

# **The Generation and Characterisation of Human Insulin-like Growth Factor Mutants**

**Steven Thomas Henderson B.Sc. (Hons.)**

**Thesis submitted for the degree of Doctor of Philosophy**

School of Biological Sciences,  
Faculty of Science and Engineering,  
The Flinders University of South Australia

**September 2011**



# Table of Contents

Table of Contents.....	i
Abstract.....	iv
Declaration.....	vi
Acknowledgments .....	vii
Abbreviations.....	viii
Chapter 1 Introduction .....	1
1.1 Thesis overview .....	1
1.2 The Insulin-like Growth Factor axis.....	2
1.2.1 IGF evolution.....	2
1.2.2 Gene structure, expression and physiological effects of the IGFs .....	5
1.2.3 Regulation of the IGFs.....	8
1.2.4 Structure of IGF-I and IGF-II .....	13
1.3 The function of the Insulin-like Growth Factor I Receptor .....	15
1.4 The structure of the Insulin-like Growth Factor I Receptor.....	15
1.5 Insulin receptor isoforms and their functions .....	16
1.5.1 Distribution and functions of the IR-A.....	17
1.5.2 Distribution and functions of the IR-B.....	17
1.6 Hybrid receptors and their functions.....	18
1.6.1 IR-A/IR-B hybrid receptors.....	18
1.6.2 IR/IGFR-IR hybrid receptors .....	18
1.7 The structure of the Insulin Receptor .....	19
1.8 Characteristics and mechanism of ligand binding to the IR and IGF-IR.....	21
1.8.1 IGF receptor binding sites are primarily located in the B and A-domains.....	25
1.8.2 Receptor binding specificity is determined by the IGF C and D-domains .....	25
1.8.3 IR and IGF-IR signal transduction.....	27
1.9 The IGF system and cancer .....	30
1.9.1 IGF ligands and the IGF-IR in cancer .....	31
1.9.2 IGF-II and the IR-A in cancer .....	31
1.9.3 Targeting the IGF system for use as potential cancer therapeutics.....	32
1.10 Directed evolution .....	37
1.10.1 Directed evolution requires protein function to be physically achievable and obtainable through evolution.....	38
1.10.2 An overview of library construction for directed evolution .....	43
1.10.3 An overview of screening and selection for directed evolution.....	44
1.10.4 Directed evolution strategies .....	46
1.10.5 Directed evolution in <i>Neurospora</i> .....	56
1.11 Aims .....	65
Chapter 2 Materials and methods.....	66
2.1 Materials .....	66
2.1.1 Chemicals, reagents & enzymes .....	66
2.1.2 <i>Neurospora crassa</i> stocks .....	66
2.1.3 Bacterial stocks .....	67
2.1.4 Plasmid stocks.....	68
2.1.5 Mammalian cell lines .....	69
2.1.6 Antibodies .....	69

2.1.7	Protein standards.....	70
2.1.8	Oligonucleotides .....	70
2.1.9	Buffers & solutions.....	74
2.1.10	<i>Neurospora</i> culture media .....	80
2.1.11	Bacterial culture media .....	86
2.1.12	Mammalian cell culture media .....	87
2.2	Methods .....	88
2.2.1	<i>Neurospora</i> methods .....	88
2.2.2	Bacterial methods .....	96
2.2.3	Mammalian cell culture methods.....	99
2.2.4	DNA methods.....	101
2.2.5	Protein methods .....	108
2.2.6	Biochemical methods.....	119
Chapter 3	Vector construction for diversification of human IGF-I in <i>Neurospora</i> ..	123
3.1	Introduction.....	123
3.1.2	Optimising the <i>IGF-I</i> CDS for gene diversification in <i>Neurospora</i> .....	125
3.1.3	Heterologous expression of IGF-I in <i>Neurospora</i> .....	136
3.2	Discussion .....	159
Chapter 4	Transformation, expression and diversification of IGF-I in <i>Neurospora</i>	164
4.1	Introduction.....	164
4.2	Ectopic transformation of <i>Neurospora</i> with pFENIGFI .....	164
4.2.1	Electroporation of <i>Neurospora</i> with pFENIGFI .....	165
4.2.2	Isolation of pFENIGFI transformed <i>Neurospora</i> homokaryons .....	165
4.2.3	PCR confirmation of <i>Neurospora</i> transformation with pFENIGFI.....	165
4.2.4	Repeat <i>Neurospora</i> transformation with pFENIGFI.....	166
4.2.5	Southern blot confirmation of pFENIGFI integration in ectopic <i>Neurospora</i> transformants .....	168
4.2.6	Western blot analysis of extracellular and intracellular protein from ectopic pFENIGFI <i>Neurospora</i> transformants .....	170
4.2.7	IGF-IR ligand binding assay of culture supernatant from ectopic pFENIGFI <i>Neurospora</i> transformants .....	174
4.2.8	IGF-IR kinase phosphorylation assay of culture supernatant from ectopic pFENIGFI <i>Neurospora</i> transformants in FRIES medium.....	181
4.2.9	Optimisation of culturing conditions for <i>Neurospora</i> IGF-I secretion ....	182
4.2.10	Characterisation of IGF-I degradation in <i>Neurospora</i> culture supernatants in the presence of protease inhibitors.....	188
4.3	Characterisation of transplacated IGF-I expression vectors .....	193
4.3.1	Targeted transplacement of the IGF-I expression constructs into <i>Neurospora</i> .....	194
4.3.2	Western blot characterisation of the IGF-I expression constructs in <i>Neurospora</i> .....	202
4.3.3	IGF-IR ligand binding assay of the IGF-I expression constructs in <i>Neurospora</i> .....	211
4.3.4	Characterisation of RIP-generated IGF-I mutants .....	219
4.4	Discussion.....	231
4.4.1	Diversifying the human IGF-I coding sequence in <i>Neurospora crassa</i> ...	231
4.4.2	IGF-I secretion in <i>Neurospora crassa</i> .....	235

4.4.3	Screening of IGF-I variants in <i>Neurospora</i> culture supernatants .....	244
4.4.4	Diversification of human IGF-I in <i>Neurospora crassa</i> : summary and conclusion .....	247
Chapter 5	Generation and characterisation of IGF-II analogues.....	249
5.1	Introduction .....	249
5.2	Construction of IGF-II analogues.....	253
5.2.1	Cloning of <i>E. coli</i> codon optimised IGF-II coding sequence .....	253
5.2.2	<i>DpnI</i> site-directed mutagenesis.....	254
5.2.3	Cloning into pGH(1-11)VNPAPM vector .....	254
5.3	Heterologous expression of IGF-II protein in <i>Escherichia coli</i> .....	255
5.3.1	<i>E. coli</i> transformation and induction .....	255
5.3.2	Purification, refolding and cleavage of IGF-II mutant proteins .....	260
5.3.3	Isolation & purification of C1 protein summary .....	261
5.3.4	Isolation & purification of C2 protein summary .....	262
5.3.5	Isolation & purification of C3 protein .....	262
5.3.6	Quantitation, mass spectrometry analysis and N-terminal sequencing of IGF-II analogues .....	266
5.4	Characterisation of IGF-II analogues.....	266
5.4.1	IR-A interaction with the IGF-II analogues.....	268
5.4.2	IR-B interaction with the IGF-II analogues.....	274
5.4.3	IGF-IR interaction with the IGF-II analogues.....	277
5.4.4	Structural modelling of the IGF analogues .....	283
5.5	Discussion.....	287
5.5.1	C1 interaction with IR-A and IR-B .....	290
5.5.2	C2 interaction with IR-A and IR-B .....	292
5.5.3	C3 interaction with the IR-A and IR-B .....	296
5.5.4	C1 and C2 interaction with the IGF-IR .....	297
5.5.5	C3 interaction with the IGF-IR .....	299
5.5.6	Comparison of assay results from this study with published reports ....	300
5.5.7	Protein structure modeling.....	303
5.5.8	Future work.....	304
5.5.9	Conclusion.....	306
Final discussion	.....	307
Appendix	.....	311
Bibliography	.....	315

## Abstract

Human insulin-like growth factor-I and -II (IGF-I and -II) are structurally homologous ligands with differential binding affinities for the insulin/IGF family of cell surface receptors. High affinity binding of IGF-I and IGF-II ligands to the insulin-like growth factor I receptor (IGF-IR), insulin receptor exon 11- (IR-A) and exon 11+ (IR-B) isoforms stimulates receptor tyrosine kinase activity regulating numerous biological responses such as cell cycle progression, proliferation and cell survival. Deregulation of the IGF-IR and IR-A signalling axis is common in cancer. Both IGF-I and IGF-II have been implicated in cancer initiation and progression due to ligand overexpression and autocrine ligand production.

The focus of this study was to diversify human IGF-I as part of the ongoing development of an *in vivo* gene diversification system associated with the filamentous fungus *Neurospora crassa* (Catchside *et al.*, 2003). The aim of this project was to generate, express, secrete and screen novel IGF-I variants to identify IGF-IR antagonists as potential cancer therapeutics. Mutant IGF-I DNA sequences were generated with high efficiency using the *Neurospora* Repeat-Induced Point Mutation (RIP) & HotSpot-Mediated Recombination (HSMR) gene diversification platforms. Multiple IGF-I expression vectors were constructed and the novel secretion of human IGF-I protein in *Neurospora* was achieved with the fusion of the IGF-I protein to the endogenous glucoamylase (*gla-1*) protein. However, the abundance of mature IGF-I protein in culture medium was relatively low due, at least in part, to extracellular proteolytic degradation. In addition, the IGF-IR binding assay was found to be unsuitable for reliable screening of secreted IGF-I protein in *Neurospora* culture medium. Resolution of these issues were not pursued in this project as the generation of protease deficient *Neurospora* expression strains and development of a more robust screening methodology were already being undertaken by Neugenesis Corporation and were unlikely to be resolved within the duration of this project. Consequently, the project direction was diverted to the characterisation of IGF residues involved in receptor binding and activation through site-directed mutagenesis.

The IGF-I and IGF-II ligands share a common domain structure and exhibit differential receptor binding and activation which is primarily due to the IGF C-domain. Understanding the role of specific residues within the C-domains could make a significant contribution to our fundamental understanding of ligand-receptor interactions in the insulin/IGF system and enable the rational design of IGF-IR and IR-A antagonists for use as anti-cancer agents. As such, a second aspect of this study aimed to delineate the C-domain residues conferring differential receptor binding affinity and activation by generating and characterising three IGF-II analogues. Mutants C1 and C2 contained partial IGF-I C-domain substitutions (IGF-I residues underlined): GYGSSRRSR & SRVSRRAPQT, respectively. Mutant C3 contained a truncated IGF-I C-domain (GSSRRAT) which approximated the size of the IGF-II C-domain. The three IGF-II analogues were characterized by competitive receptor binding assay, receptor phosphorylation assay and cell survival assay using the IR-A, IR-B and IGF-IR receptors.

The C1 analogue bound the IR-A and IGF-IR with high affinity but bound the IR-B with only moderate affinity suggesting a negative interaction between the C1 C-domain and the exon-11 encoded peptide in the IR-B. The IR-A, IR-B and IGF-IR activation and cell survival responses were generally proportional to the relative binding affinity of C1. In contrast, C2 acted as a partial agonist against the IR-A, IR-B and IGF-IR with high receptor binding affinities but substantially lower activation and cell survival responses in all three receptors. These data indicated both flanks of the IGF C-domain play important roles in receptor binding, activation and cell survival. The C3 analogue had generally poor IR-A, IR-B and IGF-IR binding affinities, activation and cell survival responses which indicated this analogue lacked critical C-domain residues required for high affinity receptor binding and activation.

## Declaration

I certify that this thesis does not incorporate without acknowledgment any material previous submitted for a degree or diploma in any university; and that to the best of my knowledge and belief it does not contain any material previously published or written by another person except where due reference is made in the text.

S.T. Henderson



## Acknowledgments

First and foremost, I would like to thank my supervisor, Prof. David Catcheside for being an outstanding mentor and role model. I don't think I can really express how grateful I am for David's unwavering support and guidance over the years. I would also like to thank my co-supervisor Dr. Leah Cosgrove for her continual support throughout my PhD.

I thank Dr. Colin Ward at the CSIRO and Dr. Dorsey Stuart at Neugenesis Corporation for supporting my project. I am very grateful to Dr. Ed Cambareri for generously sharing his expertise with heterologous expression in *Neurospora*. I thank Prof. John Wallace for allowing me to express and purify the IGF-II analogues in his laboratory and Dr. Briony Forbes for her advice and guidance with the IGF-II analogue project. I would also like to thank Peter Hoyne for his advice regarding the Europium receptor binding assay and Dr. Morry Frenkel and John Bentley for generously providing reagents.

There are a number of people who have greatly helped me throughout my candidature and to whom I am deeply appreciative. Dr. Paul Rasmussen, for his friendship over the years and for all his help when I first started in the lab. Dr. Jane Yeadon and Dr. Fred Bowring, for their support and guidance. Graham Eariss and Dr. Gemma Brierley, for sharing the joy of life at the lab bench and various other places around Australia and abroad. Special thanks to Annette Boettcher for her help and support. Ilka Priebe, for being a great friend and always making me laugh with the many spontaneous conversations about anything and everything, and for helping me through when things started getting a bit rough. Tim Gore, for always asking how the PhD was going and for actually listening when I told him. Dr. Cathy Dandie, for offering to read "Stinky – the Wonder Thesis" in her spare time (you mad fool!). Cap'n Dave & Sue Wood and Alex Stropin for the laughs, troughfests and near-death experiences!

To the numerous people who have helped, supported, encouraged and occasionally tolerated me during my PhD. At Flinders University: Dr. James Waters, Dr. Lin Koh, Russell Stainer, Angus Forgan, Dr. Keryn Simons, Dr. Simon Schmidt, Ben James, Dr. Peter Anderson, Jennie Brand and the Higher Degrees Committee for granting the odd extension or two. At the CSIRO: Dr. Tanya Lewanowitsch, Dr. Michelle Zucker, Dr. Damien Belobrajdic, Dr. Leanne Purins, Dr. Cheng Cheng Ooi, Joanne Lane, Dr. Shusuke Toden (who has a promising alternative career as an inspirational speaker), Corinna Bennett, Emma Watson, Brad Klingner, Maryam Hor, Dr. Claus Christophersen, Ben Scherer, Robb Muirhead, Dr. Erin Symonds, Kerry Nyland, Jessica Southwood, Dr. Tom Wycherley, Sharon Zyrna, Darien Sanders, Julie Dallimore, Michael Adams, Kaylene Pickering, Caroline Cooke, Dr. Trevor Lockett, Alison Tuckfield and Dr. Karsten Oelkers (who helped to push me over the finishing line!). At Adelaide University: Dr. Adam Denley, Dr. Clair Alvino, Kerrie McNeil, Carlie Delaine and Chris Cursaro. Thank you all!

Thanks to my family who have supported me over the years, particularly my Mother although she never really understood why I wanted to work on bread mould. The biggest thankyou goes to my partner Cara-Lyn and my children, Eliza and Oliver, for putting up with the many long hours I've spent in the lab and writing this dissertation – I dedicate this thesis to you.

## Abbreviations

%CV	coefficient of variation expressed as a percentage
AA	amino acid
amp	ampicillin
BCA	bicinchoninic acid
Bis-tris	bis(2-hydroxyethyl)imino-tris(hydroxymethyl)methane
boil prep	boiled preparation of <i>Neurospora</i> macroconidial gDNA for PCR
bp	base pair
BAP	biotinylated alkaline phosphatase
BCIAD	IGF-II mutant with IGF-I C-domain
BSA	bovine serum albumin
cat#	catalogue number
C1	IGF-II analogue with N-terminal flank of IGF-I C-domain
C2	IGF-II analogue with C-terminal flank of IGF-I C-domain
C3	IGF-II analogue with shortened IGF-I C-domain
CIP	calf intestinal phosphatase
conc.	concentrated
DMEM	Dulbecco's modified eagles medium
DMSO	dimethyl sulfoxide
DNA	deoxyribonucleic acid
DNase	deoxyribonuclease
dNTP	deoxyribonucleotide triphosphate
DTPA	diethylenetriaminepentaacetate
DTT	dithiothreitol
EC <sub>50</sub>	half maximal effective concentration
<i>E. coli</i>	<i>Escherichia coli</i>
ECL	enhanced chemiluminescence
EDTA	ethylenediaminetetraacetic acid
EtBr	ethidium bromide
Eu	europium
FCS	fetal calf serum
FFS	Fast Flow S (buffer)
Fm	FRIES medium
FmH	FRIES medium supplemented with histidine
FmHS	FRIES medium supplemented with histidine and sorbose
FPLC	fast protein liquid chromatography
gDNA	genomic DNA
HCl	hydrochloric acid
HEPES	N-2-hydroxyethylpiperazine-N`-2-ethanesulfonic acid
hIGF-I	human IGF-I
his	L-histidine
hr	hour(s)
HPLC	high performance liquid chromatography
HRP	horseradish peroxidase
HSMR	Hotspot-mediated recombination

hyg	hygromycin
IB/s	inclusion body/bodies
IC <sub>50</sub>	half maximal inhibitory concentration
IGF	insulin-like growth factor
IGF-I	insulin-like growth factor I
IGF-II	insulin-like growth factor II
IGF-IR	insulin-like growth factor I receptor
IGF-IIR	insulin-like growth factor II receptor
IPTG	isopropyl-beta-D-thiogalactopyranoside
IR	insulin receptor
IR-A	insulin receptor A isoform
IR-B	insulin receptor B isoform
kb	kilobase pair
LB	Luria broth
LB/amp	Luria broth with ampicillin (100µg/ml)
LG	( <i>Neurospora</i> ) linkage group
mAB	monoclonal antibody
min	minute(s)
miniprep	miniscale plasmid preparation
midiprep	midiscale plasmid preparation
mqH <sub>2</sub> O	milli-Q water
NEB	New England Biolabs
nIGF-I	<i>Neurospora</i> RIP-optimised human IGF-I coding sequence
NMR	nuclear magnetic resonance spectroscopy
NSB	non-specific binding
NSmH	<i>Neurospora</i> secretion medium supplemented with histidine
OD <sub>xxxnm</sub>	optical density at wavelength in nm
OPM	orbits per minute
PBS	phosphate buffered saline
PCR	polymerase chain reaction
PIC	Protease inhibitor cocktail (Sigma)
PMSF	phenylmethylsulfonyl fluoride
RIP	repeat-induced point mutation
RNA	ribonucleic acid
RNase	ribonuclease
HPLC	(reverse phase) high performance liquid chromatography
RT	room temperature
s	second(s)
SC	synthetic crossing medium
SDS	sodium dodecyl sulphate
SDS-PAGE	sodium dodecyl sulphate polyacrylamide gel electrophoresis
SEM	standard error of the mean
SGF	sorbose, glucose & fructose
SOC	super optimal catabolite repression medium
SSC	salt sodium citrate buffer
TAE	tris acetate EDTA buffer
TCA	trichloroacetic acid

TE	tris EDTA buffer
TBS	tris buffered saline
TBST	tris buffered saline with tween-20
TFA	trifluoroacetic acid
T <sub>m</sub>	melting temperature
Tris	tris(hydroxymethyl) aminomethane
Triton X-100	<i>iso</i> -octylphenoxypolyethoxyethanol
TSS	transformation and storage solution
Tween 20	polyoxyethylene (20)-sorbitan monolaurate
U	unit of enzyme activity (1 μmol min <sup>-1</sup> )
UV	ultra violet
V	Volts
VmH	Vogel's N medium supplemented with histidine
VmHH	Vogel's N medium supplemented with histidine and hygromycin
vol	volume

## Chapter 1 Introduction

### 1.1 Thesis overview

The aim of this project was to diversify a candidate gene as part of the ongoing development of an *in vivo* gene diversification system associated with the filamentous fungus *Neurospora crassa* (Catchside *et al.*, 2003). Human Insulin-like Growth Factor I (IGF-I) was chosen as the candidate gene for diversification as this ligand and its cognate receptor (IGF-IR) have been well characterized and are associated with many types of cancer. As such, an IGF-I variant with IGF-IR antagonist properties could potentially be used as an anti-cancer agent. The initial focus of this project was to diversify IGF-I using the *Neurospora* gene diversification system to generate and express IGF-I DNA sequence variants, and to screen secreted diversified proteins within the *Neurospora* culture medium to identify variants acting as IGF-IR antagonists.

Directed evolution in *Neurospora* is based on the generation of mutated sequences using the endogenous mutagenic process of Repeat-Induced Point mutation (RIP) coupled with HotSpot Mediated Recombination (HSMR). A number of the primary project objectives were successfully met, namely, the novel expression of human IGF-I in *Neurospora* and the generation of diversified IGF-I DNA sequences. However, key areas of the *Neurospora* platform involving a need for protease deficient expression strains and more robust screening methodology were identified which required additional development to facilitate generation and identification of IGF-IR antagonists. These matters were not pursued in this project as both issues were being addressed by Neugenes Corporation as part of their development of the *Neurospora* directed evolution platform.

An additional research strategy involving the generation and characterisation of site-directed IGF-II mutants was undertaken to further elucidate the mechanism by which the C-domains of IGF-I and IGF-II modulate the differential receptor binding affinities and activation exhibited by the IGFs. Ultimately, a complete understanding of this

mechanism could permit the rational design of receptor antagonists as an alternative to a directed evolution approach.

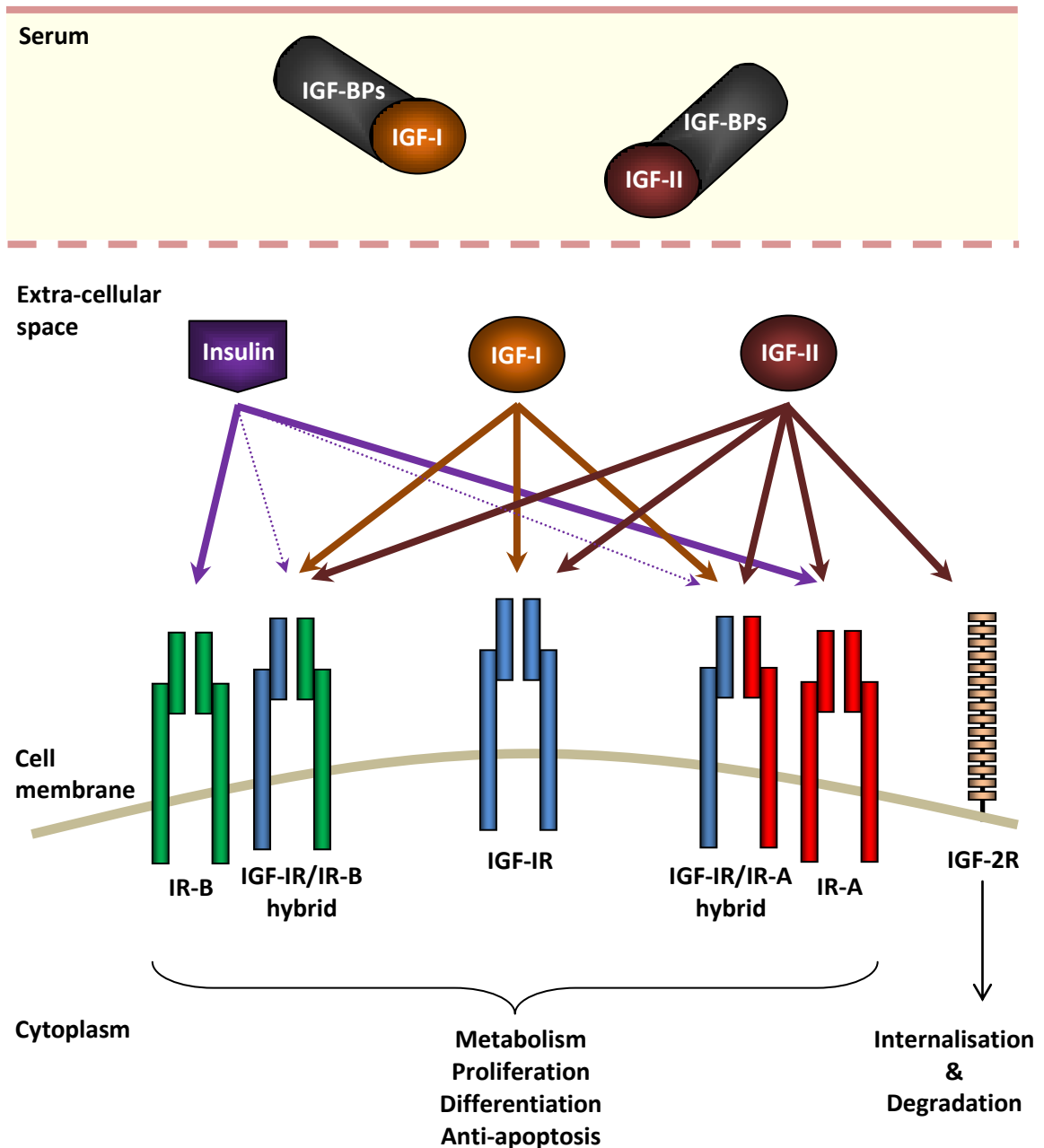
This introduction focuses on the human insulin-like growth factors (IGFs) and the interaction between the various components of the IGF system. The role of the IGF axis in cancer and strategies used to develop cancer therapeutics targeting the IGF system are discussed. In addition, the modification of protein function with directed evolution technologies will be examined with an emphasis on the *Neurospora* directed evolution platform.

## **1.2 The Insulin-like Growth Factor axis**

The insulin-like growth factor axis is a complex molecular network consisting of three peptide ligands (IGF-I, IGF-II and insulin), the insulin/IGF family of cell surface receptors (IGF-IR, IGF-IIR, the insulin receptor exon 11- (IR-A) and exon 11+ (IR-B) isoforms, and IR-A/IR-B and IGF-IR/IR hybrid receptors), six high-affinity IGF binding proteins (IGFBP-1 to IGFBP-6) and several binding protein proteases (Figure 1.1). Binding of IGF-I and IGF-II ligands to the IGF-IR, IR-A and IR-B induces conformational change in the receptor stimulating intrinsic receptor tyrosine kinase activity. Receptor activation initiates an array of various intracellular signalling cascades regulating a variety of biological responses such as cell proliferation, differentiation, apoptosis and chemotaxis. The distribution and bioavailability of IGF ligands in cellular and physiological environments is regulated by IGFBPs and IGFBP proteolysis.

### **1.2.1 IGF evolution**

Vertebrate IGFs have highly conserved amino acid sequences and contain B, C and A domains similar to those of proinsulin (Figure 1.2). Unlike insulin, the IGF C-domains are not proteolytically cleaved during post-translational processing of the prohormone. IGFs also contain D-domains and proIGF carboxy-terminal extensions known as E peptides which are cleaved during prohormone processing (LeRoith *et al.*, 1993).



**Figure 1.1 Summary of interactions in the IGF axis.**

Circulating IGF-I and IGF-II are complexed to IGFBPs (IGFBP1 to IGFBP-6) regulating ligand bioavailability in tissues. IGF-I binds to the IGF-IR and the hybrid IGF-IR/IR-A and IGF-IR/IR-B with relatively high affinity (solid arrows). IGF-II binds to the IGF-IR, IGF-IR, IR-A and the IGF-IR/IR-A and IGF-IR/IR-B hybrid receptors with relatively high affinity. Insulin binds the IR-A and IR-B with high affinity but binds the IGF-IR / IR hybrid receptors with low affinity (dotted arrows). Relative receptor binding affinities of the IGF-I, IGF-II and insulin ligands are based on data reported by Denley *et al.* (2004) and Benyoucef *et al.* (2007). Figure adapted from Sachdev & Yee (2007).





IGFs and insulin have common ancestral origins with IGFs postulated to have stemmed from duplication of a primordial insulin-like gene early in vertebrate evolution. The identification of a single insulin-like peptide (ILP) gene in an amphioxus (lancelet) (*Branchiostoma californiensis*) with 48% identity to both insulin and the IGFs and containing a D domain led to the proposal that the cephalochordate gene represented the ancestral gene to both vertebrate insulin and the IGFs (Chan *et al.*, 1990). This assertion was challenged by the discovery of separate insulin and IGF genes in the urochordate (*Chelyosoma productum*) (McRory & Sherwood 1997) which suggested the two hormones diverged earlier. However, high sequence identity between the insulin and IGF peptides, particularly in the evolutionarily highly variable C domain suggested the urochordate insulin and IGF sequences were likely to be products of a relatively recent gene duplication and not true orthologues of the vertebrate insulin and IGF genes (Chan 2000).

Modeling of IGF evolution in vertebrates and lower chordates is based on analysis of cDNA sequences from taxa representing each of the major vertebrate lineages (Collet 1997). The cloning of separate insulin and prepro-IGF genes in the agnathan (jawless fish) Atlantic hagfish indicated the gene duplication event responsible for separate insulin and IGF genes occurred prior to vertebrates, somewhere between the branching of cephalochordate (lancelets) and agnathan lineages over 550 million years ago (Nagamatsu *et al.*, 1991). Whereas the gene duplication which subsequently diverged into IGF-I and IGF-II is thought to have occurred more than 440 million years ago before chondrichthyes (cartilaginous fishes) diverged from the tetrapod lineage (Duguay *et al.*, 1995).

### **1.2.2 Gene structure, expression and physiological effects of the IGFs**

The elucidation of the amino acid sequences of IGF-I (Rinderknecht & Humbel 1978a) and IGF-II (Rinderknecht & Humbel 1978b) was followed by isolation and sequencing of cDNA clones for IGF-I (Jansen *et al.*, 1983) and IGF-II (Bell *et al.*, 1984); (Jansen *et al.*, 1985). Subsequent characterization of the genes encoding the IGF prepropeptides determined the IGF-I and IGF-II exons were dispersed over ~90kb on chromosome 12

and ~35kb on chromosome 11, respectively (Brissenden *et al.*, 1984); (Bell *et al.*, 1985); (Rotwein *et al.*, 1986); (de Pagter-Holthuisen *et al.*, 1987).

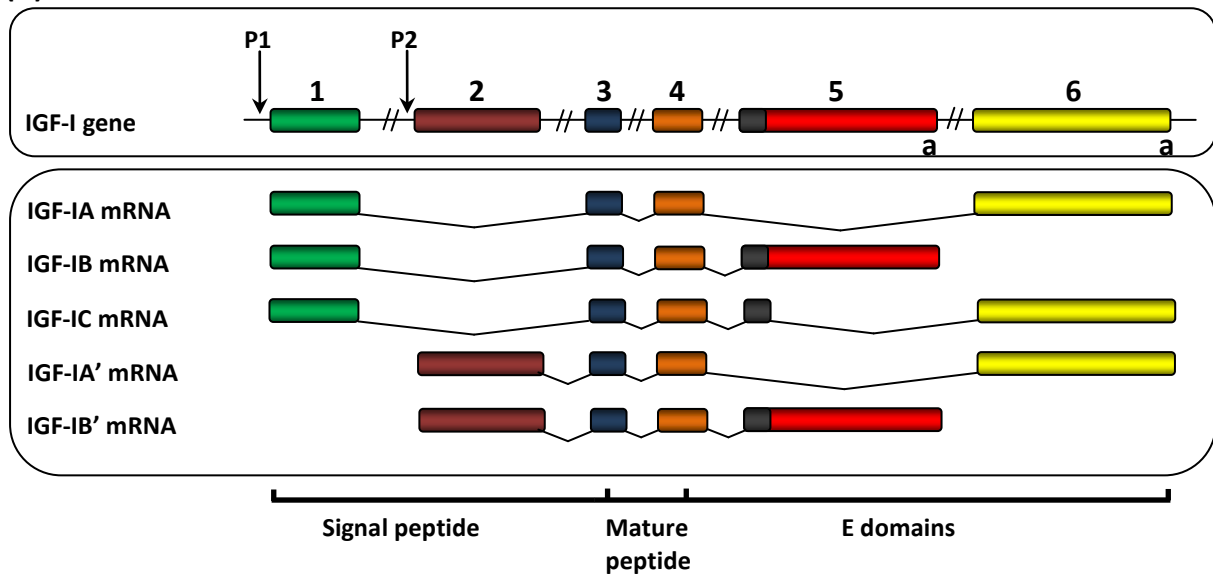
Studies of IGF gene structure and transcription in humans and rodents have shown IGF expression to be complex with multiple promoters acting on multiple initiation sites, alternate splicing of coding and noncoding exons and differential polyadenylation producing numerous transcripts specific to different tissues and developmental stages (Figure 1.3) (Jansen *et al.*, 1983); (Bell *et al.*, 1984); (Bell *et al.*, 1986); (Soares *et al.*, 1986); (Shimatsu & Rotwein 1987b); (Shimatsu & Rotwein 1987a); (Roberts *et al.*, 1987).

The IGF genes are pleiotropic, exerting a wide range of biological effects and are fundamentally involved in the regulation of somatic growth, cell proliferation, transformation and apoptosis (Sandhu *et al.*, 2002).

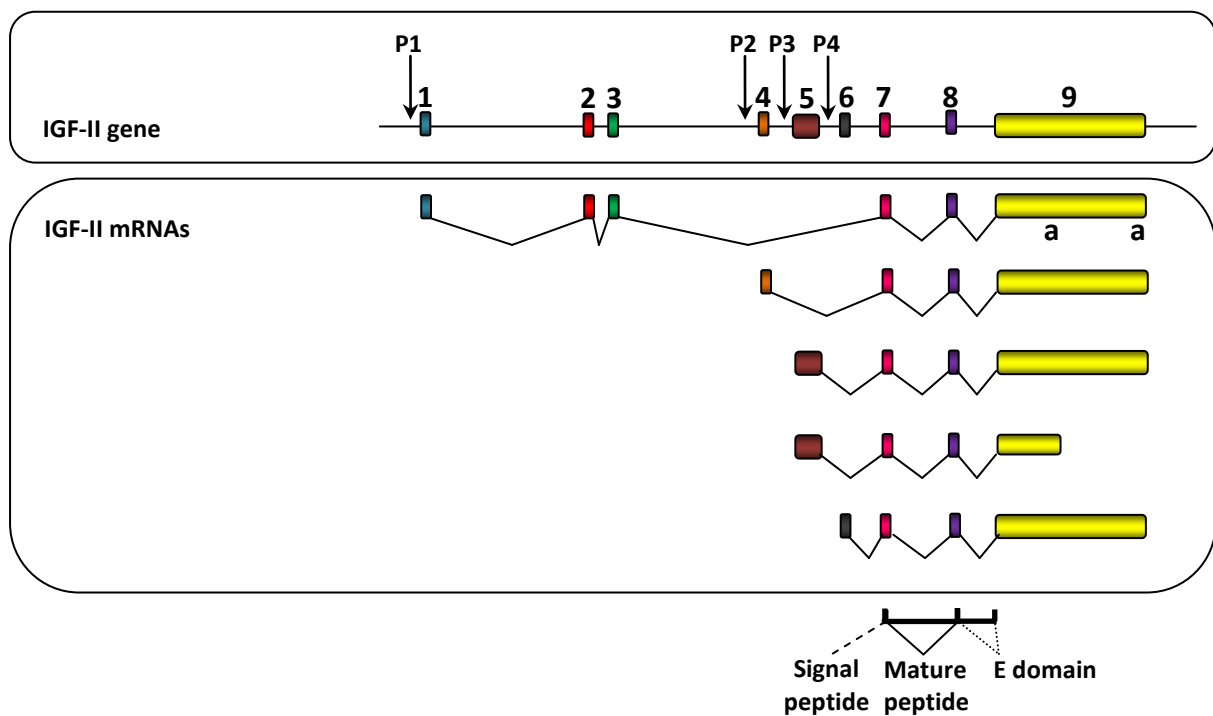
#### **1.2.2.1 IGF-I gene structure, expression and physiological effects**

The human *igfl* gene consists of six exons that undergo alternative splicing to generate five mRNA transcript isoforms each containing four exons (Figure 1.3A). Exons 1 and 2 are associated with distinct promoters and 5' non-coding sequences generating class 1 (IGF-IA, IGF-IB & IGF-IC) and class 2 (IGF-IA' & IGF-IB') mRNAs, respectively. The IGF-IC transcript is formed by the splicing of the first 49 bases of exon 5 to exon 6 through a cryptic 5' donor splice site (Chew *et al.*, 1995). Exons 1 and 2 encode the N-terminus of the signal peptide. Exons 3 and 4 encode the rest of the signal peptide and the coding sequence for mature IGF-I. Exon 4 also codes for a portion of the E-domain with the remainder of the E region encoded by exons 5 and 6 (Hameed *et al.*, 2003). IGF-I transcripts isolated in liver are predominately controlled by the P2 promoter whilst the P1 promoter is typically active in extrahepatic tissues (O'Sullivan *et al.*, 2002).

(A)



(B)



**Figure 1.3 Genomic organization and expression of the human IGF-I and IGF-II genes**

The exon-intron organization of the human IGF-I (A) and IGF-II (B) genes and alternative splicing of mRNA transcripts are depicted. Exons are represented by boxes, promoters by the letter P and vertical arrows, polyadenylation sites are shown by the letter a. Alternatively spliced transcripts are shown below the gene maps. Exons containing coding sequence are depicted by the annotated bar below the transcript boxes. Figure adapted from Rotwein (1991) & Reinecke & Collet (1998).

IGF-I is required for pre and postnatal growth and is essential for normal development and growth (O'Sullivan *et al.*, 2002). The IGF-IR mediates the effects of IGF-I ligand to regulate various aspects of cell function including cell proliferation, terminal differentiation, migration and protection from apoptosis (Khandwala *et al.*, 2000). In addition, IGF-I may play an important metabolic role either directly through the IGF-IR, or indirectly, via inhibition of Growth Hormone (GH) (Sandhu *et al.*, 2002).

### **1.2.2.2 IGF-II gene structure, expression and physiological effects**

The human *Igf2* gene contains nine exons, four different promoters and multiple 5' untranslated exons with the coding region for prepro-IGF-II located within exons 7-9 (Figure 1.3B). Exon 7 encodes a 24-residue signal peptide and the N-terminus of the mature IGF-II peptide. Exon 8 encodes the rest of the IGF-II peptide and the initial eleven amino acids of the E domain whilst exon 9 contains the remainder of the E domain and 3' untranslated exons (Rotwein 1991). The activity of the various IGF-II promoters is specific to tissue and developmental stage with P1 predominantly used in human adult liver whilst the other promoters are differentially active in various fetal tissues with P3 being the primary foetal liver promoter (Reinecke & Collet 1998).

IGF-II is an important regulator of foetal growth with a less clear role in adult human physiology (Fottner *et al.*, 1998). The IGF-II ligand is involved in numerous aspects of development involving tissues such as muscle, bone and brain (Chao & D'Amore 2008). IGF-II can also stimulate metabolic responses through the IR-B. The ability of IGF-II to signal through both the IGF-IR and the IR isoforms potentially enables a broader range of biological functions than IGF-I (Chao & D'Amore 2008).

### **1.2.3 Regulation of the IGFs**

Both IGF-I and IGF-II are produced in multiple organs and tissues throughout the human life cycle (LeRoith & Roberts 2003). IGF regulation can be specific to developmental stage and tissue-type. Regulation is mediated by complex transcriptional and post-transcriptional mechanisms and modulation of protein

stability and bioavailability (Nussbaum *et al.*, 2008). The liver is the primary source of IGF-I and IGF-II production in humans (Rall *et al.*, 1987) although many extrahepatic tissues are also capable of IGF synthesis. Thus, the IGFs exhibit endocrine, paracrine and autocrine regulatory mechanisms (Sciacca *et al.*, 1999); (Bustin & Jenkins 2001).

### **1.2.3.1 IGF-I regulation**

Hepatic production of IGF-I is largely regulated by circulatory levels of Growth Hormone secreted by the pituitary (Merimee 1982) whereby GH stimulates IGF-I gene transcription leading to increased serum IGF-I concentrations. However, the ability of GH to control IGF-I expression in extrahepatic tissues appears to be tissue-specific (Mathews *et al.*, 1986); (Roberts *et al.*, 1987). Regulation of IGF-I synthesis is also dependant on nutritional status (Underwood *et al.*, 1994), additional hormonal control and genetic factors (Cohen *et al.*, 2000).

### **1.2.3.2 IGF-II regulation**

IGF-II regulation is complex and multifaceted but unlike IGF-I, the synthesis of IGF-II is relatively GH independent. In rodents, prenatal expression of the *Igf2* gene is highly prevalent but postnatal expression is greatly diminished with *Igf2* transcripts detectable only in the leptomeninges and the choroid plexus in adult animals (DeChiara *et al.*, 1991). As such, IGF-II was initially thought to be a fetal growth factor. However, human expression studies demonstrated that both IGFs occurred pre-and postnatally in multiple human tissues with *Igf2* expression in adults primarily induced by the adult-specific P1 promoter (Li *et al.*, 1996; Li *et al.*, 1997) resulting in adult IGF-II circulatory levels being 3-4 times higher than IGF-I (Clemmons & Underwood 1991). In mice, there appears to be a non-functional remnant of the P1 promoter which presumably explains the different temporal expression patterns between mice and humans (Rotwein & Hall 1990).

The *Igf2* gene is regulated by genomic imprinting of the maternal allele in most human tissues via imprinting mechanisms involving genomic *cis*-acting DNA sequences, differential methylation and a non-translated RNA transcript. As such, only the paternally inherited *Igf2* allele is active (Ripoche *et al.*, 1997); (Thorvaldsen *et al.*, 1998); (Moore *et al.*, 1997). Monoallelic expression of *Igf2* in most tissues is required as loss of imprinting (LOI) results in increased gene dosage effects which has been implicated in some cancers (discussed in section 1.9.2).

Regulation of IGF-II does not occur solely at the level of gene expression. IGF-II binds to the IGF-IIR/Mannose-6-phosphate receptor, a 270kDa glycosylated protein with 15 extracellular repeat domains containing distinct binding sites for multiple ligands including IGF-II and mannosylated proteins such as leukemia inhibitory factor (Figure 1.1). IGF-II binds with high affinity to a localised binding site on the IGF-IIR comprised of amino acids 1508-1566 in domain 11 (Schmidt *et al.*, 1995). The cytoplasmic tail of the IGF-IIR contains recognition sequences facilitating membrane bound endocytosis to the golgi complex, endosomes and subsequent packaging into lysosomes (Rohrer *et al.*, 1995). Approximately 10% of membrane bound IGF-IIR is found at the cell surface with the remainder found within the cell (Linnell *et al.*, 2001). Thus, the internalisation and degradation of the IGF-II:IGF-IIR complex is a rapid process effectively limiting the bioavailability of IGF-II ligand.

### **1.2.3.3 IGF binding proteins**

The IGF binding proteins are a family of six multifunctional proteins (IGFBP-1 to -6) which bind with high affinity and specificity to IGF-I and IGF-II (Clemmons *et al.*, 1992); (Bach *et al.*, 1993). The six IGFBPs have an overall 40-60% amino acid sequence identity and consist of cysteine-rich N and C-terminal domains coupled together by flexible linker regions (Kim *et al.*, 1997). The IGFBPs have a higher affinity for the IGFs than the IGF receptors and modulate IGF availability, bioactivity and tissue distribution both positively and negatively (Jones & Clemmons 1995). The IGFBPs serve various biological functions and expression differs both spatially and temporally (Chao & D'Amore 2008) (Table 1.1).

<b>IGFBP</b>	<b>IGFBP characteristics</b>
<b>IGFBP-1</b>	<ul style="list-style-type: none"> <li>- Physiological levels stimulate IGF-I action</li> <li>- Molar excess inhibits mitogenic and insulin-like effects of IGF-I &amp; IGF-II</li> </ul>
<b>IGFBP-2</b>	<ul style="list-style-type: none"> <li>- Inhibits IGF-induced DNA synthesis</li> <li>- Observed to potentiate IGF function</li> </ul>
<b>IGFBP-3</b>	<ul style="list-style-type: none"> <li>- Major carrier of IGF-I &amp; IGF-II in serum</li> <li>- Modulates IGF endocrine action</li> <li>- Potentiates IGF activity</li> <li>- Excess levels are inhibitory</li> </ul>
<b>IGFBP-4</b>	<ul style="list-style-type: none"> <li>- Consistently inhibits IGF action</li> <li>- Serum concentration is generally low</li> <li>- Tissue-specific expression</li> </ul>
<b>IGFBP-5</b>	<ul style="list-style-type: none"> <li>- Inhibitory action</li> <li>- Association with extracellular matrix lowers affinity for the IGFs thereby increasing IGF activity</li> </ul>
<b>IGFBP-6</b>	<ul style="list-style-type: none"> <li>- Specifically binds IGF2</li> <li>- Generally thought to be inhibitory</li> </ul>

**Table 1.1 Insulin-Like Growth Factor Binding Protein functions**

The major functions and characteristics are presented for each individual IGFBP. Table adapted from Chao & D'Amore (2008).

Extensive studies utilizing NMR spectroscopy and site-directed mutagenesis of IGFbps have indicated that both the N and C-terminal domains contain IGF binding sites required for wildtype binding affinity (Baxter 2000); (Clemmons 2001). In contrast to insulin, which circulates freely in the bloodstream, the majority of IGFs in circulation and interstitial fluids are found complexed with the IGFbps. Most circulating IGF-I and IGF-II is bound in a 150kDa ternary complex with IGFBP-3 and an 80kDa acid-labile subunit (ALS), increasing the IGF half life and acting as a reservoir for IGF-I and IGF-II. Similarly, IGFBP-5 also forms ternary complexes with the IGFs and ALS. IGFBP-1 to -4 have similar binding affinities for both IGFs. IGFBP-5 and -6 bind IGF-II with 10-fold and 100-fold higher affinity than IGF-I, respectively (LeRoith & Roberts 2003). IGFbps typically inhibit IGF action by reducing bioavailability although under various conditions some of the IGFbps may potentiate IGF actions (Sandhu *et al.*, 2002). In addition, IGFBP-1, IGFBP-3 and IGFBP-5 exhibit IGF-IR independent effects on proliferation, migration and apoptosis (LeRoith & Roberts 2003).

The IGFbps are themselves regulated by a number of mechanisms. Post-translational phosphorylation of IGFBP-1 has been shown to increase binding affinity for IGF-I 6-fold (Jones *et al.*, 1991). The presence of various IGFBP proteases can modulate the levels of IGFBP by limited proteolytic cleavage which in turn regulates IGF bioavailability. Several of the IGFbps bind to the cell surface or to the extracellular matrix and typically have decreased binding affinity for the IGFs (Baxter 2000). Furthermore, cell surface or matrix associated IGFs may potentially influence IGF action by direct receptor interaction (LeRoith & Roberts 2003).

In summary, the complex molecular mechanisms involved in the IGFBP regulation of IGF-I and IGF-II are affected by numerous factors such as IGFBP expression levels, tissue distribution, phosphorylation, proteolysis and cell surface association.



#### 1.2.4 Structure of IGF-I and IGF-II

IGF-I and IGF-II are synthesised as prepro-proteins with 48 and 24 amino acid signal sequences, respectively, at the amino termini and an 'E domain' at the carboxy termini. There are three pro-IGF-I E-domain isoforms as a result of alternative splicing of the IGF-I mRNA (Daughaday & Rotwein 1989); (Chew *et al.*, 1995) and numerous proIGF-II isoforms resulting from differential cleavage and glycosylation within the proIGF-II E-domain (Zumstein *et al.*, 1985); (Hudgins *et al.*, 1992); (Duguay *et al.*, 1998). Mature IGFs are formed by the proteolytic cleavage of signal sequences and E-domains as the IGFs are transported through the constitutive secretory pathway (Duguay *et al.*, 1998).

Mature IGF-I and IGF-II are single chain polypeptides of 70 and 67 amino acid residues, respectively, which are ~60% identical between their amino acid sequences and share significant homology with insulin (Figure 1.2). Mature IGFs are organised into four domains designated B, C, A & D beginning from the amino termini. The B and A-domains of the IGFs are ~50% homologous to the B and A-domains of insulin whereas the D region has no counterpart in insulin (Rinderknecht & Humbel 1978b). The C-domains of the IGFs are analogous to the C-domain of proinsulin in that they connect the B and A chains together. Despite the common domain classification, the C regions of the IGFs share little sequence homology to each other and have no sequence similarity to proinsulin. Furthermore, the C-domains of the IGFs are not proteolytically removed as occurs during the maturation of proinsulin to form insulin.

Elucidation of the tertiary structures of the IGFs is a necessary prerequisite towards understanding the relationship between the structures of the IGFs and their biological functions. Seminal molecular models of the IGFs were proposed based on homology modelling using the known crystal structure of insulin (Blundell *et al.*, 1978); (Blundell *et al.*, 1983). The tertiary structure of IGF-I has been determined experimentally by NMR spectroscopy (Cooke *et al.*, 1991); (Sato *et al.*, 1993); (Schaffer *et al.*, 2003) & (Kuang *et al.*, 2009) and X-ray crystallography (Vajdos *et al.*, 2001); (Zeslawski *et al.*, 2001); (Brzozowski *et al.*, 2002). Similarly, NMR-derived IGF-II structures (Terasawa *et al.*, 1994); (Torres *et al.*, 1995) and crystal structures of IGF-II complexed with domains

11-12-13 of the IGF-IIR (Brown *et al.*, 2008), and IGF-II bound to the parental Fragment antigen-binding (Fab) region of the DX-2647 antibody (Dransfield *et al.*, 2010) have been determined. These studies have confirmed the structure of the IGFs and insulin are generally similar to each other and indicate the major secondary structural elements of the IGFs consist of three alpha helices surrounding a hydrophobic core (Figure 1.2). Alpha helix 1 (Ala8→Cys18 of IGF-I; Glu12 →Cys21 of IGF-II) is present in the B domains whilst alpha helix 2 (Gly42→Cys48 of IGF-I; Glu44→Arg49 of IGF-II) and alpha helix 3 (Leu54→Cys61 of IGF-I; Leu53→Tyr59 of IGF-II) are located in the A domains and the latter two helices are configured approximately antiparallel to each other and perpendicular to alpha helix 1. The tertiary structure of the IGFs is stabilised by three disulphide bonds (Cys6-Cys48; Cys18-Cys61; Cys47-Cys52 in IGF-I; Cys9-Cys47; Cys21-Cys60; Cys46-Cys51 in IGF-II).

The solution-state structural and crystallographic studies described above have indicated that the B-domains of the IGFs form similar stable folding motifs. In contrast, there are conformational differences in the A-domain of IGF-I and IGF-II with alpha helix 3 in IGF-II being more defined and lacking a bend in the backbone around residues Leu53 and Ala54 resulting in the A-domain of IGF-II being closer to the B chain with helix 2 shifted slightly towards the N-terminus of the B-domain (Torres *et al.*, 1995). It has been postulated that the conformational variability of the A-domain may contribute to the differential binding characteristics of the IGFs to IGF-BPs and receptors (Schaffer *et al.*, 2003).

NMR studies have indicated the C and D-domains of the IGFs lack definition which suggested these regions exhibited conformational flexibility (Cooke *et al.*, 1991); (Terasawa *et al.*, 1994); (Schaffer *et al.*, 2003). The presumed flexibility of the IGF-I C-domain was also supported by Far-UV CD spectroscopic analysis of IGF-I where little regular secondary structure was noted in the C-domain (Gill *et al.*, 1996). X-ray crystallography of IGF-I using detergents to stabilise the structure have indicated residues Tyr24-Lys27 form a type VIII  $\beta$ -turn that permits the C-domain to extend out from the core of the molecule and residues Gly30-Ser33 form a type-II  $\beta$  turn that displays Tyr31 for receptor interaction (Vajdos *et al.*, 2001); (Zeslawski *et al.*, 2001);

(Brzozowski *et al.*, 2002). However, these x-ray crystallography data should be interpreted with some caution as crystallization of the IGF-I molecule may influence the flexible C and D-domains and prevent normal solution conformations.

### **1.3 The function of the Insulin-like Growth Factor I Receptor**

The IGF-IR binds IGF-I and IGF-II with high affinity and mediates the numerous biological activities of these ligands including mitogenesis, transformation, apoptosis inhibition, cell adhesion and longevity (Baserga *et al.*, 2003). The IGF-I receptor also plays a critical role in growth and development with IGF-IR knock-out mice exhibiting severe dwarfism, muscle hypoplasia, abnormal CNS development, translucent skin, delayed bone development and an inability to survive postnatally due to respiratory failure (Liu *et al.*, 1993).

### **1.4 The structure of the Insulin-like Growth Factor I Receptor**

The IGF-I receptor is a transmembrane tyrosine kinase receptor with 70% homology to the insulin receptor (Baserga *et al.*, 2003). The IGF-IR is synthesised as a single polypeptide, glycosylated on the extracellular region and proteolytically cleaved to produce separate  $\alpha$  and  $\beta$ -subunits. The  $\alpha$ -subunit and a portion of the  $\beta$ -subunit are extracellular whilst the intracellular region of the  $\beta$  subunit possesses tyrosine kinase activity (Ullrich *et al.*, 1986). The  $\alpha$  and  $\beta$ -subunits are linked by disulfide bonds to form an  $\alpha\beta$  dimer. Two dimers are themselves disulfide bonded to form the mature IGF-IR (Bhaumick *et al.*, 1981); (Chernausek *et al.*, 1981); (Massague & Czech 1982). The IGF-IR and IR share a similar number and arrangement of structural domains (refer to insulin receptor structure in Figure 1.4).

Although the complete structure of the IGF-IR has not yet been resolved, the structure of the first three extra-cellular domains (L1-cysrich-L2) has been determined by X-ray crystallography (Garrett *et al.*, 1998). This receptor fragment was shown to have an extended bilobal conformation with the L1 & L2 domains flanking the cys-rich region at either end. The cys-rich region runs two thirds the length of the structure and

contacts along the length of the L1 domain but has minimal contact with the L2 domain. There is an  $\sim 24$  Å cavity at the centre of the molecule that is large enough to occupy the IGF ligands (Garrett *et al.*, 1998). Although the L1-CR-L2 molecule is devoid of ligand binding activity, fusion to a peptide from the C-terminus of the receptor  $\alpha$ -subunit (amino acids 692-707) results in a mini-receptor that binds IGF-I with an affinity near that of the recombinant extra-cellular domain indicating a major IGF-I binding site is formed by these regions (Kristensen *et al.*, 1999).

Ligand binding to the extra-cellular region of the IGF-IR results in a conformational shift facilitating *trans*-autophosphorylation of three tyrosine residues (Tyr1131, Tyr1135 & Tyr1136) in the activation loop of the kinase domain in one monomer by the kinase domain of the second monomer (Adams *et al.*, 2000). Receptor autophosphorylation is the initial step in the activation of a complex series of signal transduction pathways which potentiate the biological effects of IGF-IR ligands (discussed in 1.8.3)

## 1.5 Insulin receptor isoforms and their functions

The human insulin receptor exists as two isoforms, IR-A and IR-B, resulting from alternative splicing of the primary IR transcript (Seino & Bell 1989). The IR-A differs from the IR-B by the exclusion of exon 11, which encodes 12 amino acids at the carboxy terminus of the IR  $\alpha$ -subunit (Seino *et al.*, 1989). The presence or absence of the exon 11-encoded peptide in the IR modulates the receptor binding affinities of insulin and the IGFs. Although the differential ligand binding affinities for the IR isoforms reported in the literature vary widely, the overall trends are generally consistent. IGF-I & IGF-II bind the IR-A with higher affinities relative to the IR-B (Frasca *et al.*, 1999); (Denley *et al.*, 2004); (Benyoucef *et al.*, 2007). There are discrepancies within the literature regarding insulin binding to the IR isoforms with some reports indicating insulin binds the IR-B with slightly higher affinity than the IR-A (Whittaker *et al.*, 2002); (Denley *et al.*, 2004) whereas other studies suggest insulin binds the IR-A with slightly higher affinity (Mosthaf *et al.*, 1990), (Benyoucef *et al.*, 2007); (Pandini *et al.*, 2002). Binding kinetic studies showed insulin has slightly faster association and

dissociation rates for the IR-A than the IR-B (Yamaguchi *et al.*, 1993) supporting studies which indicated insulin has a higher affinity for the IR-B isoform. The IR-B has greater kinase activity and more efficient coupling to the insulin receptor substrate 1 (IRS-1) relative to the IR-A (Kosaki *et al.*, 1995). In addition, the IR-A exhibits a faster rate of receptor endocytosis and recycling than the IR-B (Vogt *et al.*, 1991). As such, it is unsurprising the two IR isoforms are differentially regulated in various tissues (Moller *et al.*, 1989) and have diverse functional properties.

### **1.5.1 Distribution and functions of the IR-A**

The IR-A is ubiquitously expressed in adults and is the only IR isoform in brain, spleen and lymphocytes (Pashmforoush *et al.*, 1994). Furthermore, the IR-A is predominately expressed in embryogenesis and cancer tissues (Belfiore *et al.*, 2009). The IR isoforms bind insulin and IGF-II and these two ligands induce partially different gene expression profiles (Pandini *et al.*, 2004) corresponding with the predominately mitogenic effects of IGF-II and metabolic effects of insulin mediated through the IR-A (Frasca *et al.*, 1999).

### **1.5.2 Distribution and functions of the IR-B**

The IR-B is predominant in the liver, skeletal muscle, adipocytes and the kidneys. These tissues are responsive to the metabolic effects of insulin such as regulation of glucose homeostasis through synthesis and storage of carbohydrates, lipids and proteins (Saltiel & Pessin 2002). In addition, IGF-II can also stimulate metabolic responses through binding to the IR-B (Chao & D'Amore 2008).

## 1.6 Hybrid receptors and their functions

When co-expressed in a cell, the high degree of sequence and structural homology between the IR-A, IR-B and IGF-IR facilitates heterodimerisation to form IR-A/IR-B, IR-A/IGF-IR and IR-B/IGF-IR hybrid receptors (Blanquart *et al.*, 2008); (Soos *et al.*, 1993).

### 1.6.1 IR-A/IR-B hybrid receptors

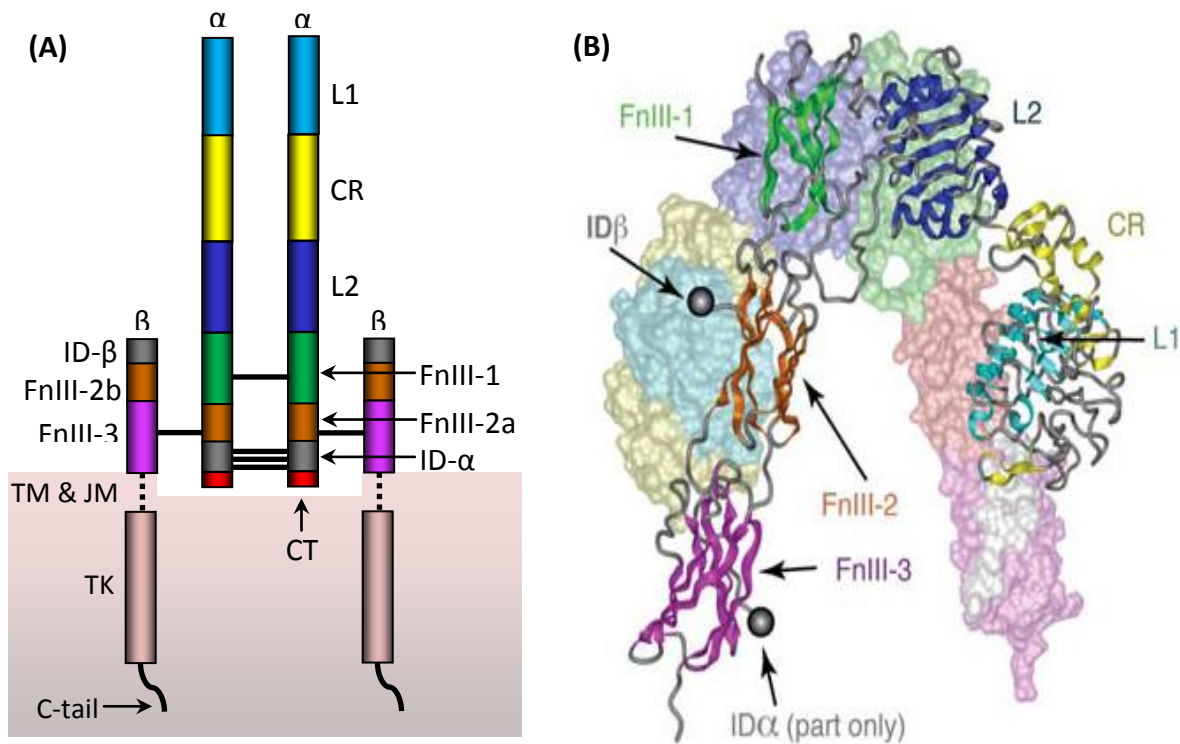
Functional studies of IR-A/IR-B hybrids have been difficult due to the lack of adequate methodologies and the concomitant cellular expression of homodimers of each isoform (Belfiore *et al.*, 2009). A study using bioluminescence resonance energy transfer (BRET) showed homodimeric and IR-A/IR-B hybrids formed with equal probability in cells suggesting the increased expression of IR-A in certain tissues results in a greater number of hybrid receptors to the detriment of IR-B homodimers (Blanquart *et al.*, 2008). Hybrid receptors bind IGF-II with the same affinity as IR-A homodimers, and can transmit intracellular signals upon ligand stimulation by recruiting interacting partners (Blanquart *et al.*, 2008). Thus, the formation of hybrid IR-A/IR-B receptors increases a cells capacity to potentiate the mitogenic effects of IGF-II.

### 1.6.2 IR/IGFR-IR hybrid receptors

The existence of IR/IGF-IR hybrid receptors was first confirmed in studies of human placenta (Soos & Siddle 1989). As with IR-A/IR-B hybrids, the formation of homodimers (IR & IGF-IR) and heterodimers (IR-A/IGF-IR & IR-B/IGF-IR) occurs by random assembly and reflects the molar ratio of IR and IGF-IR half receptors in a given cell (Baillyes *et al.*, 1997); (Pandini *et al.*, 2002); (Slaaby *et al.*, 2006). Studies with IR-A/IGF-IR and IR-B/IGF-IR receptors indicated hybrids bound IGF-I & IGF-II with an affinity similar to homodimeric IGF-IR whilst insulin bound with low affinity to both hybrids (Slaaby *et al.*, 2006); (Benyoucef *et al.*, 2007). Similarly, autophosphorylation studies of the IR/IGF-IR hybrids in intact cells indicated IGF-I elicited a strong activation response whereas the insulin responses in hybrids were near basal level. Together, these results suggest IR/IGF-IR hybrids function more like IGF-I receptors than insulin receptors (Benyoucef *et al.*, 2007).

## 1.7 The structure of the Insulin Receptor

The insulin receptor is a glycosylated homodimeric  $\alpha_2\beta_2$  tyrosine kinase receptor. Each receptor monomer consists of the following structural domains: the first leucine-rich repeat (L1) domain; the cysteine-rich (CR) region; the second leucine-rich repeat (L2) domain; the three fibronectin type III domains (FnIII-1, -2, -3); the transmembrane (TM) region, the juxta-membrane (JM) region; the tyrosine kinase (TK) domain and the C-terminal tail (C-tail). The FnIII-2 domain contains the insert domain (ID) which is proteolytically cleaved by furin to produce the N-terminal  $\alpha$ -chain and the C-terminal  $\beta$ -chain. A single disulphide bond links the  $\alpha$ - and  $\beta$ -chains and four disulphide bonds link the two monomers to form the mature receptor homodimer (Whitten *et al.*, 2009) (Figure 1.4A). The crystal structure of the IR-A ectodomain has shown each monomer assumes an inverted 'V' conformation relative to the cell membrane with the L1-CR-L2 region of one monomer forming one side and the FnIII domains forming the other side (McKern *et al.*, 2006). In the receptor dimer, the L1-CR-L2 region of one monomer is juxtaposed against the FnIII-1 & FnIII-2 domains of the other receptor monomer (Figure 1.4B). The intracellular component of the  $\beta$ -subunit of each monomer contains the tyrosine kinase domain flanked by the juxtamembrane and C-tail regulatory regions. The juxtamembrane region is involved in docking insulin-receptor substrates (IRS)1-4 and Shc, as well as receptor internalization via clathrin-coated pits. The C-tail region contains two phosphotyrosine binding sites (De Meyts & Whittaker 2002); (Belfiore *et al.*, 2009).



**Figure 1.4 Structure and domain organization of the insulin receptor**

(A) Cartoon representation of the structural domain organization within the insulin receptor heterotetramer. L1 & L2: large (leucine-rich-repeat) domains; CR: cysteine-rich domain; FnIII-1, -2, -3: fibronectin type III domains; ID-  $\alpha$  & ID-  $\beta$ : insert domains ( $\alpha$  and  $\beta$ -chain components); CT: C-terminal segment (last 16 amino acids of the  $\alpha$  – chain); TM & JM: transmembrane and juxta-membrane domains; TK: tyrosine kinase domain; C-tail: carboxy-terminal tail;  $\alpha$  &  $\beta$ : alpha and beta-chains. Black solid lines denote disulphide bonds. Intracellular receptor regions are located within shaded area. Figure adapted from Lawrence *et al.* (2007).

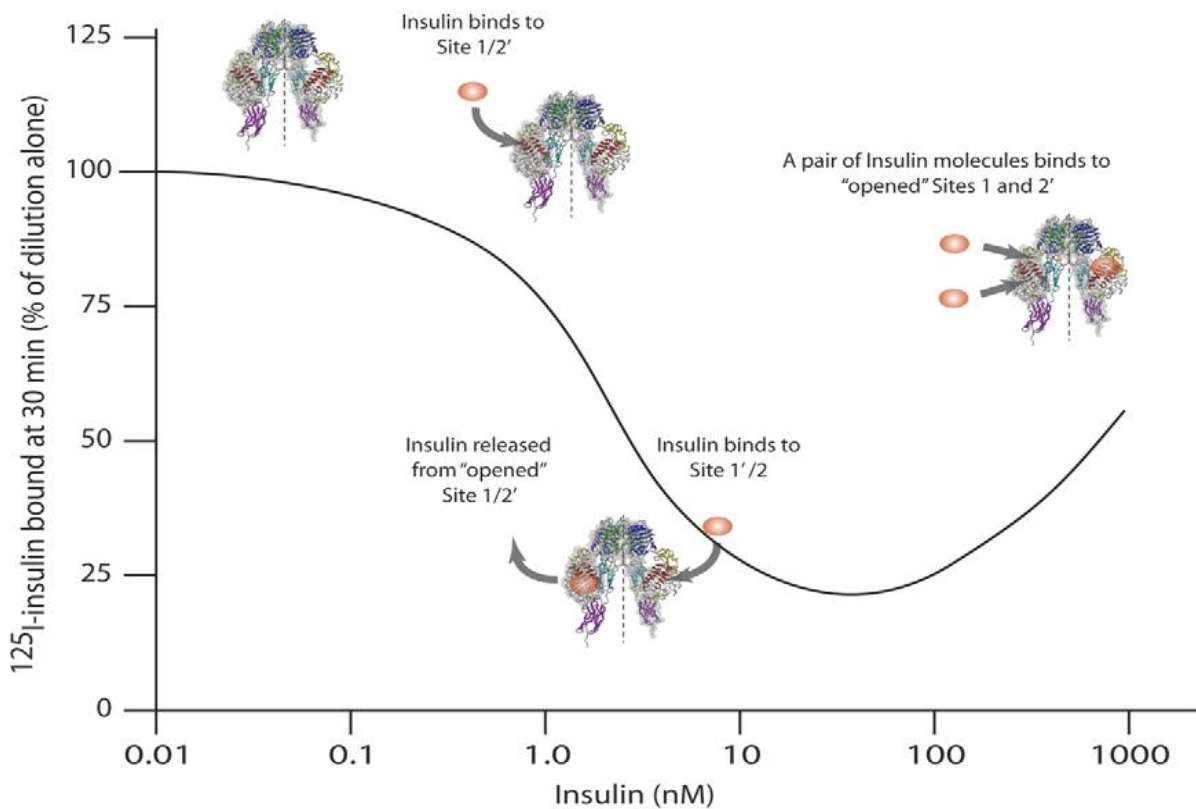
(B) Crystal structure of the insulin receptor ectodomain (PDB ID:2DTG). One receptor ectodomain monomer is represented as a ribbon cartoon, the other monomer is depicted in surface mode. The domains in both monomers are colour coded as per the scheme depicted in (A). Grey spheres indicate the observed N and C-termini of the  $\alpha$  and  $\beta$ -chains, respectively. The intervening ID-  $\alpha$  residues (656-719) & ID- $\beta$  residues (724-754) were not resolved. Figure adapted from Smith *et al.* (2010).



## 1.8 Characteristics and mechanism of ligand binding to the IR and IGF-IR

The binding of ligands to the IR and IGF-IR exhibit complex kinetic characteristics with curvilinear Scatchard plots indicating these receptors have low and high-affinity binding sites (De Meyts *et al.*, 2009). The two different ligand binding sites on each  $\alpha$ -subunit of the IR and IGF-IR are referred to as Site 1 & 2 on one monomer and Site 1' & 2' on the other monomer. Site 1 in the IR is composed of elements from two distinct regions, the central  $\beta$ -sheet of the L1 domain and the last 16 amino acids of the CT peptide. Site 2 is likely to be comprised of loops at the FnIII-1 and FnIII-2 domain junction in the opposite monomer to that which contributes the L1 domain (Ward *et al.*, 2008); (Smith *et al.*, 2010). The binding sites of the IGF-IR are generally equivalent to the IR (Adams *et al.*, 2000). However, in contrast to the IR, the central modules of the cysteine-rich region also contribute to Site 1 in the IGF-IR and are important for IGF-I binding but not IGF-II binding suggesting IGF-I and IGF-II have different mechanisms of binding to the IGF-IR (Sorensen *et al.*, 2004); (Keyhanfar *et al.*, 2007); (Surinya *et al.*, 2008).

The accelerated dissociation of pre-bound tracer from the IR and IGF-IR in the presence of cold ligand indicates a mechanism of negative co-operativity occurs in these receptors whereby the binding of one ligand molecule decreases the affinity of ligand bound to the other binding sites (De Meyts & Whittaker 2002) (Figure 1.5). The dose response curve for the accelerated dissociation of insulin from the IR has a reverse bell shape, indicating there is a loss of dissociation acceleration at insulin concentrations above 100nM (Figure 1.5) (De Meyts 1994). In contrast, the dose response curve for IGF-I accelerated dissociation from the IGF-IR is sigmoidal indicating there is no loss of accelerated dissociation (Christoffersen *et al.*, 1994). The biological advantages of negative co-operativity are unclear but the duration of ligand binding to the insulin receptor may influence outcomes in branching signalling pathways and could be important for limiting the mitogenic effects of insulin (De Meyts *et al.*, 2004).

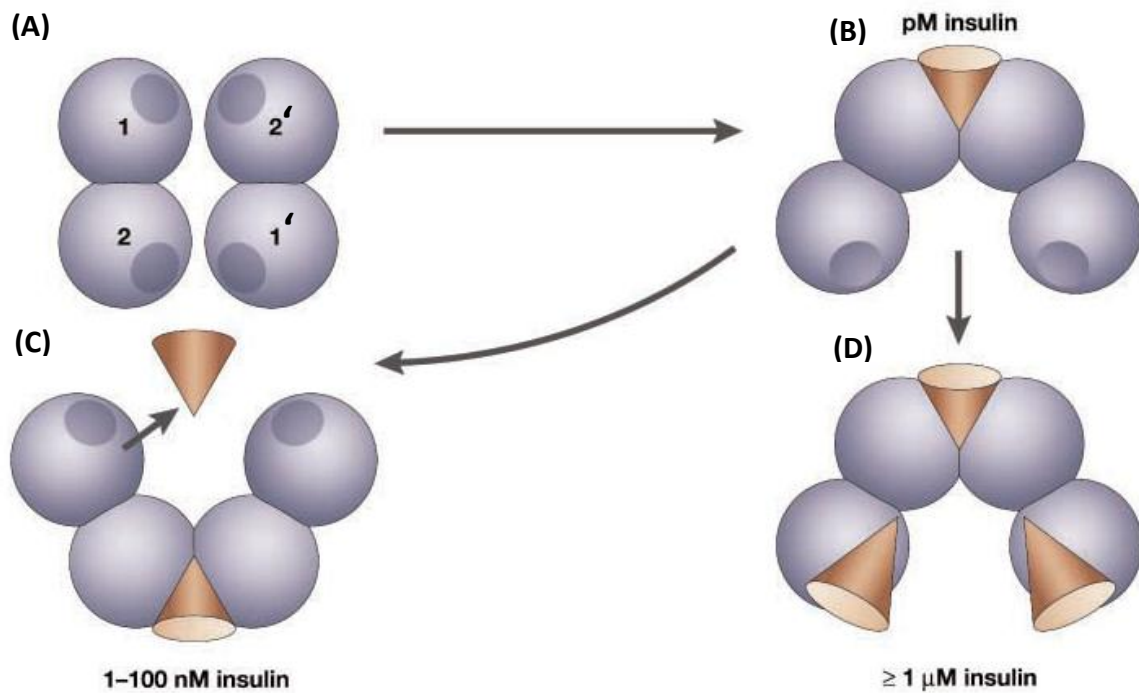


**Figure 1.5 Negative co-operativity of insulin binding to the insulin receptor**

The dose response curve of  $^{125}\text{I}$ -labelled insulin dissociation from the IR in a cell-based assay showing negative co-operativity. The mechanism underlying negative co-operativity of ligand binding is depicted with IR ectodomain structures overlaid on the curve. In the basal state, both low-affinity sites are equally accessible to ligand. An insulin molecule binds with high affinity and cross-links Sites 1 and 2' on one side of the dimer causing a conformational change that opens up the two monomers on the opposite side. On partial dissociation of the first bound insulin molecule, a second insulin molecule can crosslink the alternate Sites 1' and 2 causing the accelerated dissociation of the insulin molecule bound to Sites 1 and 2'. At high insulin concentrations, one insulin molecule cross-links a Site 1/2 pair whilst two additional insulin molecules bind to the remaining two binding sites and stabilise binding of the first bound insulin molecule. This binding mechanism accounts for the bell-shaped curve of insulin negative co-operativity. Figure adapted from Lawrence *et al.* (2007) and De Meyts & Whittaker (2002).

A cross-linking model was proposed to account for the complex binding kinetics of the IR whereby two different binding surfaces on the insulin molecule crosslink two different binding sites on the IR explaining the 1:2 stoichiometry and the presence of low and high-affinity binding sites (Schaffer 1994). Although this model accounted for much of the accumulated biochemical data associated with insulin binding to the IR, it did not explain the acceleration of tracer dissociation by cold ligand or the bell-shaped curve for negative co-operativity.

An alternative cross-linking model was postulated to fit all aspects of negative co-operativity in which the receptor binding sites 1 and 2 were oriented in an antiparallel symmetry (De Meyts 1994) (Figure 1.6). In this model, an insulin molecule binds to Sites 1 and 2' opening up Sites 1' and 2. Upon partial dissociation of the bound insulin molecule, a second insulin molecule bridging Sites 1' & 2 of the IR accelerates dissociation of the originally bound insulin molecule. At very high concentrations of insulin, monovalent binding of two additional insulin molecules saturates Sites 1' and 2 thereby stabilising the originally bound insulin molecule explaining the bell-shaped curve for negative co-operativity. Although IGF-I binding to the IGF-IR also displays negative co-operativity, the absence of accelerated dissociation of IGF-I from the IGF-IR is probably due to steric hindrance from the larger IGF-I molecule preventing simultaneous binding of two IGF-I ligands to the second set of binding sites in the IGF-IR (De Meyts 2008). Recently, mathematical models based on a harmonic oscillator have been proposed for the allosteric binding and activation of the IR and IGF-IR. These models accurately reproduce the equilibrium and kinetic properties of both receptors and may be adaptable for other receptors where ligand cross-linking occurs (Kiselyov *et al.*, 2009).



**Figure 1.6 Model of bivalent insulin cross-linking to the insulin receptor**

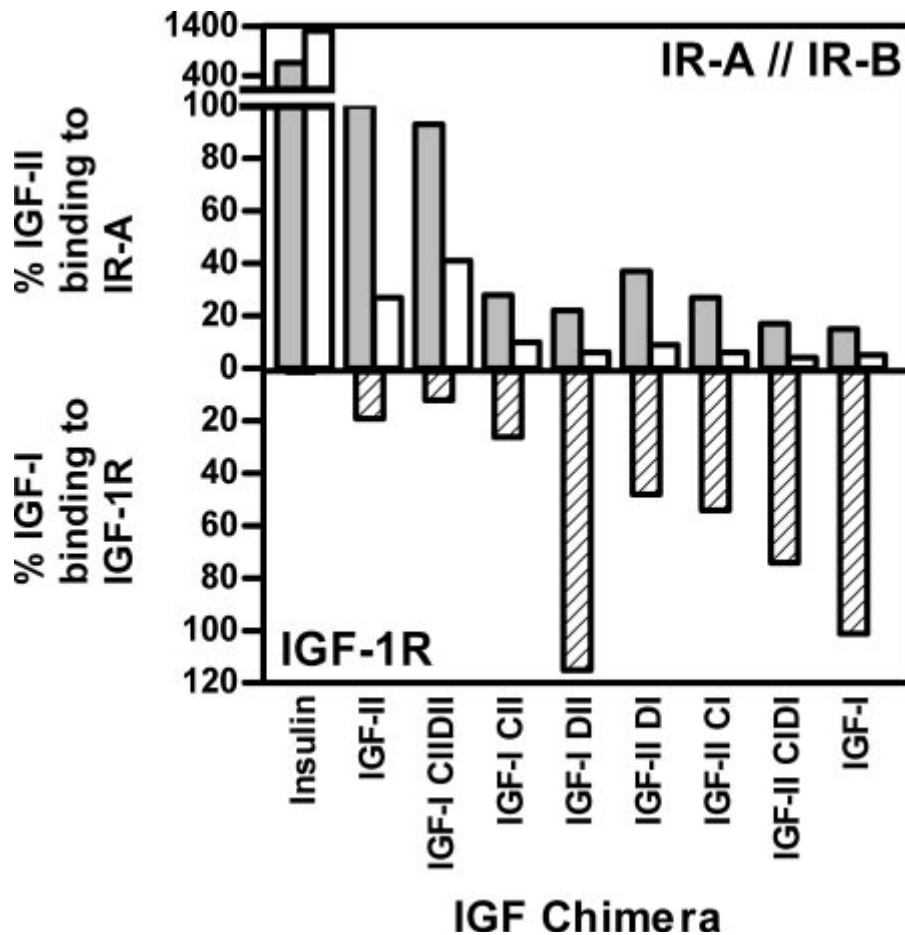
Cross-linking model of insulin binding to the insulin receptor as proposed by (De Meyts 1994). **(A)** The ligand binding Sites 1/1' and 2/2' on each insulin receptor monomer are depicted in a symmetrical antiparallel arrangement. **(B)** An insulin molecule (represented as a cone) binds with high affinity to Sites 1 and 2'. **(C)** Partial dissociation of the bound insulin molecule permits cross-linking of a second insulin molecule to the alternate binding Sites 1' and 2 accelerating dissociation of the originally bound insulin molecule. **(D)** At high concentrations of insulin, two additional insulin molecules saturate Sites 1' and 2 and stabilize the binding of the originally bound insulin molecule accounting for the bell-shaped negative co-operativity curve. Figure adapted from De Meyts & Whittaker (2002).

### **1.8.1 IGF receptor binding sites are primarily located in the B and A-domains**

The binding interaction between insulin and the insulin receptor has been extensively studied. Mutagenesis studies of the insulin ligand have confirmed the presence of two binding sites with receptor binding surface 1 consisting of residues GlyA1, IleA2, ValA3, GlnA5, ThrA8, TyrA19, AsnA21, ValB12, TyrB16, GlyB23, PheB24, PheB25 and TyrB26 whilst surface 2 involves residues SerA12, LeuA13, GluA17, HisB10, GluB13 and LeuB17 (Ward & Lawrence 2009) (Figure 1.2B). Although not as well characterised as insulin, the IGF-I and IGF-II ligands also have two binding sites comprised of residues that are primarily located in the B- and A-domains of the ligands and are generally comparable to the binding sites in insulin (Figure 1.2B). IGF-I binding surface 1 residues Phe23, Tyr24, Val44, Tyr60 and Ala62 have structural counterparts in insulin binding surface 1. In addition, IGF-I residues Ala8, Tyr31, Arg36, Arg37 and Met59 also interact with Site 1 of the IGF-IR. The second binding surface of IGF-I was identified by alanine scanning mutagenesis and consists of residues Glu9, Asp12, Leu54 and Glu58 (Gauguin *et al.*, 2008a). Site-directed mutagenesis studies have indicated IGF-II binding surface 1 consists of Val14, possibly Asp15, Gln18, Phe26, Tyr27, Phe28 and Val43 whereas binding surface 2 is comprised of Glu12, Phe19, Leu53 and Glu57 (Sakano *et al.*, 1991); (Alvino *et al.*, 2009).

### **1.8.2 Receptor binding specificity is determined by the IGF C and D-domains**

Despite their sequence and structural similarity, IGF-I and IGF-II bind with differential affinities to each of the IR-A, IR-B and IGF-IR (Denley *et al.*, 2005). The majority of free energy in the binding interaction between IGF and receptor is due to amino acids in the B and A-domains of IGF-I and IGF-II (Alvino *et al.*, 2009). However, a study of six IGF-I/IGF-II chimeras with individual and tandem exchanges of the C and/or D-domains showed swapping of both domains conferred the IR-A, IR-B & IGF-IR binding specificity of the donor molecule upon the recipient molecule (Denley *et al.*, 2004) (Figure 1.7). This study indicated receptor binding specificity is primarily determined by the C-domains, and to a lesser extent, by the D-domains of the IGFs.



**Figure 1.7** Relative binding affinities of IGF-I/IGF-II chimeras with exchanged C and/or D-domains to the IR-A, IR-B and IGF-IR.

The affinities of all ligands for the IR-A (grey bars) and IR-B (white bars) relative to percent IGF-II binding are shown in the upper histogram. The affinities of all ligands for the IGF-IR (striped bars) relative to percent IGF-I binding are shown in the lower histogram. Abbreviations used are IGF-I CIIDII: IGF-I with the C- and D-domains from IGF-II; IGF-I CII: IGF-I with C-domain from IGF-II; IGF-I DII: IGF-I with D-domain from IGF-II; IGF-II DI: IGF-II with D-domain from IGF-I; IGF-II CI: IGF-II with C-domain from IGF-I, IGF-II CIDI: IGF-II with C- and D-domains from IGF-I. Figure adapted from Denley *et al.* (2004).

Individual swapping of the C- and D-domains resulted in intermediate binding affinities for all receptors with the exception of IGF-I DII (an IGF-I chimera with the IGF-II D-domain) which bound the IGF-IR with an affinity comparable to IGF-I (Denley *et al.*, 2004). This is consistent with the observation that deletion of the D-domain in IGF-I had little effect on IGF-IR binding (Bayne *et al.*, 1989). However, the D-domain appears to play a role in IGF-II binding to the IGF-IR as deletion of the D-domain in IGF-II resulted in an ~5-fold reduction in IGF-IR binding affinity (Roth *et al.*, 1991). Furthermore, an IGF-II chimera with the IGF-I D-domain (IGF-II DI) had an ~2.5-fold increase in IGF-IR binding affinity (Denley *et al.*, 2004). As mentioned previously, IGF-I DII had the same IGF-IR binding affinity as IGF-I. However, an IGF-I chimera with both IGF-II C and D-domains (IGF-I CIIDII) had an IGF-IR binding affinity comparable to IGF-II (Denley *et al.*, 2004). Together, these results suggested the C-domain solely determined if an IGF ligand utilised the IGF-I or IGF-II mode of binding to the IGF-IR.

#### **1.8.2.1 Differential receptor activation is determined by the IGF C-domain**

The IGF-II ligand is more potent than IGF-I at inducing autophosphorylation of specific tyrosine residues in the activation loop of the kinase domain of the IR-A (residues 960, 1158, 1162 & 1163) and IR-B (residues 972, 1170, 1174 & 1175). Furthermore, IGF-II elicits greater activation of the downstream IRS-1 and Akt/PKB signalling pathways relative to IGF-I. IGF-II is also more potent than IGF-I at stimulating migration and cell survival via the IR-A (Denley *et al.*, 2006). Analyses of IGF chimeras with exchanged C and/or D-domains indicated the IGF C-domain was the sole structural determinant responsible for the differential activation of the IR isoforms, IRS-1 & Akt signalling pathways, migration & cell survival responses (Denley *et al.*, 2006).

#### **1.8.3 IR and IGF-IR signal transduction**

The functions of the IGF-IR and IR are physiologically distinct although there is still a degree of overlap as the IGF-IR can mediate metabolic actions (Di Cola *et al.*, 1997) and the IR has been shown to regulate mitogenic responses (Kimura & Ogihara 1997). The IR and IGF-IR activate a complex network of signal transduction pathways which

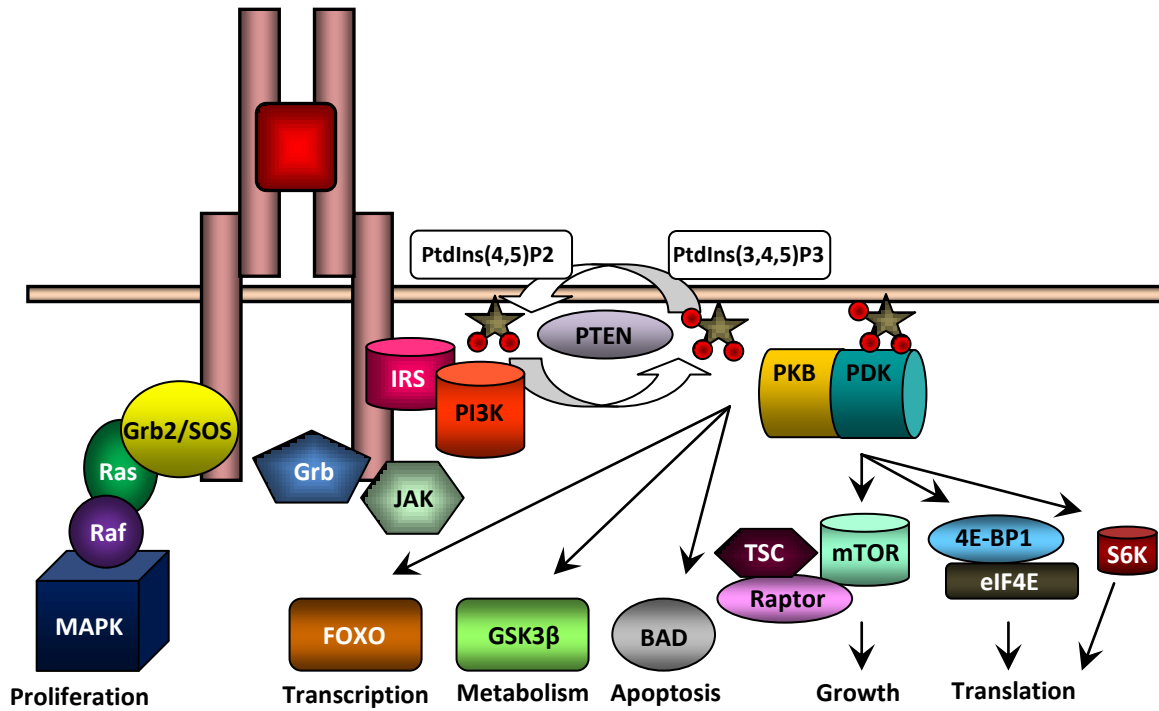
largely overlap (Kim & Accili 2002). The specificity of IGF-IR and IR signalling may be due to a number of factors such as structural variation within the receptor beta-subunits (Dupont & LeRoith 2001), differences in receptor number, tissue distribution, developmental regulation (Kim & Accili 2002) and subtle differences in ligand binding interactions between the two receptors such as ligand binding duration (Foulstone *et al.*, 2005). The primary signal transduction pathways discussed herein are pertinent to both the IR and IGF-IR.

In contrast to other receptor kinases, signalling by the IR and IGF-IR depends on the phosphorylation of intracellular substrates rather than autophosphorylation alone for recruitment of downstream signalling molecules (Siddle *et al.*, 2001).

Ligand induced activation of the IR and IGF-IR occurs via intramolecular *trans*-autophosphorylation of the activation loop increasing catalytic activity of the tyrosine kinase domain which provides docking sites for ATP and several intracellular substrates to bind. The major substrates are the insulin receptor substrates 1-4 (IRS1-4) and Shc (Src homology / collagen) proteins which trigger numerous signalling pathways including the MAPK and PI3K signalling pathways (Figure 1.8) (Chitnis *et al.*, 2008).

Phosphorylation of IRS recruits and activates class 1a phosphoinositide 3-kinase (PI3K) leading to synthesis of membrane associated  $\text{PtdIns}(3,4,5)\text{P}_3$  which serves as an allosteric regulator of phosphoinositide-dependent kinases (PDKs). PDK phosphorylates and activates protein kinase B (Akt/PKB) and p70rsk. PKB substrates include forkhead transcription factors (FOXO), glycogen synthase kinase 3 $\beta$  (GSK3 $\beta$ ) and the mammalian target of rapamycin (mTOR). These targets regulate gene expression, metabolism, cell survival and growth (Foulstone *et al.*, 2005).





**Figure 1.8 Insulin/IGF-I receptor signal transduction pathways**

Ligand binding to the IGF-IR and IR induces conformational change resulting in phosphorylation of the receptor tyrosine kinase domain and recruitment of the IRS and Shc substrates to the receptor. Binding of these substrates initiates signalling cascades which activate the PI3K and MAPK pathways regulating cell proliferation, apoptosis, growth, transcription, translation and metabolism. Figure adapted from Foulstone *et al.* (2005).

A second signalling pathway involves binding of the SH2 domain of the adapter Grb2 to phosphorylated IRS proteins leading to activation of the guanine nucleotide exchange factor SOS, Ras, Raf and the extracellular-signal-regulated kinase/mitogen activated protein kinase (ERK/MAPK). The ERK/MAPK pathway phosphorylates numerous cytoplasmic and nuclear substrates and is involved in anti-apoptotic signalling and transcriptional regulation (Saltiel & Pessin 2002).

Dysregulation of the IGF signalling axis has been strongly implicated in critical aspects of tumour formation and progression such as tumour cell proliferation, inhibition of apoptosis, angiogenesis, invasiveness and chemotherapy resistance (Breuhahn *et al.*, 2006)

## **1.9 The IGF system and cancer**

Epidemiological and laboratory studies have provided strong evidence implicating the IGF system in numerous different cancers including breast cancer, colorectal cancer, liver cancer, prostate cancer, pancreatic cancer, multiple myeloma, melanoma, glioblastoma, mesothelioma and childhood cancers (Sachdev & Yee 2007); (Werner & Bruchim 2009). Epidemiological studies have linked high circulating levels of IGF ligands and polymorphism in relevant genes to cancer risk and prognosis (Pollak 2008a). For example, a comprehensive meta-analysis of numerous case-control studies indicated high circulatory concentrations of IGF-I were associated with increased risk of prostate, colorectal and premenopausal breast cancer (Renehan *et al.*, 2004). Increased local expression of IGF-I and/or IGF-II have been consistently correlated with tumour progression in some cancers such as colon and pancreatic carcinomas although there have been inconsistencies with other types of cancer such as breast carcinoma (Samani *et al.*, 2007). Inconsistencies among studies relating to IGF levels and cancer risk could be due to technical issues associated with measurement methodology or may be due to biological factors whereby IGF levels are only related to risk in specific subsets of patients and may be influenced by variable modifying factors (Pollak 2008b).

### **1.9.1 IGF ligands and the IGF-IR in cancer**

The IGF-IR plays an important role in oncogenic transformation as shown by studies with IGF-IR deleted murine fibroblasts (R<sup>-</sup>) which were resistant to transformation by the SV40 large T-antigen and activated *ras* proto-oncogenes (Sell *et al.*, 1993); (Sell *et al.*, 1994). Unlike the EGF family of receptors, activation of the IGF-IR is ligand dependent and is not activated by receptor overexpression (Kaleko *et al.*, 1990). Early *in-vitro* experiments demonstrated proliferation of breast cancer cells was enhanced by the addition of IGF-I and IGF-II in a dose-dependent manner (Myal *et al.*, 1984). *In vivo* studies using murine cancer models indicated tumour growth and progression were accelerated with overexpression of IGF-I (Tanori *et al.*, 2010), IGF-II (Hassan & Howell 2000) and the IGF-IR (Lopez & Hanahan 2002). Small interfering RNA (siRNA) studies using human hepatocellular carcinoma cells indicated IGF-II was able to influence tumour cell viability, proliferation and anti-apoptosis through the IGF-IR (Nussbaum *et al.*, 2008). Abrogation of IGF-IR signalling with IGF-IR antibodies reduced proliferation of numerous types of cancer cells including colorectal, breast, neuroblastoma and melanoma (Werner & LeRoith 1996).

### **1.9.2 IGF-II and the IR-A in cancer**

Autocrine production of IGF-II occurs in several cancer types such as breast (Sciacca *et al.*, 1999), colorectal (Lahm *et al.*, 1994) and thyroid (Vella *et al.*, 2002) carcinoma. The *Igf2* gene is the most overexpressed gene in colorectal cancer relative to normal colorectal mucosa (Zhang *et al.*, 1997b) and implies that autocrine and paracrine supply of IGF-II may confer a selective advantage within the tumour microenvironment (Pollak 2008c). IGF-II overexpression can occur through a number of mechanisms including loss of imprinting of the maternal allele resulting in biallelic expression of *Igf2* (DeChiara *et al.*, 1991) and/or loss of heterozygosity of the gene for the IGF-IR which limits IGF-II bioavailability (Foulstone *et al.*, 2005).

A study of the SKUT-1 leiomyosarcoma cell line which predominately expresses the IR-A but not the IGF-IR had shown that the IR-A is able to exclusively mediate the biological effects of IGF-II (Sciacca *et al.*, 2002). The IR-A is predominately expressed in

a variety of cancers including breast, colon, lung, thyroid, ovary and skeletal muscle carcinomas (Belfiore *et al.*, 2009). The preferential expression of the IR-A and autocrine overexpression of IGF-II suggests a possible autocrine proliferative loop which may play an important role in tumour progression and aggressiveness.

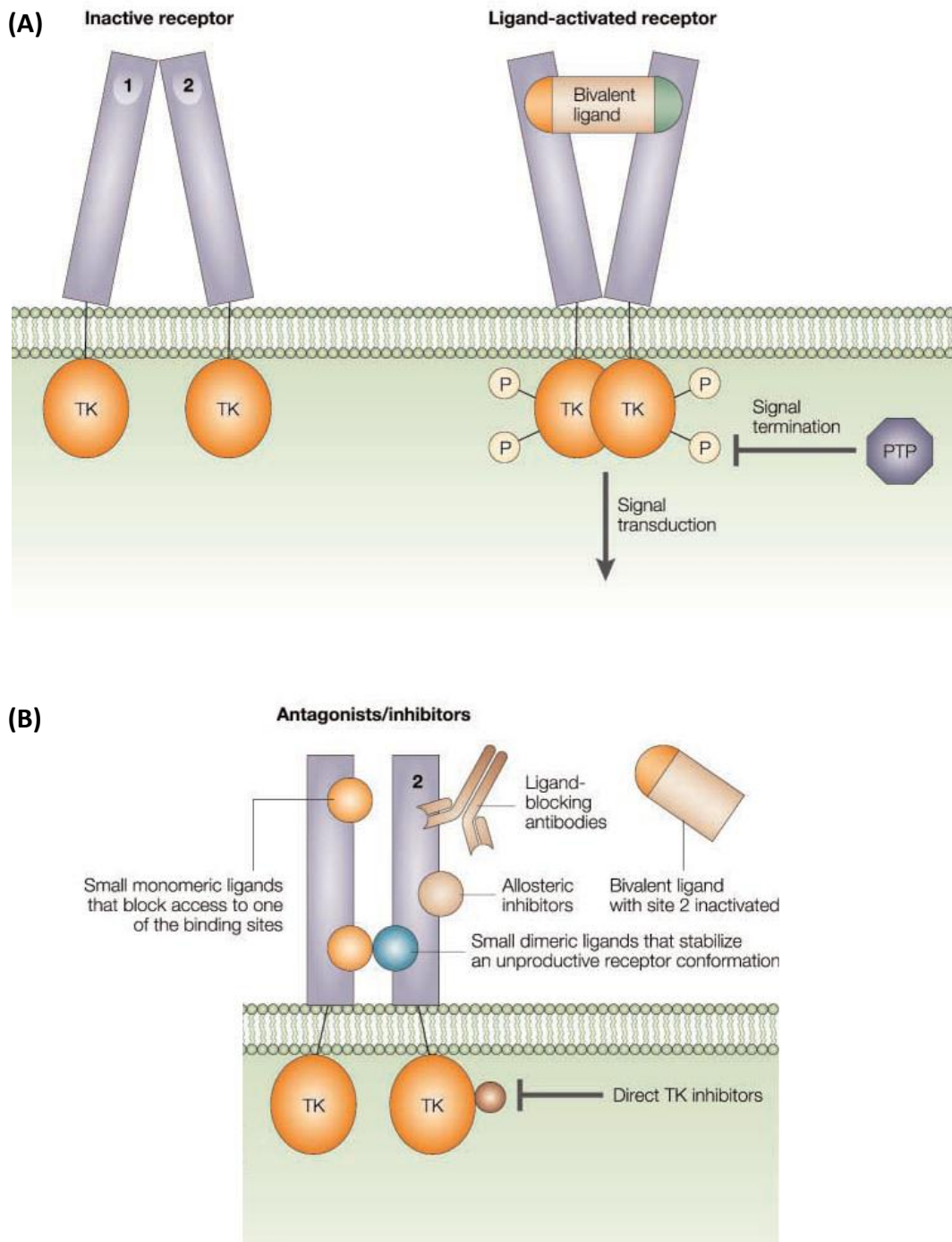
### **1.9.3 Targeting the IGF system for use as potential cancer therapeutics**

The role of the IGF axis in cancer presents a number of potential molecular targets for cancer therapy and has been the subject of considerable research interest, particularly as the cytotoxicity of many conventional treatments are enhanced when used in combination with therapies which disrupt IGF-IR signalling (Pollak 2008a). There are two primary approaches for generating IGF-based cancer therapeutics. Firstly, by reducing IGF ligand levels or bioactivity and secondly, by inhibiting receptor function (Figure 1.9).

#### **1.9.3.1 Strategies to reduce IGF levels**

##### **1.9.3.1.1 Antagonists of the growth hormone receptor reduce IGF-I levels**

Synthesis of IGF-I in the liver is regulated by growth hormone (GH). A polyethylene glycol modified analogue of GH (pegvisomant) acts as a GH receptor antagonist and has been successfully used to reduce IGF-I levels in normal individuals (Veldhuis *et al.*, 2001) and patients with acromegaly (Paisley *et al.*, 2004). Pegvisomant has been shown to inhibit growth of tumours in nude mice xenografted with colorectal cancer cell lines (Dagnaes-Hansen *et al.*, 2004), breast cancer cell lines (Divisova *et al.*, 2006) and primary human meningioma tumours (McCutcheon *et al.*, 2001).



