELTLVVSINN
Sp Alpha J   TVPILRNCQN --LTCVPDLR LAVTKNVE-- --KLFVGDRK NIIMIITITN

                    260        270        280        290        300
                 ***** * *     *  *         *         *
Sp Alpha C   REEDAYNAKL VVTYPAYLGF --ISLSSASQ AFTADCNPEP FEEGANETSI
Sp Alpha F   NGEDAYLSTL TVEVPQELQY --AGFTRVQT DQIVTCN--- --SMVQSSEI
Sp Alpha D   SGEDAFLSVL DIVEPPGLFF --VNVLRSNT GIIVSCS--- --VSTTTRIR
Sp Alpha P   RGEDAFQSTL AVYYPEGLQY -VRLERRANM DFSVTCT--- --EDSALRMI
Sp Alpha J   QGEDAYQSRL KIHTPPELLF KKAEQLRLIH DFIASCS--- --NDVQLSLV

                    310        320        330        340        350
                 *  ***                     *             *    *
Sp Alpha C   ICELGNPYSE GSVDSFVVLY DASSVPPDAE VFNITLVASS TNDAD-SNPG
Sp Alpha F   SCDVGNPFTA GTTVILKMQF SSTNLPGNGD NVHFTFGVSS VESEQ--MIA
Sp Alpha D   SCNIGNPVPA GDQITVGLQY QTATYSTFTG PAVLNISTRS IDAERAGRDV
Sp Alpha P   TCDTGNPLVG KYNLEFGLTL STLQVSGDKD NIEFYLVAGS ENNEDPNTLD
Sp Alpha J   TCTVGNPLPA GAELKLRVEF ENVNLSGQVD LLTFSMQVQS VEAERESTRR



                    360        370        380        390        400
                 **   * * *          *    *  *                     *
Sp Alpha C   NNQYIITIEV ESITDILISG KG-PEQLYFS GQVIG----- ESAMNYFEDI
Sp Alpha F   DNEFNITVPL AIAADLSFHG NVIPETVIYS EENY----LL TRAIKLEKDV
Sp Alpha D   DNEAMVSIPV YATSSLALSG NSIPDTLVIT ED-------- ---SNATDDK
Sp Alpha P   NNELNVTVAV IVDATLKLLS ASYPEIVTYR VPEDNIVPEF PTKNASEADI
Sp Alpha J   DNLMDVLIPV EIVASMELIG VSSPQEVVFR PEHY----PL DRTINSYDDV
```

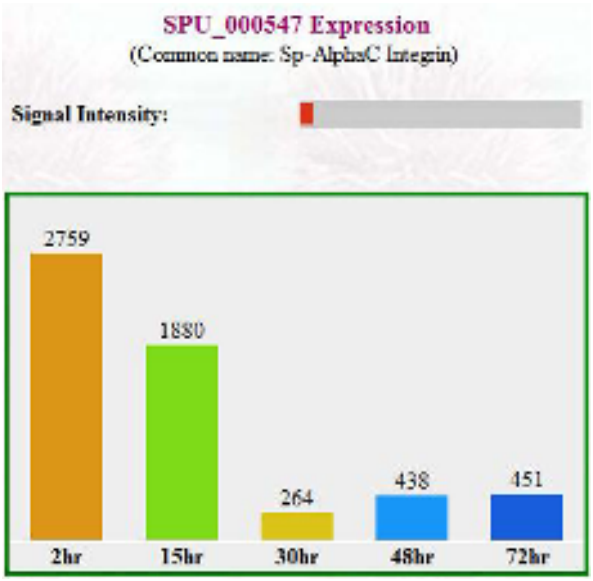
```
                  410        420        430        440        450
              *    *      * ***        * **          *          *
Sp Alpha C    GLAVNQRWTI FNEGPGAVNT ARVTIDFPYE VA--NGKWLL YMTEMPFVED
Sp Alpha F    GPDITHVFSL YNNGPSSVAR TEITIPWPMR YNGVQRNYLL YLLSATMDTG
Sp Alpha D    GPLMTHVIRL QNNGPSFIGP STIEILWPIR LQ--DGKKLM NVTEVTMDTG
Sp Alpha P    GMEVVHLYEV RNTGSSNAAE VTLNIRWPEK DE--NGDYLF YLLGIMTDEG
Sp Alpha J    GPEISHCFSL INKGPSHIGP TIVHIEWPLF TN--NGSLLL LLTRASMNNG

                  460        470        480        490        500
              *        *
Sp Alpha C    NKGSCNVTPA VYVNELGLKP KNGGGAYNPV APGTGGTTQS RKRRQAEAEA
Sp Alpha F    E--ECTID-- GPWNPEGL-- --------GS VDNSTGTPES KRRRRQAA--
Sp Alpha D    M--PCRLT-- PRIN------ ---------- RDSPKQNNVT TFVDDVSNST
Sp Alpha P    V--TCQISQ- GQANPLGVKL EAS--TKEQL SNSTTQVSGR RKREGEYAEA
Sp Alpha J    Q--MCTVN-- AYVNPAKLPM ES------GV MYEDTGTRPD RSYCKDTNTA



                  510        520
                                *  *
Sp Alpha C    VTSTRELTAG KDITLDCKTG TAQCLTIIC
Sp Alpha F    ---------- ---------- ---------
Sp Alpha D    IYTLWADCD- ---------S LPECVAIR-
Sp Alpha P    LAQAEPIFC- ---------T PESCVLIN-
Sp Alpha J    MTSHHMQCC- ---------S LGNCATITC
```

>*S. purpuratus* integrin_α2 domains - CUT

```
                  10         20         30         40         50
               * **          *  *   *    *  **     * **   *
Sp Alpha C    RPVINTTNCE YSGKSLCFTV QLCMSYNVIL KAGFTDIFRP IPINLTYYIP
Sp Alpha F    SRPVVKNTNH NTASGTAVSS FDFEVCFTAY AKPTIVDKLS PIPVSLSFGL
Sp Alpha D    SRAVVDVKDE TTSRGVLVTG FNVVVCHAAY IQPTIRDKQT PMSLRLIHSV
Sp Alpha P    TRPVVDNKTY ELPDGTMVTS FVAMACLRAY VVNTIQDKLT PIAVDLQYEL
Sp Alpha J    SRPVVKERNY RLPGGALVLG FNVTLCHVAY IPNSIQDKLT PITMRADYSL

                  60         70         80         90         100
                    **          * *              *
Sp Alpha C    EVEAVMPPFP ILQENTECAK LCITDLDVRA SVDLELKVGE QQEISMSAHI
Sp Alpha F    PDSSPAPGIL PILDESSNCM NTCVPDLRVR ASTETALLVG SSESITIEVT
Sp Alpha D    AQTSQSDSVQ PIMSNAVDCA GTCYPDLAVQ TIVSTTLMIG KAETFLLTVE
Sp Alpha P    TDESILLPIL PIINKEANCV NICIPEIGIT VTPNLPIVIG QAQELTLVVS
Sp Alpha J    VIADQVNSIQ PILNKNINCQ NTCVPDLRLA VTKNVELFVG DRKNIIMIIT