**Figure 1.9 Strategies for generating IGF-IR and IR antagonists.**

**A)** Mechanism of ligand cross-linking activation of the IGF-IR and IR. **B)** Various strategies for the generation of receptor antagonists. PTP: protein tyrosine phosphatase; TK: tyrosine kinase domain. Figure adapted from De Meyts & Whittaker (2002).

#### 1.9.3.1.2 IGF neutralising antibodies

A monoclonal antibody (KM1468) directed against IGF-I and IGF-II suppressed tumour development in metastatic human prostate cancer in bone (Goya *et al.*, 2004) and influenced survival in the late phase of hepatic metastasis of colorectal cancer (Miyamoto *et al.*, 2005). Although not addressed in these studies, it is possible the reduction in free IGF-I levels could induce GH secretion which may directly stimulate cancer cell growth and cause insulin resistance and hyperinsulinemia (Belfiore *et al.*, 2009). Recently, a human monoclonal antibody (DX-2647) was shown to bind IGF-II and pro-IGF-II (IGF-IIe) with high affinity and blocked binding of IGF-II to the IGF-IR and IR-A (Dransfield *et al.*, 2010). This antibody inhibited cell proliferation, anchorage-dependent and anchorage-independent colony formation in a number of cell lines and slowed tumour progression in a hepatic xenograft model.

#### 1.9.3.1.3 Binding protein neutralization of IGF-I and IGF-II

IGF bioactivity is regulated by high affinity interactions with the IGFBPs in circulation and in the extracellular environment. IGFBP-1 has been shown to inhibit IGF-I activation of the IGF-IR and reduce breast cancer cell proliferation (Yee *et al.*, 1994) and anchorage-independent growth (Van den Berg *et al.*, 1997). IGFBPs also exert IGF-independent effects on proliferation, migration and apoptosis sensitivity (LeRoith & Roberts 2003) and these effects need to be accounted for to ensure maximum benefit of IGFBPs in cancer therapy (Rosenzweig & Atreya 2010).

#### 1.9.3.1.4 Binding of IGF ligands to soluble receptors

The secretion of soluble extracellular domains of transmembrane receptors is a mechanism employed in nature to modulate ligand activity (Jones *et al.*, 2008). Engineering of truncated soluble IGF-IR mutants have been shown to bind circulating IGF-I and inhibit metastasis of breast (Dunn *et al.*, 1998), lung (Lee *et al.*, 2003), colon (Adachi *et al.*, 2002) and pancreatic cancer cells (Min *et al.*, 2003).

### ***1.9.3.2 Strategies to inhibit receptor function***

#### ***1.9.3.2.1 Transcriptional targeting of the IGF-IR***

Antisense oligonucleotides target the receptor mRNA and form heteroduplexes which inhibit translation and/or cause mRNA degradation (Sachdev & Yee 2007). Antisense oligonucleotides targeting IGF-IR mRNA reduced IGF-IR levels and signalling pathways in breast (Neuenschwander *et al.*, 1995), prostate (Hellowell *et al.*, 2002), lung, bladder and CNS cancers (Werner & Bruchim 2009). However, the use of phosphorothioate antisense oligonucleotides has been associated with limited efficacy and nonspecific toxicity (Rochester *et al.*, 2005) and require optimisation to improve *in vivo* specificity and delivery efficiency of IGF-IR antisense reagents to tumours (Sachdev & Yee 2007).

Silencing of the IGF-IR gene with small interfering RNA (siRNA) reduced IGF-IR levels in melanoma, breast, prostate, ovarian and lung cancer cell lines (Bohula *et al.*, 2003). siRNA-mediated attenuation of IGF-IR signalling has been demonstrated in breast (Bohula *et al.*, 2003) and prostate cancer cell lines (Rochester *et al.*, 2005).

#### ***1.9.3.2.2 Receptor targeting antibodies***

Several antibodies against the IGF-IR have been shown to inhibit IGF-IR activation and induce receptor internalization and degradation (Pollak 2008a). For example, a mouse monoclonal antibody which binds the IGF-IR  $\alpha$ -subunit inhibited growth in numerous types of cancer such as breast (Arteaga & Osborne 1989), Ewing sarcoma (Scotlandi *et al.*, 1998), rhabdomyosarcoma (Kalebic *et al.*, 1994) and non-small cell lung cancer (Zia *et al.*, 1996). An antibody directed against the IGF-IR may also affect IR signalling due to the high degree of homology between the IGF-IR and IR. Disruption of IR-B signalling could potentially lead to detrimental metabolic effects or conversely, IR-A disruption might be advantageous if the antibody could be targeted to tumours (Werner & Bruchim 2009).

#### 1.9.3.2.3 Receptor tyrosine kinase inhibitors

Several small molecule inhibitors of the IGF-IR tyrosine kinase have been shown to inhibit tumor growth in numerous cancers (Foulstone *et al.*, 2005). Tyrosine kinase inhibitors function by binding to the substrate-binding site or the ATP-binding site in the receptor kinase domain or by inhibiting binding of substrates to the activated receptor (Sachdev & Yee 2007). However, small molecule inhibitors also have specificity issues similar to those encountered with receptor-directed antibodies with concomitant disruption of insulin receptor signalling due to the high degree of homology between the IR and IGF-IR (Sachdev & Yee 2007).

#### 1.9.3.2.4 Engineering ligands to act as receptor antagonists

The high degree of binding specificity and our increasing understanding of the bivalent mechanism by which the IGFs bind to the IGF-IR and IR provide an excellent opportunity to generate receptor antagonists through rational design (De Meyts & Whittaker 2002). An engineered IGF ligand retaining high affinity for Site 1 but little or no affinity for Site 2 in the IGF-IR/IR would be expected to act as an antagonist by binding the receptor without cross-linking or activation.

A similar approach was used in the rational design of human GH receptor antagonists with single substitutions of glycine 120 in the second binding site of the GH ligand with arginine or lysine (Fuh *et al.*, 1992). The GH mutants bound the GH receptor but blocked receptor dimerization preventing receptor activation and inhibiting cell proliferation (Chen *et al.*, 1994). A directed evolution approach identified eight mutations in binding site 1 of GH which increased binding affinity to the GH receptor (Jones *et al.*, 2008). A GH analogue combining the aforementioned mutations in sites 1 and 2 resulted in a more potent antagonist of the human growth hormone receptor. The half-life of this analogue was further improved by conjugation with 4-5 molecules of polyethylene glycol to form the GH receptor antagonist, pegvisomant (Ross *et al.*, 2001). The development of pegvisomant demonstrated the potential of rational design and directed evolution approaches in engineering native ligands as receptor antagonists.



## 1.10 Directed evolution

Rational protein design is theoretically the most direct way to generate proteins with desirable characteristics, for example to produce receptor antagonists or improving the stability or function of an enzyme. However, rational protein design is an enormous challenge and requires a detailed understanding of the relationship between a protein's amino acid sequence and function (Bloom *et al.*, 2005). Understanding how function is derived from sequence requires extensive knowledge of protein tertiary structure, protein folding and an understanding of complex molecular interactions such as hydrogen bonding, Van der Waals forces as well as electrostatic, hydrophobic and ionic interactions (Arnold *et al.*, 2001). In most cases, our understanding of these factors is extremely limited. In nature, the diverse and often complex functions of proteins have resulted from adaptation to the selective pressures of evolution. Similarly, laboratory-based gene diversification and directed evolution can be used to modify and/or improve protein characteristics without the extensive knowledge required for rational design.

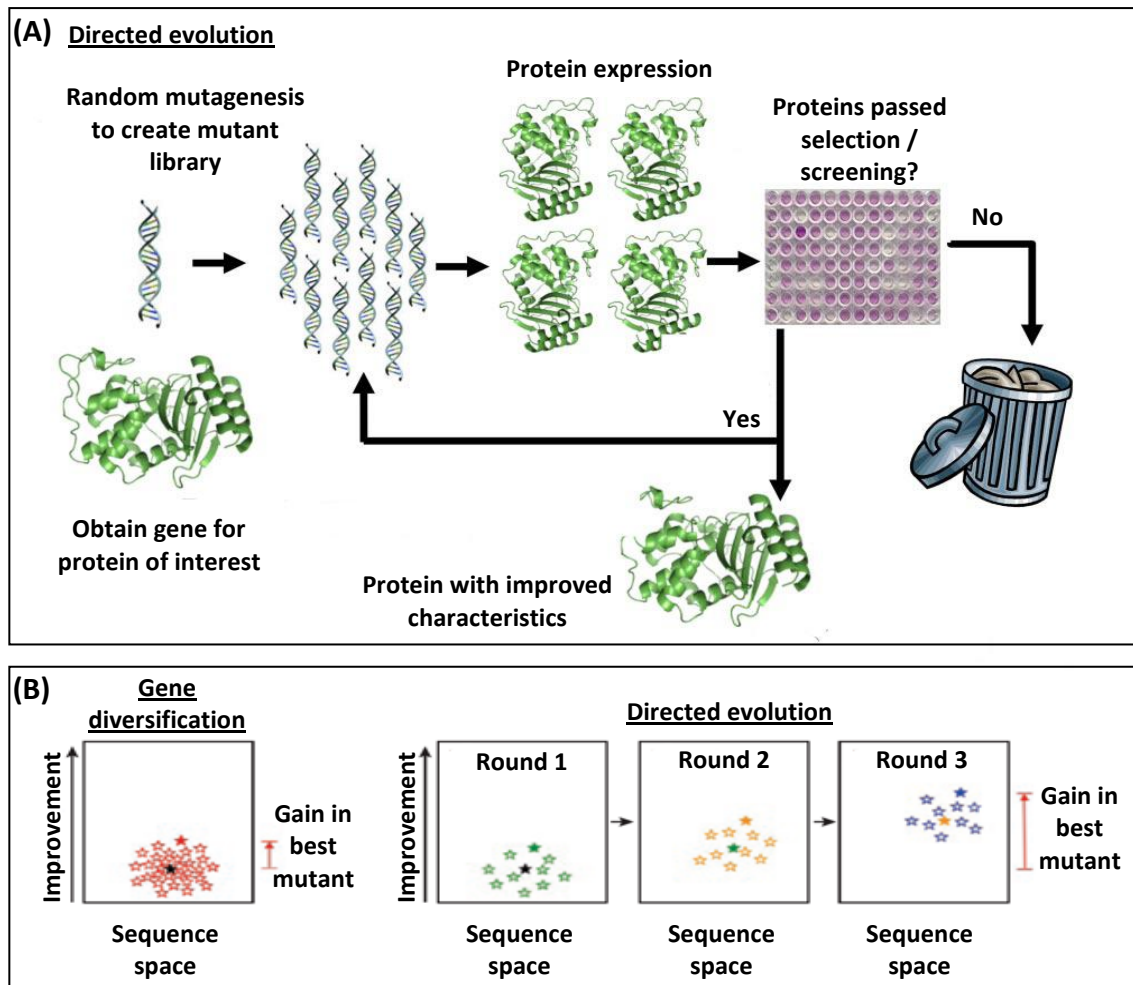
Gene diversification is a general term for the process of generating protein variants with improved or novel characteristics through a single round of mutation and selection in the laboratory. Directed evolution essentially refers to iterative rounds of gene diversification using the best variants from each round (Figure 1.10). Gene diversification and directed evolution have four fundamental requirements: Firstly, the characteristic has to be physically achievable; secondly, the characteristic has to be biologically or evolutionarily feasible; thirdly, libraries need to be constructed with sufficient complexity so as to contain variants with rare beneficial mutations; fourthly, a robust and rapid screen or selection for the desired characteristic needs to be available (Arnold 1998b). Directed evolution has been used to engineer numerous protein functions such as affinity, specificity, activity, stability and selectivity (Zhao 2007). The following sections will examine the four requirements of directed evolution and will summarise examples of *in vitro* and *in vivo* directed evolution techniques and their applications. Finally, the *Neurospora* directed evolution system utilised for much of the work presented in this thesis will be discussed in detail.

### 1.10.1 **Directed evolution requires protein function to be physically achievable and obtainable through evolution**

Protein functionality is dependent on numerous protein properties such as three dimensional structure, functional groups, structural rigidity or flexibility and interaction with other macromolecules. These properties are primarily determined by the amino acid sequence of the protein. Thus, evolutionary processes resulting in the substitution of one or more amino acids can potentially disrupt, modify or create novel functional properties in a given protein within the constraints of what is physically possible.

#### 1.10.1.1 **Sequence space**

Protein evolution has been described as a walk from one functional protein to another within a sequence space where all possible proteins exist as a connected network (Smith 1970). To put the sheer size of sequence space into perspective literally requires comparisons on an astronomical scale. For even a relatively small protein of 100 amino acids there are  $20^{100}$  ( $\sim 10^{130}$ ) possible sequence combinations which is greater than the estimated number of atoms in the known universe (Romero & Arnold 2009). Despite the enormity of sequence space there is a high degree of connectivity with each sequence having  $(A-1)N$  neighbours, where N is the number of residues and A is the total number of amino acid types (Voigt *et al.*, 2000). The majority of sequence space is devoid of function with the frequency of functional proteins in random sequence libraries postulated to occur at  $\sim 1$  in  $10^{11}$  (Keefe & Szostak 2001).



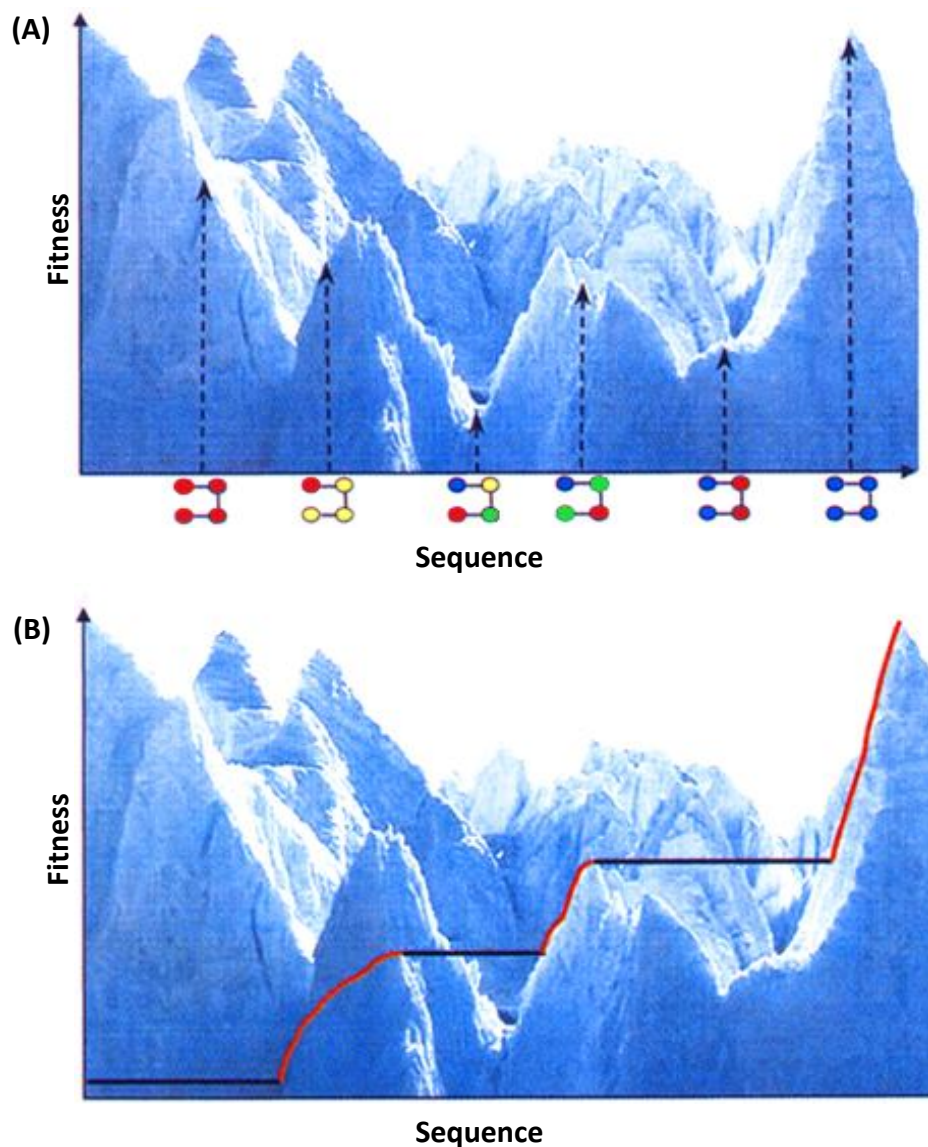
**Figure 1.10 Principles of directed evolution**

**A)** A typical format for a directed evolution experiment. The gene for the protein of interest is subjected to random mutagenesis to generate a library of mutant genes. Mutant proteins are produced and screened or selected for desired protein improvements. Improved proteins are used in the next round of mutagenesis and screening. Protein variants without improved characteristics are generally discarded. This process is repeated until mutants are obtained with the satisfactory levels of the improved characteristics. Figure adapted from Bloom & Arnold (2009). **B)** Comparison of likely functional gains in gene diversification and directed evolution experiments. Gene diversification identifies mutants with the most improved characteristics from a single library. As described in (A), directed evolution subjects the most improved proteins from gene diversification to additional rounds of diversification and screening to obtain mutants which have greater improvement in the desired characteristics and are farther away in sequence space from the original protein compared to the best mutant obtained from the gene diversification experiment. Sequence space is represented as black rectangles surrounding a protein of interest (black star). Coloured stars represent members of a mutant library characterised by similarity to the original protein in the sequence space (horizontal axis) and by improvement in the desired characteristics (vertical axis). Figure adapted from Yuen & Liu (2007).

### 1.10.1.2 Fitness landscapes

Assigning fitness to each possible protein variant within sequence space produces a theoretical fitness landscape with regions of higher elevation corresponding to increased fitness for a desirable characteristic (Figure 1.11A). Fitness landscapes contain various topological features representing sequence fitness. Mountain ranges are indicative of condensed regions of high fitness. Ridges connect ranges and small peaks are indicative of local fitness optima whereas the highest peak in the fitness landscape represents the globally fittest protein. Valleys represent regions of low fitness. An evolving protein can be envisioned as trekking through a fitness landscape with smooth uphill paths facilitating the rapid evolution of proteins whereas rougher terrains are more difficult for evolution to transverse and increase the probability of a protein becoming trapped on small peaks.

A trapped protein may require neutral mutations, simultaneous multiple beneficial mutations, recombination or even detrimental mutations to facilitate a jump to a new fitness peak (Figure 1.11B) (Dalby 2011). The overall density of functional sequences within a fitness landscape is low and uneven with functional proteins generally clustering together (Arnold 2006). The incorporation of a large number of mutations in a single step greatly increases the likelihood of a large jump that will trap a protein in a region of low fitness (Romero & Arnold 2009). Thus, directed evolution strategies usually incorporate a sequential series of single amino acid changes to move continuously uphill along networks of functional proteins (Yuan *et al.*, 2005) as single mutations often enable functional change or improvement (Tracewell & Arnold 2009). However, the effectiveness of sequentially accumulating individual beneficial substitutions (i.e. a hill-climbing approach) may be limited when epistatic coupling occurs whereby the potential benefit of a mutation requires the presence of neutral or even deleterious mutations at other positions in the protein sequence. Epistasis coupling between mutations creates rugged landscapes and a greater number of local optima in fitness landscapes (Bloom & Arnold 2009).



**Figure 1.11** Directed protein evolution through a fitness landscape in sequence space

**A)** A simplified representation of a fitness landscape within sequence space for a 4 amino acid protein with coloured circles representing different amino acid types. The highest fitness peak corresponding to the all blue sequence is the global optimum and the lower peaks represent local optima. **B)** Evolution of a protein through the fitness landscape to reach the global optimum proceeds through adaptive evolution to climb peaks (red lines) and neutral evolution to transverse local optima peaks (black lines). Figure adapted from Voigt *et al.* (2000).

### **1.10.1.3 Protein stability and the benefits of neutral mutations**

Folding to a thermodynamically stable three-dimensional structure is essential for a protein to function (Anfinsen 1973). Interestingly, the majority of natural proteins are only marginally stable at their physiologically relevant temperatures (Taverna & Goldstein 2002). Mutagenesis studies have shown mutations which improve function often decrease stability (Wang *et al.*, 2002) whereas mutations enhancing stability often decrease function which led to the hypothesis that there is a balance between protein stability and function (Shoichet *et al.*, 1995). The basis for the stability-function hypothesis is the notion that thermostability requires structural rigidity whereas enzymatic activity and protein interactions require molecular flexibility (Arnold 1998a). An alternative to the structure-function hypothesis which also accounts for the marginal stability of most proteins is the notion that evolution selects for the biochemical function of a protein once sufficient stability has been obtained for the protein to function and there is little or no selective pressure for any additional stability beyond this threshold (Bloom *et al.*, 2006).

Protein stability and functionality are relatively tolerant of mutation (Guo *et al.*, 2004) with 50-70% of random single mutations being neutral, 0.5 to 0.01% being beneficial and 30 to 50% deleterious (Bloom & Arnold 2009). There are at least two mechanisms by which neutral mutations can assist directed evolution. Firstly, functionally neutral mutations can increase protein stability and enhance tolerance for subsequent destabilizing mutations which are functionally beneficial. Secondly, neutral mutations can lead to changes in promiscuous functions, enabling new functions to evolve (Bloom & Arnold 2009). Unlike natural evolution, directed evolution experiments are usually constrained to moving uphill through a fitness landscape (Romero & Arnold 2009). Thus, the potentially beneficial impact of the long-term accumulation of neutral and even some slightly deleterious mutations are generally not realized unless they hitchhike with beneficial mutations in a selected protein (Bloom & Arnold 2009) or unless selection conditions are partially relaxed so as to accommodate neutral drift (Gupta & Tawfik 2008).

### **1.10.2 An overview of library construction for directed evolution**

Generating molecular diversity in a protein-encoding DNA library is the first step in a typical directed evolution experiment. The method utilised for gene diversification determines the quality of the mutant library and influences the efficacy of screening for the new or improved protein characteristic (Farinas *et al.*, 2001). Although increasing mutation rates will generate a more diverse library, there is an exponential decrease in the number of functional proteins due to the prevalence of deleterious mutations (Bloom *et al.*, 2005). Consequently, directed evolution experiments typically start with a DNA sequence coding for a protein with properties sufficiently close to the desired characteristics such that small improvements obtainable from low substitution rates can be identified in a high-throughput screen (Romero & Arnold 2009). The most common approaches for gene diversification are based on random mutagenesis and/or *in vitro* recombination.

#### **1.10.2.1 Random mutagenesis**

A random mutagenesis strategy requires no prior structural or mechanistic knowledge of the target protein and can generate unexpected beneficial mutations (Bloom *et al.*, 2005). For example, the catalytic efficiency of an aspartate aminotransferase was increased by the accumulation of mutations in residues outside of the active site of the enzyme (Shimotohno *et al.*, 2001). Random mutagenesis is most often achieved using some form of error prone PCR (epPCR) where the fidelity of replication is modulated by various factors such as divalent cation and dNTP concentration (Sen *et al.*, 2007). Fine tuning of epPCR experimental conditions enables the introduction of only one or two mutations in a given round of mutagenesis (Arnold *et al.*, 2001).

#### **1.10.2.2 In-vitro recombination**

*In vitro* recombination involves the fragmentation of a population of DNA sequences, for example, naturally occurring gene homologues, and reassembling them by PCR into a library of chimeric genes (Yuan *et al.*, 2005). Recombination facilitates combining beneficial mutations from multiple variants and removal of deleterious mutations

(Arnold *et al.*, 2001). Recombination can effectively overcome the inherent limitation in diversification with random mutagenesis whereby only low mutation rates can be utilised to avoid saturation with detrimental mutations. Consequently, recombination can give rise to dramatic increases in protein improvements (Yuan *et al.*, 2005).

### **1.10.3 An overview of screening and selection for directed evolution**

The greatest challenge in directed evolution is identifying the relatively small number of proteins with desirable properties within a vast population. Library variants are either screened individually in high-throughput assays or subjected to selective conditions where a protein with improved properties can be coupled to increasing the survival and growth of the host organism (Jackel *et al.*, 2008). In general, it is important to utilise sensitive and reproducible assays capable of detecting subtle improvements obtained from low levels of mutation (e.g. a two-fold increase in activity) so as to allow small steps through sequence space in each round of directed evolution (Romero & Arnold 2009). Most screens can facilitate analysis of  $10^2$  to  $10^4$  library members. If a selection approach can be utilised for the protein of interest then many more variants in a library can be analysed in a single experiment ( $>10^8$ ) compared to a typical screen (Arnold *et al.*, 2001).

#### **1.10.3.1 Display technologies**

The throughput capacity of a screen can be greatly improved through automation, miniaturisation and using display techniques such as phage display, cell surface display, mRNA display, ribosome display, and *in vitro* compartmentalization (Matsuura & Yomo 2006).



#### 1.10.3.1.1 Surface display

The most commonly used phage display technology utilises the M13 filamentous phage and generally involves fusion of exogenous proteins with the gene 3 (g3p) coat protein enabling the proteins to be presented on the surface of bacteriophage (Smith 1985). Phage display is typically used to screen for binding interactions with immobilized targets on solid surfaces but is generally not suited for complex extracellular eukaryotic proteins.

Cell surface display techniques involve fusion of exogenous proteins to integrated membrane proteins enabling them to be exported and anchored to the cell surface of various organisms such as *E. coli* and yeast (Georgiou *et al.*, 1997). One of the principle advantages of cell surface display methods is the ability to use fluorescence-activated cell sorting (FACS) for quantitative screening of up to  $\sim 10^5$  clones per second (Georgiou 2000). Bacteria have high transformation efficiencies ( $10^9$ - $10^{10}$  clones with *E. coli*) enabling diverse libraries to be generated and screened (Levin & Weiss 2006). Yeast have a lower transformation efficiency ( $10^6$ - $10^7$ ) compared to bacterial or phage systems but is better suited for the surface display of eukaryotic proteins (Boder & Wittrup 2000) (Matsuura & Yomo 2006).

#### 1.10.3.1.2 In vitro translation

The *in vivo* expression of the protein of interest is an inherent requirement for the phage and cell-surface display systems described above. However, heterologous expression of some proteins, particularly extracellular eukaryotic proteins, can be problematic at various levels including translation, folding, endomembrane trafficking and proteolytic degradation (Gouka *et al.*, 1997a). Numerous *in vitro* translation methods have been developed to overcome potential *in vivo* expression issues and to avoid limitations imposed by transformation efficiencies to produce very diverse libraries ( $>10^{14}$ ) (Matsuura & Yomo 2006).

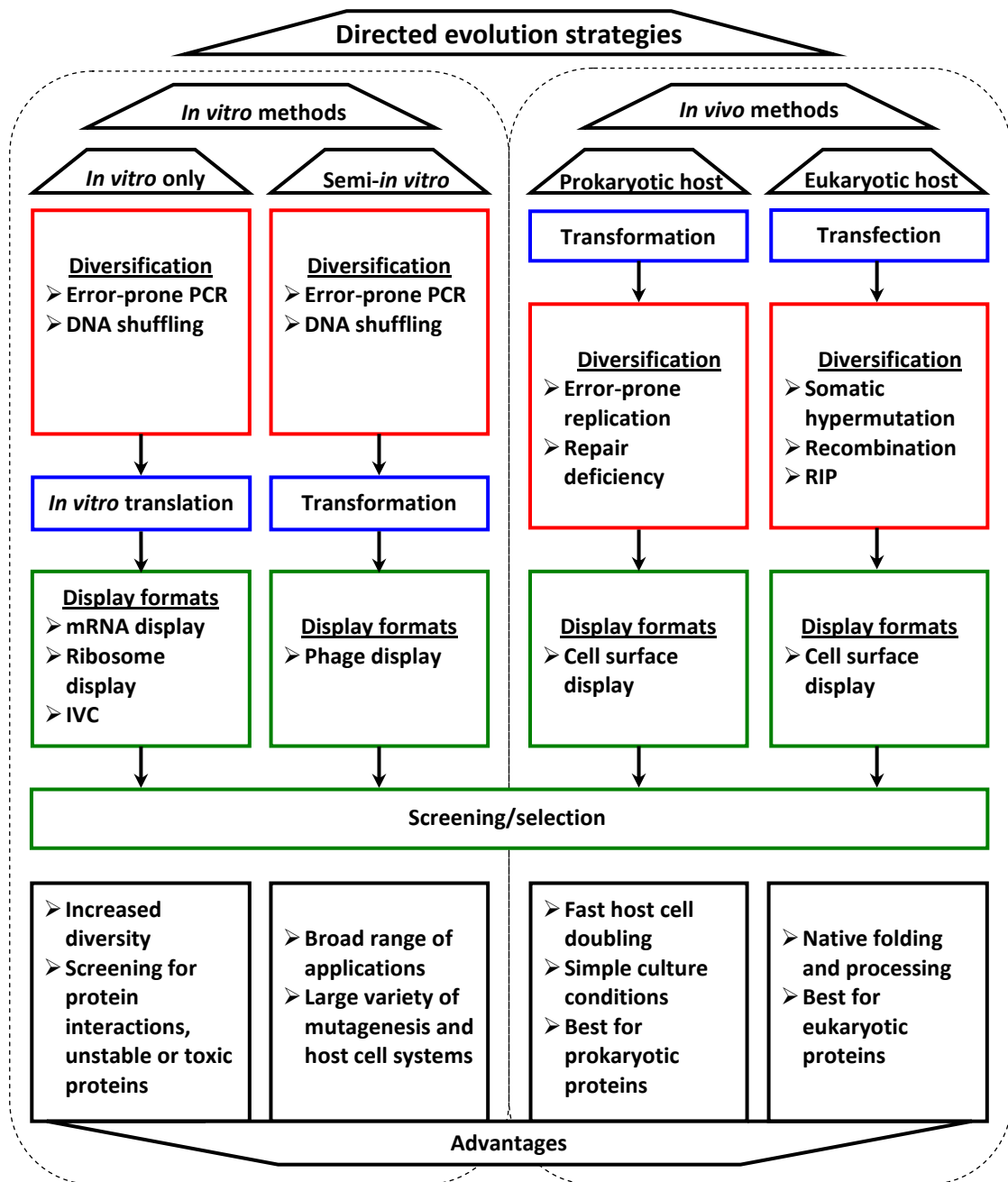
mRNA display involves the covalent fusion between mRNA and the encoded protein using the translation-terminating antibiotic, puromycin (an analogue of the aminoacyl end of tRNA). mRNA synthesized with puromycin at the 3' end is translated *in vitro*. When the puromycin enters the ribosome, a covalent bond is formed between the mRNA and the nascent protein chain thereby generating a stable linkage between the nucleic acid and the protein (Roberts & Szostak 1997).

Ribosome display links proteins and the encoding genes by eliminating stop codons at the end of the coding sequences thus capturing the nascent protein chains, ribosomes and mRNA complexes under stabilizing conditions. After screening, mRNA molecules are recovered by reverse transcription PCR (RT-PCR) (Hanes & Pluckthun 1997). mRNA display and ribosome display are primarily used for the directed evolution of protein binding interactions as selection is similar to phage display with ligands immobilized on solid surfaces (Matsuura & Yomo 2006).

*In vitro* compartmentalization (IVC) is a technique involving *in vitro* translation of a gene within the aqueous compartments of water-in-oil (w/o) emulsions thereby encapsulating the gene, transcripts and encoded protein together (Tawfik & Griffiths 1998). Re-emulsification of w/o droplets generates water-in-oil-in-water (w/o/w) emulsions. Protein variants embedded in water droplets containing fluorescent markers can be isolated in a continuous water phase by FACS enabling high-throughput screening (Bernath *et al.*, 2004). Microbeads displaying single genes and encoded proteins formed by IVC into w/o emulsions enabled direct *in vitro* selection for enzyme activity (Griffiths & Tawfik 2003).

#### **1.10.4 Directed evolution strategies**

There are a wide variety of *in vitro* and *in vivo* methods which have been utilised for directed evolution with each system having specific advantages and disadvantages (Figure 1.12). A selection of directed evolution techniques and examples of their applications are described in the following section.



**Figure 1.12 Comparison of *in vitro* and *in vivo* directed evolution strategies**

Schematic diagram depicting the main directed evolution strategies, advantages and potential display formats. IVC: *in vitro* compartmentalization. RIP: Repeat-Induced Point mutation. Figure adapted from Blagodatski & Katanaev (2011).

#### ***1.10.4.1 In vitro directed evolution techniques***

*In vitro* strategies are the most widely used for directed evolution and are typically based on epPCR and/or DNA shuffling strategies for generating diversified gene libraries which are transformed and expressed within a host organism, or used in an *in vitro* translation system prior to screening for properties of interest. There are a larger number of *in vitro* methods available compared to *in vivo* techniques and substantially larger library sizes can usually be obtained (Jackel *et al.*, 2008).

##### ***1.10.4.1.1 Error-prone PCR***

The most commonly used random mutagenesis method is error-prone PCR (Figure 1.13a) which is based on further reducing the inherently low fidelity of Taq DNA polymerase to incorporate random mutations into PCR amplicons. There are a number of ways to reduced DNA polymerase fidelity such as biasing the concentration of dNTPs, increasing magnesium concentration and adding manganese ions to the amplification reaction (Wilson & Keefe 2001).

A site-directed mutant of L-Arabinitol 4-dehydrogenase from *Neurospora crassa* was subjected to a single round of epPCR after rational design targets had been exhausted and generated a mutant with altered cofactor specificity (Bae *et al.*, 2010). epPCR was used in the directed evolution of a glutaryl acylase from *Pseudomonas* SY-77 into an adipyl acylase with improved activity for adipyl-7-aminodesacetoxycephalosporanic acid (Otten *et al.*, 2002).

##### ***1.10.4.1.2 DNA shuffling***

Recursive random mutagenesis can drive beneficial mutations into extinction but this can be avoided through recombination (Bloom *et al.*, 2005). DNA shuffling and derivative methods enable the accumulation of beneficial mutations by recombination and have been shown to be very successful for optimizing protein properties (Yuan *et al.*, 2005).

DNA shuffling (Figure 1.13B) typically involves enzymatic fragmentation of sequences from two or more homologous genes which are subjected to self-priming PCR. Homology between fragments facilitates priming with recombination occurring by template switching when fragments from one template prime on another homologous template. Non-complementary sequences form overhangs which are filled in during PCR. The resultant chimeras are reassembled into full-length shuffled genes with a final PCR step (Stemmer 1994a); (Stemmer 1994b); (Cramer *et al.*, 1998). A single parent gene format can be utilised with mutations introduced during the reassembly process by epPCR to generate a pool of variants with beneficial mutations that can then be shuffled together (Yuan *et al.*, 2005).

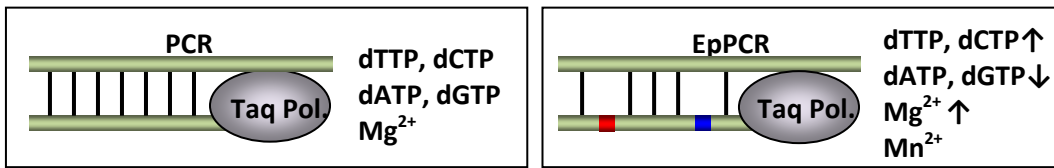
DNA shuffling has been used in the directed evolution of  $\beta$ -lactamase to increase resistance to cefotaxime (Stemmer 1994b), improve cephalosporinase activity (Cramer *et al.*, 1998), alter the substrate specificity of  $\beta$ -galactosidase (Zhang *et al.*, 1997a) and enhance the thermostability of glucose dehydrogenase (Baik *et al.*, 2003).

#### 1.10.4.1.3 StEP

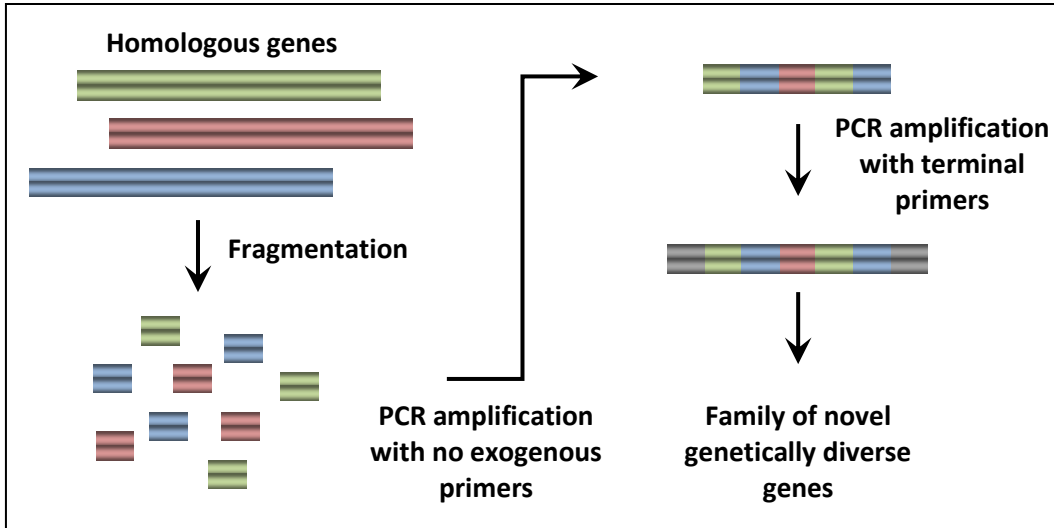
The staggered extension process (StEP) (Figure 1.13C) involves priming a set of template sequences and using very short PCR annealing and extension times to ensure extension is incomplete. Repeated cycling enables the growing fragments to anneal to different templates thereby promoting crossover events along the full length of the template sequences (Zhao *et al.*, 1998); (Zhao & Zha 2006).

StEP has been used to improve the thermostability of subtilisin E (Zhao & Arnold 1999), enhance class I enolpyruvylshikimate 3-phosphate synthase resistance to glyphosate (He *et al.*, 2001), and increase thermostability and catalytic activity of endoglucanase (Lee *et al.*, 2010).

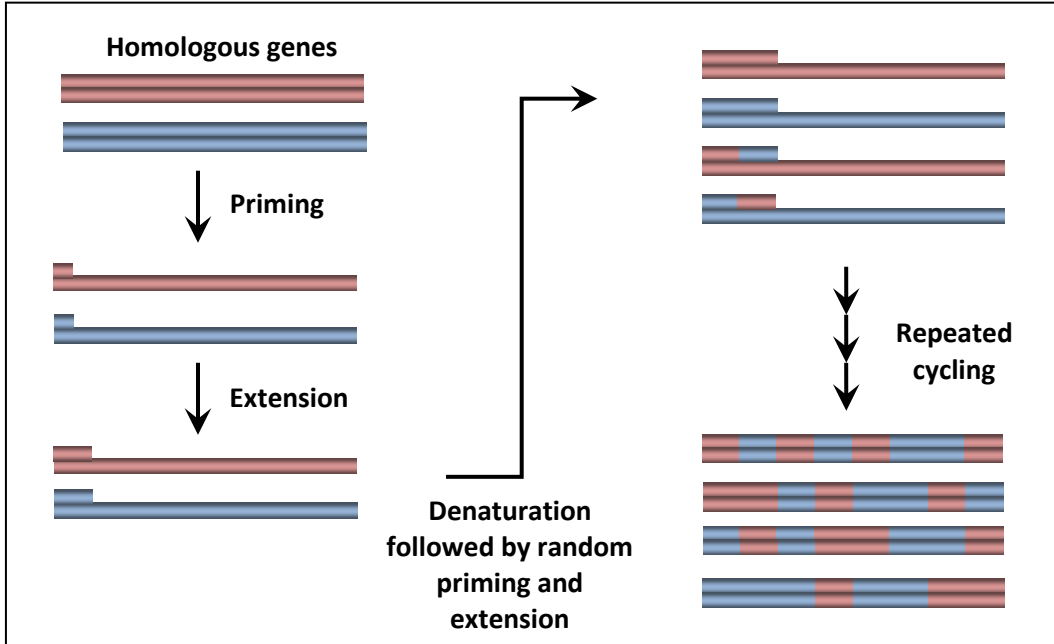
A) Error prone PCR

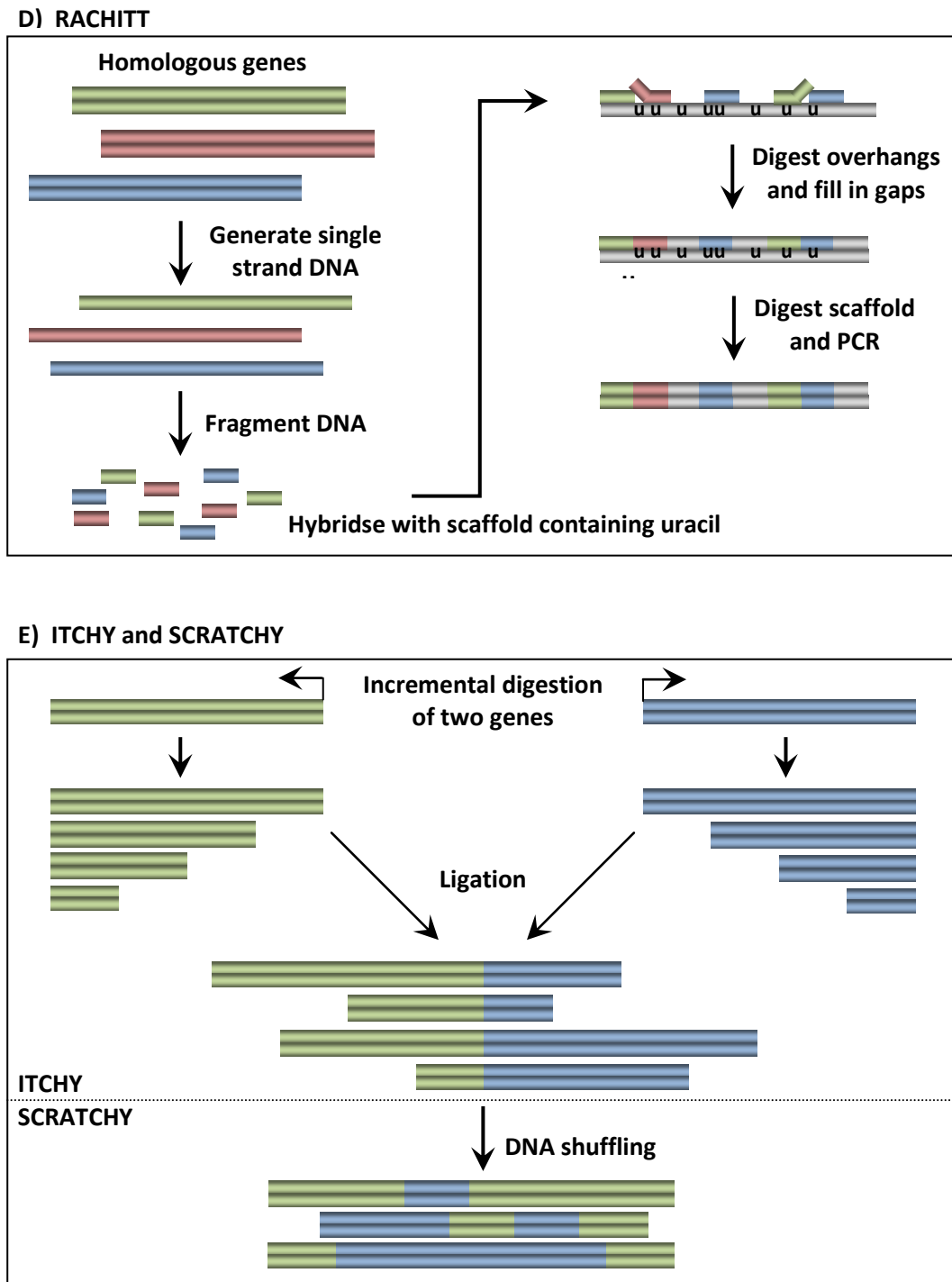


B) DNA shuffling



C) StEP





**Figure 1.13 Comparison of *in vitro* gene diversification techniques**

Schematic representation of *in vitro* strategies for gene diversification. Details of each method are described in the text. **(A)** Error prone PCR (refer section 1.10.4.1.1). **(B)** DNA shuffling (refer section 1.10.4.1.2). Figure adapted from Pelletier (2001). **(C)** Staggered extension process (refer section 1.10.4.1.3). Figure adapted from Zhao & Zha (2006). **(D)** Random chimeragenesis on transient templates (refer section 1.10.4.1.4). Figure adapted from Pelletier (2001). **(E)** Incremental truncation for the creation of hybrid enzymes & SCRATCHY (refer section 1.10.4.1.5). Figure adapted from Michnick & Arnold (1999).

#### 1.10.4.1.4 RACHITT

Random chimeragenesis on transient templates (RACHITT) (Figure 1.13D) generates chimeric genes with higher crossover frequencies than DNA shuffling or StEP (Coco *et al.*, 2002). A full length, single-stranded DNA sequence from one homologue is synthesised with uracil to serve as a transient scaffold template for the assembly of single stranded fragments from other homologues. Nuclease treatment removes unhybridised 5' and 3' fragment termini, a high fidelity DNA polymerase is used to fill gaps between hybridized fragments and nicks are repaired by ligation. The scaffold strand in the heteroduplex DNA is digested with uracil-DNA-glycosylase to ensure only the novel assembled strand is amplified by PCR to generate homoduplex chimeras (Coco *et al.*, 2001). Thus, RACHITT enables the generation of highly diverse libraries that do not contain the unshuffled parental gene which is particularly advantageous if the activity of the parental protein is problematic for library screening (Yuan *et al.*, 2005).

The RACHITT technology has been used to generate novel dibenzothiophene monooxygenases with increased reaction rates and altered substrate specificities to improve diesel biodesulfurization (Coco *et al.*, 2001).

#### 1.10.4.1.5 ITCHY and SCRATCHY

Techniques based on homologous recombination are inherently limited by a requirement for high sequence similarity (>60%) to facilitate recombination (Farinas *et al.*, 2001). Although variants generated by homologous recombination are more likely to be structurally similar to their parent proteins, non-homologous recombination has the potential to create new structural diversity and protein fold that may or may not be found in nature (Yuan *et al.*, 2005) which can be useful in elucidating structure-function relationships (Kawarasaki *et al.*, 2003). Consequently, numerous methods have been developed which recombine non-homologous sequences such as the incremental truncation for the creation of hybrid enzymes (ITCHY) technique (Figure 1.13E) (Ostermeier *et al.*, 1999). ITCHY is based on the creation of 5' and 3' fragment



libraries from the incremental truncation of two template sequences either by time-dependent exonuclease III treatment (Ostermeier *et al.*, 1999) or by the incorporation of  $\alpha$ -phosphothioate deoxynucleoside triphosphates (Lutz *et al.*, 2001a), followed by ligation of the 5' and 3' fragments to form a single crossover hybrid library. Although ITCHY negates the need for sequence homology to generate crossovers, the sequences are of variable length and may be ligated out of frame or in reverse orientation thereby generating nonsense products and impeding the search for functional variants (Yuan *et al.*, 2005). The addition of a DNA shuffling step to recombine ITCHY libraries facilitates the generation of chimeras with more than one crossover in a process termed SCRATCHY (Figure 1.13E) (Lutz *et al.*, 2001b); (Kawarasaki *et al.*, 2003); (Lutz & Ostermeier 2003).

Both ITCHY (Ostermeier *et al.*, 1999) and SCRATCHY (Lutz *et al.*, 2001b) have been used to form functional hybrids between human and *E. coli* ribonucleotide formyltransferases using parental genes with only 50% DNA identity. Similarly, SCRATCHY has been used to form functional fusions between rat theta 2 glutathione transferase and human theta 1 glutathione transferase which share 63% DNA identity (Kawarasaki *et al.*, 2003).

#### **1.10.4.2 In vivo directed evolution**

Most *in vitro* directed evolution experiments involve generating a library of diverse DNA sequences with each variant sequence typically needing to be cloned, transformed and isolated to enable expression and screening. In contrast, *in vivo* directed evolution techniques transform or transfect the host organism with the gene(s) of interest and exploit natural biological processes or defects thereof to generate sequence diversity. If the diversified gene is placed appropriately for expression and the transcribed protein can be correctly processed by the host cell, then the exogenous protein can be screened directly without the need for further manipulation. Thus, *in vivo* directed evolution methods have the advantage of generating sequence diversity and facilitating screening within a living cell culture (Blagodatski & Katanaev 2011).

1.10.4.2.1 *In vivo directed evolution in prokaryotes*

Prokaryotic *in vivo* directed evolution typically uses repair or replication-deficient bacterial strains (mutator strains) to generate sequence diversity. For example, the *E. coli* XL1-Red strain has three mutations (*mutD5*, *mutS* and *mutT*) making it deficient in three primary DNA repair pathways and exhibits an ~5,000-fold increase in random mutations (Greener *et al.*, 1996). A plasmid containing the target DNA sequence is transformed into the XL1-Red strain and mutations are accumulated with each round of DNA replication. Directed evolution with the XL1-Red *E. coli* strain had been used to increase activity of a mutant arylmalonate decarboxylase 10-fold (Terao *et al.*, 2006), enhance the activity of ADH  $\beta$ -glucuronidase at neutral pH (Callanan *et al.*, 2007) and alter the substrate specificity of a *Pseudomonas fluorescens* esterase (Bornscheuer *et al.*, 1999).

The disadvantages associated with most efficient mutator strains include slow growth and impaired transformation efficiency (Nguyen & Daugherty 2003). In addition, the susceptibility of the entire genome to indiscriminate and often deleterious mutation can decrease mutagenesis efficiency in prolonged culture and obscure phenotypic expression from mutations in the target gene (Greener *et al.*, 1997); (Long-McGie *et al.*, 2000). A system enabling random mutagenesis of targeted genes in an *E. coli* strain with a mutant low fidelity DNA polymerase I (Pol I) had been developed (Camps *et al.*, 2003); (Troll *et al.*, 2011). Pol I is predominately involved in replication of the ColE1 plasmid. Target sequences encoded in a Pol I-dependent plasmid exhibited an 80,000-fold increase in random mutation levels when cultures were maintained at stationary phase. Mutation extended at least 3kb from the plasmid origin of replication and generated a relatively balanced mutation spectrum (Camps *et al.*, 2003). This system has been used to generate variants of TEP-1  $\beta$ -lactamase capable of hydrolyzing the lactam antibiotic, aztreonam (Camps *et al.*, 2003).

#### 1.10.4.2.2 In vivo directed evolution in eukaryotes

The primary advantage of *in vivo* directed evolution in eukaryotes is the ability to diversify and express proteins within a eukaryotic cellular environment. This can be important for eukaryotic proteins affected by various intra-cellular parameters such as protein folding, post-translational modifications, pH, salt concentrations and degradation (Bulter *et al.*, 2003); (Arakawa *et al.*, 2008).

##### 1.10.4.2.2.1 *Somatic hypermutation*

Somatic hypermutation of the immunoglobulin (Ig) variable region occurs when activated B lymphocytes encounter antigen as part of the immune response and is greater than the spontaneous mutation rate by a factor of  $10^6$  (Rajewsky *et al.*, 1987). Immunoglobulin sequences are not required for hypermutation and retroviral delivery of an exogenous gene into the vicinity of the Ig locus is sufficient to access hypermutation sites (Wang *et al.*, 2004a). The directed evolution of a red fluorescent protein (RFP) in the RAMOS human B cell line with 23 rounds of FACS screening yielded variants with increased photostability and far-red emissions (Wang *et al.*, 2004b). However, targeting exogenous genes to the hypermutating *Ig* loci is difficult as integration generally occurs at random chromosomal sites (Wang *et al.*, 2004b). The directed evolution of an eGFP transgene by somatic hypermutation in the chicken DT40 B cell line was reported to have targeted integration into the Ig light chain locus in more than 70% of transfectants. Three rounds of FACS sorting were sufficient to identify eGFP variants with higher fluorescence activity (Arakawa *et al.*, 2008). One of the primary advantages of these systems is the ability to diversify and express exogenous genes in a mammalian cellular environment which could be critically important for some vertebrate proteins.

##### 1.10.4.2.2.2 *In vivo shuffling in yeast*

*In vivo* shuffling in the yeast, *Saccharomyces cerevisiae* utilises high frequencies of homologous recombination to enable fast and efficient construction of mutant libraries. In this system, mutant genes or DNA fragments are co-transformed with a

linearised vector with ends homologous to the ends of the fragment DNA. Homologous recombination between the open vector and DNA fragments regenerates a circular, replicating plasmid with chimeric exogenous genes that can be recovered from the vector or expressed within the yeast cell (Pompon & Nicolas 1989). *In vivo* shuffling has an additional advantage as it does not introduce additional mutations during recombination unlike most *in vitro* recombination techniques (Zumarraga *et al.*, 2008). *In vivo* shuffling in yeast has been used to improve a mutant fungal haem peroxidase with substantially increased thermal and oxidative stability (Cherry *et al.*, 1999).

### **1.10.5 Directed evolution in *Neurospora***

An *in vivo* DNA shuffling platform developed for the filamentous ascomycete, *Neurospora crassa* exploits high recombination frequencies for gene diversification. The *Neurospora* directed evolution system and underlying biology will be discussed in detail as it is the basis for much of the work described in this thesis.

#### **1.10.5.1 *Neurospora crassa* biology**

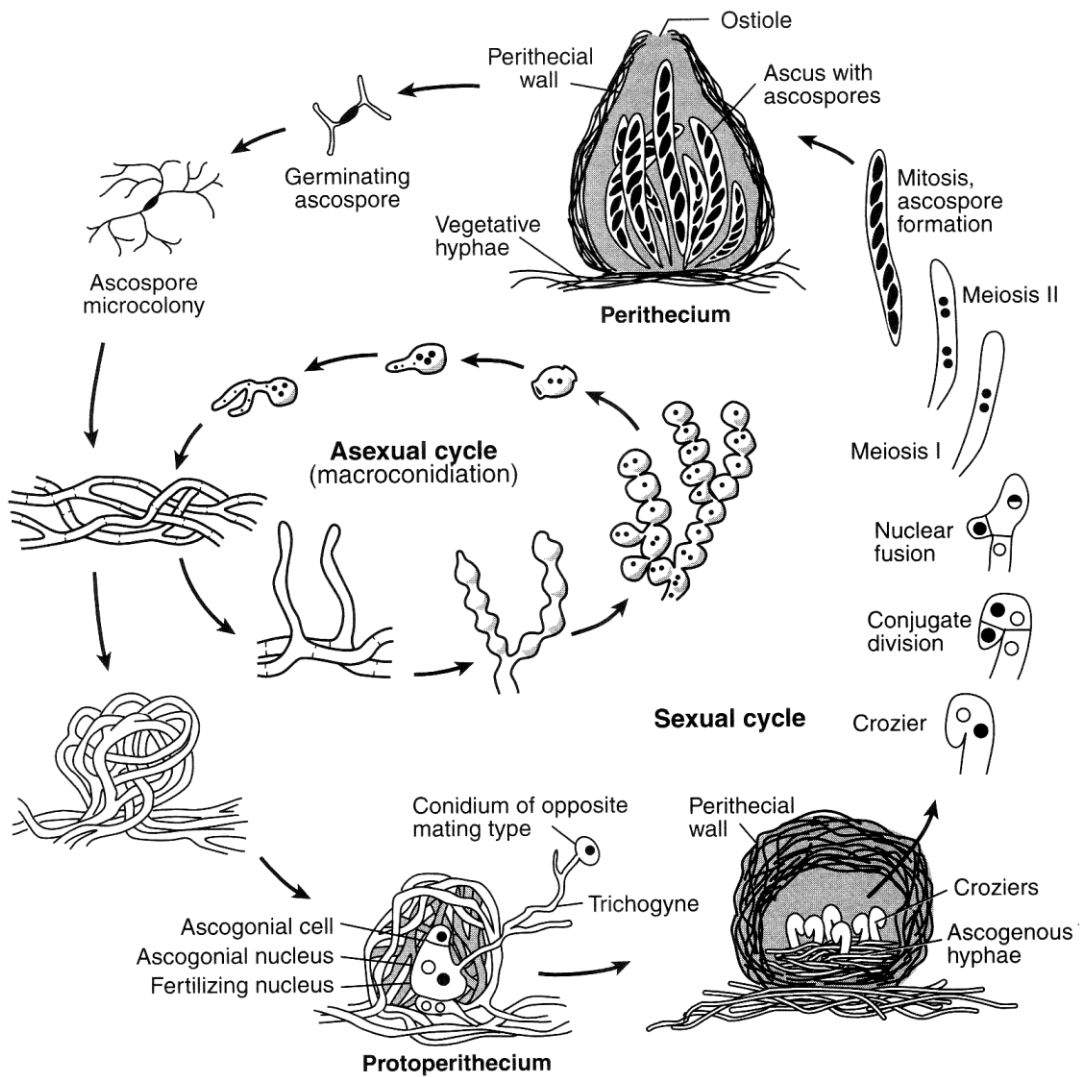
*Neurospora crassa* was developed as an experimental organism in the early twentieth century and by the latter half of the century had been developed into a highly characterised model organism for numerous biological processes such as recombination, DNA repair, genome defence, DNA methylation, gene silencing, metabolic regulation and circadian rhythms as well as cellular differentiation and development. The *N. crassa* genome is distributed among seven chromosomes (referred to as Linkage Groups (LG) I-VII) and totals ~43Mb. The genome was completely sequenced and made publicly available in 2003 (<http://www.broadinstitute.org/annotation/genome/neurospora/MultiHome.html>) and found to contain ~10,000 protein coding genes (Galagan *et al.*, 2003).

#### 1.10.5.2 The *Neurospora* lifecycle

The *Neurospora crassa* lifecycle consists of a haploid vegetative growth phase with sexual and asexual reproduction (Figure 1.14). The asexual phase consists of haploid spores which are either uninucleate (microconidia) or multinucleate (macroconidia). Germ tubes from germinating conidia grow by tip extension and branching to form multinucleate hyphae, the accumulation of hyphal mass is referred to as mycelia. Aerial hyphae undergo budding to generate macroconidia which disperse and repeat the asexual cycle (Davis 2000).

Vegetative cells can spontaneously undergo anastomosis (cell fusion) between branches of the same or different hyphae. Cell fusion facilitates cytoplasmic mixing and the formation of heterokaryons (cells containing different nuclear types). Heterokaryon viability is genetically controlled by heterokaryon incompatibility (*het*) loci. If two strains have different *het* genotypes, then a heterokaryon cell either dies or suffers severe growth inhibition. In contrast, strains with the same *het* genotypes form viable heterokaryons (Saupe 2000).

*N. crassa* is a heterothallic species requiring parents to be of opposing mating types (*A* and *a*) for sexual reproduction. The female parent forms protoperithecia which subsequently become the fruiting bodies known as perithecia when fertilized, typically by a conidium from the male parent. Nuclei from the two parents undergo pre-meiotic DNA replication in the developing mass of ascogenous hyphae and undergo nuclear fusion resulting in the diploid phase of the lifecycle. The diploid nucleus undergoes two meiotic divisions generating four haploid products which undertake an additional mitotic division to form an octad of ascospores contained within a sac called an ascus. Hundreds of asci can form within a perithecium and mature ascospores are ejected from the fruiting body by osmotic pressure within the ascus and germinate upon heat activation (Davis 2000).



**Figure 1.14** The life cycle of *Neurospora crassa*

The inner sequence portrays the asexual cycle where aerial hyphae give rise to macroconidia which form mycelium upon germination. The outer sequence depicts the sexual cycle where fertilization of a protoperithecial cell via a trichogyne with the conidium of the opposite mating type leads to nuclear fusion and meiosis (shown in the development of an individual ascus) culminating in the formation of ascospores within asci. Upon maturation, osmotic pressure within the ascus ejects ascospores from the perithecial wall. Ascospore germination leads to the formation of mycelium. Figure adapted from (Davis 2000).

#### ***1.10.5.3 Neurospora as a model organism***

*N. crassa* was adopted as a model organism for eukaryotic genetics due to a number of key characteristics and advantages. *Neurospora* is non-pathogenic, can be grown rapidly and inexpensively in the laboratory on defined solid or liquid media and strains can remain viable in storage for decades. Haploid progeny simplify genetic analyses and the ordered arrangement of the four products from a single meiosis can be isolated from an octad. Laboratory strains have high levels of sequence polymorphism which can be used as molecular markers. In addition, efficient transformation facilitates targeted gene inactivation and transplacement.

#### ***1.10.5.4 Neurospora as a model organism for recombination***

*Neurospora* is a particularly powerful model organism for elucidating the mechanism of recombination. Homologous recombination occurs during meiosis as crossovers and gene conversion. Crossovers are the exchange of DNA between homologous chromosomes whereas gene conversion refers to the nonreciprocal transfer of DNA sequence from the homologous chromosome or sister chromatid (Orr-Weaver & Szostak 1985). The local regulation of recombination has been extensively characterised in *Neurospora* (Yeadon *et al.*, 2010). This system consists of at least three *trans*-acting *rec* genes (*rec-1*, *rec-2* & *rec-3*) which modulate recombination at specific hotspots throughout the genome with the dominant *rec*<sup>+</sup> alleles acting to suppress recombination (Catcheside 1986). Recent unpublished work has shown that *rec-2* and *rec-2*<sup>+</sup> have totally different sequences; *rec* stimulates recombination at *his-3* but is silenced during meiosis if it is not homozygous (D. Catcheside: *personal communication*). The gene for *rec-2* is present on *LGV* and the *rec-2* gene product affects recombination rates in the *his-3* region on *LGI* through putative interaction with the *cis*-acting *cog* recombination hotspot located 3' of the *his-3* gene and 5' of the *lpl* (putative lysophospholipase) locus (Bowring & Catcheside 1991); (Yeadon & Catcheside 1995). *cog* has two co-dominant alleles which affect recombination within the *his-3* gene and crossing over between *his-3* and *ad-3* genes when *rec-2*<sup>+</sup> is absent

(Yeadon *et al.*, 2004a). The *his-3* gene encodes a tri-functional enzyme catalysing three distinct steps in the histidine biosynthesis pathway (Minson & Creaser 1969). There is a high degree of sequence polymorphism in the intergenic region between *his-3* and *cog* (Yeadon & Catchside 1995) and within the *his-3* gene enabling the parental origin of recombined sequences to be determined (Yeadon & Catchside 1998). Recombination between *his-3* alleles typically results from gene conversion rather than crossing over within the gene (Yeadon & Catchside 1998). Conversion tracts initiated at *cog* can extend up to at least ~6kb, are tolerant of substantial sequence heterology and are often discontinuous with patches of sequence from both parents (Yeadon & Catchside 1995); (Yeadon & Catchside 1998); (Yeadon *et al.*, 2004b). The characteristics of hot-spot mediated recombination (HSMR) at the *cog-his-3* region were utilised for the development of the *Neurospora* directed evolution platform (discussed in 1.10.5.6)

#### **1.10.5.5 Repeat-Induced Point Mutation**

Repeat-induced point mutation (RIP) is a mechanism whereby linked DNA sequence duplications greater than ~500bp and unlinked duplications greater than ~1kb are detected in the *Neurospora* genome and mutated in a pair-wise manner during the sexual cycle (Selker *et al.*, 1987). RIP acts by converting cytosine to thymine with the frequency of mutation varying according to the dinucleotide context: CpA (~64%), CpT (~18%), CpG (~13%), CpC (~5%) (Cambareri *et al.*, 1989). The efficacy of RIP in mutating duplicated sequences is demonstrated by the widespread usage of RIP in *Neurospora* gene knockout studies (Gavric *et al.*, 2007); (Koh & Catchside 2007); (Wang *et al.*, 2007). RIP is thought to act as a genome defence mechanism against foreign genetic elements such as transposons (Selker 1997).

Although RIP is typically a highly mutagenic and iterative process, variation in the degree and distribution of mutation means RIP does not always generate null alleles (Fincham 1990). This aspect of RIP had been previously exploited to obtain 70 leaky mutants of the essential endogenous *Neurospora al-3* gene (Barbato *et al.*, 1996).

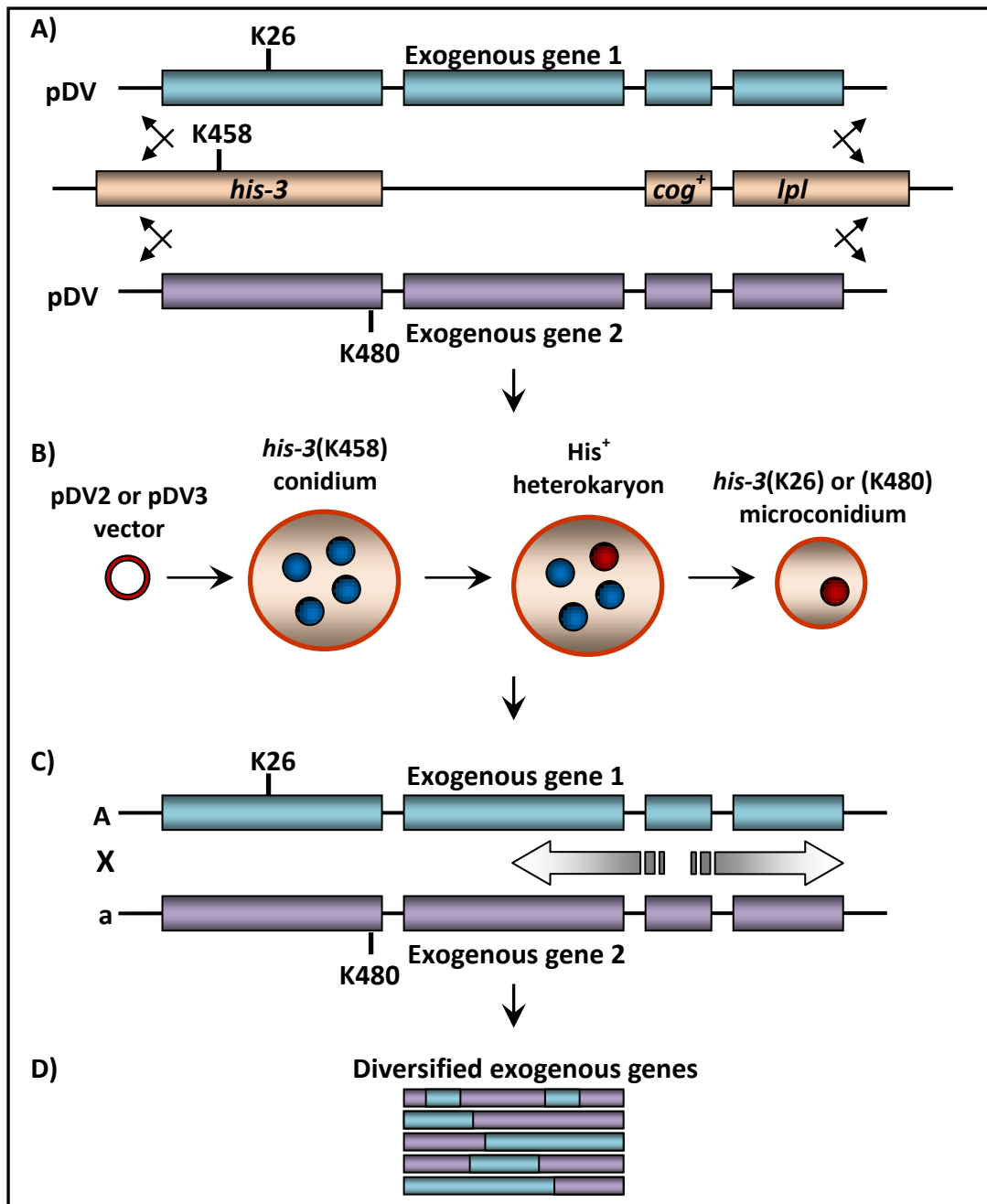


Sequence analysis indicated that no more than six point mutations and no more than three amino acid changes were present in each of the mutants. Likewise, (Fincham 1990) isolated a lightly 'RIPed' mutant of the *Neurospora am* gene encoding glutamate dehydrogenase with wildtype levels of activity but reduced thermostability. The frequency of RIP mutation can be modulated to a certain degree by altering various factors enhancing the potential of RIP for gene diversification. Reducing RIP-mediated mutations can be achieved by isolating the earliest ascospores produced by perithecia as these ascospores have spent the shortest time in the dikaryotic tissue in which RIP occurs (Singer 1995), incubating crosses below the 25°C temperature normally used (E. Cambareri: *Personal communication*) and/or increasing cellular levels of S-adenosylmethionine (Rosa *et al.*, 2004). The use of RIP for *in vivo* gene diversification of single genes has been patented by Neugenes Corporation (Cambareri 2008)

#### **1.10.5.6 The Neurospora directed evolution platform**

The directed evolution system in *N. crassa* utilises hot spot-mediated recombination (HSMR) for the *in vivo* diversification of exogenous genes (Figure 1.15) (Catchside *et al.*, 2003). Diversification vectors pDV2 and pDV3 were constructed to enable insertion of exogenous DNA between *his-3* and *cog*<sup>+</sup> in *Neurospora* strains t11630 and t11644 (Rasmussen *et al.*, 2002). The diversification strains t11630 and t11644 contain the *rec-2* allele facilitating high levels of recombination initiating at the *cog*<sup>+</sup> hotspot, *his-3* (K458) for transplacement selection and opposite mating types.

The pDV2 and pDV3 vectors contain truncated *his-3* K26 and K480 alleles, respectively, and part of *cog/lpl* to facilitate integration via homologous recombination at the *his-3* & *cog/lpl* loci. Selection for transplacement is possible as both vector borne *his-3* alleles will complement the recipient strain's K458 allele in a heterokaryotic transformant, enabling His<sup>+</sup> prototroph selection on minimal medium.



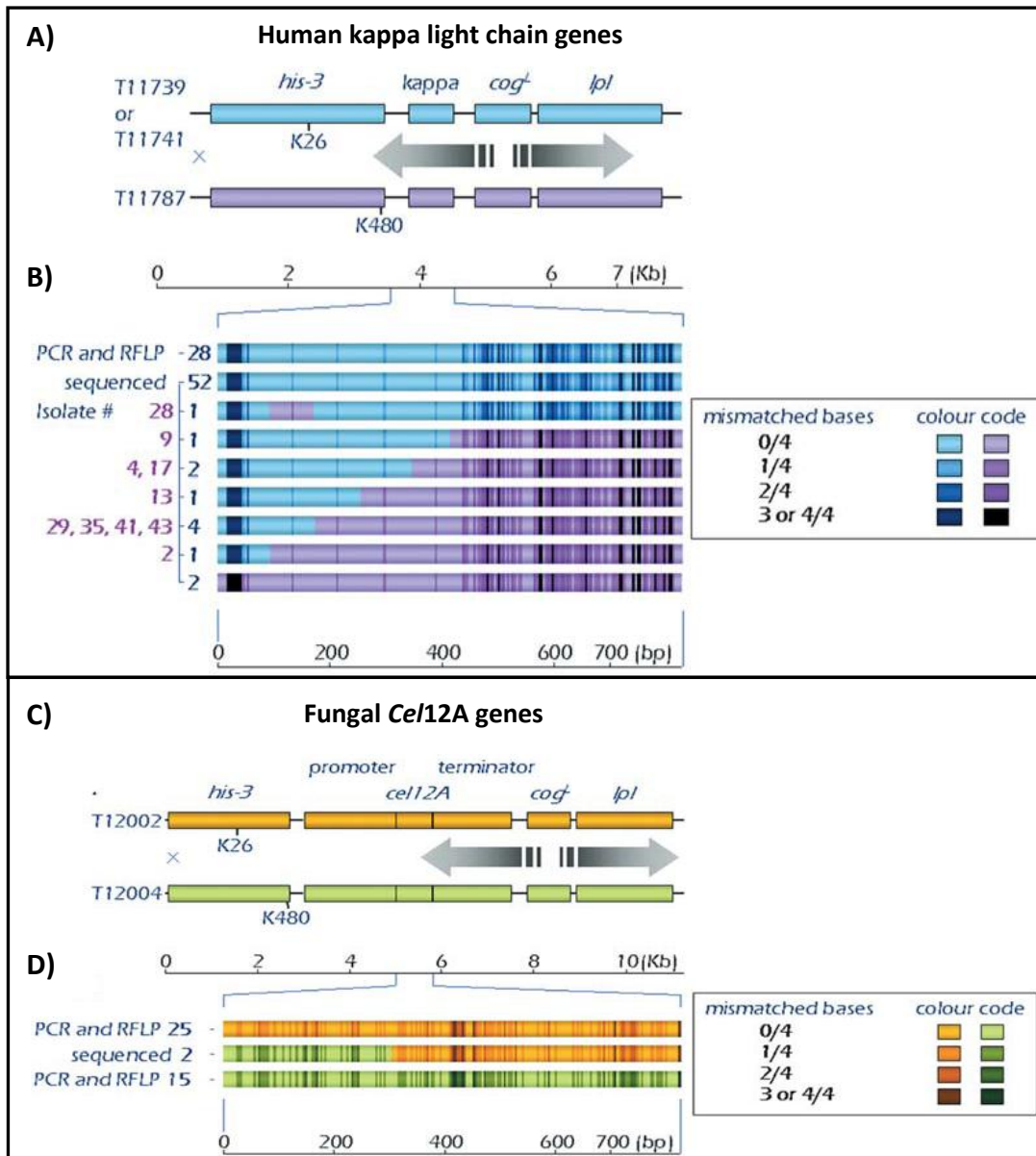
**Figure 1.15** *In vivo* shuffling of exogenous genes in *Neurospora*

**A)** A pair of homologous exogenous genes cloned into targeting vectors pDV2 and pDV3 are transplanted between *his-3* and the *cog*<sup>+</sup> recombination hotspot in Linkage Group I. **B)** His<sup>+</sup> heterokaryons are formed by transplacment in one nucleus of a multinucleate *his-3*(K458) conidium by *his-3*(K26) or *his-3*(K480) from the pDV vectors. Microconidiation facilitates recovery of nuclei containing the exogenous genes from the heterokaryon. **C)** *Neurospora* strains containing transplanted homologous genes and opposing mating types are crossed. **D)** Selection for His<sup>+</sup> recombinants enriches for progeny with exchanges between the exogenous genes (recombination shown as arrows originating at *cog*<sup>+</sup>), generating novel variants. Figure adapted from Rasmussen *et al.* (2002).

Monokaryons containing K26 or K480 are isolated from the heterokaryons and crossed together. Isolation of His<sup>+</sup> progeny from the crosses selects for recombination events originating at *cog*<sup>+</sup> that have extended through the heterologous DNA inserted between *cog* and *his-3* (Yeadon *et al.*, 2001). As described previously, recombination events initiated by *cog*<sup>+</sup> are tolerant of substantial sequence mismatch enabling them to extend through the exogenous DNA and often contain patches of sequence from both parents (Yeadon & Catcheside 1995); (Catcheside *et al.*, 2003).

The meiotic products form ascospores containing two identical haploid nuclei and are produced in virtually unlimited numbers. His<sup>+</sup> progeny enriched for novel forms of the gene of interest can be selected by plating ascospores on medium lacking histidine. Secretion of protein variants can be achieved by cloning the gene of interest into an expression cassette with appropriate signal sequences and integrating the cassette into the *Neurospora* genome.

This system has been used to diversify the human immunoglobulin kappa chain (*IgGκ*) gene (DNA variants only) and endoglucanase 3 (*Cel12A*) genes (DNA & protein variants) from different fungal species (Figure 1.16) (Catcheside *et al.*, 2003). However, there is still potential to further improve and broaden the capabilities of the *Neurospora* directed evolution platform, one such improvement is to use RIP to introduce sequence diversity into the genes of interest by mutation prior to the shuffling event of recombination.



**Figure 1.16 Gene diversification of human kappa light chain and fungal endoglucanase genes in *Neurospora crassa***

**A)** Sexual cross between strains with human kappa light chain genes (IgGkA and IgGkB) transplanted between *his-3* and *cog* in Linkage Group I. **B)** Location of exchange events between kappa sequences amongst 92 His<sup>+</sup> progeny are indicated by change in colour with the degree of mismatched sequence in each 4bp interval depicted by colour intensity. **C)** Sexual cross between strains with *Cel12A* genes from *Trichoderma reesei* and *Hypocrea schweinitzii* transplanted between *his-3* and *cog* in Linkage Group I. **D)** Location of exchange events between *Cel12A* sequences amongst 42 His<sup>+</sup> progeny are indicated by change in colour with the degree of mismatched sequence in each 4bp interval depicted by colour intensity. Figure adapted from Catchside *et al.* (2003).

## 1.11 Aims

The work described within this thesis had two primary aims. The first was to utilise the *Neurospora* directed evolution platform to diversify human IGF-I and to generate IGF-IR antagonists as potential cancer therapeutics. The second was to use site-directed mutants of IGF-II to identify residues within the IGF C-domain involved in differential receptor binding and activation.

The generation of human IGF-I mutants using *Neurospora* directed evolution had three major objectives:

- 1) To use RIP and HSMR for *in vivo* gene diversification of human IGF-I in *Neurospora crassa*.
- 2) To utilize *Neurospora* for the heterologous expression and secretion of human IGF-I protein.
- 3) To identify and characterize IGF-I variants acting as antagonists of the IGF-IR.

The generation of human IGF-II mutants using site-directed mutagenesis had two major objectives:

- 1) To generate three IGF-II analogues with partial IGF-I C-domain substitutions.
- 2) To characterise the three IGF-II analogues with respect to receptor binding, receptor phosphorylation and cell survival mediated through the IR-A, IR-B and IGF-IR.

## Chapter 2      Materials and methods

### 2.1      **Materials**

#### 2.1.1      **Chemicals, reagents & enzymes**

General laboratory chemicals were analytical grade or the best quality available and generally sourced from Sigma-Aldrich or BDH. Suppliers for various reagents and enzymes are detailed in specific methods below.

#### 2.1.2      **Neurospora crassa stocks**

Stock number refers to cataloguing system used in the Catcheside laboratory, Biological Sciences, Flinders University.

Stock number	Genotype
t2326	<i>a; fl</i>
t2327	<i>A; fl</i>
t11630	<i>a, his-3 (K458), cog<sup>+</sup>; cot-1 (C102T); rec-2</i>
t11644	<i>A, his-3 (K458), cog<sup>+</sup>; cot-1 (C102T); rec-2</i>
t12001	<i>a, his-3 (K480):: Cel12A (G. roseum); eth-1; exp-1; cog<sup>+</sup>; cot<sup>+</sup> ; rec-2</i>

2.1.3 **Bacterial stocks**

Strain	Genotype	Source
<b>DH5α</b>	<i>F' Phi80dlacZ DeltaM15 Delta(lacZYA-argF)U169 deoR recA1 endA1 hsdR17(rK-mK+ phoA supE44 lambda- thi-1</i>	Catcheside laboratory (Flinders University)
<b>DH10B</b>	<i>F' mcrA (mrr-hsdRMS-mcrBC) φ80lacZ M15 lacX74 recA1 endA1 araD139(ara, leu)7697 galU galK λ<sup>-</sup> rpsL nupG</i>	Thomas laboratory (Flinders University)
<b>JM101</b>	<i>F' traD36 proA<sup>+</sup>B<sup>+</sup> lacI<sup>q</sup> Δ(lacZ)M15/ Δ(lac-proAB) glnV thi</i>	Wallace laboratory (Adelaide University)
<b>JM109</b>	<i>F' traD36 proA<sup>+</sup>B<sup>+</sup> lacI<sup>q</sup> Δ(lacZ)M15/ Δ(lac-proAB) glnV44 e14<sup>-</sup> gyrA96 recA1 relA1 endA1 thi hdsR17</i>	Catcheside laboratory (Flinders University)
<b>TOP10F'</b>	<i>F'(lacIq Tn10 (TetR)) mcrA (mrr-hsdRMS-mcrBC) φ80lacZ M15 lacX74 recA1 araD139 (ara-leu)7697 galU galK rpsL endA1 nupG</i>	Invitrogen (cat#C3030-03)

2.1.4 Plasmid stocks

Vector name	Function in this study	Source
pGEM-T Easy	Cloning of PCR amplicons	Promega, Australia
pCR2.1-TOPO	Cloning of PCR amplicons	Invitrogen, Australia
pSH1.5EGFP	Vector backbone for ectopic integration	Created by the author in an unrelated study
pHDKXL1	Source of <i>gla-1</i> fusion expression cassette	Neugenesis Corp, California
pDV2	Vector backbone for <i>his-3 / cog</i> targeted integration	Rasmussen <i>et al.</i> (2002)
phIGFI	Source of human IGF-I cDNA	Dr Morry Frenkel (CSIRO, Melbourne)
pFEN	Vector created to facilitate fusion of proteins with partial <i>N. crassa gla-1</i>	This study
pFENIGFI	Ectopic expression of RIP optimised <i>nIGF-I</i> fused to <i>gla-1</i>	This study
pFENIGFI-GS	Ectopic expression of RIP optimised <i>nIGF-I</i> fused to <i>gla-1</i> with a glycine spacer	This study
pNEIGFI-G1	Ectopic expression of RIP optimised <i>nIGF-I</i> with a <i>gla-1</i> signal peptide	This study
pNEhIGFI-C1	Ectopic expression of human <i>IGF-I</i> with <i>egl3 (T. reesei)</i> signal peptide	This study
pNEhIGFI-H1	Ectopic expression of human <i>IGF-I</i> with native human signal peptide	This study
pFENIGFI-DV2	Targeted expression of RIP optimised <i>nIGF-I</i> fused to <i>gla-1</i>	This study
pFENIGFI-GS-DV2	Targeted expression of RIP optimised <i>nIGF-I</i> fused to <i>gla-1</i> with a glycine spacer	This study
pNEIGFI-G1-DV2	Targeted expression of RIP optimised <i>nIGF-I</i> with a <i>gla-1</i> signal peptide	This study
pNEhIGFI-C1-DV2	Targeted expression of human <i>IGF-I</i> with <i>egl3 (T. reesei)</i> signal peptide	This study
pNEhIGFI-H1-DV2	Targeted expression of human <i>IGF-I</i> with native human signal peptide	This study
pNEIGFI-G1-DV2-LD1	Targeted linked duplication of RIP optimised <i>nIGF-I</i>	This study
IGFII/pGEM	Human <i>IGF-II</i> coding sequence ( <i>E. coli</i> codon optimised) in pGEM-T easy vector	This study
pGH(1-11)VNPAPM	Vector backbone for <i>E. coli</i> optimised expression of <i>IGF-II</i>	Denley <i>et al.</i> (2004)
pGH(1-11)VNPAPM-C1	<i>E. coli</i> expression vector with 'C1' <i>IGF-II</i> mutant	This study
pGH(1-11)VNPAPM-C2	<i>E. coli</i> expression vector with 'C2' <i>IGF-II</i> mutant	This study
pGH(1-11)VNPAPM-C3	<i>E. coli</i> expression vector with 'C3' <i>IGF-II</i> mutant	This study



**2.1.5 Mammalian cell lines**

Cell line	Description	Source
P6	Mouse BALB/c3T3 fibroblast cells overexpressing IGF-IR	Dr Renato Baserga (Thomas Jefferson University)
R <sup>-</sup> IRA	Human IR-A expressing embryonic 3T3-like fibroblasts from IGF-IR knockout mouse	Generated by Dr Eric Bonython (Adelaide University) using the R <sup>-</sup> cell line obtained from Dr Renato Baserga, (Thomas Jefferson University)
R <sup>-</sup> IRB	Human IR-B expressing embryonic 3T3-like fibroblasts from IGF-IR knockout mouse	
R <sup>-</sup> IGF-IR	Human IGF-IR expressing embryonic 3T3-like fibroblasts from IGF-IR knockout mouse	

**2.1.6 Antibodies**

Antibody	Type	Working dilution/ conc.	Purpose	Ig type	Source & (Cat#)
Anti-human IGF-I monoclonal	1°	1:100	Western blotting	Mouse	Gropep (MAM1)
Anti-human IGF-I polyclonal	1°	1:100	Western blotting	Goat	R&D systems (AF-291-NA)
Anti-human IGF-I receptor monoclonal (24-31)	1°	2.5µg/ml	Receptor binding assay	Mouse	Professor Ken Siddle (University of Cambridge)
Anti-human insulin receptor monoclonal (83-7)	1°	2.5µg/ml	Receptor binding assay	Mouse	Professor Ken Siddle (University of Cambridge)
Europium labelled anti-phosphotyrosine monoclonal (PY20)	1°	1:10000	Phosphorylation assay	Mouse	Perkin Elmer (AD0038)
Anti-goat-HRP conjugate	2°	1:2000	Western blotting	Donkey	Rockland (605-703-125)
Anti-mouse-HRP conjugate	2°	1:2000	Western blotting	Goat	Anderson laboratory (Flinders University)

### 2.1.7 Protein standards

Name	Purity	Source	Cat. #
IGF-I	Receptor grade	Gropep	CM001
IGF-II	Media grade	Gropep	OM001
Insulin (Actrapid)	Medical grade	Novo NordiskPharmaceuticals	R34336
BCIAD	HPLC	Wallace laboratory, Molecular and Biomedical Sciences, Adelaide University	-

### 2.1.8 Oligonucleotides

#### 2.1.8.1 Oligonucleotides for construction of RIP optimised nIGF-I coding sequence

All oligos were 5' phosphorylated and PAGE purified by the supplier (Sigma-Aldrich).

Oligo name	Sequence (5' → 3')	Function
SO1-IGFI	GGCCCAGAGACTCTGTGGTGCAGAGCTCGTCGACG CACTCCAGTTTGTCTGTGGTGACCGGGTTTTACTT	Construction of RIP optimised IGF-I coding sequence (nIGF-I)
ASO1-IGFI	TAAAGTAAAAACCCCGGTCACCACAGACAACTGGAG TGCCTCGACGAGCTCTGCACCACAGAGTGTCTCTGGG CCTGCA	
SO2-IGFI	TAACAAGCCAACAGGTTATGGCTCATCATCACGGCGG GCACCACAGACAGGTATTGTTGATGAGTGCTGCTT	
ASO2-IGFI	CCGGAAGCAGCACTCATCAACAATACCTGTCTGTGGT GCCCCCGTGATGATGAGCCATAACCTGTTGGCTTGT	
SO3-IGFI	CCGGTCATGTGATCTCCGGCGGCTTGAGATGTACTGT GCACCACTCAAGCCTGCAAAGTCAGCATAAT	
ASO3-IGFI	CTAGATTATGCTGACTTTGCAGGCTTGAGTGGTGCAC AGTACATCTCAAGCCCGGAGATCACATGA	

2.1.8.2 Oligonucleotides for sequencing

Supplied by Geneworks

Oligo name	Sequence (5' → 3')	Function
ECS-3F	AAGCTAGATGCTAAGCGATATTGC	Sequencing of expression cassette in pFEN
ECS-4F	GCTGCCTCTGGTGTGTGG	
ECS-4R	CGACACTGTGAGCTATGCTGC	
ECS-5F	GCCCATTGTGAAGAACGACC	
ECS-5R	GCGATGGTGAAGAAAGAAGAGC	
ECS-6F	CTTCTGGGGTGTCAACTCCG	
ECS-6R	GCGGCGAGAGTAGCGAGG	
ECS-7F	ACGCCTCCTTCCTTTCCG	
ECS-7R	ACGGTGGAGGCGGAGC	
ECS <b>sb</b> fl ECS <b>Xba</b> I	CCCCTCCACGCAGACC CCTTTTCCAATCACTACTTACTACC	
LD1&2-dIGF-F	AGTAGCAAAAAAAAAACAAGAACAAAAGG	Screening and sequencing of <i>nIGF-I</i> duplication in pNEIGFI-G1-DV2-LD1
LD1&2-dIGF-R	GATGTAATAATGAAAATAAACACCCAGG	
LD1&2-nIGF-F LD1&2-nIGF-R	AGTGGTTGGTCGGGGTTCC AGCCGCAGTGCCTCGC	Screening and sequencing of <i>nIGF-I</i> in pNEIGFI-G1-DV2-LD1
SP6 T7	ATTTAGGTGACACTATAGAA TAATACGACTCACTATAGGG	Sequencing pGEM inserts

2.1.8.3 **Oligonucleotides utilized for IGF-I related work**

Oligos supplied by Geneworks. Note: \* = HPLC purified by supplier.

Oligo name	Sequence (5' → 3')	Function
EXPCASS-F1	GGTCAGCATGCGTCTTGAGC	Cloning expression cassette
EXPCASS-R1	GTACATGTTCCACTAGTAAGGCTGACAC	
IGFI-repair1	CTCCACCTCTAGATTATGCTGACTTTG	Repair of <i>nIGF-I</i> CDS
GlySpacer-1	CTAGCGACATCATCAAGAGGGGCGGCGG CGGCGGCCCTGCA	Construction of penta- glycine spacer
GlySpacer-2	GGGCCGCCGCCGCCGCCCTCTTGATGAT GTCG	
IGFI-BseRI-F*	ATCTCGTCTCTTCGCTCCTCGTCGTGGGCG CCGCCTTCCAGGCCGTGCTCGGTGGCCCA GAGACACTCTGTGG	Construction of <i>gla-1</i> ( <i>N. crassa</i> ) signal peptide
IGFI-XbaI-R	GATTTACCTCCACTCTAGATTATGCTG	
hIGFI-NS-BglII	ATCTCAGATCTCATGGGAAAAATCAGCAG TCTTC	Cloning of human <i>IGF-I</i> with native signal peptide
hIGFI-NS-XbaI	ATCTCTCTAGACCTAAGCTGACTTGGCAG GC	
hIGFI-BglII*	ATCTCAGATCTATGAAGTTCCTTCAAGTCC TCCCTGCCCTCATAACCGGCCGCCCTGGCC GGACCGGAGACGCTCTGC	Construction of <i>Cel12A</i> ( <i>T. reesei</i> ) signal peptide
GlaA-F	GAACTTAGCCTTATGAGATGAATGATG	Screening for <i>glaA</i> promoter in all <i>IGF-I</i> transformants
GlaA-R	GGAACCCCGACCAACCAC	
hGFI-checkF	CGGAGACGCTCTGCGGG	Screening of human <i>IGF-I</i> transformants
hGFI-checkR	AGCTGACTTGGCAGGCTTGAG	
IGFI-check-F	CCCAGAGACACTCTGTGGTGC	Screening of <i>nIGF-I</i> transformants
IGFI-check-R	TTATGCTGACTTTGCAGGCTTG	
nIGFI-DUD1-F	AAGGTTGTACATGACCAGAGACACTCTGT GGTGCAGAG	Construction of linked <i>nIGF-I</i> duplication
nIGFI-DUD1-R	AAGGTTGTACATTCCACTACCAAGGCTGA CACG	
B-tub-F1	GAGGTTGAGGACCAGATGCG	Boil prep PCR positive control
B-tub-R1	CCTCCTCCTCGTCAACACCAG	

**2.1.8.4 Oligonucleotides used for construction of IGF-II mutants**

Oligos supplied by Geneworks.

Oligo name	Sequence (5' → 3')	Function
EIGFII-F	GTTAACCCGGCACCGATG	Cloning of <i>E. coli</i> optimised IGF-II CDS
EIGFII-R	AAGCTTTTATTCGCTTTTTGC	
C1-F	CTCGTCCGGCGGGCTATGGCAGCAGCAGCCGGCGTTCTCGCGG	Construction of 'C1' IGF-II mutant
C1-R	CCGCGAGAACGCCGGCTGCTGCTGCCATAGCCCGCCGGACGAG	
C2-F	CTCGCGTTAGCCGGCGTGCGCCACAGACCGGAATCGTGG	Construction of 'C2' IGF-II mutant
C2-R	CCACGATTCCGGTCTGTGGCGCACGCCGGCTAACGCGAG	
C3-F	CTCGTCCGGCGGGCTCTAGCAGCCGGCGTGCGACCGGAATCGTG G	Construction of 'C3' IGF-II mutant
C3-R	CCACGATTCCGGTCGCACGCCGGCTGCTAGAGCCCGCCGGACGA G	

## 2.1.9 Buffers & solutions

### 2.1.9.1 Blocking solution, 1x

NaCl	7.3g	(125mM)
Na <sub>2</sub> HPO <sub>4</sub>	2.4g	(17mM)
NaH <sub>2</sub> PO <sub>4</sub> ·2H <sub>2</sub> O	1.25g	(8mM)
SDS	50g	(174mM)
mqH <sub>2</sub> O	up to 1L	

- Adjusted pH to 7.2
- Stored at RT

### 2.1.9.2 Blotto, 1x

Skim milk powder (Diploma brand)	50g	(5% w/v)
TBST (1x)	1L	

- Stored at 4°C and used within 2 days

### 2.1.9.3 Carbonate buffer, 1x

Na <sub>2</sub> CO <sub>3</sub>	1.59g	(15mM)
NaHCO <sub>3</sub>	2.93g	(35mM)
mqH <sub>2</sub> O	up to 1L	

- Adjusted pH to 9.5
- Stored at 4°C and used within 14 days

### 2.1.9.4 Cell lysis buffer for receptor binding assays, 1x

HEPES	4.766g	(20mM)
NaCl	8.766g	(150mM)
MgCl <sub>2</sub> ·6H <sub>2</sub> O	243mg	(1.2mM)
Glycerol	100ml	(10% v/v)
Triton X-100	10ml	(1% v/v)
EDTA (disodium salt, dihydrate)	372mg	(1mM)
mqH <sub>2</sub> O	up to 1L	

- Adjusted pH to 7.5 with 1M NaOH
- Aliquoted and stored at -20°C

**2.1.9.5 Cell lysis buffer for receptor phosphorylation assays, 1x**

HEPES	4.766g	(20mM)
NaCl	8.766g	(150mM)
MgCl <sub>2</sub> ·6H <sub>2</sub> O	243mg	(1.2mM)
Glycerol	100ml	(10% v/v)
Triton X-100	10ml	(1% v/v)
EDTA (disodium salt, dihydrate)	372mg	(1mM)
PMSF (100mM)*	10ml	(1mM)

\*Dissolved in isopropanol

Na <sub>3</sub> VO <sub>4</sub>	184mg	(1mM)
NaF	420mg	(10mM)
mqH <sub>2</sub> O	up to 1L	

- Adjusted pH to 7.5 with 1M NaOH
- Aliquoted and stored at -20°C

**2.1.9.6 Coomassie Blue staining solution, 1x**

Coomassie Blue R 250	1g	(0.1 % w/v)
Methanol	500ml	(50% v/v)
Acetic acid	100ml	(10 % v/v)
mqH <sub>2</sub> O	up to 1L	

- Stored at RT for up to 6 months

**2.1.9.7 Denaturation solution, 1x**

NaOH	20g	(500mM)
NaCl	87.7g	(1.5M)
mqH <sub>2</sub> O	up to 1L	

- Stored at RT

**2.1.9.8 Destaining solution, 1x**

Methanol	100ml	(10% v/v)
Acetic acid	100ml	(10% v/v)
mqH <sub>2</sub> O	up to 1L	

- Stored at RT

**2.1.9.9 E. coli lysis buffer**

SDS (10% w/v)	200µl	(2% v/v)
β-mercaptoethanol	1ml	(10% v/v)
mqH <sub>2</sub> O	up to 10ml	

- Used immediately

**2.1.9.10 Fast flow S (FFS) buffer A, 1x**

Urea	480.48g	(8M)
Sodium Acetate (3M)	16.6ml	(50mM)
mqH <sub>2</sub> O	up to 1L	

- Adjusted pH to 4.8 with glacial acetic acid
- Stored at RT

**2.1.9.11 Fast flow S (FFS) buffer B, 1x**

Urea	480.48g	(8M)
Sodium acetate (3M)	16.6ml	(50mM)
NaCl	58.44g	(1M)
mqH <sub>2</sub> O	up to 1L	

- Adjusted pH to 4.8 with glacial acetic acid
- Stored at RT

**2.1.9.12 Freezing solution, 1x**

Fetal calf serum	9.5ml	(95% v/v)
DMSO	500µl	(5% v/v)

- Stored at 4°C

**2.1.9.13 Inclusion body dissolution buffer, 1x**

Urea	480.48g	(8M)
Tris base	12.11g	(100mM)
Glycine	3g	(40mM)
ZnCl <sub>2</sub>	68mg	(500µM)
mqH <sub>2</sub> O	up to 1L	

- Adjusted pH to 2.0 with conc. HCl

**2.1.9.14 Inclusion body wash buffer, 1x**

KH <sub>2</sub> PO <sub>4</sub>	1.36g	(10mM)
NaCl	1.74g	(30mM)
ZnCl <sub>2</sub>	68mg	(500µM)
mqH <sub>2</sub> O	up to 1L	

- Adjusted pH to 7.8 with conc. HCl
- Stored at RT



**2.1.9.15 Ligand binding buffer, 1x**

HEPES	23.8g	(100mM)
NaCl	5.8g	(100mM)
mqH <sub>2</sub> O	900ml	
<ul style="list-style-type: none"> <li>Adjusted pH to 8.0</li> </ul>		
DTPA	800µg	(2µM)
Tween-20	500µl	(0.05%)
mqH <sub>2</sub> O	up to 1L	
<ul style="list-style-type: none"> <li>Stored at 4°C</li> </ul>		

**2.1.9.16 Neurospora genomic DNA extraction buffer, 1x**

Tris-HCl (1M, pH 8.0)	100ml	(100mM)
NaCl	5.84g	(100mM)
EDTA (0.25M)	40ml	(10mM)
N-lauroylsarcosine	10g	(37mM)
mqH <sub>2</sub> O	up to 1L	
<ul style="list-style-type: none"> <li>Stored at RT</li> </ul>		

**2.1.9.17 Neutralisation solution, 1x**

Tris base	121.1g	(1M)
NaCl	87.7g	(1.5M)
mqH <sub>2</sub> O	up to 1L	
<ul style="list-style-type: none"> <li>Adjusted pH to 7.5</li> <li>Stored at RT</li> </ul>		

**2.1.9.18 PBS, 10x**

NaCl	80g	(1.37M)
KCl	2g	(27mM)
Na <sub>2</sub> HPO <sub>4</sub>	14.4g	(100mM)
KH <sub>2</sub> PO <sub>4</sub>	2.4g	(18mM)
mqH <sub>2</sub> O	up to 1L	
<ul style="list-style-type: none"> <li>Adjusted pH to 7.4</li> <li>Sterilised by autoclaving for 15min at 121°C/103kPa</li> <li>Stored at RT</li> </ul>		

**2.1.9.19 Refold dilution buffer, 1x**

Tris base	121.11g	(1M)
EDTA (disodium salt, dihydrate)	18.6g	(50mM)
mqH <sub>2</sub> O	up to 1L	
<ul style="list-style-type: none"> <li>Adjusted pH to 9.1 with conc. HCl</li> <li>Filtered (0.1µM)</li> <li>Stored at 4°C</li> </ul>		

**2.1.9.20 rpHPLC buffer A, 1x**

Trifluoroacetic acid	1ml	(0.1%)
mqH <sub>2</sub> O	up to 1L	
<ul style="list-style-type: none"> <li>• Filtered (0.22µm)</li> <li>• Stored at RT</li> </ul>		

**2.1.9.21 rpHPLC buffer B, 1x**

Trifluoroacetic acid	800µl	(0.08%)
Acetonitrile	800ml	(80%)
mqH <sub>2</sub> O	up to 1L	
<ul style="list-style-type: none"> <li>• Filtered (0.22µm)</li> <li>• Stored at RT</li> </ul>		

**2.1.9.22 SSC, 20x**

NaCl	175.3g	(3M)
Trisodium citrate (dihydrate)	88.2g	(300mM)
mqH <sub>2</sub> O	up to 1L	
<ul style="list-style-type: none"> <li>• Adjusted pH to 7.0</li> <li>• Stored at RT</li> </ul>		

**2.1.9.23 Stringency wash solution A, 1x**

SSC, 20x (2.1.9.22)	100ml	(2x)
SDS, 10%	10ml	(1%)
mqH <sub>2</sub> O	up to 1L	
<ul style="list-style-type: none"> <li>• Stored at RT</li> </ul>		

**2.1.9.24 Stringency wash solution B, 1x**

SSC, 20x (2.1.9.22)	5ml	(0.1x)
SDS (10% w/v)	10ml	(1% v/v)
mqH <sub>2</sub> O	up to 1L	
<ul style="list-style-type: none"> <li>• Stored at RT</li> </ul>		

**2.1.9.25 TAE, 50x**

Tris base	242g	(2M)
Glacial acetic acid	57.1ml	(1M)
EDTA (disodium salt, dihydrate)	18.6g	(50mM)
mqH <sub>2</sub> O	up to 1L	
<ul style="list-style-type: none"> <li>• Stored at RT</li> </ul>		

**2.1.9.26 TBS, 10x**

Tris base	12.11g	(100mM)
NaCl	87.66g	(1.5M)
mqH <sub>2</sub> O	up to 1L	
<ul style="list-style-type: none"> <li>• Adjusted pH to 7.6</li> <li>• Stored at 4°C</li> </ul>		

**2.1.9.27 TBST, 1x**

TBS (10x, pH 7.6 – 2.1.9.26)	100ml	(1x)
Tween-20	1ml	(0.1% v/v)
mqH <sub>2</sub> O	up to 1L	
<ul style="list-style-type: none"> <li>• Stored at 4°C</li> </ul>		

**2.1.9.28 TCA, 100% (w/v)**

TCA	10g	(100% w/v)
mqH <sub>2</sub> O	4.54ml	
<ul style="list-style-type: none"> <li>• Stored at RT in dark bottle</li> </ul>		

**2.1.9.29 TE, 1x**

Tris-HCl (1M, pH 8.0)	10ml	(10mM)
EDTA (0.25M)	4ml	(1mM)
mqH <sub>2</sub> O	up to 1L	
<ul style="list-style-type: none"> <li>• Stored at RT</li> </ul>		

**2.1.9.30 TSS, 1x**

LB medium (refer 2.1.11.1)	42.5ml	
MgSO <sub>4</sub> (2M)	500µl	(20mM)
PEG 4000*	5g*	(25mM)
* dissolved in 5ml of warm LB medium		
<ul style="list-style-type: none"> <li>• Adjusted pH to 6.5-6.8</li> <li>• Filter sterilized (0.22µm)</li> </ul>		
DMSO	2.5ml	(5% v/v)
<ul style="list-style-type: none"> <li>• Stored at 4°C</li> </ul>		

**2.1.9.31 Wash solution I, 1x**

Blocking Solution (2.1.9.1)	100ml	
mqH <sub>2</sub> O	up to 1L	
<ul style="list-style-type: none"> <li>• Stored at RT</li> </ul>		

**2.1.9.32 Wash solution II, 10x**

Tris base	12.11g	(100mM)
NaCl	5.8g	(100mM)
MgCl <sub>2</sub> .6H <sub>2</sub> O	2.03g	(10mM)
mqH <sub>2</sub> O	up to 1L	

- Adjusted pH to 9.5
- Stored at RT

**2.1.10 Neurospora culture media**

**2.1.10.1 Crossing media**

**2.1.10.1.1 Plate crossing medium**

SC, 2x stock (2.1.10.11)	50 ml	(1x)
Bacto-agar	2g	(2% w/v)
Supplements (2.1.10.14)	As required	
mqH <sub>2</sub> O	up to 100ml	

- Autoclaved for 15min at 121°C/103kPa
- Aseptically dispensed as described in 2.2.1.2.3.1.