                  110        120        130        140        150
              ****          *   *     *         *      *       *
Sp Alpha C    ANREEDAYNA KLVVTYPAYL GFISSSASQA FTADCNGANE TSIICELGNP
Sp Alpha F    IDNNGEDAYL STLTVEVPQE LQYAGTRVQT DQIVTCNMVQ SSEISCDVGN
Sp Alpha D    VMNSGEDAFL SVLDIVEPPG LFFVNLRSNT GIIVSCSSTT TRIRSCNIGN
Sp Alpha P    INNRGEDAFQ STLAVYYPEG LQYVRRRANM DFSVTCTDSA LRMITCDTGN
Sp Alpha J    ITNQGEDAYQ SRLKIHTPPE LLFKKLRLIH DFIASCSDVQ LSLVTCTVGN
```

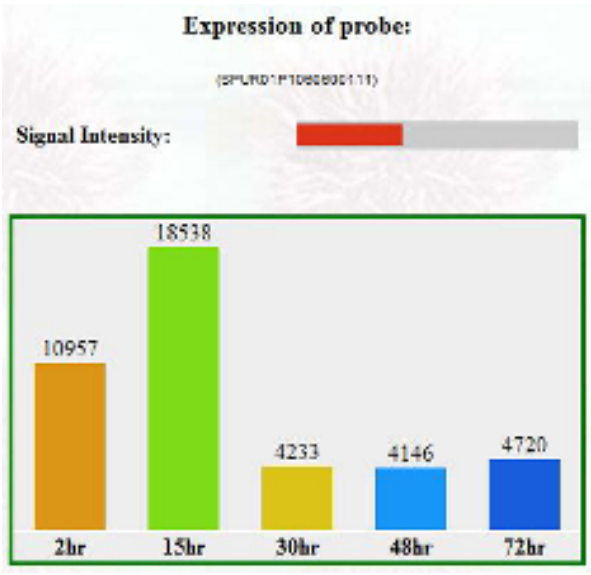
```
                    160        170        180        190        200
                  *  *        *  ***         * *             * *
Sp Alpha C    YSEGITLVAS STNDNNQYII TIFEDIGLAV NQRWTIFNEG PGAVNTARVT
Sp Alpha F    PFTAGFTFGV SSESEDNEFN ITVEKDVGPD ITHVFSLYNN GPSSVARTEI
Sp Alpha D    PVPAGLNIST RSDAEDNEAM VSITDDKGPL MTHVIRLQNN GPSFIGPSTI
Sp Alpha P    PLVGKFYLVA GSNNENNELN VTVEADIGME VVHLYEVRNT GSSNAAEVTL
Sp Alpha J    PLPAGFSMQV QSEAEDNLMD VLIYDDVGPE ISHCFSLINK GPSHIGPTIV