**2.1.10.1.2 Tube crossing medium**

SC, 2x stock (2.1.10.11)	50 ml	(1x)
Sucrose	2 g	(2% w/v)
Supplements (2.1.10.14)	As required	
mqH <sub>2</sub> O	up to 100ml	

- Dispensed as described in 2.2.1.2.3.2.
- Autoclaved for 15min at 121°C/103kPa

**2.1.10.2 *FRIES medium, 10x***

- Adapted from Davis (1970)

Sucrose	75g	(7.5% w/v)
Ammonium tartrate	50g	(270mM)
NH <sub>4</sub> NO <sub>3</sub>	10g	(125mM)
KH <sub>2</sub> PO <sub>4</sub>	10g	(74mM)
MgSO <sub>4</sub> ·7H <sub>2</sub> O	5g	(20mM)
NaCl	1g	(17mM)
CaCl <sub>2</sub> ·2H <sub>2</sub> O	1.3g	(113mM)
Bacto-yeast extract	10g	(1% w/v)
m <sub>q</sub> H <sub>2</sub> O	up to 1L	

- Stored at -20°C.
- Diluted to 1x prior to use
- Adjusted pH to 6.0
- Autoclaved for 15min at 121°C/103kPa.

**2.1.10.3 *FRIES liquid medium for protein secretion***

FRIES (10x) (2.1.10.2)	10ml	(1x)
Sucrose	7.25g	(8% w/v)
Supplements (2.1.10.14)	As required	
m <sub>q</sub> H <sub>2</sub> O	up to 100ml	

- Adjusted pH to 6.0
- Autoclaved for 15min at 121°C/103kPa then aseptically added:

Protease Inhibitor Cocktail (Sigma, cat#P8340)	100µl	(1:200)
--	-------	---------

**2.1.10.4 *FRIES + sorbose liquid medium for protein secretion***

Part 1)

FRIES (10x) (2.1.10.2)	10ml	(1x)
Sucrose	7.25g	(8% w/v)
Supplements (2.1.10.14)	As required	
m <sub>q</sub> H <sub>2</sub> O	up to 50ml	

- Adjusted pH to 6.0
- Autoclaved for 15min at 121°C/103kPa

Part 2)

Sorbose	2.5g	(2.5% w/v)
m <sub>q</sub> H <sub>2</sub> O	up to 50ml	

- Autoclaved for 15min at 121°C/103kPa
- Aseptically combined parts 1 & 2 then aseptically added:

Protease Inhibitor Cocktail (Sigma, cat#P8340)	100µl	(1:200)
--	-------	---------

**2.1.10.5 Layer agar**

Vogel's 50x stock (2.1.10.16)	2ml	(1x)
Bacto-agar	0.8g	(0.8% w/v)
Sucrose	2g	(2% w/v)
mqH <sub>2</sub> O	up to 100ml	

- Autoclaved for 15min at 121°C/103kPa
- Kept at 60°C until used

**2.1.10.6 Microconidiation medium**

SC, 2x stock (2.1.10.11)	50 ml	(1x)
Sucrose	0.5 g	(0.5% w/v)
Bacto-agar	2g	(2% w/v)
Supplements (2.1.10.14)	As required	
mqH <sub>2</sub> O	up to 100ml	

- Autoclaved for 15min at 121°C/103kPa
- Cooled to 60°C then aseptically added:

Sodium iodoacetic acid (0.1M)	60µl	(60µM)
-------------------------------	------	--------

- Dispensed 6ml of sterile medium to sterile 16 x 150mm tubes and placed on a slope to cool. Tubes were plugged with sterile non-absorbent cotton wool.

**2.1.10.7 Neurospora secretion liquid medium**

Part 1)

Vogel's 50x stock (2.1.10.16)	2ml	(1x)
Bacto-yeast extract	0.5g	(0.5% w/v)
Sucrose	8g	(8% w/v)
Supplements (2.1.10.14)	As required	
mqH <sub>2</sub> O	up to 50ml	

- Autoclaved for 15min at 121°C/103kPa

Part 2)

Sorbose	2.5g	(2.5% w/v)
mqH <sub>2</sub> O	up to 50ml	

- Autoclaved for 15min at 121°C/103kPa
- Aseptically combined parts 1 & 2 then aseptically added:

DMSO	2ml	(2% v/v)
Protease Inhibitor Cocktail (Sigma, cat#P8340)	100µl	(1:200)

**2.1.10.8 Plate medium**

Vogel's 50x stock (2.1.10.16)	2ml	(1x)
Bacto-agar	2g	(2% w/v)
Supplements (2.1.10.14)	As required	
m <sub>q</sub> H <sub>2</sub> O	up to 100ml	
<ul style="list-style-type: none"> <li>Autoclaved for 15min at 121°C/103kPa, then aseptically added:</li> </ul>		
SGF 20x (2.1.10.12)	5ml	(1x)
<ul style="list-style-type: none"> <li>Aseptically poured into petri dishes and briefly flamed surface.</li> </ul>		

**2.1.10.9 Protease plate assay medium**

Sorbose	2g	(2% w/v)
Glucose	50mg	(0.05% w/v)
Gelatine	1g	(1% w/v)
Casein (10% w/v)	10ml	(1% v/v)
Bacto-agar	1.5g	(1.5% w/v)
Protease salts (50x) (2.1.10.10)	2ml	(1x)
Supplements (2.1.10.14)	As required	
m <sub>q</sub> H <sub>2</sub> O	up to 100ml	
<ul style="list-style-type: none"> <li>Adjusted pH to 6.0-7.0</li> <li>Autoclaved for 15min at 121°C/103kPa.</li> <li>Cooled to 50°C before use.</li> </ul>		

**2.1.10.10 Protease salt solution, 50x**

NaCl	12.5g	(2.1M)
KH <sub>2</sub> PO <sub>4</sub>	5g	(368mM)
CaCl <sub>2</sub> .2H <sub>2</sub> O	0.5g	(34mM)
Biotin (10mg/ml)	25µl	(10.2µM)
m <sub>q</sub> H <sub>2</sub> O	up to 100ml	
Chloroform	200µl	(0.2% v/v)
<ul style="list-style-type: none"> <li>Stored in a dark bottle at RT</li> </ul>		

**2.1.10.11 SC medium, 2x**

KNO <sub>3</sub>	2g	(20mM)
K <sub>2</sub> HPO <sub>4</sub>	1.4g	(8mM)
KH <sub>2</sub> PO <sub>4</sub>	1g	(7.4mM)
MgSO <sub>4</sub> .7H <sub>2</sub> O	2g	(8mM)
NaCl	200mg	(3.4mM)
CaCl <sub>2</sub> .2H <sub>2</sub> O	200mg	(1.36mM)
Biotin (10mg/ml)	1ml	(40µM)
Trace element solution (2.1.10.15)	200µl	(0.02% v/v)
m <sub>q</sub> H <sub>2</sub> O	up to 1L	
Chloroform	2ml	(0.2% v/v)
<ul style="list-style-type: none"> <li>Stored at 4°C in a dark bottle</li> </ul>		

**2.1.10.12 SGF medium, 20x**

Sorbose	20g	(20% w/v)
Glucose	500mg	(0.5% w/v)
Fructose	500mg	(0.5% w/v)
mqH <sub>2</sub> O	up to 100ml	

- Sterilised by autoclaving for 15min at 121°C/103kPa
- Stored at 4°C

**2.1.10.13 Slope medium**

Vogel's 50x stock (2.1.10.16)	2ml	(1x)
Sucrose	2g	(2% w/v)
Bacto-agar	2g	(2% w/v)
Supplements (2.1.10.14)	As required	
mqH <sub>2</sub> O	up to 100ml	

- Medium was melted by autoclaving for 5min at 121°C/103kPa and 1.5ml dispensed per 13 x 100mm tube. Tubes were plugged with non-absorbent cottonwool.
- Autoclaved 15min at 121°C/103kPa and placed on a slope of ~80° to cool.

**2.1.10.14 Supplements**

L-histidine	20mg/100 ml of solution
Hygromycin	200µg/ml of solution

**2.1.10.15 Trace element solution**

Citric acid (monohydrate)	5g	(200mM)
ZnSO <sub>4</sub> .7H <sub>2</sub> O	5g	(174mM)
Fe(NH <sub>4</sub> ) <sub>2</sub> (SO <sub>4</sub> ) <sub>2</sub> .6H <sub>2</sub> O	1g	(25.5mM)
CuSO <sub>4</sub> .5H <sub>2</sub> O	250mg	(10mM)
MnSO <sub>4</sub> .1H <sub>2</sub> O	50mg	(3mM)
H <sub>3</sub> BO <sub>3</sub>	50mg	(8mM)
mqH <sub>2</sub> O	up to 100ml	
Chloroform	1ml	(1% v/v)

- Stored at 4°C in dark bottle



**2.1.10.16 Vogel's N medium, 50x**

Trisodium citrate (dihydrate)	130g	(442mM)
KH <sub>2</sub> PO <sub>4</sub>	250g	(1.8M)
NH <sub>4</sub> NO <sub>3</sub>	100g	(1.25M)
MgSO <sub>4</sub> ·7H <sub>2</sub> O	10g	(40mM)
CaCl <sub>2</sub> ·2H <sub>2</sub> O (dissolved in 50ml MqH <sub>2</sub> O)	5g	(106mM)
Trace element solution (2.1.10.15)	5ml	(0.5% v/v)
Biotin	25mg	(102µM)
mqH <sub>2</sub> O	900ml	
• Dissolve ingredients in order listed		
mqH <sub>2</sub> O	up to 1L	
Chloroform	2ml	(0.2% v/v)
• Stored at RT in dark bottle		

**2.1.10.17 Vogel's liquid medium for mycelia growth**

Vogel's 50x stock (2.1.10.16)	2ml	(2% v/v)
Sucrose	2g	(2% w/v)
Supplements (2.1.10.14)	As required	
mqH <sub>2</sub> O	up to 100ml	
• Autoclaved for 15min at 121°C/103kPa		

**2.1.10.18 Vogel's liquid medium for protein secretion**

Vogel's 50x stock (2.1.10.16)	2ml	(2% v/v)
Sucrose	8g	(8% w/v)
Supplements (2.1.10.14)	As required	
mqH <sub>2</sub> O	up to 100ml	
• Autoclaved for 15min at 121°C/103kPa then added:		
Protease Inhibitor Cocktail (Sigma, cat#P8340)	100µl	(1:200)

## 2.1.11 Bacterial culture media

### 2.1.11.1 Luria Broth liquid medium

Bacto-tryptone	10g	(1% w/v)
Bacto-yeast extract	5g	(0.5% w/v)
NaCl	5g	(85.6mM)
mqH <sub>2</sub> O	up to 1L	

- Adjusted pH to 7.0
- Autoclaved for 15min at 121°C/103kPa
- Supplemented with antibiotic when required
- Stored at 4°C

### 2.1.11.2 Luria Broth agar medium

Bacto-tryptone	10g	(1% w/v)
Bacto-yeast extract	5g	(0.5% w/v)
NaCl	5g	(85.6mM)
Bacto-agar	15g	(1.5% w/v)
mqH <sub>2</sub> O	up to 1L	

- Adjusted pH to 7.0
- Autoclaved for 15min at 121°C/103kPa
- Supplemented with antibiotic when required
- Poured ~20ml into sterile 90mm Petri dishes and allowed to set
- Stored at 4°C

### 2.1.11.3 SOC liquid medium

Bacto-tryptone	20g	(2% w/v)
Bacto-yeast extract	5g	(0.5% w/v)
NaCl	0.5g	(8.56mM)
KCl (1M)	2.5ml	(2.5mM)
MgSO <sub>4</sub> (1M)	10ml	(10mM)
MgCl <sub>2</sub> (1M)	10ml	(10mM)
Glucose (1M)	20ml	(20mM)
mqH <sub>2</sub> O	up to 1L	

- Sterilised by autoclaving for 15min at 121°C/103kPa
- Stored at RT

### **2.1.12 Mammalian cell culture media**

DMEM , penicillin & streptomycin cocktail (100x) and geneticin (50mg/ml) were sourced from Invitrogen and fetal calf serum from JRH Biosciences.

#### **2.1.12.1 Cell culture medium for thawing P6 cells**

DMEM	392.5ml	
Fetal calf serum	100ml	(20% v/v)
Penicillin & streptomycin cocktail (100x)	5ml	(1x)
Geneticin (50mg/ml)	4ml	(0.4mg/ml)

- Stored at 4°C

#### **2.1.12.2 Cell culture medium for culturing P6 cells**

DMEM	442.5ml	
Fetal calf serum	50ml	(10% v/v)
Penicillin & streptomycin cocktail (100x)	5ml	(1x)
Geneticin (50mg/ml)	4ml	(0.4mg/ml)

- Stored at 4°C

#### **2.1.12.3 Cell culture medium for thawing R<sup>-</sup>IRA & R<sup>-</sup>IRB cells**

As per 2.1.12.2

#### **2.1.12.4 Cell culture medium for culturing R<sup>-</sup>IRA & R<sup>-</sup>IRB cells**

DMEM	467.5ml	
Fetal calf serum	25ml	(5% v/v)
Penicillin & streptomycin cocktail (100x)	5ml	(1x)
Geneticin (50mg/ml)	2.5ml	(0.25mg/ml)

- Stored at 4°C

## 2.2 Methods

### 2.2.1 Neurospora methods

#### 2.2.1.1 Neurospora culture methods

##### 2.2.1.1.1 Small scale macroconidial slope cultures

*Neurospora* stocks were cultured on 1.5ml of slope medium (2.1.10.13). Slope tubes were aseptically inoculated with a wet-loop of macroconidia and incubated at 25°C in 12hr light/dark conditions for 3-4 days. Tubes were wrapped in plastic and stored at -20°C.

##### 2.2.1.1.2 Large scale macroconidial cultures

Sterile 2L flasks containing 100ml of agar slope medium (2.1.10.13) were aseptically inoculated with 500µl of sterile water inoculated with a wet-loop of macroconidia from *Neurospora* stocks (2.1.2). Flasks were plugged with sterile cottonwool bungs and the macroconidial suspension dispersed over the agar surface by slanting and slowly rotating flask. Flasks were incubated at 25°C in 12hr light/dark conditions for 7-9 days.

##### 2.2.1.1.3 Liquid cultures for mycelia growth

Sterile 100ml flasks containing 20ml of Vogel's liquid medium (2.1.10.17) were aseptically inoculated with a wet loop of macroconidia from *Neurospora* stocks. Flasks were plugged with sterile cottonwool bungs and placed in an orbital shaker and incubated at 25°C in dark conditions at 160OPM for 2-3 days.

**2.2.1.1.4** *Liquid cultures for protein secretion*

Sterile, acid-rinsed 125ml flasks containing 25ml of Vogel's liquid medium (2.1.10.1.18), *Neurospora* secretion liquid medium (2.1.10.7) or FRIES based liquid media (2.1.10.3 or 2.1.10.1.4) were aseptically inoculated with a wet loop of macroconidia from *Neurospora* stocks. Flasks were plugged with sterile cottonwool bungs and placed in an orbital shaker and incubated at 25°C in dark conditions at 160OPM. Flasks were typically incubated for up to 96 hours.

**2.2.1.1.5** *Culturing Neurospora ascospores*

Harvested ascospores (2.2.1.2.3) were heat-shocked for 45min at 60°C in a waterbath, added to layer agar (2.1.10.5), vortexed briefly and immediately poured onto Vogel's agar plates (2.1.10.8) and incubated in dark conditions at 25°C for 1-2 days before moving to 34°C for an additional 1-2 days. Single, well isolated colonies were then picked to slope tubes (2.2.1.1.1).

**2.2.1.2** *Neurospora genetic techniques*

**2.2.1.2.1** *Stock culturing from long-term storage*

*Neurospora* strains used in this project (2.1.2) were sourced from long-term storage stocks in the Catcheside laboratory in the School of Biological Sciences, Flinders University. Approximately 5-10 silica gel granules with adhering macroconidia were transferred to an agar slope and incubated at 25°C in 12hr light/dark conditions for 3-4 days. A wet loop of macroconidia was streaked onto a fully supplemented plate (2.1.10.8) and single colonies were isolated and used to inoculate agar slopes (2.2.1.1.1). Macroconidia from these slopes were used to test relevant genotypes (2.2.1.2.2).

2.2.1.2.2 Genotype testing of Neurospora strains

Macroconidia isolated from single colonies (2.2.1.2.1) were tested as described below:

Gene	Test setup	Wildtype phenotype	Mutant phenotype
Mating type	Crossed with mating type tester strains T2326 and T2327	Formation of perithecia with T2326 indicated tested stock was <i>A</i> .  Formation of perithecia with T2327 indicated tested stock was <i>a</i>	
<i>his-3</i>	Streaked macroconidia onto plates without L-histidine and then onto plates with L-histidine	Growth on minimal media	Growth on L-histidine supplemented media only
<i>cot-1</i>	Spread plated macroconidia onto plates and grew at 25°C and 34°C	Normal colony morphology at 25°C and 34°C	Colonial morphology at 34°C and normal colony morphology at 25°C
<i>eth-1</i>	Streaked macroconidia onto plates without ethionine and then onto plates with ethionine	No growth in presence of ethionine	Growth in presence of ethionine
<i>exp-1</i>	Added macroconidia to casein /gelatine layer agar and poured onto supplemented plates (as per 2.2.5.1.3)	Large colony halos on casein/gelatine protease plate assay	Small colony halos on casein/gelatine protease plate assay

2.2.1.2.3 *Crossing Neurospora strains*

2.2.1.2.3.1 *Plate crosses*

Fifty millilitres of autoclaved plate crossing medium (2.1.10.1.1) were poured into sterile 90mm glass Petri dishes and allowed agar to solidify. A sterile Whatman #1 filter disc (90mm) was placed on top of the agar as a carbon source. A wet loop of macroconidia to be used as female parent was placed into 1ml of sterile water and entire volume was gently spread plated onto filter disc. Plate was incubated at 25°C for ~5 days in dark conditions before inoculating male parent as described above for the female parent. Incubation at 25°C in dark conditions was continued for up to 4 weeks.

Ascospores were harvested at various time points by aseptically removing and replacing glass Petri dish lid. Ascospores were washed from the lid by pipetting 1ml of sterile mqH<sub>2</sub>O onto surface of lid and transferring ascospore suspension into 1.5ml microfuge tubes. Stored ascospores at 4°C and continued incubating plates at 25°C in dark conditions.

2.2.1.2.3.2 *Tube crosses*

Five millilitres of tube crossing medium (2.1.10.1.2) were aliquoted into 16 x 150mm tubes containing 6x4cm Whatman #1 filter papers folded lengthwise into a W shape. Tubes were plugged with non-absorbent cottonwool and autoclaved for 15min at 121°C/103kPa. Crossing medium was inoculated with a wet loop of macroconidia from both parent strains and the macroconidia were mixed and spread evenly over the filter paper by inclination and rotation of the tube. Crosses were incubated at 25°C for up to 4 weeks with a 12hr light/dark cycle.

Ascospores were harvested by adding 2ml of sterile mqH<sub>2</sub>O to the crossing tube, vortexing for 30s and pouring the contents through a sterile muslin filter into a sterile 2ml microfuge tube. Ascospores were stored at 4°C.

2.2.1.2.4 Isolation of homokaryons using microconidiation

- Adapted from Ebbole (1990)

Microconidiation slopes (2.1.10.6) supplemented with L-histidine (2.1.10.14) were inoculated with macroconidia from a single isolated colony (note: for hygromycin selection, slopes were also supplemented with hygromycin (2.1.10.14)). Microconidiation slopes were incubated at 25°C for 9-12 days with a 12hr light/dark cycle. Microconidia were harvested by adding 2ml of sterile mqH<sub>2</sub>O and vortexing for 30s. Macroconidia and mycelia fragments were removed from the suspension by filtration using Millipore Millex Durapore 5µm filters attached to 3ml syringes. The filtrate was collected in a sterile 2ml centrifuge tube and the microconidia pelleted by centrifugation for 5min at 20,000g at RT. The supernatant was removed and the pellet resuspended in 250µl sterile water. Aliquots of 200µl & 50µl were spread plated on Vogel's plates with L-histidine (2.1.10.8) and hygromycin if required for antibiotic selection. Plates were incubated in dark conditions at 25°C for 3-4 days before moving to 34°C for an additional 2-3 days.

Macroconidia from three primary transformants were inoculated into test tubes with Westergaard and Mitchell synthetic crossing medium (SC), iodoacetate, L-histidine and hygromycin supplemented agar slopes (2.1.10.6). Iodoacetate is required for increased production of microconidia (Rossier *et al.*, 1973) through inhibition of glycolysis and upregulation of the pentose shunt pathway (Maheshwari 1999). Microconidiation slopes were incubated at 25°C for 9-12 days with a 12hr light/dark cycle. Microconidia were harvested by adding 2ml of sterile mqH<sub>2</sub>O and vortexing for 30s. Macroconidia and mycelia fragments were removed from the suspension by filtration using Millipore Millex Durapore 5µm filters attached to 3ml syringes. The filtrate was collected in a sterile 2ml centrifuge tube and the microconidia pelleted by centrifugation for 5min at 20,000g at RT. The supernatant was removed and the pellet resuspended in 250µl sterile water. Aliquots of 200µl & 50µl were spread plated on Vogel's plates with L-histidine (2.1.10.8) and hygromycin if required for antibiotic selection. Plates were incubated in dark conditions at 25°C for 3-4 days before moving to 34°C for an additional 2-3 days.



For hygromycin selection, 3 colonies from the spread plates were picked onto agar slopes (2.2.1.1.1) supplemented with L-histidine and hygromycin and incubated at 25°C for 3 days with a 12hr light/dark cycle. Macroconidia from an agar slope were then streaked onto half a Vogel's plate supplemented with L-histidine only and incubated at 25°C for 3 days in dark conditions. 6 isolates from the streak plates were picked onto Vogel's plates supplemented with L-histidine and hygromycin and incubated at 25°C for 3 days in dark conditions. Isolates that were hygromycin resistant in all 6 test colonies were considered to be homokaryons.

For histidine selection, 16 colonies were picked onto agar slopes supplemented with L-histidine and incubated at 25°C for 3 days with a 12hr light/dark cycle. Isolation of homokaryons with the desired *his-3* allele was determined by complementation testing (2.2.1.2.5).

Although microconidia are predominately uninucleate, a small percentage of microconidia contain multiple nuclei (Baylis 1967). Thus, both heterokaryon and homokaryon microconidia are capable of growing on selective plates and require complementation testing to ensure transformed homokaryons have been obtained.

#### *2.2.1.2.5 Complementation testing*

*his-3* targeted transplacements resulted in His<sup>+</sup> transformants that were subsequently resolved into His<sup>-</sup> monokaryons by microconidiation (2.2.1.2.4) and subjected to complementation testing with the recipient strain to ensure they were homokaryons containing the vector-borne *his-3* (K26) allele and not the *his-3* (K458) allele present in the recipient strain. A wet-loop with a relatively low concentration of macroconidia from the 16 colonies isolated after microconidiation were patched onto non-tester minimal plates (2.1.10.8) and then onto tester minimal plates inoculated with macroconidia from the recipient strain (t11630 or t11644) spread plated at a concentration of 100 macroconidia /mm<sup>2</sup>. Plates were incubated at 25°C in dark conditions for 4-6 days. Isolates with strong growth on the tester plates but little or no growth on the nontester plates were selected as homokaryons with the desired *his-3* (K26) allele.

### 2.2.1.3 *Neurospora molecular techniques*

#### 2.2.1.3.1 *Neurospora electroporation*

- Adapted from Vann (1995)

Macroconidia were harvested from a 9-12 day old culture (2.2.1.1.2) by the addition of 60ml of 1M sorbitol and gently swirled so macroconidia were suspended in solution without disrupting the agar. Macroconidial suspension was poured through a sterile muslin filter into a 50ml falcon tube and centrifuged at 2,800g for 5min at 4°C. The pellet was washed three times with 50ml of 1M sorbitol and centrifuged between washes at 2,800g for 5min at 4°C. The pellet was initially resuspended in 500µl of 1M sorbitol and the number of macroconidia counted with a haemocytometer. The volume of the macroconidial suspension was subsequently adjusted to give  $3 \times 10^9$  macroconidia/ml.

Ten microlitres of linearised plasmid DNA (>500ng/µl) was added to 30µl of washed macroconidia in a 1.5ml microfuge tube and stood on ice for at least 5min. The macroconidial/DNA suspension was transferred to a prechilled 0.2 cm electroporation cuvette (Biorad GenePulser/*E. coli* Pulser cuvette, cat# 165-2086) and electroporated in a BioRad Gene Pulser (BioRad) with the following settings: voltage gradient: 1.5 kV/cm; capacitance: 50 uF; resistance: 200 ohms.

For His<sup>+</sup> selection, cuvette contents were transferred to 400µl of 1M sorbitol in a 1.5ml microfuge tube at RT and pipette mixed. One hundred microlitres of electroporated suspension was spread plated onto four minimal plates (2.1.10.8) and incubated at 25°C for 2-4 days before being transferred to 34°C for an additional 2-3 days.

For hygromycin selection, cuvette contents were transferred to 400µl of Vogel's liquid medium (2.1.10.17) supplemented with L-histidine (2.1.10.14) in a 1.5ml microfuge tube at RT and pipette mixed. The electroporated suspension was incubated at 25°C for 3hrs before 100µl of suspension was spread plated onto four plates supplemented with L-histidine and hygromycin and incubated at 25°C for 2-4 days before being transferred to 34°C for an additional 2-3 days. Transformants were subsequently microconidiated to obtain homokaryons (2.2.1.2.4).

2.2.1.3.2 Small scale preparation of *Neurospora* genomic DNA from macroconidia for PCR

- Adapted from Henderson *et al.* (2005)

A wet loop of 3 - 7 day old macroconidia from a slope tube (2.2.1.1.1) was transferred to 100µl of sterile 1x TE (2.1.9.29) in a 1.5ml microfuge tube and vortexed briefly. The macroconidial suspension was boiled in a waterbath for 10min and then placed on ice for 5min. The cellular debris was pelleted by centrifugation at 20,000g for 5min at RT. Seventy microlitres of supernatant was transferred to a new 1.5ml microfuge tube and used directly in a PCR reaction (2.2.4.6.4) or stored at –20°C.

2.2.1.3.3 Mid scale preparation of *Neurospora* genomic DNA

- Adapted from Ireland (1993)

Mycelium was harvested by pouring liquid culture (2.2.1.1.3) through Whatman #1 filter paper in a Büchner funnel mounted on a Büchner flask under vacuum for ~1-2min. Mycelium was removed from the filter paper and ground to a fine powder in liquid nitrogen using a mortar and pestle.

Ground mycelium was transferred to a 2ml microfuge tube and 1ml of *Neurospora* genomic DNA extraction buffer (2.1.9.16) was added. Ground mycelia was resuspended with gentle pipetting and incubated at 65°C for 10min. Six hundred microlitres of 7.5M ammonium acetate was added to the suspension and mixed by tube inversion. The tube was left on ice for 10min before centrifugation at 20,000g for 5min at RT. 750µl of supernatant was transferred to 2x 1.5ml microfuge tubes containing 525µl of isopropanol, the tubes were mixed by inversion and centrifuged at 20,000g for 5min at RT. The supernatant was aspirated and the pellets air dried in a heating block at 50°C for ~20min.

The pellets were resuspended in 125µl of 1x TE (2.1.9.29), 5µl of 10mg/ml RNase A was added and the tubes were incubated at RT for ~30min. The duplicate suspensions were pooled together in a single 1.5ml microfuge tube containing 440µl of DNase free

mqH<sub>2</sub>O. One volume (700µl) of phenol was added to the suspension, mixed by tube inversion and centrifuged at 20,000g for 2 minutes at RT. The top layer (~600µl) was transferred to a new 1.5ml microfuge tube, 1 volume (600µl) of water saturated chloroform was added, mixed by tube inversion and centrifuged at 20,000g for 2 min at RT. The top layer (~500µl) was transferred to a new 2ml tube containing 2.5 volumes (1250µl) of ice cold 99% ethanol, mixed by tube inversion and incubated from 1hr to overnight at -20°C. Samples were centrifuged at 20,000g for 5min at 4°C, supernatant was aspirated, tubes were inverted on paper towel and air dried for ~30min. Genomic DNA pellets were resuspended in 100µl of 1x TE (2.1.9.29).

## **2.2.2 Bacterial methods**

### **2.2.2.1 *E. coli* overnight culture**

*E. coli* strains were streaked or spread plated on LB agar plates (2.1.11.2) and incubated at 37°C for ~16hr. Single, well isolated colonies were selected and typically grown in 3ml LB medium (2.1.11.1) supplemented with 100µg/ml of ampicillin in a 10ml centrifuge tube and incubated at 37°C for ~16hr with vigorous shaking.

### **2.2.2.2 Preparation of competent *E. coli* cells**

#### **2.2.2.2.1 Preparation of chemically competent *E. coli* cells**

JM109 or DH5α *E. coli* strains were streaked onto an LB agar plate (2.1.11.2) and incubated at 37°C for ~16hr. A well isolated colony was used to inoculate 5ml of LB medium (2.1.11.1) and grown at 37°C for ~16hr with vigorous shaking. One hundred millilitres of LB medium in a 1L flask was inoculated with 5ml of ~16hr culture and incubated at 37°C with vigorous shaking for ~4hr or until the absorbance at 600nm reached 0.4-0.5. The cells were chilled on ice for ~1hr and pelleted by centrifugation at 4°C for 15min at 2,800g. The supernatant was decanted and the pellet resuspended in 10ml of chilled TSS buffer (2.1.9.30). 100µl of the cell suspension was aliquoted into sterile, prechilled 1.5ml microfuge tubes and snap frozen in liquid nitrogen. Competent cells were stored at -80°C.

**2.2.2.2.2 Preparation of electrocompetent *E. coli* cells**

DHB10 cells were streaked onto an LB agar plate (2.1.11.2) and incubated at 37°C for ~16hr. A well isolated colony was used to inoculate 10ml of LB medium (2.1.11.1) and grown at 37°C for ~16hr with vigorous shaking. Five hundred millilitres of LB medium in a 1L flask was inoculated with 5ml of ~16hr culture and incubated at 37°C with vigorous shaking for ~4hrs or until the absorbance at 600nm reached 0.4-0.5.

The cells were chilled on ice for ~1hr and pelleted by centrifugation at 4°C for 15min at 2,800g. The supernatant was decanted and the pellet resuspended in 25ml of sterile mqH<sub>2</sub>O by vortexing before centrifugation at 4°C for 5min at 2,800g. The supernatant was decanted and the pellet resuspended in 25ml of sterile 10% (v/v) glycerol by vortexing before centrifugation at 4°C for 5min at 2,800g. Removed supernatant by pipetting and resuspended pellet in 10% (v/v) glycerol to a total volume of ~2ml. 100µl of the cell suspension was aliquoted into sterile, prechilled 1.5ml microfuge tubes and snap frozen in liquid nitrogen. Competent cells were stored at -80°C.

**2.2.2.3 Heat shock transformation of competent *E. coli* cells**

**2.2.2.3.1 Heat shock transformation of competent *DH5α*, *JM101* & *JM109* cells**

Typically, 1-10µl of transforming DNA was added to 100µl of thawed competent cells (2.2.2.2.1) on ice and gently mixed by swirling with a pipette tip. Cells were incubated on ice for 30min, heated to 42°C in a waterbath for 45s then placed on ice for 2min. One millilitre of LB medium (2.1.11.1) was added and the cells incubated for 45-90min at 37°C with vigorous shaking. Aliquots of 150µl and 300µl were spread plated onto LB agar plates (2.1.11.2) supplemented with 100µg/ml ampicillin and incubated for ~16hr at 37°C.

**2.2.2.3.2 Heat shock transformation of competent TOP10F' cells**

A 50µl aliquot of competent cells was thawed on ice and 25µl was transferred to prechilled 1.5ml microfuge tubes on ice. Typically, 1-10µl of transforming DNA was added to 25µl of thawed competent cells and gently mixed by swirling with a pipette tip. Cells were incubated on ice for 30 min, heated to 42°C in a waterbath for 30s then placed on ice for 2 min. Two hundred and fifty microlitres of SOC medium (2.1.11.3) was added and the cells incubated for 60min at 37°C with vigorous shaking. Aliquots of 50µl and 100µl were spread plated onto LB agar plates (2.1.11.2) supplemented with 100µg/ml ampicillin and incubated for ~16hr at 37°C.

**2.2.2.4 Electroporation of competent DHB10 E. coli cells**

A 100µl aliquot of competent cells (2.2.2.2.2) was thawed on ice and 50µl was transferred to prechilled 1.5ml microfuge tubes on ice. Typically, 1-2µl of transforming DNA was added to 50µl of electrocompetent cells and gently mixed by swirling with a pipette tip. The DNA/cell suspension was transferred to prechilled 0.2 cm electroporation cuvette (Biorad GenePulser/E. coli Pulser cuvette, cat# 165-2086) and electroporated in a BioRad Gene Pulser (BioRad) with the following settings: voltage gradient: 1.8 kV/cm; capacitance: 25 uF; resistance: 200 ohms. The cells were immediately transferred to 200µl of LB medium (2.1.11.1) in 1.5ml microfuge tubes and incubated at 37°C for 45-90min with vigorous shaking. Aliquots of 50µl and 100µl were spread plated onto LB agar plates (2.1.11.2) supplemented with 100µg/ml ampicillin and incubated for ~16hr at 37°C.

**2.2.2.5 Preparation of glycerol stocks**

Seven hundred microlitres of an overnight *E. coli* culture (2.2.2.1) was added to 300µl of sterile 50% v/v glycerol in a 2ml cryogenic tube (Nunc, cat#264300) and stored at -80°C.

#### **2.2.2.6 Plasmid DNA purification (miniprep)**

All small scale plasmid preparations were done using an Adbiotech Plasmid Purification Kit (Adbiotech, cat#PPK-1) as per the manufacturer's instructions. 1.5ml of overnight culture (2.2.2.1) was processed and the plasmid DNA was eluted in 50 $\mu$ l of sterile mqH<sub>2</sub>O heated to 65°C.

#### **2.2.2.7 Plasmid DNA purification (midiprep)**

All medium scale plasmid preparations were done using a Qiagen Midiprep Kit (Qiagen, Australia, cat#12143) as per the manufacturer's instructions. 25ml of overnight culture (2.2.2.1) was processed and the plasmid DNA pellet was resuspended in 200 $\mu$ l of sterile mqH<sub>2</sub>O heated to 65°C.

### **2.2.3 Mammalian cell culture methods**

#### **2.2.3.1 Thawing mammalian cell lines stored in liquid nitrogen**

A cell vial stored in liquid nitrogen was thawed by gently swirling the tube in a 37°C waterbath before 1ml of cell culture medium (2.1.12.1 or 2.1.12.3) was added. The cell suspension was transferred to 10ml of cell culture medium in a 15ml centrifuge tube. Cells were centrifuged at 200g for 5min at RT and the supernatant aspirated. The cell pellet was resuspended in 1ml of cell culture medium and transferred to 20ml of cell culture medium in a sterile T75 tissue culture flask. The flask was incubated at 37°C with 5% CO<sub>2</sub> until ~80% confluent.

#### **2.2.3.2 Passaging of mammalian cell lines**

Adherent cell monolayers in tissue culture flasks were grown to ~80% confluency and then washed once with 1x PBS (diluted from 2.1.9.18) before treatment with trypsin (1ml per 75cm<sup>2</sup>) for 5min at 37°C. Detached cells were resuspended in cell culture medium (2.1.12.2 or 2.1.12.4), transferred to a sterile 15ml centrifuge tube and gently pipette mixed to disrupt any clumps. Typically a 1:30 split of P6 cells or a 1:60 split of R1RA and R1RB cells was transferred to 20ml of medium in a T75 or 30ml of medium in a T175 tissue culture flask. The flask was incubated at 37°C with 5% CO<sub>2</sub> until ~80% confluent.

### **2.2.3.3 Determination of cell number**

Viable cell counts were performed with trypsinised cells (2.2.3.2) from a T175 tissue culture flask made to a total volume of 10ml with cell culture medium (2.1.12.2 or 2.1.12.4). Twenty microlitres of cell suspension was added to 180 $\mu$ l of trypan blue (diluted 1:10 in 1x PBS (diluted from 2.1.9.18) and counted using a cytometer and a phase contrast microscope (Olympus, CK2).

### **2.2.3.4 Storage of mammalian cell lines**

Cells were grown to ~80% confluency in T175 flasks, trypsinised and cell culture medium added (2.2.3.2). Cells were centrifuged at 200g for 5 minutes at RT. The trypsin/medium was aspirated, the cells resuspended in 4ml of freezing solution (2.1.9.12) and 1ml aliquoted into Cryotube vials (Nunc, cat# 366656). The vials were kept at -80°C for two days and then stored in liquid nitrogen.

### **2.2.3.5 Preparing mammalian cell lysates for receptor binding assay**

Cells grown to ~80% confluency were washed with sterile 1xPBS (diluted from 2.1.9.18) and serum starved by incubating for 4hr in DMEM minimal medium at 37°C, 5% CO<sub>2</sub>. Cells were washed twice with 1xPBS and 11ml of cell lysis buffer (2.1.9.4) containing 1/1000 dilution of Protease Inhibitor Cocktail (Sigma, cat#P8340) was added to the flask and incubated at 4°C for 1hr. The lysate was transferred to a 15ml centrifuge tube and centrifuged at 2,800g for 5min at 4°C. The supernatant was pooled together if multiple flasks were lysed and aliquoted for use or storage at -20°C.



## 2.2.4 DNA methods

### 2.2.4.1 DNA Electrophoresis

A 13.5 x 25cm horizontal tank powered by a Biorad Power Pac 200 was used for DNA electrophoresis. Typically, agarose gels ranging from 0.7% - 1.5% (w/v) were made with 1x TAE buffer (diluted from 2.1.9.25) and 10µg/ml ethidium bromide. DNA fragments were electrophoresed through the TAE gels at voltages between 1.2 – 6 V/cm. Gels were visualized on a UV transilluminator at 302nm. The size of DNA fragments were estimated by comparison of their relative mobility to known size markers (Table 2.1).

### 2.2.4.2 DNA restriction digestion

All restriction enzymes used for DNA digestion were from NEB. Typically, 300-600ng of DNA was digested with 2-5 units of enzyme for 1 to 24hr under conditions specified by the manufacturer. Digestion reactions were terminated by heat inactivation where possible or by processing through a spin column (2.2.4.3.2).

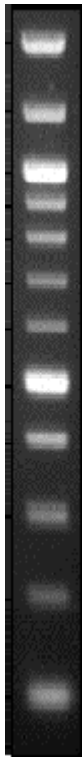



### 2.2.4.3 Purification of DNA fragments

#### 2.2.4.3.1 Gel extraction of DNA

Gel extraction of DNA fragments were performed using a Qiagen QIAquick Gel Extraction Kit (Qiagen, cat# 28704) or a Qiagen Minelute Gel Extraction Kit (Qiagen, cat# 28604). Both kits were used as per the manufacturer's instructions.

#### 2.2.4.3.2 Spin column purification of DNA

Spin column purification of DNA and PCR amplicons were performed using an Abdiotech PCR DNA Cleanup Kit (Abdiotech, cat# PCK-1) as per the manufacturer's instructions.

100bp ladder	1kb ladder	2-log ladder	SPP1/EcoRI ladder																																																																																																																				
 <table border="1"> <thead> <tr> <th>bp</th> <th>ng</th> </tr> </thead> <tbody> <tr><td>1517</td><td>45</td></tr> <tr><td>1200</td><td>35</td></tr> <tr><td>1000</td><td>95</td></tr> <tr><td>900</td><td>27</td></tr> <tr><td>800</td><td>24</td></tr> <tr><td>700</td><td>21</td></tr> <tr><td>600</td><td>18</td></tr> <tr><td>500/517</td><td>97</td></tr> <tr><td>400</td><td>38</td></tr> <tr><td>300</td><td>29</td></tr> <tr><td>200</td><td>25</td></tr> <tr><td>100</td><td>48</td></tr> </tbody> </table>	bp	ng	1517	45	1200	35	1000	95	900	27	800	24	700	21	600	18	500/517	97	400	38	300	29	200	25	100	48	 <table border="1"> <thead> <tr> <th>kb</th> <th>ng</th> </tr> </thead> <tbody> <tr><td>10.0</td><td>42</td></tr> <tr><td>8.0</td><td>42</td></tr> <tr><td>6.0</td><td>50</td></tr> <tr><td>5.0</td><td>42</td></tr> <tr><td>4.0</td><td>33</td></tr> <tr><td>3.0</td><td>125</td></tr> <tr><td>2.0</td><td>48</td></tr> <tr><td>1.5</td><td>36</td></tr> <tr><td>1.0</td><td>42</td></tr> <tr><td>0.5</td><td>42</td></tr> </tbody> </table>	kb	ng	10.0	42	8.0	42	6.0	50	5.0	42	4.0	33	3.0	125	2.0	48	1.5	36	1.0	42	0.5	42	 <table border="1"> <thead> <tr> <th>kb</th> <th>ng</th> </tr> </thead> <tbody> <tr><td>10.0</td><td>40</td></tr> <tr><td>8.0</td><td>40</td></tr> <tr><td>6.0</td><td>48</td></tr> <tr><td>5.0</td><td>40</td></tr> <tr><td>4.0</td><td>32</td></tr> <tr><td>3.0</td><td>120</td></tr> <tr><td>2.0</td><td>40</td></tr> <tr><td>1.5</td><td>57</td></tr> <tr><td>1.2</td><td>45</td></tr> <tr><td>1.0</td><td>122</td></tr> <tr><td>0.9</td><td>34</td></tr> <tr><td>0.8</td><td>31</td></tr> <tr><td>0.7</td><td>27</td></tr> <tr><td>0.6</td><td>23</td></tr> <tr><td>0.5</td><td>124</td></tr> <tr><td>0.4</td><td>49</td></tr> <tr><td>0.3</td><td>37</td></tr> <tr><td>0.2</td><td>32</td></tr> <tr><td>0.1</td><td>61</td></tr> </tbody> </table>	kb	ng	10.0	40	8.0	40	6.0	48	5.0	40	4.0	32	3.0	120	2.0	40	1.5	57	1.2	45	1.0	122	0.9	34	0.8	31	0.7	27	0.6	23	0.5	124	0.4	49	0.3	37	0.2	32	0.1	61	 <table border="1"> <thead> <tr> <th>bp</th> <th>ng</th> </tr> </thead> <tbody> <tr><td>8557</td><td>97</td></tr> <tr><td>7427</td><td>84</td></tr> <tr><td>6106</td><td>69</td></tr> <tr><td>4899</td><td>56</td></tr> <tr><td>3639</td><td>41</td></tr> <tr><td>2799</td><td>32</td></tr> <tr><td>1953</td><td>22</td></tr> <tr><td>1882</td><td>21</td></tr> <tr><td>1515</td><td>17</td></tr> <tr><td>1412</td><td>16</td></tr> <tr><td>1164</td><td>13</td></tr> <tr><td>992</td><td>11</td></tr> <tr><td>710</td><td>8</td></tr> </tbody> </table>	bp	ng	8557	97	7427	84	6106	69	4899	56	3639	41	2799	32	1953	22	1882	21	1515	17	1412	16	1164	13	992	11	710	8
bp	ng																																																																																																																						
1517	45																																																																																																																						
1200	35																																																																																																																						
1000	95																																																																																																																						
900	27																																																																																																																						
800	24																																																																																																																						
700	21																																																																																																																						
600	18																																																																																																																						
500/517	97																																																																																																																						
400	38																																																																																																																						
300	29																																																																																																																						
200	25																																																																																																																						
100	48																																																																																																																						
kb	ng																																																																																																																						
10.0	42																																																																																																																						
8.0	42																																																																																																																						
6.0	50																																																																																																																						
5.0	42																																																																																																																						
4.0	33																																																																																																																						
3.0	125																																																																																																																						
2.0	48																																																																																																																						
1.5	36																																																																																																																						
1.0	42																																																																																																																						
0.5	42																																																																																																																						
kb	ng																																																																																																																						
10.0	40																																																																																																																						
8.0	40																																																																																																																						
6.0	48																																																																																																																						
5.0	40																																																																																																																						
4.0	32																																																																																																																						
3.0	120																																																																																																																						
2.0	40																																																																																																																						
1.5	57																																																																																																																						
1.2	45																																																																																																																						
1.0	122																																																																																																																						
0.9	34																																																																																																																						
0.8	31																																																																																																																						
0.7	27																																																																																																																						
0.6	23																																																																																																																						
0.5	124																																																																																																																						
0.4	49																																																																																																																						
0.3	37																																																																																																																						
0.2	32																																																																																																																						
0.1	61																																																																																																																						
bp	ng																																																																																																																						
8557	97																																																																																																																						
7427	84																																																																																																																						
6106	69																																																																																																																						
4899	56																																																																																																																						
3639	41																																																																																																																						
2799	32																																																																																																																						
1953	22																																																																																																																						
1882	21																																																																																																																						
1515	17																																																																																																																						
1412	16																																																																																																																						
1164	13																																																																																																																						
992	11																																																																																																																						
710	8																																																																																																																						
NEB (cat# N3231s)	NEB (cat#N3232s)	NEB (cat#N3200s)	Geneworks (cat#DMW-S1)																																																																																																																				

**Table 2.1 DNA size markers used in this project.** Mass of DNA in bands when 1µl of DNA ladder is loaded on a gel. Biotinylated versions of 2-log (NEB, cat#N7554s) and SPP1/EcoRI (Geneworks, cat#DMW-S1) DNA ladders were used in Southern blots (2.2.4.7)

#### **2.2.4.4 Dephosphorylation of DNA fragments**

Calf intestinal phosphatase (CIP) (NEB, cat# M0290S) was used to dephosphorylate restricted DNA fragments (2.2.4.2) with complementary ends. The CIP was added directly to the heat inactivated digestion reaction or if column purified, buffer and CIP were added. CIP reactions typically incubated DNA with 1000U of phosphatase at 37°C for 1hr. Reactions were terminated by spin column purification (2.2.4.3.2).

#### **2.2.4.5 Ligation of DNA fragments**

DNA fragments with complementary ends were ligated with T4 DNA ligase (Promega, cat# M180A). Ligations were carried out as per the manufacturer's instructions with a typical ligation using 3 Weiss Units per reaction in a total volume of 10µl with a 1:3 molar ratio of vector: insert DNA as calculated by the following formula:

$$\text{ng of insert required} = (\text{ng of vector} \times \text{kb size of insert}) / \text{kb size of vector} \times \text{molar ratio of insert/vector}$$

Ligations were incubated for one hour at RT or overnight at 4°C.

#### **2.2.4.6 Polymerase chain reaction**

All PCR amplifications were performed with either an Applied Biosystems 2720 or a Perkin Elmer GeneAmp 2400 thermal cycler using 0.2ml PCR tubes.

##### **2.2.4.6.1 Oligonucleotide design**

All PCR, sequencing and other oligonucleotides used in this project are listed in section 2.1.8. Primers were designed manually and where possible included a 3' GC clamp, had ~50% GC content and a T<sub>m</sub> between 60-62°C. Potential secondary structures and dimer formation of candidate primers were determined using the NETPRIMER program (<http://www.premierbiosoft.com/netprlaunch.html>).

#### 2.2.4.6.2 General PCR

PCR amplification typically used the following reaction conditions: 200µM dNTPs (NEB, cat# N7552S), 1x reaction buffer IV (Abgene, cat# AB-0406b), 1.5mM MgCl<sub>2</sub> (Abgene, cat# AB-0406b), 200nM of each primer (Geneworks), 1.25-2.5U of RedHot DNA polymerase (Abgene, cat# AB-0406b) and 1-10ng of template DNA in a total reaction volume of 50µl. Thermal cycling was typically 94°C, 5min, followed by 35 cycles of 94°C, 45s; 58°C, 30s; 72°C at 1min per kb of amplicon; then 72°C, 7min before finishing at 4°C.

#### 2.2.4.6.3 E. coli colony PCR

Colony PCR was typically done as a preliminary screen of plated transformants (2.2.2.3 & 2.2.2.4). PCR reaction and cycle conditions were as described in 2.2.4.6.2 except template DNA was supplied by using a sterile 200µl pipette tip to scrape a bacterial colony and the tip was gently scraped onto a premarked LB agar plate (2.1.11.2) supplemented with 100µg/ml ampicillin before placing the tip into the PCR tube with the reaction mix. The tip was gently stirred to mix the PCR reagents and bacterial cells before the tip was discarded. The inoculated plate was incubated at 37°C until the PCR cycle was completed and the amplicons electrophoresed (2.2.4.1) to determine positive transformants to be used for inoculating overnight cultures (2.2.2.1).

#### 2.2.4.6.4 Neurospora macroconidia PCR

- Adapted from Henderson *et al.* (2005)

Two to five microlitres of macroconidial DNA 'Boil-prep' solution (2.2.1.3.2) was used as a template in a 50µl PCR reaction with 0.5U of Red Hot<sup>®</sup> DNA polymerase as per 2.2.4.6.2.

#### **2.2.4.7 Southern blot**

##### **2.2.4.7.1 DNA transfer**

DNA was transferred from agarose gels to Hybond N+ nylon membranes (Amersham, cat#RPN303B) as follows. The gels were exposed to UV light for 30s to nick the DNA before transfer. The gels were soaked in 250mM HCl for 8min to partially depurinate the DNA. The gels were rinsed in water then the DNA was denatured by soaking the gel in denaturation solution (2.1.9.7) for 30min with gentle shaking. The denaturation solution was drained and the gels soaked twice in neutralisation solution (2.1.9.17) for 15min with gentle shaking. The membranes were cut to the size of the gels and soaked in 2x SSC solution (diluted from 2.1.9.22). DNA was transferred onto membranes by standard capillary transfer for 12-18hrs using 2x SSC transfer buffer. The membranes were dried by placing in a 80°C incubator for 15min before crosslinking the DNA to the membranes by exposure to 50mJ of UV radiation at 314nm using a GS genelinker™ UV chamber (Biorad).

##### **2.2.4.7.2 Probe labelling**

Probe was made by random biotinylated labelling of linearised plasmid DNA or column purified PCR product using the NEBlot™ Phototope™ Kit (NEB, cat#N7550S) as per the manufacturer's instructions.

##### **2.2.4.7.3 Probe hybridization**

All hybridization and washing steps were done in a Bartelt Xtron Hi-2002 incubator. The membranes were placed in sterile hybridisation bottles and prehybridised with sufficient ULTRAhyb (Ambion, cat#8669) to keep the membranes uniformly wet at 42°C with gentle rotation. The biotinylated DNA probe (2.2.4.7.2) was denatured by boiling for 5min before chilling on ice for 3min. Approximately 1-5ng of denatured probe was added per ml of ULTRAhyb solution and the membranes were hybridised for approximately 12hrs at 42°C with gentle rotation. The membranes were washed twice in stringency wash solution A (2.1.9.23) for 5min at RT with gentle rotation and then washed twice with stringency wash solution B (2.1.9.24) for 15min at 42°C with gentle rotation.

2.2.4.7.4 *Probe detection*

The biotinylated probe was detected with the Phototope™-Star Detection Kit (NEB, cat# 7020) with the following modifications to the manufacturer's instructions. The membranes were washed once with ~100ml of blocking solution (2.1.9.1) and incubated for 10min at RT with gentle rotation and then drained.

The volumes of streptavidin to be added to the membranes were determined with the formula: volume of streptavidin (ml) = size of membranes (cm<sup>2</sup>) x 0.05ml/cm<sup>2</sup>. The stock streptavidin was diluted 1:1000 with blocking solution and the appropriate volume added to the membranes, incubated for 5min at RT with gentle rotation and then drained. The membranes were washed 3x with wash solution I (2.1.9.31) for 15s, 15min and 15min at RT with gentle rotation.

The volumes of biotinylated alkaline phosphatase (BAP) to be added to the membranes were determined with the formula: volume of BAP (ml) = size of membranes (cm<sup>2</sup>) x 0.05ml/cm<sup>2</sup>. The stock BAP was diluted 1:1000 with blocking solution and the appropriate volume added to the membranes, incubated for 5min at RT with gentle rotation and then drained. The membranes were washed 3x with wash solution II (diluted to 1x from 2.1.9.32) for 15s, 15min and 15min at RT with gentle rotation.

The volume of CDP-star reagent to be added to the membranes was determined with the formula: volume of CDP-star reagent (ml) = size of membranes (cm<sup>2</sup>) x 0.015ml/cm<sup>2</sup>. The stock CDP-star reagent was diluted 1:100 with 1x CDP-star diluent. The membranes were transferred from the hybridization bottle to an A4 plastic sheet protector, the appropriate volume was added to the membranes and the plastic sheet protector was carefully lowered onto the membranes. Any airbubbles in the CDP-star reagent were pushed off the membrane surface by gently rubbing the plastic cover and the membranes was incubated for 5min at RT. The membranes were removed and the CDP-star reagent drained before transferring the membranes to a plastic sleeve in a transfer cassette. X-ray film was loaded into the cassette, exposed for 10-30min and the film detected using a Kodak X-OMAT 1000 film processor.

**2.2.4.8 DpnI site directed mutagenesis**

- Adapted from Fisher (1997)

Two complimentary oligonucleotides were designed with the desired mutation(s) flanked by unmodified DNA sequence (2.2.4.6.1). The primers included silent mutations to create *NgoMIV*/*NaeI* restriction sites to facilitate mutant screening. The primers were HPLC purified by the manufacturer and used in a thermal cycling reaction with the vector containing the coding sequence of interest in the following reaction conditions: 350µM dNTPs (NEB, cat# N7552S), 1x Pfx Amplification Buffer, 2mM MgSO<sub>4</sub>, 200nM of each primer (Geneworks), 2.5U of Platinum Pfx DNA Polymerase (Invitrogen, cat# 11708-013) and 10ng of template DNA in a total reaction volume of 50µl. Thermal cycling conditions were 95°C, 30s, followed by 18 cycles of 95°C, 30s; 52/55/58°C, 1min; 68°C at 3.5min then 20°C, 2min. This reaction generated nicked circular double stranded DNA containing the mutations of interest.

After thermal cycling, 10µl of the reaction was removed for agarose gel electrophoresis and 20U of *DpnI* restriction enzyme (NEB, cat#RO176) was added to the remaining DNA solution and incubated at 37°C for ~1hr to digest the methylated plasmid DNA leaving nicked circular mutated DNA. After digestion, 10µl was removed for agarose gel electrophoresis and 2µl of the DNA solution was transformed into TOP10F' *E. coli* cells by heat shock and spread plated onto LB agar plates supplemented with 100µg/ml ampicillin (2.2.2.3.2). The *E. coli* host cells repair the nicked double stranded DNA forming a mutated DNA plasmid. Colony PCR of the putative transformed *E. coli* colonies was performed (2.2.4.6.3) and the amplicons were subsequently restricted with *NgoMIV* (2.2.4.2) to screen for the presence of mutated plasmid DNA. Final confirmation of site-directed mutagenesis and general sequence integrity was achieved by DNA sequencing of the regions of interest in the mutated plasmid (2.2.4.9).

#### **2.2.4.9 DNA sequencing**

DNA sequencing was performed by the Flinders Sequencing Facility at the Flinders Medical Centre using an Applied Biosystems 3100 Genetic Analyser with BigDye Terminator v3.1 chemistry as per the service provider's instructions.

##### **2.2.4.9.1 DNA sequence analysis and manipulation**

DNA sequences (2.2.4.9) were compiled and edited using Sequencher 4.6 (Genecodes). Vector maps were constructed using Clone Manager 7.03 (Sci Ed Central). Bioinformatic analyses were performed with the suite of programs available in Biomanager ([www.angis.org.au](http://www.angis.org.au)). The *Neurospora crassa* genome database was accessed at <http://www.broad.mit.edu/annotation/genome/neurospora/Home.html>.

#### **2.2.5 Protein methods**

##### **2.2.5.1 Neurospora protein expression**

###### **2.2.5.1.1 Harvesting of secreted protein**

Supernatant from *Neurospora* liquid culture (2.2.1.1.4) was removed by pipetting 1ml aliquots into 1.5ml microfuge tubes with 1/1000 dilution of Protease Inhibitor Cocktail (Sigma, cat#P8340). Samples not used immediately for subsequent experimental work were snap frozen in liquid nitrogen and stored at -80°C.

###### **2.2.5.1.2 Isolation of intracellular Neurospora protein**

Mycelium from liquid culture (2.2.1.1.4) was harvested by pouring liquid culture through Whatman #1 filter paper in a Büchner funnel mounted on a Büchner flask under vacuum for 2-3s. Mycelium was transferred to 1.5ml microfuge tube up to 100µl graduation on tube. Three hundred microlitres of 1x NuPAGE sample buffer (2.2.5.5) was added before boiling the sample for 10min in a waterbath and centrifuging at 20,000g for 3min at RT. The supernatant was transferred to a new 1.5ml microfuge tube with 1/1000 dilution of Protease Inhibitor Cocktail (Sigma, cat#P8340) and centrifuged at 20,000g for 3min at RT. Samples were immediately used in SDS-PAGE (2.2.5.5) and western blot (2.2.5.6) analysis and the remaining sample snap frozen in liquid nitrogen and stored at -80°C.



#### 2.2.5.1.3 Neurospora protease plate assay

A wet loop of macroconidia was added to 500µl of sterile mqH<sub>2</sub>O and the macroconidia concentration was determined using a haemocytometer. The macroconidial suspension was diluted to 200 macroconidia/ml and 100µl was added to 10ml of protease plate assay medium in a falcon tube kept at 50°C (2.1.10.9). The tube was mixed by inversion and poured into a 150mm Petri dish and the macroconidial suspension dispersed over entire plate surface by gentle shaking. The plate was placed in a large plastic bag, ~40ml of sterile mqH<sub>2</sub>O was added to the bottom of the bag and the top of the bag folded loosely and taped down. The plate was incubated at 25°C for 6-7 days in dark conditions. Protease activity was determined by relative colony halo size.

#### 2.2.5.2 E. coli protein expression

Note: at each purification stage, the protein concentration and/or integrity was determined by analytical HPLC (2.2.5.9.1).

##### 2.2.5.2.1 Induction of protein expression in E. coli

One hundred microlitres of *E. coli* JM101 cells (2.2.2.3.1) were transformed with pGH(1-11)VNPAPM-C1, pGH(1-11)VNPAPM-C2 & pGH(1-11)VNPAPM-C3 *E. coli* expression vectors (2.1.4) and plated onto an LB agar plate (2.1.11.2) supplemented with 100µg/ml ampicillin and incubated at 37°C for ~16hr. A single colony was used to inoculate 5ml of LB medium (2.1.11.1) supplemented with 100µg/ml ampicillin and incubated for ~7hrs at 37°C with vigorous shaking. Five hundred microlitres was subcultured into each of 4 x 50ml of LB medium supplemented with 100µg/ml ampicillin and incubated at 37°C with vigorous shaking for ~16hr. Each 50ml culture was used to inoculate one of 4 x 500ml of prewarmed LB medium supplemented with 100µg/ml ampicillin and incubated with vigorous shaking at 37°C until the absorbance at 600nm reached 0.6. One millilitre of culture was taken as a pre-induction sample for SDS-PAGE analysis (2.2.5.5.1) and then 0.1mM IPTG was added to 500ml cultures to induce protein expression. The induced cultures were incubated at 37°C with vigorous shaking for ~5hrs. One millilitre of culture was taken as a post-induction sample for

SDS-PAGE analysis and the cells were pelleted by centrifugation at 7,000g for 15min at 4°C and the supernatant decanted. Pellets were stored at -80°C until used. Successful expression was determined by SDS-PAGE and rpHPLC analysis (2.2.5.9.1) of pre- and post-induction samples.

#### 2.2.5.2.2 *French press lysis of E. coli*

Pelleted *E. coli* cells (2.2.5.2.1) were resuspended in a total volume of 80 ml of inclusion body wash buffer (2.1.9.14). The *E. coli* cells were lysed by 2 passes through the French Press at 45-60MPa. Lysed cells were centrifuged at 9,000g for 15min at 4°C and the supernatant decanted. The pellet was resuspended in 5% (w/v) inclusion body wash buffer and incubated at RT for 15min before centrifugation at 4,000g for 15min at 4°C. The supernatant was decanted and the pellet resuspended in 5% (w/v) inclusion body wash buffer and centrifuged at 2,800g for 10min at 4°C. The centrifuge tube was weighed before and after pellet processing to determine the weight of IBs and the IB pellet was stored at -80°C.

#### 2.2.5.2.3 *Purification and processing of IGF-II mutants from E. coli inclusion bodies*

##### 2.2.5.2.3.1 *Gel filtration chromatography*

Note: all solutions were filtered (1µm) before use and a flow rate of 1ml/min was used for all washes, loading and running of the column.

An XK-26 column (Length: 100cm, internal diameter 2.6cm) (Pharmacia) packed with 400ml of sephacryl S-200 resin (Pharmacia) was connected to an FPLC (2.2.5.8). The column was washed with 0.5 column volume of mqH<sub>2</sub>O, sanitized with 0.5 column volume of 0.5M NaOH, washed with 0.5 column volume of mqH<sub>2</sub>O before being equilibrated with 1 column volume of inclusion body dissolution buffer (2.1.9.13) with 1.6mM DTT.

IBs were dissolved in IB dissolution buffer with 20mM DTT at a final concentration of 10ml/g (wet weight) of IBs for 30min at RT before loading on to the equilibrated gel filtration column. The inclusion body dissolution buffer with 1.6mM DTT was used as a running buffer and the elution profile was monitored using a UV detector at 280nm.

5ml fractions were collected and alternate fractions were analysed by SDS-PAGE (2.2.5.5.1).

#### 2.2.5.2.3.2 *Protein refolding*

The dissolved IB solution was diluted 1:4 with refold dilution buffer (2.1.9.19) to create the following refold conditions: a protein concentration of 200 $\mu$ g (or less)/ml, 2M Urea, 0.775M Tris, 10mM glycine, 37.5mM EDTA and 0.4mM DTT at pH 9.1. Refolding was initiated by the addition of 1mM 2-Hydroxyethylidisedisulphide (Sigma, cat#1892-29-1) as an oxidizing agent to promote formation of disulphide bonds. The refolding reaction was incubated for 120min at RT with gentle stirring and the status of the refolded isomers was determined at various time points by analytical rpHPLC (2.2.5.9.1). The refolding reaction was terminated by acidification with concentrated HCl to pH 2.5.

#### 2.2.5.2.3.3 *Ion exchange chromatography*

Note: all solutions were filtered (1 $\mu$ m) before use.

An XK-16 column (Length: 20cm, internal diameter 1.6cm) (Pharmacia) packed with 15ml of Fast Flow S (FFS) cation exchange resin (Pharmacia) was connected to an FPLC (2.2.5.8). The column was washed with 1 column volume of mqH<sub>2</sub>O, sanitized with 1 column volume of 0.5M NaOH, washed with 10 column volumes of mqH<sub>2</sub>O before being charged with 10 column volumes of 0.5M acetic acid.

The refolded protein was filtered (1 $\mu$ m) and loaded onto the column at a flow rate of 8 ml/min. The column was then washed with FFS buffer A (2.1.9.10) until all unbound protein was removed and a stable baseline on the chart recorder was reached. The bound protein was eluted with FFS buffer B (2.1.9.11).

#### 2.2.5.2.3.4 *Alpha-lytic protease cleavage*

The IGF-II mutant protein was cleaved from the porcine leader sequence by a mutant of  $\alpha$ -lytic protease called Prag A9 (Lien *et al.*, 2001). This enzyme cleaves between the methionine of the PAPM linker and alanine which is the first residue of mature human IGF-II. The cleavage reaction was performed at a protein concentration of 0.2 mg (or less) /ml in 2M urea, 0.1M Tris, pH 8 at 37°C for up to 215min. The Prag A9 enzyme (obtained from Wallace laboratory, Adelaide University) was added at a ratio of 1:500 fusion protein:enzyme. Monitoring of the cleavage reaction was by rpHPLC (2.2.5.9.1) and the reaction was terminated was by acidification to pH 2.5 with concentrated HCl. Final purification of IGF-II mutants was achieved by preparative rpHPLC (refer 2.2.5.9.2).

### 2.2.5.3 *Protein quantitation*

#### 2.2.5.3.1 *Bradford assay*

Bradford assays were performed with Bradford reagent (Sigma, cat#B6916) using BSA for standard curves as per the manufacturer's instructions.

#### 2.2.5.3.2 *BCA assay*

BCA assays were performed with the BCA Protein Assay Kit (Pierce, cat# 23227) using BSA for standard curves as per the manufacturer's instructions.

#### 2.2.5.3.3 *Quantitation of IGF-II mutants by rpHPLC*

Lyophilised aliquots of mutant IGF-II protein were resuspended in 1x PBS (diluted from 2.1.9.18) and pooled together. Triplicate samples of the pool were acidified with 0.1% (v/v) TFA and analysed by rpHPLC (2.2.5.9.1) with a 30 minute 15-80 % (v/v) acetonitrile gradient. The area under the peak on the rpHPLC profile was compared to a reference standard of receptor grade, human IGF-I standard (2.1.7). To account for the extinction coefficient differences between the IGF-I standard and the mutant IGF-II proteins, a correction value (Q-value relative to IGF-I) was determined for the three mutant proteins (refer to Table 2.2).

Protein	Q-value relative to IGF-I
IGF-I	1
IGF-II	0.9659
C1	0.9981
C2	0.9805
C3	0.9901

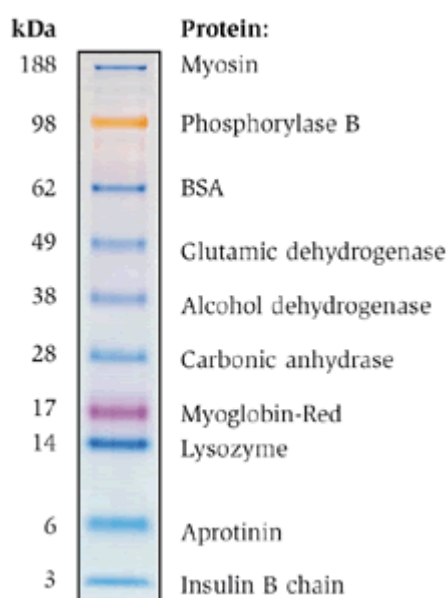
**Table 2.2 Q-value of all IGF-II mutant proteins relative to IGF-I.** Q-values relative to IGF-I for the three IGF-II mutant proteins were used to determine the protein concentration from analytical rpHPLC profiles as described in Buck *et al.* (1989).

#### **2.2.5.4 TCA precipitation of proteins**

Ten microlitres (1/100 vol) of 2% (w/v) sodium deoxycholate (Sigma, cat#D-6750) was added to 1ml of *Neurospora* supernatant (2.2.1.1.4) in a 1.5ml microfuge tube. The solution was vortexed briefly and incubated for 30min at 4°C. 100µl (1/10 vol) of 100% (w/v) TCA (2.1.9.28) was added and the solution incubated overnight at 4°C. The solution was centrifuged at 20,000g for 15min at 4°C. The supernatant was aspirated, the pellets washed twice with 1ml (1 vol) of ice cold acetone by vortexing the pellet and centrifuging samples at 20,000g for 5min at 4°C between washes. Pellets were dried by tube inversion on paper towel for ~15min. Pellets were resuspended in 40µl of 1x NuPAGE sample buffer for SDS-PAGE (2.2.5.5.2). Note: if yellowing of the NuPAGE buffer occurred due to acidification by residual TCA, then the samples were titrated with 1M NaOH until sample buffer turned blue.

### 2.2.5.5 SDS PAGE

All SDS-PAGE experiments were performed using the XCell SureLock Minicell pre-cast gel system (Invitrogen, cat#E10001) with 12% Bis-Tris NuPAGE gels (cat#NP0343). All required buffers and solutions were sourced from Invitrogen. Specifically, 4x NuPAGE LDS sample buffer (cat#NP0007), 10x NuPAGE sample reducing agent (1M DTT) (cat#NP0004), 20x NuPAGE MES SDS running buffer (cat#NP0002) and NuPAGE antioxidant (cat#NP0005). All gels included a lane with 12 $\mu$ l of SeeBlue Plus2 pre-stained protein standard (Invitrogen, cat#LC5925) (refer to Fig. 2.1). The gel apparatus was assembled and run under reducing conditions as per the manufacturer's instructions. Gels were typically run at 160V for ~1hr or until the dye front had electrophoresed off the end of the gel.



**Fig 2.1 SeeBlue Plus2 pre-stained standard (Invitrogen, cat#LC5925)**

2.2.5.5.1 Preparation of *E. coli* induction samples for SDS-PAGE

The 1ml pre- and post-induction samples from each of the 4 flasks (2.2.5.2.1) were centrifuged at 20,000g for 1min at RT and the supernatant discarded. Each pellet was resuspended in *E. coli* lysis buffer (2.1.9.9) at a volume of 50µl per unit OD<sub>600nm</sub> to enable the relative amount of protein expression between flasks to be determined. 2µl of 4x NuPAGE sample buffer was added to 6µl of resuspended sample, incubated for 5min at 95°C in a heating block, centrifuged at 20,000g for 15s at RT and then 8µl of sample was loaded onto a 12% Bis-Tris NuPAGE gel and run under standard reducing conditions (2.2.5.5).

2.2.5.5.2 Preparation of *Neurospora* samples for SDS-PAGE

Harvested *Neurospora* supernatant (2.2.5.1.1) was either processed immediately after culturing or thawed on ice from -80°C stocks. To 19.5µl of *Neurospora* supernatant in a 1.5ml microfuge tube, 7.5µl of prewarmed 4x NuPAGE sample buffer and 3µl of 10x reducing agent was added. The solution was flick mixed and incubated for 5min at 95°C in a heating block before centrifugation at 20,000g for 15s at RT. 18µl of prepared sample was loaded onto a 12% Bis-Tris NuPAGE gel and receptor grade human IGF-I protein (2.1.7) was used as a positive control where 2ng and 0.5ng total protein were typically loaded and the gel was run under standard reducing conditions (2.2.5.5).

2.2.5.5.3 Staining & destaining SDS-PAGE gels

Pre-cast NuPAGE gels were removed from their cases after electrophoresis (2.2.5.5) and placed in Coomassie Blue staining solution (2.1.9.6) for ~1hr at RT with gentle shaking.

Destaining solution (2.1.9.8) was added to stained NuPAGE gels for ~30min before being discarded and replaced with fresh destaining solution. A small foam block was placed in the tray to absorb the Coomassie stain as it leached from the gel.

#### 2.2.5.6 **Western blot**

Western blots were performed using the XCell SureLock Minicell pre-cast gel system (Invitrogen, cat#E10001) with XCell II blot module (Invitrogen, cat#E19051). 20x transfer buffer (cat#NP0006) and antioxidant (cat#NP0005) were also sourced from Invitrogen. The transfer apparatus was assembled and run under reducing conditions as per the manufacturer's instructions.

The electrophoresed gel was transferred onto Hybond Nylon-C extra membrane (Amersham, cat#RPN303E) at 30V for 1hr in transfer buffer with 10% (v/v) methanol (1 gel) or 20% (v/v) methanol (2 gels). The transfer apparatus was disassembled and the membrane blocked in blotto (2.1.9.2) for 1hr at RT with shaking or overnight at 4°C without shaking. The blotto was drained, the membrane was transferred to a plastic sheet protector, 0.5ml of primary antibody (2.1.6) diluted 1:100 in blotto was added to the membrane and the plastic sheet protector was carefully lowered onto the membrane. Any air bubbles in the primary antibody solution were pushed off the membrane surface by gently rubbing the plastic cover. The membrane was incubated with primary antibody for 1hr at RT or overnight at 4°C. The membrane was transferred to a tray and washed 3x in blotto for 5min each.

The membrane was transferred to a new plastic sheet protector, 0.5ml of appropriate secondary antibody (2.1.6) diluted 1:2000 in blotto was added to the membrane and the plastic sheet protector was carefully lowered onto the membrane. Any air bubbles in the secondary antibody solution were pushed off the membrane surface by gently rubbing the plastic cover. The membrane was incubated with secondary antibody for 30min at RT. The membrane was transferred to a tray and washed 3x in 1x TBST (2.1.9.27) for 2.5min, 5min and 5min.

The membrane was transferred to a new plastic sheet protector, 0.5ml of Lightning Chemiluminescence reagent *Plus* (Perkin Elmer, cat#NEL103) was added to the membrane and the plastic sheet protector was carefully lowered onto the membrane. Any air bubbles in the ECL reagent were pushed off the membrane surface by gently



rubbing the plastic cover. The membrane was incubated with ECL reagent for 3min at RT.

The membrane was removed from the sheet protector and the ECL reagent drained by holding the membrane vertical over tissue. The membrane was transferred to a clear plastic sleeve in a transfer cassette. X-ray film was loaded into the cassette, exposed for 15s-15min and the film detected using a Kodak X-OMAT 1000 processor.

#### **2.2.5.7 Spectrometry**

A Wallac Victor<sup>3</sup>V 1420 Multilabel Counter (Perkin Elmer) was used for all spectrometry.

#### **2.2.5.8 FPLC**

Gel filtration (2.2.5.2.3.1) and ion exchange (2.2.5.2.3.3) columns were connected to a P-500 Pump, LCC-500 CI Controller, UV-1 monitor & controller, chart recorder and LKB Frac-100 Fraction Collector (Pharmacia).

#### **2.2.5.9 rpHPLC**

Note: all solutions were filtered (0.22 $\mu$ m) before use.

##### **2.2.5.9.1 Analytical rpHPLC**

Samples were acidified with 1% (v/v) TFA in a 1.5ml microfuge tube, centrifuged at 20,000g for 1 min at RT and supernatant transferred to HPLC bottles.

Analytical rpHPLC was performed with an 1100 series rpHPLC (Agilent) with a Zorbax (300 SB-CN), 2.1x150mm, 5 $\mu$ m column (Agilent, cat# 883700-905) using a 20-50% acetonitrile gradient over 30min. in the presence of 0.1 % (v/v) TFA (refer to 2.1.9.20 for rpHPLC buffer A and 2.1.9.21 for rpHPLC buffer B) at a flow rate of 0.5 ml/min.

#### **2.2.5.9.2 Preparative HPLC of IGF-II mutants**

The final purification step of the cleaved IGF-II mutants (2.2.5.2.3.4) was preparative HPLC to remove the cleaved porcine leader sequence and other impurities. The acidified cleavage reaction was filtered (0.22µm) and loaded onto a V-25 C4, 10µm, 300Å, 10mm x 250mm column (Grace Vydac, cat#214TP1010) connected to a Waters 510 HPLC pump, Waters 6486 tunable absorbance detector, Waters system interface module and an LKB Frac-100 Fraction Collector (Pharmacia). A 20-50% acetonitrile gradient over 220min in the presence of 0.1 % (v/v) TFA (refer to 2.1.9.20 for HPLC buffer A and 2.1.9.21 for HPLC buffer B) at a flow rate of 5 ml/min was used to separate the recombinant IGF-II protein from other protein contaminants. 5ml fractions were collected, analysed by HPLC (2.2.5.9.1) and all fractions containing IGF-II mutant protein were pooled, aliquoted, lyophilized (2.2.5.10) and quantified (2.2.5.3.3). Final purity and confirmation of protein identity was determined by mass spectrometry (2.2.5.11) and N-terminal sequencing (2.2.5.12).

#### **2.2.5.10 Lyophilisation of IGF-II mutant protein**

Pooled purified protein fractions from the preparative rpHPLC (2.2.5.9.2) were aliquoted at approximately 2µg, 10µg and 100µg in siliconised microfuge tubes. The proteins were then lyophilized using a Heto Maxidry Plus CT110 freeze dryer and stored at -20°C. A 10µg aliquot of each lyophilized protein was thawed and resuspended in 10mM HCl and accurately quantified by rpHPLC using receptor grade human IGF-I as a standard (2.2.5.3.3).

#### **2.2.5.11 Mass spectrometry of IGF-II mutant protein**

Two micrograms of lyophilized IGF-II mutant protein was analysed by an API-100 mass spectrometer (Perkin Elmer) operated by Dr. Chris Bagley at the Hanson Institute, Adelaide. The protein was resuspended in 2% (v/v) acetic acid and 0.5µg injected into the mass spectrometer.

#### **2.2.5.12 N-terminal sequencing**

One microgram of lyophilized IGF-II mutant protein was analysed by Edman degradation N-terminal sequencing of the 6 N-terminal amino acid residues using a 492 Procise Protein Sequencer (Applied Biosystems) operated by Mr Chris Cursaro at the School of Molecular and Biomedical Science, Adelaide University.

### **2.2.6 Biochemical methods**

#### **2.2.6.1 Coating of 96-well lumitrac plates with anti-receptor antibody**

Lumitrac 600 96-well plates (Greiner, cat#655074) were coated with anti-IGF-IR (24-31) or anti-IR (83-7) antibodies (2.1.6). Briefly, antibody stocks were diluted to 2.5µg/ml with 1x carbonate buffer (2.1.9.3), 100µl was added to each well and plates were incubated at 4°C overnight. Plates were washed 3x with 300µl of 1xTBST (2.1.9.27) and blocked with 200µl of 1xTBST with 0.5% (w/v) BSA. Plates were incubated for 2hr at RT before covering with aluminium foil and storage at -20°C.

#### **2.2.6.2 Competitive receptor binding assay with europium labelled IGF-I or insulin**

Note: all washes were performed with a 1296-026 Delphia automatic plate washer (Wallac).

The binding affinities of *Neurospora* produced IGF-I and purified IGF-II mutant proteins to the IRA, IRB and/or IGF-IR were determined by a competitive binding assay with europium labelled insulin or IGF-I. 96-well plates coated with anti-receptor antibodies (2.2.6.1) were washed 3x with 1xTBST (2.1.9.27). Lysates of R<sup>-</sup>IRA, R<sup>-</sup>IRB and P6 cells (2.2.3.5) were used as sources of IRA, IRB and IGF-IR respectively (2.1.5). 50µl of P6 lysate diluted 1/8 with 1x PBS (diluted from 2.1.9.18) or 100µl of undiluted R<sup>-</sup>IRA and R<sup>-</sup>IRB lysates were added to 93 wells of the coated 96-well plates and the appropriate volume of 1x PBS with 0.1% (w/v) BSA was added to the remaining 3 wells on the plates as non-specific binding controls. The plates were incubated overnight at 4°C.

Europium labelled receptor grade human IGF-I and insulin were prepared by Ms Joanne Lane (CSIRO Molecular and Health Technologies, Adelaide) and Dr Peter Hoyne (CSIRO Molecular and Health Technologies, Parkville) using a DELFIA<sup>®</sup> Eu-labelling kit (Perkin Elmer, cat#1244-302).

IGF-II analogue assays used a dilution series of unlabelled ligands in ligand binding buffer with 0.1% (w/v) BSA whereas assays of *Neurospora* produced IGF-I used undiluted *Neurospora* supernatants and a dilution series of IGF-I standard in liquid culture medium (2.2.1.1.4). Fifty microlitres of sample and standard were added to each well. Eu-ligand was diluted in ligand binding buffer (2.1.9.15) with 0.1% (w/v) BSA such that  $\sim 1 \times 10^6$  counts were present when 50 $\mu$ l of Eu-ligand was added to each well.

The plates were covered and incubated at 4°C overnight. The plates were washed 3x with 1xTBST and 2x with mqH<sub>2</sub>O before the addition of 100 $\mu$ l/well of Delphia enhancement solution (Perkin Elmer, cat#1244-105) and incubation for 30min at RT. Time-resolved fluorescence was measured using a Wallac Victor<sup>3</sup>V 1420 Multilabel Counter (Perkin Elmer) with 340nm excitation and 615nm emission filters. The data were represented as %specific binding (Bound – NSB/total bound – NSB x100) and plotted using Prism 4.00 with IC<sub>50s</sub> calculated by curve-fitting with a one-site competition model.

### **2.2.6.3 IGF-IR tyrosine phosphorylation assay of *Neurospora* secreted IGF-I protein**

- Adapted from Chen *et al.* (2003)

Note: all washes were performed with a 1296-026 Delphia automatic plate washer (Wallac).

One hundred microlitres of P6 cells (2.1.5) in tissue culture medium (2.1.12.2) were plated in a Falcon 96 well flat bottom plate at  $2.5 \times 10^4$  cells/well and grown overnight at 37°C, 5% CO<sub>2</sub>. Cells were washed with 1x PBS (diluted from 2.1.9.18) and incubated for 4 hours in (serum free) DMEM medium at 37°C, 5% CO<sub>2</sub> to facilitate internalization of ligand bound receptors. Two IGF-I standard dilution series were made with DMEM

minimal medium and FRIES liquid medium (2.1.10.3) and 100µl of diluted standards, undiluted samples or medium only were added to each well and incubated at 37°C, 5% CO<sub>2</sub> for 10min. The sample/medium was aspirated and 120µl of lysis buffer (2.1.9.5) was added to the cells and incubated at 4°C for 60min. A 96-well lumitrac plate coated with anti-IGF-IR antibody (24-31) (2.2.6.1) was thawed and washed 3x with 1x TBST (2.1.9.27). One hundred microlitres of cell lysate was added to 93 wells of the coated 96-well plate and the appropriate volume of 1x PBS with 0.1% (w/v) BSA was added to the remaining 3 wells on the plate as a non-specific binding control and the plate was incubated overnight at 4°C. The plate was washed 3x with 1xTBST and 100µl of 1:10000 diluted Eu-labelled antiphosphotyrosine antibody (PY-20) (2.1.6) was added to each well and the plate incubated at RT for 2 hours. The plate was washed 3x with 1xTBST and 6x with mqH<sub>2</sub>O before the addition of 100µl/well of Delphia enhancement solution (Perkin Elmer, cat#1244-105) and incubated for 10min at RT. Time-resolved fluorescence was measured with a Wallac Victor<sup>3</sup>V 1420 Multilabel Counter (Perkin Elmer) using 340nm excitation and 615nm emission filters. The data were plotted using Prism 4.00.

#### **2.2.6.4 Receptor phosphorylation assay of IGF-II analogues**

- Adapted from Chen *et al.* (2003)

Note: all washes were performed with a 1296-026 Delphia automatic plate washer (Wallac).

R<sup>-</sup>IRA, R<sup>-</sup>IRB and R<sup>-</sup>IGF-IR cells were used as sources of IRA, IRB and IGF-IR, respectively (2.1.5). One hundred microlitres of cells in tissue culture medium (2.1.12.4) were plated in a Falcon 96 well flat bottom plate at 8 x 10<sup>3</sup> cells/well and grown overnight at 37°C, 5% CO<sub>2</sub>. Cells were washed with 1x PBS (diluted from 2.1.9.18) and incubated for 4 hours in (serum free) DMEM medium at 37°C, 5% CO<sub>2</sub> to facilitate internalization of ligand bound receptors. Ligand dilution series were prepared using DMEM minimal medium as diluent. One hundred microlitres of diluted standards or DMEM minimal medium only were added to each well and incubated at 37°C, 5% CO<sub>2</sub> for 10min. The medium was aspirated and 120µl of lysis buffer (2.1.9.5) was added to the cells and incubated at 4°C for 60min. A 96-well lumitrac plate coated with anti-IR (83-7) or anti-

IGF-IR antibody (24-31) (2.2.6.1) was thawed and washed 3x with 1x TBST (2.1.9.27). One hundred microlitres of cell lysate was added to 93 wells of the coated 96-well plate and the appropriate volume of 1x PBS with 0.1% (w/v) BSA was added to the remaining 3 wells on the plate as a non-specific binding control and the plate was incubated overnight at 4°C. The plate was washed 3x with 1xTBST and 100µl of 1:10,000 diluted Eu-labelled antiphosphotyrosine antibody (PY-20) (2.1.6) was added to each well and the plate incubated at RT for 2 hours. The plate was washed 3x with 1xTBST and 6x with mqH<sub>2</sub>O before the addition of 100µl/well of Delphia enhancement solution (Perkin Elmer, cat#1244-105) and incubated for 10min at RT. Time-resolved fluorescence was measured with a Wallac Victor<sup>3</sup>V 1420 Multilabel Counter (Perkin Elmer) using 340nm excitation and 615nm emission filters. The data were plotted using Prism 4.00.

#### **2.2.6.5 Cell survival assay of IGF-II analogues**

Note: all washes were performed with a 1296-026 Delphia automatic plate washer (Wallac).

R<sup>-</sup>IRA, R<sup>-</sup>IRB and R<sup>-</sup>IGF-IR cells were used as sources of IRA, IRB and IGF-IR, respectively (2.1.5). One hundred microlitres of cells in tissue culture medium (2.1.12.4) were plated in a Nunc black Nunclon Delta 96 well flat bottom plate at  $2.5 \times 10^3$  cells/well and grown for 48hrs at 37°C, 5% CO<sub>2</sub>. Cells were washed with DMEM medium and incubated for 4 hours in (serum free) DMEM medium at 37°C, 5% CO<sub>2</sub> to facilitate internalization of ligand bound receptors. Ligand dilution series were prepared using DMEM + 0.1% (w/v) BSA + 5mM butyrate as diluent. The well medium was aspirated and 100µl of diluted standards or DMEM + 0.1% (w/v) BSA medium (as no butyrate control) were added to each well. The plate was incubated at 37°C, 5% CO<sub>2</sub> for ~48hrs after which 20µl of Promega CellTitre-Blue Cell Viability Assay reagent (Cat# G8081) was added to each well. The well contents were mixed by placing the plate on a plate shaker for ~20s. The plate was incubated for 2hrs at 37°C, 5% CO<sub>2</sub> and then the fluorescent signal indicative of viable cells was measured using a Wallac Victor<sup>3</sup>V 1420 Multilabel Counter (Perkin Elmer) using 550nm excitation and 615nm emission filters. The data were plotted using Prism 4.00.

## **Chapter 3      Vector construction for diversification of human IGF-I in *Neurospora***

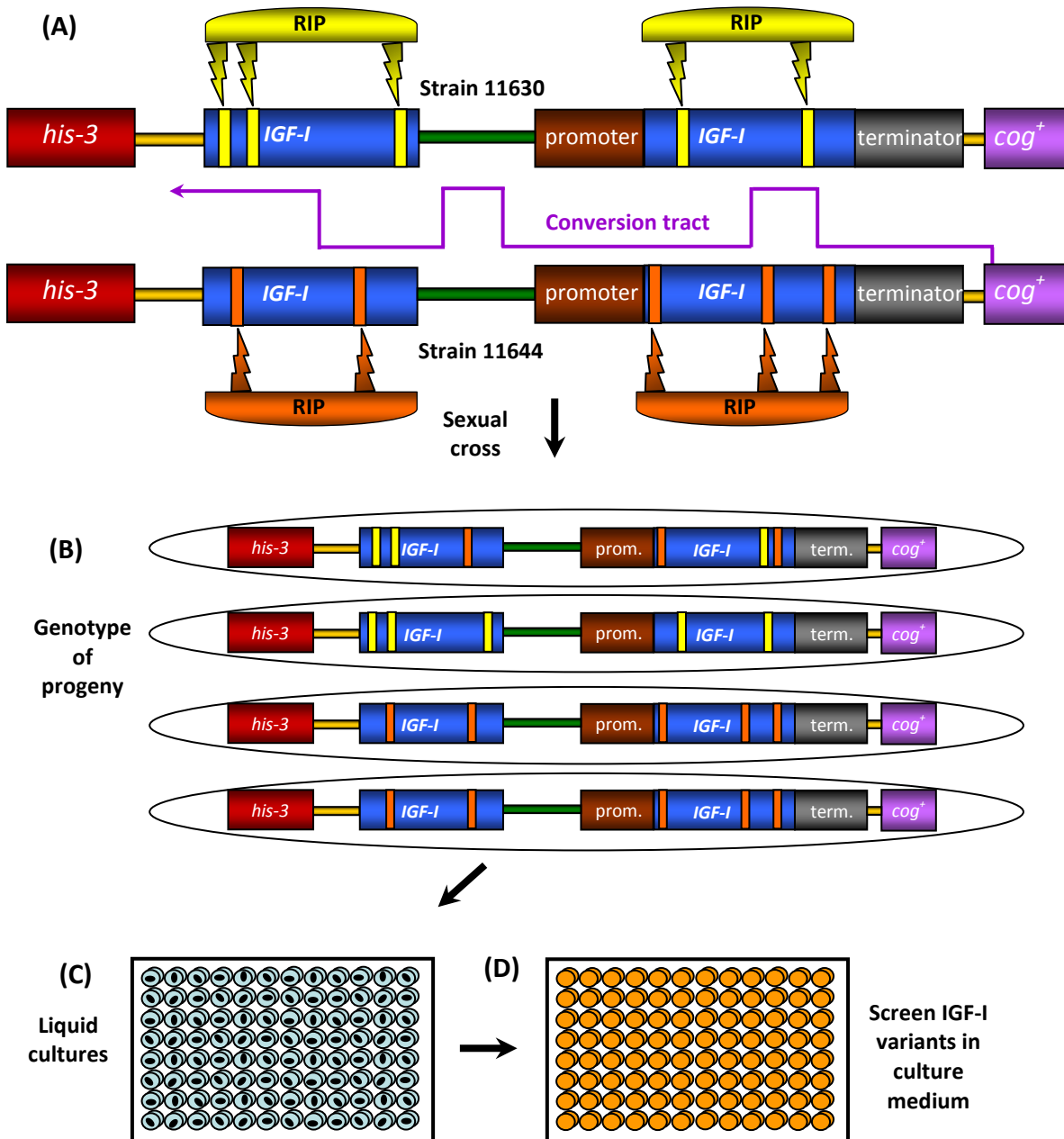
### **3.1      Introduction**

The ultimate goal of diversifying the human IGF-I ligand with the *Neurospora* directed evolution system was to generate IGF-IR antagonists as potential cancer therapeutics. To achieve this aim, an IGF-I coding sequence within a suitable expression vector was required to facilitate the generation, expression and secretion of novel IGF-I variants in *Neurospora*. This chapter describes the engineering of an IGF-I coding sequence optimised for gene diversification in *Neurospora*, the construction of a number of expression vectors to facilitate secretion of mature IGF-I protein into culture supernatant and construction of an additional IGF-I duplication vector to facilitate RIP-mutagenesis of the IGF-I coding sequence.

Although both the IGF-I and IGF-II ligands were desirable candidates for diversification, IGF-I was selected for the ongoing development of the *Neurospora* directed evolution system as this ligand had been reported to be less susceptible than IGF-II to proteolytic degradation in other heterologous expression systems (Moks *et al.*, 1987); (Hammarberg *et al.*, 1989).

#### **3.1.1.1      Strategy for RIP mediated diversification of IGF-I**

The strategy for diversification of IGF-I involved the use of RIP (permission to use the patented RIP process for gene diversification was kindly provided by Neugensis Corporation) and Hotspot Mediated Recombination (HSMR) (Figure 3.1) with a linked *IGF-I* duplication expression vector transplaced between *his-3* and *cog<sup>+</sup>* in *LGI* of *Neurospora* diversification strains t11630 and t11644. The *IGF-I* expression cassette contained promoter and terminator sequences flanking a functional *IGF-I* sequence with a signal peptide fused to the N-terminus of the IGF-I protein to facilitate secretion of the novel peptide.



**Figure 3.1 Strategy for RIP & HSMR diversification of human IGF-I in *Neurospora crassa* using linked duplications juxtaposed to the *cog*<sup>+</sup> recombination hotspot.**

(A) A linked IGF-I duplication construct transplanted between *his-3* and *cog* in diversification strains t11630 and t11644 are crossed with each other to facilitate RIP-induced mutagenesis and recombination of the duplicated IGF-I sequences. (B) Haploid ascospores containing single copies of the mutated IGF-I sequences are isolated from the cross. Linked duplications in both parents are expected to result in most progeny carrying mutated IGF-I. For simplicity, the depicted progeny genotype assumes heteroduplex formed during meiosis is corrected to the converted alleles and the correction is uniform. (C) Protein from the functional IGF-I coding sequence is secreted into liquid culture medium and (D) screened by IGF-IR receptor binding assay.



RIP differentially mutates both copies of a duplicated sequence (Grayburn & Selker 1989). As the presence of two IGF-I variants in each isolate could potentially confound the screening assay, only one of the *IGF-I* duplications contained upstream promoter or signalling elements to ensure only one IGF-I variant was expressed and secreted from each of the progeny. Crossing of the diversification strains induced pre-meiotic RIP mutagenesis of the *IGF-I* sequences in virtually all progeny due to linked *IGF-I* duplications in both parent strains. During meiosis, the relatively high frequency of recombination events initiated at *cog*<sup>+</sup> would have resulted in interrupted conversion tracts in the juxtaposed *IGF-I* sequences which would have further diversified the RIPed sequences and provided a means of separating advantageous mutations from deleterious ones generated by RIP. Liquid cultures of individual progeny were used to facilitate secretion of IGF-I variant protein into the medium to be screened by a competitive IGF-IR binding assay to identify mutants with high receptor binding affinity which would be subsequently screened by an IGF-IR phosphorylation assay to identify IGF-IR antagonists.

### **3.1.2 Optimising the *IGF-I* CDS for gene diversification in *Neurospora***

As RIP makes specific changes to a DNA sequence, it is possible to identify the spectrum of mutations that can be potentially generated in both DNA and protein sequences. The missense mutations that could occur in the mature human *IGF-I* coding sequence (Figure 3.2) extend throughout all four domains of the mature IGF-I protein (Table 3.1B) and enable a total of 95 possible amino acid substitutions in 53 of the 70 amino acids of the IGF-I protein (Table 3.1A). Substitution in 36 of these amino acids requires RIP of the less frequently mutated CpT, CpG, CpC or multiple CpX sequences, resulting in an under-representation of these mutations in the diversified IGF-I proteins. Increasing the frequency and total number of RIP-mediated missense mutations can be achieved by introducing silent mutations into the *IGF-I* coding sequence to create additional CpA targets. Although RIP optimisation of the *IGF-I* coding sequence is desirable for gene diversification, heterologous protein expression requires consideration of additional factors such as codon bias.

*CHAPTER 3 – Vector construction for diversification of human IGF-I in Neurospora*

**A)**

C	GGA	CCG	GAG	ACG	CTC	TGC	GGG	GCT	GAG	CTG	GTG	GAT	GCT	CTT	CAG
G	CCT	GGC	CTC	TGC	GAG	ACG	CCC	CGA	CTC	GAC	CAC	CTA	CGA	GAA	GTC

1	<b>REK</b>	<b>SL</b> P	<b>E</b> K	<b>M</b> T <sub>(I)</sub>	<b>F</b> L	<b>C</b> Y	<b>G</b> REK	<b>V</b> T	<b>E</b> K	<b>L</b>	<b>V</b> MI	<b>D</b> N	<b>V</b> A <sub>(I)</sub>	<b>F</b> L	<b>*</b> Q	15
---	------------	----------------	---------------	------------------------------	---------------	---------------	-----------------	---------------	---------------	----------	----------------	---------------	------------------------------	---------------	---------------	----

	TTC	GTG	TGT	GGA	GAC	AGG	GGC	TTT	TAT	TTC	AAC	AAG	CCC	ACA	GGG
	AAG	CAC	ACA	CCT	CTG	TCC	CCG	AAA	ATA	AAG	TTG	TTC	GGG	TGT	CCC

16	<b>F</b>	<b>V</b> MI	<b>C</b> Y	<b>G</b> REK	<b>D</b> N	<b>R</b> K	<b>G</b> SDN	<b>F</b>	<b>Y</b>	<b>F</b>	<b>N</b>	<b>K</b>	<b>SLF</b> P	<b>I</b> T	<b>G</b> REK	30
----	----------	----------------	---------------	-----------------	---------------	---------------	-----------------	----------	----------	----------	----------	----------	-----------------	---------------	-----------------	----

	TAT	GGC	TCC	AGC	AGT	CGG	AGG	GCG	CCT	CAG	ACA	GGC	ATC	GTG	GAT
	ATA	CCG	AGG	TCG	TCA	GCC	TCC	CGC	GGA	GTC	TGT	CCG	TAG	CAC	CTA

31	<b>Y</b>	<b>G</b> SDN	<b>S</b> N	<b>S</b> N	<b>S</b> Q	<b>R</b> <sup>(*)</sup> K	<b>R</b> T	<b>V</b> A <sub>(IM)</sub>	<b>SLF</b> P	<b>*</b> Q	<b>I</b> T	<b>G</b> SDN	<b>I</b>	<b>V</b> MI	<b>D</b> N	45
----	----------	-----------------	---------------	---------------	---------------	------------------------------	---------------	-------------------------------	-----------------	---------------	---------------	-----------------	----------	----------------	---------------	----

	GAG	TGC	TGC	TTC	CGG	AGC	TGT	GAT	CTA	AGG	AGG	CTG	GAG	ATG	TAT
	CTC	ACG	ACG	AAG	GCC	TCG	ACA	CTA	GAT	TCC	TCC	GAC	CTC	TAC	ATA

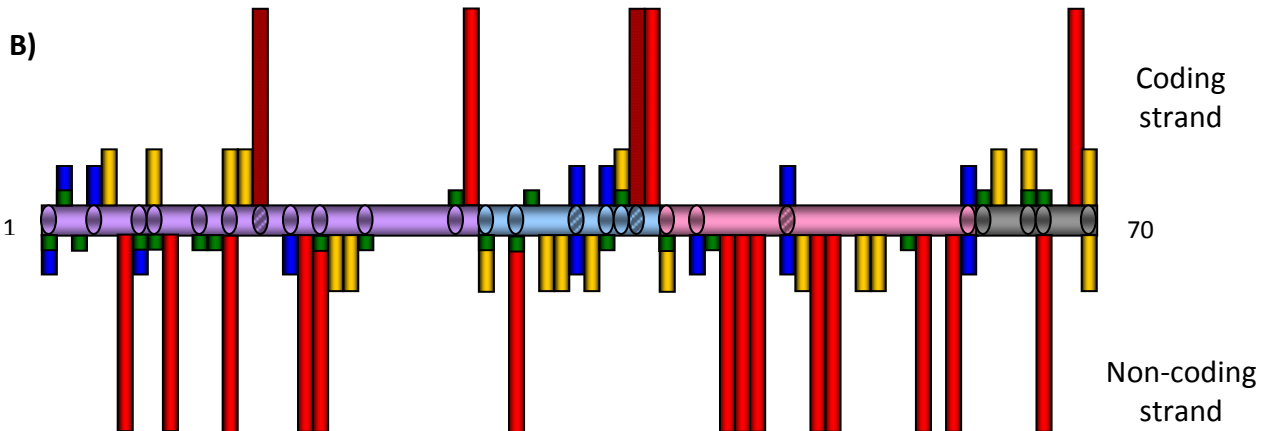
46	<b>E</b> K	<b>C</b> Y	<b>C</b> Y	<b>F</b>	<b>R</b> <sup>(*)</sup> Q	<b>S</b> N	<b>C</b> Y	<b>D</b> N	<b>L</b>	<b>R</b> K	<b>R</b> K	<b>L</b>	<b>E</b> K	<b>M</b> I	<b>Y</b>	60
----	---------------	---------------	---------------	----------	------------------------------	---------------	---------------	---------------	----------	---------------	---------------	----------	---------------	---------------	----------	----

	TGC	GCA	CCC	CTC	AAG	CCT	GCC	AAG	TCA	GCT	TAG
	ACG	CGT	GGG	GAG	TTC	GGA	CGG	TTC	AGT	CGA	ATC

61	<b>C</b> Y	<b>V</b> A <sub>(I)</sub>	<b>SLF</b> P	<b>F</b> L	<b>K</b>	<b>SLF</b> P	<b>V</b> A <sub>(I)</sub>	<b>K</b>	<b>L</b> S	<b>V</b> A <sub>(I)</sub>	<b>*</b>	70
----	---------------	------------------------------	-----------------	---------------	----------	-----------------	------------------------------	----------	---------------	------------------------------	----------	----



**Figure 3.2 Potential RIP-mediated mutations in the human IGF-I sequence.**

**(A)** The DNA coding sequence of human IGF-I is colour coded to highlight the four potential dinucleotide target sites for RIP-induced mutations: CpA are red, CpT are gold, CpG are blue and CpC are green. The mature human IGF-I amino acid sequence (shown in black) is presented below the relevant codons. Amino acid substitutions that could occur due to RIP of either coding or non-coding strands are shown above and below the protein sequence respectively. The colour of the potential amino acid substitution matches the relevant RIP dinucleotide target sequence. Amino acid mutations that require RIP of more than one dinucleotide target in a given codon are coloured grey. Missense mutations resulting from RIP of multiple dinucleotide targets within a given codon on both coding and non-coding strands are indicated by bracketed subscripts in the peptide sequence and are also coloured grey. Stop codons are represented by a \*. The colour coded bars flanking the DNA sequence delineate the BCAD domains of IGF-I.

**(B)** Histogram representing potential RIP-mediated amino acid substitutions. The horizontal bar represents the IGF-I mature peptide and is colour coded to delineate the BCAD domains. The vertical columns represent potential amino acid substitutions resulting from RIP missense mutations with the bar colour coded for the specific dinucleotides targeted by RIP whilst the height of the columns are proportional to their relative RIP frequency: CpA (~64%) are red, CpT (~18%) are gold, CpG (~13%) are blue and CpC (~5%) are green. Columns above and below the bar indicate mutations in the coding and non-coding strands respectively. RIP-mutations that introduce stop codons are represented as columns with a diagonal pattern. Overlaying of columns indicates an amino acid residue that may be altered by RIP targeting of alternative dinucleotide targets within a single codon. Amino acid substitutions that could occur when multiple dinucleotide targets within a codon are mutated by RIP are represented as ovals within the horizontal bar but are not scaled to reflect relative frequency.

(A)

Potential dinucleotide RIP target	Number of amino acid residues that can be mutated due to RIP <sup>*†</sup>			Number of amino acid substitutions potentially created by RIP <sup>*</sup>		
	Human IGF-I	RIP enhanced IGF-I	Codon bias optimised IGF-I	Human IGF-I	RIP enhanced IGF-I	Codon bias optimised IGF-I
CpA	17	35	8	17	38	8
CpT	17	12	8	19	13	7
CpG	10	8	28	13	15	34
CpC	9	1	12	24	14	30
Multiple CpX	NA	NA	NA	22	14	24
<b>Total</b>	53	56	56	95	94	103

(B)

IGF-I domain	Number of amino acids in each domain	Number of amino acid residues that can be mutated due to RIP <sup>*</sup>			Number of amino acid substitutions potentially created by RIP <sup>*</sup>		
		Human IGF-I	RIP enhanced IGF-I	Codon bias optimised IGF-I	Human IGF-I	RIP enhanced IGF-I	Codon bias optimised IGF-I
<b>B</b>	29	21	22	22	38	37	38
<b>C</b>	12	10	10	10	20	19	22
<b>A</b>	21	16	18	18	22	25	28
<b>D</b>	8	6	6	6	14	13	14

**Table 3.1 Comparison of the number of amino acid residues with the potential to be altered due to RIP and number of possible substitutions in human, RIP-enhanced and codon bias-optimised IGF-I coding sequences.**

(A) Number of missense mutations resulting from RIP-mutation of specific dinucleotide targets. Multiple CpX refers to missense mutations resulting from RIP of multiple dinucleotide targets within a given codon. \*Tally does not include nonsense mutations. †Where multiple missense mutations are possible in a codon, residue change is assigned to the RIP dinucleotide target of highest frequency. NA: not applicable as substitutions can occur by RIP of a single dinucleotide target within a relevant codon. (B) Distribution of potential RIP mutations throughout the BCAD domains of the IGFI protein. \*Tally does not include nonsense mutations.

**3.1.2.1 Codon optimisation: RIP vs. codon bias**

Codon bias is the preferential use of specific synonymous codons in coding sequences for highly abundant proteins and has resulted from co-evolution between tRNA abundance and codon usage to maximize protein translation efficiency for highly expressed genes (Gustafsson *et al.*, 2004). Codon bias in highly expressed genes has been reported in numerous organisms such as *Neurospora crassa* (Kinnaird *et al.*, 1991) & (Borkovich *et al.*, 2004), *Escherichia coli* (Ikemura 1981), *Saccharomyces cerevisiae* (Bennetzen & Hall 1982), *Caenorhabditis elegans*, *Drosophila melanogaster* and *Arabidopsis thaliana* (Duret & Mouchiroud 1999).

Bioinformatic analysis of *Neurospora* codon usage indicates a bias towards G+C content in the third position of synonymous codons in highly expressed genes with the base preference being C, G, T then A, in descending order (Edelmann & Staben 1994); (Borkovich *et al.*, 2004). In contrast, RIP-mediated missense mutations have the opposite preference in the third base of codons where the preferred order is A, T, G then C. Significantly, the number of amino acids that can be altered as a result of RIP-preferred CpA mutation drops from 35 residues in the RIP-enhanced coding sequence (Figure 3.3; Table 3.1) to 8 in the codon-bias optimised sequence (Figure 3.4). The opposing requirements for codons ideally suited for RIP or codon bias presented a quandary when codon optimising for the purposes of gene diversification in *Neurospora*.

A)

A	GGC	CCA	GAG	ACA	CTC	TGT	GGT	GCA	GAG	CTC	GTC	GAC	GCA	CTC	CAG
T	CCG	GGT	CTC	TGT	GAG	ACA	CCA	CGT	CTC	GAG	CAG	CTG	CGT	GAG	GTC

1	G	P	E	T	L	C	G	A <sub>(I)</sub>	E	L	V	D	A <sub>(I)</sub>	L	Q	15
	SDN		K			Y	SDN	T	K		I	N	T			

	TTT	GTC	TGT	GGT	GAC	CGG	GGT	TTT	TAC	TTT	AAC	AAG	CCA	ACA	GGT
	AAA	CAG	ACA	CCA	CTG	GCC	CCA	AAA	ATG	AAA	TTG	TTC	GGT	TGT	CCA

16	F	V	C	G	D	R <sub>(*)</sub>	G	F	Y	F	N	K	P	T	G	30
		I	Y	SDN	N	Q	SDN								SDN	

	TAT	GGC	TCA	TCA	TCA	CGG	CGG	GCA	CCA	CAG	ACA	GGT	ATT	GTT	GAT
	ATA	CCG	AGT	AGT	AGT	GCC	GCC	CGT	GGT	GTC	TGT	CCA	TAA	CAA	CTA

31	Y	G	S	S	S	R <sub>(*)</sub>	R <sub>(*)</sub>	A <sub>(I)</sub>	P	Q	T	G	I	V	D	45
		SDN				Q	Q	T				SDN		I	N	

	GAG	TGC	TGC	TTC	CGG	TCA	TGT	GAT	CTC	CGG	CGG	CTT	GAG	ATG	TAC
	CTC	ACG	ACG	AAG	GCC	AGT	ACA	CTA	GAG	GCC	GCC	GAA	CTC	TAC	ATG

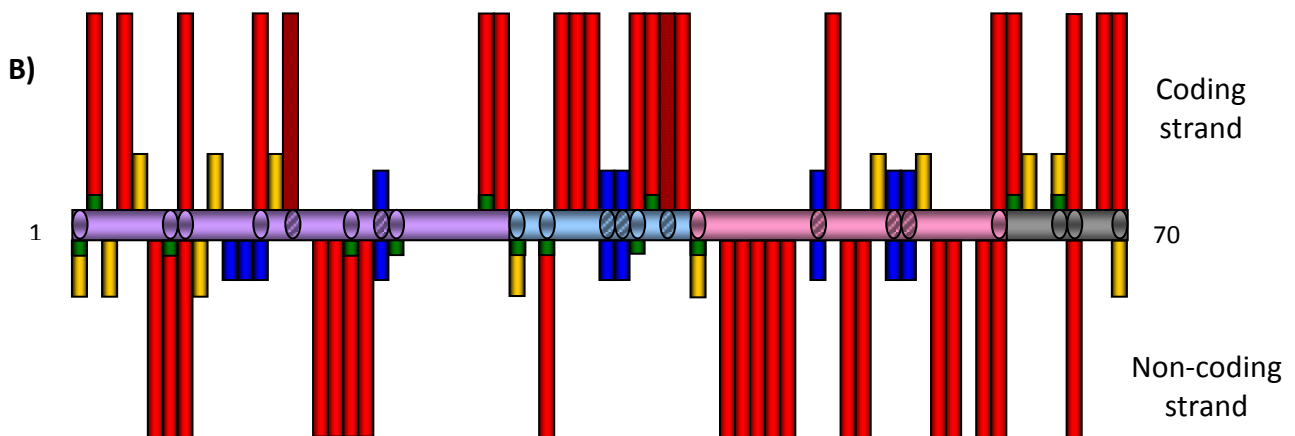
46	E	C	C	F	R <sub>(*)</sub>	S	C	D	L	R <sub>(*)</sub>	R <sub>(*)</sub>	L	E	M	Y	60
	K	Y	Y		Q		Y	N		Q	Q		K	I		

	TGT	GCA	CCA	CTC	AAG	CCT	GCA	AAG	TCA	GCA	TAA				
	ACA	CGT	GGT	GAG	TTC	GGA	CGT	TTC	AGT	CGT	ATT				

61	C	A <sub>(I)</sub>	P	L	K	P	A <sub>(I)</sub>	K	S	A <sub>(I)</sub>	*	70
	Y	T					T			T		



**Figure 3.3 Potential RIP-mediated mutations in the RIP-enhanced IGF-I sequence.**

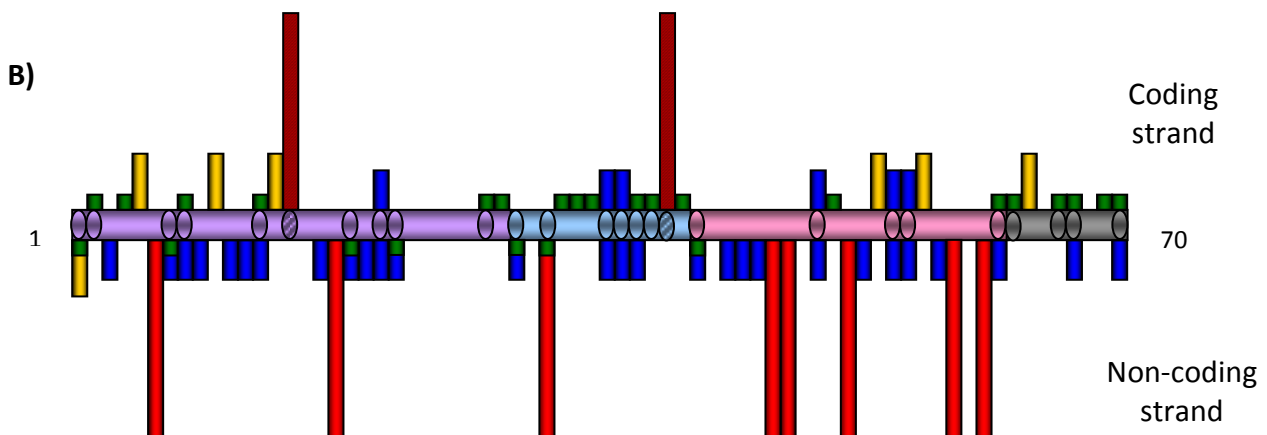
**(A)** The human IGF-I coding sequence re-engineered to enhance its potential for RIP-induced mutations in *Neurospora* is colour coded to indicate the four potential dinucleotide target sites for RIP mutations: CpA are red, CpT are gold, CpG are blue and CpC are green. The mature human IGF-I amino acid sequence (shown in black) is presented below the relevant codons. Amino acid substitutions that could occur due to RIP of either coding or non-coding strands are shown above and below the protein sequence, respectively. The colour of the potential amino acid substitution matches the relevant RIP dinucleotide target sequence. Amino acid mutations that require RIP of more than one dinucleotide target in a given codon are coloured grey. Missense mutations resulting from RIP of multiple dinucleotide targets within a given codon on both coding and non-coding strands are indicated by bracketed subscripts in the peptide sequence and are also coloured grey. Stop codons are represented by a \*. The colour coded bars flanking the DNA sequence delineate the BCAD domains of IGFI.

**(B)** Histogram representing potential RIP-mediated amino acid substitutions. The horizontal bar represents the IGFI mature peptide and is colour coded to delineate the BCAD domains. The vertical columns represent potential amino acid substitutions resulting from RIP missense mutations with the bar colour coded for the specific dinucleotides targeted by RIP whilst the height of the columns are proportional to their relative RIP frequency: CpA (~64%) are red, CpT (~18%) are gold, CpG (~13%) are blue and CpC (~5%) are green. Columns above and below the bar indicate mutations in the coding and non-coding strands respectively. RIP-mutations that introduce stop codons are represented as columns with a diagonal pattern. Overlaying of columns indicates an amino acid residue that may be altered by RIP targeting of alternative dinucleotide targets within a single codon. Amino acid substitutions that could occur when multiple dinucleotide targets within a codon are mutated by RIP are represented as ovals within the horizontal bar but are not scaled to reflect relative frequency.

*CHAPTER 3 – Vector construction for diversification of human IGF-I in Neurospora*

**A)**

A	GGC	CCC	GAG	ACC	CTC	TGC	GGC	GCC	GAG	CTC	GTC	GAC	GCC	CTC	CAG
T	CCG	GGG	CTC	TGG	GAG	ACG	CCG	CGG	CTC	GAG	CAG	CTG	CGG	GAG	GTC
1	<b>G</b>	<b>P</b>	<b>E</b>	<b>T</b>	<b>L</b>	<b>C</b>	<b>G</b>	<b>A<sub>(I)</sub></b>	<b>E</b>	<b>L</b>	<b>V</b>	<b>D</b>	<b>A<sub>(I)</sub></b>	<b>L</b>	<b>Q</b>
	<b>SDN</b>		<b>K</b>			<b>Y</b>	<b>SDN</b>	<b>T</b>	<b>K</b>		<b>I</b>	<b>N</b>	<b>T</b>		
	TTC	GTC	TGC	GGC	GAC	CGC	GGC	TTC	TAC	TTC	AAC	AAG	CCC	ACC	GGC
	AAG	CAG	ACG	CCG	CTG	GCG	CCG	AAG	ATG	AAG	TTG	TTC	GGG	TGG	CCG
16	<b>F</b>	<b>V</b>	<b>C</b>	<b>G</b>	<b>D</b>	<b>R<sub>(Y)</sub></b>	<b>G</b>	<b>F</b>	<b>Y</b>	<b>F</b>	<b>N</b>	<b>K</b>	<b>P</b>	<b>T</b>	<b>G</b>
		<b>I</b>	<b>Y</b>	<b>SDN</b>	<b>N</b>	<b>H</b>	<b>SDN</b>								<b>SDN</b>
	TAC	GGC	TCC	TCC	TCC	CGC	CGC	GCC	CCC	CAG	ACC	GGC	ATC	GTC	GAC
	ATG	CCG	AGG	AGG	AGG	GCG	GCG	CGG	GGG	GTC	TGG	CCG	TAG	CAG	CTG
31	<b>Y</b>	<b>G</b>	<b>S</b>	<b>S</b>	<b>S</b>	<b>R<sub>(Y)</sub></b>	<b>R<sub>(Y)</sub></b>	<b>A<sub>(I)</sub></b>	<b>P</b>	<b>Q</b>	<b>T</b>	<b>G</b>	<b>I</b>	<b>V</b>	<b>D</b>
		<b>SDN</b>				<b>H</b>	<b>H</b>	<b>T</b>				<b>SDN</b>		<b>I</b>	<b>N</b>
	GAG	TGC	TGC	TTC	CGC	TCC	TGC	GAC	CTC	CGC	CGC	CTC	GAG	ATG	TAC
	CTC	ACG	ACG	AAG	GCG	AGG	ACG	CTG	GAG	GCG	GCG	GAG	CTC	TAC	ATG
46	<b>E</b>	<b>C</b>	<b>C</b>	<b>F</b>	<b>R<sub>(Y)</sub></b>	<b>S</b>	<b>C</b>	<b>D</b>	<b>L</b>	<b>R<sub>(Y)</sub></b>	<b>R<sub>(Y)</sub></b>	<b>L</b>	<b>E</b>	<b>M</b>	<b>Y</b>
	<b>K</b>	<b>Y</b>	<b>Y</b>		<b>H</b>		<b>Y</b>	<b>N</b>		<b>H</b>	<b>H</b>		<b>K</b>	<b>I</b>	
	TGC	GCC	CCC	CTC	AAG	CCC	GCC	AAG	TCC	GCC	TAA				
	ACG	CGG	GGG	GAG	TTC	GGG	CGG	TTC	AGG	CGG	ATT				
61	<b>C</b>	<b>A<sub>(I)</sub></b>	<b>P</b>	<b>L</b>	<b>K</b>	<b>P</b>	<b>A<sub>(I)</sub></b>	<b>K</b>	<b>S</b>	<b>A<sub>(I)</sub></b>	<b>*</b>				<b>70</b>
	<b>Y</b>	<b>T</b>				<b>T</b>				<b>T</b>					





**Figure 3.4 Potential RIP-mediated mutations in the codon bias optimised IGF-I sequence.**

**(A)** The *Neurospora* codon-optimised DNA coding sequence of human IGF-I is colour coded to highlight the four potential dinucleotide target sites for RIP-induced mutations: CpA are red, CpT are gold, CpG are blue and CpC are green. The mature human IGF-I amino acid sequence (shown in black) is presented below the relevant codons. Amino acid substitutions that could occur due to RIP of either coding or non-coding strands are shown above and below the protein sequence, respectively. The colour of the potential amino acid substitution matches the relevant RIP dinucleotide target sequence. Amino acid mutations that require RIP of more than one dinucleotide target in a given codon are coloured grey. Missense mutations resulting from RIP of multiple dinucleotide targets within a given codon on both coding and non-coding strands are indicated by bracketed subscripts in the peptide sequence and are also coloured grey. Stop codons are represented by a \*. The colour coded bars flanking the DNA sequence delineate the BCAD domains of IGFI.

**(B)** Histogram representing potential RIP-mediated amino acid substitutions. The horizontal bar represents the IGFI mature peptide and is colour coded to delineate the BCAD domains. The vertical columns represent potential amino acid substitutions resulting from RIP missense mutations with the bar colour coded for the specific dinucleotides targeted by RIP whilst the height of the columns are proportional to their relative RIP frequency: CpA (~64%) are red, CpT (~18%) are gold, CpG (~13%) are blue and CpC (~5%) are green. Columns above and below the bar indicate mutations in the coding and non-coding strands respectively. RIP-mutations that introduce stop codons are represented as columns with a diagonal pattern. Overlaying of columns indicates an amino acid residue that may be altered by RIP targeting of alternative dinucleotide targets within a single codon. Amino acid substitutions that could occur when multiple dinucleotide targets within a codon are mutated by RIP are represented as ovals within the horizontal bar but are not scaled to reflect relative frequency.

The primary objective of this project was to use RIP to generate a large number of IGF-I variants and express them in *Neurospora* for high-throughput screening. Novel IGF-I protein secreted from progeny of a diversification cross were to be screened by an IGF-IR ligand binding assay. Thus, RIP-optimisation of the IGF-I coding sequence would enable a more diverse library to be generated. Conversely, optimisation of the IGF-I coding sequence for codon bias would have reduced potential IGF-I diversity. This was likely to significantly increase the number of progeny secreting unaltered IGF-I or common IGF-I variants resulting in a highly undesirable drain on screening time and resources. Initially, a fusion vector was constructed with the endogenous glucoamylase *glaA* CDS fused to the 5' end of the IGF-I coding sequence (described in 3.1.3.1). As rare codons tend to limit protein levels if the rare codons appear in the N-terminal part of the protein (Gustafsson *et al.*, 2004), it was reasoned that the fusion strategy might mitigate some of the potentially detrimental effects of the rare codons in the RIP-optimised coding sequence. In addition, the availability of a highly sensitive IGF-IR binding assay with a typical resolution range of 10pM to 1nM (76.5pg/ml to 7.65ng/ml) meant potentially high expression levels were initially considered to be of lesser importance than obtaining the greatest number of mutated IGF-I sequences with the greatest distribution of mutations. As such, optimisation of the *IGF-I* coding sequence was focused towards RIP-enhancement.

Codon positions suitable for RIP optimisation were identified using the following three characteristics of the *Neurospora* genetic code as guidelines:

- 1) C to T and G to A mutations in the first and/or second base of a codon are always missense mutations with the exception of C to T transition in the first base of CTA and CTG codons (both encoding leucine) where mutation is silent.
- 2) C to T transitions in the third base of a codon are always silent mutations.
- 3) G to A transitions in the third base of the codon are always silent with the exception of ATG (methionine) to ATA (isoleucine) and TGG (tryptophan) to TGA (stop codon).

With minor exceptions, RIP-optimisation of the human *IGF-I* coding sequence was achieved by introducing silent mutations to maximize the number of CpX dinucleotides (with a preference for CpA>CpT>CpG>CpC) in both strands where RIP would result in missense mutations (Figure 3.3). When RIP enhancement was not possible, codons were modified to optimise for codon bias using *Neurospora* codon usage (Edelmann & Staben 1994). This strategy was utilized for all codons with the exception of those encoding arginine. In this instance, the CGG codon was used in place of the codon-bias preferred CGC codon as the latter codon could be mutated to encode a cysteine residue which could potentially disrupt the global IGF-I structure by interfering with normal disulphide bond formation. RIP optimisation of the *IGF-I* coding sequence resulted in 56 amino acids that could be mutated by RIP with 94 possible amino acid substitutions (Table 3.1).

### **3.1.2.2 Construction of a RIP-enhanced coding sequence**

RIP-enhancement of *IGF-I* required 45bp changes throughout the 213bp coding sequence (Figure 3.3). To this end, large complementary synthetic oligonucleotides were used to construct a RIP-enhanced *IGF-I* coding sequence (*nIGF-I*). Three pairs of complementary oligonucleotides (SO1/ASO1, SO2/ASO2 & SO3/ASO3) ranging from 55 to 89-bases were designed (Table 2.1.8.1) such that complementary hybridisation produced single stranded overhangs of up to 6 bases at either end to facilitate annealing between the double stranded oligonucleotide pairs. The oligonucleotides were analysed by OligoAnalyzer 3.0 (<http://207.32.43.70/biotools/oligoalc/oligoalc.asp>) to ensure the free energy of potential homodimers and secondary structures such as hairpins was small relative to the free energy of hybridisation between complementary oligonucleotides.

The manufacturer supplied the oligonucleotides with 5' termini phosphorylated to facilitate ligation and were also PAGE purified to remove truncated sequences due to DNA synthesis inefficiencies. Each of the six oligonucleotides were combined at an individual concentration of 20.25µM in complementary pairwise combinations and annealed together by heating to 100°C for 10' and slowly cooling to room temperature. Two sequential ligation reactions were performed as follows: Firstly, 8.1µM of SO1/ASO1 and SO2/ASO2 were ligated together overnight at 4°C to form a 154bp oligonucleotide designated as O1&2. Secondly, ligation of 3µM of O1&2 and SO3/ASO3 overnight at 4°C formed the 222bp double stranded oligonucleotide encoding the *IGF-I* coding sequence flanked by single stranded overhangs corresponding to *SbfI* and *XbaI* sticky ends for cloning into the pFEN expression vector.

### **3.1.3 Heterologous expression of IGF-I in Neurospora**

Filamentous fungi are widely utilised for the production of homologous and heterologous proteins due to the naturally high capacity of their protein secretion machinery (Yin *et al.*, 2007). Whilst the yield of homologous proteins can be very high, production levels of heterologous proteins are often relatively low (Iwashita 2002). Potential limitations depend on the specific heterologous protein and/or the fungal host strain and can occur at numerous stages such as transcription, translation, intracellular protein trafficking, secretion and/or extracellular proteolytic degradation (Gouka *et al.*, 1997a).

#### **3.1.3.1 Fusion of glucoamylase to heterologous genes**

A commonly used strategy to improve secretion yield in filamentous fungi is to fuse an endogenous, highly expressed carrier protein such as glucoamylase (*glaA*) to the N-terminus of the heterologous protein of interest (Archer *et al.*, 1994). Fusion of the glucoamylase protein is believed to improve heterologous protein secretion by assisting translocation into the endoplasmic reticulum, subsequent protein folding and

processing through the endomembrane pathway with typical yield increases from 5 to 1000-fold (Gouka *et al.*, 1997b). In addition, *glaA* fusion proteins can also have a beneficial effect at the transcriptional level as well, for example, a 25-fold increase in mRNA transcript levels was reported with the fusion of *glaA* protein to the N-terminus of an  $\alpha$ -galactosidase expressed in *Aspergillus awamori* (Gouka *et al.*, 1997b). The glucoamylase proteins from *Aspergillus niger* or *A. awamori* are typically used and consist of catalytic and starch-binding domains linked by a flexible linker region. Fusion strategies generally use a truncated glucoamylase gene (*glaA*) consisting of the secretion signal, catalytic domain and most of the linker region fused to the 5' end of the gene of interest (Jeenes *et al.*, 1993). The flexible linker is presumed to facilitate folding of the heterologous protein independently of the catalytic domain of the glucoamylase protein (Gouka *et al.*, 1997b). Incorporation of a KEX2 recognition site (lysine-arginine) in the linker region enables cleavage of the fusion protein by a KEX2-like endopeptidase in the Golgi apparatus prior to secretion (Broekhuijsen *et al.*, 1993).

The presence of a glucoamylase gene (*gla-1*) in *Neurospora crassa* homologous to the *Aspergillus glaA* gene (Stone *et al.*, 1993) facilitates the use of a fusion strategy using the endogenous *gla-1* sequence to secrete heterologous proteins at relatively high levels in *Neurospora* (E. Cambareri: *Personal communication*); (Rasmussen-Wilson *et al.*, 1997). The presence of a 63bp intron in the *Neurospora gla-1* sequence may be an additional factor contributing to the increased protein yields obtained with *gla-1* fusions as the presence of introns has been shown to improve heterologous protein levels in numerous organisms such as yeast (Juneau *et al.*, 2006), plants and mammalian cells (Gouka *et al.*, 1997a). The mechanisms by which introns increase protein yield vary, for example, (Liu *et al.*, 1995) demonstrated that intron removal from the rat Growth Hormone gene (rGH) disrupted nucleosome binding and (Sleckman *et al.*, 1996) identified transcriptional enhancers in the introns of human antigen receptor variable region genes. An endogenous glucoamylase-IGF-I fusion strategy was employed in this project to facilitate secretion of IGF-I in *Neurospora crassa*.

**3.1.3.2 Construction of pFEN: a glucoamylase fusion expression vector (Figure 3.5)**

A glucoamylase fusion vector (pHDKXL1) (Figure 3.5) was kindly provided by Neugenes Corporation (California). This vector contained the following key features: the *glaA* promoter from *Aspergillus niger* (note: the *glaA* promoter is inducible in *Aspergillus niger* (Verdoes *et al.*, 1994) but is constitutively expressed in *Neurospora crassa* *gla-1* sequence encoding 521 amino acids of the N-terminus of the glucoamylase protein; a KEX2 cleavage site in the truncated glucoamylase linker region; a cloning site for the gene of interest followed by the *Neurospora crassa gla-1* terminator. Because pHDKXL1 lacked an antibiotic resistance gene suitable for selection in fungi, the expression cassette was cloned into an alternative vector backbone as described in Figure 3.5.

RedHot DNA polymerase (Abgene) was used for high-fidelity PCR amplification of the ~3400bp expression cassette from the pHDKXL1 plasmid using primers EXPCASS-F1 and EXPCASS-R1 (Table 2.1.8.3), the latter of which included specific base mismatches to introduce an *SphI* restriction site 5' of the *glaA* promoter and *SpeI* and *PciI* restriction sites 3' of the *gla-1* terminator. The *SphI* and *PciI* restriction sites facilitated cloning into the vector backbone whilst the *SpeI* site was included in preparation for subsequent cloning into the pDV2 and pDV3 diversification vectors.

Plasmid pSH1.5EGFP (Figure 3.5) (created by the author in an unrelated study) was sequentially digested with *SphI* & *PciI* and the plasmid fragments resolved by gel electrophoresis prior to gel extraction of the ~3.9kb vector backbone. The purified vector sequence contained an ampicillin resistance gene (*amp*) and an autonomous replication sequence (*ori*) to enable selection and propagation in *E. coli*. The sequence also contained the coding sequence for hygromycin resistance (*hph*) expressed by a truncated *gpdA* promoter from *Aspergillus nidulans* to confer selection in *Neurospora*.

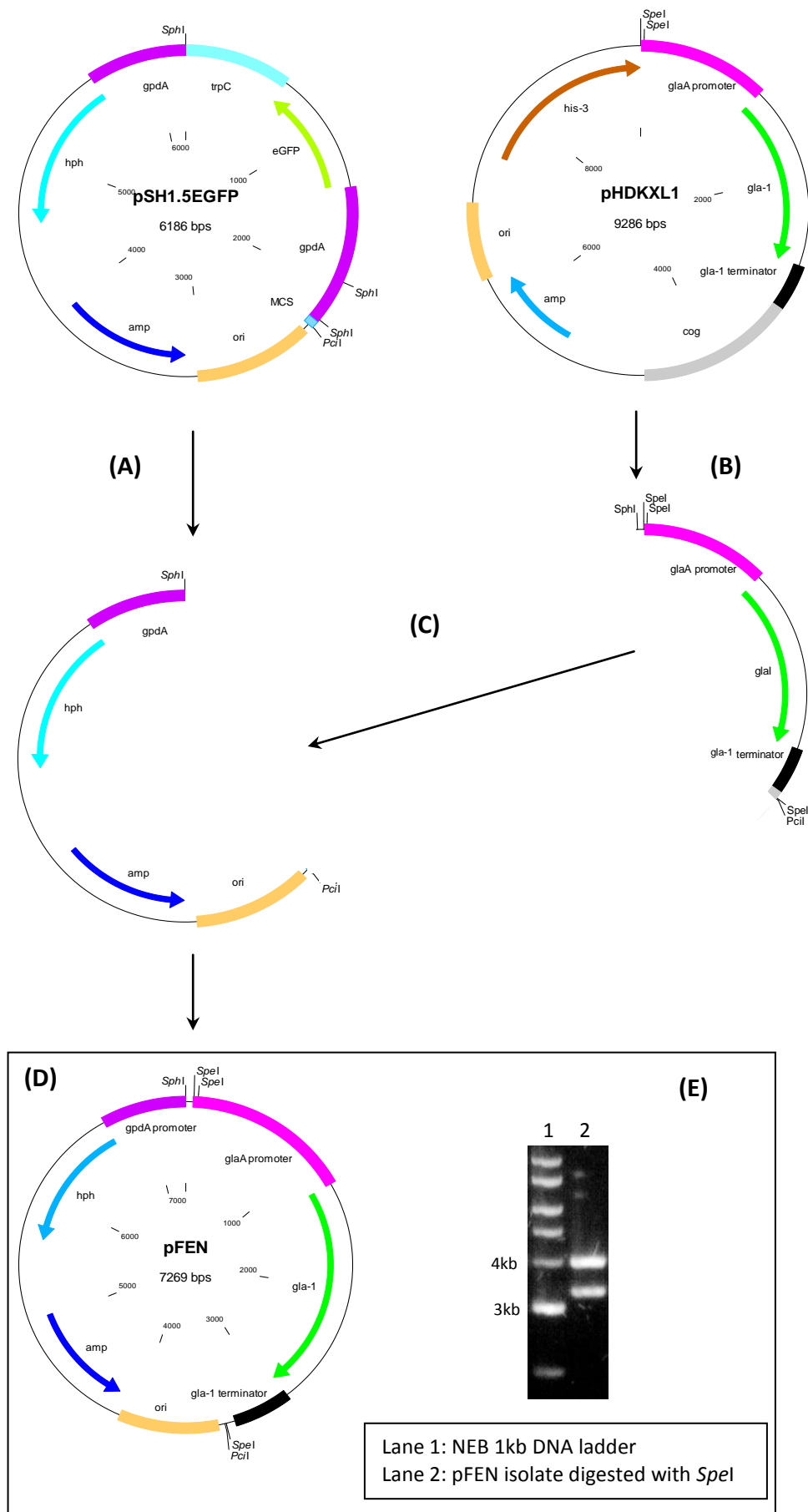
The amplified expression cassette was digested with *SphI* and *PciI*, purified with a spin column, ligated into the pSH1.5EGFP vector backbone to form pFEN (Fusion Expression in *Neurospora*) and transformed into *E. coli*. Screening for pFEN positive transformants was performed by restriction digest of plasmid DNA with *SpeI* (Figure 3.5B) and the expression cassette integrity was subsequently confirmed by sequencing.

### **3.1.3.3 Construction of pFENIGFI: a *nIGF-I*:glucoamylase fusion expression vector (Figure 3.6)**

The single stranded overhang 5' of the *nIGF-I* coding sequence start codon was designed to be complementary to the sticky end of an *SbfI* restriction site whilst the overhang 3' of the *nIGF-I* stop codon matched an *XbaI* sticky end facilitating directional cloning into the pFEN expression vector. The insertion of *nIGF-I* into the *SbfI* site 3' of the Kex2 recognition site resulted in the addition of DNA sequence encoding proline and alanine residues to the N-terminus of the mature IGF-I protein. The pFEN expression vector was digested with *SbfI* and *XbaI*, purified with a spin column, ligated with the double stranded *nIGF-I* coding sequence oligonucleotides and transformed into *E. coli*. Putative pFENIGFI transformants were screened by *BstEII* digestion of plasmid DNA prior to DNA sequencing of the *nIGF-I* coding sequence.

#### **3.1.3.3.1 Repair of the *nIGF-I* coding sequence**

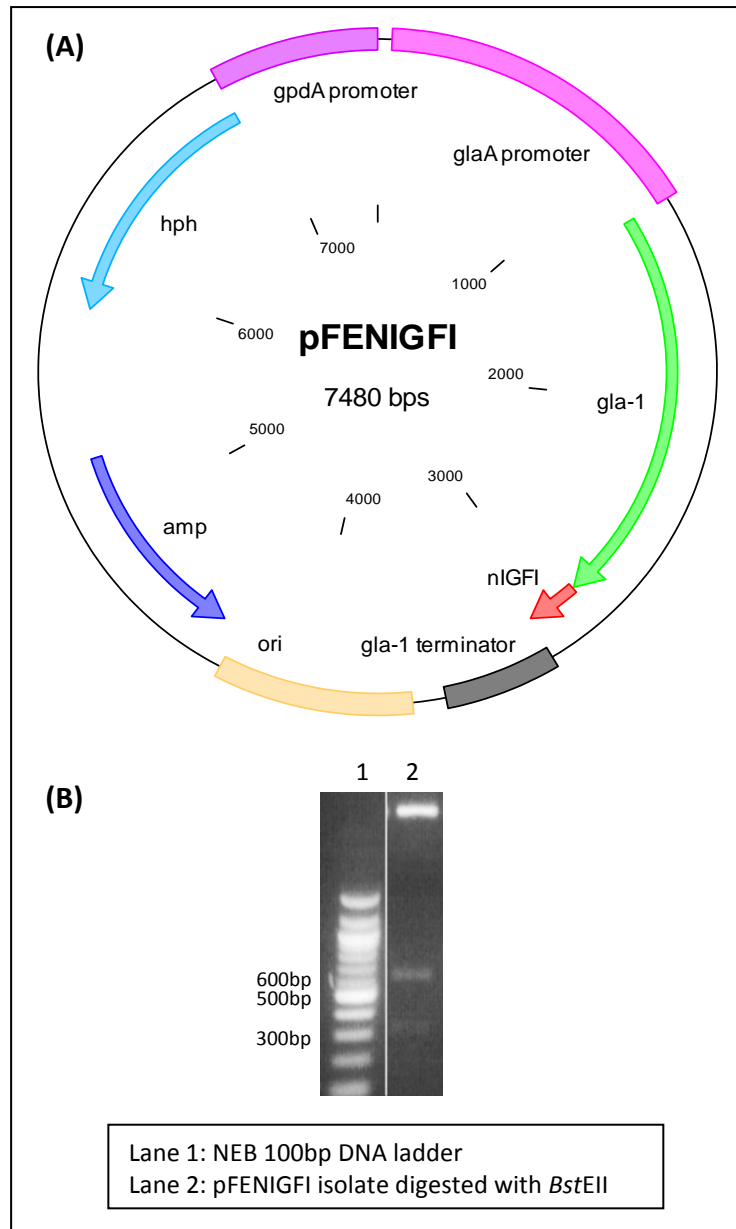
The stop codon of the *nIGF-I* coding sequence within pFENIGFI needed to be repaired as DNA sequencing of the region had revealed a deletion in the third position of the *nIGF-I* stop codon (TAA) forming a tyrosine codon (TAT). The *nIGF-I* coding sequence was PCR amplified with primers EXPCASS-F1 and IGFI-repair1 (Table 2.1.8.3). The latter primer contained an additional mismatched thymine base incorporating a complementary adenine into the sense strand upon PCR amplification, thereby reforming the TAA stop codon. The PCR amplicon and pFEN vector were digested with *SbfI* and *XbaI*, spin column purified and ligated together to form pFENIGFI. Putative pFENIGFI *E. coli* transformants were screened by *BstEII* digestion of plasmid DNA (Figure 3.6B) prior to DNA sequencing of the *nIGF-I* coding sequence to confirm integrity.





**Figure 3.5 Construction of a fusion expression vector for ectopic integration in *Neurospora*.**

**(A)** Plasmid pSH1.5EGFP was digested with *SphI* & *PciI* and the vector backbone containing hygromycin (*hph*) and ampicillin (*amp*) resistance genes was isolated by gel extraction. **(B)** The expression cassette in pHDKXL1 consisting of a truncated *Neurospora crassa gla-1* coding sequence flanked by *glaA* promoter (*A. niger*) and *gla-1* terminator sequences (*N. crassa*) was PCR amplified with primers that introduced an *SphI* restriction site upstream and *PciI* and *SpeI* sites downstream. **(C)** The PCR amplicon was digested with *SphI* & *PciI* and ligated with the gel extracted fragment of pSH1.5EGFP to form pFEN **(D)**. Confirmation of pFEN formation was determined by *SpeI* restriction digest of plasmid DNA from candidate *E. coli* transformants yielding ~4kb and ~3.2kb fragments when resolved by gel electrophoresis **(E)**.



**Figure 3.6** Map of pFENIGFI, an ectopic insertion vector for expression of glucoamylase-IGFI fusion protein in *Neurospora*.

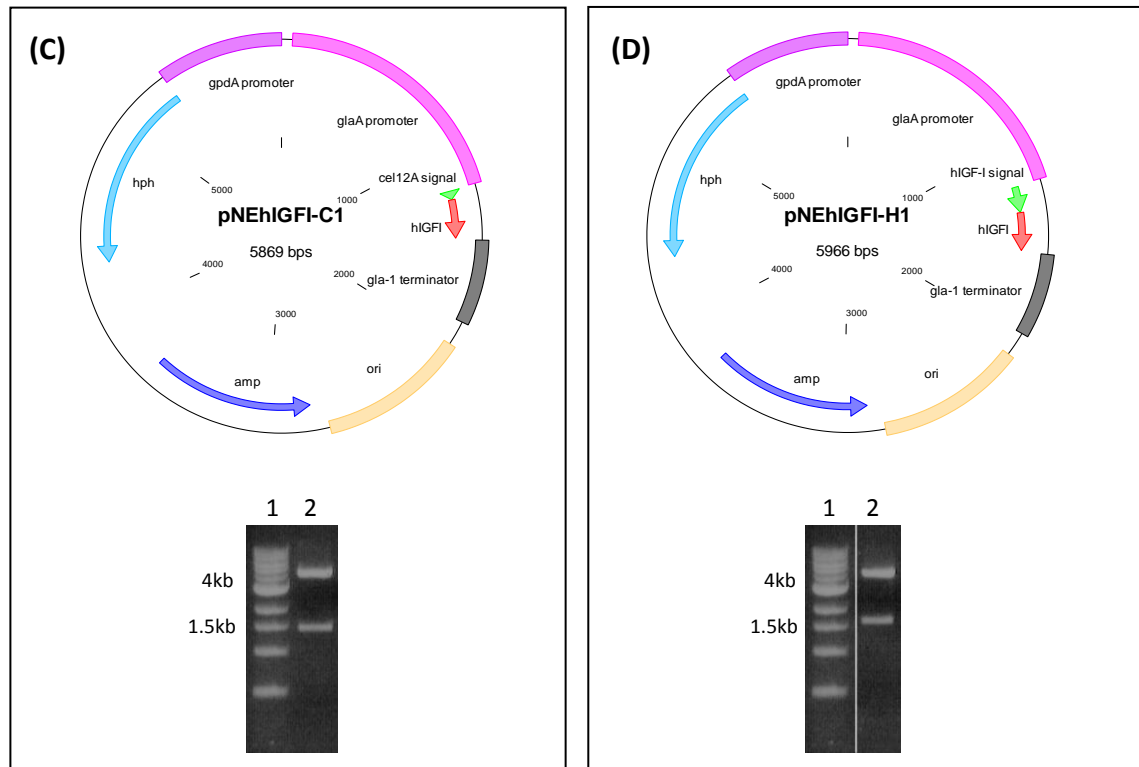
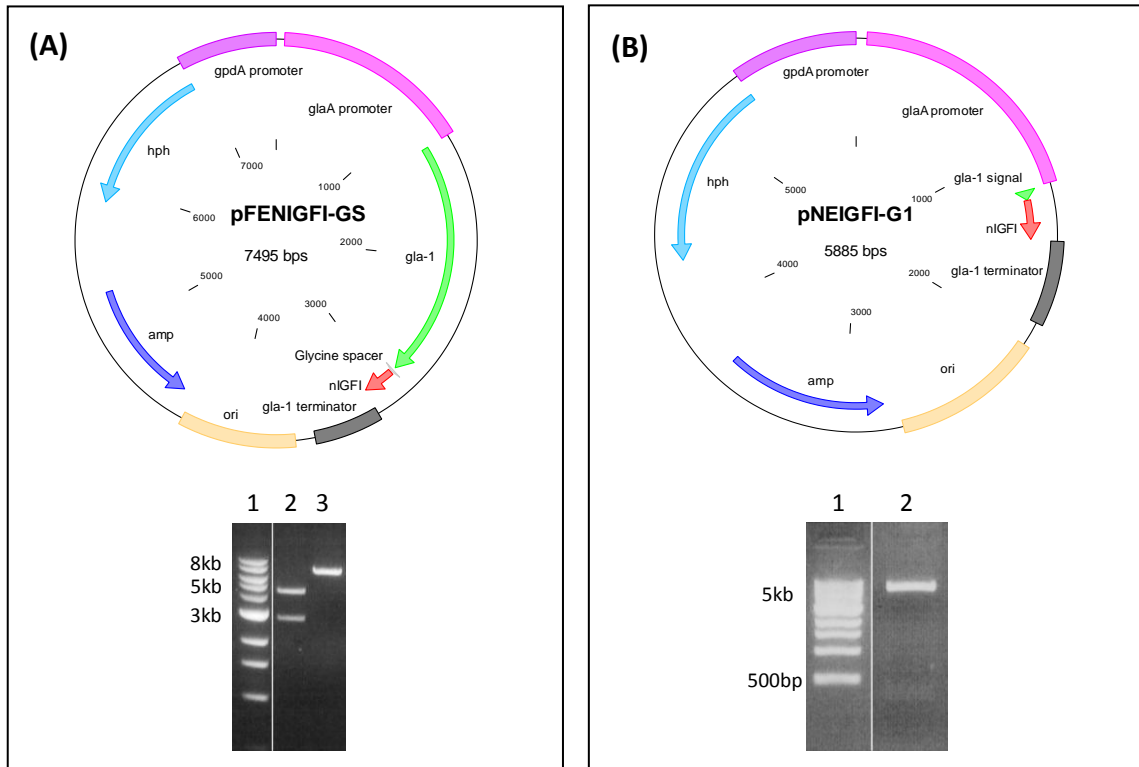
(A) pFENIGFI contains a RIP-enhanced *IGF-I* coding sequence (*nIGFI*) fused to the C-terminus of a truncated endogenous *gla-1* coding sequence. A Kex2-like cleavage site is present in the linker region between the fused proteins enables endopeptidase-mediated separation of the IGF-I from the glucoamylase protein prior to secretion. *glaA* promoter (*A. niger*) and *gla-1* terminator (*N. crassa*) sequences flank the *gla-1* & *nIGFI* coding sequence. A *gpdA* promoter (*A. nidulans*) coupled to a *hph* coding sequence confers hygromycin resistance for fungal selection. *E. coli* propagation and selection is facilitated by an origin of replication (*ori*) and an *amp* gene encoding  $\beta$ -Lactamase provided ampicillin resistance. (B) Confirmation of pFENIGFI construction was determined by *BstEII* restriction digestion of plasmid DNA from candidate *E. coli* transformants yielding ~6.6kb, ~600bp & ~300bp fragments when resolved by agarose gel electrophoresis. IGF-I coding sequence integrity was determined by DNA sequencing.

**3.1.3.4 Construction of alternative IGF-I expression vectors pFENIGFI-GS, pNEIGFI-G1, pNEhIGFI-C1 & pNEhIGFI-H1 (Figure 3.7)**

Four additional IGF-I expression plasmids were constructed for the purpose of increasing IGF-I secretion in *Neurospora*. The requirement for these vectors is described in chapter 4. As gene expression levels are regulated by chromosomal location (Chiou *et al.*, 2002), direct comparison between the IGF-I yields resulting from the five expression vectors required standardization of the chromosomal context to ensure consistent expression levels. As such, all the variant expression cassettes were to be placed into a *his-3* targeting vector. However, the paucity of suitable restriction sites in the pDV2 or pDV3 diversification vectors hindered a simple cloning strategy. Thus, variant *IGF-I* expression sequences were constructed in the pFEN vector background prior to cloning the expression cassettes into a diversification vector.

**3.1.3.4.1 Construction of pFENIGFI-GS (Figure 3.7A)**

pFENIGFI-GS was created to determine if the close proximity of the Kex2-like recognition site to the globular IGF-I protein could be blocking endopeptidase accessibility and thereby inhibiting cleavage of the IGF-I protein from the glucoamylase fusion protein. Glycine is the smallest amino acid with the absence of a side chain permitting a large degree of conformational freedom. Insertion of a polyglycine linker sequence between a protease recognition site and adjacent protein sequence can facilitate increased protease accessibility to the cleavage site (Melcher 2000); (Kavoosi *et al.*, 2007). For example, (Guan & Dixon 1991) reported increased thrombin cleavage of several GST fusion proteins in *E. coli* with the use of glycine rich linker sequences.



**Gel (A).** Lane 1: NEB 1kb DNA ladder; Lane 2: pFENIGFI (-ve control) digested with *SphI* & *StuI*; Lane 3: pFENIGFI-GS isolate digested with *SphI* & *StuI*.

**Gel (B).** Lane 1: NEB 1kb DNA ladder; Lane 2: pNEIGFI-G1 isolate digested with *BstEII*.

**Gel (C).** Lane 1: NEB 1kb DNA ladder; Lane 2: pNEhIGFI-C1 isolate digested with *SphI* & *XbaI*.

**Gel (D).** Lane 1: NEB 1kb DNA ladder; Lane 2: pNEhIGFI-H1 isolate digested with *SphI* & *XbaI*.

**Figure 3.7 Maps of additional IGF-I expression vectors.**

(A) pFENIGFI-GS contains a RIP-enhanced IGF-I coding sequence (*nIGF-I*) fused to the C-terminus of a truncated endogenous *gla-1* coding sequence. A Kex2-like cleavage site is present in the linker region between the fusion proteins to facilitate enzymatic separation of the IGF-I and the glucoamylase proteins prior to secretion. A flexible penta-glycine spacer has been inserted between the Kex2-like recognition site and the *nIGF-I* coding sequence to facilitate protease access to the cleavage site. (B) pNEIGFI-G1 contains the *Neurospora* glucoamylase signal peptide fused directly to the N-terminus of the IGF-I protein. (C) pNEhIGFI-C1 has a fungal *Cel12A* secretion signal (*Trichoderma reesei*) fused directly to the N-terminus of the mature human IGF-I coding sequence (*hIGF-I*). (D) pNEhIGFI-H1 contains the mature *hIGF-I* coding sequence with its native secretion signal.

All vectors contain *glaA* promoter (*A. niger*) and *gla-1* terminator (*N. crassa*) sequences flanking the coding sequences. A *gpdA* promoter (*A. nidulans*) coupled to a *hph* gene confers hygromycin resistance for fungal selection. *E. coli* propagation and selection is facilitated by an origin of replication (*ori*) and an *amp* gene encoding  $\beta$ -Lactamase for ampicillin resistance.

Confirmation of vector construction was determined by restriction digestion of plasmid DNA from candidate *E. coli* transformants. Agarose gel electrophoresis of the digested plasmids resulted in the following expected fragment sizes:

- (A) pFENIGFI-GS: ~7.5kb; pFENIGFI (-ve control): ~4.6kb & ~2.9kb
- (B) pNEIGFI-G1: ~5.6kb & ~300bp
- (C) pNEhIGFI-C1: ~4.4kb & ~1.5kb
- (D) pNEhIGFI-H1: 4.4kb & 1.6kb

IGF-I coding sequence integrity was determined by DNA sequencing in all constructs.

A pentaglycine spacer was inserted between the Kex2-like cleavage site and the N-terminus of the IGF-I peptide in pFENIGFI to facilitate efficient cleavage of the IGF-I peptide from the catalytic domain of the glucoamylase protein. Complementary oligonucleotides GlySpacer-1 & GlySpacer-2 (Table 2.1.8.3) were designed to encode a penta-glycine repeat (using the *Neurospora* preferred GGC codon) downstream of the Kex2-like cleavage site. Annealing of GlySpacer-1 and GlySpacer-2 formed a double stranded oligonucleotide with single stranded overhangs corresponding to *NheI* and *SbfI* sticky ends. pFENIGFI was digested with *NheI* and *SbfI*, spin column purified and ligated with the double stranded glycine spacer oligonucleotide to form pFENIGFI-GS (Figure 3.7A). Integration of the glycine spacer into pFENIGFI resulted in the loss of a *StuI* restriction site enabling digestion with *StuI* to be used to screen for pFENIGFI-GS transformants. DNA sequencing was performed to confirm integration of the glycine linker and fidelity of the *nIGF-I* coding sequence.

#### 3.1.3.4.2 Construction of pNEIGFI-G1 (Figure 3.7B)

The sequence encoding the truncated glucoamylase protein was removed from pFENIGFI to determine if the fusion protein was responsible for the low abundance of mature *nIGF-I* (described in chapter 4). The endogenous glucoamylase signal peptide was fused directly to the first amino acid of the mature IGF-I peptide to facilitate secretion.

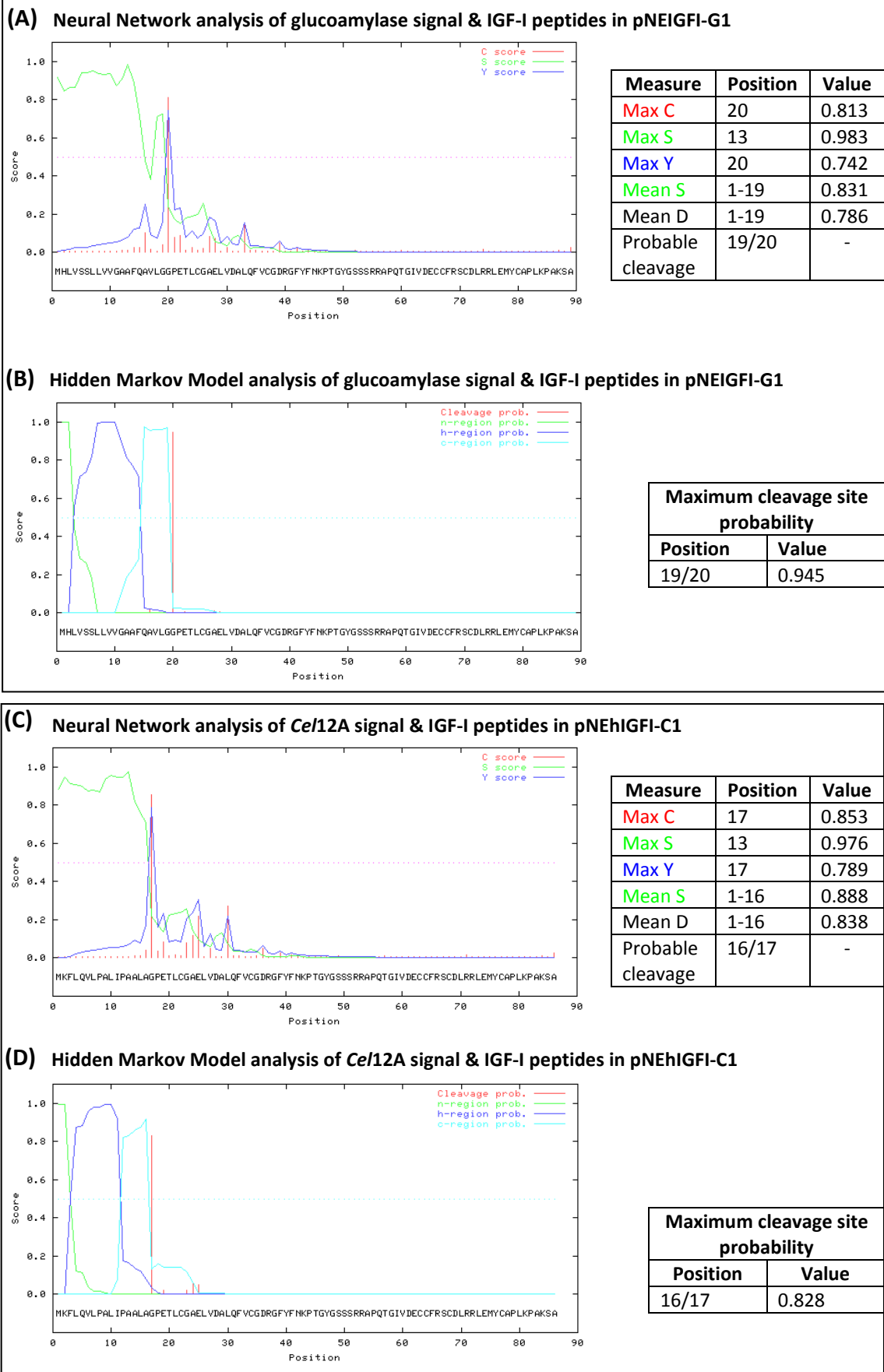
The IGF-I coding sequence was PCR amplified from pFENIGFI using RedHot DNA polymerase and primers IGFI-BseRI-F & IGFI-XbaI-R (Table 2.1.8.3). The IGFI-BseRI-F oligonucleotide was 73bases long with ~20bases of the 3' end complementary to the *nIGF-I* coding sequence whilst the remaining 5' sequence contained the non-complementary glucoamylase signal peptide coding sequence and an upstream *BseRI* restriction site. The PCR amplicon was digested with *BseRI* and *XbaI* restriction enzymes and purified with a spin column. The pFENIGFI vector was also digested with *BseRI* and *XbaI* to excise the *gla-1* & *nIGF-I* coding sequences and the vector backbone was purified by gel extraction. Ligation of the vector with the PCR amplicon encoding the glucoamylase signal peptide attached directly to the N-terminus of IGF-I formed

pNEIGFI-G1 (Figure 3.7B). Putative pNEIGFI-G1 *E. coli* transformants were screened by *Bst*EII digestion of plasmid DNA prior to DNA sequencing to confirm fidelity of the glucoamylase signal peptide and *nIGF-I* coding sequence.

#### *3.1.3.4.2.1 In-silico analysis of signal peptide cleavage in pNEIGFI-G1*

Signal peptides are N-terminal peptides typically 15-30 amino acids long which, in eukaryotes, facilitate the translocation of newly synthesized proteins into the endoplasmic reticulum (ER). Upon entry into the ER, the signal peptide is cleaved and the mature protein is secreted from the cell *via* the Golgi apparatus or vacuoles (Derby & Gleeson 2007). Signal peptides have highly divergent primary structures and are thought to mediate transport through the cell based on charge, hydrophobicity and conformational properties (Tsuchiya *et al.*, 2003). As such, signal peptides are divided into three structurally and functionally distinct regions: a positively charged n-terminal region, a central hydrophobic core (h-region) consisting of at least six residues and a c-terminal region of polar uncharged residues (Emanuelsson *et al.*, 2007). Substitutions in the amino acid sequence flanking the signal peptide cleavage site can result in aberrant or inhibited cleavage (Tsuchiya *et al.*, 2007).

An *in-silico* sequence analysis was performed using the SignalP 3.0 program (<http://www.cbs.dtu.dk/services/SignalP/>) to ascertain if the direct fusion of the glucoamylase signal peptide to the N-terminus of the IGF-I peptide encoded by pNEIGFI-G1 was likely to affect proper cleavage of the signal sequence. The SignalP 3.0 program utilises Neural Network and Hidden Markov Model machine learning methods trained on the SWISS-PROT database to predict the probable signal cleavage sites in a given sequence (Emanuelsson *et al.*, 2007). SignalP 3.0 analysis of the glucoamylase signal:IGF-I amino acid sequence predicted a high probability of the signal peptide being cleaved correctly (Figure 3.8).





**Figure 3.8. SignalP 3.0 analyses of probable signal peptide cleavage in pNEIGFI-G1 & pNEhIGFI-C1**

The predicted cleavage of glucoamylase or Ce/12A signal peptides from the N-terminus of IGF-I (encoded by pNEIGFI-G1 & pNEhIGFI-C1, respectively) were analysed using SignalP 3.0 with Neural Networks and Hidden Markov Models.

**(A)** and **(C)** Neural Network model analyses of pNEIGFI-G1 & pNEhIGFI-C1.

Neural Network graph key:

- C score:** A signal peptide cleavage site score for each amino acid. High scores are indicative of a possible cleavage site
- S score:** A high score indicates an amino acid may be a residue in a signal peptide.
- Y Score:** Derivative of combined C-scores & S-scores for each amino acid. High scores are indicative of a possible cleavage site.

Neural Network table key:

- Max C:** Residue with the maximum C score. The position value indicates the first amino acid of the mature protein.
- Max S:** Residue with the maximum S score.
- Max Y:** Derived from the maximum C-score and S-score to predict the probable signal peptide cleavage site.
- Mean S:** Mean of the S-scores for the length of the predicted signal peptide.
- Mean D:** Mean of the S-mean and Y-max scores for the length of the predicted signal peptide.

**(B)** and **(D)** Hidden Markov Model analyses of pNEIGFI-G1 & pNEhIGFI-C1.

Hidden Markov Model graph key:

- Cleavage prob:** Probability of an amino acid being a signal peptide cleavage site.
- n-region prob:** Probability of an amino acid being part of the positively charged n-region.
- h-region prob:** Probability of an amino acid being part of the hydrophobic h-region.
- c-region prob:** Probability of an amino acid being part of the neutral but polar c-region.

3.1.3.4.3 Construction of pNEhIGFI-C1 (Figure 3.7C)

As described in 3.1.2.1, the RIP-enhancement of the *IGF-I* coding sequence necessitated using non-optimal codons for *Neurospora* codon bias which may account for the low abundance of IGF-I obtained using pFENIGFI (described in Chapter 4). Comparison of codon usage between the human *IGF-I* (*hIGF-I*) coding sequence, RIP-enhanced *nIGF-I* coding sequence and the *Neurospora* codon bias-optimised *IGF-I* coding sequence are shown in Figure 3.9. The *hIGF-I* sequence consists of 26 codons more preferred by *Neurospora*, 12 that are less preferred and 32 codons that are the same in both sequences. Given the higher number of preferred codons in the human versus RIP-enhanced *IGF-I* coding sequences, the availability of a plasmid with a *hIGF-I* coding sequence and relative simplicity of construction, a vector was assembled with the *Cel12A* signal peptide and the *hIGF-I* coding sequences.

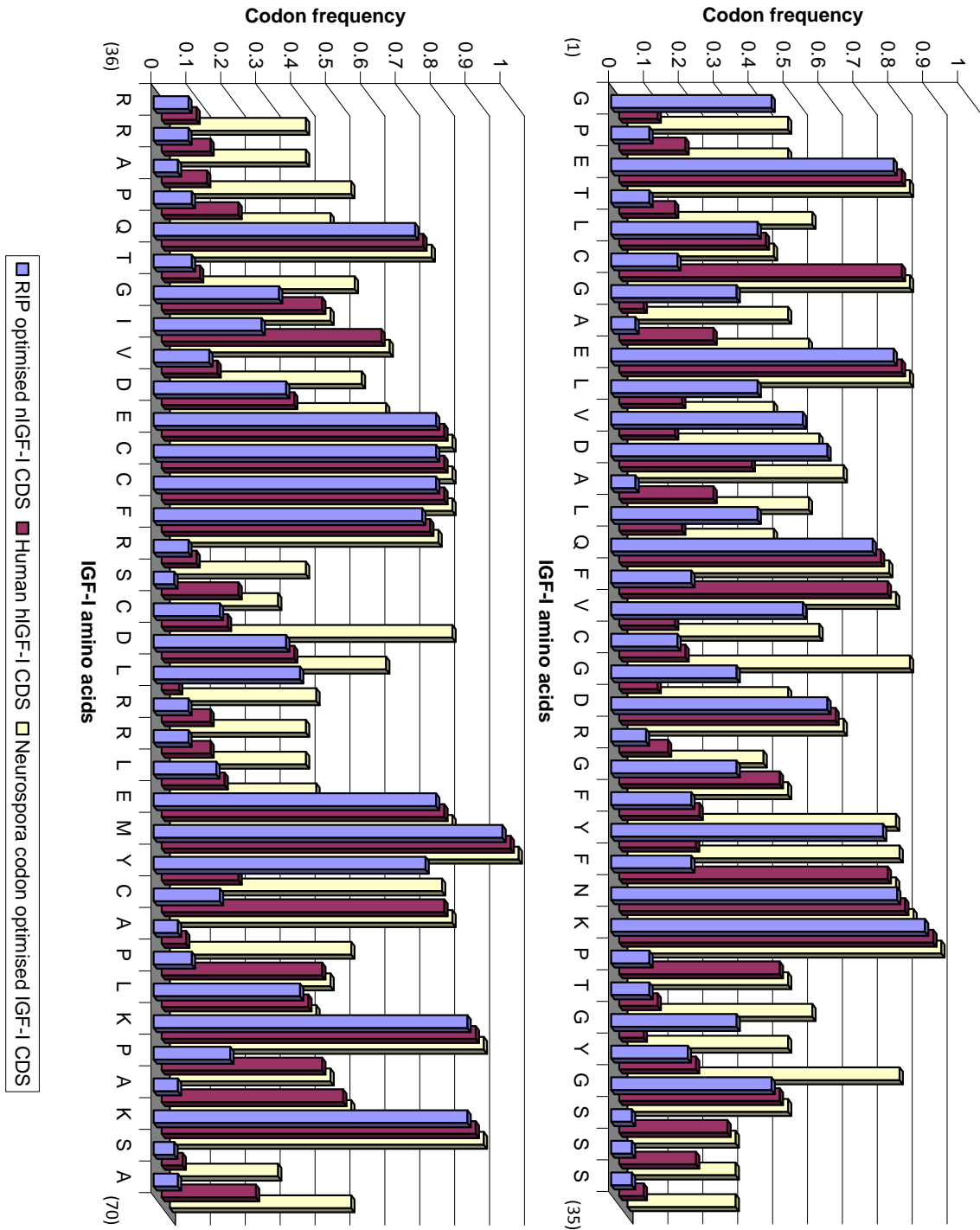
The presence of positively charged amino acids in the n-region of eukaryote signal peptides can be important for effective secretion of proteins (Tsuchiya *et al.*, 2003).

The glucoamylase signal peptide (MHLVSSLLVVGAAFQAVLG) contains a basic histidine residue in the n-region. However, as histidine has a pKa of ~6.1 and the pH of the *Neurospora* cytosol is ~7.15 (Legerton & Yanofsky 1985), the histidine in the signal sequence will predominately have a neutral charge which may not be optimal for secretion. The *Cel12A* (also known as EGIII) endoglucanase from *Trichoderma reesei* has been successfully secreted in *N. crassa* (Catcheside *et al.*, 2003). The *Cel12A* signal peptide (MKFLQVLPALIPAALA) contains a lysine residue in the n-region. As lysine has a pKa of 10.5, it will be positively charged in the cytosol of a *Neurospora* cell. As such, the *Cel12A* signal peptide was used in an additional vector as this signal sequence might be more effective for secretion than the glucoamylase signal peptide. SignalP 3.0 analysis predicted the expected cleavage point of both the glucoamylase and *Cel12A* signal peptides when fused to the *hIGF-I* coding sequence (Figure 3.8).

*hIGF-I* was PCR amplified from a plasmid kindly supplied by Dr. Morry Frenkel (CSIRO, Parkville, Australia) using RedHot DNA polymerase and primers hIGFI-BglII & hIGFI-NS-XbaI (Table 2.1.8.3). The hIGFI-BglII primer was 77 bases long with 18 bases of the 3' end complementary to the *hIGF-I* coding sequence whilst the remaining 5' sequence of the oligonucleotide contained the *Cel12A* signal peptide coding sequence and an upstream *BglII* restriction site. The PCR amplicon was digested with *BglII* and *XbaI* restriction enzymes and purified with a spin column. The pFENIGFI vector was also digested with *BglII* and *XbaI* to excise the *gla-1* & *nIGF-I* coding sequences and the vector backbone purified by gel extraction. Ligation of the vector with the PCR amplicon formed pNEhIGFI-C1 (Figure 3.7C). Putative pNEhIGFI-C1 *E. coli* transformants were screened by *SphI* & *XbaI* digestion of plasmid DNA prior to DNA sequencing to confirm fidelity of the *Cel12A* signal peptide and *hIGF-I* coding sequences

#### 3.1.3.4.4 Construction of pNEhIGFI-H1 (Figure 3.7D)

Heterologous signal peptides have been used to successfully secrete proteins in *N. crassa*. For example, (Nakano 1993) expressed and secreted enzymatically active bovine chymosin in *Neurospora* under the direction of its native secretion signal. As such, an additional expression vector was constructed with the coding sequence for the human IGF-I signal peptide fused directly to the mature IGF-I protein. The *hIGF-I* coding sequence was PCR amplified from a plasmid using primers hIGFI-NS-BglII & hIGFI-NS-XbaI (Table 2.8) with 5' non-complementary sequences containing *BglII* and *XbaI* restriction sites, respectively. The PCR amplicon was digested with *BglII* and *XbaI* restriction enzymes and spin column purified. The *gla-1* & *nIGF-I* coding sequences in the pFENIGFI vector were excised by *BglII* and *XbaI* digestion and the vector backbone was gel purified. Ligation of the vector with the PCR amplicon encoding the hIGF-I signal peptide and mature protein formed the vector pNEhIGFI-H1 (Figure 3.7D). Putative pNEhIGFI-H1 *E. coli* transformants were screened by *SphI* & *XbaI* digestion of plasmid DNA prior to DNA sequencing to confirm *hIGF-I* sequence fidelity.



**Figure 3.9 Comparison of codon usage frequencies between RIP-optimised *nIGF-I*, human *hIGF-I* and *Neurospora* codon bias-optimised *IGF-I* coding sequences.**

*Neurospora crassa* codon usage frequencies based on Edelman & Staben (1994).

3.1.3.4.5 Cloning of expression cassettes from pFENIGFI, pFENIGFI-GS, pNEIGFI-G1, pNEhIGFI-C1 & pNEhIGFI-H1 into pDV2 (Figure 3.10)

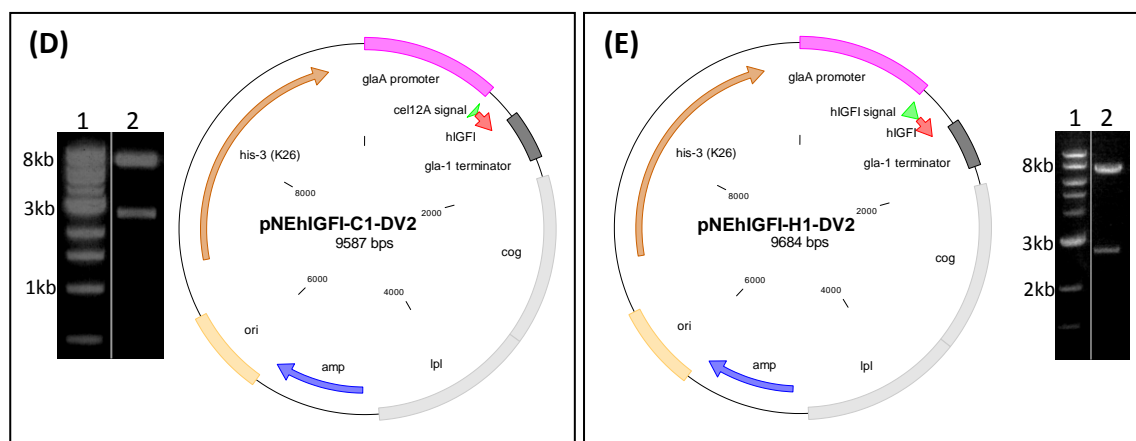
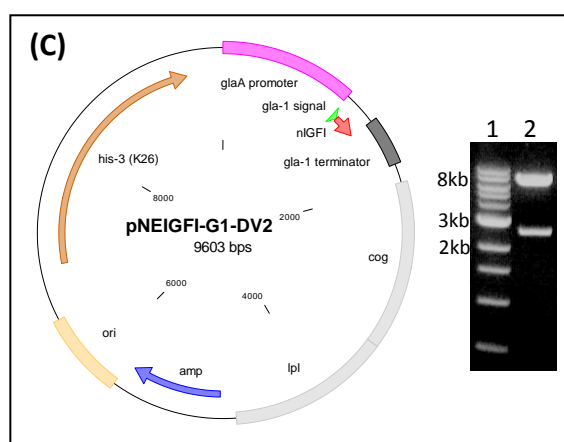
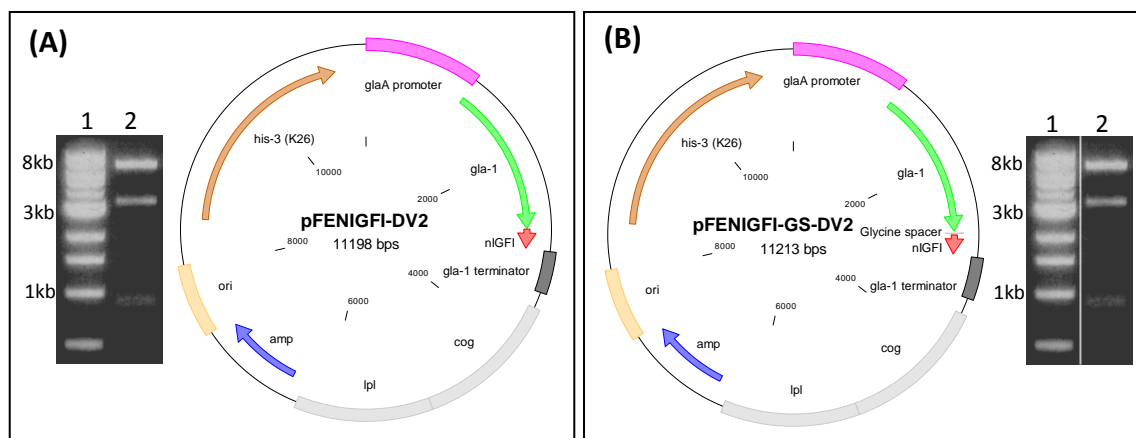
Direct comparison between the five IGF-I expression constructs generated during the course of this study required standardization of the chromosomal context of the expression cassettes once integrated into the *Neurospora* genome. Therefore, the five IGF-I vectors were cloned into the pDV2 diversification vector to enable *his-3* targeted transplacement. The locus encompassing *his-3* to *cog*<sup>+</sup> is highly suitable for heterologous gene expression because it is transcriptionally active (Catcheside *et al.*, 2003) and single copies of transforming DNA inserted downstream of *his-3* exhibit very low or no methylation (Windhofer *et al.*, 2000) which can potentially silence genes in *Neurospora* by inhibiting transcript elongation (Rountree & Selker 1997).

The expression cassettes in the pFENIGFI, pFENIGFI-GS, pNEIGFI-G1, pNEhIGFI-C1 and pNEhIGFI-H1 plasmids were excised by cleavage with *SpeI* restriction endonuclease and gel purified. The pDV2 vector was digested with *SpeI*, spin column purified and treated with phosphatase prior to ligation with the five purified expression cassettes. Plasmid DNA from *E. coli* transformants were cleaved with *Bam*HI and resolved by gel electrophoresis to screen for vectors with the expression cassettes oriented such that the promoter and terminator regions were juxtaposed to *his-3* and *cog*, respectively (Figure 3.10). This orientation facilitated downstream construction of a linked duplication vector for RIP mutagenesis of *nIGF-I* (described in 3.1.3.5). The appellation of these newly created IGF-I transplacement vectors were designated through the addition of the ‘-DV2’ suffix to the plasmid name.

**3.1.3.5 Construction of RIP tandem vector (pNEIGFI-G1-DV2-LD1) (Figure 3.11)**

A *nIGF-I* duplication was inserted into the pNEIGFI-G1-DV2 construct to characterize the frequency and distribution of RIP-mediated mutation of *nIGF-I* at the DNA sequence level. An ~800bp segment containing the *nIGF-I* coding sequence, *gla-1* terminator and ~200bp of the 5' end of the *cog* region was PCR amplified from pNEIGFI-G1-DV2 using primers nIGFI-DUD-F & nIGFI-DUD-R (Table 2.1.8.3). Both primers contained non-complementary 5' sequences creating *BsrGI* restriction sites on either flank of the amplicon to facilitate cloning between *his-3* and the *glaA* promoter in pNEIGFI-G1-DV2. The *nIGF-I* duplication was specifically integrated ~1.5kb upstream of the existing *nIGF-I* sequence as RIP between linked duplications is most efficient when separated by ~1kb or more (E. Cambareri: *Personal communication*).

Both the PCR product and the pNEIGFI-G1-DV2 vector were digested with *BsrGI*, spin column purified and ligated to create pNEIGFI-G1-DV2-LD1. Plasmids from putative *E. coli* transformants were screened by *PmeI* digestion and a construct with *nIGF-I* direct repeats was obtained and designated pNEIGFI-G1-DV2-LD1 (Figure 3.11). The sequence integrity of both duplicated regions was confirmed by DNA sequencing.



**Gel (A).** Lane 1: NEB 1kb DNA ladder; Lane 2: pFENIGFI-DV2 isolate digested with *Bam*HI.  
**Gel (B).** Lane 1: NEB 1kb DNA ladder; Lane 2: pFENIGFI-GS-DV2 isolate digested with *Bam*HI.  
**Gel (C).** Lane 1: NEB 1kb DNA ladder; Lane 2: pNEIGFI-G1-DV2 isolate digested with *Bam*HI.  
**Gel (D).** Lane 1: NEB 1kb DNA ladder; Lane 2: pNEhIGFI-C1-DV2 isolate digested with *Bam*HI.  
**Gel (E).** Lane 1: NEB 1kb DNA ladder; Lane 2: pNEhIGFI-H1-DV2 isolate digested with *Bam*HI.



**Figure 3.10 Maps of targeted transplacement vectors for expression of IGF-I in Neurospora.**

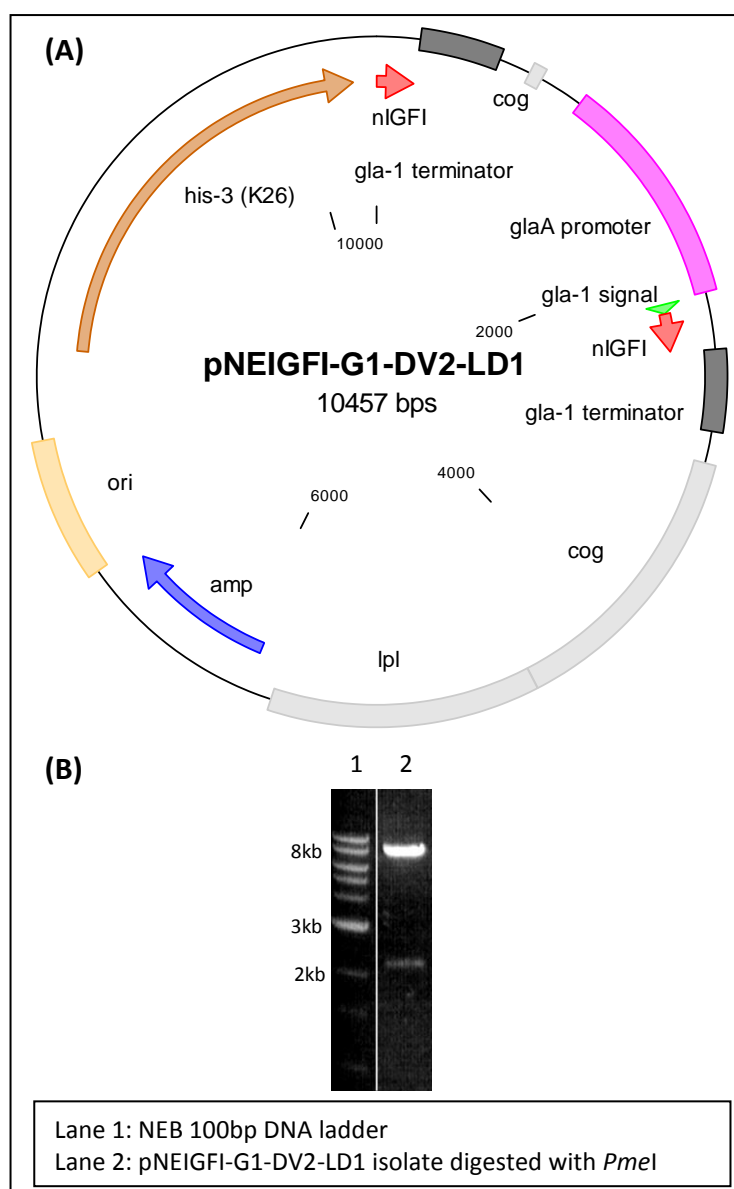
Vectors used to transplace *IGF-I* expression cassettes between *his-3* and *cog<sup>+</sup>*. **(A)** pFENIGFI incorporates a RIP-enhanced *IGF-I* coding sequence (*nIGF-I*) fused to the C-terminus of a truncated endogenous *gla-1* coding sequence. A Kex2-like cleavage site is present in the linker region between the fusion proteins to facilitate enzymatic separation of IGF-I and glucoamylase proteins prior to secretion. **(B)** pFENIGFI-GS consists of the same sequences as pFENIGFI with a flexible penta-glycine spacer inserted between the Kex2-like cleavage site and the *nIGF-I* coding sequence to facilitate protease access to the cleavage site. **(C)** pNEIGFI-G1 encodes a *Neurospora* glucoamylase secretion signal peptide attached to the N-terminus of the IGF-I protein. **(D)** pNEhIGFI-G1 contains a human IGFI (*hIGFI*) coding sequence with a fungal *Ce/12A* secretion signal (*Trichoderma reesei*). **(E)** pNEhIGFI-H1 contains the mature *hIGF-I* coding sequence and the native *hIGF-I* secretion signal.

All vectors contain *glaA* promoter (*A. niger*) and *gla-1* terminator (*N. crassa*) sequences flanking the coding sequences. A *gpdA* promoter (*A. nidulans*) coupled to a *hph* gene confers hygromycin resistance for fungal selection. *E. coli* propagation and selection is facilitated by an origin of replication (*ori*) and an *amp* gene encoding  $\beta$ -Lactamase for ampicillin resistance.

Confirmation of vector construction was determined by *Bam*HI restriction digest of plasmid DNA from candidate *E. coli* transformants. Agarose gel electrophoresis of the digest plasmids resulted in the following expected fragment sizes:

<b>(A)</b> pFENIGFI-DV2:	~7kb, ~3.2kb & ~0.88kb
<b>(B)</b> pFENIGFI-GS-DV2:	~7kb, ~3.2kb & ~0.88kb
<b>(C)</b> pNEIGFI-G1-DV2:	~7kb & ~2.5kb
<b>(D)</b> pNEhIGFI-C1-DV2:	~7kb & ~2.5kb
<b>(E)</b> pNEhIGFI-H1-DV2:	~7kb & ~2.6kb

*IGF-I* coding sequence integrity was determined by DNA sequencing in all constructs.



**Figure 3.11 Map of linked *nIGF-I* duplication vector for targeted transplacement between *his-3* and *cog*<sup>+</sup>.**

**(A)** The pNEIGFI-G1-DV2-LD1 construct was used to transplace linked RIP-enhanced *IGF-I* (*nIGF-I*) duplications between *his-3* and *cog*<sup>+</sup>. The vector contains an ~800bp duplication consisting of a *nIGF-I* coding sequence, a *gla-1* terminator (*N. crassa*) and ~200bp of *cog*<sup>+</sup> sequence. This duplication is upstream of the *nIGF-I* expression cassette which encodes the *Neurospora* glucoamylase signal peptide fused directly to the N-terminus of the IGF-I protein. The coding sequences are flanked by a *glaA* promoter (*A. niger*) and *gla-1* terminator (*N. crassa*) sequences. The duplicated *nIGF-I* & *gla-1* terminator sequences are separated by ~1400bp of non-duplicated sequence to facilitate efficient RIP-induced mutation. The vector also contains endogenous *his-3* (K26) and *cog/lpl* sequences to enable targeted transplacement of the *nIGF-I* duplication cassette into the *Neurospora* genome. *E. coli* propagation and selection is facilitated by an origin of replication (*ori*) and an *amp* gene encoding  $\beta$ -Lactamase for ampicillin resistance. **(B)** Confirmation of pNEIGFI-G1-DV2-LD1 construction was determined by *PmeI* restriction digestion of plasmid DNA from candidate *E. coli* transformants yielding ~8.2kb & ~2.3kb fragments when resolved by agarose gel electrophoresis.

### 3.2 Discussion

Five expression vectors were constructed for the purposes of heterologous expression of human *IGF-I* in *Neurospora*. All vectors contained an expression cassette with the *glaA* promoter and *gla-1* terminator as these transcription elements have been successfully used for the strong expression of heterologous coding sequences in *Neurospora* (E. Cambareri: *Personal communication*). The *glaA* promoter sequence was used in a BLASTn search of the *Neurospora crassa* genome sequence database (<http://www.broad.mit.edu/annotation/genome/neurospora/Home.html>) to ensure homologous sequences were not present as these could potentially disrupt promoter function in subsequent crosses through the action of RIP. The 400bp *gla-1* terminator is endogenous to *Neurospora* but is below the ~500bp minimum threshold for RIP.

Although desirable, it was not possible to optimise the *IGF-I* coding sequence for both RIP-mediated mutation and *Neurospora* codon bias as the codon sequence requirements were diametrically opposite. The decision to use RIP-optimised codons in the *nIGF-I* coding sequence was based on the preference for generating maximum *IGF-I* sequence diversity over potentially increasing protein abundance. The latter was initially considered to be of lesser importance due to the availability of a highly sensitive *IGF-I* receptor binding assay. As such, the frequency and distribution of potential RIP-mediated missense mutations in *nIGF-I* were greatly enhanced by re-engineering the *IGF-I* coding sequence with silent mutations to optimise the RIP preferred (CpA>>CpT>CpG>>CpC) dinucleotide target sites.

As the same *nIGF-I* allele was present in both strains of the diversification cross, *nIGF-I* sequence diversity was primarily generated by RIP prior to meiosis with HSMR expected to make a minor but important contribution during meiosis. HSMR subsequently recombines the RIP-mutagenised alleles facilitating potential recovery of advantageous mutations that would be lost if deleterious mutations were also present. For example, RIP mutagenesis of the CAG or CAA codons encoding glutamine invariably results in a nonsense mutation. As there are two glutamine residues in the *IGF-I* peptide and RIP preferentially targets CpA dinucleotides, the frequency of

truncated IGF-I proteins being produced may be relatively high. Thus, recombination within the *IGF-I* coding sequence could separate useful from deleterious mutations and increase diversity by separating other common mutations.

The initial aim of this project was to secrete and characterise heterologous IGF-I in *Neurospora*. A subsequent objective was the diversification of *IGF-I* by insertion of a linked *IGF-I* duplication expression vector between *his-3* and *cog<sup>+</sup>*. Thus, it would have been preferable to transplace the *IGF-I* expression cassettes to the locus between *his-3* and *cog<sup>+</sup>* to ensure consistency in the IGF-I secretion levels at all stages of the project. Unfortunately, the standard laboratory method used for transforming *Neurospora* was no longer possible as the procedure required spheroplasting of *Neurospora* macroconidia with Novozyme (Novolabs) and this enzyme was no longer being manufactured. This necessitated the use of electroporation as an alternative technique for *Neurospora* transformation. Early attempts at *Neurospora* electroporation by various laboratory members were relatively inefficient and the technique required optimisation over a period of time. Given transforming DNA typically integrates at ectopic sites in *Neurospora* (Rasmussen *et al.*, 2002) and the need for an *IGF-I* expression vector coincided with the period of transformation difficulties, pFENIGFI was constructed with features amenable for non-targeted transformation (the non-targeted integration of exogenous IGF-I sequences into the *Neurospora* genome are henceforth referred to as ectopic integrations) to expedite preliminary characterisation of IGF-I secretion.

The pFENIGFI(-DV2) and pFENIGFI-GS-DV2 vectors encoded the truncated glucoamylase protein attached directly to the IGF-I peptide as fusion to an endogenous protein can often increase secretion of a heterologous protein. A Kex2-like protease recognition site is present to facilitate cleavage of the fusion proteins prior to secretion. In pFENIGFI(-DV2), insertion of the *nIGF-I* coding sequence into the cloning site downstream of the Kex2-like cleavage site resulted in an additional 6bp of in-frame DNA sequence being added 5' of the *nIGF-I* coding sequence. The additional sequence was predicted to add proline and alanine residues to the N-terminus of the mature IGF-I protein after Kex2 cleavage. Similarly, the flexible glycine spacer used to

separate the Kex2 cleavage site from IGF-I in pFENIGFI-GS-DV2 added seven amino acids (GGGGGPA) to the N-terminus of IGF-I. These amino terminal extensions were not expected to significantly affect receptor binding as the eight amino acid FLAG epitope has previously been fused to the N-terminus without affecting receptor binding specificity (Zhang *et al.*, 1994). However, the amino terminal extensions, particularly the inflexible proline residue, may have affected cleavage efficiency at the Kex2 site.

The glucoamylase fusion protein strategy is potentially complicated by an unlinked duplication of the *gla-1* sequence as an endogenous *gla-1* sequence is present in *Neurospora*. Thus, the presence of a 1.6kb unlinked segment of duplicated *gla-1* sequence in the expression cassette will result in RIP-mediated mutation of *gla-1* in approximately 50% of the progeny. RIP of the *gla-1* sequence could result in nonsense mutations thereby truncating the glucoamylase protein before *IGF-I* translation. In addition, missense mutations in key glucoamylase residues could disrupt proper folding and/or protein trafficking of the fusion protein resulting in intracellular degradation. Efficient RIP-based gene diversification and secretion using a *gla-1* fusion strategy requires modification of the endogenous *gla-1* sequence using either molecular biology or genetic techniques. Deletion of the *gla-1* gene would require a purpose-built transplacement vector containing a selectable marker flanked by homologous sequence 5' and 3' of the *gla-1* locus. A genetic approach could exploit RIP's requirement for greater than ~80% sequence identity between duplications for detection (Cambareri *et al.*, 1991). An additional copy of the *gla-1* sequence could be ectopically transformed into the parent strains and repeatedly crossed together to induce RIP of *gla-1* until the sequences were sufficiently divergent so as to no longer be recognized by the RIP machinery. This would require considerable time to do the extensive crossing and sequence analysis required since RIP-mediated sequence divergence in repeated crosses with unlinked duplications is less severe than linked duplications (Cambareri *et al.*, 1991).

Although glucoamylase fusion to heterologous proteins generally results in increased protein yields, this is not invariably the case and is dependent upon the specific

heterologous protein (Gouka *et al.*, 1997a). As such, pNEIGFI-G1-DV2 was constructed with the truncated glucoamylase fusion protein removed and the glucoamylase signal peptide fused directly to the N-terminus of IGF-I. As discussed earlier, the *nIGF-I* coding sequence was not optimised for codon bias which could have hindered translation efficiency and potentially resulted in suboptimal protein yields. pNEhIGFI-C1-DV2 contained the *Cel12A* signal peptide fused to the *hIGF-I* coding sequence and contained a greater number of *Neurospora* preferred codons relative to the RIP-optimised *nIGF-I* sequence. In addition, pNEhIGFI-H1-DV2 was constructed with the *hIGF-I* coding sequence with its native signal peptide. All these vectors were made in an attempt to increase secretion levels of mature IGF-I. The five vectors were cloned into the pDV2 plasmid to facilitate targeted transformation between the *his-3* and *cog*<sup>+</sup> loci enabling a direct comparison between the constructs as described in chapter 4.

The pNEIGFI-G1-DV2-LD1 vector was constructed to characterize the RIP effects on linked *nIGF-I* repeats. As the 213bp *nIGF-I* coding sequence was less than the ~500bp minimum size required for recognition by RIP, the duplication was expanded to ~800bp by inclusion of the *nIGF-I* coding sequence, the *gla-1* terminator and ~200bp of the *cog*<sup>+</sup> region. RIP in the duplicated *cog* segment is unlikely to impair high levels of recombination initiating at this site as this region does not contain the 10bp sequence motif shown to be essential for the elevated recombination frequency phenotype (Yeadon *et al.*, 2001). Conversion tracts originating at *cog* decline exponentially with respect to the distance travelled (Yeadon *et al.*, 2002). Thus, the expression cassettes were specifically oriented within the pDV2 diversification vector to place the ‘expressed’ *nIGF-I* repeat nearest to *cog*<sup>+</sup> to exploit the higher recombination frequencies.

Expression of a RIP-mutated *IGF-I* coding sequence needs to account for RIP’s association with frequent but not invariable methylation of the remaining cytosines in duplicated sequences (Cambareri *et al.*, 1989). Methylation in *Neurospora* can reduce mRNA transcript levels by blocking transcription elongation resulting in gene silencing (Rountree & Selker 1997). This issue may be overcome in at least two ways. Firstly, methylation is completely eliminated in strains defective for *dim-2* (Kouzminova &

Selker 2001) or greatly reduced in strains bearing the *eth-1* thermosensitive mutation (Foss *et al.*, 1993). The absence of methylation has no effect on the mutagenic action of RIP (Kouzminova & Selker 2001). As such, either *dim-2* or *eth-1* mutant alleles could be crossed into the diversification strains to eliminate or reduce methylation overcoming this potential gene silencing problem. Alternatively, *Neurospora* could be grown in medium containing 5-azacytidine which substantially reduces DNA methylation (Selker & Stevens 1985). Although this approach is technically simpler, the use of 5-azacytidine is not ideal as this compound retards growth of liquid cultures (E. Cambareri: *Personal communication*) and has pleiotopic effects on gene expression (Christman *et al.*, 1985).

The re-engineering of the *IGF-I* coding sequence and construction of expression vectors facilitated the transformation and subsequent characterisation of IGF-I secretion and RIP-mediated mutagenesis in *Neurospora*.

## **Chapter 4      Transformation, expression and diversification of IGF-I in *Neurospora***

### **4.1      Introduction**

Heterologous protein secretion and screening of protein variants in culture supernatant was required for directed evolution of IGF-I within *Neurospora*. However, there are no previous reports of human IGF-I being expressed in *Neurospora crassa* or of the competitive IGF-IR ligand binding assay being used with complex mixtures of secreted proteins in culture media. The work described within this chapter focused on the following:

- Transformation of *Neurospora* with the IGF-I fungal expression vectors.
- Characterisation and optimisation of IGF-I heterologous expression in *Neurospora*.
- Optimisation of *Neurospora* growth media and culturing conditions for IGF-I secretion.
- Preliminary validation of the IGF-IR ligand binding assay for screening IGF-I in *Neurospora* culture supernatants
- Generation and characterisation of diversified IGF-I DNA sequences within *Neurospora*.

### **4.2      Transformation of *Neurospora* with pFENIGFI**

*Neurospora* strains t11630, t11644 and t12001 were utilised as hosts for the IGF-I expression vectors generated in this study. The former two strains are diversification strains generated in the Catcheside laboratory whilst the latter is a protease deficient strain (kindly provided by Neugenesis Corporation, California). Strain t12001 was generated by random mutagenesis and selection for reduced protease activity using a generic casein and gelatine plate assay. The specific mutation is referred to as *exp-1* although the proteolytic action of this protease is currently unknown. (E. Cambareri: *Personal communication*).



#### **4.2.1 Electroporation of *Neurospora* with pFENIGFI**

Initial attempts at IGF-I expression in *Neurospora* utilised the pFENIGFI expression vector (refer to 3.1.3.3) in an ectopic integration strategy as discussed in 3.2. The pFENIGFI construct contains the RIP-enhanced *nIGF-I* coding sequence fused to the 3' end of the truncated *gla-1* coding sequence and a hygromycin resistance gene for antibiotic selection. Approximately 1µg of pFENIGFI vector was electroporated into  $1 \times 10^8$  macroconidia from *Neurospora* strains t11630, t11644 & t12001 and plated onto selective medium (2.2.1.3.1). The number of primary transformants obtained for each strain was relatively low (<10 in each case) as this phase of the project coincided with the recent adaption and optimisation of the *Neurospora* electroporation technique (Vann 1995) in the Catcheside laboratory (as discussed in 3.2). Three primary transformants were selected for each transformed strain and subjected to microconidiation to obtain homokaryons.

#### **4.2.2 Isolation of pFENIGFI transformed *Neurospora* homokaryons**

Transformation of *Neurospora* macroconidia results in the formation of heterokaryons as integration into the same locus in multiple nuclei within a conidium is highly improbable. Thus, *Neurospora* homokaryons were obtained to ensure all nuclei within an isolate contained the transforming DNA. Isolation of homokaryons was achieved by induction of microconidiogenesis (2.2.1.2.4). A minimum of three putative transformed homokaryons were identified for each transformed strain. However, unequivocal determination of pFENIGFI integration in the putative transformants required subsequent molecular analyses.