                    210        220        230        240        250
                   * **        *
Sp Alpha C    IDFPYEVAWL LYMTEMPFVE DNKGSCNV.. .......... ..........
Sp Alpha F    TIPWPMRYNR NYLLYLLSAT MDTGEECTI. .......... ..........
Sp Alpha D    EILWPIRLQG KKLMNVTEVT MDTGMPCRL. .......... ..........
Sp Alpha P    NIRWPEKDEG DYLFYLLGIM TDEGVTCQI. .......... ..........
Sp Alpha J    HIEWPLFTNG SLLLLLTRAS MNNGQMCTV. .......... ..........
```

Appendix III
Temporal Expression Patterns

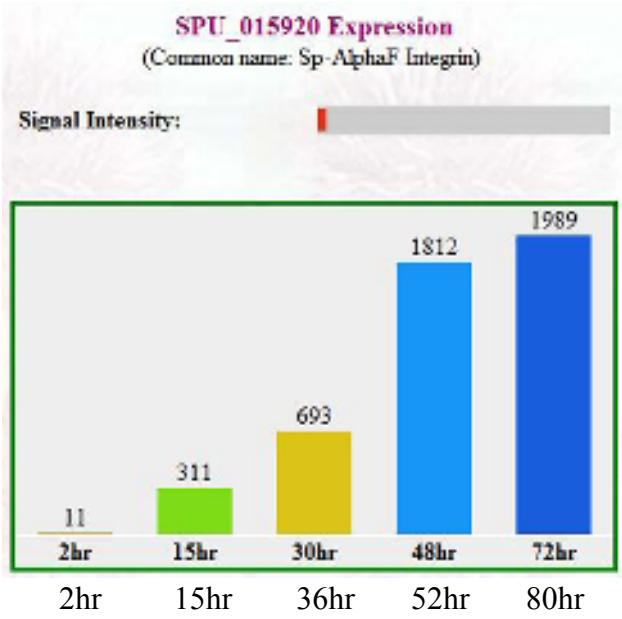


2hr 15hr 36hr 52hr 80hr

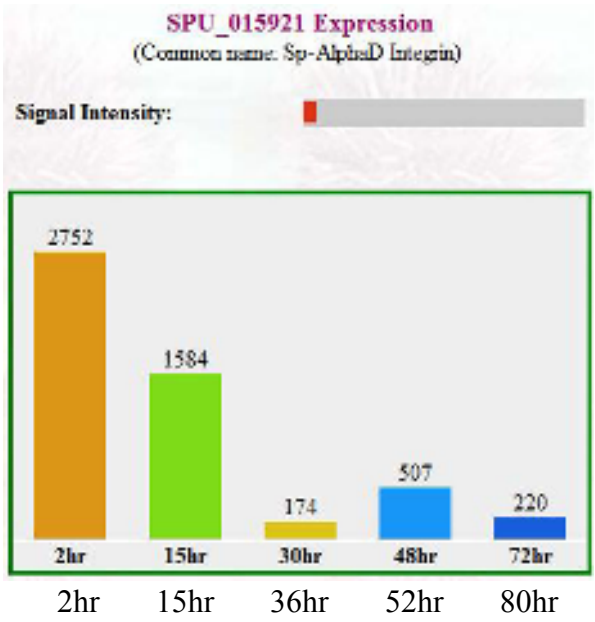
Temporal expression patterns of αC from chip hybridization data from the *S. purpuratus* Genome Search (http://urchin.nidcr.nih.gov/blast/index.html). Five chip-bound hybridization probes based on GLEAN3_00547 were hybridized with fluorescently labelled mRNA and the level of fluorescence was measured. Ubiquitin was used as a control to standardize measurements. The numbers below the graph represent development timing that would be similar to the embryos used in this study as the experiments were conducted at slightly different temperatures. The values above each bar represent the signal intensity, which was normalized to the amount of signal at 48 hours. The overall signal for this gene was only ~10% of the possible signal (red bar). αC is expressed at significantly higher levels during early development (Wei et al., 2006).



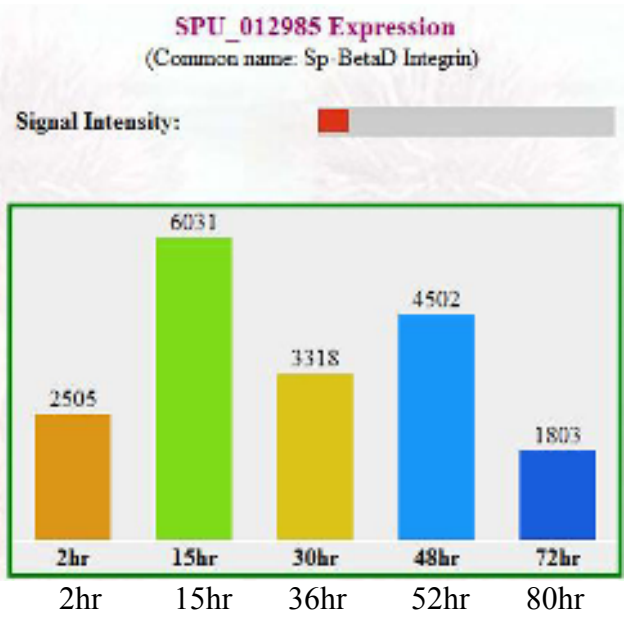
Temporal expression patterns of αC from chip hybridization data from the *S. purpuratus* Genome Search (http://urchin.nidcr.nih.gov/blast/index.html). This was one probe of five designed for GLEAN3_00547 that was hybridized with fluorescently labelled mRNA and the average level of fluorescence was measured. Ubiquitin was used as a control to standardize measurements. The numbers below the graph represent development timing that would be similar to the embryos used in this study as the experiments were conducted at slightly different temperatures. The values above each bar represent the signal intensity, which was normalized to the amount of signal at 48 hours. The overall signal for this gene was almost 50% of the possible signal (red bar). αC is expressed at significantly higher levels during early development (Wei et al., 2006).



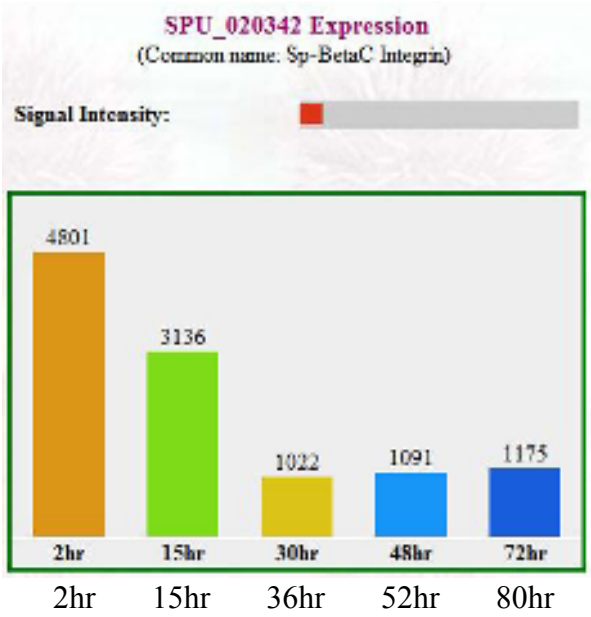
Temporal expression patterns of αF from chip hybridization data from the *S. purpuratus* Genome Search (http://urchin.nidcr.nih.gov/blast/index.html). Five chip-bound hybridization probes based on GLEAN3_00547 were hybridized with fluorescently labelled mRNA and the average level of fluorescence was measured. Ubiquitin was used as a control to standardize measurements. The numbers below the graph represent development timing that would be similar to the embryos used in this study as the experiments were conducted at slightly different temperatures. The values above each bar represent the signal intensity, which was normalized to the amount of signal at 48 hours. The overall signal for this gene was only ~5% of the possible signal (red bar). Expression of αF is upregulated as development continues (Wei et al., 2006).