#### **4.2.3 PCR confirmation of *Neurospora* transformation with pFENIGFI**

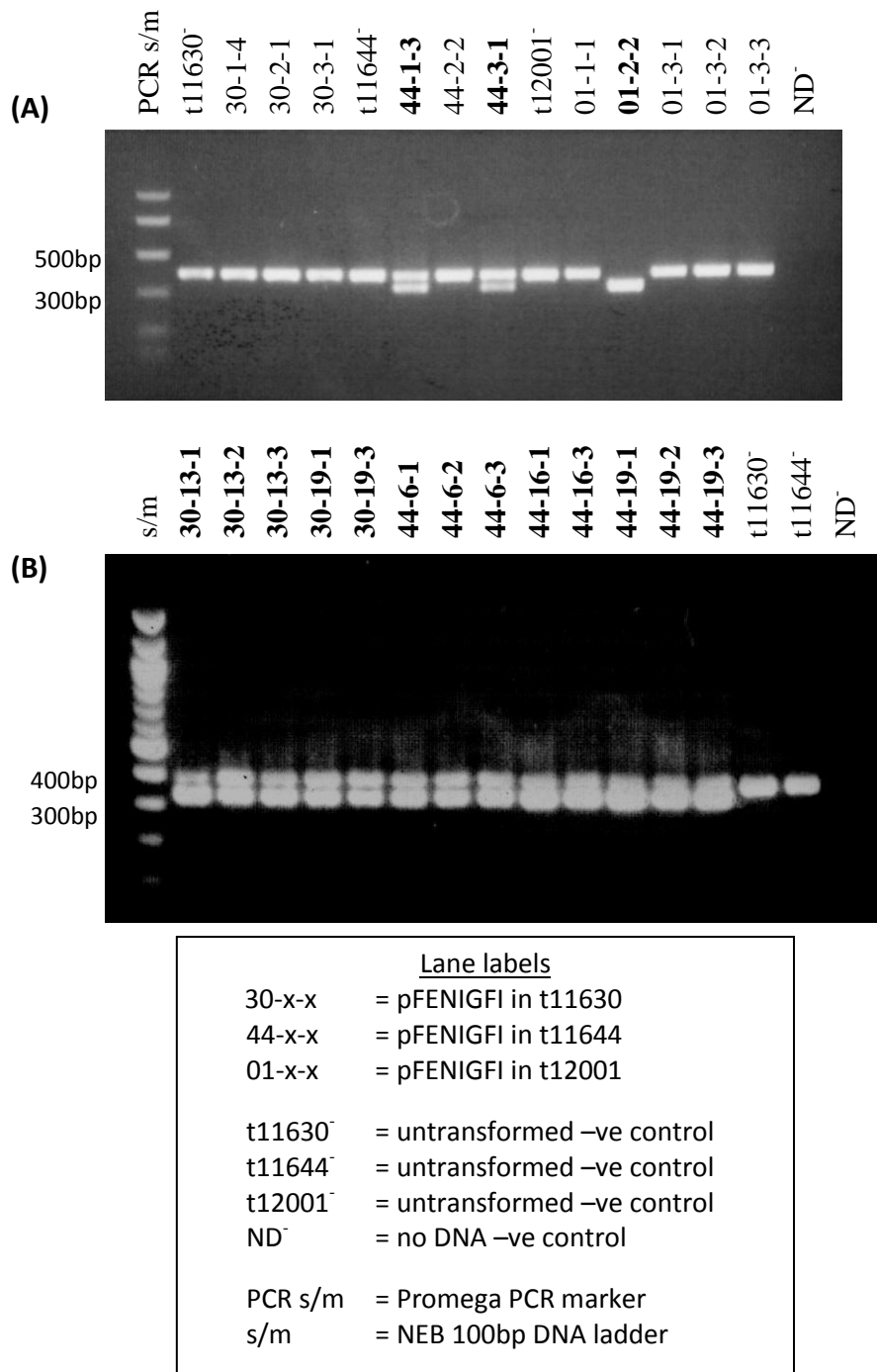
A PCR screen for pFENIGFI integration in the putative t11630, t11644 and t12001 transformants was performed using primers IGFI-check-F and IGFI-check-R (2.1.8.3). The complementary sequences to these oligonucleotides were located in the *gla-1* coding and terminator sequences flanking the *IGF-I* coding sequence resulting in an

~330bp amplicon from pFENIGFI transformants. These primers also amplified the endogenous *gla-1* sequence resulting in an additional ~400bp PCR product that served as an internal positive amplification control.

The PCR screen of the putative transformants isolated in 4.2.2 indicated none of the t11630 transformants, two of the t11644 (44-1-3 & 44-3-1) and one of the t12001 (01-2-2) transformants contained pFENIGFI (Figure 4.1A). The absence of the ~400bp endogenous *gla-1* positive control band in the 01-2-2 transformant and its presence in the t12001 host strain suggested the pFENIGFI vector may have undergone homologous transplacement into the endogenous *gla-1* locus.

#### **4.2.4 Repeat *Neurospora* transformation with pFENIGFI**

The absence of t11630 transformants and paucity of t11644 & t12001 transformants obtained from the PCR screen necessitated a repeat transformation of pFENIGFI into these host strains. Fungal electroporation was performed as described in 4.2.1 except the amount of pFENIGFI DNA electroporated into  $1 \times 10^8$  macroconidia was increased from ~1µg to ~4µg in an attempt to increase transformation efficiency. Unfortunately, whilst preparing macroconidia for electroporation a tube cracked during centrifugation resulting in the loss of the t12001 macroconidia. Given the 7 to 10-day lead time to obtain another large scale t12001 macroconidia culture, it was decided to proceed with electroporation of the t11630 and t11644 strains only. The repeated transformation resulted in hundreds of putative transformants for both strains. Twelve of the t11630 and t11644 primary transformants were subjected to microconidiation and screened to identify 5 and 8 hygromycin resistant homokaryons, respectively. PCR screening indicated all of these putative transformants contained the pFENIGFI construct (Figure 4.1B).



**Figure 4.1 PCR screen for pFENIGFI ectopic insertion in *Neurospora* t11630, t11644 & t12001 transformants.**

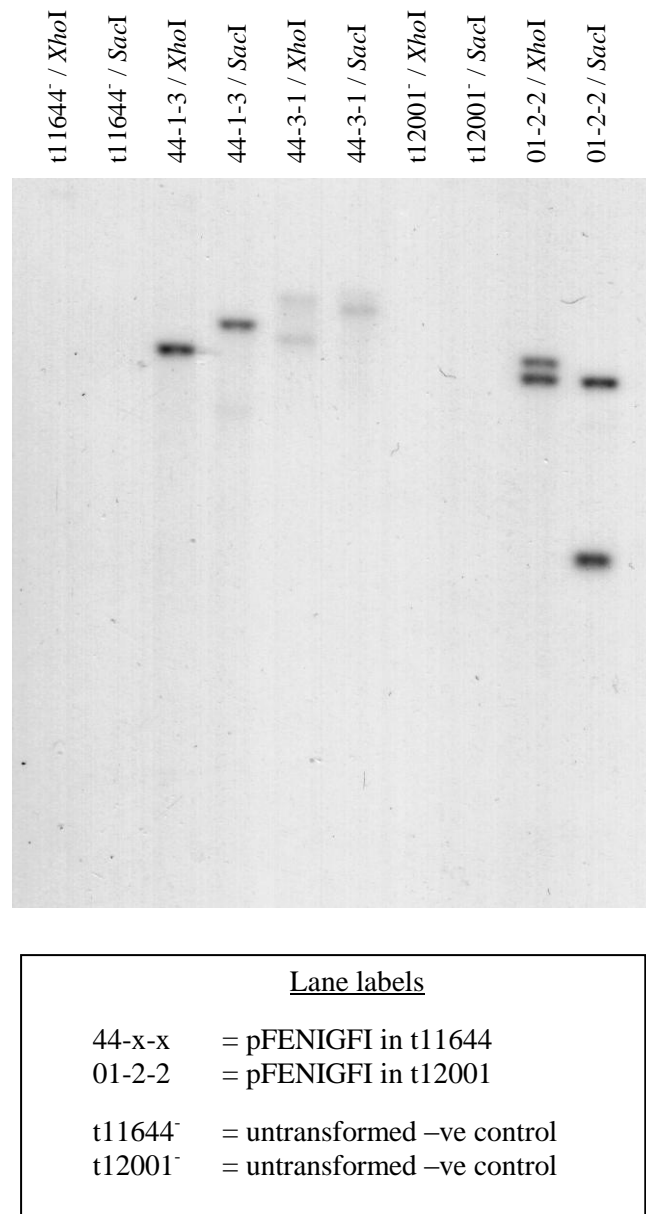
Genomic DNA extracted from hygromycin resistant *Neurospora* homokaryons isolated after transformation with pFENIGFI were PCR screened with primers located in the *gla-1* coding and terminator sequences flanking the *nIGF-I* coding sequence. PCR with these primers amplified regions in both the transformed pFENIGFI and endogenous *gla-1* sequences resulted in ~330bp and ~400bp amplicons, respectively (lane labels bolded). **(A)** A PCR screen of the initial pFENIGFI transformants identified three positive candidates: 44-1-3, 44-3-1 and 01-2-2. **(B)** PCR screening of transformants from a second transformation experiment indicated all homokaryons analysed were positive pFENIGFI ectopic transformants.

#### **4.2.5 Southern blot confirmation of pFENIGFI integration in ectopic *Neurospora* transformants**

The processes of *Neurospora* electroporation, microconidiation and genetic isolation of transformed homokaryons required a period of ~40 days to complete. As such, the initial batch of ectopic t11644 and t12001 transformants obtained in 4.2.3 were characterised by Southern blot whilst additional t11630 and t11644 transformants were being generated (4.2.4).

PCR screening of putative transformants indicated the presence of the pFENIGFI vector but not the number of plasmid copies integrated into the *Neurospora* genome. To determine copy number, the 44-1-3, 44-3-1 & 01-2-2 transformants were analysed by Southern blot (Figure 4.2) (2.2.4.7). The Southern blot strategy consisted of digesting genomic DNA from each transformant separately with *Xho*I and *Sac*I and probing with <sup>32</sup>P labelled 200bp IGF-I. As the pFENIGFI vector does not contain any *Xho*I restriction sites, digestion with this enzyme cleaves *Xho*I sites in the genomic DNA sequences flanking the vector resulting in a single band of variable size for each plasmid copy integrated at different loci within the *Neurospora* genome. The *Sac*I enzyme cuts within ~30bp of the 5' end of the *IGF-I* coding sequence and was utilised to determine if transformants had tandem integrations. In the absence of a tandem integration event, cleavage with *Sac*I was expected to result in two bands for each vector copy. *Sac*I digestion of a transformant with a tandem integration should have produced three bands.

The Southern blot of the 44-1-3/*Xho*I and *Sac*I digests resulted in one and two bands, respectively, indicating transformant 44-1-3 had a single copy of pFENIGFI. However, there was considerable difference between the densities of the two bands in the *Sac*I digest. As *Sac*I digestion of the pFENIGFI vector cleaved the IGF-I probe target sequence into ~170bp and ~30bp fragments, the differential band density probably reflected higher affinity of the *IGF-I* probe for the ~170bp target sequence.



**Figure 4.2 Southern blot analysis of putative ectopic pFENIGFI *Neurospora* transformants.**

Genomic DNA from putative transformants obtained by PCR screening of the first pFENIGFI/*Neurospora* electroporation experiment were digested with *XhoI* and *SacI* and probed with a <sup>32</sup>P-labelled *nIGF-I* coding sequence to confirm pFENIGFI integration.

The 44-3-1 *Xho*I digest resulted in two bands which was indicative of two pFENIGFI copies. However, the *Sac*I digest of 44-3-1 only produced a single band when at least four bands were predicted as two vectors were present. Both the *Xho*I & *Sac*I digests of 44-3-1 genomic DNA resulted in fainter bands than the 44-1-3 and 01-2-2 transformants which suggested less DNA was present on the gel. This would have further reduced the relatively low signal obtained from the 30bp fragment of the IGF-I target DNA accounting for the absence of two of the four expected bands. Furthermore, the higher density of the single *Sac*I band relative to the two bands in the *Xho*I digest indicated this band was probably two co-migrating DNA fragments of similar size.

The presence of two bands in the 01-2-2/*Xho*I digest suggested this transformant had two copies of pFENIGFI. The 01-2-2/*Sac*I digestion resulted in two strong bands instead of the four bands predicted. As described previously, the absence of the additional two bands was attributed to poor signal from the 30bp IGF-I target sequence. Although the Southern blot analysis of *Xho*I digested transformant DNA enabled the number of pFENIGFI insertion sites within the three transformants to be determined, attempts to identify possible tandem integrations with *Sac*I digested DNA resulted in equivocal data. As such, it was only possible to ascertain the minimum number of pFENIGFI copies within the genome of the *Neurospora* transformants.

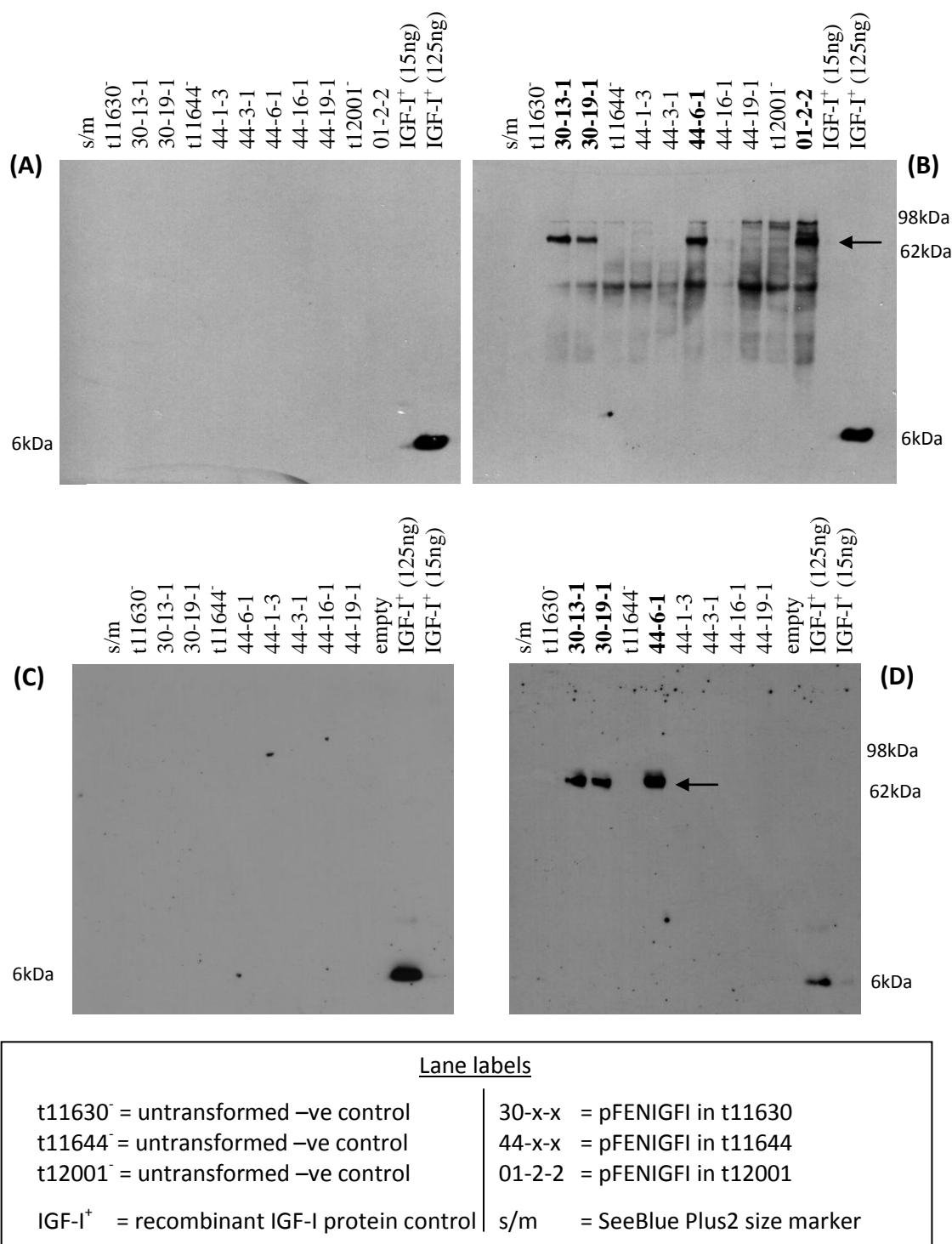
#### **4.2.6 Western blot analysis of extracellular and intracellular protein from ectopic pFENIGFI *Neurospora* transformants**

Western blot analysis of culture medium and intracellular protein from ectopic pFENIGFI *Neurospora* transformants was performed to determine if these transformants were correctly processing the glucoamylase:IGF-I fusion protein and secreting mature IGFI protein (Figure 4.3). pFENIGFI integration in transformants 01-2-2, 44-1-3 and 44-3-1 had been characterised by PCR and Southern blot. pFENIGFI integration in transformants 30-13-1, 30-19-1, 44-6-1, 44-16-1 & 44-19-1 had been demonstrated by PCR (Figure 4.1B), however no information on plasmid copy number was available as these transformant were obtained in a second transformation experiment (4.2.4) and Southern blot analysis had not been conducted. These

transformants were included to facilitate a more comprehensive analysis of *Neurospora* IGF-I secretion as the different chromosomal context of the ectopic pFENIGFI integrations were likely to have affected the *nIGF-I* expression levels (Collins *et al.*, 1991).

Duplicate *Neurospora* liquid cultures were prepared in 125ml Erlenmeyer flasks with 25ml of either VmH or *Neurospora* secretion medium (NSm) supplemented with L-histidine (NSmH) (2.2.1.1.4). Vm is widely used for *Neurospora* liquid cultures whilst NSm is a variant formulation of Vm typically increasing *Neurospora* protein secretion by 2 to 4-fold relative to Vm (E. Cambareri: *Personal communication*). Liquid media were inoculated with  $2 \times 10^5$  of 6-10 day-old macroconidia and incubated at 25°C in the absence of light for 72 hours on an orbital shaker. Culture media and mycelia were harvested as described in 2.2.5.1.1 and 2.2.5.1.2, respectively. The t12001 control strain and 01-2-2 transformant were nonviable in NSmH yet produced strong growth in VmH cultures which suggested a strain-specific sensitivity to one or more components in the NSmH.

Extra and intra-cellular protein samples from each of the transformants and controls were resolved by SDS-PAGE (2.2.5.5.2) and analysed by western blot (2.2.5.6) using an anti-human IGF-I monoclonal antibody (2.1.6). The total protein from the transformants and controls were not standardised as the objective of this particular experiment was not to directly compare IGF-I levels between transformants but to determine if mature protein was being processed and secreted at detectable levels by any of the transformants. No mature protein was detected in the supernatant or mycelia with either medium (Figure 4.3). The 125ng IGF-I positive controls produced a strong band but the 15ng IGF-I controls were barely detectable. As 11.7µl of protein sample was loaded into each well, this indicated the western blot readily detected secretion levels of 10.7µg of IGF-I per ml of supernatant with a maximum sensitivity of 1.3µg/ml. As the total protein produced by *Neurospora* cultures in the VmH and NSmH after 72 hours typically ranged from 10-40µg/ml, IGF-I secretion would need to have constituted a minimum of ~3-13% of total secreted protein to be detected by western blot.



**Figure 4.3 Western blot of secreted and intracellular protein from ectopic pFENIGFI *Neurospora* transformants.**

*Neurospora* liquid cultures were grown for 72 hours in VmH or NSmH media. Western blots of secreted protein and intracellular protein were probed with a monoclonal anti-human IGF-I antibody. **(A)** Liquid medium from transformants cultured in VmH. **(B)** Mycelia protein extracted from transformants culture in VmH. **(C)** Liquid medium from transformants cultured in NSmH. **(D)** Mycelia protein extracted from transformants culture in NSmH. Note: putative uncleaved gla-1:IGF-I fusion protein bands are indicated by arrows and associated lane labels are bolded.



Western analysis of intracellular protein in VmH cultures resulted in some non-specific antibody binding as demonstrated by the presence of similar bands in the t11644 and t12001 negative control strains. However, there were ~68kDa bands present in the Vm cultures of the 30-13-1, 30-19-1, 44-6-1 & 01-2-2 transformants that were absent in the t11630, t11644 & t12001 control strains (Figure 4.3B). Similarly, the NSmH cultures had an ~68kDa band present in the same samples with the exception of 01-2-2 which was not viable in the NSm (Figure 4.3D). The ~68kDa band in the NSm cultured samples was distinctive as there was no non-specific antibody binding. The ~68kDa band approximated the ~64kDa predicted molecular weight of the gla-1:IGF-I fusion protein and the absence of this band in the negative controls indicated this protein was probably uncleaved gla-1:IGF-I fusion protein.

**4.2.6.1 Western blot of TCA precipitated extracellular protein from ectopic pFENIGFI Neurospora transformants**

Culture supernatants of pFENIGFI transformants 30-13-1, 30-19-1, 44-6-1 & 01-2-2 were TCA precipitated to determine if mature IGF-I was being secreted by these transformants at levels below the sensitivity of the western blot. These transformants were selected as the most promising candidates as they contained detectable intracellular levels of putative uncleaved gla-1:IGF-I protein.

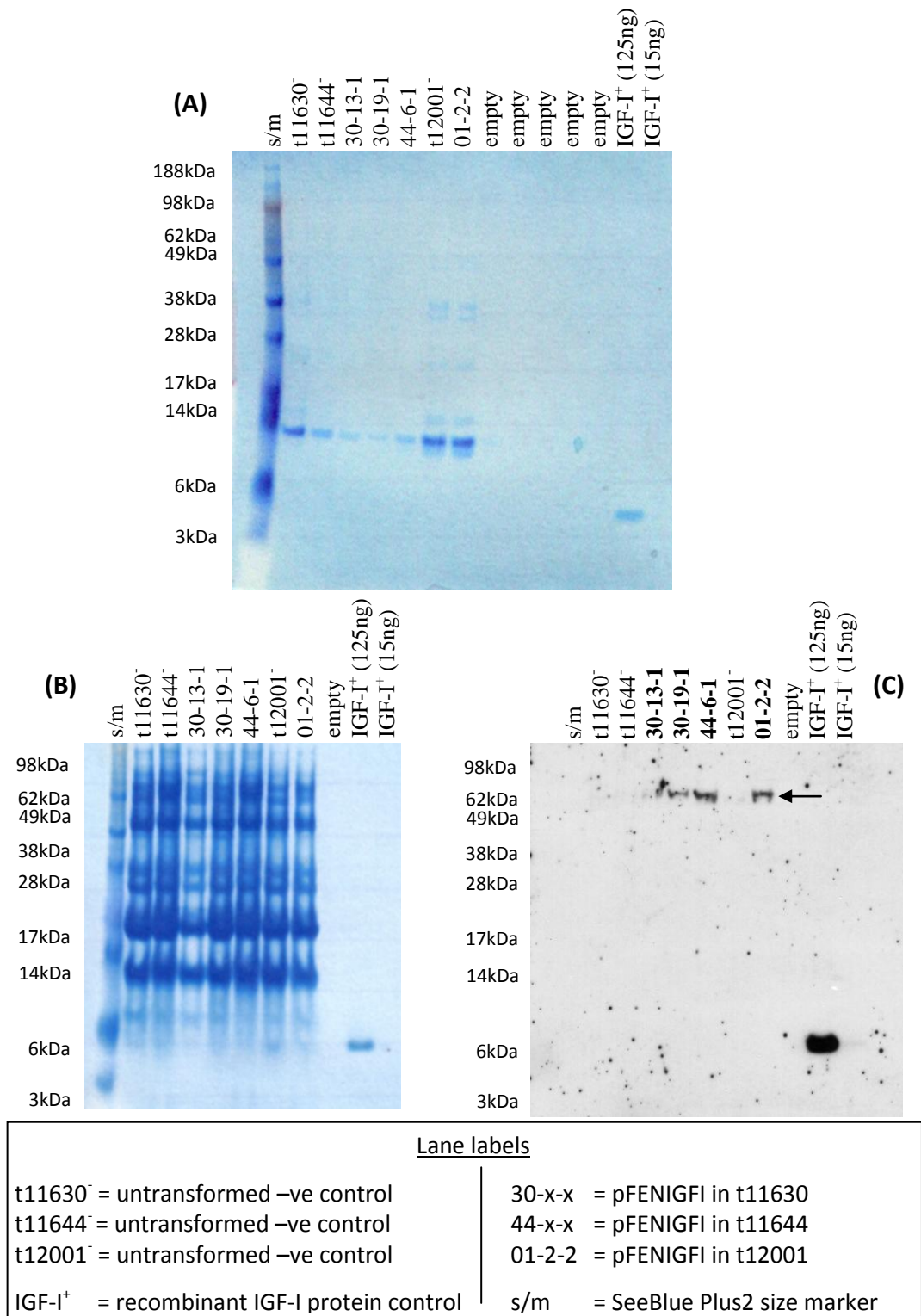
Liquid cultures of these transformants and controls were prepared as detailed in 4.2.6 using VmH only as the t12001 strain and 01-2-2 transformant were non-viable in NSmH. One millilitre of culture supernatant was TCA precipitated and resuspended in 40µl of SDS-PAGE sample buffer (2.2.5.4), resolved by SDS-PAGE (2.2.5.5.2) and analysed by western blot (2.2.5.6) using a monoclonal anti-human IGF-I antibody (2.1.6) (Figure 4.4). Although no mature IGF-I protein was visible in the western blot, all the transformants contained the putative uncleaved gla-1:IGF-I fusion protein. This result indicated that uncleaved fusion protein was being secreted. TCA precipitation of the culture medium concentrated the protein by up to a maximum of 25-fold. The minimum sensitivity of the western blot was ~1.3µg/ml, therefore, a conservative estimation of secreted uncleaved fusion protein was determined to be in the order of

40ng/ml. As the levels of IGF-I secretion were below the sensitivity of the western blot, the culture supernatants were subjected to a more sensitive IGF-IR ligand binding assay.

#### **4.2.7 IGF-IR ligand binding assay of culture supernatant from ectopic pFENIGFI *Neurospora* transformants**

The IGF-IR ligand binding assay used in this study was based on dissociation enhanced time resolved fluorescence. The assay utilised a 96-well plate coated with an anti-IGF-IR antibody (24-31) (kindly supplied by Professor Ken Siddle, University of Cambridge, United Kingdom) (2.1.6) to capture human IGF-IR in P6 cell lysates. The IGF-I ligand was covalently labelled with a non-fluorescent europium-N<sup>1</sup>-(p-isothiocyanatobenzyl)-diethylenetriamine-N<sup>1</sup>,N<sup>2</sup>,N<sup>3</sup>-tetraacetic acid (Eu-DTTA) chelate to form Eu-IGF-I (kindly supplied by Dr. Peter Hoyne, CSIRO, Parkville, Australia). Eu-IGF-I was used in competitive IGF-IR binding with unlabelled IGF-I ligand present in the *Neurospora* culture supernatants. The plate wells were washed to remove unbound ligand prior to the addition of an acidic chelating detergent solution which dissociated Eu<sup>3+</sup> ions from the Europium chelate-labelled IGF-I forming a homogeneous fluorescent Eu-(2-NTA)-<sub>3</sub>(TOPO)<sub>2-3</sub> micellar chelate solution. This fluorescent chelate enhanced the lanthanide fluorescence by up to 10<sup>7</sup>-fold which enabled highly sensitive measurements to be obtained with a time-resolved fluorometer (Handl & Gillies 2005).

IGF-IR ligand binding assays of *Neurospora* supernatant from cultures grown in VmH resulted in minimal fluorescence in all *Neurospora* samples assayed including VmH controls where the absence of unlabelled IGF-I should have resulted in maximum fluorescent signal (data not shown). Ligand binding buffer (LBB) only controls exhibited substantial fluorescence suggesting one or more components of the VmH were interfering with the assay. As such, alternative growth media formulations were tested to identify a medium compatible with the ligand binding assay.



**Figure 4.4** SDS-PAGE and western blot of TCA-precipitated secreted protein in media from ectopic pFENIGFI *Neurospora* transformants.

(A) *Neurospora* liquid cultures were grown for 72 hours in VmH and samples were electrophoresed on a 12% Bis-Tris SDS-PAGE gel. (B) Secreted proteins in culture media were TCA-precipitated and resolved on a 12% Bis-Tris SDS-PAGE gel and (C) subjected to western blot analysis with a monoclonal anti-human IGF-I antibody. Note: putative uncleaved gla-1:IGF-I fusion protein bands are indicated by an arrow and associated lane labels are bolded. The fusion band in transformant 30-13-1 was very faint in the scanned image but was more clearly visible on the original blot.

**4.2.7.1 Adaption of Neurospora culture media for use in the IGF-IR ligand binding assay**

Lanthanide chelates are generally robust and stable although they are vulnerable to chelating agents that alter the chelate structure such as EDTA and sodium citrate (Kohek *et al.*, 2002). In addition, transition metal ions such as  $Mn^{2+}$ ,  $Fe^{2+/3+}$  and  $Cu^{2+}$  can quench fluorescence by a mechanism presumed to involve electronic energy transfer from the donor lanthanide (Arakawa & Akamine 2003); (Barja 2005). As the VmH included both sodium citrate and a trace element solution containing  $CuSO_4$ ,  $MnSO_4$  and  $Fe(NH_4)_2(SO_4)_2$  (refer to 2.1.10.18), an IGF-IR ligand binding assay was performed using different formulations of VmH with and without sodium citrate and the trace element solution. An alternative culture medium, FRIES medium (Fm) supplemented with L-histidine (FmH), contained neither chelating agents nor transition metals (2.1.10.3) and was assayed to determine if this medium inhibited the IGF-IR ligand binding assay.

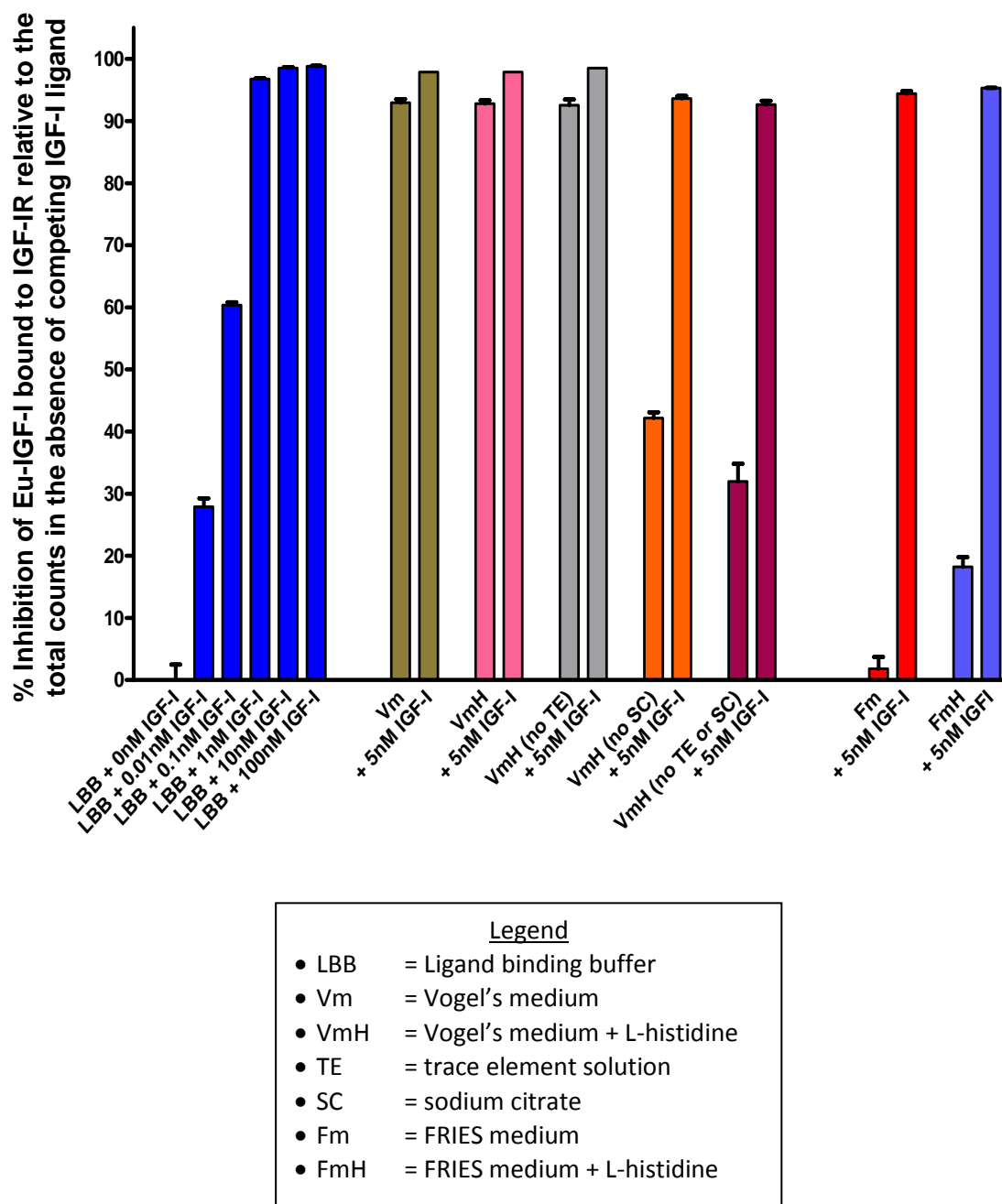
The L-histidine added to the medium was also considered to be a potential problem as this amino acid is capable of forming coordinated complexes with  $Eu^{3+}$  in solution (de Andrade Leite *et al.*, 1999). Although the total coordination number for  $Eu^{3+}$  is approximately  $8.6 \pm 0.3$  (Choppin 1997), DTTA-derivatives are only 7-dentate ligands (Mukkala *et al.*, 1989). Thus, it was theorised the L-histidine could have potentially bound to the  $Eu^{3+}$  ion in the Eu-DTTA chelate and may have subsequently interfered with the correct formation of the fluorescent  $Eu-(2-NTA)_3(TOPO)_{2-3}$  micellar chelate. To investigate this possibility, Vm and Fm without L-histidine supplementation were also assayed.

The results of the IGF-IR ligand binding assay of the various VmH formulations (Figure 4.5) indicated the sodium citrate was primarily responsible for the interference with the assay with the trace elements having a lesser but significant effect. The background fluorescence level in the FmH control approximated the level seen in the LBB control. The FmH control had a lower fluorescent signal relative to the Fm control which indicated L-histidine also interfered with the IGF-IR ligand binding assay.

However, L-histidine was an essential requirement for all auxotrophic *Neurospora* strains used in this study and could not be removed from the media formulations. The fluorescence signal in FmH was greater than VmH (no SC or TE) suggesting there were additional components in the VmH that interfered with the assay. As such, the FmH was deemed to be more suitable than VmH (no SC or TE) for subsequent IGF-IR ligand binding assays. Duplicate IGF-I standard curves were prepared separately in LBB and FmH for comparison of the standard IGF-IR ligand binding assay with the *Neurospora*-optimised assay (Figure 4.6). The standard curve in FmH still exhibited the expected sigmoidal dose-response curve. However, the linear dynamic range of the standard curve in FmH relative to LBB was reduced by ~3-fold (~2nM : ~50nM vs. ~25pM : ~2nM) whilst the sensitivity was determined to be ~80-fold less (~2nM vs. ~25pM). Despite the sensitivity loss, the IGF-IR ligand binding assay still exhibited at least an 85-fold greater sensitivity than the western blot assay with the anti-IGF-I monoclonal antibody (as described in 4.2.6).

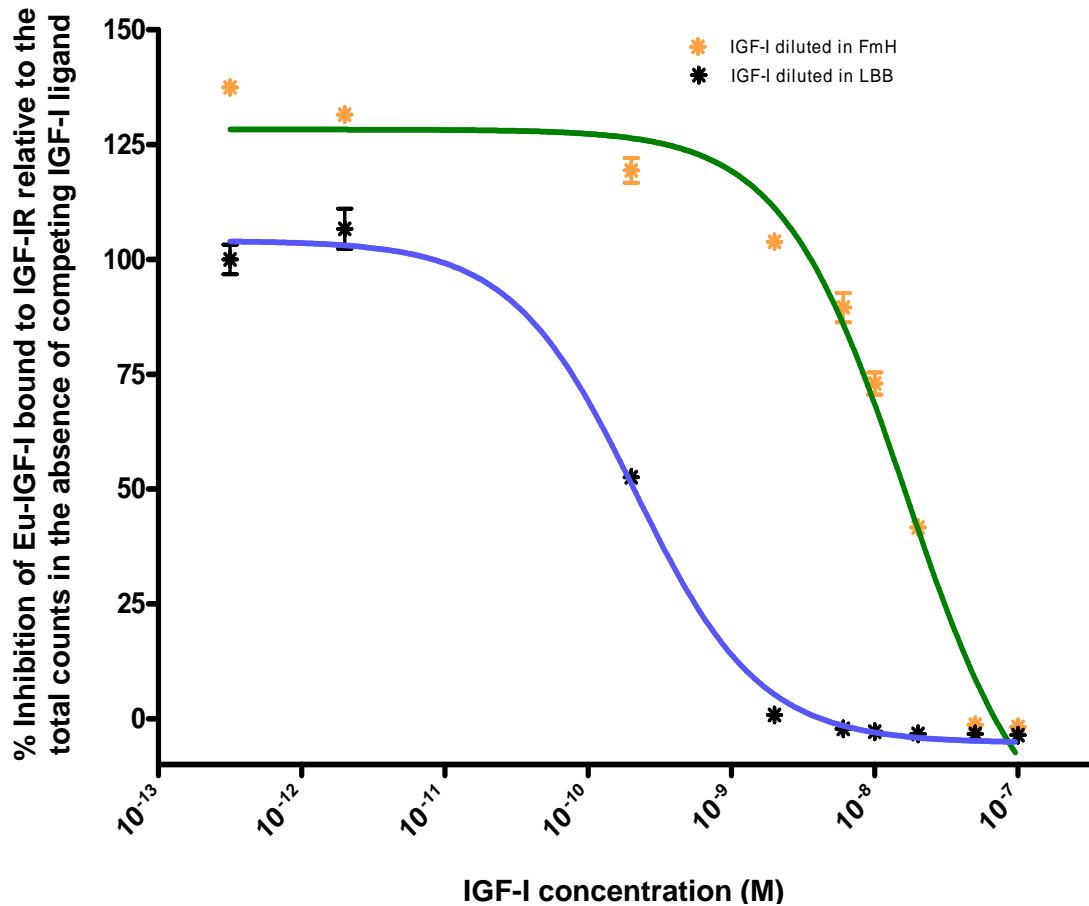
#### **4.2.7.2 IGF-IR ligand binding assay of culture supernatant from ectopic pFENIGFI *Neurospora* transformants in FRIES medium**

Liquid cultures of ectopic pFENIGFI *Neurospora* transformants 30-13-3, 30-19-1, 44-6-1 & 01-2-2 in FmH were grown for 72 hours and the culture supernatants were analysed by IGF-IR ligand binding assay with an IGF-I standard curve diluted in FmH (Figure 4.7). The IGF-I ligand binding assay involved the addition of equal volumes of sample/standard and Eu-IGFI solutions to each well. As the concentration of IGF-I standards were based on the total volume in each well, the measured IGF-I concentrations in the *Neurospora* samples were doubled to correct for the assay dilution factor. The background fluorescence levels of the t11630, t11644 & t12001 negative controls were subtracted from the relevant transformant measurements. Transformants 30-19-1 and 01-2-2 had measured IGF-I concentrations of 3.1ng/ml and 19.1ng/ml, respectively. However, transformants 30-13-1 & 44-6-1 did not produce detectable levels of IGF-I. A receptor phosphorylation assay was performed in an attempt to confirm the results of the IGF-IR ligand binding assay and to determine if the *Neurospora* IGF-I activated the kinase function of the IGF-IR.



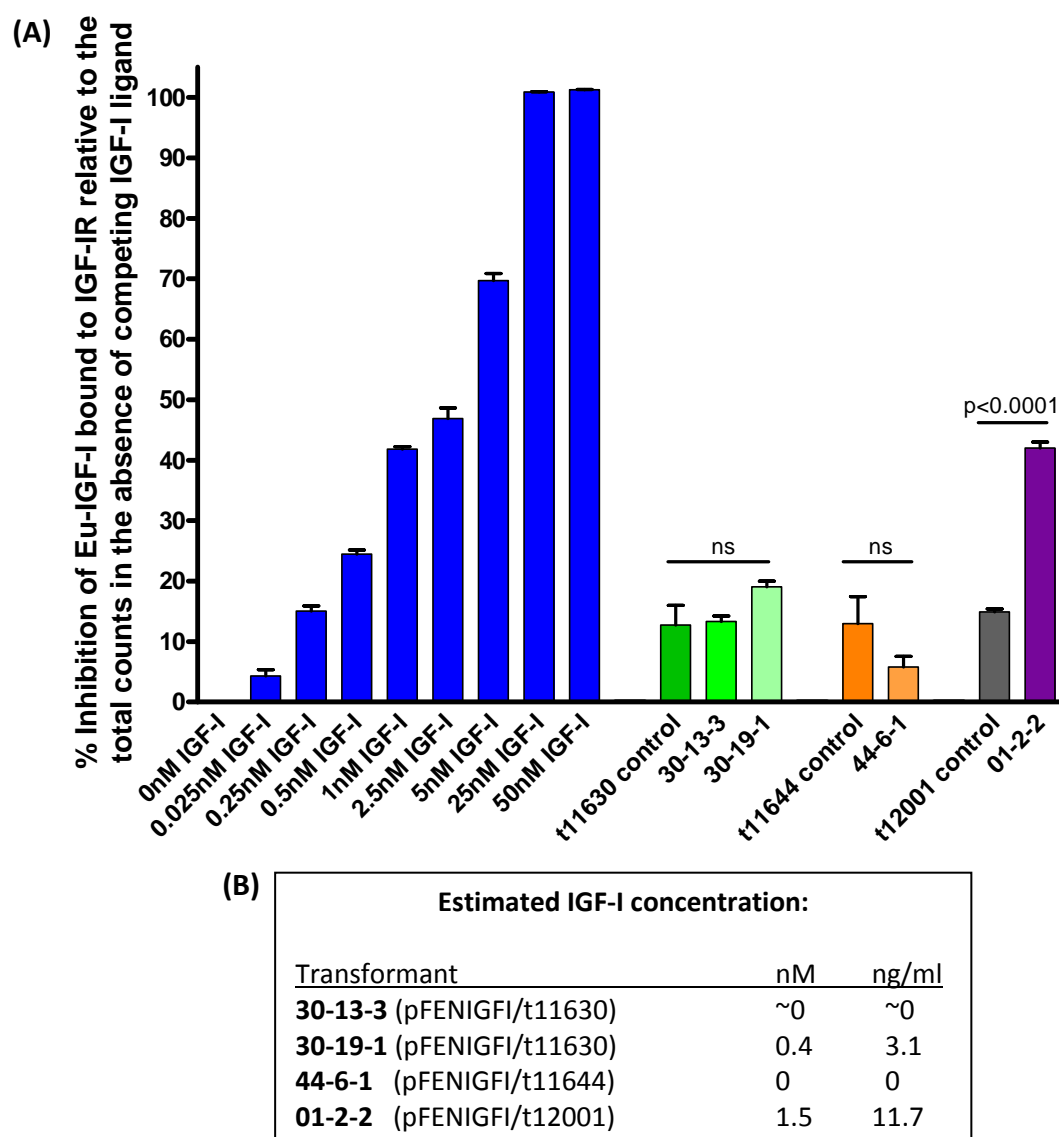
**Figure 4.5** Effects of various *Neurospora* growth media formulations on the competitive IGF-IR ligand binding assay.

A 96-well plate coated with anti-IGF-IR (24-31) antibody was used to capture human IGF-IRs from P6 cell lysates. Captured IGF-IRs were incubated overnight with a combination of Eu-IGF-I and different variants of *Neurospora* liquid media or replicate media samples spiked with 5nM IGF-I. An IGF-I standard curve in LBB was included as a positive control. The plate was washed and remaining Eu-IGF-I bound to the IGF-IR was measured by time-resolved fluorescence. The standard curve was fitted to a one-site competition binding model using Prism. The results are presented as percentage inhibition of Eu-IGF-I binding to IGF-IR in the absence of competing IGF-I in LBB.



**Figure 4.6 Comparison of IGF-I standard curves diluted in ligand binding buffer and FRIES + L-histidine medium in a competitive IGF-IR ligand binding assay.**

A 96-well plate coated with anti-IGF-IR (24-31) antibody was used to capture human IGF-IRs from P6 cell lysates. Two standard curves were created by diluting IGF-I in LBB or FmH. Captured IGF-IRs were incubated overnight with a combination of Eu-IGF-I and IGF-I standard curve dilutions. The plate was washed and the remaining Eu-IGF-I bound to the IGF-IR was measured by time-resolved fluorescence. The standard curves were fitted to a one-site competition binding model using Prism. The results are presented as a percentage relative to the maximum Eu-IGF-I binding to IGF-IR in the absence of competing IGF-I in ligand binding buffer.



**Figure 4.7 Competitive IGF-IR binding assay of secreted protein in culture media from ectopic pFENIGFI *Neurospora* transformants.**

**(A)** A 96-well plate coated with anti-IGF-IR (24-31) antibody was used to capture human IGF-IRs from P6 cell lysates. Captured IGF-IRs were incubated overnight with a combination of Eu-IGF-I and *Neurospora* culture supernatant. IGF-I protein was diluted in FmH to generate a standard curve. The plate was washed and remaining Eu-IGF-I bound to the IGF-IRs was measured by time-resolved fluorescence. The standard curve was based on a one-site competition binding model and the *Neurospora* sample measurements were fitted to the curve by non-linear regression using Prism. The results are presented as percentage inhibition of Eu-IGF-I binding to IGF-IR in the absence of competing IGF-I in LBB. The statistical significance of differences between transformants relative to untransformed controls was determined by ANOVA for the t11630 transformants and t-test for the t11644 and t12001 transformants. Note: ns = not significant ( $p > 0.05$ ). **(B)** The estimated IGF-I concentration in media of *Neurospora* transformants was corrected by subtraction of non-transformed control background and the estimate doubled to account for assay dilution factor.



**4.2.8 IGF-IR kinase phosphorylation assay of culture supernatant from ectopic pFENIGFI *Neurospora* transformants in FRIES medium**

The IGF-IR ligand binding assay results indicated that *Neurospora* transformants 30-19-1 and 01-2-2 produced IGF-I capable of binding to IGF-IR (4.2.7.2). An IGF-IR kinase assay was performed to determine if the secreted IGF-I was capable of activating the IGF-IR. Ligand binding to the IGF-IR induces conformational changes activating the intracellular kinase domain of the receptor which subsequently autophosphorylates IGF-IR tyrosine residues. As receptor autophosphorylation is the first step in the intracellular signalling cascade, phosphorylated IGF-IR levels can be used to determine ligand bioactivity (Chen *et al.*, 2003).

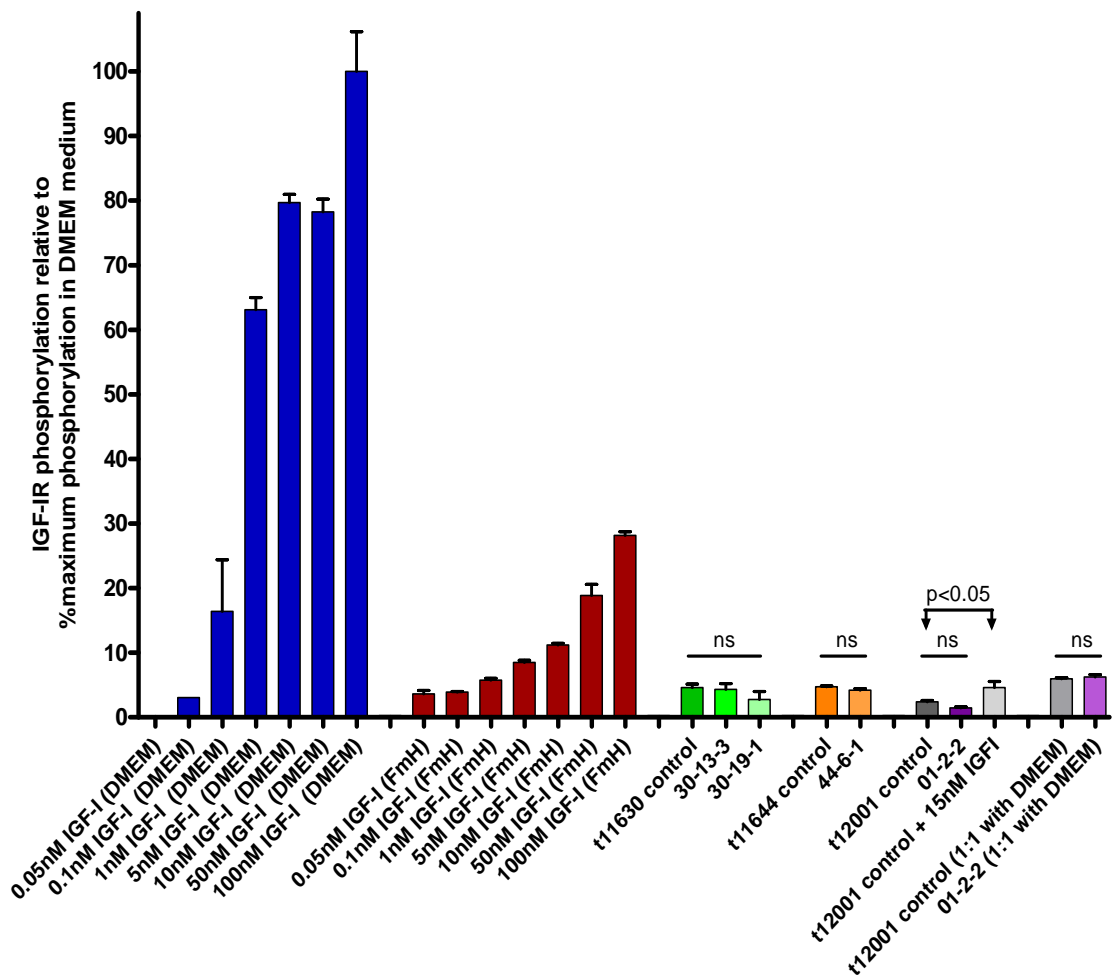
The IGF-IR kinase assay involved seeding P6 cells in a 96 well plate at a density of  $2.5 \times 10^4$  cells/well and incubating overnight. Cells were treated for 10min with 72hr culture supernatants from *Neurospora* pFENIGFI transformants 30-13-1, 30-19-1, 44-6-1 and 01-2-2 cultured in FmH and IGF-I standards prepared in FmH or DMEM medium. P6 cells were lysed and the lysates transferred to a 96-well plate coated with anti-IGF-IR antibody (24-31) (2.1.6). A Eu-labelled antiphosphotyrosine antibody (PY-20) (2.1.6) was added to each well 2 hours prior to the addition of Delphia enhancement solution and subsequent fluorescence measurement (2.2.6.3). Comparison of the IGF-I standard curves in FmH and DMEM medium (Figure 4.8) indicated the FmH interfered significantly with the IGF-IR kinase assay with maximum fluorescence decreased by ~75%. None of the transformants produced a signal greater than the untransformed controls and all assayed samples were located at or below the bottom of the IGF-I standard curve in FmH. Comparison of the measured value for the t12001 control spiked with 15nM of recombinant IGF-I with the IGF-I/FmH standard curve resulted in an estimate of <1nM which indicated the *Neurospora* strains were secreting one or more compounds that were inhibiting the IGF-IR phosphorylation assay. A 1:1 dilution of the t12001 and 01-2-2 samples with DMEM medium resulted in minor fluorescence increases but these were proportional in both the transformed and untransformed strains which suggested dilution of the fungal culture supernatants was unlikely to resolve the problem. As it was not clear which components of the FmH might have

been affecting the IGF-IR phosphorylation assay, optimisation of the FmH formulation or identification of an alternative medium suited for both the IGF-IR ligand and phosphorylation assays would have required a significant investment of time and resources with no guarantee of success. Given the IGF-IR ligand binding assay was functional with FmH medium, it was opted not to pursue optimisation of *Neurospora* supernatants with the IGF-IR phosphorylation assay.

#### **4.2.9 Optimisation of culturing conditions for *Neurospora* IGF-I secretion**

Initial characterisation of *Neurospora* pFENIGFI transformants by western blot (4.2.6) and IGF-IR ligand binding assay (4.2.7) was performed with liquid cultures incubated for 72 hours. However, the 72-hour incubation period was arbitrarily chosen and may not have been optimal for IGF-I secretion in *Neurospora*. Consequently, alternative incubation conditions were trialled to identify the most appropriate for IGF-I secretion in *Neurospora*.

Cultures of transformant 01-2-2 were used in this experiment as 01-2-2 secreted the highest level of IGF-I as determined by IGF-IR ligand binding assay (Figure 4.7). Liquid cultures of 01-2-2 and the t12001 untransformed control strain were grown in FmH under three different regimes. Firstly, liquid cultures were grown for 48, 72 or 96 hours (subsequently referred to as 2, 3 and 4 days respectively) prior to harvesting of culture media to determine the optimal duration of incubation for IGF-I secretion. Secondly, to increase fungal biomass and remove accumulated extracellular proteases, culture media was replaced after 2, 3 or 4 days of culture growth and incubated for a further 2 days prior to harvesting. Thirdly, cultures were grown for 2, 3 or 4 days before the media was replaced and the cultures were incubated for a further 2 days after which 10ml of media was added to ensure an adequate carbon source was available for the extended fermentation. These cultures were incubated for an additional 2 days before harvesting.



**Figure 4.8** IGF-IR phosphorylation assay of ectopic pFENIGFI *Neurospora* transformants.

P6 cells were grown overnight in a 96-well plate, washed and serum starved for ~4 hours. The following were added to the plates and incubated for 10 minutes at 37°C: an IGF-I dilution series in DMEM or FmH media; *Neurospora* ectopic pFENIGFI transformants and untransformed controls, t12001 & 01-2-2 diluted 1:1 with DMEM and a 15nM IGF-I spiked t12001 control. Cells were lysed and the lysate transferred to a 96-well plate coated with anti-IGF-IR (24-31) antibody for IGF-IR immunocapture. Eu-PY20 was added to the plate to bind phosphorylated tyrosine residues in the activated IGF-IR and measured by time-resolved fluorescence. The statistical significance of differences between transformants relative to untransformed controls was determined by ANOVA with Tukey’s multiple comparison test for the t11630 and t12001-derived transformants and by t-test for the t11644-derived transformant. Note: ns = not significant ( $p>0.05$ ).

*Neurospora* protein secretion increases when sorbose is present in the culture medium (E. Cambareri: *Personal communication*). Sorbose affects the structural composition of the *Neurospora* cell walls resulting in dense mycelia with greatly increased branching and thus a larger number of hyphal tips (De Terra & Tatum 1961). These morphological changes are likely to be the mechanism by which sorbose enhances extracellular protein levels as protein secretion in filamentous fungi occurs predominately at actively growing hyphal tips and is improved by increasing actively growing surface area (Wosten *et al.*, 1991); (Lee *et al.*, 1998). To determine the effect of sorbose on IGF-I yield, FmH media with and without sorbose were trialled for each of the three incubation conditions.

The secreted IGF-I was analysed by both western blot and the IGF-IR ligand binding assay. The *Neurospora* cultures were grown solely in FRIES medium because Vogel's based media had been shown to inhibit the IGF-IR ligand binding assay (refer 4.2.7.1)

Liquid cultures consisted of 25ml of FmH or FmH with 2.5% (w/v) sorbose (FmHS) in 125ml Erlenmeyer flasks and were inoculated with  $2 \times 10^5$  of 6-10 day-old 01-2-2 or t12001 macroconidia. Cultures were incubated at 25°C on an orbital shaker in the absence of light and maintained under the three culturing regimes outlined previously.

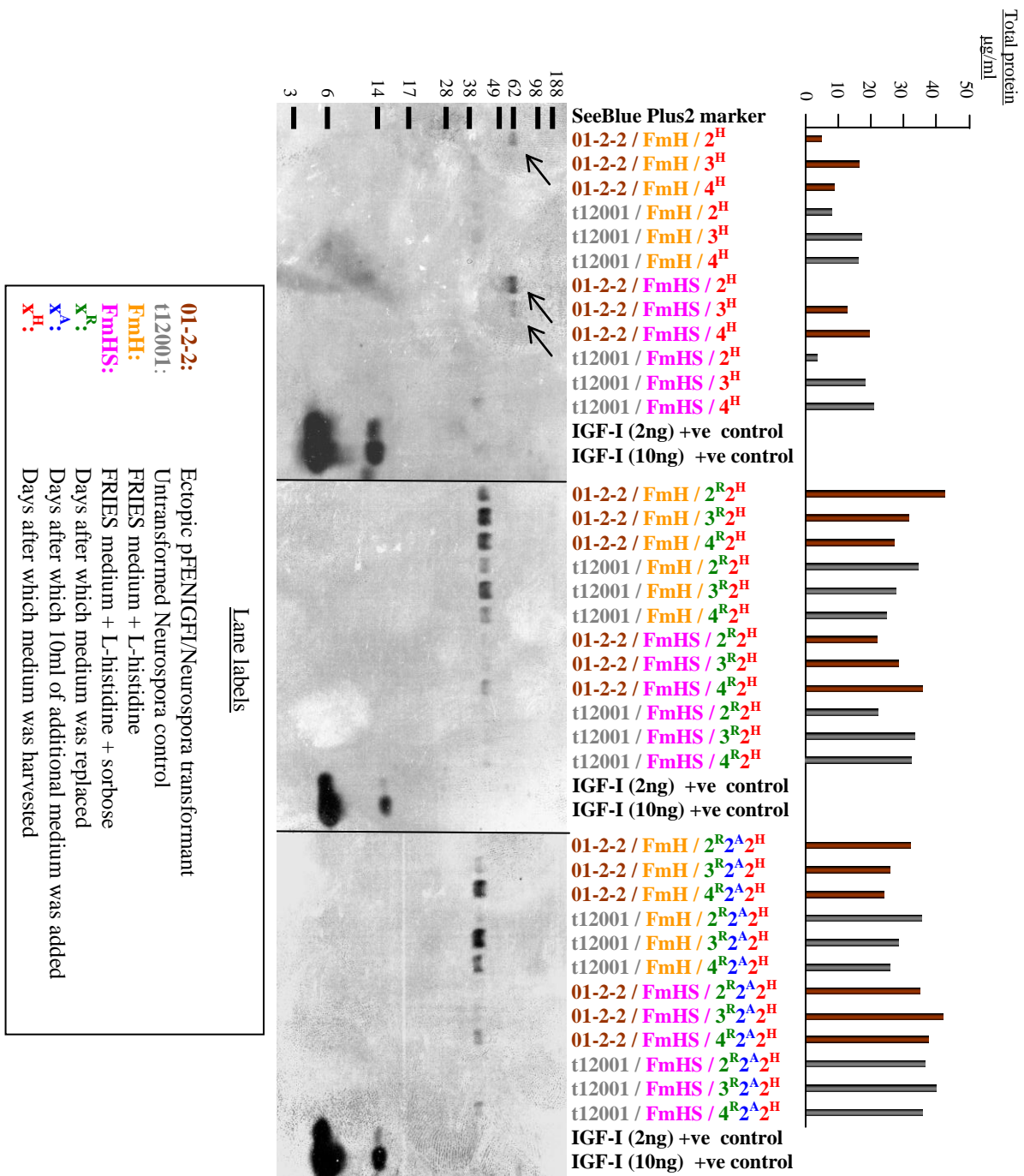
#### **4.2.9.1 Optimisation of Neurospora culture conditions by western blot**

Supernatants from *Neurospora* liquid cultures (4.2.9) were harvested (2.2.5.1.1), the total protein in the supernatants were determined by Bradford assay (2.2.5.3.1) and analysed by western blot (2.2.5.6) using an anti-human IGF-I polyclonal antibody (2.1.6) (Figure 4.9). The samples were not standardised for total protein prior to loading on the SDS-PAGE gels used for western blotting as it was desirable to load as much protein as possible from each sample to determine if mature protein was being processed and secreted.

Although mature protein was not detected in any of the culture supernatants, the putative gla-1:IGF-I fusion protein was detected in the medium of the 2 & 3 day FmHS cultures and the 2 day FmH culture which suggested shorter incubation periods were better for secretion of the fusion protein and possibly for mature IGF-I. The putative fusion protein was not detected in previous western blots of supernatants from cultures grown in VmH and NSmH (Figure 4.4) although this may have been due to the greater sensitivity of the polyclonal anti-IGF-I antibody which readily detected the 2ng IGF-I positive control whereas previous westerns with a monoclonal anti-IGF-I antibody barely detected 15ng of the IGF-I control. Alternatively, the FRIES based media may have decreased *Neurospora* protein secretion and thereby reduced the level of extracellular proteases. It was unlikely the FRIES medium had increased secretion of gla-1:IGF-I fusion protein as the total protein yield in 72 hour FmH/FmHS and VmH cultures ranged from 5 to 20µg/ml and 10 to 40µg/ml, respectively. As the relative intensity of the putative fusion protein band was substantially less than the 2ng IGF-I positive control band and 11.7µl of sample was loaded into each well, the concentration of uncleaved fusion protein detected in the culture medium must have been less than 170ng/ml.

#### **4.2.9.2 Optimisation of *Neurospora* culture conditions for the IGF-IR ligand binding assay**

Supernatants from the *Neurospora* liquid cultures (4.2.9) were harvested (2.2.5.1.1) and analysed by IGF-IR ligand binding assay (Figure 4.10) (2.2.6.2). The t12001 untransformed control in the FmHS and FmH media produced varying degrees of background fluorescence which typically increased with culture age. Generally, the FmHS control cultures consistently produced less background than the FmH control cultures. The FmHS culture harvested after 2 days incubation yielded the highest amount of IGF-I (22ng/ml) although it was unknown if this measurement reflected the IGF-IR binding with mature IGF-I, uncleaved gla-1:IGF-I fusion protein or both. The t12001/2<sup>H</sup> controls spiked with 5mM of IGF-I in FmH and FmHS media resulted in measurements of 3.1nM and 3.3nM, respectively. The discrepancy between expected and observed results for the spiked IGF-I control was postulated to have resulted from proteolysis in the culture medium and required further investigation.



**Figure 4.9 Western blot optimisation of *Neurospora* incubation conditions for IGF-I secretion.**

Ectopic pFENIGFI *Neurospora* transformant 01-2-2 and untransformed t12001 control strain were grown in FmH or FmHS media for 2, 3 or 4 days before media was replaced, added or harvested. Secreted protein in culture media was analysed by western blot using a polyclonal anti-human IGFI primary antibody. Note: putative uncleaved gla-1:IGF-I fusion protein bands are indicated by arrows. The total protein concentrations in each of the culture supernatants were determined by Bradford assay and are shown in the bar graph above the relevant lanes.

**4.2.10 Characterisation of IGF-I degradation in *Neurospora* culture supernatants in the presence of protease inhibitors.**

*Neurospora crassa* is a saprotroph and naturally secretes a range of proteases into the extracellular environment (Heiniger & Matile 1974). IGF-I is known to be susceptible to proteolytic degradation in human serum, however, serine protease inhibitors such as aprotinin and phenylmethylsulfonyl fluoride (PMSF) have been shown to significantly reduce IGF-I proteolysis (Khosravi *et al.*, 2005). The Protease Inhibitor Cocktail (PIC) from Sigma-Aldrich (cat# P8340) facilitates broad ranging inhibition of serine, cysteine, aspartic proteases and aminopeptidases and is widely used in IGF cell-based assays. As such, the PIC was added to the *Neurospora* culture supernatants upon harvesting at a 1:1,000 dilution in an attempt to inhibit IGF-I degradation. However, the estimated ~3nM concentration in controls spiked with 5nM IGF-I in the IGF-IR ligand binding assay (Fig 4.10) suggested IGF-I was still susceptible to degradation by proteases present in the culture medium. Thus, it was desirable to determine if higher concentrations of PIC would inhibit IGF-I degradation and if the PIC would affect *Neurospora* viability if added to the culture medium during the early stages of culturing.

To determine if culture medium containing PIC would impair *Neurospora* viability, various dilutions of PIC (1:200, 1:500, 1:1,000 and neat) were added to FmHS culture medium which was subsequently inoculated with the ectopic pFENIGFI transformant 01-2-2 or the untransformed t12001 control strain and cultured for 48 hours. The *Neurospora* growth rate and biomass was reasonably consistent between all dilutions of the PIC and no PIC control cultures indicating the PIC had no major inhibitory effect on growth. The culture supernatants were harvested after 48 hours and analysed by western blot (Figure 4.11A) using a standardised quantity of total protein from each treatment group to determine if the addition of protease inhibitors to the culture medium prevented proteolytic degradation of secreted mature IGF-I. Although no mature IGF-I protein was detected with the western blot, the density of the putative gla-1:IGF-I fusion band in the 01-2-2 transformant increased in the presence of PIC. This result suggested the addition of protease inhibitors reduced proteolytic



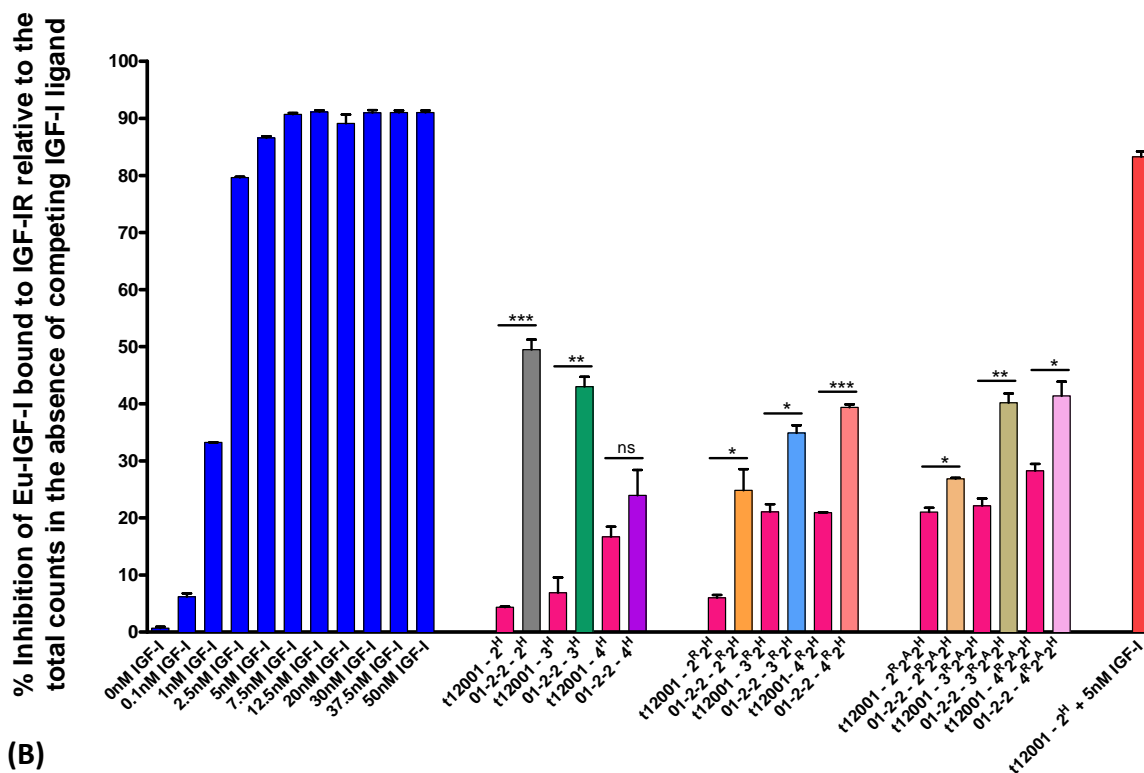
degradation in the culture supernatant. However, it was not possible to determine if the increased band density reflected reduced degradation of IGF-I, gla-1 or both components of the fusion protein.

An IGF-I spiking experiment was performed to determine if the PIC was effective in preventing proteolytic degradation of IGF-I in *Neurospora* cultures. Three hundred and twenty nanograms (41.8nM) of recombinant IGF-I was spiked into liquid cultures of the untransformed t12001 strain grown in FmHS medium with and without PIC (1:200) at 48hrs, 24hrs and 30 minutes prior to harvesting. Spiked FmHS controls were also included to ensure IGF-I was stable in the *Neurospora* culture medium. The *Neurospora* cultures were harvested after 48 hours and the supernatant analysed by western blot (Figure 4.11B). The blot results indicated the spiked recombinant IGF-I was stable in FmHS yet the absence of an IGF-I band in all the *Neurospora* cultures supported the notion that *Neurospora* secreted proteases were degrading IGF-I in the culture supernatant even in the presence of PIC.

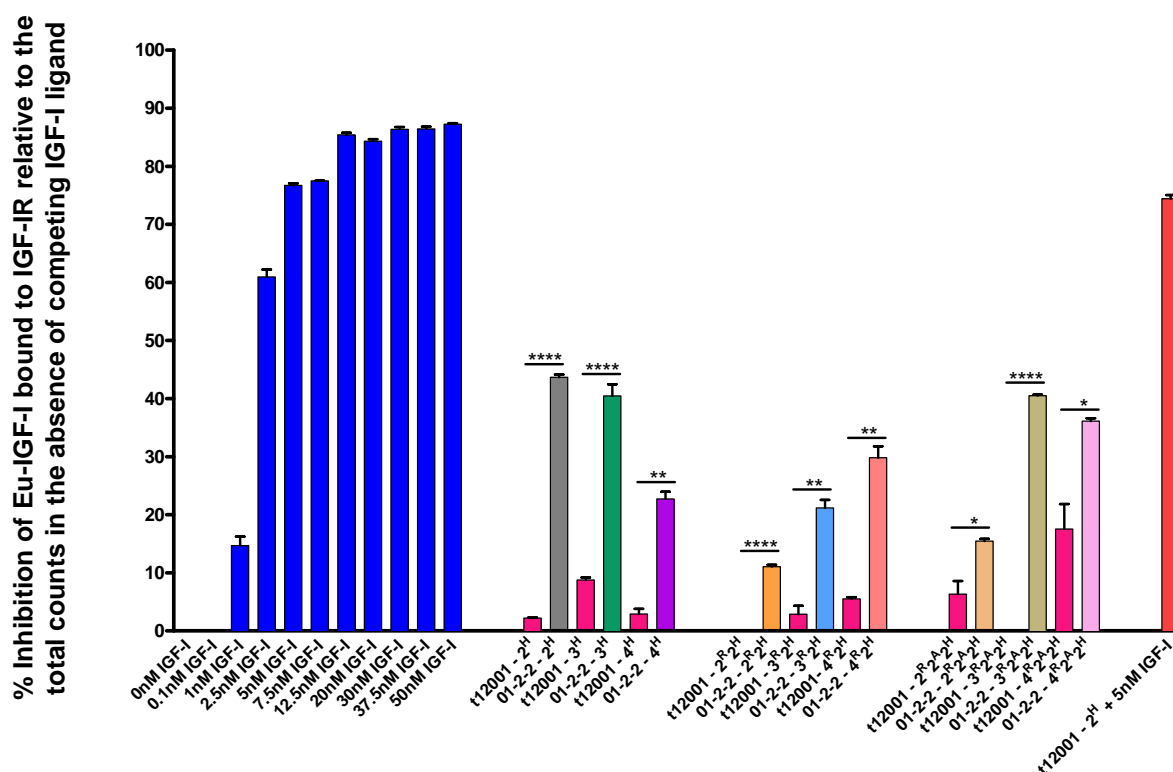
Proteolytic degradation of IGF-I in *Neurospora* cultures presented a significant hurdle for screening *Neurospora* supernatants as the ligand binding assay could be confounded by fluctuations in the relatively small yield of IGF-I. Ideally, the engineering of protease-deficient *Neurospora* strains would have been undertaken to reduce the impact of proteolytic degradation. However, this work was not pursued in the course of this project as the development of improved *Neurospora* expression strains was already being undertaken at Neugenesis Corporation.

The presence of uncleaved gla-1:IGF-I fusion protein in the culture medium introduced uncertainty as to which molecules were binding to the IGF-IR in the ligand binding assays. As such, alternative expression vectors were used in an attempt to increase the yield of IGF-I secretion in *Neurospora* to a level where the effects of proteolytic degradation and uncleaved fusion protein would be relatively small.

(A)



(B)



**01-2-2:** Ectopic pFENIGFI/Neurospora transformant  
**t12001:** Untransformed Neurospora control  
**x<sup>R</sup>:** Days after which medium was replaced  
**x<sup>A</sup>:** Days after which 10ml of medium was added  
**x<sup>H</sup>:** Days after which medium was harvested

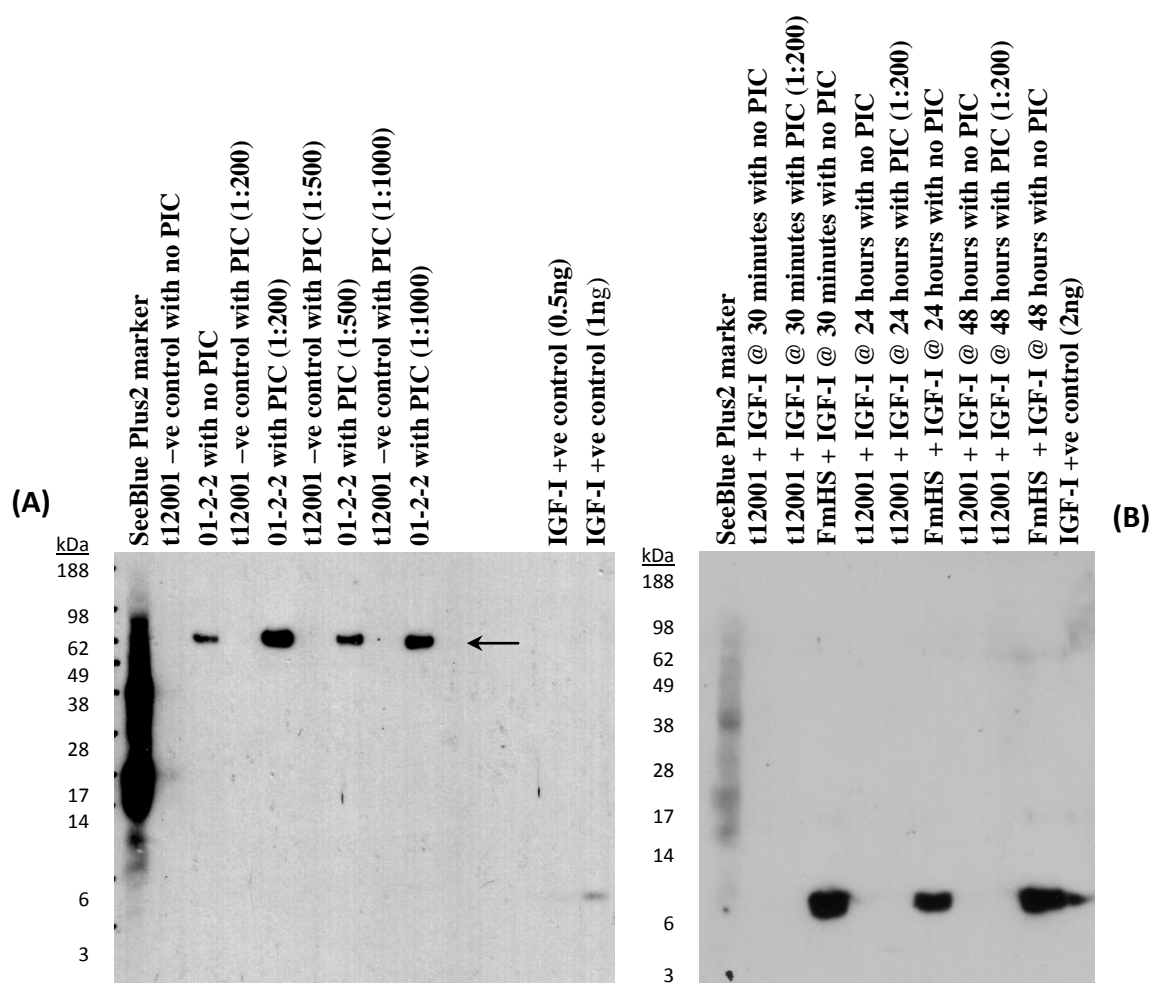
(C)

01-2-2 sample	FmH		FmHS	
	nM	ng/ml	nM	ng/ml
2 <sup>H</sup>	2.0	15.6	2.8	22.0
3 <sup>H</sup>	1.4	11.0	2.2	17.2
4 <sup>H</sup>	0.2	1.6	1.0	7.8
2 <sup>R</sup> 2 <sup>H</sup>	0.6	4.8	0.6	4.8
3 <sup>R</sup> 2 <sup>H</sup>	0.6	4.8	0.8	6.2
4 <sup>R</sup> 2 <sup>H</sup>	0.8	6.2	1.4	11.0
2 <sup>R</sup> 2 <sup>A</sup> 2 <sup>H</sup>	0.2	1.6	0.4	3.2
3 <sup>R</sup> 2 <sup>A</sup> 2 <sup>H</sup>	0.8	6.2	2.6	20.4
4 <sup>R</sup> 2 <sup>A</sup> 2 <sup>H</sup>	0.6	4.8	1.6	12.6
t12001 / 2 <sup>H</sup> + 5nM IGF-I	3.1	23.3	3.3	25.3

**Figure 4.10** IGF-IR ligand binding assay optimisation of *Neurospora* incubation conditions for IGF-I secretion.

Ectopic pFENIGFI *Neurospora* transformant 01-2-2 and untransformed t12001 control strain were grown in (A) FmH or (B) FmHS for 2, 3 or 4 days before media was replaced, added or harvested. Secreted protein in culture media was analysed by a competitive IGF-IR ligand binding assay. A 96-well plate coated with anti-IGF-IR (24-31) antibody was used to capture human IGF-IRs from P6 cell lysates. IGF-IRs were incubated overnight with a combination of Eu-IGF-I and *Neurospora* media samples. The plate was washed and remaining Eu-IGF-I bound to the IGF-IRs was measured by time-resolved fluorescence. The standard curve was based on a one-site competition binding model and the *Neurospora* sample measurements were fitted to the curve by non-linear regression using Prism. The results are presented as a percentage inhibition of Eu-IGF-I binding to IGF-IR in the absence of competing IGF-I in LBB. The statistical significance of differences between transformants relative to untransformed controls was determined by t-test. Note: \* = p<0.05; \*\* = p<0.01; \*\*\* = p<0.001; \*\*\*\* = p<0.0001 & ns = not significant (p>0.05).

(C) The secreted IGF-I concentrations were corrected by subtraction of background measurements from non-transformed strain controls and the estimate doubled to account for the inherent assay dilution factor with the exception of the t12001 controls spiked with 5nM of recombinant IGF-I protein as dilution factor had been accounted for during assay preparation.



**Figure 4.11** Western blot characterisation of IGF-I stability in *Neurospora* cultures in the presence of a protease inhibitor cocktail (PIC).

**(A)** Effects of protease inhibitors during incubation of *Neurospora* cultures: ectopic pFENIGFI *Neurospora* transformant 01-2-2 and untransformed t12001 control strain were cultured in FmHS medium with 1:200, 1:500 & 1:1000 dilutions of PIC for 48 hours. The ability of the PIC to potentially reduce IGF-I and/or gla-1 degradation was assessed by the relative intensity of the putative uncleaved gla-1:IGF-I fusion protein band (indicated by an arrow).

**(B)** IGF-I spiking in *Neurospora* cultures: Untransformed t12001 strain was cultured in FmHS with and without PIC (1:200) for 48 hours. FmHS with no PIC was also included as a control. Cultures were spiked with 41.8nM (320ng/ml) of IGF-I at 48 hours, 24 hours and 30 minutes before harvesting.

The media in both experiments were analysed by western blot using a polyclonal anti-human IGF-I primary antibody.

### 4.3 Characterisation of transplaced IGF-I expression vectors

An additional four IGF-I expression vectors were constructed for the purposes of increasing IGF-I yield in *Neurospora* cultures. The rationale for each vector is described in 3.1.3.4 and the key features of the constructs are compared in Table 4.1. All five IGF-I expression constructs were cloned into pDV2 to form pFENIGFI-DV2, pFENIGFI-GS-DV2, pNEIGFI-G1-DV2, pNEhIGFI-C1-DV2 and pNEhIGFI-H1-DV2. Subcloning into pDV2 facilitated targeted transplacement between *his-3* and *cog*<sup>+</sup> in the *Neurospora* genome. Transplacement of the IGF-I expression constructs between the *his-3* and *cog*<sup>+</sup> loci was done for two reasons. Firstly, this region of the *Neurospora* genome is known to be transcriptionally active (Catcheside *et al.*, 2003) and secondly, integration of the different constructs into a single locus eliminated any potential chromosomal position effects on IGF-I expression and thereby facilitated a direct comparison between the IGF-I yields associated with the different vectors.

Construct name	Primary features of construct	Design rationale chapter reference
<b>pFENIGFI-DV2</b>	Fusion of <i>gla-1</i> peptide to nIGF-I	3.1.3.2 & 3.1.3.3
<b>pFENIGFI-GS-DV2</b>	Fusion of <i>gla-1</i> peptide to nIGF-I incorporating a glycine spacer to facilitate Kex2 cleavage of the fusion protein	3.1.3.4.1
<b>pNEIGFI-G1-DV2</b>	nIGF-I with the <i>gla-1</i> signal peptide	3.1.3.4.2
<b>pNEhIGFI-C1-DV2</b>	hIGF-I with the <i>Cel12A</i> signal peptide	3.1.3.4.3
<b>pNEhIGFI-H1-DV2</b>	hIGF-I with the human signal peptide	3.1.3.4.4

**Table 4.1 Key features of the five IGF-I expression vectors used in this study.**

All IGF-I expression vectors were placed into the pDV2 vector backbone to facilitate targeted transplacement between the *his-3* and *cog*<sup>+</sup> loci in Linkage Group I of *Neurospora crassa*.

### **4.3.1 Targeted transplacement of the IGF-I expression constructs into *Neurospora***

The five constructs were transformed into the t11630 and t11644 diversification strains as the genotypes of these strains were amenable to the project requirements. Specifically, both strains contained the *his-3* (K458) allele suitable for selection of vectors transplaced 3' of the *his-3* locus using the pDV2 targeting vector, opposing mating types to facilitate sexual crossing and the *rec-2* and *cog*<sup>+</sup> alleles required for HSMR.

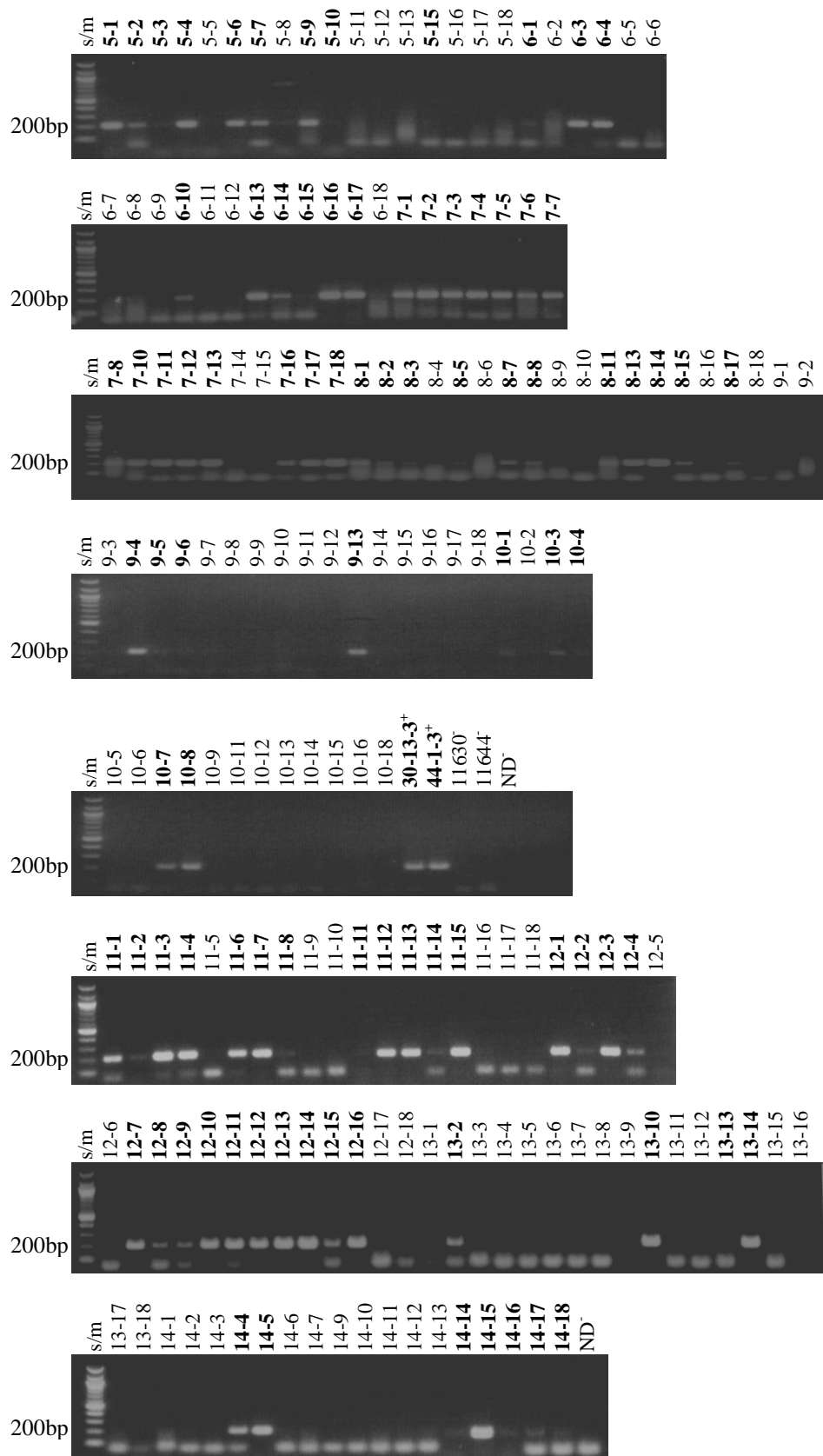
#### **4.3.1.1 Electroporation of the IGF-I expression constructs into *Neurospora***

Approximately 5µg of DNA from each vector was linearised by restriction digestion with *Ssp*I and electroporated into 1x10<sup>8</sup> macroconidia from *Neurospora* strains t11630 & t11644 and plated onto minimal medium (2.2.1.3.1). Eighteen well-isolated primary transformants were selected for each of the ten vector/strain combinations and picked onto agar slopes with minimal medium to produce macroconidiating cultures. Genomic DNA was obtained from macroconidia (2.2.1.3.2) and vector integration in the putative transformants screened for by PCR (2.2.4.6.4). Putative transformants with pFENIGFI-DV2, pFENIGFI-GS-DV2 and pNEIGFI-G1-DV2 were screened using primers IGFICheck-F and IGFICheck-R and the presence of pNEhIGFI-C1-DV2 and pNEhIGFI-H1-DV2 were screened for using primers hIGFICheck-F & hIGFICheck-R (2.1.8.3). Both primer sets were located within the *IGF-I* coding sequences and positive transformants produced amplicons of ~200bp when the PCR reactions were resolved by agarose gel electrophoresis (Figure 4.12).

Multiple positive primary transformants were obtained for each vector/strain combination and at least three of these were subjected to microconidiation to obtain homokaryons (2.2.1.2.4). Complementation testing indicated a number of transformed homokaryons had been obtained for each of the vector/strain combinations. Confirmation of correct transplacement in these homokaryon isolates required Southern blot analysis.

**4.3.1.2 Southern blot analysis of the transplaced IGF-I expression constructs in Neurospora**

Southern blot analysis of genomic DNA from selected transformed homokaryons identified in the genetic complementation screen (4.3.1.1) was performed to ensure the constructs were correctly transplaced between *his-3* and *cog<sup>+</sup>* in the *Neurospora* genome (2.2.4.7). The DNA sequence contig containing the *his-3* and *cog<sup>+</sup>* loci was obtained from the *Neurospora* genome database (<http://www.broadinstitute.org/annotation/genome/neurospora/MultiHome.html>). The *Clal* restriction enzyme was found to be suitable for the Southern blot analysis as transformants with targeted integration of the IGF-I vectors were predicted to eliminate a *Clal* site and alter the restriction pattern (Figure 4.13). Using the pDV2 vector backbone as a probe, the untransformed strains were predicted to produce ~3, ~2.1 & ~3.4kb bands upon cleavage with *Clal* whereas transplaced pFENIGFI-DV2 and pFENIGFI-GS-DV2 vectors were expected to result in ~6.8 & ~3.4kb bands and the pNEIGFI-G1-DV2, pNEhIGFI-C1-DV2 and pNEhIGFI-H1-DV2 constructs should have produced ~5.2 and ~3.4kb bands. However, Southern blot analysis of the putative transformants did not produce the expected banding patterns in either of the transformed or untransformed strains (Figure 4.14). The untransformed strain contained ~3, ~2.1 & ~6.8kb bands. The presence of a 6.8kb band instead of the expected 3.4kb band suggested the *Clal* site 3' of the *cog<sup>+</sup>* sequence was not present (Figure 4.13). The absence of the *Clal* site was probably a sequence polymorphism between strains or there may have been a sequencing error in the genome database. Alternatively, the *Clal* restriction site may have been methylated which could have prevented the methylation sensitive *Clal* enzyme from cleaving.



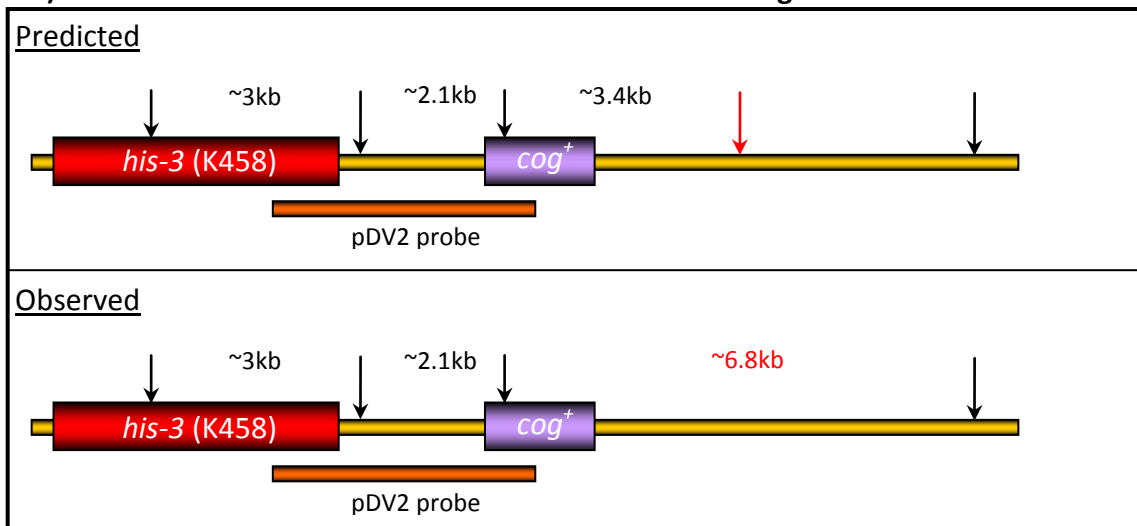


<u>Lane labels</u>	
5-x	= pFENIGFI-DV2 in t11630
6-x	= pFENIGFI-DV2 in t11644
7-x	= pFENIGFI-GS-DV2 in t11630
8-x	= pFENIGFI-GS-DV2 in t11644
9-x	= pNEIGFI-G1-DV2 in t11630
10-x	= pNEIGFI-G1-DV2 in t11644
11-x	= pNEhIGFI-C1-DV2 in t11630
12-x	= pNEhIGFI-C1-DV2 in t11644
13-x	= pNEhIGFI-H1-DV2 in t11630
14-x	= pNEhIGFI-H1-DV2 in t11644
30-13-3 <sup>+</sup>	= pFENIGFI in t11630 +ve control
44-1-3 <sup>+</sup>	= pFENIGFI in t11644 +ve control
t11630 <sup>-</sup>	= untransformed –ve control
t11644 <sup>-</sup>	= untransformed –ve control
ND <sup>-</sup>	= no DNA –ve control

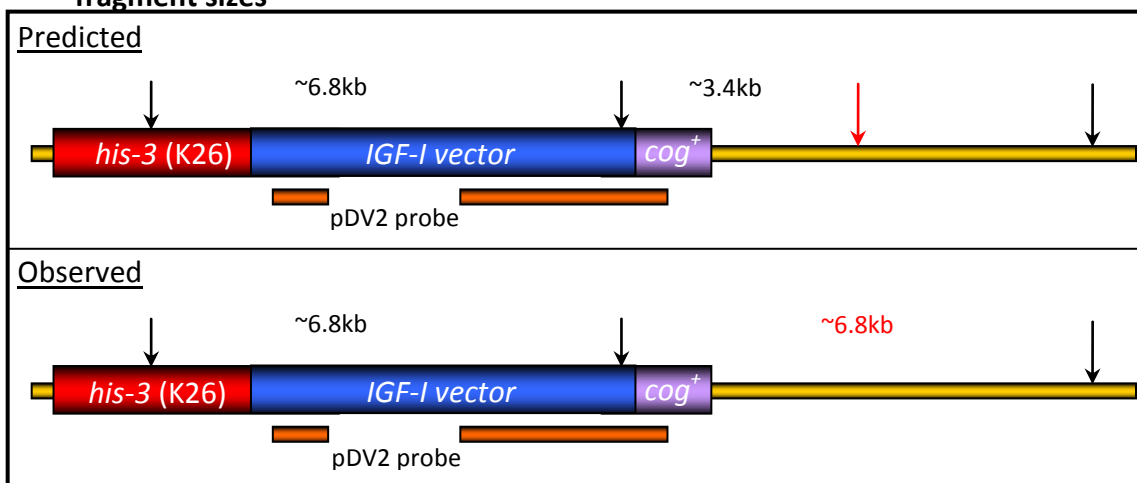
**Figure 4.12 PCR screen for pFENIGFI-DV2, pFENIGFI-GS-DV2, pNEIGFI-G1-DV2, pNEhIGFI-C1 & pNEhIGFI-H1-DV2 integration in *Neurospora* t11630 and t11644 transformants.**

Genomic DNA extracted from primary *Neurospora* transformants were PCR screened with primers specific for *nIGF-I* or *hIGF-I*. Positive candidates contained an ~200bp PCR amplicon (lane labels bolded). Please note: in some instances image digitization had failed to reproduce faint bands that were evident in the original image.

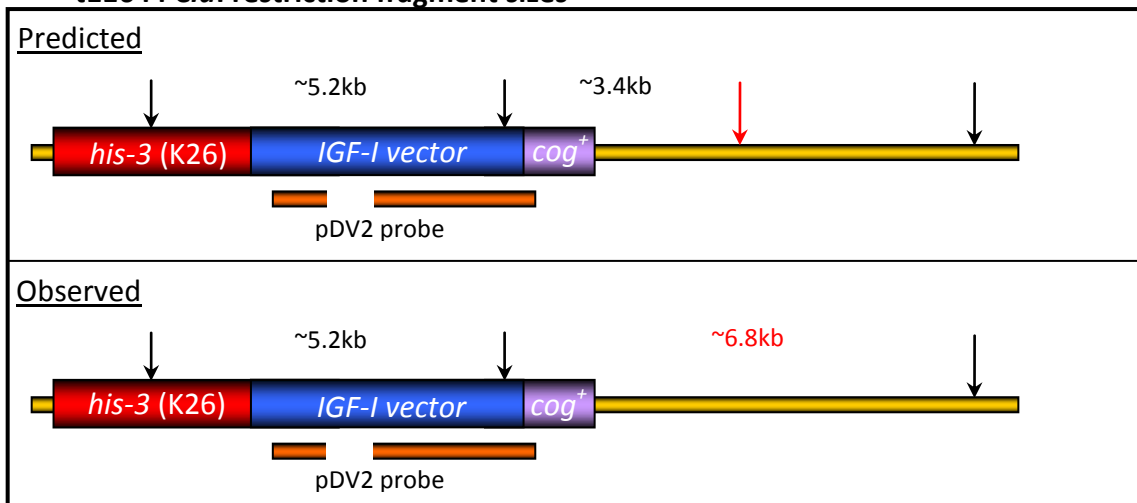
**A) Untransformed t11630 and t11644 *Clal* restriction fragment sizes**



**B) pFENIGFI-DV2 & pFENIGFI-GS-DV2 transformed t11630 & t11644 *Clal* restriction fragment sizes**



**C) pNEIGFI-G1-DV2, pNEhIGFI-C1-DV2 & pNEhIGFI-H1-DV2 transformed t11630 & t11644 *Clal* restriction fragment sizes**

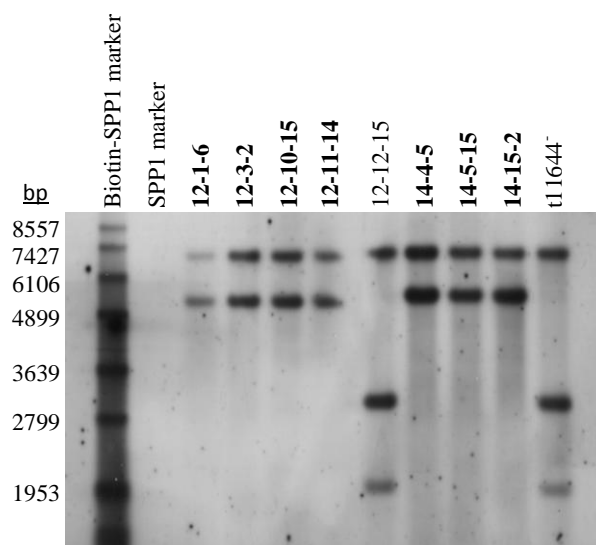
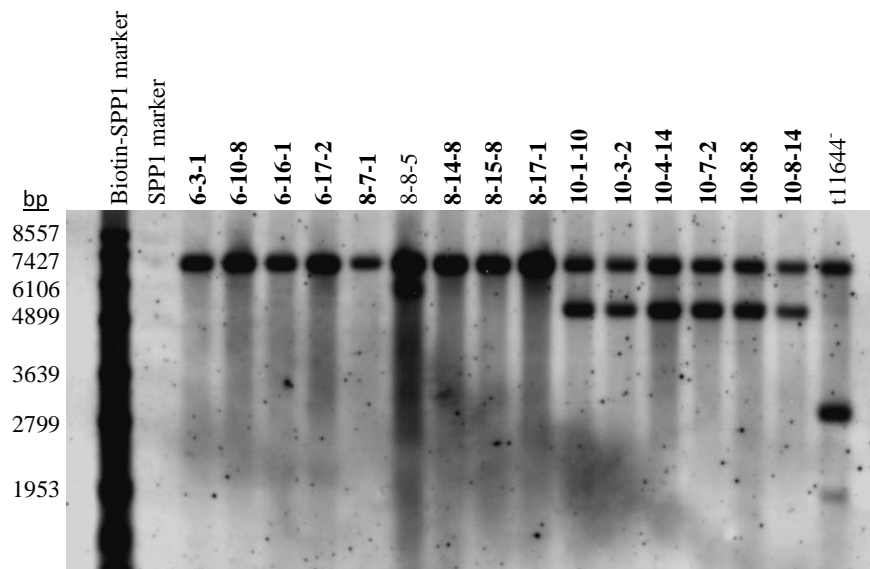
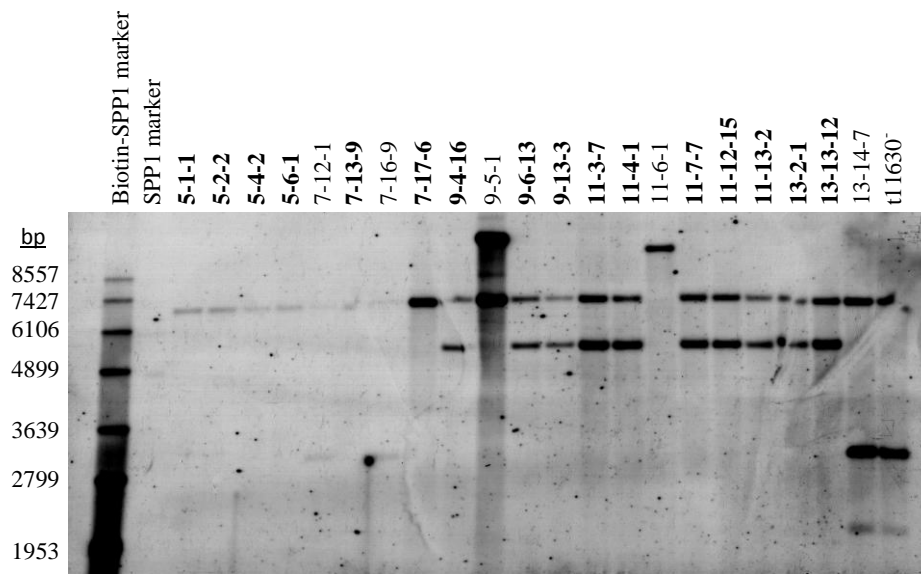


**Figure 4.13 Southern blot strategy for confirmation of pFENIGFI-DV2, pFENIGFI-GS-DV2, pNEIGFI-G1-DV2, pNEhIGFI-C1 & pNEhIGFI-H1-DV2 transplacement between the *his-3* & *cog*<sup>+</sup> region in *Neurospora* strains t11630 & t11644.**

Restriction digestion of *Neurospora* genomic DNA with *Clal* at the *his-3* to *cog*<sup>+</sup> region in untransformed and transplanted *Neurospora* strains probed with the pDV2 vector was predicted to result in specific Southern blot banding patterns that differentiated between **(A)** untransformed t11630 & t11644 strains, **(B)** pFENIGFI-DV2 & pFENIGFI-GS-DV2 transplacements and **(C)** pNEIGFI-G1-DV2, pNEhIGFI-C1-DV2 and pNEhIGFI-H1-DV2 transplacements.