Temporal expression patterns of αD from chip hybridization data from the *S. purpuratus* Genome Search (http://urchin.nidcr.nih.gov/blast/index.html). Five chip-bound hybridization probes based on GLEAN3_00547 were hybridized with fluorescently labelled mRNA and the average level of fluorescence was measured. Ubiquitin was used as a control to standardize measurements. The numbers below the graph represent development timing that would be similar to the embryos used in this study as the experiments were conducted at slightly different temperatures. The values above each bar represent the signal intensity, which was normalized to the amount of signal at 48 hours. The overall signal for this gene was only ~10% of the possible signal (red bar). αD is expressed at significantly higher levels during early development (Wei et al., 2006).



Temporal expression patterns of βD from chip hybridization data from the *S. purpuratus* Genome Search (http://urchin.nidcr.nih.gov/blast/index.html). Five chip-bound hybridization probes based on GLEAN3_00547 were hybridized with fluorescently labelled mRNA and the average level of fluorescence was measured. Ubiquitin was used as a control to standardize measurements. The numbers below the graph represent development timing that would be similar to the embryos used in this study as the experiments were conducted at slightly different temperatures. The values above each bar represent the signal intensity, which was normalized to the amount of signal at 48 hours. The overall signal for this gene was only ~10% of the possible signal (red bar). βD expression peaks at 15 and 48 hours of development (Wei et al., 2006).



Temporal expression patterns of βC from chip hybridization data from the *S. purpuratus* Genome Search (http://urchin.nidcr.nih.gov/blast/index.html). Five chip-bound hybridization probes based on GLEAN3_00547 were hybridized with fluorescently labelled mRNA and the average level of fluorescence was measured. Ubiquitin was used as a control to standardize measurements. The numbers below the graph represent development timing that would be similar to the embryos used in this study as the experiments were conducted at slightly different temperatures. The values above each bar represent the signal intensity, which was normalized to the amount of signal at 48 hours. The overall signal for this gene was only ~10% of the possible signal (red bar). βC is expressed at significantly higher levels during early development (Wei et al., 2006).