The observed Southern blot band sizes did not match the expected band sizes but could be best explained if the *Clal* restriction site 3' of the *cog*<sup>+</sup> site was non-existent or uncleaved as shown by the maps in the 'observed' boxes. The *Clal* sites are indicated by black arrows and the erroneous or uncleaved *Clal* site is shown by a red arrow. Probed restriction fragment sizes are indicated between the arrows. The observed restriction fragments contradictory to the predicted fragment sizes are shown in red.

*CHAPTER 4 – Transformation, expression and diversification of IGF-I in Neurospora*



<u>Transformant labels</u>	
5-x-x	= pFENIGFI-DV2 in t11630
6-x-x	= pFENIGFI-DV2 in t11644
7-x-x	= pFENIGFI-GS-DV2 in t11630
8-x-x	= pFENIGFI-GS-DV2 in t11644
9-x-x	= pNEIGFI-G1-DV2 in t11630
10-x-x	= pNEIGFI-G1-DV2 in t11644
11-x-x	= pNEhIGFI-C1-DV2 in t11630
12-x-x	= pNEhIGFI-C1-DV2 in t11644
13-x-x	= pNEhIGFI-H1-DV2 in t11630
14-x-x	= pNEhIGFI-H1-DV2 in t11644
t11630	= untransformed –ve control
t11644	= untransformed –ve control

<u>Expected band sizes</u>	
Untransformed control:	2.1, 3 & 6.8kb
pFENIGFI-DV2:	6.8kb
pFENIGFI-GS-DV2:	6.8kb
pNEIGFI-G1-DV2:	5.2 & 6.8kb
pNEhIGFI-C1-DV2:	5.2 & 6.8kb
pNEhIGFI-H1-DV2:	5.2 & 6.8kb

**Figure 4.14** Southern blot of putative *Neurospora* transformants to confirm vector transplacement between *his-3* & *cog*<sup>+</sup> loci in Linkage Group I.

Genomic DNA from homokaryotic putative transformants were restricted with *Clal* and probed with a biotin-labelled pDV2 vector. Correct transplacements are indicated by bolded labels.

The presence of the ~5.2kb band and loss of ~3kb and ~2.1kb bands in the pNEIGF-G1-DV2, pNEhIGFI-C1-DV2 and pNEhIGFI-H1-DV2 transformants indicated correct transplacement had occurred. However, the observed 6.8kb band common to all untransformed and transplaced strains approximated the size of another restriction fragment in the pFENIGFI-DV2 & pFENIGFI-GS-DV2 *Neurospora* transformants. As such, correct transplacement was identified in these transformants by the absence of the 3 & 2.1kb bands (Figure 4.13). The Southern blot analysis identified transformants with the five IGF-I expression vectors correctly transplaced between *his-3* and *cog*<sup>+</sup> in the t11630 and t11644 strains enabling direct comparison of the constructs.

#### **4.3.2 Western blot characterisation of the IGF-I expression constructs in *Neurospora***

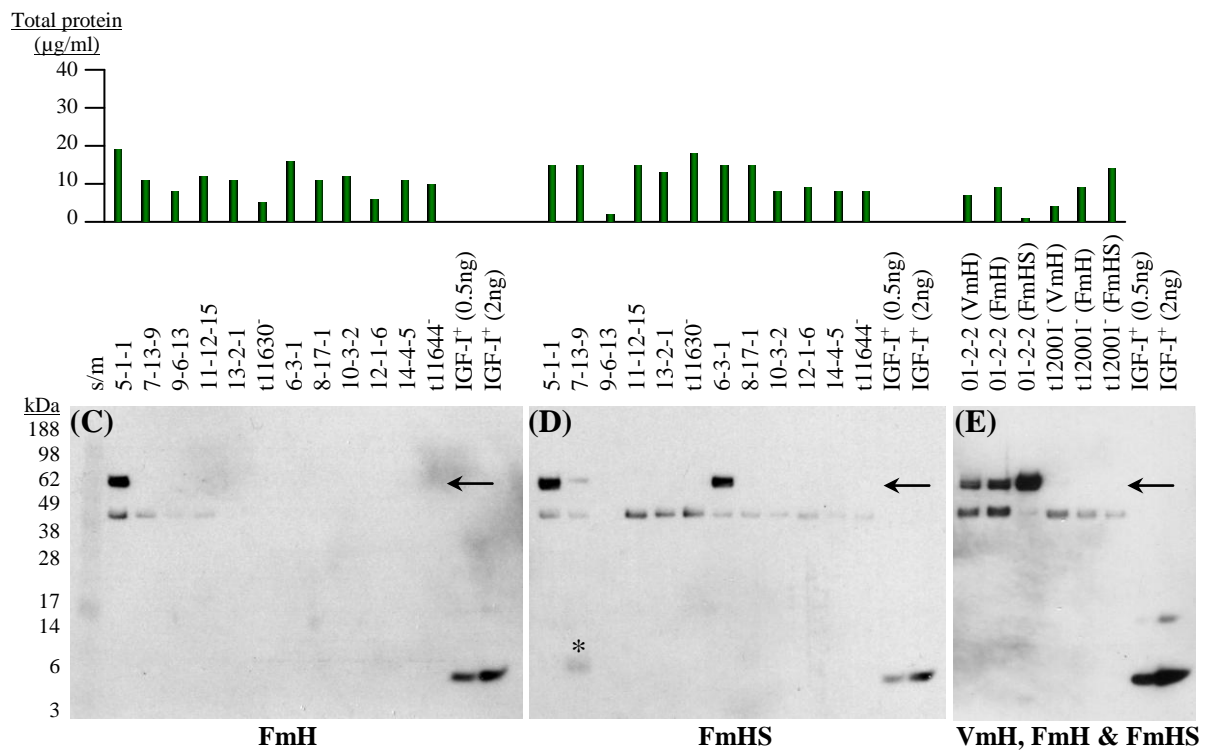
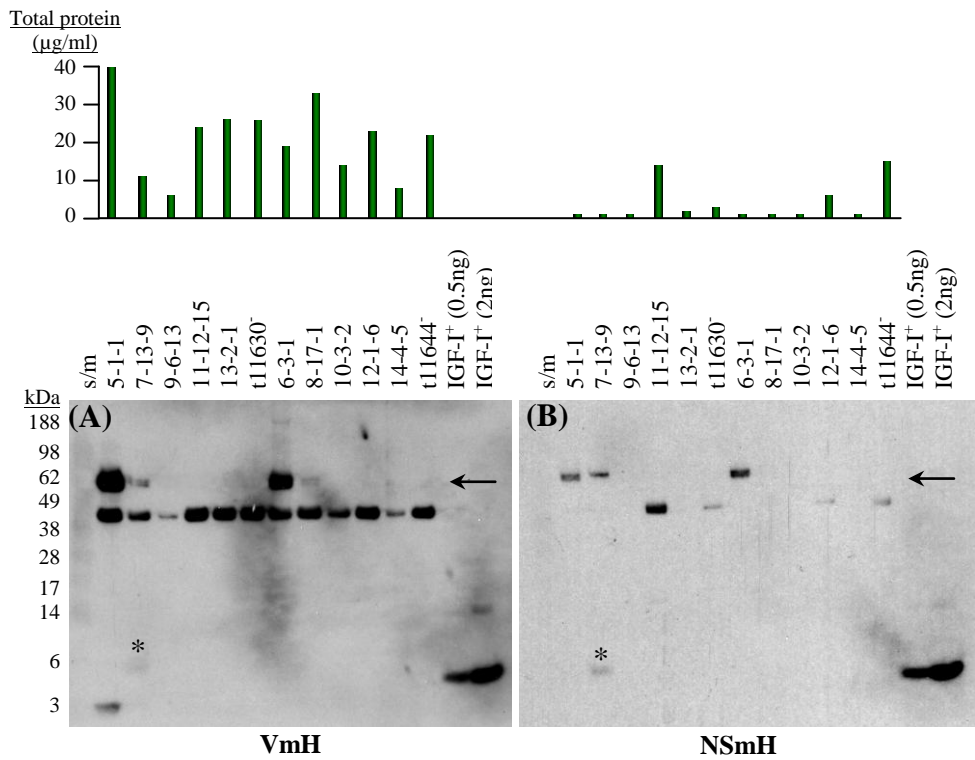
Initial characterisation of the different IGF-I expression strains was performed by western blot analysis. The culture conditions previously used for pFENIGFI ectopic transformants may not have been optimal for the other IGF-I expression strains as the inherent differences in the IGF-I expression vectors could have potentially affected the intracellular trafficking and/or secretion of IGF-I in *Neurospora*. In an attempt to address this potential issue, western blots of the IGF-I expression strains were performed on liquid cultures grown for various times (48, 72 & 96hrs) and in different media (VmH, NSmH, FmH or FmHS).

Three replicate *Neurospora* liquid cultures were prepared in 125ml Erlenmeyer flasks with 25ml of VmH, NSmH, FmH or FmHS (2.2.1.1.4). A 1:200 dilution of PIC was added to the medium in each culture flask prior to inoculation with  $2 \times 10^5$  of 10 day-old macroconidia from the IGF-I expression strains and untransformed controls. For comparative purposes, liquid cultures of the pFENIGFI/t12001 ectopic transformant 01-2-2 were similarly prepared except for the omission of NSmH cultures as the 01-2-2 and t12001 strains were non-viable in NSmH. The liquid cultures were incubated at 25°C in the absence of light on an orbital shaker and media from the replicate liquid cultures were harvested at 48, 72 & 96 hours (2.2.5.1.1). The total protein in the culture supernatant was quantified by Bradford assay (2.2.5.3.1) and analysed by

western blot (2.2.5.6) using an anti-human IGF-I polyclonal antibody from R&D systems (2.1.6) (Figures 4.15, 4.16 & 4.17). As described previously, the total protein for each sample loaded into the SDS-PAGE gels used for western blotting was not standardised as maximum sensitivity was required to determine if mature protein was being processed and secreted at detectable levels.

The western blot results for the 48hr cultures (Figure 4.15) indicated the supernatant from the 7-13-9 (pFENIGFI-GS-DV2/t11630) transformant contained a putative mature IGF-I band of ~7kDa in the VmH, NSmH & FmHS but not in FmH. The putative mature IGF-I band was not present in the 72hr (Figure 4.16) and 96hr (Figure 4.17) cultures. Relative comparison of band intensities indicated the putative mature IGF-I bands in the 48hr cultures were substantially less than the 0.5ng positive IGF-I control which suggested the mature IGF-I concentration in the supernatants of these isolates were less than 43ng/ml (5.25nM).

Comparison of transformant 7-13-9 in the four types of media indicated the putative mature IGF-I band did not appear to be correlated with the total protein amount as the 7-13-9/NSmH culture had a mature IGF-I band yet had ~10-fold less total protein compared to the 7-13-9/FmH sample in which no mature IGF-I band was observed. Furthermore, the 7-13-9/FmH had similar total protein levels to 7-13-9/VmH yet the latter contained a putative mature IGF-I band whilst the former did not. The presence or absence of mature IGF-I could have been due to variation in the types of extracellular proteins secreted by *Neurospora* in the different culture media. For example, derepression of extracellular protease gene expression can be effected by limitations in carbon, nitrogen and/or sulphur (Hanson & Marzluf 1975); (Lindberg *et al.*, 1982). Thus, variation in media composition potentially could have led to differential nutrient depletion by the *Neurospora* cultures, altering the composition of extracellular protein in one medium relative to another which may have directly or indirectly affected the secreted IGF-I protein.

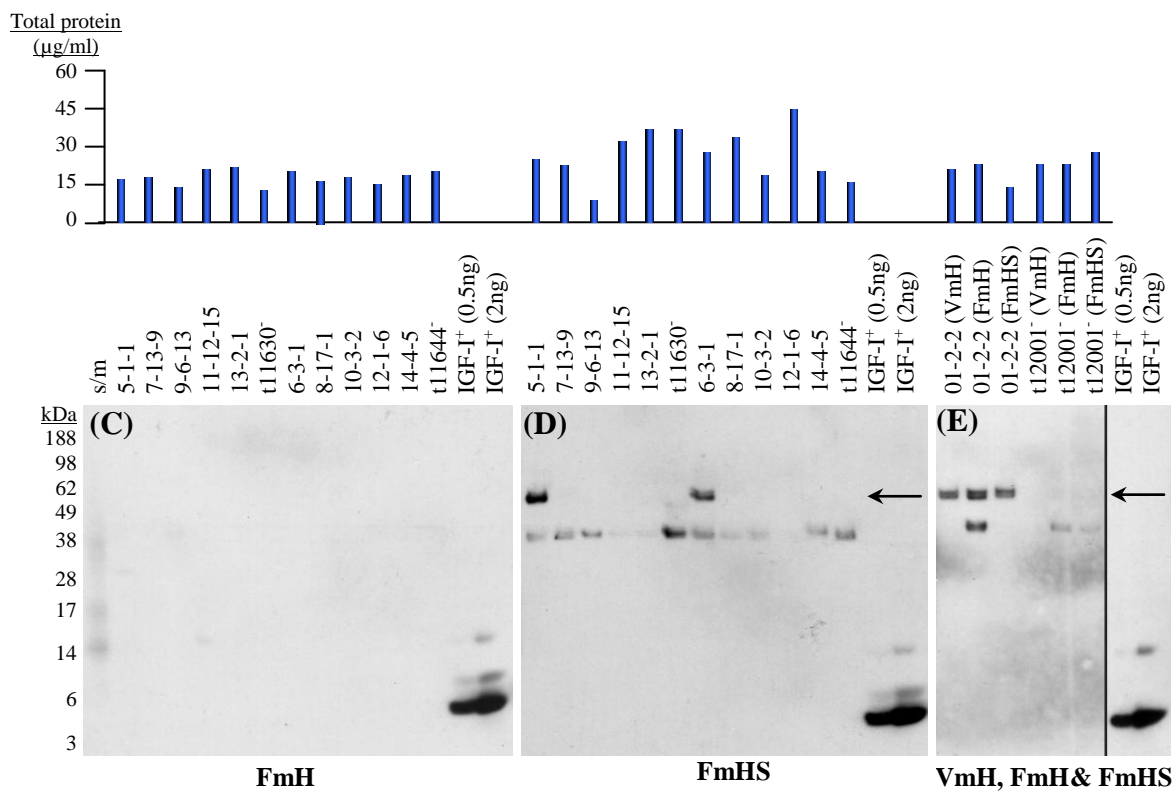
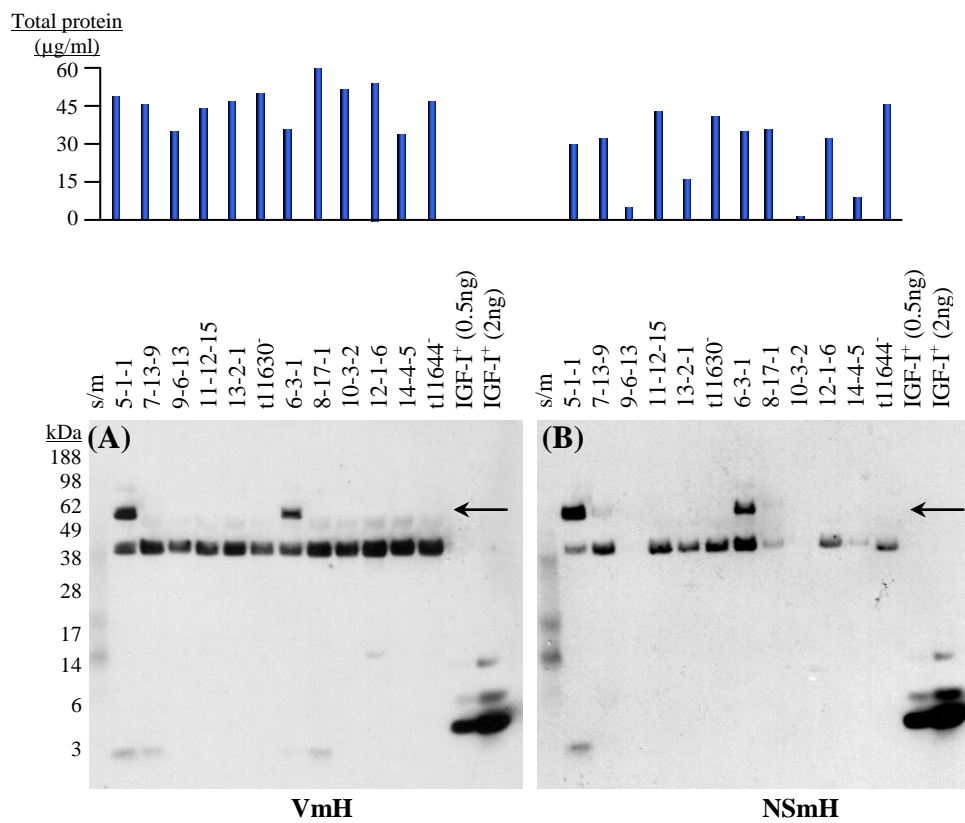




<b><u>Transformant labels</u></b>	
5-1-1	= pFENIGFI-DV2 in t11630
6-3-1	= pFENIGFI-DV2 in t11644
7-13-9	= pFENIGFI-GS-DV2 in t11630
8-17-1	= pFENIGFI-GS-DV2 in t11644
9-6-13	= pNEIGFI-G1-DV2 in t11630
10-3-2	= pNEIGFI-G1-DV2 in t11644
11-12-15	= pNEhIGFI-C1-DV2 in t11630
12-1-6	= pNEhIGFI-C1-DV2 in t11644
13-2-1	= pNEhIGFI-H1-DV2 in t11630
14-4-5	= pNEhIGFI-H1-DV2 in t11644
01-2-2	= pFENIGFI in t12001 (ectopic)
s/m	= SeeBlue Plus2 size marker
IGF-I <sup>+</sup>	= recombinant IGF-I protein
t11630 <sup>-</sup>	= untransformed –ve control
t11644 <sup>-</sup>	= untransformed –ve control
t12001 <sup>-</sup>	= untransformed –ve control

**Figure 4.15 Western blot of secreted protein in various 48-hour culture media from pFENIGFI-DV2, pFENIGFI-GS-DV2, pNEIGFI-G1-DV2, pNEhIGFI-C1-DV2, pNEhIGFI-H1-DV2 & pFENIGFI *Neurospora* transformants.**

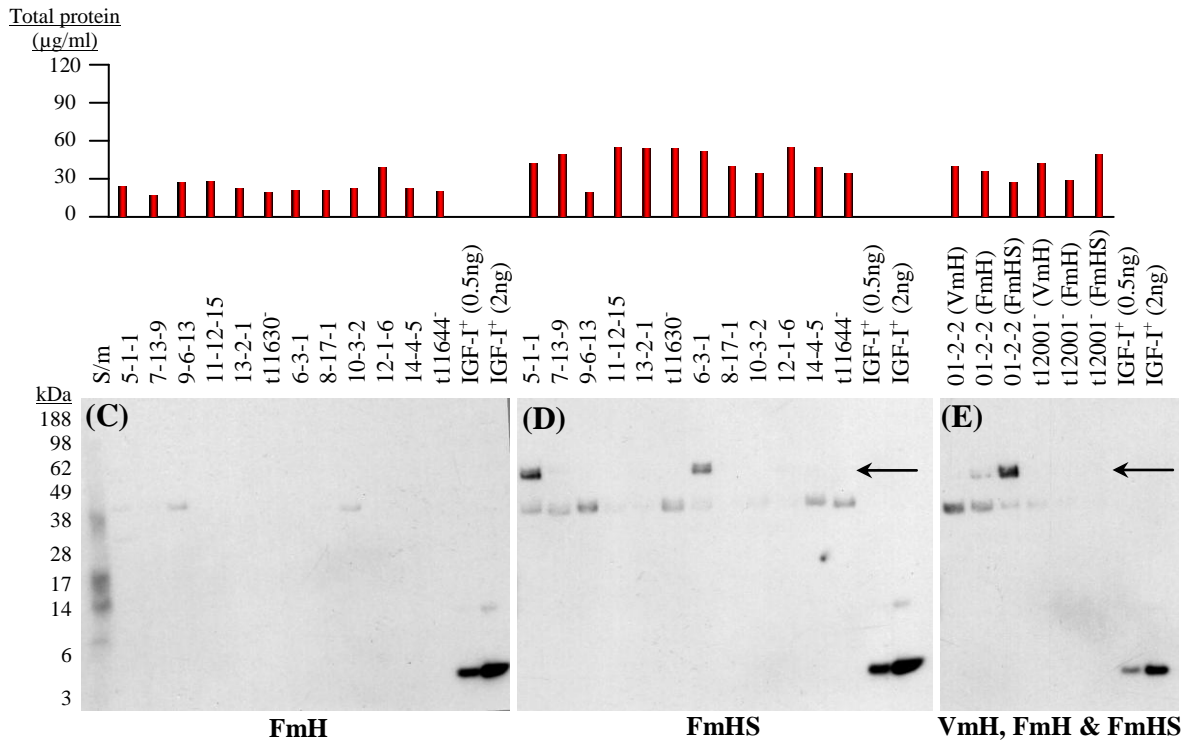
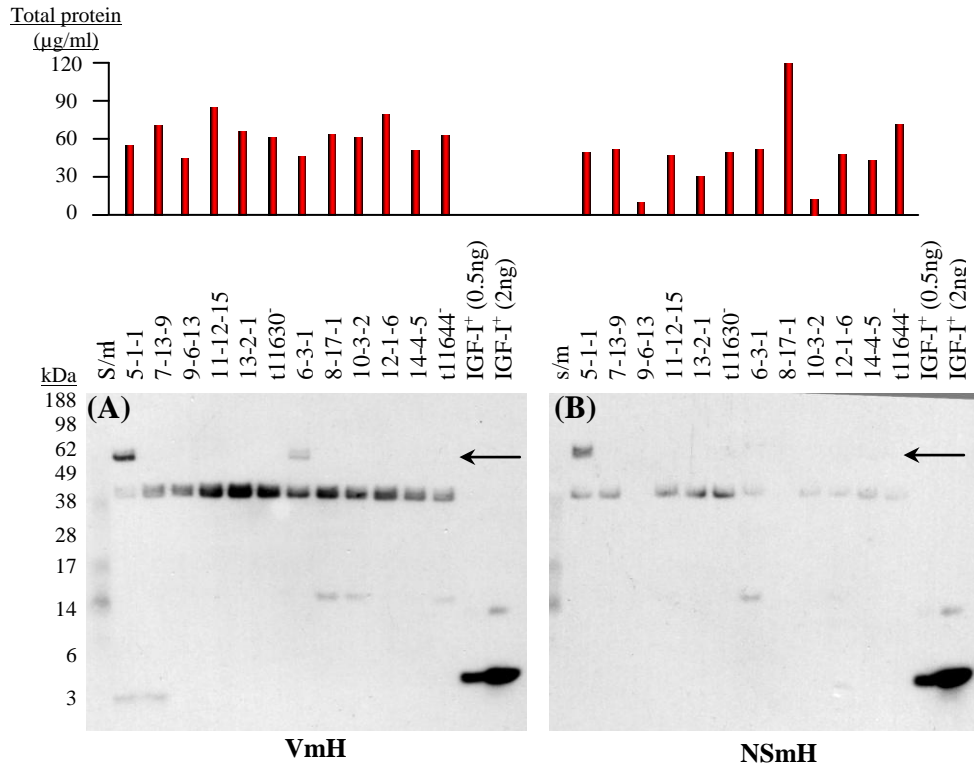
Culture flasks containing VmH, NSmH, FmH or FmHS were inoculated with  $1.65 \times 10^5$  macroconidia from each *Neurospora* transformant. A 1:200 dilution of protease inhibitor cocktail (Sigma) was added to each flask. Cultures were incubated for 48 hours and the media were analysed by western blot using a polyclonal anti-human IGF-I primary antibody. Note: putative mature IGF-I bands are indicated by \* and putative gla-1:IGF-I fusion bands by arrows. The total protein concentrations in the media were determined by Bradford assay and are indicated in the bar graphs above the western blots.



<b><u>Transformant labels</u></b>	
5-1-1	= pFENIGFI-DV2 in t11630
6-3-1	= pFENIGFI-DV2 in t11644
7-13-9	= pFENIGFI-GS-DV2 in t11630
8-17-1	= pFENIGFI-GS-DV2 in t11644
9-6-13	= pNEIGFI-G1-DV2 in t11630
10-3-2	= pNEIGFI-G1-DV2 in t11644
11-12-15	= pNEhIGFI-C1-DV2 in t11630
12-1-6	= pNEhIGFI-C1-DV2 in t11644
13-2-1	= pNEhIGFI-H1-DV2 in t11630
14-4-5	= pNEhIGFI-H1-DV2 in t11644
01-2-2	= pFENIGFI in t12001 (ectopic)
s/m	= SeeBlue Plus2 size marker
IGF-I <sup>+</sup>	= recombinant IGF-I protein
t11630 <sup>-</sup>	= untransformed –ve control
t11644 <sup>-</sup>	= untransformed –ve control
t12001 <sup>-</sup>	= untransformed –ve control

**Figure 4.16** Western blot of secreted protein in various 72-hour culture media from pFENIGFI-DV2, pFENIGFI-GS-DV2, pNEIGFI-G1-DV2, pNEhIGFI-C1-DV2, pNEhIGFI-H1-DV2 & pFENIGFI *Neurospora* transformants.

Culture flasks containing VmH, NSmH, FmH or FmHS were inoculated with  $1.65 \times 10^5$  macroconidia from each *Neurospora* transformant. A 1:200 dilution of protease inhibitor cocktail (Sigma) was added to each flask. Cultures were incubated for 72 hours and the media were analysed by western blot using a polyclonal anti-human IGF-I primary antibody. Note: putative gla-1:IGF-I fusion bands are indicated by arrows. The total protein concentrations in the media were determined by Bradford assay and are indicated in the bar graphs above the western blots.



<b>Transformant labels</b>	
5-1-1	= pFENIGFI-DV2 in t11630
6-3-1	= pFENIGFI-DV2 in t11644
7-13-9	= pFENIGFI-GS-DV2 in t11630
8-17-1	= pFENIGFI-GS-DV2 in t11644
9-6-13	= pNEIGFI-G1-DV2 in t11630
10-3-2	= pNEIGFI-G1-DV2 in t11644
11-12-15	= pNEhIGFI-C1-DV2 in t11630
12-1-6	= pNEhIGFI-C1-DV2 in t11644
13-2-1	= pNEhIGFI-H1-DV2 in t11630
14-4-5	= pNEhIGFI-H1-DV2 in t11644
01-2-2	= pFENIGFI in t12001 (ectopic)
s/m	= SeeBlue Plus2 size marker
IGF-I <sup>+</sup>	= recombinant IGF-I protein
t11630 <sup>-</sup>	= untransformed –ve control
t11644 <sup>-</sup>	= untransformed –ve control
t12001 <sup>-</sup>	= untransformed –ve control

**Figure 4.17** Western blot of secreted protein in various 96-hour culture media from pFENIGFI-DV2, pFENIGFI-GS-DV2, pNEIGFI-G1-DV2, pNEhIGFI-C1-DV2, pNEhIGFI-H1-DV2 & pFENIGFI *Neurospora* transformants.

Culture flasks containing VmH, NSmH, FmH or FmHS were inoculated with  $1.65 \times 10^5$  macroconidia from each *Neurospora* transformant. A 1:200 dilution of protease inhibitor cocktail (Sigma) was added to each flask. Cultures were incubated for 96 hours and the media were analysed by western blot using a polyclonal anti-human IGF-I primary antibody. Note: putative gla-1:IGF-I fusion bands are indicated by arrows. The total protein concentrations in the media were determined by Bradford assay and are indicated in the bar graphs above the western blots.

#### CHAPTER 4 – Transformation, expression and diversification of IGF-I in Neurospora

The equal loading of total protein in the western blot of transformants 5-1-1 (pFENIGFI-DV2/t11630) and 7-13-9 (pFENIGFI-GS-DV2/11630) cultured in FmHS (Figure 4.15D) enabled a direct comparison between the relative densities of the putative gla-1:IGF-I fusion bands in these two vectors. The presence of the putative mature IGF-I band and lesser density of the putative fusion protein band in transformant 7-13-9 suggested the glycine spacer in pFENIGFI-GS-DV2 enabled, or increased, the cleavage efficiency of the gla-1-IGF-I fusion protein liberating sufficient mature IGF-I into the culture medium to be detected by western blot.

The western blot of 7-13-9 (pFENIGFI-GS-DV2/t11630) and 8-17-1 (pFENIGFI-GS-DV2/t11644) cultures in NSmH (Figure 4.15B) and FmHS (Figure 4.15D) contained equivalent quantities of total protein yet the putative mature IGF-I band was only present in 7-13-9. Similarly, there was a greater amount of putative gla-1:IGF-I fusion protein in 7-13-9 relative to 8-17-1 which indicated IGF-I yield varied between the t11630 and t11644 strains.

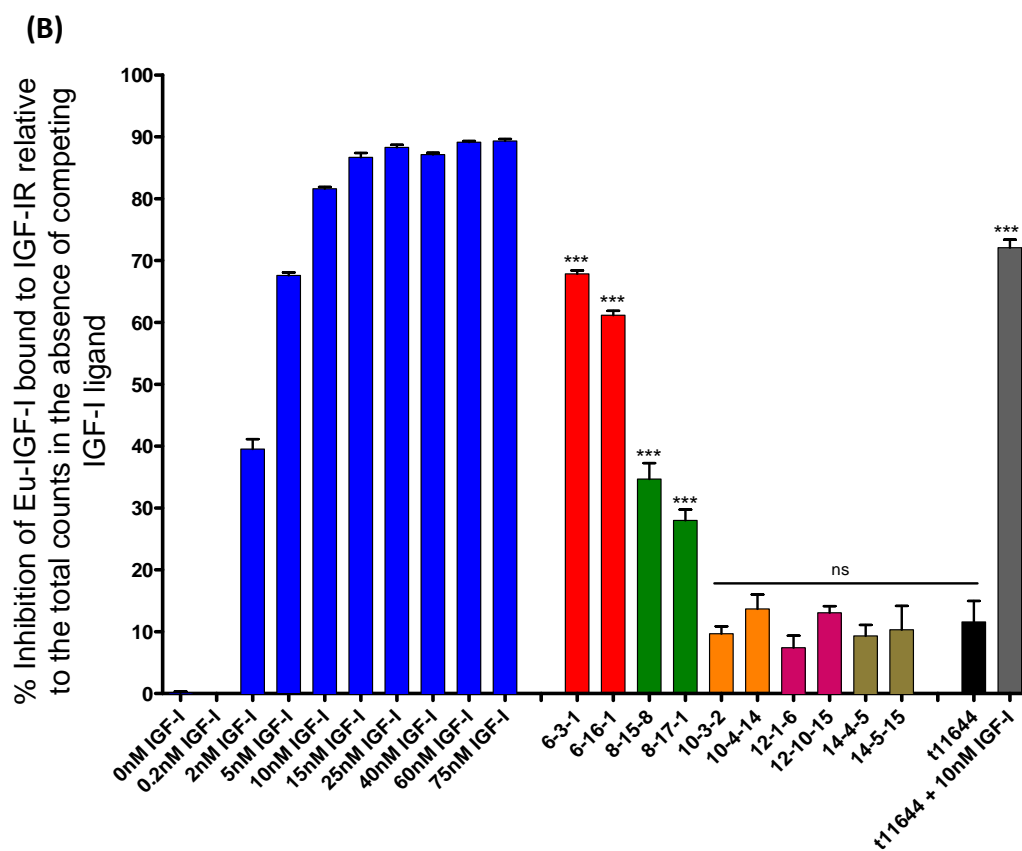
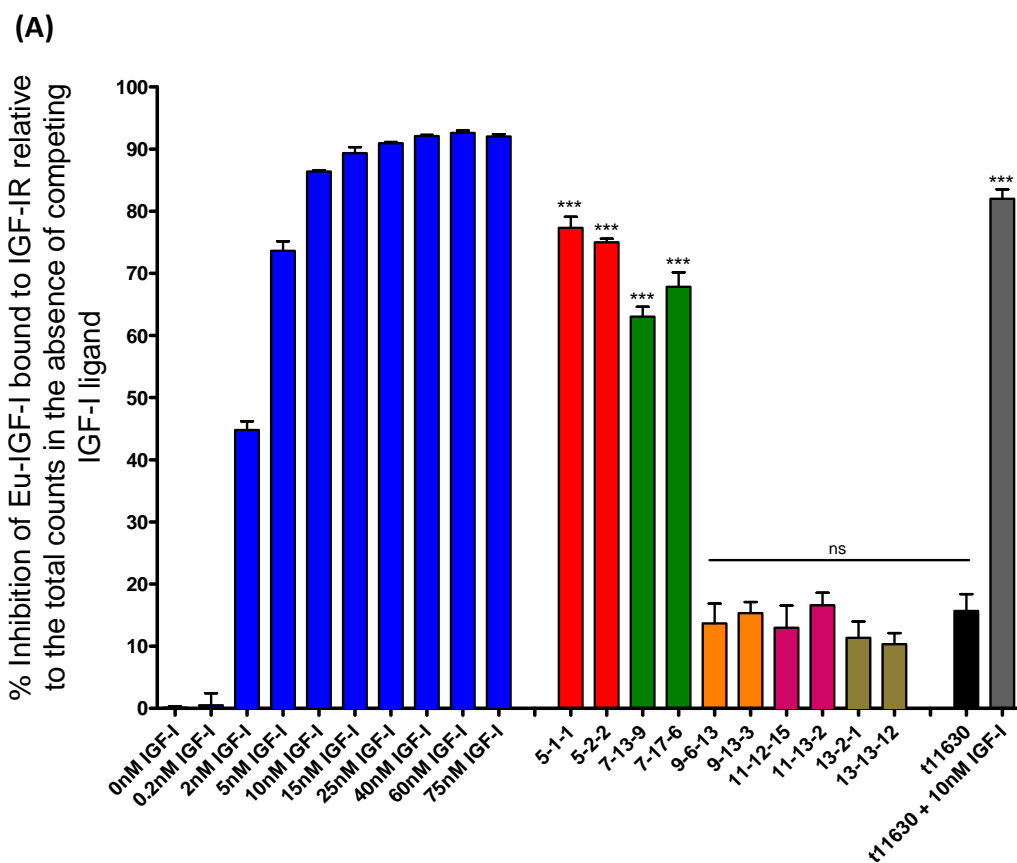
An ~3kDa band was present in pFENIGFI, pFENIGFI-DV2 & pFENIGFI-GS-DV2 transformants 5-1-1 / VmH / 48, 72, & 96hrs; 7-13-9 / VmH / 72 & 96hrs; 6-3-1 / VmH / 72hrs; 8-17-1 / VmH / 72 hrs and 01-2-2 / FmHS / 48hrs (Figures 4.15, 4.16 & 4.17). The ~3kDa band was smaller than the expected ~7.5kDa molecular weight of mature IGF-I and could have been a fragment of proteolytically cleaved IGF-I, secretion of a truncated IGF-I protein or may simply have been a protein cross-reacting with the polyclonal IGF-I antibody. However, the latter explanation seemed unlikely as there was an equivalent or greater amount of total protein in the negative controls t11630 / VmH / 72hrs and t12001 / FmHS / 48 hours relative to the corresponding transformants 5-1-1, 7-13-9, 8-17-1 / VmH / 72hrs and 01-2-2 / FmHS / 48hrs, respectively, yet the negative controls did not contain the ~3kDa band.

**4.3.3 IGF-IR ligand binding assay of the IGF-I expression constructs in *Neurospora***

Frozen aliquots of 48hr FmHS culture supernatants from ten *Neurospora* strains transformed with the five targeted expression constructs (4.3.2) were thawed and analysed by IGF-IR ligand binding assay (Figure 4.18) (2.2.6.2). The IGF-I concentration in the *Neurospora* supernatants was determined by non-linear regression with a standard IGF-I curve diluted in FmHS. The measured concentration was doubled to correct for the assay dilution factor. The background fluorescence levels of the t11630 and t11644 negative controls were subtracted from the relevant transformant measurements. Upon correction, only transformants 5-1, 5-2, 6-3 & 6-16 (pFENIGFI-DV2) and 7-13, 7-17, 8-15 & 8-17 (pFENIGFI-GS-DV2) produced detectable quantities of IGF-I ranging from 0.9nM to 11.7nM (Figure 4.18C). Although assaying of transformants 10-4-14 (pNEIGFI-G1-DV2) and 11-13-2 & 12-10-15 (pNEhIGFI-C1) resulted in IGF-I measurements of ~0.1nM, these values probably represented small fluctuations in background fluorescence rather than true IGF-I measurements.

Western blot analysis of FmHS / 48hr culture supernatants indicated 5-1-1 (pFENIGFI-DV2/t11630) did not have a detectable mature IGF-I band but had a relatively dense gla-1:IGF-I fusion band whereas 7-13-9 (pFENIGFI-GS-DV2/t11630) contained a mature IGF-I band but had a weak gla-1:IGF-I fusion band (Figure 4.15C). However, the IGF-IR ligand binding assay of the same 5-1-1 and 7-13-9 culture supernatants (Figure 4.18C) suggested 5-1-1 supernatant had 11.7nM of IGF-I whilst 7-13-9 had 4.3nM. The high IGF-IR binding assay measurement for 5-1-1 relative to 7-13-9 suggested the gla-1:IGF-I fusion protein was possibly competitively displacing Eu-IGF-I from the IGF-IR in the ligand binding assay.

Western blot analysis indicated the yield of mature IGF-I was lower in the t11644 transformants relative to the t11630 transformants (as described in 4.3.2). The IGF-IR ligand binding assay measurements of 7-13-9 (4.3nM) & 7-17-6 (6.8nM) (pFENIGFI-GS-DV2/t11630) were higher than 8-15-8 (1.3nM) & 8-17-1 (0.9nM) (pFENIGFI-GS-DV2/t11644) (Figure 4.18) supporting the assertion of differential IGF-I yield in the t11630 and t11644 strains.





(C)

<b>Estimated concentration of IGF-I secreted by Neurospora transformants:</b>					
<b>t11630 transformants</b>			<b>t11644 transformants</b>		
	nM	ng/ml		nM	ng/ml
<b>5-1-1</b> (pFENIGFI-DV2)	11.7	91.4	<b>6-3-1</b> (pFENIGFI-DV2)	9.1	71.2
<b>5-2-2</b> (pFENIGFI-DV2)	9.2	72.0	<b>6-16-1</b> (pFENIGFI-DV2)	6.8	52.9
<b>7-13-9</b> (pFENIGFI-GS-DV2)	4.3	35.3	<b>8-15-8</b> (pFENIGFI-GS-DV2)	1.3	10.7
<b>7-17-6</b> (pFENIGFI-GS-DV2)	6.8	55.7	<b>8-17-1</b> (pFENIGFI-GS-DV2)	0.9	7.4
<b>9-6-13</b> (pNEIGFI-G1-DV2)	0	0	<b>10-3-2</b> (pNEIGFI-G1-DV2)	0	0
<b>9-13-3</b> (pNEIGFI-G1-DV2)	0	0	<b>10-4-14</b> (pNEIGFI-G1-DV2)	0.1	0.8
<b>11-12-15</b> (pNEhIGFI-C1-DV2)	0	0	<b>12-1-6</b> (pNEhIGFI-C1-DV2)	0	0
<b>11-13-2</b> (pNEhIGFI-C1-DV2)	0.1	0.8	<b>12-10-15</b> (pNEhIGFI-C1-DV2)	0.1	0.8
<b>13-2-1</b> (pNEhIGFI-H1-DV2)	0	0	<b>14-4-5</b> (pNEhIGFI-H1-DV2)	0	0
<b>13-13-12</b> (pNEhIGFI-H1-DV2)	0	0	<b>14-5-15</b> (pNEhIGFI-H1-DV2)	0	0

**Figure 4.18 IGF-IR ligand binding assay of secreted protein from pFENIGFI-DV2, pFENIGFI-GS-DV2, pNEIGFI-G1-DV2, pNEhIGFI-C1-DV2 & pNEhIGFI-H1-DV2 Neurospora t11630 & t11644 transformants.**

Culture flasks containing FmHS with a 1:200 dilution of protease inhibitor cocktail (Sigma) were inoculated with  $1.65 \times 10^5$  of macroconidia from transformants in the (A) t11630 and (B) t11644 strains and incubated for 48 hours. 96-well plates coated with anti-IGF-IR (24-31) antibody was used to capture human IGF-IRs from P6 cell lysates. IGF-IRs were incubated overnight with a combination of Eu-IGF-I and medium from each transformant culture. The plates were washed and remaining Eu-IGF-I bound to IGF-IRs was measured by time-resolved fluorescence. The standard curve was based on a one-site competition binding model and the *Neurospora* sample measurements were fitted to the curve by non-linear regression using Prism. The results are presented as a percentage inhibition of Eu-IGF-I binding to IGF-IR in the absence of competing IGFI in LBB. The statistical significance of differences between transformants relative to the untransformed controls was determined by ANOVA with Tukey's multiple comparison test. Note: \*\*\* =  $p < 0.001$  & ns = not significant ( $p > 0.05$ ).

(C) The secreted IGF-I concentrations were corrected by subtraction of background measurements from non-transformed strain controls and the estimate doubled to account for the inherent assay dilution factor with the exception of the untransformed controls spiked with 10nM of recombinant IGF-I as the dilution factor had been accounted for during assay preparation.

The pFENIGFI-GS-DV2 fusion vector appeared to facilitate secretion of mature IGF-I in *Neurospora* through increased, albeit incomplete, cleavage of the gla-1:IGF-I fusion protein. However, the inability of the IGF-IR ligand binding assay to distinguish between mature IGF-I and uncleaved gla-1:IGF-I fusion protein could have confounded resolution of IGF-IR binding affinities of IGF-I variants. Thus, additional refinement of the *Neurospora*/IGF-I expression system was required before *in-vivo* screening could be undertaken. However, before continuing attempts at improving *Neurospora* secretion of IGF-I, additional issues arose during the characterisation of the *Neurospora* transformants that needed to be addressed first. Specifically, substantial variation in the levels of secreted total protein existed between transformants of the same genetic background despite standardisation of inoculum size, culturing, harvesting and processing conditions, (Figures 4.15, 4.16 & 4.17). Although laborious in the context of high-throughput screening, it was feasible to perform total protein assays to standardise the total protein assayed from each transformant.

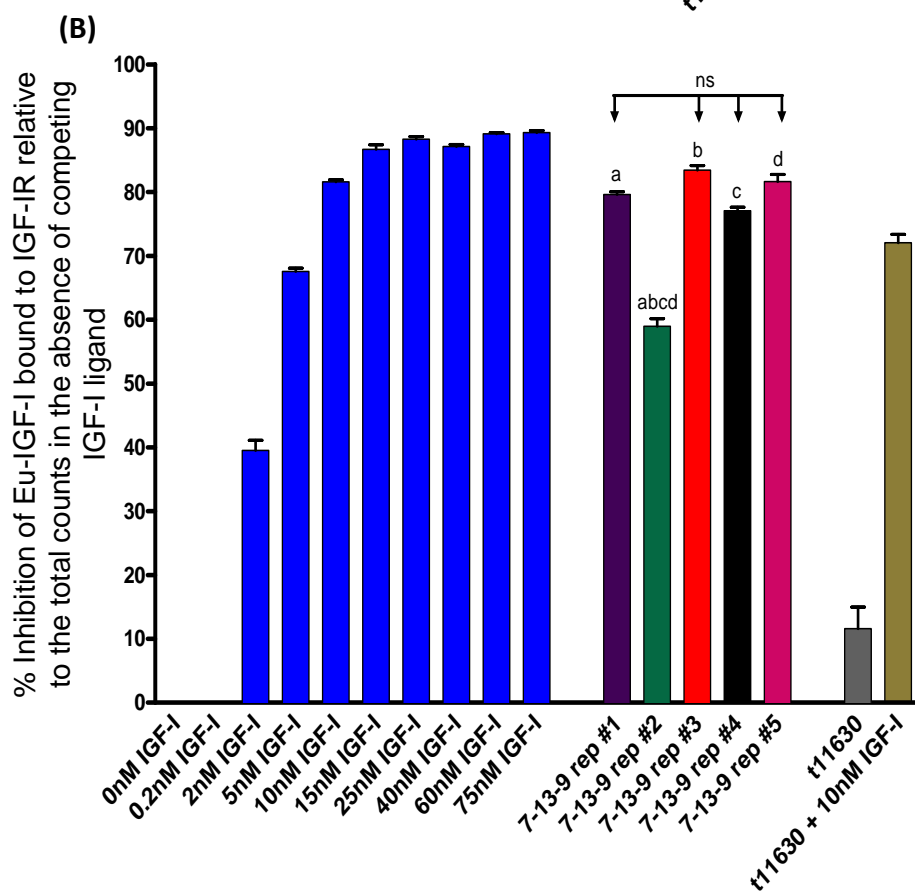
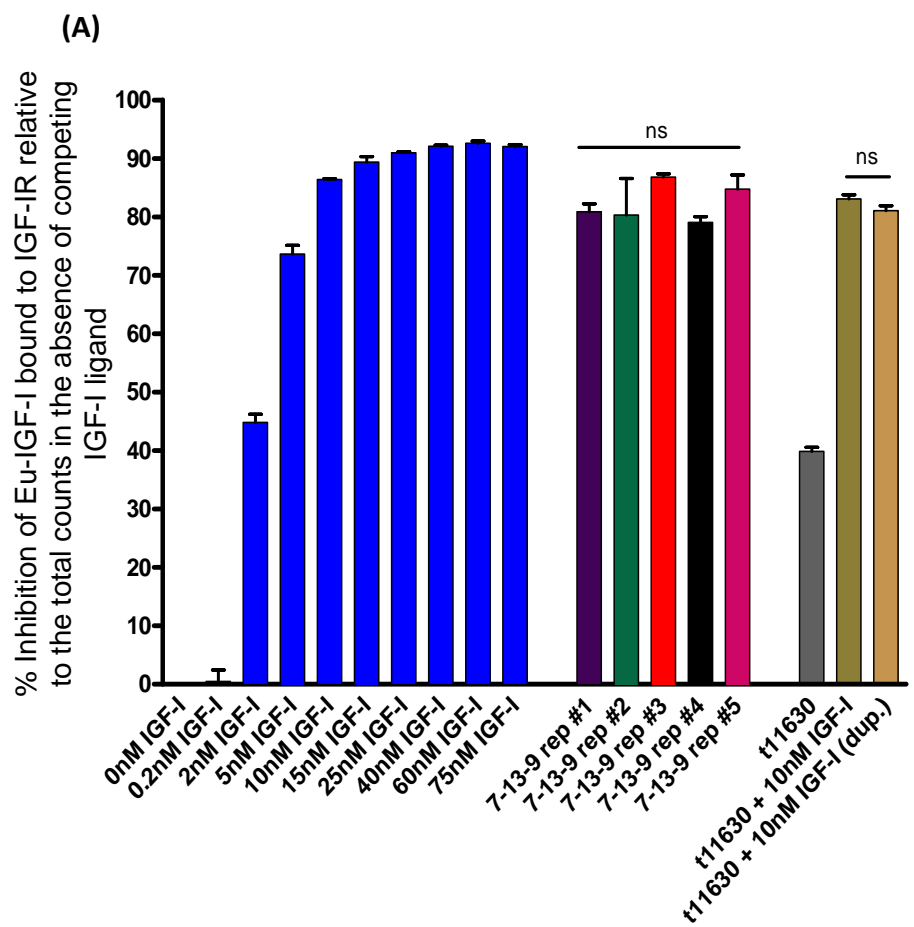
#### **4.3.3.1 Characterisation of IGF-I variability in *Neurospora* cultures and the IGF-IR ligand binding assay**

It was necessary to determine if standardisation of *Neurospora* culture supernatants based on total protein levels would reduce measured IGF-I variation between samples and enable reasonably robust screening of secreted IGF-I variants using the IGF-IR ligand binding assay. In addition, a cursory validation of the IGF-IR ligand binding assay was performed to determine the assay's suitability for screening IGF-I variants secreted by *Neurospora*. To this end, five replicate cultures of *Neurospora* transformant 7-13-9 (pFENIGFI-GS-DV2/t11630) and 10nM IGF-I spiked t11630 controls were used to characterise the intra and inter-plate variation of the IGF-IR ligand binding plate assay. Transformant 7-13-9 was selected for assay validation as western blot analysis detected mature IGF-I protein (Figure 4.15D) and a previous IGF-IR ligand binding assay of 7-13-9 measured 4.3nM of IGF-I (Figure 4.18) although it was likely this measurement also included uncleaved gla-1:IGF-I fusion protein.

CHAPTER 4 – Transformation, expression and diversification of IGF-I in *Neurospora*

Five replicate *Neurospora* liquid cultures of 7-13-9 and a single culture of the untransformed t11630 control were prepared in 125ml Erlenmeyer flasks with 25ml of FmHS (2.2.1.1.4). A 1:200 dilution of PIC was added to the medium in each culture flask prior to inoculation with  $1.65 \times 10^5$  of 7 day-old macroconidia. The liquid cultures were incubated at 25°C in the absence of light on an orbital shaker. The culture supernatants were harvested after 48 hours (2.2.5.1.1). The total protein in the culture supernatant was quantified by Bradford assay (2.2.5.3.1) and standardised to 13µg/ml which equalled the replicate with the lowest total protein concentration. Duplicate t11630 controls spiked with 10nM of recombinant IGF-I were also prepared. The culture supernatants were analysed by IGF-IR ligand binding assay in duplicate plates (Figure 4.19) (2.2.6.2). IGF-I standard curves were generated in FmHS and the measured IGF-I concentration in the *Neurospora* supernatants were doubled to correct for the assay dilution factor. The mean background fluorescence of the t11630 untransformed control was subtracted from the transformant measurements.

The t11630 controls spiked with 10nM IGF-I measured as 9.1nM & 7.6nM in assay 'A' and 6.2nM in assay 'B' (assay B did not contain a duplicate spiked sample) (Figure 4.19). These controls had less than the 10nM of IGF-I spiked into the samples, however, the discrepancies in absolute terms were not unexpected given the presence of extracellular proteases in the culture supernatants. In relative terms, the IGF-I spiked t11630 controls in assay A had a 9.8 %CV between duplicates indicating the intra-plate variation of the spiked samples was acceptable.



(C)

<b>Estimated concentration of IGF-I secreted by <i>Neurospora</i> transformants:</b>					
<b>Samples</b>	<b>Assay A</b>		<b>Assay B</b>		<b>Inter-assay variation</b>
	nM	ng/ml	nM	ng/ml	(%CV)
7-13-9 rep #1	14.0	115	21.3	174	17.4
7-13-9 rep #2	13.3	109	6.1	50	35.5
7-13-9 rep #3	26.0	213	30.7	252	8.9
7-13-9 rep #4	11.9	97	17.4	142	14.4
7-13-9 rep #5	20.5	168	25.6	210	21.4
<b>Intra assay variation (%CV)</b> (between 7-13-9 replicates)	30.9		38.7		
<b>Controls</b>					
t11630 +10nM IGF-I	9.1	70	6.2	47.3	9.8
t11630 +10nM IGF-I (dup.)	7.6	58	nd	nd	

**Figure 4.19** Intra and inter-assay variation of secreted IGF-I in replicate *Neurospora* 7-13-9 transformant cultures analysed by the IGF-IR ligand binding assay.

**(A)** Assay A: replicate culture variability was determined using the IGF-IR ligand binding assay with 5 replicate cultures of *Neurospora* transformant 7-13-9 (pFENIGFI-GS-DV2/t11630). Replicate culture flasks containing FmHS medium with a 1:200 dilution of protease inhibitor cocktail (Sigma) were inoculated with  $1.65 \times 10^5$  of macroconidia from transformant 7-13-9 and incubated for 48 hours. Duplicate 96-well plates coated with anti-IGF-IR (24-31) antibody were used to capture human IGF-IRs from P6 cell lysates and were incubated overnight with a combination of Eu-IGF-I and medium from each replicate culture, t11630 untransformed control supernatant or t11630 supernatant spiked with 10nM of recombinant IGF-I. The plates were washed and remaining Eu-IGF-I bound to IGF-IRs was measured by time-resolved fluorescence. The standard curve was based on a one-site competition binding model and the *Neurospora* sample measurements were fitted to the curve by non-linear regression using Prism GraphPad. The results are presented as a percentage inhibition of Eu-IGF-I binding to IGF-IR in the absence of competing IGF-I. **(B)** Assay B: duplicate assay with the exception of the t11630 +10nM IGF-I duplicate sample which was omitted from this plate. Note: The statistical significance of differences between transformants relative to the untransformed controls was determined by ANOVA with Tukey's multiple comparison test. Note: a, b, c & d =  $p < 0.001$  & ns = not significant ( $p > 0.05$ ).

**(C)** The secreted IGF-I concentrations were corrected by subtraction of background measurements from the t11630 non-transformed strain control and the estimate doubled to account for the inherent assay dilution factor with the exception of the t11630 untransformed control spiked with 10nM of recombinant IGF-I as the dilution factor had been accounted for during assay preparation. Note: Intra-assay variation is the %CV of the combined 7-13-9 replicate samples within each plate. Inter-assay variation is the %CV of the individual 7-13-9 replicates on both plates. nd = not determined as only assay A contained the duplicate 10nM IGF-I spiked control.

The intra-assay variation as determined by the %CV of the combined 7-13-9 replicates in assays A and B were 30.9% and 38.7%, respectively (Figure 4.19C). The high level of intra-assay variation indicated there was substantial variability between the replicate cultures despite the standardization of inoculum size and total protein prior to assaying. The variation in measured IGF-I levels was probably due, at least in part, to the effects of proteases within the culture supernatants. Various factors such as minor differences in secreted protease concentrations and slight variation in processing time could have facilitated differential proteolytic degradation levels in the replicate cultures prior to assaying. The substantial variation in the IGF-I levels in the culture supernatants suggested standardisation of total protein levels was unlikely to be an effective way of normalising the samples prior to screening.

There was a variable degree of inter-plate variation in the IGF-IR ligand binding assays determined by the %CV of each 7-13-9 replicate from both plates which ranged from 8.9 to 35.5% (Figure 4.19C). Receptor binding assays are typically utilised to characterise the interactions between purified proteins and immunocaptured receptors in an optimal buffer solution as opposed to assaying the interaction of receptors and a specific protein within a complex mixture that may contain inhibitory molecules and proteases. As such, the intra and inter-assay variation observed in this experiment was attributed to inherent limitations of the IGF-IR ligand binding assay format for measuring IGF-I in *Neurospora* culture supernatants.

There were numerous technical obstacles that still needed to be overcome before *in-vivo* screening of diversified IGF-I proteins could be attempted. These obstacles included reducing IGF-I proteolysis in *Neurospora* cultures, increasing secretion of mature IGF-I, minimizing strain-specific differences in IGF-I yield, reducing variation in the IGF-IR ligand binding assay or adaption of an alternative screening assay. However, these issues were unlikely to be resolved readily so the research focus was directed towards achieving other project goals.

A core objective of the *Neurospora* IGF-I diversification project was the utilization of RIP to generate IGF-I variants. Thus, the final aspect of IGF-I diversification investigated in this project was the generation and characterisation of IGF-I variants at the DNA level.

#### **4.3.4 Characterisation of RIP-generated IGF-I mutants**

RIP generation of IGF-I mutants required the insertion of two copies of IGF-I into the *Neurospora* genome and sexual crossing of two parent strains. The most expedient way to achieve this would have been to ectopically transform pNEIGFI-G1 into existing pNEIGFI-G1-DV2/*Neurospora* transformants to generate strains with two copies of IGF-I. However, RIP mutagenesis of unlinked and linked duplication sequences occurs at frequencies of ~50% and ~100%, respectively. To maximise the number of progeny with mutated IGF-I sequences, a linked IGF-I duplication vector was constructed and cloned into both t11630 and t11644 strains. To this end, an ~800bp duplication containing the *nIGF-I* coding sequence, *gla-1* terminator and part of the *cog/lpl* region was PCR amplified and cloned into pNEIGFI-G1-DV2 to form pNEIGFI-G1-DV2-LD1 (Figure 3.11) (3.1.3.5).

##### **4.3.4.1 Electroporation of the IGF-I duplication vector into *Neurospora***

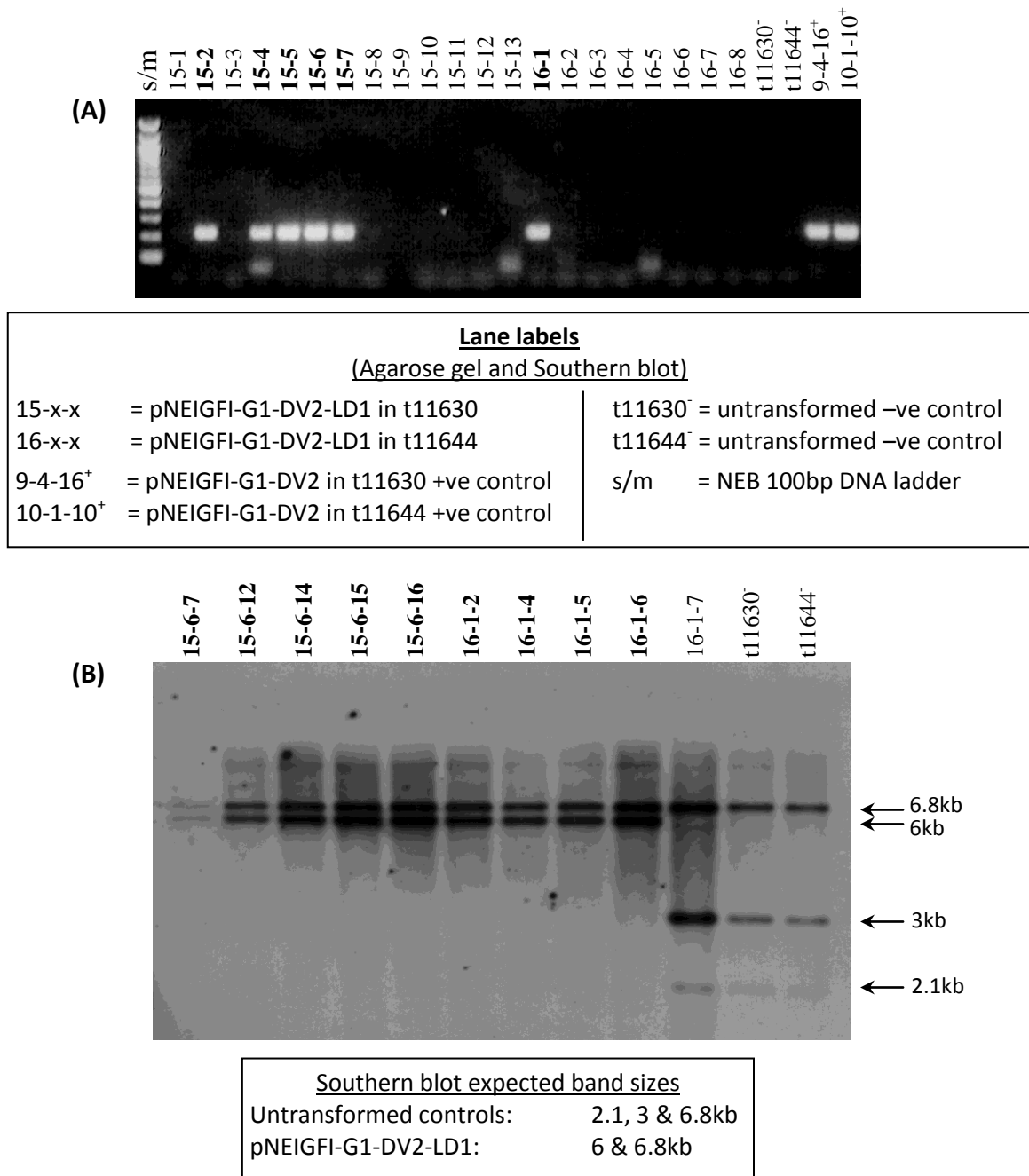
Approximately 5µg of plasmid DNA was linearised by restriction digestion with *PvuII* and electroporated into  $1 \times 10^8$  macroconidia from *Neurospora* strains t11630 & t11644 and plated onto minimal medium (2.2.1.3.1). Thirteen and eight well-isolated primary transformants were selected for strains t11630 & t11644, respectively, and picked onto agar slopes with minimal medium to produce macroconidiating cultures. Genomic DNA was obtained from macroconidia (2.2.1.3.2) and vector integration in the putative transformants was screened for by PCR (2.2.4.6.4) using primers IGFICheck-F and IGFICheck-R (2.1.8.3) located on either flank of the *nIGF-I* coding sequence. Transformants produced amplicons of ~200bp when the PCR reactions were resolved by agarose gel electrophoresis (Figure 4.20A). The PCR screen identified five

transformants from t11630 and one transformant from t11644. All transformants were subjected to microconidiation to obtain homokaryons (2.2.1.2.4). Complementation testing confirmed transformed homokaryons had been obtained for each of the vector/strain combinations. Five of the isolated homokaryons for each of the t11630 and t11644 transformants were analysed by Southern blot for confirmation of correct transplacement.

**4.3.4.2 Southern blot of IGF-I expression vectors transplaced between *his-3* and *cog* in *Neurospora***

Southern blots of selected transformed homokaryons identified in the genetic complementation screen were performed to ensure the constructs were correctly transplaced between *his-3* and *cog*<sup>+</sup> in the *Neurospora* genome. The Southern blot strategy was similar to that outlined in Figure 4.13 with the untransformed strains predicted to produce ~3, ~2.1 & ~6.8kb bands upon cleavage with *Clal* whereas transplaced pNEIGFI-G1-DV2-LD1 was expected to produces band sizes of 6 & 6.8kb. Liquid cultures of the transformed homokaryons were prepared, harvested (2.2.1.1.3) and genomic DNA was extracted from the mycelium (2.2.1.3.3). Approximately 1µg of genomic DNA from each transformant was digested with *PvuII* (2.2.4.2) and analysed by Southern blot using biotin-labelled pDV2 vector as a probe (2.2.4.7). All 5 of the pNEIGFI-G1-DV2-LD1/t11630 and 4 of the pNEIGFI-G1-DV2-LD1/t11644 transformants had the vectors correctly transplaced between *his-3* and *cog*<sup>+</sup> (Figure 4.20B).





**Figure 4.20 PCR screen and Southern blot of putative pNEIGFI-G1-DV2-LD1 transformants in *Neurospora* t11630 and t11644 strains.**

(A) PCR screen of genomic DNA from putative pNEIGFI-G1-DV2-LD1 *Neurospora* transformants with *nIGF-I* specific primers. Positive candidates (indicated by bolded labels) contained an ~200bp PCR amplicon when resolved by agarose gel electrophoresis. (B) Southern blot of genomic DNA from homokaryotic putative pNEIGFI-G1-DV2-LD1 *Neurospora* transformants. Genomic DNA from homokaryotic putative transformants were restricted with *Cla*I and probed with a biotin-labelled pDV2 vector. Correct transplacements are indicated by bolded labels.

**4.3.4.3 Crossing of Neurospora strains bearing the IGF-I duplication vector**

Crosses were performed between 15-6-14 & 15-6-15 (pNEIGFI-G1-DV2-LD1/t11630) and 16-1-2 & 16-1-4 (pNEIGFI-G1-DV2-LD1/t11644). Ascospores were harvested after two and three weeks incubation as cross maturation time is known to affect the extent of RIP (Singer 1995). Initially, crosses were performed by co-inoculation of parent strains in test tubes (2.2.1.2.3.2), however, these were found to be relatively infertile. An alternative glass petri dish crossing method was employed whereby one strain was spread plated onto an agar plate and used as the protoperithecial (maternal) parent. After allowing the maternal parent to grow for 5 days, the second fertilising (paternal) parent was spread plated and incubated at 25°C (2.2.1.2.3.1). This crossing method produced large numbers of ascospores and was used for all subsequent crosses.

The following crosses were performed in duplicate with each transformant used as both the maternal and paternal parent: 15-6-14 x 16-1-2, 15-6-14 x 16-1-4, 15-6-15 x 16-1-2 & 15-6-15 x 16-1-4. Crosses using 15-6-x as the maternal parent were relatively barren, however, crosses using 16-1-x as the maternal parent and 15-6-x as the paternal parent produced an abundance of ascospores. After the crosses were incubated for two weeks, the petri dish lid was removed under aseptic conditions and ascospores washed off using sterile water and stored at 4°C. The lid was replaced with a new sterile lid and the process was repeated 1 week later.

**4.3.4.4 DNA sequencing of IGF-I in RIPed progeny from the cross of Neurospora strains bearing the IGF-I duplication vector**

The 2 and 3 week old ascospores obtained from the cross of 15-6-15 with 16-1-2 were germinated by heat-shock at 60°C for 45min in layer agar prior to pouring onto VmH agar plates. Single, well isolated colonies were picked to VmH agar slope tubes to facilitate production of macroconidia (2.2.1.1.5).

Genomic DNA was extracted from macroconidia of 5 progeny from each of the 2 and 3 week maturation periods (2.2.1.3.2) and the IGF-I duplication downstream of the *glaA* promoter was PCR amplified (2.2.4.6.4) using primers LD1&2-nIGF-F and LD1&2-nIGF-R

(2.1.8.2). The PCR amplicons were purified by spin column and sequenced with the LD1&2-nIGF-F primer used for PCR amplification (2.2.4.9).

4.3.4.4.1 Characterisation of nIGF-I RIP DNA sequences

As expected for linked duplications, all sequenced progeny were affected by RIP to varying degrees (Figure 4.21). The distribution of RIP mutations occurred at frequencies of 76% for CpA, 15% for CpT, 8% for CpG and 1% for CpC dinucleotides approximating the RIP frequencies typically reported in the literature (Cambareri *et al.*, 1991).

Progeny harvested after 2 weeks contained a total of 26 mutations with individual progeny containing from 1 to 9 RIP mutations with a median of 6 mutations in the 210bp *nIGF-I* coding sequence. Ascospores isolated after 3 weeks maturation contained a total of 65 mutations with individuals ranging from 7 to 19 RIP events each with a median of 13 mutations in the *nIGF-I* coding sequence. The sequence analysis indicated the 3 week maturation period resulted in significantly more RIP mutations than the 2 week period ( $\chi^2_{\text{Yates}} = 16.725$ , 1df,  $p < 0.0001$ ).

The combined data from the 2 and 3 week progeny had 17 mutations in the first 105bp of *nIGF-I* and 74 mutations in the following 105bp which demonstrated the distribution of RIP mutations was significantly skewed towards the 3' end of the *nIGF-I* coding sequence ( $\chi^2_{\text{Yates}} = 17.936$ , 1df,  $p = <0.0001$ ). RIP frequencies typically drop sharply towards the ends of duplicated sequences (Cambareri *et al.*, 1989); (Grayburn & Selker 1989); (Singer *et al.*, 1995); (Watters *et al.*, 1999), presumably due to detection of the end of the duplication by the RIP machinery. Thus, the paucity of mutations in the first half of the *nIGF-I* coding sequence was attributed to the physical positioning of the *nIGF-I* sequence with the 5' end of *nIGF-I* delineating one side of the duplicated sequences.



(B)

Progeny isolate#	Mutation type				Mutation in dinucleotide target sequence							
	Total	Missense	Silent	Nonsense	CpA		CpT		CpG		CpC	
					CS	NS	CS	NS	CS	NS	CS	NS
2weeks #1	6	5	1	0	1	4		1				
2weeks #2	8	8	0	0		6		1		1		
2weeks #3	9	8	1	0		8		1				
2weeks #4	1	1	0	0	1							
2weeks #5	2	2	0	0	1	1						
2week total	26	24	2	0	3	19	0	3	0	1	0	0
3weeks #1	13	12	1	0	2	8		2				1
3weeks #2	7	4	3	0	1	4		2				
3weeks #3	13	13	0	0	4	3	3	1	1	1		
3weeks #4	19	18	0	1	12	3	1		1	2		
3weeks #5	13	12	1	0		10		2		1		
3week total	65	59	5	1	19	28	4	7	2	4	0	1

**Figure 4.21** DNA sequencing of RIPed *nIGF-I* coding sequences in progeny from a cross of *Neurospora* strains bearing the pNEIGFI-G1-DV2-LD1 linked *nIGF-I* duplication vector.

*Neurospora* strains 15-6-15 and 16-1-2 contained the pNEIGFI-G1-DV2-LD1 linked *nIGF-I* duplication vector in the t11630 and t11644 strains, respectively. Ascospores from sexual crosses between strains 15-6-15 and 16-1-2 were harvested after 2 and 3 weeks maturation. The *nIGF-I* coding sequence downstream of the *glaA* promoter was PCR amplified from 5 individual progeny for each of the 2 and 3 week incubation periods and sequenced.

**(A)** The coding strand of the wild-type *nIGF-I* coding sequence is shown and the RIP alterations of the mutated progeny harvested after 2 and 3 weeks of cross maturation are depicted above and below the *nIGF-I* coding sequence, respectively. The mutations are colour coded to indicate the targeted dinucleotide sites: CpA are red, CpT are gold, CpG are blue and CpC are green. Missense mutations are bolded, silent mutations are underlined and nonsense mutations are italicised. Positions of identity between a mutated gene and the wild-type gene are indicated with dashes. C to T and G to A transitions result from RIP alterations in the coding and non-coding strands, respectively. **(B)** Summary of *nIGF-I* RIP mutations in the DNA coding sequences of progeny isolated from the sexual cross of *Neurospora* strains 15-6-15 and 16-1-2. CS = Coding strand and NS = Non-coding strand

RIP mutation ‘hotspots’ can form within a given locus suggesting there are factors affecting RIP in addition to the CpN context (Grayburn & Selker 1989) & (Rosa *et al.*, 2004). Identification of genuine RIP hotspots in the *nIGF-I* coding sequence was complicated by the skewed distribution of mutations resulting from the physical location of *nIGF-I* at the end of the duplicated sequences. As such, the first 105bp of the *nIGF-I* sequence was excluded from statistical analysis to identify potential RIP hotspots. There were a total of 22 CpA dinucleotides in both strands of the second half of the *nIGF-I* coding sequence. Chi-Square analysis indicated there was no statistically significant bias in the distribution of the 56 CpA-associated RIP events within this region ( $\chi^2 = 20.286$ , 21df,  $p = 0.0963$ ). However, this statistical analysis was performed with a relatively small number of sequences and RIP hotspots in the *nIGF-I* coding sequence might be evident in a larger data set.

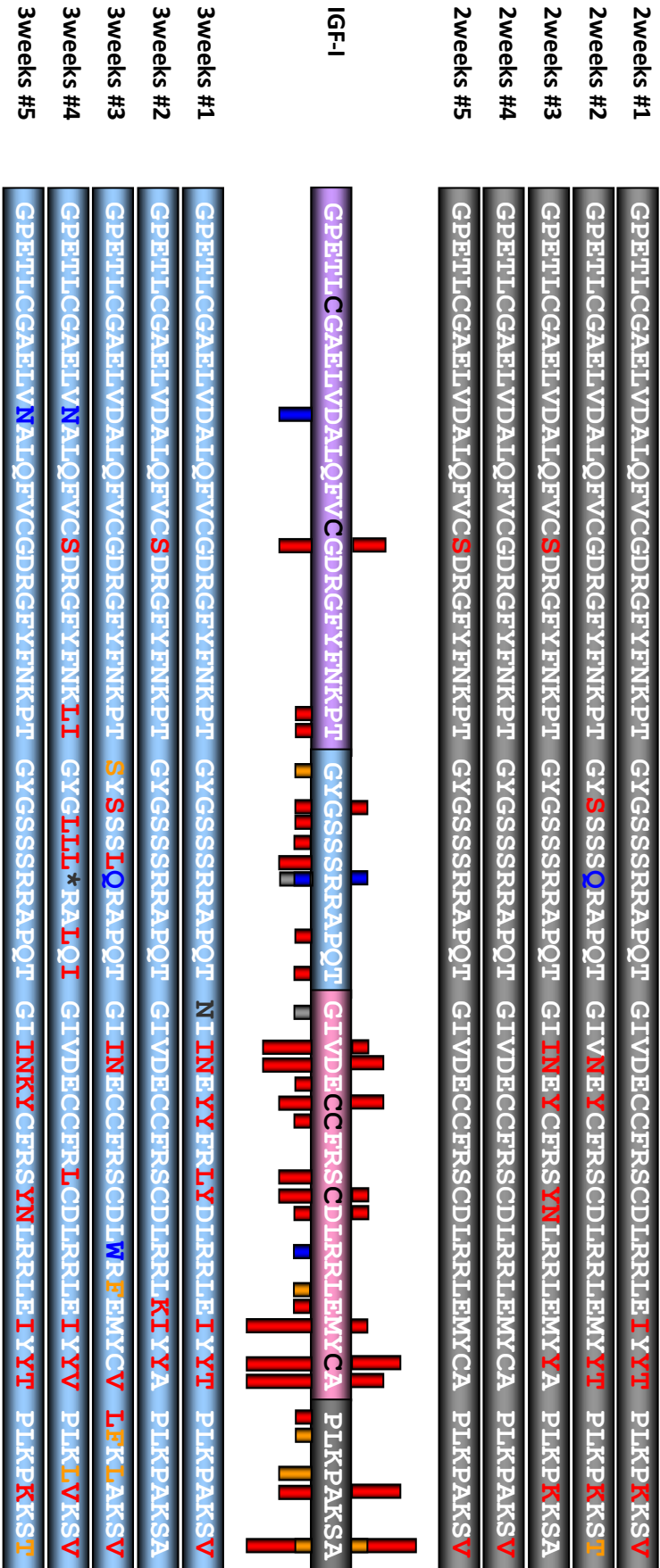
#### 4.3.4.4.2 Characterisation of translated IGF-I RIP sequences

The *nIGF-I* coding sequences in the *Neurospora* progeny were translated *in silico* to determine the potential amino acid substitutions resulting from RIP (Figure 4.22). IGF-I contains 70 amino acids of which a total of 32 were mutated by RIP in the 10 isolates sequenced. The B, C, A & D domain distribution of these 32 mutated residues consisted of 4, 8, 15 & 5 sites, respectively.

Amino acid substitutions can be classified as being deleterious, neutral or beneficial with respect to a protein’s function(s) and/or structure. Substitution scoring matrices such as BLOSUM62 can be used to determine similarities between amino acids and are typically used for the generation of protein sequence alignments (Henikoff & Henikoff 1992). However, substitution scoring matrices do not take into account the sequence context of an amino acid exchange and may provide inaccurate predictions of the effect of a substitution on protein structure and/or function (Ng & Henikoff 2001). The SIFT (Sorting Intolerant From Tolerant) program (<http://blocks.fhcrc.org/sift/SIFT.html>) is a sequence homology-based tool used to predict if an amino acid substitution will affect protein function. The premise of SIFT is that protein evolution is correlated with protein function and therefore, functionally important positions should be conserved

in a multiple alignment of the protein family, whereas diverse residues should be present in unimportant positions in an alignment. (Ng & Henikoff 2003). Amino acids not present at the substitution sites in the multiple alignment can still be predicted by SIFT to be tolerated if there are amino acids with similar physiochemical properties present in the alignment (Ng & Henikoff 2006). SIFT has been shown to greatly increase the prediction accuracy of a proteins tolerance to amino acid substitutions. For example, SIFT correctly predicted twice as many neutral substitutions compared to a substitution scoring matrix (BLOSUM62) even with only one homologous sequence aligned to a test protein (bacteriophage T4 lysozyme) (Ng & Henikoff 2001).

SIFT returns a score for the normalised probability of an amino acid change affecting protein function ranging from 0.00 to 1.00. Substitutions predicted to affect protein function are represented by scores  $\leq 0.05$ . The confidence in the prediction of a substitution's affect on protein function is dependent on sequence diversity within the protein sequence alignment. However, if the sequences within the alignment are closely related and/or are of a recently evolved protein then amino acids may be conserved due to limited exposure to selective pressures rather than being conserved by evolution. In these instances, SIFT may provide a high false positive error where neutral substitutions are predicted to affect protein structure and/or function based on limited sequence diversity at a given residue. To this end, SIFT calculates the median sequence conservation score (MSC score) representing the diversity of sequences at a given position within the alignment. The MSC score ranges from 0, when all 20 amino acids are present in the alignment at a single position to 4.32, when a position is completely conserved. Predictions based on sequence alignments with higher MSC scores (e.g.  $\geq 3.25$ ) are less diverse and typically have a higher false positive error (Ng & Henikoff 2003). However, the SIFT predictions for the IGF-I substitutions are likely to be robust as the IGF/Insulin axis has an ancient evolutionary origin with the primary structure of IGF-I being highly conserved among vertebrates (Chan 1993) (Reinecke & Collet 1998).





**Figure 4.22 Amino acid sequence translated from RIPed *nIGF-I* coding sequences in progeny obtained from crossing *Neurospora* strains bearing the pNEIGFI-G1-DV2-LD1 linked IGF-I duplication vector.**

*Neurospora* strains 15-6-15 and 16-1-2 contained the pNEIGFI-G1-DV2-LD1 linked *nIGF-I* duplication vector in the t11630 and t11644 strains, respectively. Ascospores from sexual crosses between strains 15-6-15 and 16-1-2 were harvested after 2 and 3 weeks maturation. The *nIGF-I* coding sequence downstream of the *glaA* promoter was PCR amplified from 5 individual progeny for each of the 2 and 3 week incubation periods and sequenced.

The RIPed *nIGF-I* coding sequences were translated into IGF-I amino acid sequences. Amino acid substitutions resulting from RIP targeting of the four dinucleotide target sites are colour coded: CpA are red, CpT are gold and CpG are blue (note: missense mutations arising solely from RIP of CpC dinucleotides did not occur in the isolates sampled). Mutations arising from RIP of multiple dinucleotide targets with a given codon are coloured grey. Protein truncation points resulting from nonsense mutations are represented by a \*, however, the downstream peptide sequence is still shown. The depiction of a mutation arising from RIP of a specific dinucleotide target is predicated on the absence of preceding RIP events altering adjacent residues.

The IGF-I bar is colour coded to delineate the BCAD domains. The cysteines involved in disulphide bond formation within the IGF-I protein are coloured black. The histogram either side of the IGF-I sequence represents the number of amino acid substitutions at a given amino acid position after 2 weeks (above the bar) and 3 weeks (below the bar) incubation. The bars are also colour coded for the specific dinucleotides targeted by RIP whilst the heights of the columns are proportional to their relative RIP frequency.

A BLASTp search of the mature human IGF-I protein sequence was performed against the SWISS-PROT 56.1 and TREMBL 39.1 databases. A ClustalW multiple alignment of 69 homologous sequences consisting of human IGF-I, 67 sequences from the BLASTp search and the IGF-I sequence of the grey kangaroo (Yandell *et al.*, 1998) which was not located in the protein databases (refer to Appendix Table A1) was performed. The IGF-I sequence alignment and the 35 amino acid substitutions in the RIPed progeny were analysed by SIFT to predict the probability of the individual mutations being neutral or affecting protein function. An arbitrary cut-off of  $\leq 0.05$  was used for a substitution to be considered to affect protein function and scores  $> 0.05$  were considered to be neutral substitutions.

Neutral mutations comprised 16 of the 90 total missense mutations in the IGF-I sequences. However, none of the 10 IGF-I variants contained solely neutral mutations. Although the effects of the amino acid substitutions were not experimentally determined, the majority of random mutations were expected to be deleterious (Arnold 2006). Of the 35 amino acid residues mutated by RIP, 8 were predicted by SIFT to be neutral and 27 were predicted to affect the IGF-I protein function and/or structure (Table 4.2). Cysteine residues are particularly vulnerable to RIP as cysteine is encoded by codons TGT and TGC, both of which contain a RIP-preferred CpA dinucleotide in the non-coding strand. One or more cysteine mutations occurred in 7 of the 10 IGF-I variants sequenced (Figure 4.23). As the tertiary structure of IGF-I is stabilised by three disulphide bonds (Cys<sup>6</sup>-Cys<sup>48</sup>; Cys<sup>18</sup>-Cys<sup>61</sup>; Cys<sup>47</sup>-Cys<sup>52</sup>), cysteine substitution would most likely disrupt the global IGF-I structure rendering these variants non-functional. One of the cysteine residue mutants (3weeks #4) also contained a non-sense mutation in a codon encoding arginine, which would have truncated the IGF-I protein to 35 amino acids.

## 4.4 Discussion

Directed evolution of human IGF-I in *Neurospora* had three practical requirements: first, variation needed to be introduced into the IGF-I DNA sequence; second, IGF-I needed to be expressed, processed and secreted; third, a rapid screen or selection for the desired function needed to be available with low inherent variability.

### 4.4.1 Diversifying the human IGF-I coding sequence in *Neurospora crassa*

The introduction of diversity into the IGF-I sequence was achieved through RIP. The IGF-I coding sequence was re-engineered to maximise the number of RIP preferred dinucleotides so as to enhance the distribution of mutation. Genomic integration of a linked IGF-I duplication vector facilitated high efficiency mutagenesis with all sequenced progeny having altered IGF-I sequences. However, there was a significant bias of mutations in the 3' end of the *nIGF-I* coding sequence, presumably due to construction of the duplication vector with the 5' end of the *IGF-I* coding sequence delineating one end of the duplications. Although inclusion of the 65bp *gla-1* signal peptide within the duplicated segment probably would have resulted in a more even distribution of mutations in the majority of the IGF-I coding sequence, the *gla-1* signal sequence was omitted from the duplication to reduce the likelihood of IGF-I secretion being adversely affected by RIP of the signal peptide. This was of particular importance as the DNA sequence encoding the *gla-1* signal peptide contained a total of 5 CpA dinucleotides with the potential for 1 nonsense and 2 missense RIP mutations including substitution of the methionine start codon for an isoleucine codon. Despite the substantial reduction in the mutation rate in the 5' end of the IGF-I sequence, 5 of the 10 progeny sequenced still contained at least one missense mutation in the first 65bp of the IGF-I coding sequence.

Residue position	Wildtype amino acid	Mutated amino acid	Function affected?		SIFT score	MCS score	Mutated residue present in progeny:												
			YES	NO			2w #1	2w #2	2w #3	2w #4	2w #5	3w #1	3w #2	3w #3	3w #4 <sup>+</sup>	3w #5			
12	D	N	✓		0.00	4.01													
19	G	S	✓		0.00	4.01							✓						
28	P	L	✓		0.00	4.01													
29	T	I		x	0.08	4.01													x
30	G	S	✓		0.00	4.01													✓
32	G	S	✓		0.01	4.01													✓
33	S	L		x	0.10	4.01												x	
34	S	L	✓		0.00	4.01							✓						✓
35	S	L	✓		0.01	4.01													✓
36	R	Q	✓		0.02	4.01							✓						✓
39	P	L	✓		0.05	3.97													✓
41	T	I	✓		0.01	4.01													✓
42	G	N	✓		0.00	4.01													✓
44	V	I	✓		0.01	4.01													✓
45	D	N		x	0.11	4.01												x	✓
46	E	K		x	0.13	3.12												x	✓
47	C	Y	✓		0.00	4.01													✓
48	C	Y	✓		0.00	4.01													✓
51	S	L	✓		0.00	4.01													✓
52	C	Y	✓		0.00	4.01													✓
53	D	N	✓		0.01	4.01													✓
55	R	W	✓		0.00	4.01													✓
57	L	F	✓		0.01	4.01													✓
58	E	K	✓		0.00	4.01													✓
59	M	I	✓		0.00	4.01													✓
61	C	Y	✓		0.00	4.01													✓
62	A	V	✓		0.00	4.01													✓
62	A	T	✓		0.00	4.01													✓
63	P	L	✓		0.00	4.03													✓
64	L	F		x	0.28	4.03													✓
66	P	L	✓		0.01	4.01													✓
67	A	K		x	0.09	4.01													✓
67	A	V		x	0.06	4.01													✓
70	A	T		x	0.37	4.01													✓
70	A	V	✓		0.03	4.01													✓

**Table 4.2 SIFT analysis of predicted amino acid substitutions resulting from RIP of the *nIGF-I* coding sequence in progeny obtained from crossing *Neurospora* strains bearing the pNEIGFI-G1-DV2-LD1 linked IGF-I duplication vector.**

The predicted amino acid substitutions in the RIPed *nIGF-I* coding sequences from progeny of cross 15-6-15 and 16-1-2 were analysed by the SIFT (Sorting Intolerant From Tolerant) program to identify mutations likely to affect protein function. The SIFT score represents the normalised probability of an amino acid affecting protein function and ranged from 0.00 to 1.00. A score  $\leq 0.05$  was considered to represent a substitution that probably affected the protein function. The median sequence conservation score (MSC score) reflected the diversity of the sequences in the alignment. The MCS score would be 0 when all 20 amino acids were observed at a single position in the sequence alignment and ranged to 4.32 if the position was completely conserved.

Residues predicted to affect protein function are indicated by a ☒ whereas tolerated substitutions are indicated by an x. Progeny containing mutations in the cysteine substitutions involved in IGF-I disulphide bond formation are shown in red as these mutants are highly unlikely to be functional due to disruption of the tertiary IGF-I structure. The residue positions and amino acids are colour coded to indicate the location of the residues within the BCAD domains of IGF-I. Note: Isolate 3w4<sup>+</sup> contained a nonsense mutation at residue 36 in the IGF-I protein sequence. The remainder of the translated sequence is included for comparative purposes only.

SIFT analysis of the IGF-I mutants indicated all 10 progeny sequenced contained 1 or more substitutions in the 27 'RIPable' residues predicted to affect IGF-I function. Most random amino acid substitutions are likely to be deleterious to a protein's function, primarily through destabilisation of the native protein structure (Bloom *et al.*, 2005). Indeed, the high probability of introducing deleterious mutations supports the putative role of RIP as an efficient genome defence mechanism. Clearly, the introduction of deleterious mutations into a peptide sequence is not desirable in terms of gene diversification. However, 8 of the 35 residues potentially mutated by RIP are predicted to be neutral substitutions with little or no significant effect on protein function. The neutral theory of molecular evolution predicts that a large proportion of amino acid substitutions in a population will be neutral, resulting from random drift rather than selection (Kimura 1983); (Kimura 1968b); (Kimura 1968a). Neutral mutations are considered important in terms of evolution and directed evolution as they may confer beneficial effects through synergistic interaction with other combinations of mutations even though they themselves have little or no phenotypic effect (Axe 2000).

Deleterious mutations are an unavoidable consequence of RIP and are likely to comprise a large percentage of the mutations created by RIP. For example, 7 of the 10 progeny from the linked duplication vector cross contained one or more substitutions in cysteine residues involved in disulphide bond formation required for stabilisation of the tertiary IGF-I structure. Therefore, obtaining progeny with a relatively small number of mutations is desirable as these are more likely to yield proteins containing only neutral and/or beneficial substitutions. The two week cross maturation period was shown to have a significantly lower mutation frequency relative to the three week incubation period demonstrating that the cross maturation period can be used as an effective means of modulating the extent of RIP. Singer (1995) demonstrated RIP inactivation of the *am* gene occurred in 2% of ascospores isolated 9-10 days after fertilisation compared to 28% in progeny harvested after 12-15 days suggesting incubation periods shorter than two weeks are likely to further decrease the frequency of RIP.

Meiotic recombination takes place after RIP in *Neurospora*'s sexual cycle. As such, recombination events occurring within the *nIGF-I* coding sequence may separate deleterious RIP mutations from those that are beneficial or neutral. Characterisation of the effects of recombination required sequencing of ascospores isolated from octads such that all the products of single meiosis events could be analysed. However, octad analysis was not performed due to the substantial input of labour and time required. Particularly as recombination within the IGF-I coding sequence was not the primary source of diversification and was expected to occur in an important but relatively small subset of progeny.

#### **4.4.2 IGF-I secretion in *Neurospora crassa***

Various filamentous fungi, such as *Aspergillus* and *Trichoderma* have been extensively used as hosts for heterologous protein production. These fungi have been the recipients of extensive research, particularly from the commercial sector, elucidating and manipulating the genetics and biochemistry of the host strains to increase heterologous protein production (Archer *et al.*, 1994).

*Neurospora crassa* has been utilised for heterologous protein expression (Nakano 1993); (Rasmussen-Wilson *et al.*, 1997) although not as widely as other 'mainstream' fungal production systems. As such, *Neurospora* has been the subject of considerably less research into improving the efficacy of heterologous protein secretion. However, *Neurospora* secretion of IGF-I protein was a necessary prerequisite for screening IGF-I variants as it was not practical to diversify the *nIGF-I* coding sequences, PCR amplify and subclone into an alternative expression system for high throughput screening.

Western blot characterisation of transformants with ectopically integrated pFENIGFI did not identify any transformants producing detectable levels of mature IGF-I. However, the presence of a putative uncleaved *gla-1:IGF-I* fusion band suggested the *gla-1:nIGF-I* coding sequence was being expressed and uncleaved fusion protein was being secreted. The absence of mature protein and the presence of uncleaved fusion protein presented numerous possibilities that could be further investigated to improve

IGF-I production: first, the media and incubation conditions may not have been optimal for IGF-I secretion. Second, engineering of the *nIGF-I* coding sequence to optimise for RIP bias rather than codon bias may have had a detrimental effect on translation efficiency. Third, the *gla-1*:IGF-I may have been incorrectly folded and/or processed in the protein trafficking pathway. Fourth, the secreted IGF-I protein may have been proteolytically degraded by extracellular proteases. Fifth, ectopic transformants contained expression constructs that may have integrated into suboptimal loci for secretion. Sixth, the *gla-1*:*nIGF-I* fusion vector configuration may not have been optimal for IGF-I secretion. Seventh, the globular IGF-I tertiary structure may have blocked the cleavage site from *Neurospora*'s Kex-2 like protease thereby inhibiting cleaving of the mature IGF-I protein from the *gla-1* fusion protein.

Optimising recombinant protein production is not a trivial task as there are many variables which are interdependent and can be both host strain and protein product-dependent (Hohenblum *et al.*, 2004). Although many of the potential causes for the inefficient IGF-I secretion were scrutinised, it was not feasible to address all possible areas. Key areas were selected in terms of those that could be realistically investigated within the project, were most likely to produce mature IGF-I protein and were best suited for IGF-I directed evolution.

#### **4.4.2.1 Optimisation of *Neurospora* growth conditions**

Initially, Vogel's medium was utilised for culturing *Neurospora* strains although this was subsequently found to inhibit the IGF-IR ligand binding assay. Further investigation revealed assay inhibition was primarily due to chelating agents within the medium. FRIES medium was shown to be less detrimental to the receptor assay and was used for the initial optimisation of incubation period using the pFENIGFI/t12001 ectopic transformant 01-2-2. Although mature protein was not detected by western blot, the presence of putative *gla-1*:IGF-I fusion protein suggested the 48 hour incubation period was probably best suited for IGF-I secretion at least when fused to the *gla-1* protein. Similarly, the 48 hour time period yielded the highest concentration of IGF-I as determined by IGF-IR ligand binding assay although it was not possible at



this stage to determine if these measurements reflected mature IGF-I protein or the putative gla-1:IGF-I fusion protein.

Subsequent western blot optimisation of media formulation and incubation period with t11630 & t11644 transformants comprising all 5 targeted IGF-I expression vectors indicated the 7-13-9 (pFENIGFI-GS-DV2/t11630) transformant produced mature IGF-I in the 48 hour incubation period in VmH, NSmH & FmHS but not FmH. The absence of mature IGF-I protein in transformant 8-17-1 (pFENIGFI-GS-DV2/t11644) indicated there were t11630 & t11644 strain specific differences which affected IGF-I yield. There is substantial polymorphism in the laboratory strains of *Neurospora crassa* as they were originally derived from four wild strains Abbott 4A, Abbot12a, Lindegren A and Lindegren  $\alpha$  (Yeadon & Catchside 1999). The variable IGF-I yield could have been due to numerous factors such as differences in protein folding, protein trafficking, secretion and/or extracellular protease profiles.

The variation in secreted IGF-I levels in the t11630 & t11644 host strains presented a potential complication for screening IGF-I variants in progeny supernatants as crossing of the t11630 and t11644 parent strains would result in polymorphic progeny as a consequence of chromosome random assortment and recombination during meiosis. Thus, it would be possible the IGF-I yield from these progeny could be highly variable.

Differences in measured IGF-IR binding affinity with the IGF-IR ligand binding assay is based on the presumption that IGF-I concentrations are equal in all measured samples. As such, variation in IGF-I levels between progeny would yield a large number of false positives and negatives. Variation of IGF-I yield in progeny could be reduced by the generation of near isogenic host strains that differed solely by mating type. Although it would be technically possible to create near isogenic strains through extensive backcrossing or substitution of the mating type genes using molecular techniques, both methods would be laborious and time consuming. Furthermore, the initial choice of host strain from which to create an isogenic counterpart ideally would have required substantial characterisation of numerous IGF-I transformed strains to identify those with the highest yields of mature IGF-I.

**4.4.2.2 IGF-I processing through the *Neurospora* protein trafficking pathway**

Eukaryotic secretion pathways are complicated multi-step processes involving the generation and transportation of cargo carriers between intracellular compartments and the fusion of these carriers with specific target membranes. Protein trafficking is connected with numerous other cellular processes such as cytoskeletal dynamics and quality control in the endoplasmic reticulum (Derby & Gleeson 2007). The gla-1:IGF-I fusion strategy was utilised in an attempt to assist endomembrane trafficking of IGF-I in *Neurospora*. Although improper folding and/or processing could have been responsible for the relatively low yield of IGF-I, thorough investigation of IGF-I processing in the *Neurospora* protein trafficking pathway was unlikely to have been a trivial exercise. As such, characterisation of intracellular IGF-I processing was considered to be a relatively low priority and was not pursued further. However, alternative strategies such as the removal of the gla-1 fusion protein from IGF-I and the use of alternative signal peptides were subsequently employed in this project in an attempt to overcome problems in the *Neurospora* secretory pathway should they exist (as discussed in 4.4.2.5).

A potential method for increasing production of some proteins is the upregulation of key components of the unfolded protein response (UPR) pathway. The UPR is an intracellular signalling pathway that responds to the accumulation of unfolded proteins in the ER by regulating the expression of genes such as ER-associated chaperones and foldases and various other genes involved in secretion (Travers *et al.*, 2000). Overexpression of the UPR-regulating *HAC1* transcription factor in *S. cerevisiae* has been demonstrated to improve heterologous production of  $\alpha$ -amylase (*Bacillus amyloliquefaciens*) (Valkonen *et al.*, 2003a). Similarly, the *hacA* transcription factor from *Aspergillus awamori* is homologous to *HAC1* in yeast and overexpression of activated *hacA* in *A. awamori* has been shown to increase the yield of heterologous laccase from the basidiomycete *Trametes versicolor* and bovine preprochymosin (Valkonen *et al.*, 2003b). Although a *HAC1/hacA* homologue has been identified in *Neurospora* (E. Cambareri: *Personal communication*), this protein's putative role in

regulating the UPR has yet to be characterised but could prove to be a useful tool in future for improving heterologous protein abundance in *Neurospora*.

#### **4.4.2.3 Proteolytic degradation of IGF-I**

Proteolytic degradation of secreted heterologous and homologous proteins is a common problem in filamentous fungi (van den Hombergh *et al.*, 1997). Western blot analysis of untransformed t12001 culture supernatants spiked with recombinant IGF-I at various time periods during a 48-hour incubation revealed a high degree of proteolysis occurred despite the presence of a broad ranging protease inhibitor cocktail.

*Neurospora* protease synthesis is known to be repressed at sucrose concentrations of 0.5 to 2% (w/v) (Drucker 1972). However, the final Vogel's and FRIES based media formulations used in this study contained 8% (w/v) sucrose which is the maximum concentration before fungal growth inhibition occurs (E. Cambareri: *Personal communication*) yet there was no discernable effect on mature IGF-I yield.

Strain improvement programs aimed at reducing proteolytic degradation of secreted proteins have traditionally consisted of random and/or targeted mutagenesis although the latter strategy has often been restricted by limited availability of sequence information (Gouka *et al.*, 1997a). The publication of the *Neurospora crassa* genome sequence (Galagan *et al.*, 2003) would enable a targeted but broad ranging approach to be taken to inactivate putative protease encoding genes. Borkovich *et al.* (2004) performed an *in silico* analysis of the *Neurospora* genome sequence and predicted a total of approximately 55 proteases. Of these, 10 aspartyl proteases, 1 serine protease and 1 metallo-protease were postulated to be extracellular based on the presence of signal sequences necessary for secretion. These 12 putative extracellular proteases are ideal candidates for an extensive *Neurospora* strain protease inactivation strategy.

Development of protease deficient strains could be achieved in a number of ways. Targeted gene inactivation using gene replacement can be done but the homologous

integration frequency in *Neurospora* is relatively low (typically 10% to 30%) as DNA double stranded breaks are preferentially repaired independently of sequence homology through a non-homologous end joining (NHEJ) process. Recent work has shown that *mus-51* and *mus-52* *Neurospora* mutants have suppressed NHEJ activity with the frequency of homologous recombination increasing to ~100% when homologous sequences of 1000bp or greater are present on the 5' and 3' flanks of a selectable marker gene (Ninomiya *et al.*, 2004). The homology flanked selectable marker sequences could be used to disrupt any given gene and could be generated relatively easily through fusion PCR. However, the number of protease genes that can be inactivated using this approach is effectively limited to the number of unique selectable marker genes available.

Alternatively, RIP could be used to inactivate all 12 putative extracellular protease genes using a single ectopically integrated vector containing 500-700bp of duplicated sequence from each of the putative proteases. This approach would require extensive strain crossing and DNA sequencing of progeny to ensure extensive mutation in the target protease sequences and functional assay screening to verify reduced protease activities in the *Neurospora* strain.

Inactivation of the 12 putative extracellular proteases would be expected to substantially improve the efficacy of heterologous protein secretion in *Neurospora*. However, developing and characterising improved expression strains would be a labourious and time-consuming endeavour. Furthermore, Neugenesis Corporation was in the process of generating protease deficient *Neurospora* strains. Thus, to avoid duplication of research effort, no strain improvement work was conducted in this project.

**4.4.2.4 Neurospora codon bias**

Translational inefficiencies associated with codon bias could have been potentially responsible for the relatively low levels of IGF-I protein observed. Construction of an additional *IGF-I* CDS to account for codon bias was considered but not undertaken given the significant reduction of RIP preferred CpA non-silent target sites from 35 to 8 limiting the potential for generating sequence diversity.

The human *IGF-I* CDS contained 17 non-silent CpA RIP target sites and a greater number of *Neurospora* preferred codons compared to the RIP-optimised CDS. A modest reduction in the number of rare codons within a gene can increase the likelihood of obtaining higher protein yields (Kane 1995); (Gustafsson *et al.*, 2004). For example, heterologous production of human RAP74 in *E. coli* was improved 10-fold by replacement of a subsection of human *RAP74* cDNA with *E. coli* optimised codons (Wang *et al.*, 1994). As such, the human IGF-I CDS represented a possible compromise between overcoming *Neurospora* codon bias whilst retaining versatility as a template for RIP mutagenesis. Accordingly, the human *IGF-I* CDS was chosen as an alternative *IGF-I* coding sequence for diversification.

**4.4.2.5 IGF-I expression vector configurations**

Five IGF-I expression vectors were constructed in an effort to obtain a reasonable yield of secreted mature IGF-I protein. To enable a direct comparison between the different constructs and to eliminate variation in expression associated with ectopic integration, all 5 vectors were cloned into the pDV2 vector and transplaced between *his-3* and *cog*<sup>+</sup> in *LG I*. All the vectors contained the constitutive *glaA* promoter from *A. niger* as this had been previously shown to result in strong expression and had been successfully used for heterologous protein production in *Neurospora* (E. Cambareri: *Personal communication*). Despite being under control of the same promoter, sequence differences between the coding regions could have affected mRNA synthesis and/or stability (Gouka *et al.*, 1996) which may have had a subsequent affect on IGF-I levels. Ideally, characterisation of mRNA levels of the different *IGF-I* variants by real time PCR would have been undertaken. However, by this stage of the project there was little

practical benefit in undertaking expression analysis as the research priorities had shifted away from IGF-I directed evolution in *Neurospora* towards a rational IGF-II mutagenesis strategy.

#### 4.4.2.5.1 Fusion vector strategies

The pFENIGFI-DV2 construct encoded the fusion of a truncated glucoamylase to the N-terminus of the IGF-I protein. The basis of this fusion strategy was to attempt to exploit the normally highly expressed and abundantly secreted endogenous gla-1 protein to assist transcription and trafficking of IGF-I through the secretory pathway. Although, putative gla-1:IGF-I fusion protein was produced in transformants of *Neurospora* strains t11630 and t11644, no mature IGF-I protein was detected by western blot. The more sensitive IGF-IR ligand binding assay indicated IGF-I was being produced by these transformants but it was probable that the assay was measuring the uncleaved gla-1:IGF-I protein in addition to any mature IGF-I that may have been produced (discussed in 4.4.3).

The presence of uncleaved gla-1:IGF-I protein suggested the Kex2 cleavage site in the linker bridging gla-1 and IGF-I was inaccessible to the Kex2-like protease. pFENIGFI-GS-DV2 was generated by the insertion of a flexible penta-glycine spacer between the N-terminus of IGF-I and the Kex2 cleavage site to facilitate unhindered access to the protease target site. Western blot analysis of culture media from transformant 7-13-9 (pFENIGFI-GS-DV2/t11630) indicated mature IGF-I protein was present in the supernatants albeit at relatively low levels (<43ng/ml). However, the t11644 counterpart (8-17-1) did not have a detectable mature IGF-I band. Comparison of the putative gla-1:IGF-I fusion bands of 5-1-1 (pFENIGFI-DV2/t11630) & 7-13-9 (pFENIGFI-GS-DV2/t11630) in the 48hr FmHS western blot indicated the introduction of the glycine spacer had substantially increased cleavage at the Kex2 cleavage site. However, the density of the mature IGF-I band was not proportional to the difference in fusion band densities between 5-1-1 & 7-13-9, presumably due to proteolytic degradation of the liberated mature IGF-I protein. Although the pFENIGFI-GS-DV2 construct was a substantial improvement over pFENIGFI-DV2, the presence of residual

uncleaved fusion protein still potentially confounded the use of the IGF-IR ligand binding assay.

4.4.2.5.2 Non-fusion vector strategies

Fusing heterologous proteins to a truncated glucoamylase protein is often used to increase secretion yield, however, the effectiveness of this strategy is dependent on the specific heterologous protein. Consequently, it was desirable to ascertain if a non-fusion strategy would be more suitable for IGF-I secretion. The pNEIGFI-G1-DV2 vector was constructed to address this issue by fusing the *gla-1* signal peptide coding sequence directly to the *nIGF-I* coding sequence. However, western blot and IGF-IR ligand binding assays did not detect any IGF-I protein in the supernatants of these transformants.

The pNEhIGFI-C1-DV2 plasmid contained the human IGF-I coding sequence downstream of the *Ce/12A* signal sequence whereas the pNEhIGFI-H1-DV2 vector had the human mature *IGF-I* coding sequence and the human *IGF-I* signal. However, western blot and IGF-IR ligand binding assays did not detect any IGF-I protein in the supernatants of transformants with either construct.

The results obtained with the different vector configurations indicated the glucoamylase fusion strategy increased IGF-I yield. It was clear proteolytic degradation of IGF-I was substantially reducing the yield of mature IGF-I in the culture supernatant. It is possible there were additional factors limiting the production of mature IGF-I. However, identifying these problematic areas would have required potentially extensive characterisation of transcription, translation, protein folding, intracellular trafficking and secretion in the *Neurospora* transformants and was beyond the scope of this project.

### **4.4.3 Screening of IGF-I variants in *Neurospora* culture supernatants**

Successful engineering of a protein variant with desirable properties through directed evolution occurs relatively rarely and requires a diverse mutant library and an efficient screening assay that is robust, sensitive, adaptable for high throughput and has low inherent variability (Boder & Wittrup 2000) (Arnold *et al.*, 2001). The suitability of the IGF-IR ligand binding for screening IGF-I variants secreted by *Neurospora* into the culture supernatant was assessed. Although the IGF-IR ligand binding assay involved numerous steps, the standard 96-well plate format, 2.5 day turn-around time for a complete assay and the potential adaptability for robotic handling suggested the assay would be capable of reasonable sample throughput.

#### **4.4.3.1 *Neurospora media incompatibility with the IGF-IR ligand binding assay***

Initial experiments indicated the ligand binding assay was completely inactivated by chelating agents present in Vm. Optimisation with various media formulations demonstrated the assay was functional when used with FmH albeit with a 3-fold and 80-fold reduction in dynamic range and minimum sensitivity, respectively. Untransformed culture supernatants produced a small degree of background fluorescence that resulted from *Neurospora* secreted molecules that most likely affected the Eu-IGFI chelate, degraded or potentially even bound to the IGF-IR which further diminished sensitivity and accuracy.

#### **4.4.3.2 IGF-IR ligand binding assay quantification of *Neurospora* IGF-I**

The IGF-IR ligand binding assay was used to estimate the concentration of IGF-I in the *Neurospora* culture supernatants although this was predicated on the assumption the *Neurospora* secreted IGF-I was correctly folded and had no aberrant post-translational modifications and therefore had a binding affinity equivalent to recombinant IGF-I used as a standard. Should this assumption be incorrect and the *Neurospora* IGF-I had greater or lesser affinity for the IGF-IR relative to the IGF-I standard then any quantitation would be an overestimate or underestimate, respectively. Determining



the binding affinity of *Neurospora* IGF-I would first require purification, quantitation and characterisation of the IGF-I protein which was not attempted due to the relatively low levels of mature IGF-I obtained.

The IGF-IR ligand binding assay analysis of culture supernatant from *Neurospora* transformants comprising all five IGF-I expression vectors in the t11630 and t11644 strains indicated the pFENIGFI-DV2 and pFENIGFI-GS-DV2 transformants had produced detectable quantities of IGF-I. However, these transformants also produced putative uncleaved gla-1:IGF-I fusion bands. The pFENIGFI-DV2 transformants had higher IGF-I assay measurements than their pFENIGFI-GS-DV2 counterparts. However, only the pFENIGFI-GS-DV2/t11630 transformant (7-13-9) produced a detectable putative mature IGF-I band on a western blot, suggesting the gla-1:IGF-I fusion band was capable of binding to the IGF-IR. The presence of uncleaved fusion protein could confound the interpretation of IGF-I variant binding affinities and potentially lead to a high rate of false positives and negatives.

The IGF-IR ligand binding assay was not primarily intended for absolute measurements of IGF-I concentration but rather for comparative screening of IGF-I mutants. A meaningful comparison of the binding affinities of IGF-I variants in *Neurospora* culture supernatants would require equal concentrations of IGF-I in all the samples assayed. However, *Neurospora* cultures exhibited substantial variation in total protein concentrations despite rigorous attempts at standardisation of inoculum concentrations and incubation conditions. Although quantitation of total protein in a small number of samples is a relatively simple procedure using techniques such as Bradford or BCA assays, the large number of samples that would need to be assayed and standardised prior to high-throughput screening would make this additional step very laborious and time consuming.

**4.4.3.3 Variation of IGF-I in Neurospora supernatants and the IGF-IR ligand binding assay**

The IGF-IR ligand binding assay of five replicate 7-13-9 (pFENIGFI-GS-DV2/t11630) culture supernatants indicated there was a high degree of variation between the five samples on the same plate (up to ~39%CV) and the same sample on two separate plates (up to ~35%CV). This level of sample and/or assay variation is not ideal for directed evolution as small changes in binding affinities could be missed with false negatives whilst valuable time and resources would need to be spent verifying IGF-I variants identified in a screen were not false positives.

**4.4.3.4 Surface display as an alternative screening assay?**

Surface display could potentially be used as an alternative method of screening for *Neurospora* generated IGF-I variants. Surface display of proteins is achieved by the fusion of a library of coding sequences of interest to an endogenous gene encoding a surface protein resulting in the anchoring of the heterologous-surface fusion protein to the surface of the host organism. Screening is then used to identify protein variants with desired characteristics. Surface display has been successfully utilised with various micro-organisms such as bacteriophage (Smith 1985), bacteria (Francisco *et al.*, 1992) (Francisco *et al.*, 1993) and yeast (Schreuder *et al.*, 1993) for numerous applications including affinity maturation of receptor-ligand interactions (Levin & Weiss 2006). It might be possible to develop a system whereby IGF-I variants generated by RIP are fused to proteins on the surface of *Neurospora* macroconidia which could then be screened by fluorescent-activated cell sorting (FACS) enabling quantitative screening of ligand binding kinetics and equilibria. In addition, FACS provides a quantitative measure of cell surface expression and can sort large numbers of cells quickly (Wittrup 2001). Screening under these conditions can facilitate high-throughput screening with the ability to resolve small improvements in mutants with high statistical certainty (Boder & Wittrup 2000).

**4.4.4 Diversification of human IGF-I in *Neurospora crassa*: summary and conclusion**

The *in vivo* diversification of human IGF-I in *Neurospora* required the ability to create a library of diverse variants, the ability to secrete IGF-I into the culture medium and the use of an effective screen to identify mutants with altered IGF-IR binding characteristics.

The use of RIP in conjunction with a RIP-optimised coding sequence was shown to be an efficient means of generating IGF-I mutants although *in-silico* analysis suggested reducing the extent of RIP by harvesting crosses after shorter incubation periods would be more likely to produce mutants with desirable IGF-IR binding characteristics.

Mature IGF-I was only obtained with the pFENIGFI-GS-DV2 construct albeit with a very low yield which was due, at least in part, to the presence of extracellular proteases. The quantity of mature IGF-I secreted into the culture media was not sufficient for purification and characterisation of the *Neurospora*-produced IGF-I protein. An unforeseen complication with the fusion protein strategy was the presence of residual uncleaved fusion protein that appeared to be capable of binding to the IGF-IR thereby potentially confounding the IGF-IR ligand binding assay. The secretion levels of the *Neurospora* cultures varied significantly despite attempts at standardising the culturing conditions and would likely complicate downstream screening of IGF-I variants in the culture supernatant.

The IGF-IR ligand binding assay was not ideally suited for screening IGF-I in *Neurospora* culture supernatants due to incompatibilities with the *Neurospora* liquid culture media, background signal from *Neurospora*-produced proteins and a relatively high degree of inter- and intra-plate variability.

CHAPTER 4 – Transformation, expression and diversification of IGF-I in *Neurospora*

This study successfully secreted human IGF-I and generated IGF-I DNA variants in *Neurospora* thereby demonstrating potential of the *Neurospora* diversification system for the directed evolution of IGF-I. However, difficulties with proteolytic degradation and screening of secreted protein reduced the probability of identifying IGF-I antagonists within the duration of this study. As the proteolysis and screening issues were being addressed by Neugenesis Corporation, the direction of this project was altered so as to avoid duplication of research effort. Instead, the project focus was diverted to the characterisation of IGF residues involved in receptor binding and activation through rational mutagenesis.

## **Chapter 5 Generation and characterisation of IGF-II analogues**

### **5.1 Introduction**

The IGF-I and IGF-II ligands share a common BCAD domain structure and exhibit differential receptor binding and activation characteristics which are partly modulated by the IGF C-domain. Elucidation of the mechanism by which the IGFs differentially bind and activate the same receptor would make a significant contribution to our fundamental understanding of IGF-receptor interactions and could assist in the rational design of an IR and/or IGF-IR antagonist as a potential cancer therapeutic. The work described in this chapter involved the generation and purification of three IGF-II analogues designed to delineate the structural elements within the IGF C-domain responsible for high affinity receptor binding and activation. The IGF-II analogues were characterized by competitive receptor binding assays, receptor phosphorylation assays and cell survival assays using the IR-A, IR-B and IGF-IR receptors.

The IGFs and insulin bind with high affinity to their cognate receptors and with lower affinity to their noncognate receptors. Interestingly, IGF-II binds with high affinity to both the IR-A and IGF-IR (Frasca *et al.*, 1999); (Denley *et al.*, 2004), with both receptors mediating IGF-II bioactivity (Krywicky & Yee 1992). There is an *~2-fold* increase in insulin binding affinity to the IR-B relative to the IR-A, whilst IGF-II and IGF-I have *~3-4-fold* lower affinities for the IR-B compared to the IR-A (Denley *et al.*, 2004). The mechanism for the differential binding of ligands to the two insulin receptor isoforms has yet to be elucidated.

The two binding surfaces and mechanism of ligand-receptor interaction between IGF-I and the IGF-IR are comparable to the insulin and IR interaction (Adams *et al.*, 2000). Although high affinity binding of IGF-I to the IGF-IR also requires interaction of the IGF-I C-domain with module 6 of the cysteine-rich domain (Hoyne *et al.*, 2000); (Keyhanfar *et al.*, 2007).

The molecular interactions between the IGF-II ligand and the IR & IGF-IR are considerably less well understood relative to our understanding of how the insulin and IGF-I ligands interact with their cognate receptors. Site-directed mutagenesis studies have identified two receptor binding sites in IGF-II composed of amino acids in the B and A-domains (Site 1: Val14, possibly Asp15, Gln18, Phe26, Tyr27, Phe28 and Val43; Site 2: Glu12, Phe19, Leu53 and Glu57) analogous to the binding sites in IGF-I and insulin although there are subtle differences in the contribution of particular residues (Sakano *et al.*, 1991); (Alvino *et al.*, 2009). Alanine-scanning mutagenesis studies have shown that IGF-II also binds to elements in the L1 domain and the CT peptide of the IGF-IR but unlike IGF-I, IGF-II does not interact with elements in the CR domain suggesting IGF-I and IGF-II utilize different mechanisms to bind to the same receptor binding site (Sorensen *et al.*, 2004).

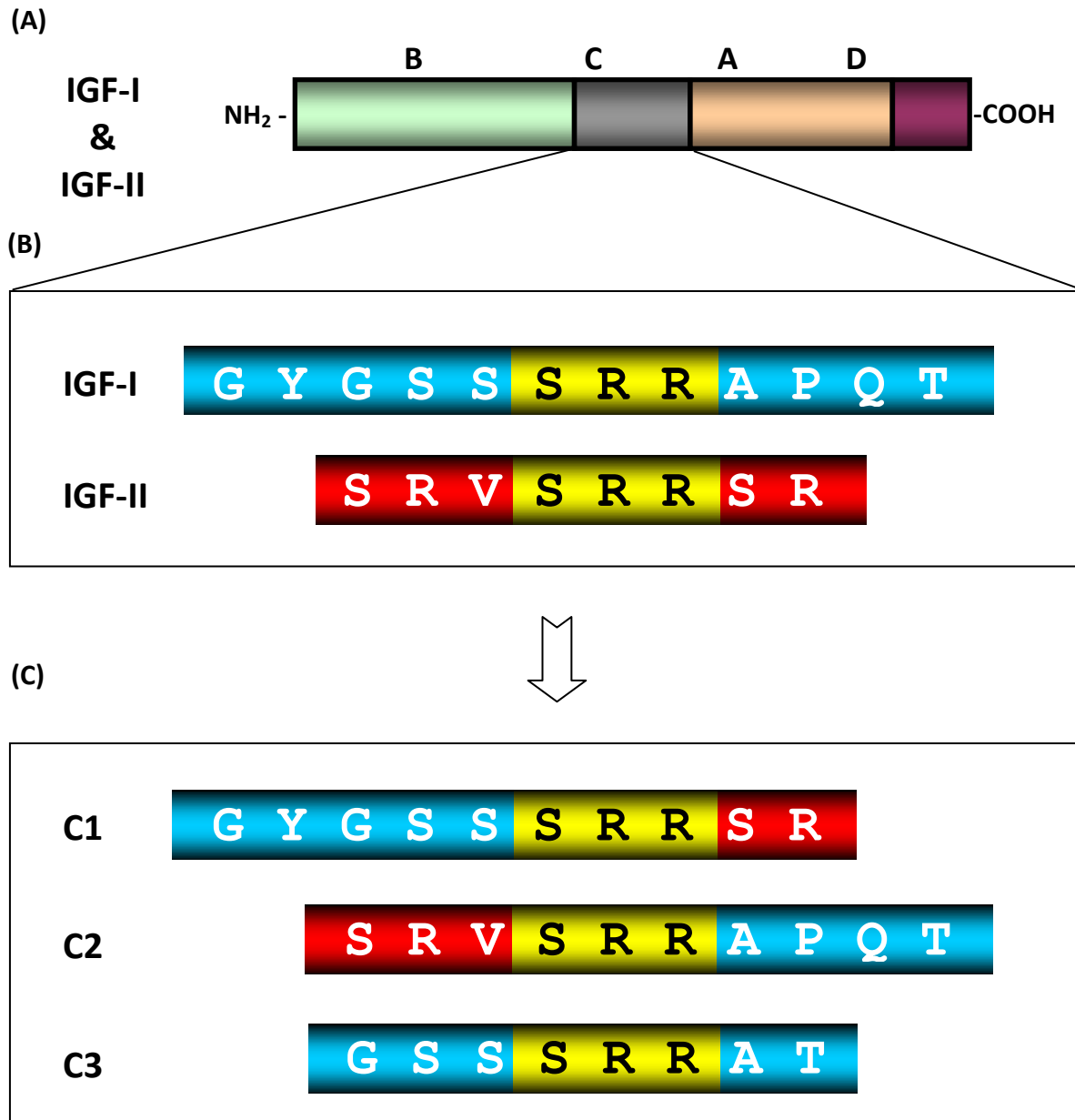
The majority of free energy of IGF-I and IGF-II binding is due to the two binding sites located primarily in the B and A domains (Alvino *et al.*, 2009). However, IGF-II and IGF-I chimeras with reciprocally swapped C-domains and/or D-domains exhibited receptor binding affinities similar to those of the donor molecules thereby demonstrating the IR and IGF-IR binding specificity of the IGFs is primarily determined by the C-domains and to a lesser extent, the D-domains of the IGFs (Denley *et al.*, 2004). Furthermore, an IGF-I chimera with the IGF-II C-domain induced autophosphorylation of specific tyrosine residues in the kinase domain and juxtamembrane region of both IR-A & IR-B and activated downstream IRS-1 and akt/PKB signalling pathways to a similar extent as IGF-II (Denley *et al.*, 2006). These studies demonstrated the importance of the IGF C-domain in modulating IGF-receptor interactions.

The mechanism by which the IGF-II C-domain determines receptor binding specificity and activation is unknown. The IGF-II C-domain consists of 8 amino acids whilst the IGF-I C-domain has 12 amino acids. Thus, it is possible the smaller IGF-II C-domain permits greater steric accessibility to the ligand binding pocket of the IR. However, a study by (Gauguin *et al.*, 2008a) showed that an IGF-I analogue with four insulin substitutions in the B and A-domains had insulin-like binding affinity to the IR-A

suggesting the IR-A ligand binding pocket is capable of accepting the bulkier IGF-I C-domain. This assertion is supported by the finding that a single chain insulin analogue with the IGF-I C-domain was able to bind to the IR with a similar affinity to insulin (Kristensen *et al.*, 1995). Alternatively, the specificity conferred by the IGF C-domains may be determined by ligand-receptor amino acid side chain interactions such as hydrogen bonding, hydrophobic, electrostatic and Van der Waals forces.

Identifying the IGF C-domain residues involved in differential receptor binding and activation could provide valuable insights into how IGF-II and IGF-I exert their biological effects. Elucidating the IGF-II mechanism is of particular interest as the biological effects of this ligand are mediated through the IR-A and IGF-IR and both receptors are implicated in cancer (Belfiore *et al.*, 2009). As such, the work described herein aimed to delineate the residues within the C-domain of IGF-II which confer high affinity binding and potent activation of the IR isoforms. In addition, this study aimed to identify C-domain residues within an IGF-II chimera containing the IGF-I C-domain (Denley *et al.*, 2004) (henceforth referred to as BCIAD) conferring IGF-I-like binding and activation of the IGF-IR. To this end, three IGF-II chimeras (C1, C2 & C3) were constructed with chimeric IGF-II/IGF-I C-domains or a shortened IGF-I C-domain (Figure 5.1). Specifically:

- C1 substituted IGF-II residues Ser33, Arg34 & Val35 with IGF-I residues Gly30, Tyr31, Gly32, Ser33 & Ser34.
- C2 substituted IGF-II residues Ser39 & Arg40 with IGF-I residues Ala38, Pro39, Gln40 & Thr41.
- C3 substituted IGF-II residues Ser33, Arg34, Val35, Ser39 & Arg40 with IGF-I residues Gly32, Ser33, Ser34, Ala38 & Thr41.



**Figure 5.1** Amino acid sequences of the C-domains of IGF-I, IGF-II and IGF-II analogues C1, C2 & C3.

**(A)** Domain structure of mature IGF-I and IGF-II protein. **(B)** Amino acid sequences of IGF-I and IGF-II C-domains. The yellow region highlights the conserved residues. **(C)** Amino acid sequences of C1, C2 & C3. IGF-I residues are highlighted in blue, IGF-II residues in red and residues conserved in both IGF-II and IGF-I in yellow.



All three analogues included the C-domain residues Ser36/35, Arg37/36 & Arg39/37 (numbering refers to IGF-II/IGF-I positions, respectively) conserved in both IGF-II and IGF-I. The design rationale for C1 and C2 was to investigate the role of C-domain residues flanking the conserved SRR region in IGF-II by individually substituting each IGF-II flank with equivalent IGF-I residues. The C-domains of the C1 and C2 chimeras each consisted of 10 amino acids. C3 was designed to contain a shortened IGF-I C-domain so as to be a similar size to the IGF-II C-domain whilst retaining the charge characteristics of the IGF-I C-domain.

This chapter describes vector construction, site-directed mutagenesis, heterologous expression and purification of the IGF-II analogues. In addition, the interaction of the IGF-II mutants with the IR-A, IR-B and IGF-IR were characterized through receptor binding, receptor phosphorylation and cell survival assays using IGF-II, IGF-I, insulin and BCIAD as control ligands.

## **5.2 Construction of IGF-II analogues**

### **5.2.1 Cloning of *E. coli* codon optimised IGF-II coding sequence**

An *E. coli* codon optimised human IGF-II coding sequence was used as a base template for IGF-II analogue construction to facilitate optimal protein yield during heterologous expression of the analogues in *E. coli* (Denley *et al.*, 2004). The *E. coli* optimised IGF-II coding sequence preceded by coding sequence for the PAPM alpha lytic protease recognition site was PCR amplified from a plasmid kindly supplied by Dr Adam Denley (School of Molecular & Biomedical Science, Adelaide University) using primers EIGFII-F and EIGFII-R (Table 2.1.8.4) and ligated into the Promega pGEM-T easy vector (Table 2.1.4) to form IGFII/pGEM. Putative IGFII/pGEM transformants were screened by colony PCR (2.2.4.6.3) using primers EIGFII-F and EIGFII-R. Plasmid DNA from putative transformants was screened by *EcoRI* digestion prior to DNA sequencing with SP6 and T7 primers (Table 2.1.8.2) to confirm IGF-II coding sequence fidelity.

### **5.2.2 DpnI site-directed mutagenesis**

Three IGF-II chimeras were created using *DpnI* site-directed mutagenesis (2.2.4.8). Oligonucleotide pairs C1-F/C1-R, C2-F/C2-R & C3-F/C3-R (Table 2.1.8.4) were designed to introduce the desired C-domain sequence changes for C1, C2 & C3, respectively into the *E. coli* codon optimised IGF-II coding sequence. The primers were designed to introduce a silent mutation at position +139 of the PAMP-IGFII coding sequence to create an *NgoMIV* restriction site for screening putative transformants with mutated IGF-II coding sequences. The oligonucleotides were combined with IGFII/pGEM plasmid DNA in a thermal cycling reaction using *Pfx* high fidelity DNA polymerase to facilitate linear amplification of plasmid DNA with nicked circular strands containing the C-domain mutations (Figure 5.2A). The methylated plasmid DNA was then digested using *DpnI* restriction enzyme to leave intact nicked circular DNA (Figure 5.2B), which was transformed into *E. coli*. Plasmid DNA from putative transformants was screened by *NgoMIV* digestion and gel electrophoresis (Figure 5.2C) and the integrity of the desired mutations and the remaining coding sequence were confirmed by DNA sequencing with SP6 and T7 primers. The verified C1, C2 & C3 constructs were referred to as C1-pGEM, C2-pGEM & C3-pGEM, respectively.

### **5.2.3 Cloning into pGH(1-11)VNPAPM vector**

The C1, C2 and C3 coding sequences were cloned into the pGH(1-11)VNPAPM expression vector (kindly supplied by Prof. John Wallace, School of Molecular & Biomedical Science, Adelaide University) (as described in section 5.3). The mutant IGF-II coding sequences were cleaved from the C1-pGEM, C2-pGEM & C3-pGEM constructs, the pGH(1-11)VNPAPM vector was linearised by double digestion with *HpaI* and *HindIII* and the desired DNA fragments were gel purified. The C1, C2 & C3 coding sequences were individually ligated into the pGH(1-11)VNPAPM expression vector and transformed into *E. coli*. Putative transformants were screened by colony PCR with EIGFII-F and EIGFII-R primers and digestion of the PCR products with *NgoMIV*. The integrity of the leader and mutant IGF-II coding sequences were determined by DNA sequencing and the C1, C2 & C3 expression constructs were referred to as pGH(1-11)VNPAPM-C1, -C2 & -C3, respectively (Figure 5.3).

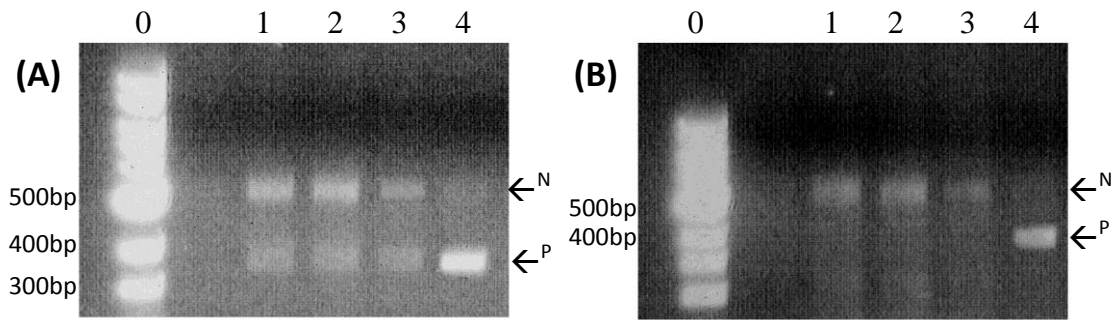
### 5.3 Heterologous expression of IGF-II protein in *Escherichia coli*

The pGH(1-11) expression vector facilitates efficient heterologous expression of wildtype and mutant IGF-I & IGF-II proteins in *E. coli* (King *et al.*, 1992) (Francis *et al.*, 1992) (Lien *et al.*, 2001) (Denley *et al.*, 2004). The C1, C2 & C3 coding sequences were fused in-frame to the coding sequence for the first eleven amino acids of porcine growth hormone and a PAPM linker sequence. The addition of the hydrophobic N-terminal porcine sequence promotes correct folding and improves IGF solubility (Francis *et al.*, 1992). The PAPM leader sequence is recognized by the  $\alpha$ -lytic protease mutant Prag A9 which cleaves the leader sequences between M<sup>-1</sup> and A<sup>+1</sup> to produce mature IGF-II protein (Lien *et al.*, 2001). The IPTG-inducible promoter *trc* mediates inducible IGF-II expression and the  $\beta$ -lactamase gene facilitates ampicillin selection.

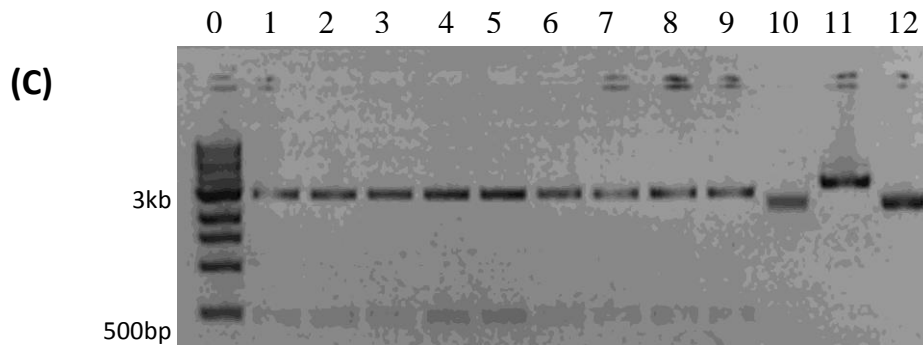
Heterologous expression of recombinant proteins in *E. coli* often results in cytoplasmic aggregation of high levels of intermediately folded recombinant protein in the form of inclusion bodies. Solubilisation of inclusion bodies and refolding protein to obtain native tertiary structure and regain full biological activity can often prove challenging (Burgess 2009). A method for refolding IGF protein from inclusion bodies has been developed and routinely used to obtain high protein yields (King *et al.*, 1992); (Francis *et al.*, 1993); (Delaine *et al.*, 2007). Thus, heterologous expression of the IGF-II analogues required isolation, purification and solubilisation of inclusion bodies, refolding protein, cleavage of leader sequences and purification of mature protein.

#### 5.3.1 *E. coli* transformation and induction

*E. coli* JM101 cells transformed with the pGH(1-11)VNPAPM-C1, -C2 & -C3 vectors were used to inoculate 4x 500ml of LB/amp medium in 2L flasks (2.2.5.2.1). Expression was induced by the addition of 100 $\mu$ M IPTG and the cultures were incubated for a further 5hrs with vigorous shaking. The *E. coli* cells were pelleted by centrifugation at 7,000g for 15min at 4°C and stored at -80°C. Pre and post-induction samples of the ferments were analysed by SDS-PAGE (2.2.5.5.1) and the presence of the inclusion bodies in the post-induction samples indicated expression was successful (Figure 5.4).



**Lane 0:** 100bp DNA ladder; **Lane 1:** IGF-II-pGEM amplified with C1-F/C1-R primers; **Lane 2:** IGF-II-pGEM amplified with C2-F/C2-R primers; **Lane 3:** IGF-II-pGEM amplified with C3-F/C3-R primers; **Lane 4:** Undigested IGF-II-pGEM vector control.



**Lane 0:** 1kb DNA ladder; **Lane 1:** digested C1-pGEM isolate #1; **Lane 2:** digested C1-pGEM isolate #2; **Lane 3:** digested C1-pGEM isolate #3; **Lane 4:** digested C2-pGEM isolate #1; **Lane 5:** digested C2-pGEM isolate #2; **Lane 6:** digested C2-pGEM isolate #3; **Lane 7:** digested C3-pGEM isolate #1; **Lane 8:** digested C3-pGEM isolate #2; **Lane 9:** digested C3-pGEM isolate #3; **Lane 10:** undigested C3-pGEM isolate #3 control; **Lane 11:** digested IGF-II-pGEM control; **Lane 12:** undigested IGF-II-pGEM control.

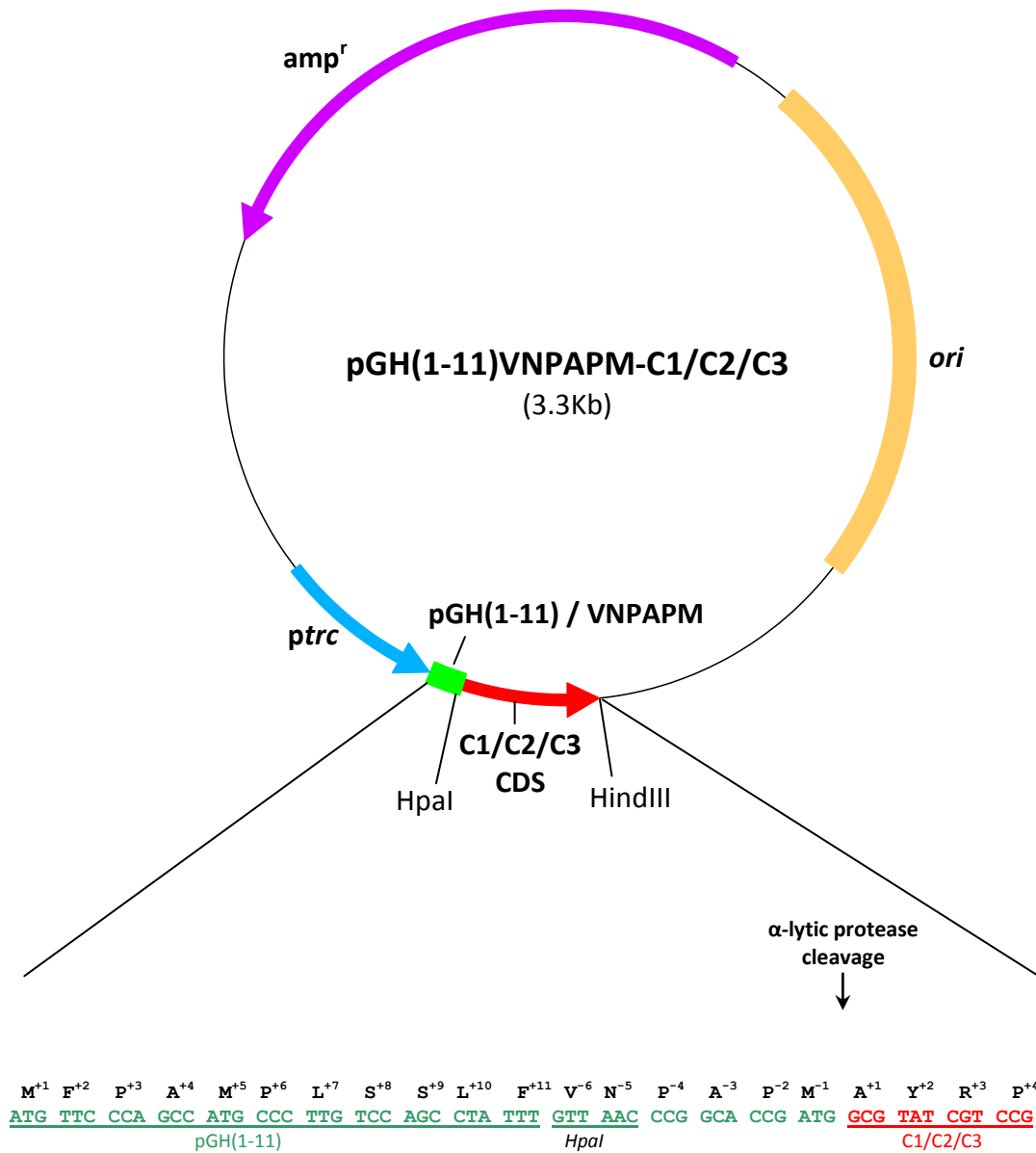
**Figure 5.2 Gel electrophoresis of *DpnI* digested DNA to create C1, C2, C3 mutants and restriction screening.**

**(A) Pre-*DpnI* treatment:** IGF-II-pGEM was used as a template for linear amplification reactions using primers with the required codon changes to create nicked circular DNA with the C1, C2 & C3 mutations. Plasmid DNA is indicated by a  $\leftarrow^P$  and nicked circular DNA is indicated by a  $\leftarrow^N$ .

**(B) Post-*DpnI* treatment:** The PCR products were subsequently treated with *DpnI* restriction endonuclease to digest the methylated plasmid DNA thereby isolating the amplified nicked circular DNA for subsequent cloning into *E. coli*. Plasmid DNA is indicated by a  $\leftarrow^P$  and nicked circular DNA is indicated by a  $\leftarrow^N$ .

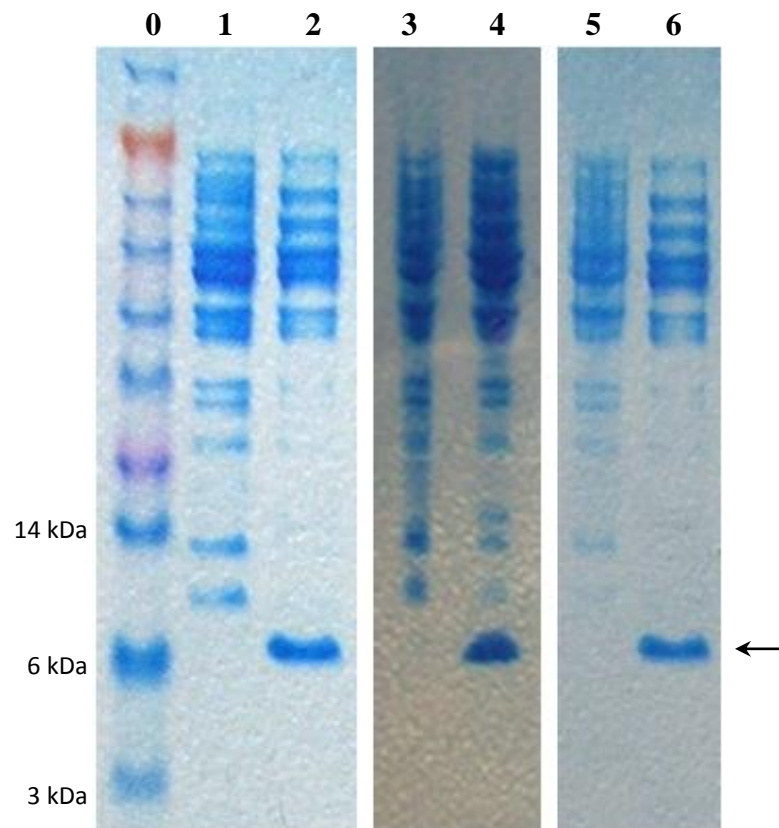
**(C) Restriction enzyme screen of *DpnI* site-directed mutants.** Each of the C1, C2 & C3 coding sequences contained an engineered *NgoMIV* restriction site to facilitate restriction enzyme screening of mutant sequences from template sequences with cleavage of positive mutants into 2 fragments. Putative C1, C2 & C3 nicked circular DNA were cloned into *E. coli* and plasmid DNA from single colony isolates were extracted, digested with *NgoMIV* and the resulting DNA fragments were visualized by gel electrophoresis to identify the following expected fragment sizes:

C1/C2/C3-pGEM:	~2.75kb & ~470/500bp (depending on insert orientation)
Undigested C3-pGEM:	~3.25kb (supercoiled)
Digested IGFII-pGEM:	~3.2kb (linear)
Undigested IGFII-pGem:	~3.2kb (supercoiled)



**Figure 5.3 Expression vector used to express C1, C2 & C3 proteins in *E. coli*.**

The pGH(1-11)VNPAPM expression vector utilized for expression of wildtype and mutant IGF-I & IGF-II proteins in *E. coli*. C1, C2 & C3 peptide sequences were fused to the first eleven amino acids of porcine growth hormone and a VNPAPM linker sequence. The α-lytic protease mutant Prag A9 cleaves the PAMM recognition site between M<sup>-1</sup> and A<sup>+1</sup> thereby liberating the mature IGF-II analogue. Gene expression is regulated by the IPTG-inducible promoter *trc* and the β-lactamase gene provides ampicillin selection.



**Lane 0:** SeeBlue Plus2 pre-stained standard; **Lane 1:** C1 pre-induction cell lysate; **Lane 2:** C1 post-induction cell lysate; **Lane 3:** C2 pre-induction cell lysate; **Lane 4:** C2 post-induction cell lysate; **Lane 5:** C3 pre-induction cell lysate; **Lane 6:** C3 post-induction cell lysate.

**Figure 5.4 Induction of C1, C2 & C3 protein expression in *E. coli*.**

Overnight liquid cultures of *E. coli* JM101 cells transformed with pGH(1-11)VNPAPM-C1/C2/C3 vectors were induced with 100µM IPTG for 5 hours. Verification of induction was obtained by electrophoresis of pre- & post-induction samples through 12% bis-tris polyacrylamide gels and staining with Coomassie blue dye. Successful induction was determined by the appearance of ~9kDa bands in the post-induction samples as indicated with an arrow.

### **5.3.2 Purification, refolding and cleavage of IGF-II mutant proteins**

Please note: samples of the C1, C2 and C3 proteins were monitored by analytical HPLC (2.2.5.9.1) at each of the following purification, refolding and cleavage steps.

The pelleted *E. coli* cells stored at -80°C were thawed on ice and resuspended in 13ml of IB wash buffer (2.1.9.14). The four replicate samples were pooled together in a small beaker on ice and cells were lysed by French press and IBs were pelleted by centrifugation (2.2.5.2.2.)

#### **5.3.2.1 Gel filtration chromatography**

The IBs were dissolved in IB dissolution buffer with 20mM DTT (2.1.9.13) at a concentration of 10ml/g of IB and purified by gel filtration chromatography using an XK26 column packed with Sephacryl S-200 resin gel filtration column at a flow rate of 1ml/min (2.2.5.2.3.1). The elution profile was monitored using a UV detector at 280nm, 5ml fractions were collected and the presence of fusion protein in alternate fractions was monitored by SDS-PAGE.

#### **5.3.2.2 Refolding of fusion protein**

Gel filtration fractions containing fusion protein were pooled in a small beaker on ice. Refolding reactions were prepared (2.2.5.2.3.2) and initiated by the addition of 1mM 2-hydroxyethylidysulphide as an oxidizing agent to promote disulphide bond formation. The refolding reaction was incubated at 120min at room temperature with gentle stirring. The status of the refolded isomers was monitored at various time points by analytical HPLC and the reaction was terminated by acidification with concentrated HCl to pH 2.5.



### **5.3.2.3 Ion exchange chromatography**

The refolded protein was further purified by ion exchange chromatography (2.2.5.2.3.3) using an XK16 column packed with Fast Flow S cation exchange resin at a flow rate of 8ml/min. The column was washed with FFS buffer A (2.1.9.10) and then FFS buffer B (2.1.9.11) was added to elute protein. Although this was an established protocol for the purification of IGF-II protein, HPLC analysis of the load and eluate indicated the IGF-II analogues were not binding to the column and were being eluted in the load. As relatively high quantities of recombinant protein were present in the load and other material was present in the eluate, a degree of purification had occurred and subsequent  $\alpha$ -lytic protease cleavage was performed using the protein in the load solution.

### **5.3.2.4 $\alpha$ -lytic protease cleavage**

The  $\alpha$ -lytic protease Prag A9 mutant enabled efficient cleavage of the N-terminal leader sequence from the mature IGF-II analogues (Lien *et al.*, 2001). The cleavage reaction was performed (2.2.5.2.3.4) and the efficiency of the reaction was monitored at periodic time points by analytical HPLC. The reaction was terminated by acidification by HCl to pH 2.5 and final purification was achieved by preparative HPLC.

### **5.3.2.5 Preparative reverse phase HPLC**

Preparative HPLC was used for the final purification of the mature IGF-II analogues from the cleaved leader peptides using a V-25 10mm x 250mm C4 column (2.2.5.9.2.) Fractions of 5ml were collected and analysed by analytical HPLC to identify fractions with IGF-II protein. These fractions were pooled, aliquoted, lyophilized (2.2.5.10) and stored at -20°C.

### **5.3.3 Isolation & purification of C1 protein summary**

The purification of C1 protein was achieved by using the methodology outlined in 5.3.2. Processing ~64mg of inclusion bodies resulted in a yield of 8.65mg (13%) of

mature protein. A summary of the purification process is shown in Figure 5.5. The purified protein produced a single peak when analysed by analytical HPLC indicating a high degree of purity.

#### **5.3.4 Isolation & purification of C2 protein summary**

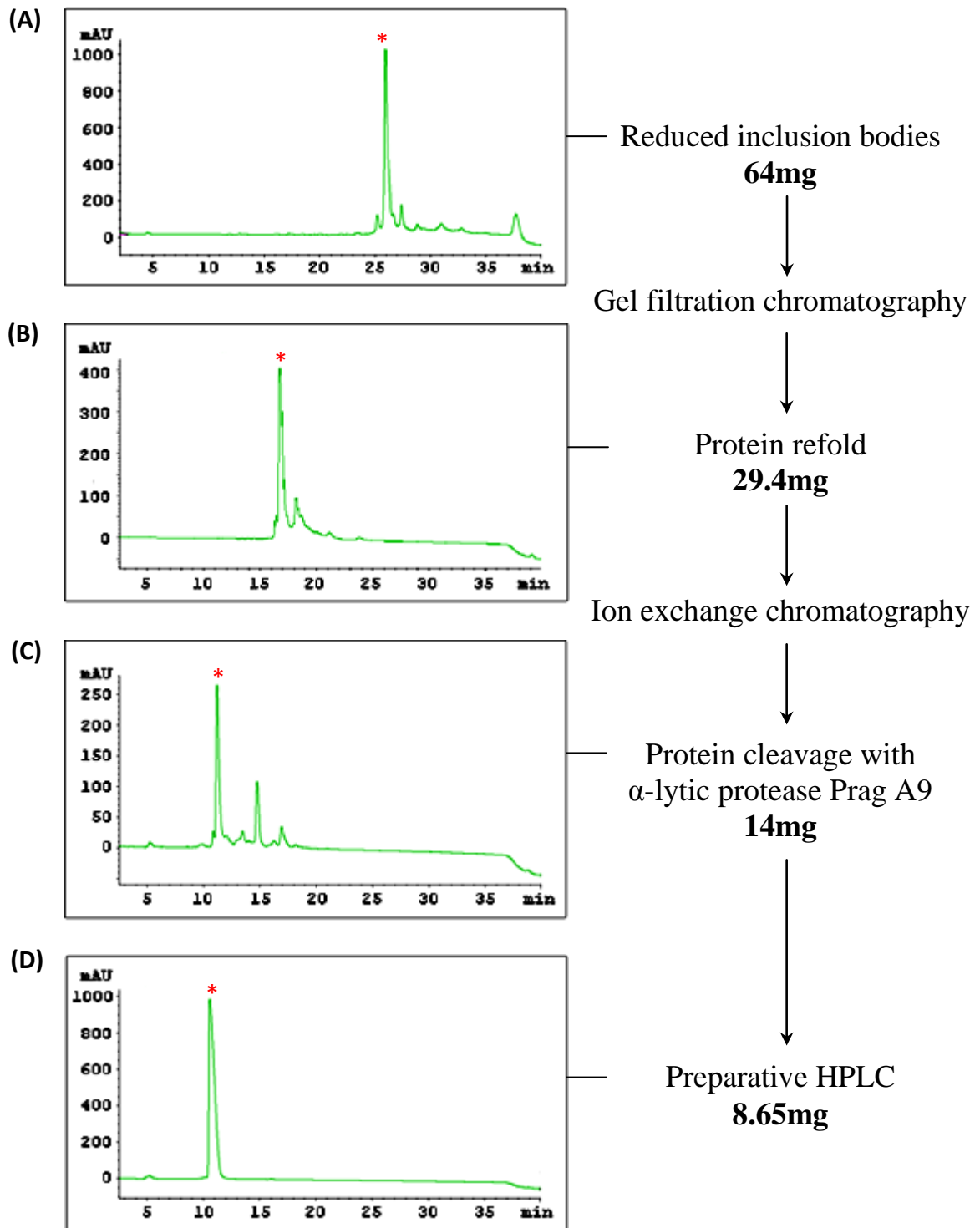
An initial attempt at purifying C2 protein was unsuccessful as multiple peaks were present in HPLC analysis of sample after the final purification step. The experiment was repeated using a faster, smaller scale procedure which is similar to the method outlined in 5.3.2 with the following exceptions:

1. A smaller volume of overnight culture of *E. coli* JM101 cells transformed with pGH(1-11)VNPAPMN-C2 was prepared (2x 500ml) and a quarter of these inclusion bodies were processed with the remainder being stored at -80°C in case method optimisation was required for purifying C2 protein.
2. A smaller gel filtration column (10x300mm with 30ml of Superdex 75 resin) was used at a slower flow rate of 0.5ml/min and 1ml fractions were collected.
3. The ion exchange chromatography step was omitted.

This smaller scale procedure resulted in the successful purification of C2 protein with a single peak present after the final purification step. The C2 protein recovery was 10% with ~270µg of mature protein obtained from 2.7mg of inclusion bodies (Figure 5.6)

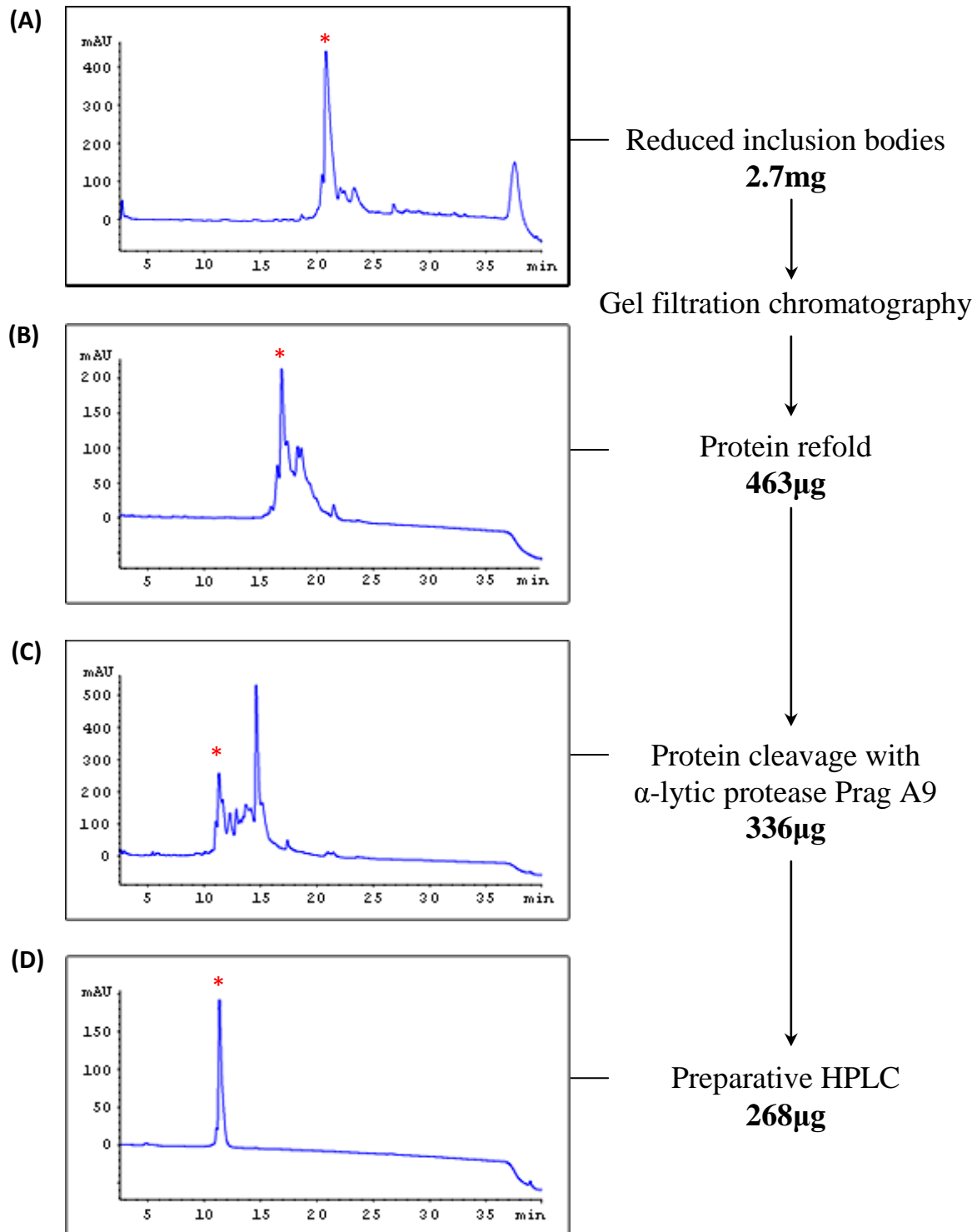
#### **5.3.5 Isolation & purification of C3 protein**

Purification of C3 protein utilised the same methodology as the C1 protein (5.3.2). From 62mg of inclusion bodies, 2.1mg (3.4%) of mature C3 protein was recovered and produced a single peak when analysed by HPLC (Figure 5.7)



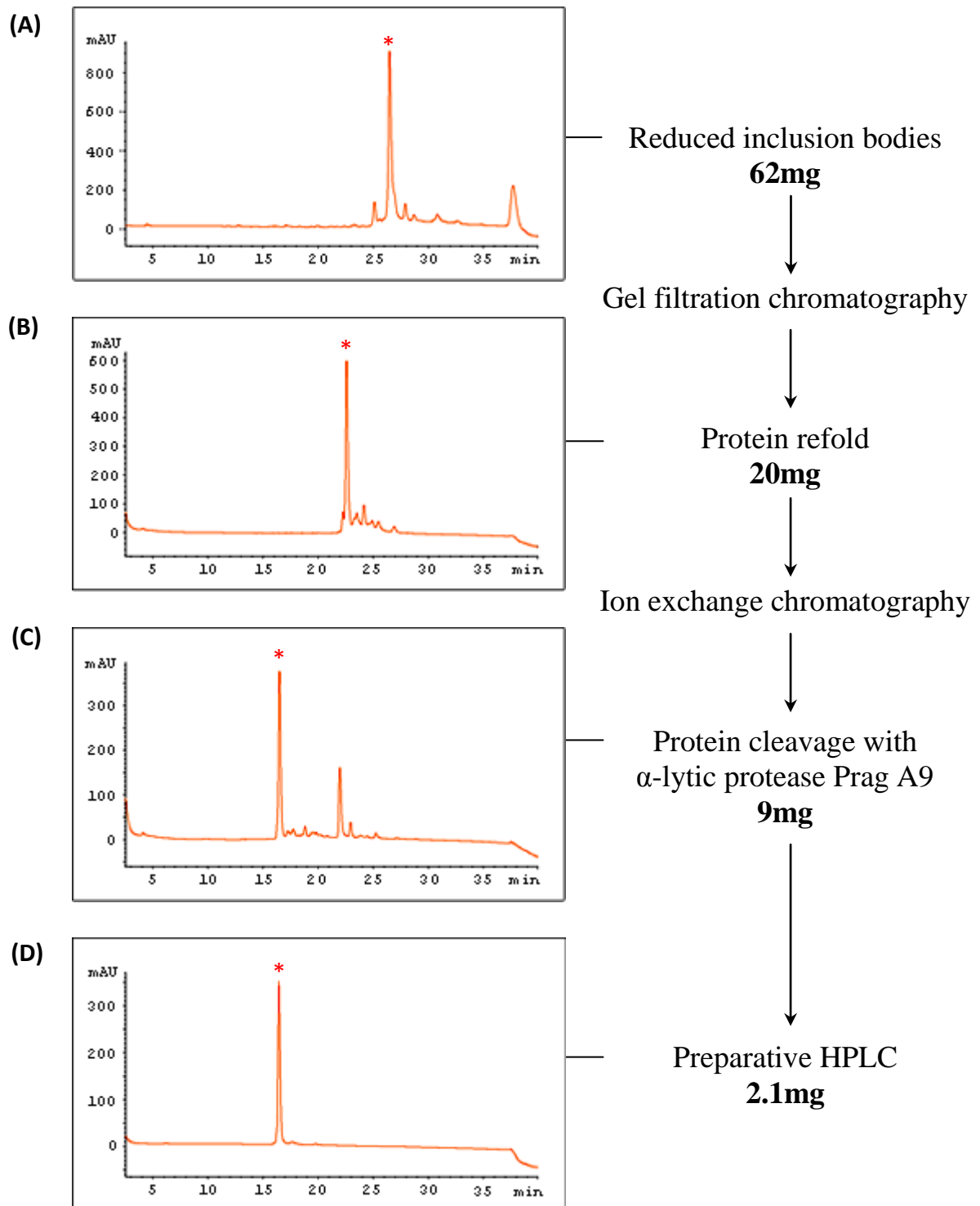
**Figure 5.5 Purification summary of C1 protein.**

The purification steps and associated yield of the C1 protein is shown in the flowchart. Protein purification was monitored at the end of each stage by analytical HPLC and profiles are shown for (A) the reduced inclusion bodies; (B) protein refolded under oxidative conditions; (C)  $\alpha$ -lytic protease cleavage of leader sequence and (D) purification of mature C1 protein by preparative HPLC. The C1 protein peak is indicated with a \*.



**Figure 5.6 Small scale purification summary of C2 protein.**

The purification steps and associated yield of the C2 protein is shown in the flowchart. Protein purification was monitored at the end of each stage by analytical HPLC and profiles are shown for **(A)** the reduced inclusion bodies; **(B)** protein refolded under oxidative conditions; **(C)**  $\alpha$ -lytic protease cleavage of leader sequence and **(D)** purification of mature C2 protein by preparative HPLC. The C2 protein peak is indicated with a \*.



**Figure 5.7 Purification of C3 protein.**

The purification steps and associated yield of the C3 protein is shown in the flowchart. Protein purification was monitored at the end of each stage by analytical HPLC and profiles are shown for **(A)** the reduced inclusion bodies; **(B)** protein refolded under oxidative conditions; **(C)**  $\alpha$ -lytic protease cleavage of leader sequence and **(D)** purification of mature C3 protein by preparative HPLC. The C3 protein peak is indicated with a \*.

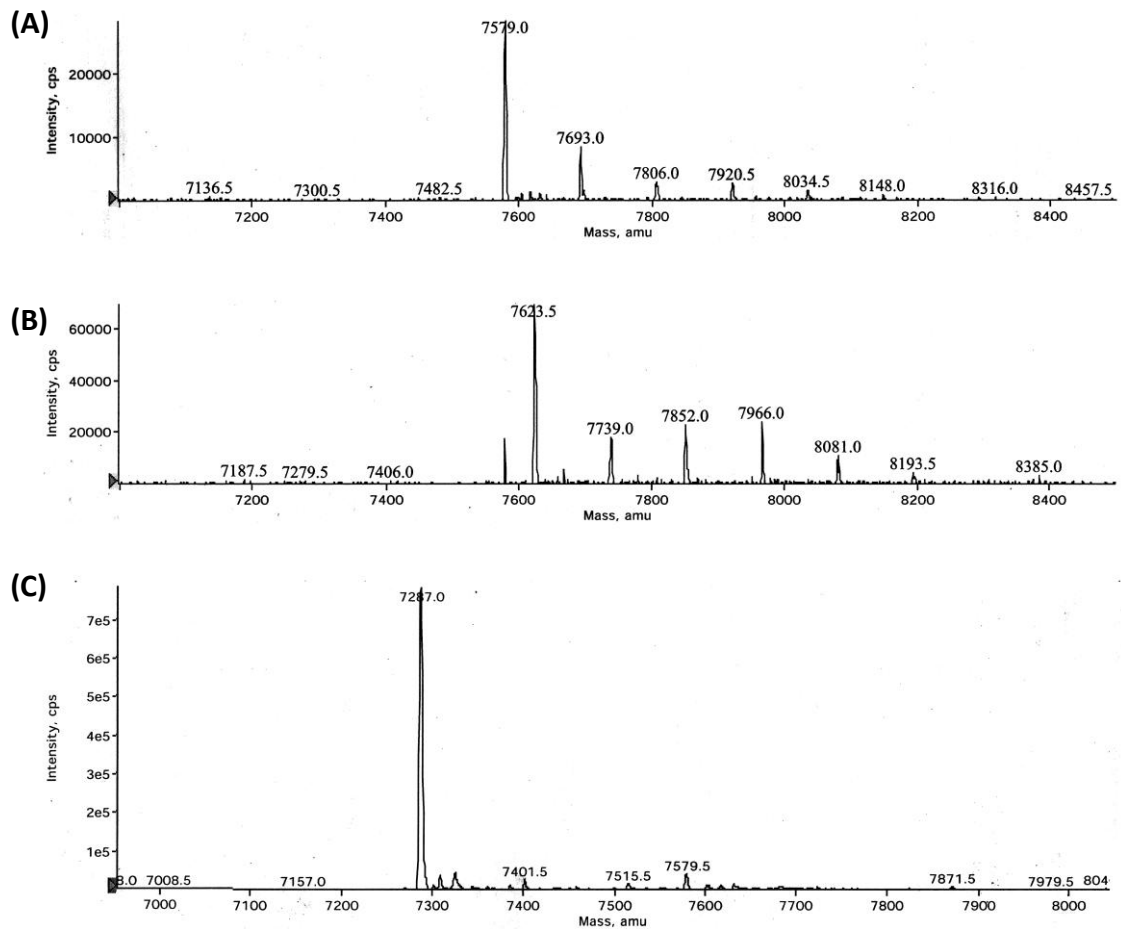
### **5.3.6 Quantitation, mass spectrometry analysis and N-terminal sequencing of IGF-II analogues**

Lyophilised C1, C2 & C3 proteins were reconstituted in 1x PBS and quantified by HPLC using receptor grade human IGF-I as a standard (2.2.5.3.3). The molecular weight of the proteins was determined by mass spectrometry using an Applied Biosystems API-100 LC/MS at the Hanson Institute Protein Core Facility by Dr Chris Bagley and all proteins were found to have molecular masses within 0.03% of the predicted mass (Figure 5.8). In addition to protein peaks, the mass spectra contained evenly spaced higher mass peaks which were presumed to be non-protein contaminants such as TFA adducts (C. Bagley: *Personal communication*). N-terminal sequencing of the first five amino acids in the C1, C2 & C3 proteins was performed with an Applied Biosystems 492 Procise Protein Sequencer by Mr. Chris Cursaro (School of Molecular and Biomedical Science, Adelaide University) and all proteins had the expected AYRPS N-terminal sequence of the mature protein (data not shown).

## **5.4 Characterisation of IGF-II analogues**

Interactions between the C1, C2 & C3 mutants and the IR-A, IR-B and IGF-IR were characterized by competitive receptor binding assays, receptor phosphorylation assays and cell survival assays as summarized below.

- The ligand binding assay measured the ability of various ligand concentrations to outcompete europium labelled insulin (Eu-insulin) (IR-A & IR-B assays) or IGF-I (Eu-IGF-I) (IGF-IR assay) bound to immunocaptured IR-A, IR-B or IG-IR from cell lysates (2.2.6.2).
- The receptor phosphorylation assay determined IR-A, IR-B or IGF-IR autophosphorylation in response to various ligand concentrations through binding of a europium labelled antibody (PY20) to phosphorylated tyrosine residues in immunocaptured receptors from cell lysates (2.2.6.4). Receptor activation is essentially the first step in the receptor signalling pathways.



Protein	Molecular weight (kDa)	
	Predicted	Observed
C1	7578.5	7579.0 ±1
C2	7623.6	7623.5 ±1
C3	7287.2	7286.0 ±1

**Figure 5.8** Mass spectrometry analysis of purified C1, C2 & C3 proteins.

Lyophilised purified proteins were resuspended in 2% (v/v) acetic acid and analysed by mass spectrometry and the data were transformed onto a true-mass scale. The mass spectrometer analysis of C1 is shown in (A), C2 in (B) and C3 in (C).

- The cell survival assay determined the effect of various ligand concentrations on cell viability in R<sup>1</sup>IRA, R<sup>1</sup>IRB or R<sup>1</sup>IGF-IR cells challenged with 5mM of butyrate (2.2.6.5). Butyrate is a short chain fatty acid which induces apoptosis and inhibits cell proliferation in culture (Singh *et al.*, 1997) possibly through a mechanism involving inhibition of histone deacetylase activity (Shao *et al.*, 2004). This assay was used to determine if the IGF-II analogues elicited a biological response through the IR-A, IR-B and IGF-IR.

All assays were performed in 96-well plate format using triplicate samples and the combined data from at least three independent experiments are presented as concentration-response curves. Data from a single concentration are presented as column graphs to facilitate comparison at physiological levels of ligand. The IGF-II, IGF-I and Insulin ligands were obtained from commercial suppliers and the BCIAD ligand was kindly supplied by Professor John Wallace and Dr Briony Forbes, School of Molecular & Biomedical Science, Adelaide University (2.1.7).

#### **5.4.1 IR-A interaction with the IGF-II analogues**

##### **5.4.1.1 IR-A binding**

IGF-II, C1 and C2 bound with similar binding affinities to the IR-A ( $IC_{50}$ s of  $5.9 \pm 0.7$ nM,  $7.5 \pm 0.8$ nM and  $6.8 \pm 1.3$ nM, respectively) and with greater affinity than BCIAD ( $IC_{50}$ :  $12.0 \pm 1.8$ nM) indicating the C-domain configurations of both C1 and C2 facilitated high affinity binding to the IR-A (Table 5.1 and Figure 5.9A). In contrast, C3 ( $IC_{50}$ :  $13.8 \pm 3.0$ nM) bound with a 13% lower affinity than BCIAD.

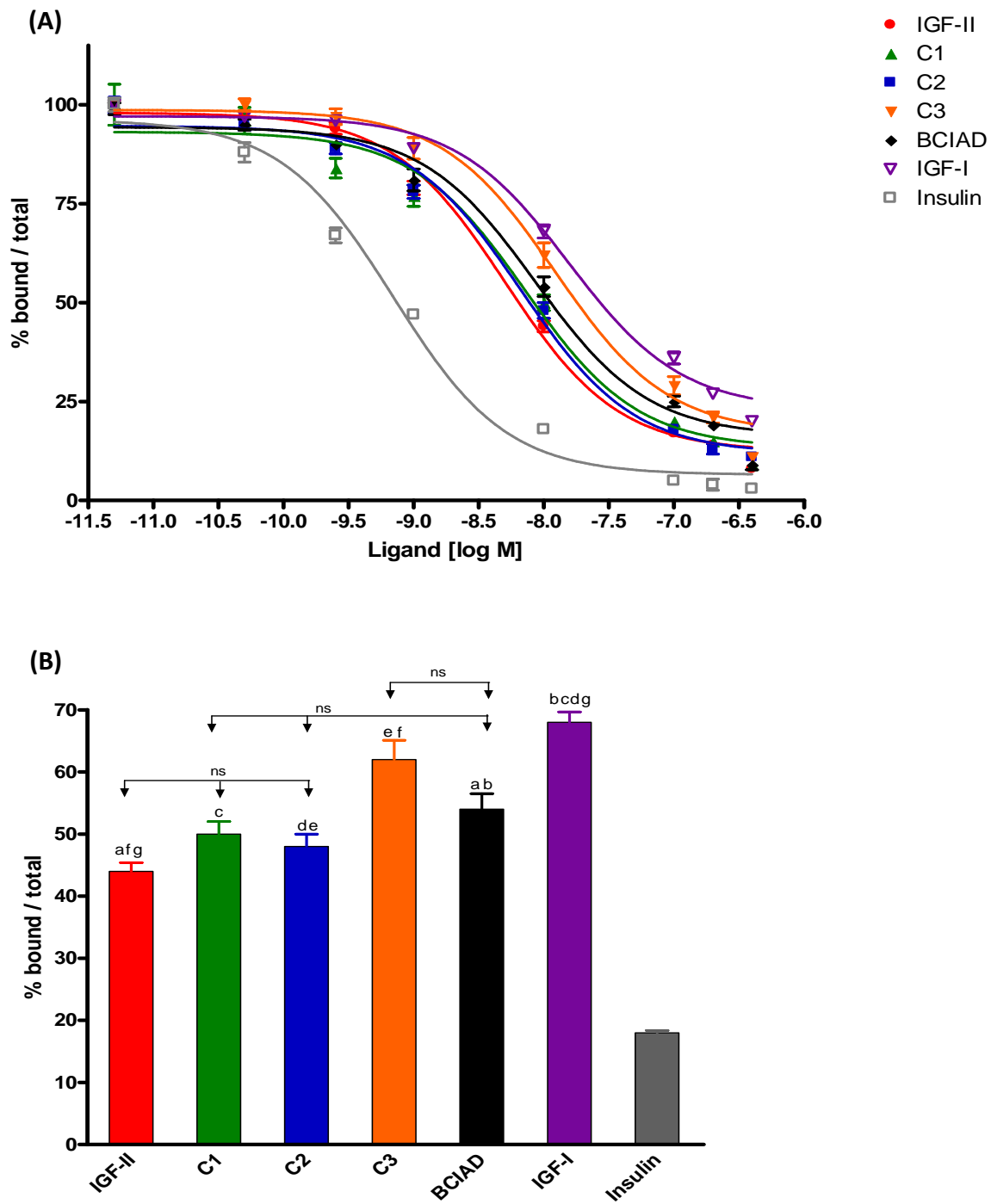
At a ligand concentration of 10nM, ANOVA analysis indicated there was a significant difference ( $p < 0.001$ ) between IGF-II and BCIAD binding affinities. Although C1 and C2 were more IGF-II-like in terms of binding affinity, there was no significant difference ( $P > 0.05$ ) between the binding affinity of C1 and C2 to BCIAD (Figure 5.9B) reflecting the intermediate binding characteristics of these two analogues. As expected, insulin and IGF-I bound to the IR-A with the highest ( $IC_{50}$ :  $1.2 \pm 0.2$ nM) and lowest ( $IC_{50}$ :  $26.2 \pm 4.3$ nM) affinities, respectively.



Ligand	IR-A		IR-B		IGF-IR	
	IC <sub>50</sub> (nM)	Rel. IC <sub>50</sub> (%)	IC <sub>50</sub> (nM)	Rel. IC <sub>50</sub> (%)	IC <sub>50</sub> (nM)	Rel. IC <sub>50</sub> (%)
<b>IGF-II</b>	5.9 ± 0.7	100	11.5 ± 0.9	100	0.19 ± 0.02	100
<b>C1</b>	7.5 ± 0.8	79	23.0 ± 3.6	50	0.12 ± 0.02	158
<b>C2</b>	6.8 ± 1.3	87	15.8 ± 2.3	73	0.10 ± 0.03	190
<b>C3</b>	13.8 ± 3.0	43	20.0 ± 2.7	58	0.91 ± 0.48	21
<b>BCIAD</b>	12.0 ± 1.8	49	24.5 ± 2.2	47	0.14 ± 0.02	136
<b>IGF-I</b>	26.2 ± 4.3	23	40.6 ± 4.5	28	0.09 ± 0.01	211
<b>Insulin</b>	1.2 ± 0.2	492	0.6 ± 0.03	1917	18.0 ± 7.0	1

**Table 5.1 Summary of IC<sub>50</sub> values for ligands analysed in competitive IR-A, IR-B & IGF-IR ligand binding assays.**

The IC<sub>50</sub> values are the means ± SEM from three independent experiments. The Rel. IC<sub>50</sub> refers to the relative binding affinity of each ligand compared to IGF-II (Rel. IC<sub>50</sub> (%) = IC<sub>50</sub> of IGF-II / IC<sub>50</sub> ligand x 100)



**Figure 5.9 IR-A competitive ligand binding assay of IGF-II analogues.**

**(A)** Plates with immunocaptured IR-A were treated with increasing concentrations of unlabelled ligands in the presence of europium-labelled insulin. Results are expressed as a percentage of europium labelled insulin bound in the absence of competing ligand and the data points are the mean  $\pm$  S.E.M of triplicate samples from at least three independent experiments. Error bars are shown when greater than the size of the symbols. **(B)** IR-A competitive binding at 10nM of ligand. The statistical significance of differences between all ligand responses except insulin was determined by ANOVA with Tukey’s multiple comparison test. Note: a, b, c, d, f & g =  $p < 0.001$ ; e =  $p < 0.05$ . ns = not significant ( $p > 0.05$ ).

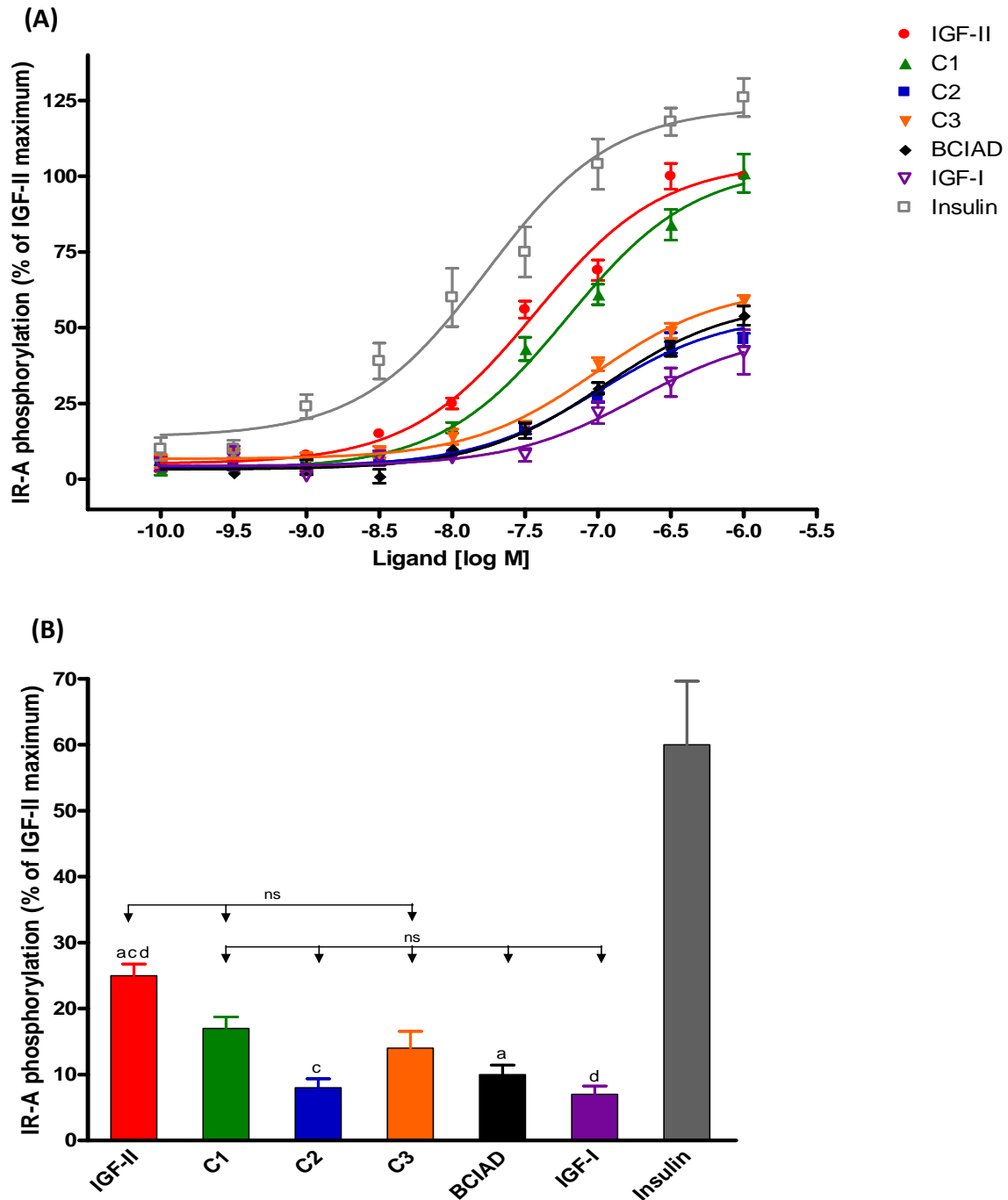
#### **5.4.1.2 IR-A phosphorylation**

Despite the IR-A binding affinities of C1 and C2 being equipotent (refer 5.4.1.1), the concentration-phosphorylation response and maximum phosphorylation induced by C1 was comparable to IGF-II yet C2 generated only 50% maximum phosphorylation relative to IGF-II and had a concentration-response similar to C3, BCIAD and IGF-I (Figure 5.10A).

Although the ligand concentrations generating the most disparate activation responses were at supraphysiological levels, there was a significant difference ( $P < 0.01$ ) between IGF-II and C2 induced phosphorylation at the 10nM concentration (Figure 5.10B). Insulin elicited the highest levels of phosphorylation and IGF-I induced the lowest phosphorylation consistent with their relative IR-A binding affinities.

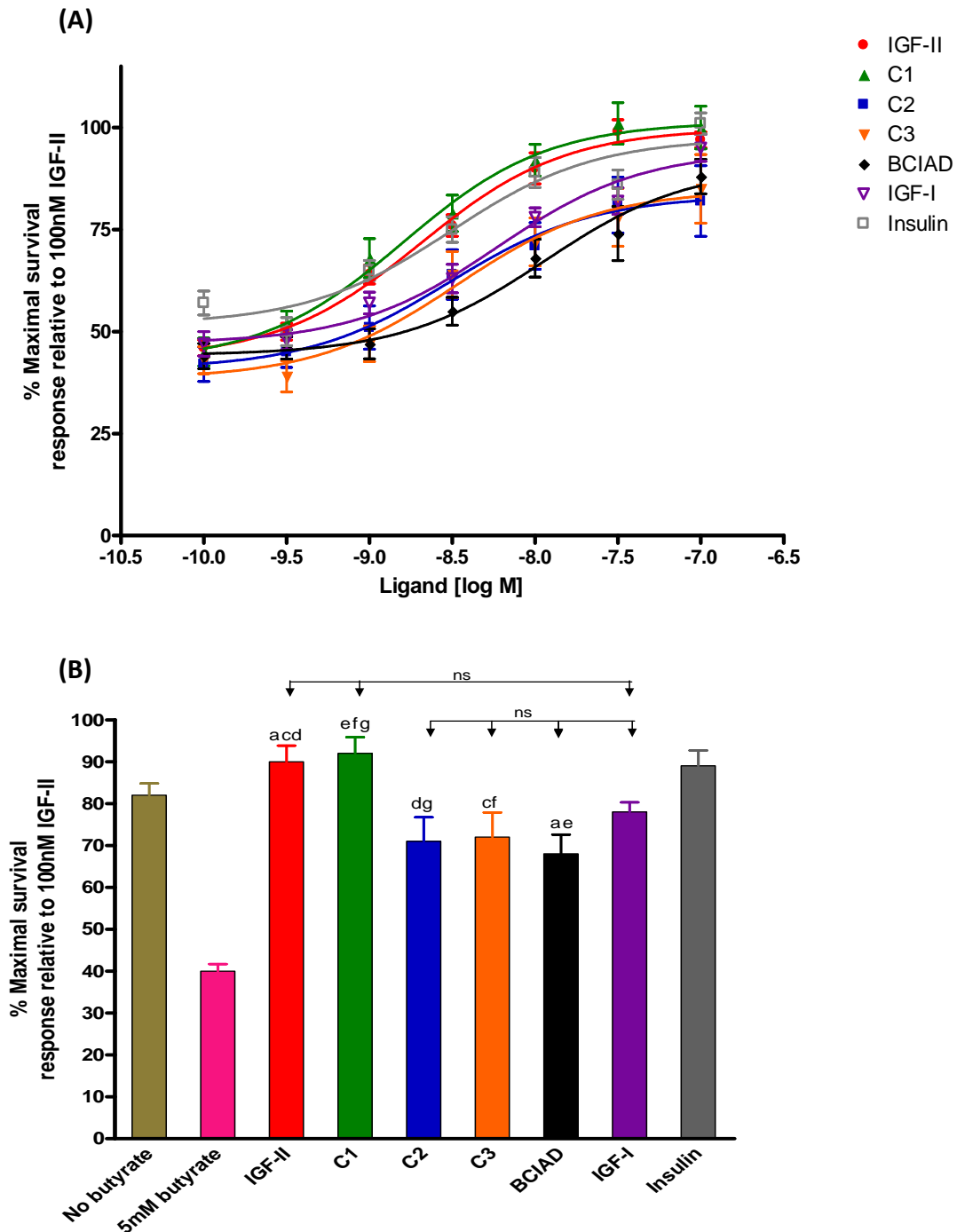
#### **5.4.1.3 R<sup>1</sup>IR-A cell survival**

The addition of 5mM butyrate resulted in a 2-fold reduction in R<sup>1</sup>IRA cell survival and all ligands assayed were able to increase cell viability (Figure 5.11B). The R<sup>1</sup>IRA cell survival assay data were generally consistent with the IR-A phosphorylation assay. Specifically, IGF-II, C1 and insulin elicited similar response curves whereas the C2, C3 and BCIAD responses were slightly lower than IGF-I (Figure 5.11A). At 10nM, the survival response induced by IGF-II and C1 was significantly higher ( $p < 0.05$ ) than C2, C3 and BCIAD (Figure 5.11B).



**Figure 5.10 IR-A phosphorylation by IGF-II analogues.**

**(A)** R1RA cells overexpressing IR-A were stimulated with increasing concentration of ligand for 10 minutes. Cells were lysed and activated receptors were immunocaptured on 96-well plates and incubated with Eu-PY20 to detect phosphorylated tyrosines using time-resolved fluorescence. The data points are the mean  $\pm$  S.E.M of triplicate samples from at least three independent experiments. Error bars are shown when greater than the size of the symbols. **(B)** IR-A phosphorylation at 10nM of ligand. The statistical significance of differences between all ligand responses except insulin was determined by ANOVA with Tukey's multiple comparison test. Note: a & c =  $p < 0.01$ ; d =  $p < 0.001$ . ns = not significant ( $p > 0.05$ ).



**Figure 5.11 R<sup>1</sup>R-A cell survival assay of IGF-II analogues.**

**(A)** R<sup>1</sup>R-A cells overexpressing IR-A were incubated with increasing concentration of ligand in the presence of 5mM butyrate. Cell viability was determined using the Promega Cell-titre Blue Reagent. Results are expressed as % survival relative to the maximum response obtained with 100nM of IGF-II. The data points are the mean  $\pm$  S.E.M of triplicate samples from at least three independent experiments. Error bars are shown when greater than the size of the symbols. **(B)** R<sup>1</sup>R-A cell survival at 10nM of ligand in the presence of 5mM butyrate. The statistical significance of differences between all ligand responses except insulin was determined by ANOVA with Tukey's multiple comparison test. Note: a, & e =  $p < 0.01$ ; c, d, f & g =  $p < 0.05$ . ns = not significant ( $p > 0.05$ ).

## **5.4.2 IR-B interaction with the IGF-II analogues**

### **5.4.2.1 IR-B binding**

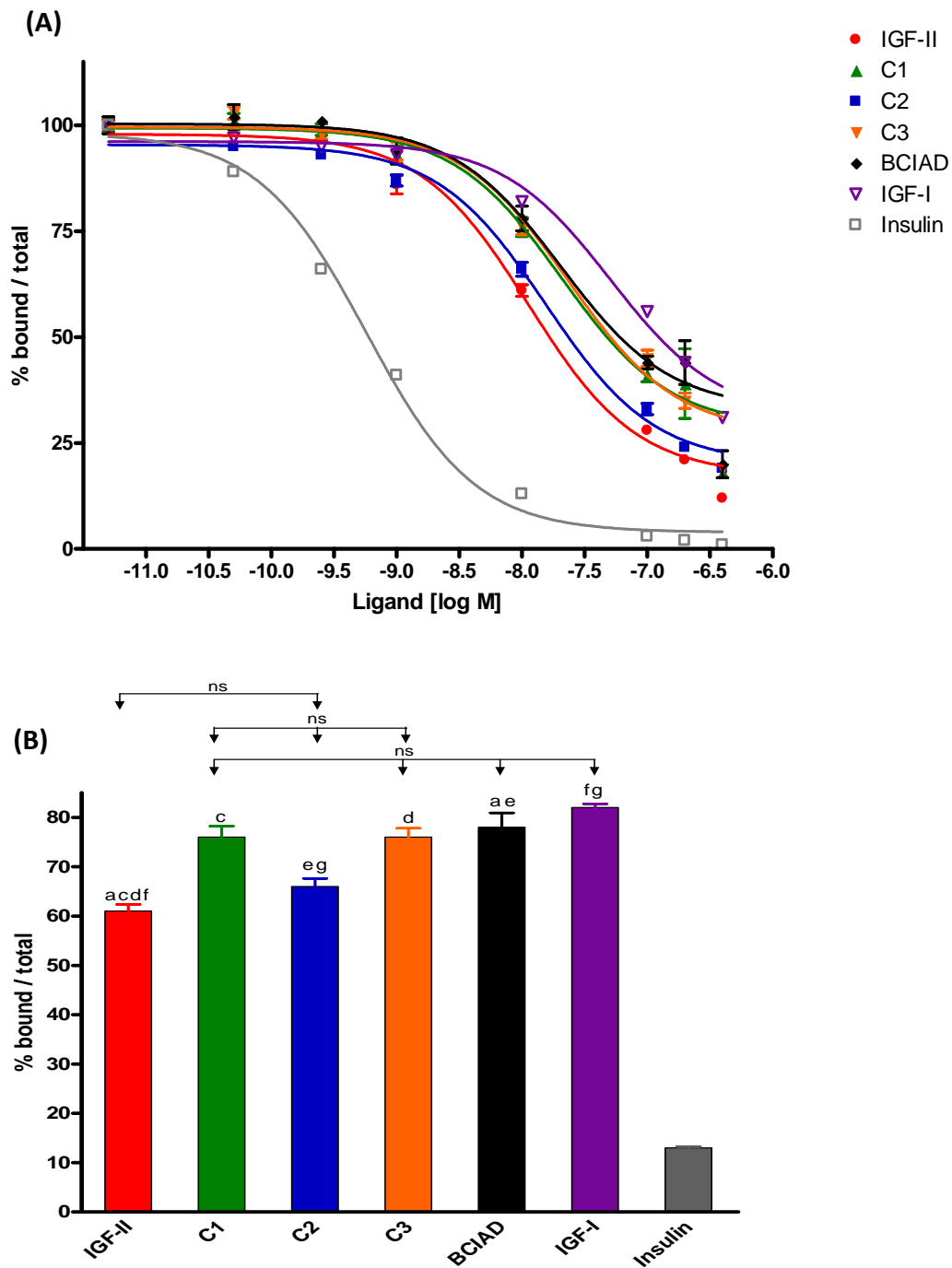
Insulin bound the IR-B with an  $\sim 2$ -fold greater affinity than the IR-A whereas all other ligands had  $\sim 1.5$  to  $3$ -fold lower affinity for the IR-B (Table 5.1 and Figure 5.12A). The IGF-II and C2 ligands had similar IR-B binding affinities ( $IC_{50}$ s of  $11.5 \pm 0.9$ nM and  $15.8 \pm 2.3$ nM, respectively). Interestingly, C1 ( $IC_{50}$ :  $23.0 \pm 3.6$ nM) bound with similar affinity to C3 ( $IC_{50}$ :  $20.0 \pm 2.7$ nM) and BCIAD ( $24.5 \pm 2.2$ nM), the collective binding affinities of these three proteins were  $\sim 2$ -fold lower than IGF-II and  $\sim 2$ -fold higher than IGF-I ( $IC_{50}$ :  $40.6 \pm 4.5$ nM).

The IR-B binding affinities of ligands at 10nM were consistent with the overall concentration-response curves described above. The IGF-II and C2 proteins bound the IR-B with similar affinities. The C1, C3, BCIAD and IGF-I bound with similar affinities, which were significantly lower ( $p < 0.001$ ) than the binding affinity of IGF-II (Figure 5.12B).

### **5.4.2.2 IR-B phosphorylation**

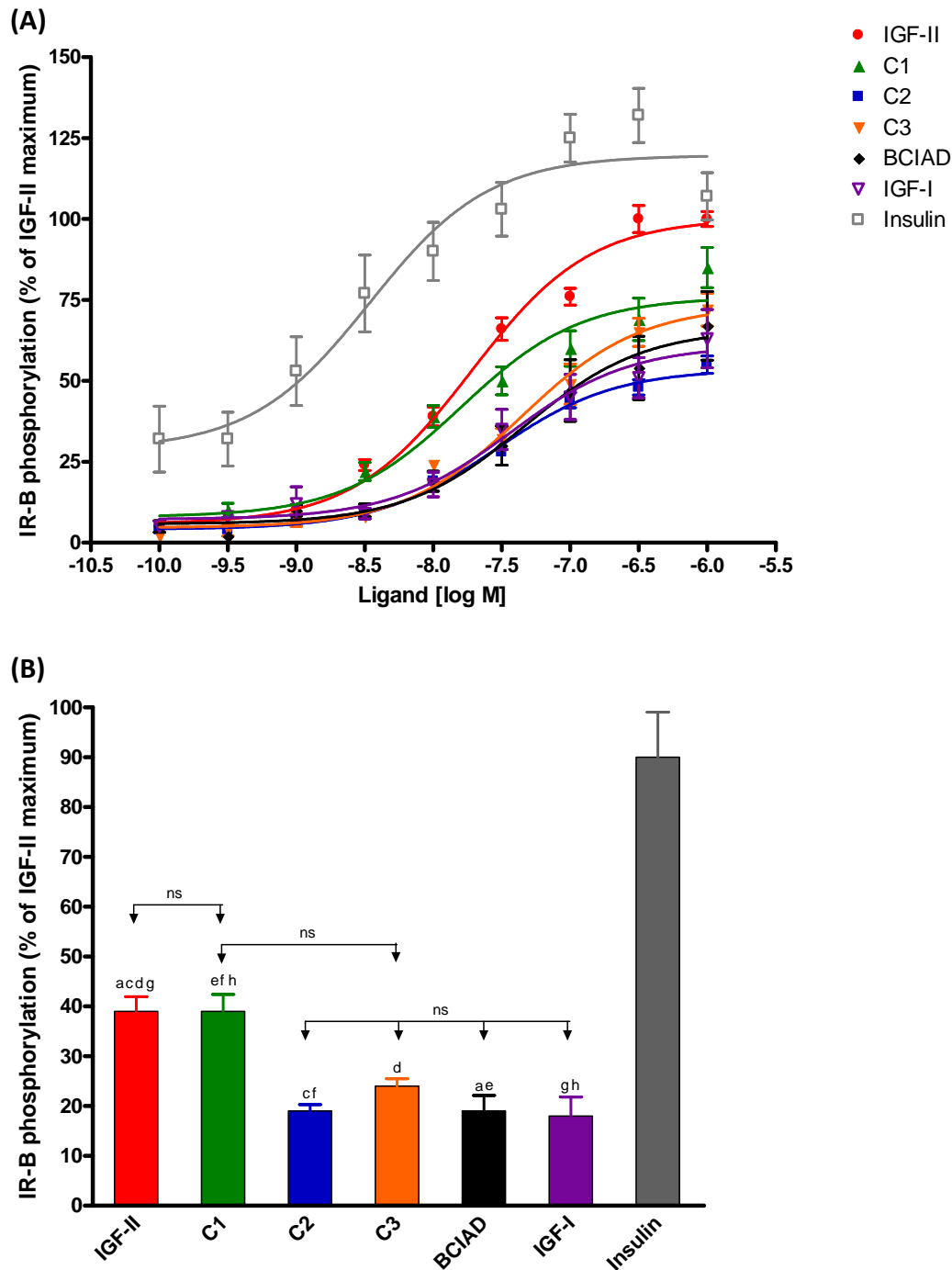
Despite the relatively high IR-B binding affinity of C2, this ligand induced the lowest levels of IR-B kinase activity with maximum phosphorylation being  $\sim 50\%$  of the IGF-II peak level (Figure 5.13A). The C2 phosphorylation response curve was similar to the IGF-I, BCIAD and C3 curves. C1 bound the IR-B with BCIAD-like affinity yet receptor activation at concentrations up to 10nM was comparable to IGF-II. However, at concentrations above 30nM the C1 phosphorylation response increased at a progressively lesser rate relative to IGF-II and generated a phosphorylation maximum that was  $\sim 75\%$  of the IGF-II maximum. Insulin induced the highest phosphorylation levels corresponding with the high binding affinity of this ligand to the IR-B.

At 10nM, the C1 induced phosphorylation levels was significantly higher than C2 levels ( $p < 0.05$ ). C1 and IGF-II induced equivalent IR-B phosphorylation levels whilst the C2 response was similar to the lower activation induced by C3, BCIAD and IGF-I (Figure 5.13B).



**Figure 5.12** IR-B Competitive ligand binding assay of IGF-II analogues.

**(A)** Plates with immunocaptured IR-B were treated with increasing concentrations of unlabelled ligands in the presence of europium-labelled insulin. Results are expressed as a percentage of europium labelled insulin bound in the absence of competing ligand and the data points are the mean  $\pm$  S.E.M of triplicate samples from at least three independent experiments. Error bars are shown when greater than the size of the symbols. **(B)** IR-B competitive binding at 10nM of ligand. The statistical significance of differences between all ligand responses except insulin was determined by ANOVA with Tukey's multiple comparison test. Note: a, c, d & f =  $p < 0.001$ ; e =  $p < 0.05$ ; g =  $p < 0.01$ . ns = not significant ( $p > 0.05$ ).



**Figure 5.13 IR-B phosphorylation by IGF-II analogues.**

**(A)** R<sup>1</sup>IRB cells overexpressing IR-B were stimulated with increasing concentration of ligand for 10 minutes. Cells were lysed and activated receptors were immunocaptured on 96-well plates and incubated with Eu-PY20 to detect phosphorylated tyrosines using time-resolved fluorescence. The data points are the mean  $\pm$  S.E.M of triplicate samples from at least three independent experiments. Error bars are shown when greater than the size of the symbols. **(B)** IR-B phosphorylation at 10nM of ligand. The statistical significance of differences between all ligand responses except insulin was determined by ANOVA with Tukey's multiple comparison test. Note: a & c =  $p < 0.01$ ; d, e, f & h =  $p < 0.05$ ; g =  $p < 0.001$ . ns = not significant ( $p > 0.05$ ).



### **5.4.2.3 R<sup>1</sup>IR-B cell survival**

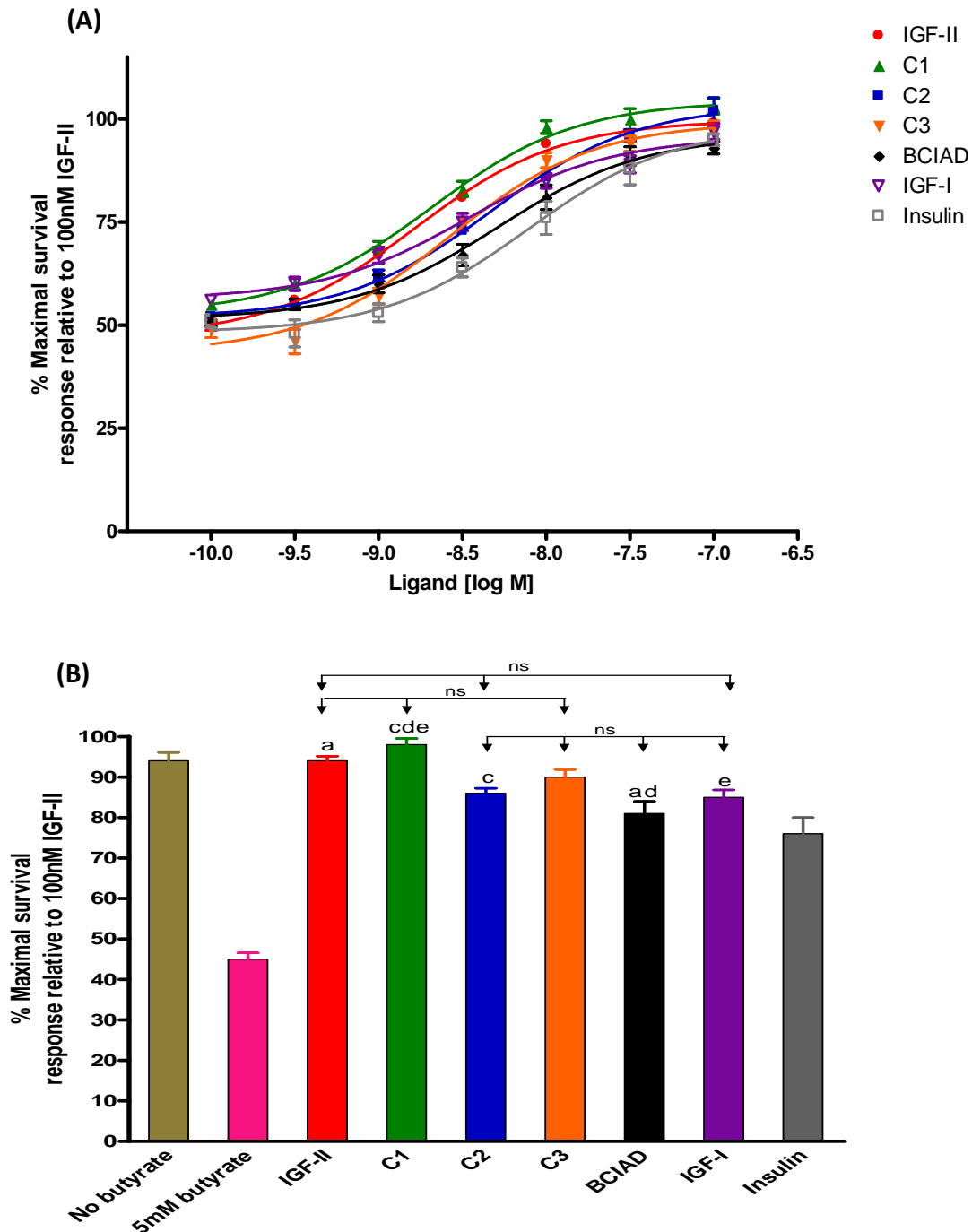
The C1 ligand had a concentration-response curve similar to IGF-II and had the greatest effect on R<sup>1</sup>IRB cell survival (Figure 5.14A) even at high concentrations where IR-B phosphorylation levels dropped relative to IGF-II (Figure 5.13A). At concentrations below 10nM, the C2 and C3 ligands elicited similar responses to IGF-I and BCIAD although at higher concentrations the effects on cell survival were similar to C1 and IGF-II. Interestingly, insulin produced the lowest cell survival response despite insulin's high affinity binding and induced phosphorylation of the IR-B. At 10nM, C1 had a significantly higher ( $p < 0.05$ ) response than C2 although neither were significantly different from IGF-II or C3 (Figure 5.14B).

### **5.4.3 IGF-IR interaction with the IGF-II analogues**

#### **5.4.3.1 IGF-IR binding**

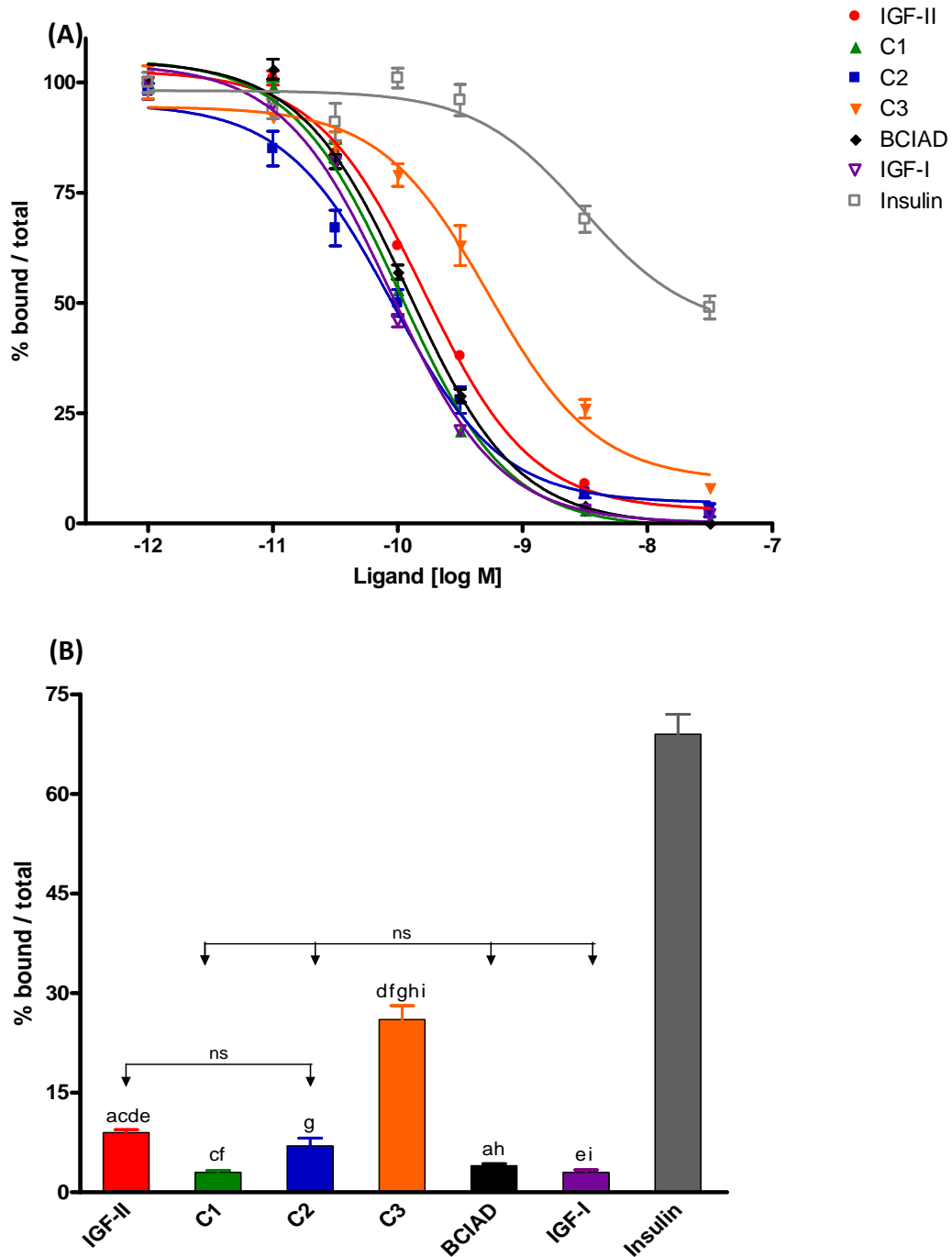
The C1 ( $IC_{50}$ :  $0.12 \pm 0.02$ nM) and C2 ( $IC_{50}$ :  $0.10 \pm 0.03$ nM) analogues bound the IGF-IR with similar affinity to IGF-I ( $IC_{50}$ :  $0.09 \pm 0.01$ nM) and higher affinity than BCIAD ( $IC_{50}$ :  $0.14 \pm 0.02$ nM) and IGF-II ( $IC_{50}$ :  $0.19 \pm 0.02$ nM). The C3 ( $IC_{50}$ :  $0.91 \pm 0.48$ nM) ligand and insulin ( $IC_{50}$ :  $18.0 \pm 7.0$ nM) ligands bound the IGF-IR with relatively poor affinities (Table 5.1 and Figure 5.15A).

At a ligand concentration of 3nM, the C1, C2, BCIAD and IGF-I proteins bound the IGF-IR with similar affinity (Figure 5.15B). IGF-II bound with lower affinity than C1 ( $p < 0.01$ ), BCIAD & IGF-I ( $p < 0.001$ ) whilst the difference between IGF-II and C2 binding affinities were statistically insignificant ( $p > 0.05$ ). C3 and insulin bound the IGF-IR poorly relative to the other ligands.



**Figure 5.14** R<sup>1</sup>R-B survival assay of IGF-II analogues.

**(A)** R<sup>1</sup>R-B cells overexpressing IR-B were incubated with increasing concentration of ligand in the presence of 5mM butyrate. Cell viability was determined using the Promega Cell-titre Blue Reagent. Results are expressed as % survival relative to the maximum response obtained with 100nM of IGF-II. The data points are the mean  $\pm$  S.E.M of triplicate samples from at least three independent experiments. Error bars are shown when greater than the size of the symbols. **(B)** R<sup>1</sup>R-B cell survival at 10nM of ligand in the presence of 5mM butyrate. The statistical significance of differences between all ligand responses except insulin was determined by ANOVA with Tukey's multiple comparison test. Note: a & d =  $p < 0.001$ ; c =  $p < 0.05$ ; e =  $p < 0.01$ . ns = not significant ( $p > 0.05$ ).



**Figure 5.15 IGF-IR competitive ligand binding assay of IGF-II analogues.**

**(A)** Plates with immunocaptured IGF-IR were treated with increasing concentrations of unlabelled ligands in the presence of europium-labelled insulin. Results are expressed as a percentage of europium labelled insulin bound in the absence of competing ligand and the data points are the mean  $\pm$  S.E.M of triplicate samples from at least three independent experiments. Error bars are shown when greater than the size of the symbols. **(B)** IGF-IR competitive binding at 3nM of ligand. The statistical significance of differences between all ligand responses except insulin was determined by ANOVA with Tukey's multiple comparison test. Note: c =  $p < 0.01$ ; a, d, e, f, g, h, & i =  $p < 0.001$ . ns = not significant ( $p > 0.05$ ).

#### 5.4.3.2 IGF-IR phosphorylation

At sub-nanomolar concentrations, IGF-I was the most potent activator of the IGF-IR, whereas at higher concentrations, the C1 ligand induced the greatest receptor phosphorylation with a maximum phosphorylation level of ~120% relative to IGF-II (Figure 5.16A). BCIAD elicited an intermediate activation response between IGF-I and IGF-II except at concentrations over 30nM where phosphorylation levels converged. The C2 ligand induced lower levels of IGF-IR activation than BCIAD at most concentrations contrasting with the receptor binding results where C2 bound the IGF-IR with an affinity equal to IGF-II (Figure 5.15A and Table 5.1). The C3 ligand was a relatively poor activator of the IGF-IR in line with the relatively low binding affinity of this ligand (Figure 5.15A and Table 5.1). The C3 response curve at concentrations of 10nM or less was only slightly less than that of IGF-II although the phosphorylation maximum of C3 at 100nM was approximately 25% less than the IGF-II maximum.

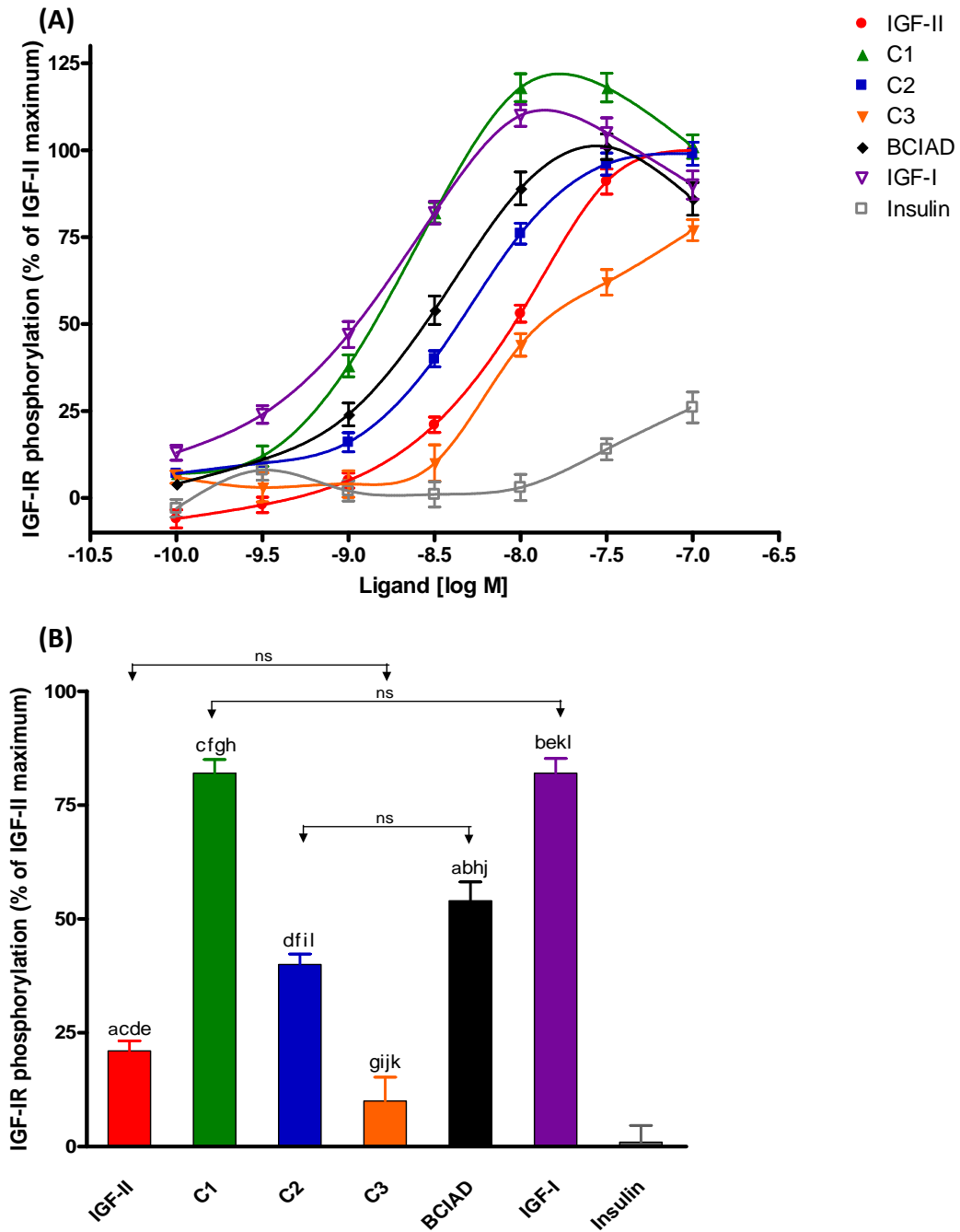
Comparison of IGF-IR activation at 3nM of ligand (Figure 5.16B) mirrored the overall IGF-IR phosphorylation response. The activation of the IGF-IR induced by the C1 and IGF-I ligands was ~3-fold higher than IGF-II induced phosphorylation. The BCIAD and C2 responses were ~2-fold greater than IGF-II. Both IGF-II and C3 induced similar levels of phosphorylation whilst insulin elicited the smallest activation of the IGF-IR.

#### 5.4.3.3 R IGF-IR cell survival

The IGF-IR cell survival assay indicated all ligands with the exception of insulin had a similar effect on cell survival at concentrations higher than 10nM (Figure 5.17A).

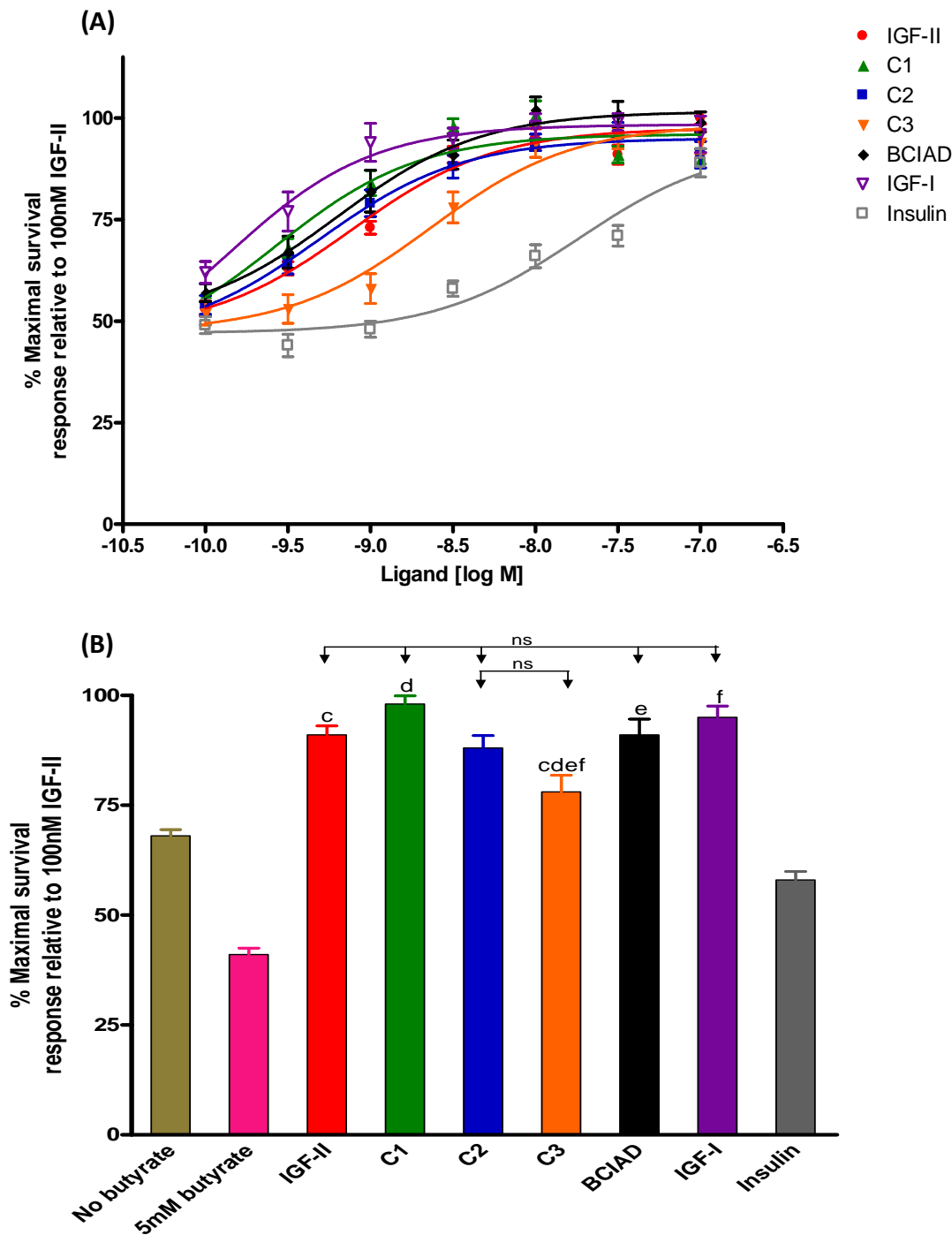
At lower concentrations, IGF-I had the greatest increase in cell survival. C1, BCIAD and C2 had intermediate responses between IGF-I and IGF-II which reflected the IGF-IR phosphorylation curves in Figure 5.15A. The C3 and insulin ligands had substantially lesser effects on cell survival than IGF-II.

At 3nM, there was no statistically significant ( $P>0.05$ ) difference in cell survival induced by the IGF-I, BCIAD, IGF-II, C1 and C2 ligands whereas C3 was significantly ( $p<0.01$ ) less potent than IGF-II (Figure 5.17B).



**Figure 5.16 IGF-IR phosphorylation by IGF-II analogues.**

**(A)** RIGF-IR cells expressing IGF-IR were stimulated with increasing concentration of ligand for 10 minutes. Cells were lysed and activated receptors were immunocaptured on 96-well plates and incubated with Eu-PY20 to detect phosphorylated tyrosines using time-resolved fluorescence. The data points are the mean  $\pm$  S.E.M of triplicate samples from at least three independent experiments. Error bars are shown when greater than the size of the symbols. **(B)** IGF-IR phosphorylation at 3nM of ligand. The statistical significance of differences between all ligand responses except insulin was determined by ANOVA with Tukey's multiple comparison test. Note: a, b, c, d, e, f, g, h, i, j, k & l =  $p < 0.001$ ; ns = not significant ( $p > 0.05$ ).



**Figure 5.17 RIGF-IR cell survival assay of IGF-II analogues.**

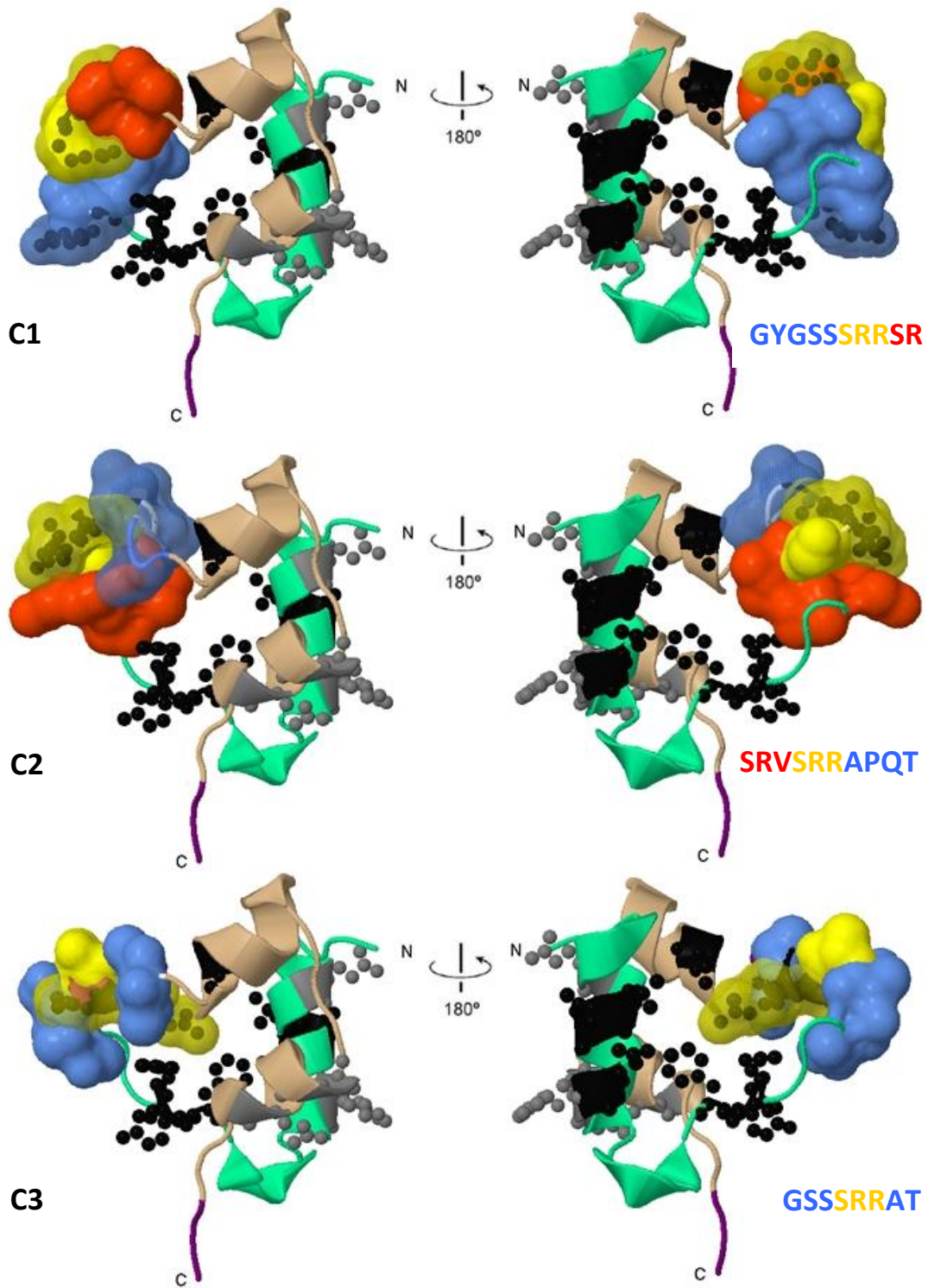
**(A)** RIGF-IR cells expressing IGF-IR were incubated with increasing concentration of ligand in the presence of 5mM butyrate. Cell viability was determined using the Promega Cell-titre Blue Reagent. Results are expressed as % survival relative to the maximum response obtained with 100nM of IGF-II. The data points are the mean  $\pm$  S.E.M of triplicate samples from at least three independent experiments. Error bars are shown when greater than the size of the symbols. **(B)** RIGF-IR survival at 3nM of ligand in the presence of 5mM butyrate. The statistical significance of differences between all ligand responses except insulin was determined by ANOVA Tukey's multiple comparison test. Note: c =  $p < 0.01$ ; d & f =  $p < 0.001$ ; e =  $p < 0.05$ . ns = not significant ( $p > 0.05$ ).









#### **5.4.4 Structural modelling of the IGF analogues**

Homology modelling is the process of generating three dimensional models of proteins based on structural templates obtained from related proteins with experimentally determined structures using high resolution techniques such as x-ray crystallography or nuclear magnetic resonance (NMR) spectroscopy (Bordoli *et al.*, 2009). The SWISS-PROT protein structure homology modelling server (<http://swissmodel.expasy.org>) was used to compare the predicted tertiary structures of the C1, C2 & C3 analogues with IGF-II, BCIAD and IGF-I (Figure 5.18).

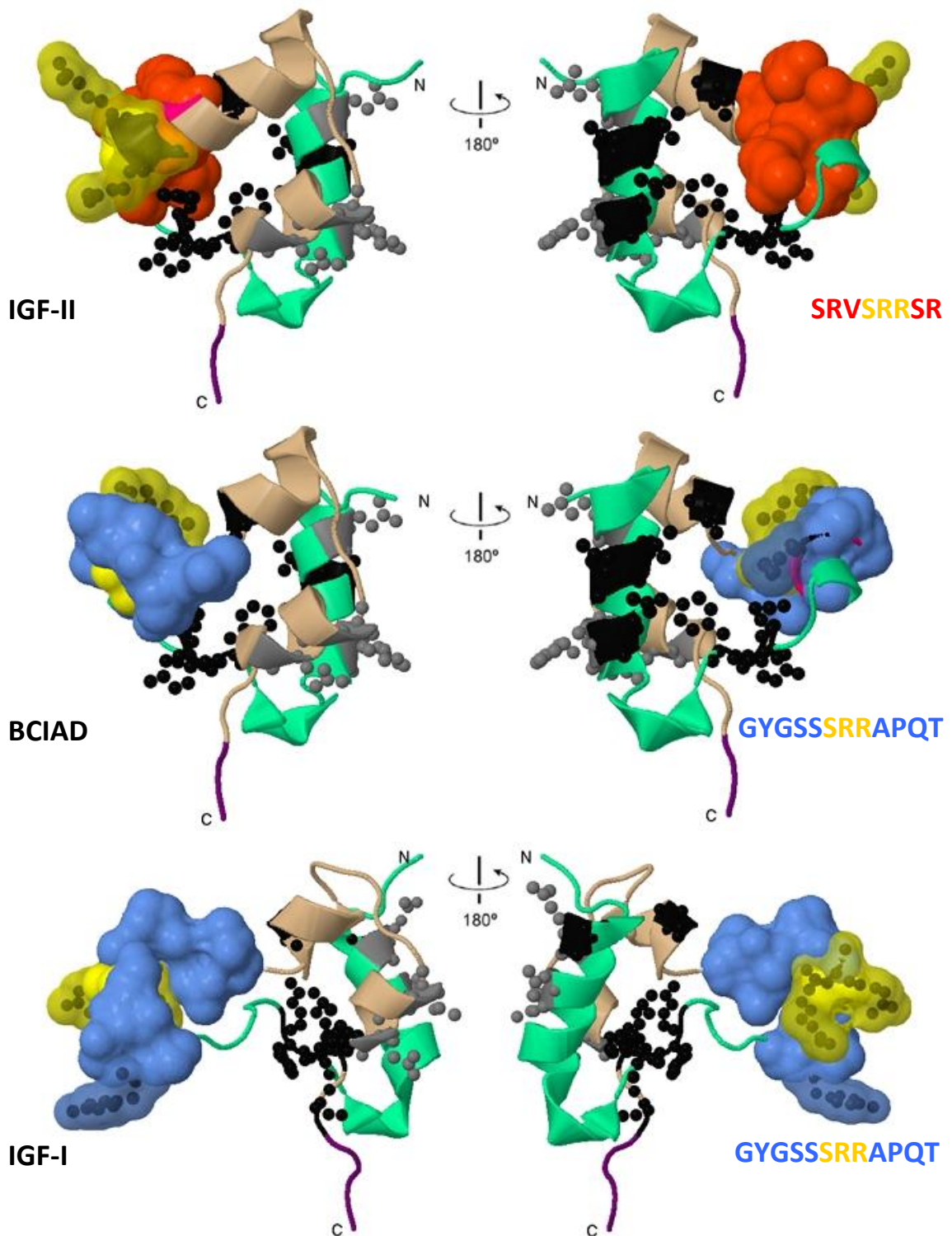
The models of IGF-II and the IGF-II analogues were constructed using coordinates from a crystal structure of human IGF-II complexed with a monoclonal antibody fragment (PDB id: 3kr3D) (Dransfield *et al.*, 2010) as a template. Similarly, the IGF-I model was built upon a structural template obtained using the crystal structure coordinates of human IGF-I ligand in the presence of the detergent SB12 (PDB id: 1gzs) (Brzozowski *et al.*, 2002). Both crystal structures were missing residues at the amino and carboxy-termini and the IGF-I structure was missing Arg36, Arg37 & Ala38 from the IGF-I C-domain. Thus, the IGF-II structure covered residues Ser5 to Ala64 and the IGF-I structure incorporated residues Glu3 to Ser35 and Pro39 to Pro66.

The IGF C-domains are omega loop motifs with the beginning and ends of the loop located close together. Accurately predicting the conformation of protein loops is difficult with comparative modelling due to the flexibility inherent in these structural motifs (Michalsky *et al.*, 2003). As such, the C-domains in the C1, C2, C3 and BCIAD structures obtained from the SWISS-PROT server were subsequently analysed with the ArchPRED server (<http://manaslu.aecom.yu.edu/loopred/>) using a fragment-search based method for predicting loop conformations (Fernandez-Fuentes *et al.*, 2006). Candidate loops were selected from the Search Space loop library and the loop conformations with the top ranking Z-scores (C1: 3.4; C2: 4.5; C3: 5.3 & BCIAD: 4.0) were incorporated into the models.



	B domain
	C domain (IGF-II residues)
	C domain (IGF-II/IGF-I conserved SRR residues)
	C domain (IGF-I residues)
	A domain
	D domain
	Binding Site 1 residues
	Binding Site 2 residues





**Figure 5.18 C-domain structural modelling comparisons between IGF-II, BCIAD, IGF-I and the IGF-II chimeras.**

Homology-based structural models of IGF-II, BCIAD & IGF-I (above) and C1, C2 & C3 (opposite) are shown in two orientations rotated 180° around the y-axis with C-domains depicted in surface mode and the B, A & D domains as ribbon cartoons. Colour coding is used to distinguish domains, to identify IGF-II, IGF-I and the conserved Ser-Arg-Arg residues within the C-domain. The IGF-II and IGF-I residues associated with binding Sites 1 & 2 are depicted as van der Waals spheres. Amino and carboxyl termini are labelled N and C, respectively. The C-domain residues are noted beside each structure. See section 5.4.4 in text for details of model construction.

Although homology modelling is a useful tool for visualizing the potential effects of mutation on the protein's tertiary structure, caution should be exercised when interpreting the IGF-II analogue models as there is currently no experimental structural data supporting the predicted C-domain conformations.

The IGF-II and IGF-I models show the difference in C-domain size due to the additional four amino acids present in the IGF-I C-domain (Figure 5.18). In contrast to IGF-II, the C-domain loop in IGF-I extends away from the core of the molecule. Interestingly, the IGF-I C-domain in the BCIAD model adopted a much more compact conformation and altered orientation to fit into the IGF-II molecule. As expected given their sequence diversity, the four IGF-II analogue models were predicted to adopt different C-domain conformations from each other and their parent molecules which could potentially disrupt or enable direct interactions between C-domain and receptor residues. In addition, local perturbations caused by the C-domain mutations could directly or indirectly affect residue interactions in the juxtaposed binding site 1 (Figure 5.18). It is also possible the mutants could alter the ligand-receptor docking orientation thereby having a global effect on the interaction between the receptor and unmutated ligand residues located away from the C-domain such as binding site 2 residues which are located on the opposite side of the molecule (Figure 5.18).

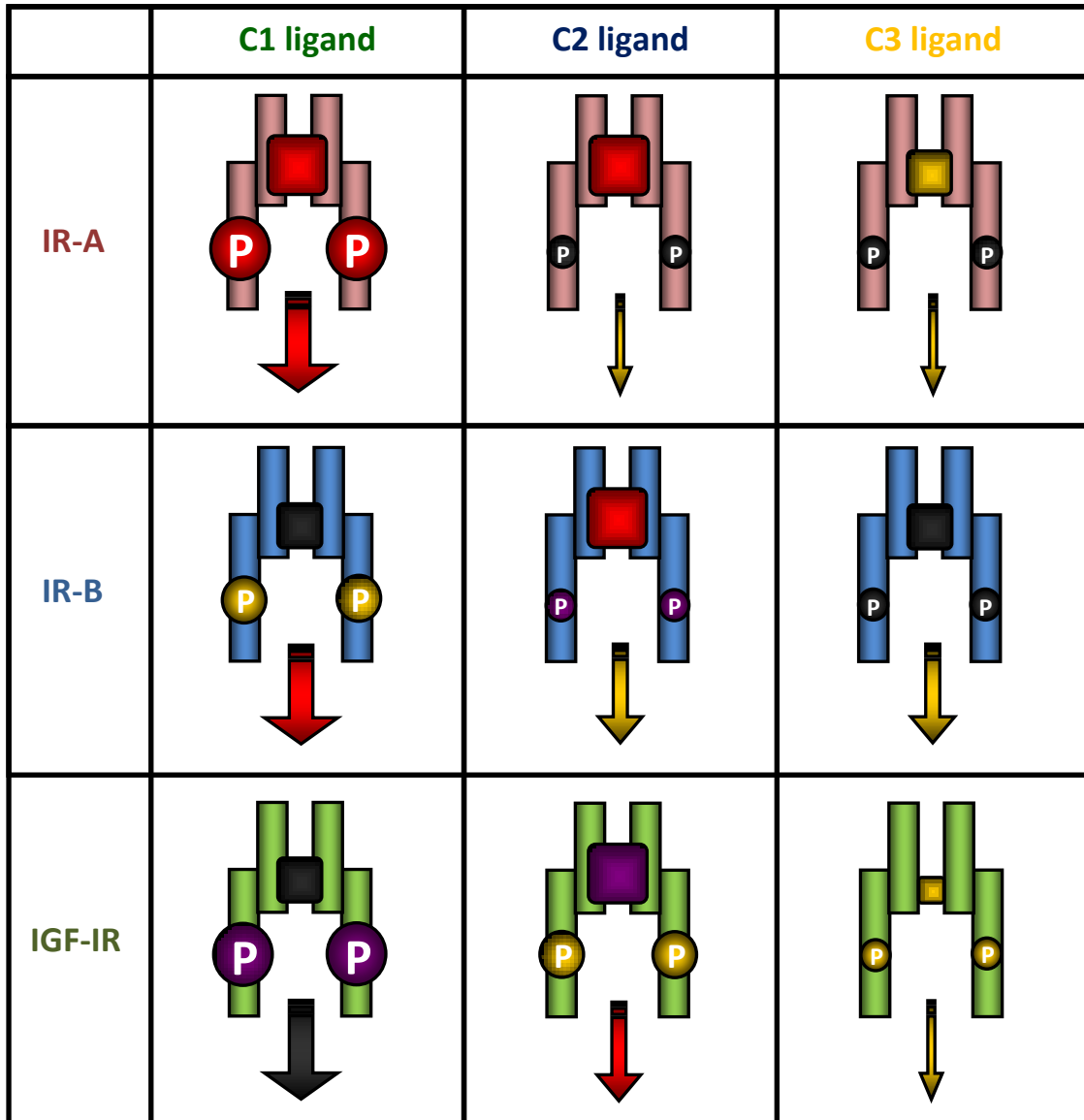
## **5.5 Discussion**

The aim of the work described within this chapter was to delineate the C-domain of the IGF ligands to identify the regions in IGF-II and BCIAD which conferred differential IR-A, IR-B and IGF-IR binding affinities and activation. For clarity, discussion of specific residues within the IGF analogue C-domains will refer to the position of equivalent residues within the parent IGF-I and IGF-II molecules. The overall receptor binding, activation and cell survival characteristics of the IGF-II analogues relative to the IGF-II, IGF-I and BCIAD control proteins are summarized in Figure 5.19.

The IR-A binding affinity, receptor activation and cell survival responses of C1 were comparable to those obtained with IGF-II. In contrast, C1 bound the IR-B isoform with BCIAD-like binding affinity but was more potent than BCIAD at inducing IR-B phosphorylation whilst the cell survival response was equivalent to that of IGF-II. The IGF-IR binding affinity of C1 was intermediate between the IGF-I and BCIAD responses. The IGF-IR phosphorylation and cell survival responses induced by C1 were equivalent to those of IGF-I.

The C2 analogue bound both the IR-A and IR-B with affinities comparable to that of IGF-II but acted as a partial agonist against both receptor isoforms with lower activation and cell survival responses similar to those obtained with BCIAD and IGF-I. Interestingly, C2 also acted as a partial agonist against the IGF-IR with a binding affinity equivalent to that of IGF-I yet receptor activation and cell survival responses were lower than those obtained with BCIAD.

The C3 analogue bound with low affinity to the IR isoforms and with poor affinity to the IGF-IR. Similarly, receptor phosphorylation and cell survival in response to C3 were relatively low with all three receptors.



Symbol key:

Low, medium & high receptor binding affinity of ligand

Low, medium & high receptor phosphorylation

Low, medium & high cell survival

Colour code key:

- Response similar to IGF-II control protein
- Response similar to BCIAD control protein
- Response similar to IGF-I control protein
- Response dissimilar to control proteins
- IR-A
- IR-B
- IGF-IR

**Figure 5.19 Comparison of the IGF-II analogue-mediated receptor binding, phosphorylation and cell survival responses through the IR-A, IR-B and IGF-IR.**

The C1, C2 & C3 responses obtained for the three assays are depicted in cartoon form to summarise their relative magnitude and similarity to the IGF control proteins. The receptor binding affinities of the IGF-II analogues are portrayed as squares binding to the receptor symbols. Receptor phosphorylation levels are shown as circles and downstream cell survival responses in butyrate-treated cells are represented as arrows. The relative magnitude of the responses are categorized as low, medium or high and are depicted by small, medium or large symbols, respectively. The symbols are colour coded to indicate similarity of the IGF-II analogue responses to those obtained with the IGF-II, BCIAD and IGF-I control proteins.

### **5.5.1 C1 interaction with IR-A and IR-B**

The C1 ( $IC_{50}$ :  $7.5 \pm 0.8nM$ ) mutant bound the IR-A with near IGF-II ( $IC_{50}$ :  $5.9 \pm 0.7nM$ ) affinity and elicited IGF-II like receptor phosphorylation and cell survival responses (Figure 5.19). The high IR-A binding affinity of C1 was most likely facilitated by one or more of the partial IGF-II C-domain residues (Ser36, Arg37, Arg38, Ser39 & Arg40). The BCIAD ligand contained the complete IGF-I C-domain which conferred lower IR-A binding affinity ( $IC_{50}$ :  $12.0 \pm 1.8nM$ ) than C1 and IGF-II. However, the high affinity binding of C1 suggests the partial IGF-I C-domain in C1 (Gly30, Tyr31, Gly32, Ser33 & Ser34) does not contain the elements responsible for the lower binding affinity of BCIAD. Indeed, the possibility of the partial IGF-I C-domain contributing to high affinity IR-A binding cannot be discounted as the contribution of the IGF-II and IGF-I residues within the chimeric C-domain cannot be delineated without further experimentation.

All of the IGF ligands exhibited  $\sim 1.5$  to  $3$ -fold lower binding affinities to the IR-B relative to the IR-A. In contrast, insulin had a  $2$ -fold increase in binding affinity to the IR-B compared to the IR-A. The  $\sim 3$ -fold decrease in the IR-B binding affinity of C1 relative to the IR-A was greater than the reduction observed with the other IGF ligands which ranged from  $\sim 1.5$  to  $\sim 2$ -fold. The C1 analogue had an IR-B binding affinity equivalent to BCIAD (C1  $IC_{50}$ :  $23.0 \pm 3.6nM$  vs. BCIAD  $IC_{50}$ :  $24.5 \pm 2.2nM$ ). This result suggested either one or more IGF-II C-domain residues absent in C1 (Ser35, Arg36 and Val37) contributed to IR-B binding or the presence of the partial IGF-I C-domain (Gly30, Tyr31, Gly32, Ser33 & Ser34) reduced the C1 binding affinity to the IR-B but not the IR-A. The latter possibility would presumably involve steric hindrance and/or incompatible sidechain interactions and could occur directly between the partial IGF-I C-domain and the additional 12 amino acids encoded by exon-11 in the IR-B. Alternatively, the C-domain may not be directly interacting with the exon-11 encoded peptide but may alter the orientation of C1 in the IR-B binding pocket resulting in unfavourable ligand-receptor interactions.

As mentioned previously, a single chain insulin with the IGF-I C-domain (Kristensen *et al.*, 1995) and an IGF-I mutant containing four B and A-domain substitutions with equivalent insulin residues (Gauguin *et al.*, 2008b) exhibited high affinity binding to the IR-A, suggesting the IGF-I C-domain is unlikely to sterically hinder binding to the IR-A. Although equivalent studies using the IR-B have not been reported, it is unlikely the IGF-I C-domain residues in C1 are incompatible or sterically hindering IR-B interaction as the complete IGF-I C-domain present within BCIAD would be expected to also result in a proportionally greater reduction in IR-B binding affinity. However, the *2-fold* reduction in IR-B binding affinity of BCIAD relative to the IR-A is proportional to the reduced affinity of IGF-II. Thus, the *3-fold* reduction in IR-B binding affinity exhibited by C1 is probably due to the absence of one or more of the IGF-II C-domain residues Ser35, Arg36 and Val37 and would imply these amino acids make a differential contribution to IGF-II binding to the IR-A and IR-B.

IGF-II, IGF-I and insulin have similar IR-A and IGF-IR binding surfaces although individual amino acids make variable contributions to ligand binding affinities (Alvino *et al.*, 2009). Similarly, certain amino acid sidechains in the insulin binding sites of the IR-A and IR-B make different energetic contributions to insulin binding (Whittaker *et al.*, 2002). The IR-B binding surfaces of the IGF-II ligand are likely to be comparable to the IR-A binding surfaces although it is probable there are differential amino acid interactions between IGF-II and the IR-A and IR-B. Therefore, disruption of specific sidechain interactions between ligand and receptor could have a negligible effect on IR-A binding but a more substantial effect on IR-B binding.

Interestingly, despite having equivalent binding affinities, C1 elicited greater IR-B phosphorylation and cell survival responses than BCIAD which is possibly due to C1 facilitating a conformational change within the receptor that is more amenable to kinase domain autophosphorylation.

Although the IR-B activation response to C1 was intermediate between IGF-II and BCIAD at concentrations above 10nM, the IR-B survival response to C1 was equivalent to IGF-II at all concentrations indicating that the magnitude of receptor

activation does not invariably result in a proportional effect on downstream pathways. This was also seen with insulin which had the highest IR-B binding affinity and activation yet produced the lowest cell survival response. The potential disconnect between receptor binding, activation and biological outcome was also evident with the relatively low IGF-I binding and activation of the IR-A and IR-B inducing moderate cell survival responses in both R<sup>-</sup>IR-A and R<sup>-</sup>IR-B cells. It has been shown that IGF-I can activate the IRS-2 signalling pathway in both IR isoforms without significant phosphorylation of the IR and can induce downstream actions such as protection from butyrate-induced apoptosis (Denley *et al.*, 2007). Thus, it is likely the same or a similar mechanism enabled IGF-I-induced cell survival in R<sup>-</sup>IR-A and R<sup>-</sup>IR-B cells in the presence of butyrate.

### **5.5.2 C2 interaction with IR-A and IR-B**

The C2 analogue bound the IR-A (IC<sub>50</sub>: 6.8 ± 1.3nM) and IR-B (IC<sub>50</sub>: 15.8 ± 2.3nM) with affinities similar to IGF-II (IR-A IC<sub>50</sub>: 5.9 ± 0.7nM and IR-B IC<sub>50</sub>: 11.5 ± 0.9nM) indicating the partial IGF-II C-domain in C2 (Ser33, Arg34, Val35, Ser36, Arg37 & Arg38) contains residues which confer high affinity binding to both the IR-A and IR-B. The presence of the partial IGF-I C-domain (Ala38, Pro39, Gln40 & Thr41) in C2 had no apparent detrimental effect on ligand binding to either IR isoform.

The similar IR-A binding affinities of C2, C1 and IGF-II indicate either flank of the IGF-II C-domain is sufficient for high affinity binding to the IR-A. The presence of either the N-terminal or C-terminal flanks of the IGF-I C-domain in C1 and C2 had no apparent negative influence on ligand binding to the IR-A. As mentioned previously, the entire IGF-I C-domain can be accommodated by the IR-A ligand binding pocket (Kristensen *et al.*, 1995); (Gauguin *et al.*, 2008b). Taken together, these findings suggest it was probably the absence of critical IGF-II elements that resulted in the lower binding of BCIAD to the IR-A relative to IGF-II rather than a direct incompatibility between IGF-I C-domain residues or steric hindrance of the bulkier IGF-I C-domain. It would be interesting to generate IGF-I counterparts to C1 & C2 with substitution of either flank of the IGF-I C-domain with equivalent IGF-II residues in the IGF-I ligand and determine



if these analogues exhibited IGF-II-like binding affinity to the IR-A and differential binding affinity to the IR-B.

The IR-A and IR-B phosphorylation and cell survival responses induced by C2 were similar to those obtained with BCIAD and were discordant with the high affinity binding of this IGF-II analogue to both IR isoforms indicating C2 is a partial agonist. The diminished tyrosine kinase activity of the IR isoforms in response to C2 could be due to IGF-II residues Ser39 and/or Arg40 being necessary to facilitate the correct conformational changes required for receptor activation. Alternatively, the presence of the partial IGF-I C-domain (Ala38, Pro39, Gln40 & Thr41) may have prevented the adoption of the necessary conformation for receptor phosphorylation.

Further delineation of IGF-II C-domain residues conferring high affinity binding and activation could be achieved by alanine scanning of the IGF-II C-domain flanks. The arginine residues in the C-domains of IGF-II and IGF-I are highly conserved in many vertebrate species (Table 5.2) implying these positively charged amino acids are likely to serve important but not necessarily the same functions. Given the inherent flexibility of the C-domain, it is unlikely these residues contribute to the overall stability of the ligand structure. Alanine substitution of Arg36 & Arg37 in IGF-I (corresponding to residues Arg37 & Arg38 in IGF-II) resulted in an *~30-fold* increase in binding affinity to the IR and a *15-fold* decrease in IGF-IR binding affinity (Zhang *et al.*, 1994). Furthermore, an IGF-II mutant with glutamine substitutions at Arg37 and Arg38 exhibited an *~3-fold* loss in IGF-IR and *~10-fold* loss in IR binding affinities (Edwards 1998). These studies demonstrated that the presence of charged residues in the IGF C-domain can have significant effects on receptor binding affinity, presumably due to electrostatic interactions between ligand and receptor. As such, the IGF-II Arg34 and Arg40 residues are prime candidates for alanine scanning mutagenesis both individually and together to determine if these residues contribute to IR binding and activation. Although not definitive, it is interesting to note the shortened IGF-I C-domain in the C3 analogue had neither Arg34 nor Arg40 and bound relatively poorly to the IR-A and IR-B.

CHAPTER 5 – IGF-II analogues

**IGF-I**

Human	1	10	20	30	40	50	60	70
Cow	1	10	20	30	40	50	60	70
Dog	1	10	20	30	40	50	60	70
Horse	1	10	20	30	40	50	60	70
Guinea Pig	1	10	20	30	40	50	60	70
Pig	1	10	20	30	40	50	60	70
Sheep	1	10	20	30	40	50	60	70
Deer	1	10	20	30	40	50	60	70
Goat	1	10	20	30	40	50	60	70
Rat	1	10	20	30	40	50	60	70
Mouse	1	10	20	30	40	50	60	70
Kangaroo	1	10	20	30	40	50	60	70
Chicken	1	10	20	30	40	50	60	70
Frog	1	10	20	30	40	50	60	70
Salmon / Trout	1	10	20	30	40	50	60	70
Eel	1	10	20	30	40	50	60	70
Carp	1	10	20	30	40	50	60	70
Zebrafish	1	10	20	30	40	50	60	70
Barramundi	1	10	20	30	40	50	60	70
Tuna / Flounder	1	10	20	30	40	50	60	70
Gilthead Bream	1	10	20	30	40	50	60	70
Tilapia	1	10	20	30	40	50	60	70
Catfish (brain)	1	10	20	30	40	50	60	70
Shark	1	10	20	30	40	50	60	70

**Domains:**



**IGF-II**

Human	1	10	20	30	40	50	60	67
Pig	1	10	20	30	40	50	60	67
Guinea Pig	1	10	20	30	40	50	60	67
Cow	1	10	20	30	40	50	60	67
Deer	1	10	20	30	40	50	60	67
Sheep	1	10	20	30	40	50	60	67
Mink	1	10	20	30	40	50	60	67
Rat	1	10	20	30	40	50	60	67
Mouse	1	10	20	30	40	50	60	67
Kangaroo	1	10	20	30	40	50	60	67
Chicken	1	10	20	30	40	50	60	67
Salmon / Trout	1	10	20	30	40	50	60	67
Eel	1	10	20	30	40	50	60	67
Gilthead Bream	1	10	20	30	40	50	60	67
Tilapia	1	10	20	30	40	50	60	67
Shark	1	10	20	30	40	50	60	67

**Table 5.2 IGF-I and IGF-II peptide sequence divergence between vertebrates.**

The IGF-I and IGF-II peptide sequences from numerous vertebrates are aligned with full stops indicating residues conserved with the human IGF sequence and asterixes indicating frame shifts required for maximal sequence alignment. The colour coded central bar denotes the IGF domain structure.

Figure adapted from the GroPep IGF comparative sequence chart ([http://www.gropep.com/index.php/filemanager/download/578/IGF\\_Sequence\\_Chart\(locked\).pdf](http://www.gropep.com/index.php/filemanager/download/578/IGF_Sequence_Chart(locked).pdf)). Details of the specific species and reference sources used for each group are available at this URL.

Also of interest in the context of potential C-domain sidechain interactions are the IGF-II residues Ser33 and Ser39 as these residues could potentially form hydrogen bonds due to the presence of a hydroxyl group in the serine sidechain. However, the relatively poor receptor binding affinity and activation characteristics of the C3 analogue suggests the serine residues are probably not the sole critical elements for IR binding and activation as the C3 mutant contains two serine residues and a threonine residue (also containing a hydroxylated sidechain capable of hydrogen bonding) in the N-terminal and C-terminal flanks of the C-domain, respectively.

### **5.5.3 C3 interaction with the IR-A and IR-B**

The design rationale for C3 was to create an IGF-I-like C-domain with the same number of amino acids as the IGF-II C-domain. The C3 C-domain lacks IGF-I Tyr31 and Pro39 as it had been postulated these residues may sterically hinder the IGF-I C-domain interaction with the IR isoforms (Denley *et al.*, 2004) although alanine substitution of Tyr31 in IGF-I had been previously shown to have little effect on IR binding (Bayne *et al.*, 1990). The C3 C-domain contains eight IGF-I C-domain residues with the exclusion of IGF-I amino acids Gly30, Tyr31, Pro39 & Gln40. The absence of IGF-II residues Arg34 & Arg40 was thought likely to provide an indication of the potential involvement of electrostatic interactions in IGF-II binding to the IR isoforms.

The C3 receptor binding, phosphorylation and cell survival responses to the IR-A and IR-B were generally similar to the BCIAD responses in these assays (Figure 5.19) supporting the notion that the size of the IGF-I C-domain is not solely responsible for the reduced IR affinity of BCIAD and IGF-I relative to IGF-II. In addition, the lower receptor binding and activation responses compared to IGF-II suggested there were important elements missing from the C3 C-domain, supporting the possibility that Arg34 and/or Arg40 may be important residues required for IGF-II-like binding and activation.

A potential issue with all mutagenesis studies is the inherent possibility that the effect of a mutation may not be due to direct disruption of the mutated amino acid but an

indirect artefact resulting from the mutation creating global or local intramolecular perturbations in the protein structure. Crystal structure analysis of IGF-I indicated the C-domain serves a minor structural role ensuring the Gly22 to Phe25  $\beta$ -strand adopts an IGF specific conformation whilst maintaining an 'insulin-like' fold across A domain residues Gly42 and Ile43 (Brzozowski *et al.*, 2002). These IGF-I residues are conserved in IGF-II and the C-domain may serve similar structural functions in IGF-II thereby illustrating how C-domain substitutions could potentially disrupt the IGF-II structure. IGF mutagenesis studies often utilize techniques such as circular dichroism spectroscopy (Gill *et al.*, 1996); (Magee *et al.*, 1999) or characterization of the interaction between IGF mutants and IGF binding proteins such as IGF-BP3 (Denley *et al.*, 2004) to verify the overall structure has retained native conformation. Unfortunately, experiments to verify protein structure prior to characterization were omitted due to project timeframe limitations. Thus, it is possible the relatively low receptor binding affinity and receptor activation characteristics of C3 could be due to local conformational perturbations. It is highly unlikely there are gross perturbations in the protein tertiary structure as C3 is generally as potent as BCIAD and always more potent than IGF-I at receptor binding, activation and induction of cell survival through the IR-A and IR-B.

#### **5.5.4 C1 and C2 interaction with the IGF-IR**

The IGF-IR binding affinities of C1 ( $IC_{50}$ :  $0.12 \pm 0.02nM$ ) and C2 ( $IC_{50}$ :  $0.10 \pm 0.03nM$ ) were similar to the binding affinity of IGF-I ( $IC_{50}$ :  $0.09 \pm 0.01nM$ ) and were slightly greater than BCIAD ( $IC_{50}$ :  $0.14 \pm 0.02nM$ ) and IGF-II ( $IC_{50}$ :  $0.19 \pm 0.02nM$ ). In relative terms, IGF-II binds with a 2.1-fold lower affinity than IGF-I with BCIAD, C1 & C2 having intermediate binding affinities within 2-fold of IGF-I and IGF-II. Differences in IGF receptor binding affinities of less than 2-fold are considered by some to be insignificant (Whittaker *et al.*, 2002); (Alvino *et al.*, 2009). However, the significance of the binding assay results obtained in this study is supported by the differential receptor activation clearly induced by the C1, C2 & BCIAD ligands. The C1 analogue exhibited IGF-I-like levels of IGF-IR activation whereas the phosphorylation level induced by C2 was substantially lower with an intermediate response between BCIAD

and IGF-II. The differences in IGF-IR phosphorylation levels were statistically significant ( $p < 0.001$ ) between C1 and C2, and between C1 and BCIAD at 3nM.

It cannot be determined from this study if the slightly increased IGF-IR binding affinities of C1 and C2 relative to BCIAD were due to the IGF-II or IGF-I C-domain residues or a combination of both. It is possible the presence of the entire native C-domain in IGF-II reduced the efficiency of ligand docking with the IGF-IR. For example, ligand binding affinity may be reduced by unfavourable electrostatic interactions between incompatible amino acid sidechains in the IGF-II C-domain and the binding pocket of the IGF-IR. The reduced number of IGF-II C-domain residues and/or the presence of IGF-I residues in C1 and C2 may partly, or fully, alleviate these constraints and enable higher affinity IGF-IR binding.

The slight increase in the IGF-IR binding affinity of C1 and C2 may be due solely to the presence of partial IGF-I residues within the C-domains. However, this seems unlikely to be the case for at least the C2 analogue as this mutant binds with IGF-I-like affinity yet lacks the Tyr31 residue which is required for IGF-I high affinity binding to the IGF-IR (Bayne *et al.*, 1990).

The IGF-I Arg36 & Arg37 C-domain residues are known to interact with the cysteine rich region of the IGF-IR (Zhang *et al.*, 1994). Despite the conservation of these arginine residues in IGF-II, IGF-IR binding studies using monoclonal antibodies with binding epitopes within the IGF-IR CR region showed impaired binding of IGF-I whilst IGF-II was unaffected (Keyhanfar *et al.*, 2007);(Surinya *et al.*, 2008). These studies demonstrated IGF-I and IGF-II utilized different modes of IGF-IR binding. Interestingly, BCIAD binding to the IGF-IR was also inhibited by the mABs which demonstrated the C-domain is responsible for the mode of binding employed by the IGFs (Keyhanfar *et al.*, 2007). Thus, characterizing C1 and C2 binding to the IGF-IR in the presence of these mABs could be very useful in determining if IGF-I, IGF-II or a combination of both C-domain elements are modulating ligand interaction with the IGF-IR.

The C1 and C2 ligands had greater IGF-IR binding affinities than BCIAD yet contained fewer IGF-I C-domain residues. Models of the BCIAD and IGF-I proteins have substantially different IGF-I C-domain conformations and highlight the possibility that substitution of the IGF-II C-domain with the larger IGF-I counterpart could significantly alter the conformation of the C-domain when in the IGF-II molecule. An altered C-domain configuration could lower binding affinity of BCIAD by various mechanisms such as interfering with the interactions between binding site residues or changing the orientation of the BCIAD ligand in the receptor binding pocket. It is possible the partial IGF-I C-domain residues in C1 and C2 permit a more optimal C-domain configuration for IGF-II binding to the IGF-IR. However, the predictive C-domain configurations in the IGF-II analogue models are based solely on computational analyses and the validity of the models cannot be substantiated without structural experimental data.

#### **5.5.5 C3 interaction with the IGF-IR**

The C3 analogue bound the IGF-IR with an ~5-fold lower affinity than IGF-II (C3 IC<sub>50</sub>: 0.91 ± 0.48nM vs. IGF-II IC<sub>50</sub>: 0.19 ± 0.02nM) and the IGF-IR phosphorylation and cell survival responses were relatively low with substantially less potency than IGF-II. These results suggest the C3 analogue is missing beneficial binding elements and indicates the conserved IGF-I Arg36 and Arg37 residues present in C3 are not sufficient to facilitate high affinity IGF-IR binding by themselves.

The C-domain residues in C3 are derived solely from IGF-I with omission of residues Gly30, Tyr31, Pro39 and Glu40 so as to contain the same number of amino acids as the IGF-II C-domain. The C3 ligand peptide sequence is most similar to BCIAD which has been shown to employ an IGF-I mode of IGF-IR binding as discussed previously. BCIAD has a higher IGF-IR binding affinity than IGF-II suggesting the presence of IGF-I residues and/or the absence of incompatible IGF-II C-domain amino acids enhances IGF-IR binding. However, the relatively poor IGF-IR binding affinity of C3 indicates it is unlikely the higher IGF-IR binding affinity of BCIAD is due solely to the absence of unfavourable IGF-II C-domain residues. The ~6.5-fold reduction in C3 affinity for the IGF-IR could be due solely to the absence of Tyr31 as alanine substitution of Tyr31 in

an IGF-I mutant was reported to have a similar *6-fold* reduction in IGF-IR binding affinity (Bayne *et al.*, 1990). However, the other absent IGF-I residues (Gly30, Pro39 and Glu40) may also be important for IGF-IR binding.

The IGF-I-like binding affinities of C1 and C2 indicate, at least in the context of the C-domains of these particular mutants, that IGF-I residues Gly30, Tyr31, Pro39 & Glu40 are not specifically required for high affinity IGF-IR binding of IGF-II analogues to the IGF-IR. As discussed earlier, the high IGF-IR binding affinities of C1 & C2 may reflect an IGF-II-like mode of binding with higher affinity due to the partial IGF-II C-domains. The possibility also exists that the presence of IGF-I residues Gly30 and/or Tyr31 in C1 and Pro39 and/or Glu40 in C2 (residues absent in C3) are sufficient by themselves or in conjunction with IGF-II C-domain residues to confer high affinity IGF-IR binding upon these IGF-II analogues.

As discussed in 5.5.3, it is possible the substitution of the shortened IGF-I C-domain in C3 created a local perturbation which impaired receptor binding and activation. It is unlikely the C3 protein had gross conformational alterations due to misfolding during the refolding step of the purification process as the IGF-IR binding affinity of misfolded IGF-II is typically *30-fold* lower (Delaine *et al.*, 2007) whereas C3 exhibited only a *5-fold* reduction in IGF-IR binding affinity relative to IGF-II.

#### **5.5.6 Comparison of assay results from this study with published reports**

There is substantial variation in competitive receptor binding assay formats and absolute receptor binding affinities of the IGFs reported in the literature. The different assay formats typically used in receptor binding studies include the use of intact cells or immunopurified receptors and use of iodine-125 (<sup>125</sup>I) or europium-labelled homologous or heterologous tracers. A study which directly compared the binding affinities of IGF-II and IGF-II analogues in IR-A and IGF-IR binding assays with whole cells and immunocaptured receptors indicated there was up to a *~10-fold* variation in binding affinities of ligands between the two types of assays (Alvino *et al.*, 2009). Whereas direct comparison of <sup>125</sup>I-labelled and europium-labelled IGF-I and IGF-II



tracers in immunopurified IGF-IR homologous competitive binding assays indicated there was little difference between the binding curves for each label (Surinya *et al.*, 2008). On this basis, the competitive receptor binding assay results obtained in this study were compared with published IGF binding studies using immunopurified receptors and europium or <sup>125</sup>I-labelled ligand tracers.

The IR-A and IR-B binding affinities for IGF-II, IGF-I and insulin from this study fell within the broad range reported in the literature (Table 5.3A). The IGF-IR binding affinities of the IGF-II, IGF-I and insulin ligands obtained in this study were ~2.5-fold greater than the highest binding affinities in published studies (Table 5.3A). However, there is an ~9-fold range of IGF-IR binding affinities reported in the literature for IGF-I and IGF-II. Importantly, the relative binding affinities of IGF-II, BCIAD, IGF-I and insulin for the IR-A, IR-B and IGF-IR in this study are similar to the trends reported in the literature (Table 5.3B).

The IGF-II, IGF-I, BCIAD & insulin-mediated IR-A, IR-B & IGF-IR phosphorylation responses were similar in potency and efficacy to those reported by (Denley *et al.*, 2004) with the exception of BCIAD activation of the IR-A and IR-B which have not been reported. Insulin was found to be more potent at activating the IR-B than the IR-A, a result consistent with previous observations (Kosaki *et al.*, 1995) (Denley *et al.*, 2004). The efficacy of IGF-I activation of the IR-A and IR-B relative to IGF-II was ~30% and ~50%, respectively, and is similar to the relative differences reported by (Denley *et al.*, 2004). The potency and efficacy of the IGF-IR phosphorylation response curves induced by IGF-II, BCIAD, IGF-I and insulin ligands are also comparable to the study reported by (Denley *et al.*, 2004).

(A)

Ligand	IR-A	IR-B	IGF-IR	Study reference
IGF-II	5.9	11.5	0.19	Current study
	4	29.7	3.4	Alvino <i>et al.</i> (2009) <sup>*</sup>
	-	-	0.49	Surinya <i>et al.</i> (2008) <sup>*</sup>
	2.2	10	-	Benyoucef <i>et al.</i> (2007) <sup>+</sup>
	18.2	68	4.4	Denley <i>et al.</i> (2004)
	0.9	11	0.6	Pandini <i>et al.</i> (2002) <sup>+</sup>
BCIAD	12.0	24.5	0.14	Current study
	66.3	310	1.5	Denley <i>et al.</i> (2004)
IGF-I	26.2	40.6	0.09	Current study
	-	-	0.29	Gaughin <i>et al.</i> (2008)
	-	-	0.25	Surinya <i>et al.</i> (2008)
	9	90	0.5	Benyoucef <i>et al.</i> (2007) <sup>+</sup>
	120.4	366	0.8	Denley <i>et al.</i> (2004)
	>30	>30	0.2	Pandini <i>et al.</i> (2002) <sup>+</sup>
Insulin	1.2	0.6	18	Current study
	0.25	0.51	>1000	Benyoucef <i>et al.</i> (2007) <sup>+</sup>
	2.8	1.4	>100	Denley <i>et al.</i> (2004)
	0.2	0.3	>30	Pandini <i>et al.</i> (2002) <sup>+</sup>

(B)

Ligand	IR-A		IR-B		IGF-IR	
	Rel. IC <sub>50</sub> (%) from current study	Rel. IC <sub>50</sub> (%) from Denley <i>et al.</i> (2004)	Rel. IC <sub>50</sub> (%) from current study	Rel. IC <sub>50</sub> (%) from Denley <i>et al.</i> (2004)	Rel. IC <sub>50</sub> (%) from current study	Rel. IC <sub>50</sub> (%) from Denley <i>et al.</i> (2004)
IGF-II	100	100	100	100	100	100
BCIAD	49	27	47	22	134	293
IGF-I	23	15	28	19	211	550
Insulin	492	654	1917	4857	1	<4

**Table 5.3 Comparison of the IR-A, IR-B and IGF-IR binding affinities of IGF-II, BCIAD, IGF-I and insulin obtained in this study with published data.**

(A) Comparison of IC<sub>50</sub> values (nM) as measured in immunocaptured receptor binding assays. All values were obtained with europium labelled ligand with the exception of studies marked with a <sup>+</sup> where <sup>125</sup>I was used as a label. Studies marked with a <sup>\*</sup> used Eu-IGF-II. (B) Receptor binding assay IC<sub>50</sub> values obtained in this study and the work of Denley *et al.* (2004) are presented as the relative binding affinity of each ligand compared to IGF-II (Rel. IC<sub>50</sub> (%) = IC<sub>50</sub> of IGF-II / IC<sub>50</sub> ligand x 100).

Assessing R<sup>1</sup>IR-A cell survival in the presence of butyrate has previously been used as a means of assessing a biological response to IGF analogues (Denley *et al.*, 2006). In this study, IGF-II and insulin were equipotent at inducing survival in butyrate treated R<sup>1</sup>IR-A cells and were more potent than IGF-I which had a greater survival response than BCIAD. In the work of (Denley *et al.*, 2006), IGF-II and insulin were also found to be more potent than BCIAD and IGF-I. However, BCIAD was reported to be more potent than IGF-I, whilst the IGF-II and insulin ligands were equipotent at concentrations up to 10nM after which the response curves diverged with the IGF-II maximum being ~20% greater than insulin. The Promega CellTiterBlue<sup>®</sup> cell viability assay used in this study is based on viable cells reducing the resazurin dye into the fluorescent compound resorufin (O'Brien *et al.*, 2000). In contrast, the study by (Denley *et al.*, 2006) utilized the Promega CellTiterGlo<sup>®</sup> assay which measured adenosine triphosphate (ATP) levels as an indicator of cell viability (Crouch *et al.*, 1993). Thus, it is possible the discrepancies between the survival assay results obtained in this study and those reported by Denley *et al.* (2006) may simply reflect a differential effect of the ligands on the two biomarkers used to assess cell viability.

Overall, the data obtained from the competitive receptor binding, receptor activation, and cell survival assays used in this study were generally consistent with other studies reported in the literature.

### **5.5.7 Protein structure modeling**

Homology modelling of the IGF ligands used in this study were undertaken to visualize the potential effects of the introduced mutations on the ligand tertiary structure. The models of BCIAD and IGF-I predicted a substantial difference in C-domain conformation between these two ligands (Figure 5.18) which could play a role in the greater binding affinity of BCIAD to the IR-A & IR-B relative to IGF-I. Differences in C-domain conformation may also contribute fully or partially to the increased receptor binding affinity of C1 and C2 relative to BCIAD.

There are NMR-derived structures for IGF-II (Torres *et al.*, 1995) and IGF-I (Cooke *et al.*, 1991); (Sato *et al.*, 1993). However, the conformational flexibility in the C and D domains limit definition in these regions of the NMR structures thereby limiting their usefulness as templates for homology modelling. As such, crystal structures of IGF-II and IGF-I were used as they had better defined C and D-domains. The crystal structure of IGF-II was obtained with IGF-II bound to the parental Fab of the DX-2647 antibody (Dransfield *et al.*, 2010) whereas IGF-I crystals were obtained using SB12 detergent as an additive (Brzozowski *et al.*, 2002). However, these ligand complexes could have potentially altered the native conformation of the IGF-II and IGF-I proteins. In addition, the structures adopted in the solid-state during crystallization may also have been influenced by crystal packing effects (Tiekink 2000).

As stressed previously, the predictive nature of the homology models are not supported by structural data and should be interpreted with caution as there are potential inaccuracies in modelling the inherently flexible IGF C-domain loops.

### **5.5.8 Future work**

The generation and preliminary characterization of the IGF-II analogues C1, C2 & C3 described herein provides the basis for a number of areas of further investigation.

#### **5.5.8.1 Confirmation of structural integrity in IGF-II analogues**

The relatively poor receptor binding affinity of C3 could be due to structural perturbations in the ligand. Far-UV circular dichroism spectroscopy analysis and/or characterization of the interaction between C3 and IGF binding proteins could be used to verify the structural integrity of the IGF-II analogues. The latter approach has the added benefit of characterizing the interaction between the IGF-II analogues and IGF binding proteins.

#### **5.5.8.2 Kinetic analysis of the ligand-receptor interaction**

Surface plasmon resonance analysis of the interaction between the C1, C2 & C3 ligands and the IGF-IR and IR isoforms would enable real time characterization of binding kinetic parameters such as association and dissociation rate constants.

#### **5.5.8.3 Additional IGF ligand mutants**

Identifying the C1 & C2 C-domain residues involved in receptor binding and activation would provide useful insights into the binding mechanism of IGF-II. Furthermore, elucidating how C2 acts as a partial agonist against the IR isoforms and the IGF-IR could be particularly useful for the design of receptor antagonists. The competitive receptor binding assays of C1 & C2 suggest either IGF-II C-domain flank is sufficient for high affinity binding to the IR-A, IR-B and IGF-IR. Mutagenesis of residues within either flank of C1, C2 & IGF-II could be used to delineate the C-domain and identify residues which are major contributors to receptor binding. In particular, the charged IGF-II Arg34 and Arg40 residues could be involved in electrostatic interactions which could be determined by alanine scanning mutagenesis of these amino acids both individually and together.

An IGF-I analogue with substitution of the IGF-II C-domain into the IGF-I ligand increased binding affinity to the IR-A and IR-B whilst simultaneously decreasing binding affinity to the IGF-IR (Denley *et al.*, 2004). Generating IGF-I counterparts to C1 and C2 with partial IGF-II C-domains substituted into the IGF-I ligand could prove useful in delineating IGF elements that contribute to receptor specificity.

#### **5.5.8.4 IGF-II analogue characterisation with monoclonal antibodies**

The precise mechanism by which the IGF C-domains influence receptor binding affinity is unknown although binding studies with monoclonal antibodies 7C2 and 9E11 have demonstrated the IGF-I C-domain interacts with the cysteine rich region of the IGF-IR whereas the IGF-II ligand does not (Keyhanfar *et al.*, 2007). Using these monoclonal antibodies in binding studies of C1 & C2 could indicate if the high IGF-IR binding affinities of these ligands are due to an IGF-I or IGF-II mode of interaction.

#### **5.5.8.5 Characterisation of IGF-II analogues with receptor mutants**

Numerous studies have mapped ligand binding sites using an extensive collection of alanine scanned mutants of the IR-A, IR-B and IGF-IR (Mynarcik *et al.*, 1997); (Whittaker *et al.*, 2001); (Whittaker *et al.*, 2002); (Sorensen *et al.*, 2004); (Whittaker *et al.*, 2008). Binding studies with these mutant receptors and the C1 & C2 analogues would be useful in elucidating the receptor regions involved in ligand interaction.

#### **5.5.8.6 Delineation of signalling pathways**

IGF-II but not IGF-I was found to induce potent autophosphorylation of specific tyrosine residues in both IR isoforms and to induce the downstream signalling pathways IRS-1 and Akt/PKB in the IR-A (Denley *et al.*, 2006). Interestingly, the C-domain was found to be responsible for this differential receptor activation. IGF-II and IGF-I analogues containing the IGF-II C-domain were generally more potent at inducing autophosphorylation of specific IR-A and IR-B tyrosine residues relative to IGF-I and IGF-II analogues with the IGF-I C-domain (Denley *et al.*, 2006). Given the importance of the C-domain in determining the downstream signalling properties of the IGFs, similar studies investigating signalling pathways induced by C1 & C2 could provide valuable insights into how the C-domain modulates this differential signalling response.

#### **5.5.9 Conclusion**

This study aimed to identify regions within the IGF C-domain responsible for differential receptor binding and activation, with a particular focus on the IGF-II ligand. The results presented in this chapter indicate the C-domain modulation of receptor binding is complex with both flanks of the IGF C-domain playing roles in IR-A, IR-B and IGF-IR interaction. The data obtained in this study contribute to the delineation of the IGF C-domain and the ongoing elucidation of how the C-domain modulates IGF-receptor interaction. Ultimately, a comprehensive understanding of this mechanism may facilitate the rational design of receptor antagonists with tailored modulation of downstream effects which would have particular relevance to the treatment of various cancers.

## **Final discussion**

The work described within this thesis investigated the application of directed evolution and rational site-directed mutagenesis to generate human insulin-like growth factor mutants with the ultimate aim of producing variants acting as IGF-IR/IR antagonists.

The diversification of human IGF-I was undertaken as part of the ongoing development of the *Neurospora* directed evolution platform. The primary advantage of directed evolution is that no mechanistic understanding of the interaction between ligand and receptor is required to generate antagonists. The criteria for directed evolution are for the protein function to be physically achievable and obtainable through evolution, the system must be capable of generating a diverse library of variants and the screening assay needs to be robust and sensitive. The advantage of using IGF-I as the starting point for directed evolution is that the ligand binds the IGF-IR with high affinity and specificity, therefore, relatively small changes to the IGF-I molecule through directed evolution could be reasonably expected to generate IGF-IR antagonists. The RIP process in *Neurospora* was shown to be effective at inducing mutation in the IGF-I coding sequence, demonstrating the potential of RIP to generate a diverse library of IGF-I variants. RIP is not a completely random process with a strong preference for targeting CpA dinucleotides. The re-engineering of the IGF-I coding sequence was undertaken to increase the number of CpA targets and facilitate a broader distribution of mutations throughout the IGF-I sequence. A previous study had utilised HSMR for gene diversification in *Neurospora* {Catcheside, 2003 #273}. Although not investigated in this study, shuffling of RIPed sequences with HSMR would be expected to make an important contribution to generating diversity and by separating beneficial and deleterious mutations. RIP and HSMR have a demonstrated potential in generating diverse sequence libraries in *Neurospora*.

A requirement for the directed evolution of IGF-I in *Neurospora* was to express and secrete IGF-I protein so as to enable screening of IGF-I variants in culture supernatants. IGF-I was successfully secreted into the culture medium although the yield of mature IGF-I protein was low, due at least in part, to the presence of

extracellular proteases. A total of five different expression vectors were constructed and the *Neurospora* culture conditions were optimised in unsuccessful attempts to increase mature IGF-I yields to enable purification and characterisation of the recombinant protein.

High-throughput screening with a robust and sensitive assay is a fundamental requirement and usually the single greatest limitation in directed evolution. The IGF-IR binding assay for initial screening of IGF-I variants was originally developed for use with purified proteins and was found to be adversely affected by the standard *Neurospora* culture medium. Although a more compatible medium was identified, the receptor binding assay was deemed to be unacceptable for screening due to the high degree of variation when assaying IGF-I protein in *Neurospora* culture supernatants. Screening of IGF-I mutants would likely be best achieved through surface display in *Neurospora*. A display system could be developed by fusing IGF-I variants to a conidium surface protein and would enable the use of FACS to facilitate high-throughput screening and would likely circumvent degradation issues with extracellular proteases as fungal liquid cultures would not be required for assaying secreted protein.

Whilst this study was being undertaken, Neugenesis Corporation was engineering improved expression strains and had done some preliminary investigation into developing a surface display system for *Neurospora*. Consequently, the project focus was shifted to avoid duplication of research effort. However, the development of a surface display system was not pursued by Neugenesis due to a subsequent change in the direction of the company.

The work described in this thesis has demonstrated *Neurospora* is capable of heterologous expression and secretion of human IGF-I, and that RIP has the potential to generate a diverse library of IGF-I sequences. The development of *Neurospora* protease-deficient strains and a surface display system would likely overcome the limiting issues identified in this study and permit directed evolution of IGF-I with the *Neurospora* system.



Rational design is theoretically the most direct way of generating a receptor antagonist but typically requires extensive knowledge of the interaction between ligand and receptor. Our understanding of the mechanisms by which the IGFs interact with receptors is gradually becoming clearer through a variety of techniques including biochemical studies of ligands and receptors mutated in key regions. This study used a rational site-directed mutagenesis approach to delineate important regions of the IGF-II C-domain modulating differential receptor binding and activation.

Three IGF-II analogues (C1, C2 & C3) were designed with C1 & C2 containing partial C-domain swaps with IGF-I and C3 containing a shortened IGF-I C-domain. The coding sequences of the IGF-II mutants were constructed by site-directed mutagenesis and were cloned into an *E. coli* expression vector. Recombinant protein was purified from *E. coli* host cells and characterised by IR-A, IR-B and IGF-IR binding assays, phosphorylation assays and butyrate-treated cell survival assays.

The C2 analogue was found to be a partial agonist of the IR-A, IR-B and IGF-IR with high receptor binding affinities but relatively low receptor activation. The partial agonist properties of C2 suggested the C-domain of this analogue permitted high affinity binding to all three receptors but did not allow the conformational changes required for proportional receptor activation. Further investigation and characterisation of this mutant is likely to provide informative insights into how an antagonist might be generated through rational design. Furthermore, this mutant could be useful as a starting point for directed evolution to enhance receptor binding affinity and/or further decrease receptor activation potency.

The C1 analogue bound and activated the IR-A with IGF-II-like characteristics but bound the IR-B with a lower, BCIAD-like affinity and activated the IR-B to an intermediate level between BCIAD and IGF-II. The characterisation of the C1 interaction with the IR-B suggested the C-domain of this analogue either lacked IGF-II residues which contributed to IR-B binding, contained IGF-I residues incompatible with the IR-B or the C-domain of this analogue adversely affected the alignment of important residues or the entire ligand in the IR-B binding pocket. The results obtained for the C1 and C2 analogues indicate both flanks of the IGF C-domain play important

roles in receptor binding and activation. Alanine scanning could be used to identify the key residues within both flanks of the C-domain.

The C3 analogue was found to have relatively poor binding and activation of all three receptors. C3 contained a shortened IGF-I C-domain suggesting the larger physical size of the IGF-I C-domain was unlikely to be solely responsible for the reduced IR-A and IR-B binding affinity of IGF-I and BCIAD relative to IGF-II. Rather, it appears likely the C-domain of C3 lacked positively charged residues important for IR-A and IR-B binding.

The results of this study have provided a number of insights into the contributions made by the C-domain flanks in modulating IGF-II binding and activation of the IR-A, IR-B and IGF-IR. The work described herein provides the basis for future investigation involving additional characterisation of the C1, C2 & C3 analogues and generation of additional C-domain mutants to further elucidate the mechanism by which the C-domain modulates IGF-receptor interactions.

### **Final conclusion**

The primary objective of this thesis was to utilise *Neurospora* for the directed evolution of IGF-I. This study has demonstrated the *Neurospora* gene diversification system is capable of human IGF-I heterologous expression and secretion, and the endogenous RIP process can be utilised to generate a library of IGF-I variants. However, improved expression strains and development of a surface display system is required to fully realise the potential of *Neurospora* for *in vivo* directed evolution of human IGF-I.

The other aim of this thesis was to use rational site-directed mutagenesis to identify C-domain residues within IGF-II required to modulate receptor binding and activation specificity. The biochemical characterisation of three IGF-II analogues has indicated both flanks of the IGF-II C-domain play important roles in the interaction with the IR-A, IR-B and IGF-IR. In addition, the C2 mutant was found to act as a partial agonist and further characterisation of this mutant could provide useful insights into generating an IGF-IR/IR antagonist.

## Appendix

**Table A.1 IGF-I amino acid sequences from various organisms used for SIFT analysis (refer 4.3.4.4.2)**

Swiss prot entry	Organism (common)	Organism (scientific)	IGF-I sequence
Q14620	Human	<i>Homo sapiens</i>	GPETLLCGAELVDALQIFVCGDRGFYFNKPTGYGSSSRRAPQTGIVDECCFRSCDLRRLLEMYCAPLKPAAKSA
Q95222	Rabbit	<i>Oryctolagus cuniculus</i>	GPETLLCGAELVDALQIFVCGDRGFYFNKPTGYGSSSRRAPQTGIVDECCFRSCDLRRLLEMYCAPLKPAAKAA
Q0PPT2	Pig	<i>Sus scrofa</i>	GPETLLCGAELVDALQIFVCGDRGFYFNKPTGYGSSSRRAPQTGIVDECCFRSCDLRRLLEMYCAPLKPAAKSA
Q9NIC1	Cow	<i>Bos taurus</i>	GPETLLCGAELVDALQIFVCGDRGFYFNKPTGYGSSSRRAPQTGIVDECCFRSCDLRRLLEMYCAPLKPAAKSA
P17647	Guinea pig	<i>Cavia porcellus</i>	GPETLLCGAELVDALQIFVCGDRGFYFNKPTGYGSSSRRAPQTGIVDECCFRSCDLRRLLEMYCAPLKPAAKSA
A0N8V2	Sheep	<i>Ovis aries</i>	GPETLLCGAELVDALQIFVCGDRGFYFNKPTGYGSSSRRAPQTGIVDECCFRSCDLRRLLEMYCAPLKPAAKSA
Q6IVA5	Red panda	<i>Ailurus fulgens</i>	GPETLLCGAELVDALQIFVCGDRGFYFNKPTGYGSSSRRAPQTGIVDECCFRSCDLRRLLEMYCAPLKPAAKSA
Q68IC0	Golden snub-nosed monkey	<i>Pygathrix roxellana</i>	GPETLLCGAELVDALQIFVCGDRGFYFNKPTGYGSSSRRAPQTGIVDECCFRSCDLRRLLEMYCAPLKPAAKSA
Q5RK13	Rat	<i>Rattus norvegicus</i>	GPETLLCGAELVDALQIFVCGDRGFYFNKPTGYGSSIRRAPQTGIVDECCFRSCDLRRLLEMYCAPLKPAAKSA
Q6JLX1	Giant Panda	<i>Ailuropoda melanoleuca</i>	GPETLLCGAELVDALQIFVCGDRGFYFNKPTGYGSSSRRAPQTGIVDECCFRSCDLRRLLEMYCAPLKPAAKSA
Q6GUL6	Siberian tiger	<i>Panthera tigris altaica</i>	GPETLLCGAELVDALQIFVCGDRGFYFNKPTGYGSSSRRAPQTGIVDECCFRSCDLRRLLEMYCAPLKPAAKSA
P33712	Dog	<i>Canis familiaris</i>	GPETLLCGAELVDALQIFVCGDRGFYFNKPTGYGSSSRRAPQTGIVDECCFRSCDLRRLLEMYCAPLKPAAKSA
Q4VJB9	Mouse	<i>Mus musculus</i>	GPETLLCGAELVDALQIFVCGDRGFYFNKPTGYGSSIRRAPQTGIVDECCFRSCDLRRLLEMYCAPLKPAAKSA
P51457	Goat	<i>Capra hircus</i>	GPETLLCGAELVDALQIFVCGDRGFYFNKPTGYGSSSRRAPQTGIVDECCFRSCDLRRLLEMYCAPLKPAAKSA
Q000X1	Red deer	<i>Cervus elaphus</i>	GPETLLCGAELVDALQIFVCGDRGFYFNKPTGYGSSSRRAPQTGIVDECCFRSCDLRRLLEMYCAPLKPAAKSA
O93380	Turkey	<i>Meleagris gallopavo</i>	GPETLLCGAELVDALQIFVCGDRGFYFNKPTGYGSSSRRLHKKGIVDECCFQSCDLRRLLEMYCAPIKPPKSA
Q14WA7	Goose	<i>Anser anser</i>	GPETLLCGAELVDALQIFVCGDRGFYFNKPTGYGSSSRRLHKKGIVDECCFQSCDLRRLLEMYCAPIKPPKSA
P18254	Chicken	<i>Gallus gallus</i>	GPETLLCGAELVDALQIFVCGDRGFYFNKPTGYGSSSRRLHKKGIVDECCFQSCDLRRLLEMYCAPIKPPKSA
P51462	Japanese quail	<i>Coturnix coturnix japonica</i>	GPETLLCGAELVDALQIFVCGDRGFYFNKPTGYGSSSRRLHKKGIVDECCFQSCDLRRLLEMYCAPIKPPKSA

Appendix

P16501	African clawed frog	<i>Xenopus laevis</i>	GPETLLCGAELVDTLLQFVCCGDRGFYFSKPTGYGSSNNRRSHHNGIIVDECCFQSCDFRRLLEMYCAPAKQAKSA
P51458	Horse	<i>Equus caballus</i>	GPETLLCGAELVDALQFVCCGDRGFYFNKPTGYGSSSRRAPQGTIVDECCFRSCDLRRLLEMYCAPLKPAKSA
Q5BQH3	Scale-less Carp	<i>Gymnocypris przewalskii</i>	GPETLLCGAELVDTLLQFVCCGDRGFYFSKPTGYGSSRRSHNRGIVDECCFQSCDLRRLLEMYCAPVKPKGKSP
Q1PCW8	Fish	<i>Spinibarbus sinensis</i>	GPETLLCGAELVDTLLQFVCCGDRGFYFSKPTGYGSSRRSHNRGIVDECCFQSCDLRRLLEMYCAPVKPKGKTP
Q9IBI0	Common carp	<i>Cyprinus carpio</i>	GPETLLCGAELVDTLLQFVCCGDRGFYFSKPTGYGSSRRSHNRGIVDECCFQSCDLRRLLEMYCAPVKPKGKTP
Q02815	Rainbow trout	<i>Oncorhynchus mykiss</i>	GPETLLCGAELVDTLLQFVCCGDRGFYFSKPTGYGSSRRSHNRGIVDECCFQSCDLRRLLEMYCAPVKSSGKAA
Q28933	Musk shrew	<i>Suncus murinus</i>	GPETLLCGAELVDALQFVCCGDRGFYFNKPTGYGSSSRRAPQGTIVDECCFRSCDLRRLLEMYCAPLKPAPAKAA
Q90VY9	Zebrafish	<i>Danio rerio</i>	GPETLLCGAELVDTLLQFVCCGDRGFYFSKPTGYGSSRRSHNRGIVDECCFQSCDLRRLLEMYCAPVKTKGKSP
Q506L8	Mud carp	<i>Cirrhinus molitorella</i>	GPETLLCGAELVDTLLQFVCCGDRGFYFSKPTGYGSSRRSHNRGIVDECCFQSCDLRRLLEMYCAPVKPKGKTP
Q98SR6	Bream carp	<i>Megalobrama amblycephala</i>	GPETLLCGAELVDTLLQFVCCGDRGFYFSKPTGYGSSRRSHNRGIVDECCFQSCDLRRLLEMYCAPVKTKGKTP
Q91476	Atlantic salmon	<i>Salmo salar</i>	GPETLLCGAELVDTLLQFVCCGERGFYFSKPTGYGSSRRSHNRGIVDECCFQSCDLRRLLEMYCAPVKSSGKAA
Q91965	King salmon	<i>Oncorhynchus tshawytscha</i>	GPETLLCGAELVDTLLQFVCCGERGFYFSKPTGYGSSRRSHNRGIVDECCFQSCDLRRLLEMYCAPVKSSGKAA
P17085	Coho salmon	<i>Oncorhynchus kisutch</i>	GPETLLCGAELVDTLLQFVCCGERGFYFSKPTGYGSSRRSHNRGIVDECCFQSCDLRRLLEMYCAPVKSSGKAA
Q6V293	Channel catfish	<i>Ictalurus punctatus</i>	GPETLLCGAELVDTLLQFVCCGDRGFYFSKPTGYGPNRRLHNRGIVDECCFQSCDLRRLLEMYCAPVKSSGKAP
Q9I9I4	Grass carp	<i>Ctenopharyngodon idella</i>	GPETLLCGAELVDTLLQFVCCGDRGFYFSKPTGYGSSRRSHNRGIVDECCFQSCDLRRLLEMYCAPVKTKGKSP
Q9YI82	Goldfish	<i>Carassius auratus</i>	GPETLLCGAELVDTLLQFVCCGDRGFYFSKPTGYGPNRSHNRGIVDECCFQSCDLRRLLEMYCAPVKPKGKTP
Q800D5	Black Amur bream	<i>Megalobrama terminalis</i>	GPETLLCGAELVDTLLQFVCCGDRGFYFSKPTGYGSSRRSHNRGIVDECCFQSCDLRRLLEMYCAPVKTKGKTP
Q27YE0	Sterlet	<i>Acipenser ruthenus</i>	GSETLLCGAELVDTLLQFVCCGERGFYFNKPTGYGASSRRPHHNGIIVNECCFQSCDLRRLLEMYCAPVKPAKAS
Q3HRT3	Caspian sturgeon	<i>Acipenser fuldenstadti</i>	GSETLLCGAELVDTLLQFVCCGERGFYFNKPTGYGPATRRPHHNGIIVNECCFQSCDLRRLLEMYCAPVKPAKAS

Appendix

Q93527	Japanese flounder	<i>Paralichthys olivaceus</i>	GPETLLCGAELVDTLLQFVCCGERGFYFSKPTGYGNARRS--RGI VDECCFQSCELRRLLEMYCAPAKTSKAA
Q5U8T6	Orange- spotted grouper	<i>Epinephelus coiodes</i>	GPETLLCGAELVDTLLQFVCCGERGFYFSKPTGYGNARRS--RGI VDECCFQSCELRRLLEMYCAPAKTSKAA
Q70BT6	Perch	<i>Perca fluviatilis</i>	GPETLLCGAELVDTLLQFVCCGERGFYFSKPTGYGNARRS--RGI VDECCFQSCELRRLLEMYCAPAKTSKAA
Q800Y5	Yellow perch	<i>Perca flavescens</i>	GPETLLCGAELVDTLLQFVCCGERGFYFSKPTGYGNARRS--RGI VDECCFQSCELRRLLEMYCAPAKTSKAA
Q29U99	Rabbitfish	<i>Siganus guttatus</i>	GPETLLCGAELVDTLLQFVCCGERGFYFSKPTGYGNARRS--RGI VDECCFQSCELRRLLEMYCAPAKTSKAA
Q6T8S2	Shi drum	<i>Umbrina cirrosa</i>	GPETLLCGAELVDTLLQFVCCGERGFYFSKPTGYGNARRS--RGI VDECCFQSCELRRLLEMYCAPAKTSKAA
Q3YKS1	Flathead mullet	<i>Mugil cephalus</i>	GPETLLCGAELVDTLLQFVCCGERGFYFSKPTGYGNARRS--RGI VDECCFQSCELRRLLEMYCAPAKTNSX
Q0Q005	European sea bass	<i>Dicentrarchus labrax</i>	GPETLLCGAELVDTLLQFVCCGERGFYFSKPTGYGNARRS--RGI VDECCFQSCELRRLLEMYCAPAKTGTKAA
Q2TU60	Largemouth bass	<i>Micropterus salmoides</i>	GPETLLCGAELVDTLLQFVCCGERGFYFSKPTGYGNARRS--RGI VDECCFQSCELRRLLEMYCAPAKTSKPA
Q42336	Fathead minnow	<i>Pimephales promelas</i>	GPETLLCGAELVDTLLQFVCCGERGFYFSKPTGYGNARRS--RGI VDECCFQSCELRRLLEMYCAPVKTGKTP
Q800M7	Shorthorn sculpin	<i>Myoxocephalus scorpius</i>	GPETLLCGAELVDTLLQFVCCGERGFYFSKPTGYGNARRS--RGI VDECCFQSCELRRLLEMYCAPPKTSKAA
Q800M8	White bass	<i>Morone americana</i>	GPETLLCGAELVDTLLQFVCCGERGFYFSKPTGYGNARRS--RGI VDECCFQSCELRRLLEMYCAPAKTGKAA
Q800M9	White bass	<i>Morone chrysops</i>	GPETLLCGAELVDTLLQFVCCGERGFYFSKPTGYGNARRS--RGI VDECCFQSCELRRLLEMYCAPAKTGKAA
Q4JIT74	Striped bass	<i>Morone saxatilis</i>	GPETLLCGAELVDTLLQFVCCGERGFYFSKPTGYGNARRS--RGI VDECCFQSCELRRLLEMYCAPAKTGKAA
P79824	Water buffalo	<i>Bubalus bubalis</i>	XXXXLCCGAELVDALQFVCCGERGFYFNKPTGYGSSRRAPQGTGIVDECCFRSCDLRRLLEMYCAPLKPASKX
Q0ZND0	Mozambique tilapia	<i>Oreochromis mossambicus</i>	GPETLLCGAELVDTLLQFVCCGERGFYFNKPTGYGSSARRS--RGI VDECCFQSCELRRLLEMYCAPVKTPKIS
Q9YI57	Black toothcarp	<i>Poeciliopsis elongata</i>	XXETLLCGGELVDALQFVCCEDRGRGFYFSRPNRSNSHRRPQNRGIVEERCFRSCDNLLEQYCA--KPAKXSX
Q0WY75	Black porgy	<i>Acanthopagrus schlegelii</i>	SPETLLCGAELVDTLLQFVCCGERGFYFSKPTGYGNARRS--RGI VDECCFQSCELRRLLEMYCAPAKTSKAA
Q0WY75	Sole	<i>Solea senegalensis</i>	GPETLLCGAELVDTLLQFVCCGERGFYFSKPTGYGNARRS--RGI VDECCFQSCELRRLLEMYCAPAKTSKAA

Appendix

Q2KK35	Nile tilapia	<i>Oreochromis niloticus</i>	GPETLLCGAELVDTLQ̄FVCCGERGFYFNKPTGYGSSARRS--RGI VDECCFQ̄SCELQR̄LLEMYYCAPVKTPKIS
Q7LZC6	Thai catfish (brain)	<i>Clarias macrocephalus</i>	GPPYTLGRAELVDSLQ̄FVCCGQ̄RGRGFYFSRPMGYSSSTCWTHNRGILDKCCFQ̄SCQLR̄WIEMYYCAAFEP SLVT
Q91442	Spiny dogfish shark	<i>Squalus acanthias</i>	SPETLLCGAELVDTLQ̄FVCCGDRGFYFNKPAGYRSSVRR--PHRGI VNECCFQ̄SCDLK̄LLEMYYCA--KPQ̄RAD
Q2LCL9	Blue discus	<i>Symphysodon aequifasciata</i>	GPETLLCGAELVDTLQ̄FVCCGERGFYFNKP--GYGPSARRT--RGI VDECCFQ̄SCELQR̄LLEMYYCAPAKTPKIS CA--TPAKSX
Q9N1S6	Roe deer	<i>Capreolus capreolus</i>	XXXXXXXXXXXXDALQ̄FVCCGDRGFYFNKPTGYGSSSRRAPQ̄TGIVDECCFRSCDLR̄RLEMYYCAPLKPAA
Q8SQC4	Brush-tailed possum	<i>Trichosurus vulpecula</i>	GPETLLCGAELVDTLQ̄FVCCGERGFYFSKPTGYGSSSRRLHHTGIVDECCXXXXXXXXXXXXXXXXXXXX
B0FBP3	Arctic Fox	<i>Alopex lagopus</i>	GPETLLCGAELVDTLQ̄FVCCGDRGFYFNKPTGYGSSSRRAPQ̄TGIVDECCFRSCDLR̄RLEMYYCAPLKPAA
B3GQC7	Black-lipped pika	<i>Ochotona curzoniae</i>	APETLLCGAELVDTLQ̄FVCCGDRGFYFNKPIGYGSSSRRAQ̄TGIVDECCFRSCDLR̄RLEMYYCAPLKSAA
B3GQC6	Mole rat	<i>Eospalax baileyi</i>	APETLLCGAELVDTLQ̄FVCCGDRGFYFNKPIGYGSSSRRAQ̄TGIVDECCFRGCDLRRPEMYYCAPLKSAA
A7LKM7	Domestic duck	<i>Anas platyrhynchos</i>	GPETLLCGAELVDTLQ̄FVCCGDRGFYFSKPTGYGSSSRRLHKKGIVDECCFQ̄SCDLR̄RLEMYYCAPIKPKSA
A7BK20	Japanese eel	<i>Anguilla japonica</i>	GTELLCGAELVDTLQ̄FVCCGDRGFYFSKPTGYGSSSRSHNRGIVDECCFQ̄SCELR̄RLEMYYCAPVKPGKAA
Q38JX3	Giant danio	<i>Brachydanio aequipinnatus</i>	GPETLLCGAELVDTLQ̄FVCCGDRGFYFSKPTGYGSSSRSHNRGIVDECCFQ̄SCELR̄RLEMYYCAPVKTGKTP
Not present	Grey kangaroo	<i>Macropus fulliginosus</i>	GPETLLCGAELVDTLQ̄FVCCGERGFYFSKPTGYGSSSRRLHHTGIVDECCFRSCDLR̄RLEMYYCAPIKPAKSA

## Bibliography

- Adachi, Y., Lee, C. T., Coffee, K., Yamagata, N., Ohm, J. E., Park, K. H., Dikov, M. M., Nadaf, S. R., Arteaga, C. L. and Carbone, D. P.** (2002). "Effects of genetic blockade of the insulin-like growth factor receptor in human colon cancer cell lines." *Gastroenterology* 123(4): 1191-1204.
- Adams, T. E., Epa, V. C., Garrett, T. P. and Ward, C. W.** (2000). "Structure and function of the type 1 insulin-like growth factor receptor." *Cellular and Molecular Life Sciences : Cmls* 57(7): 1050-1093.
- Alvino, C. L., McNeil, K. A., Ong, S. C., Delaine, C., Booker, G. W., Wallace, J. C., Whittaker, J. and Forbes, B. E.** (2009). "A novel approach to identify two distinct receptor binding surfaces of insulin-like growth factor II." *J Biol Chem* 284(12): 7656-7664.
- Anfinsen, C. B.** (1973). "Principles that govern the folding of protein chains." *Science* 181(96): 223-230.
- Arakawa, H., Kudo, H., Batrak, V., Caldwell, R. B., Rieger, M. A., Ellwart, J. W. and Buerstedde, J. M.** (2008). "Protein evolution by hypermutation and selection in the B cell line DT40." *Nucleic Acids Res* 36(1): e1.
- Arakawa, T. and Akamine, M.** (2003). "Determination of transition metal ions based on quenching of the rare earth luminescence." *Sensors and Actuators B: Chemical* 91(1-3): 252-255.
- Archer, D. B., Jeenes, D. J. and Mackenzie, D. A.** (1994). "Strategies for improving heterologous protein production from filamentous fungi." *Antonie Van Leeuwenhoek* 65(3): 245-250.
- Arnold, F. H.** (1998a). "Design by Directed Evolution." *Acc. Chem. Res.* 31: 125-131.
- Arnold, F. H.** (1998b). "When blind is better: protein design by evolution." *Nat Biotechnol* 16(7): 617-618.
- Arnold, F. H.** (2006). "Fancy footwork in the sequence space shuffle." *Nat Biotechnol* 24(3): 328-330.
- Arnold, F. H., Wintrode, P. L., Miyazaki, K. and Gershenson, A.** (2001). "How enzymes adapt: lessons from directed evolution." *Trends Biochem Sci* 26(2): 100-106.
- Arteaga, C. L. and Osborne, C. K.** (1989). "Growth inhibition of human breast cancer cells in vitro with an antibody against the type I somatomedin receptor." *Cancer Research* 49(22): 6237-6241.
- Axe, D. D.** (2000). "Extreme functional sensitivity to conservative amino acid changes on enzyme exteriors." *J Mol Biol* 301(3): 585-595.
- Bach, L. A., Hsieh, S., Sakano, K., Fujiwara, H., Perdue, J. F. and Rechler, M. M.** (1993). "Binding of mutants of human insulin-like growth factor II to insulin-like growth factor binding proteins 1-6." *J Biol Chem* 268(13): 9246-9254.
- Bae, B., Sullivan, R. P., Zhao, H. and Nair, S. K.** (2010). "Structure and engineering of L-arabinitol 4-dehydrogenase from *Neurospora crassa*." *J Mol Biol* 402(1): 230-240.
- Baik, S. H., Ide, T., Yoshida, H., Kagami, O. and Harayama, S.** (2003). "Significantly enhanced stability of glucose dehydrogenase by directed evolution." *Appl Microbiol Biotechnol* 61(4): 329-335.

- Bailyes, E. M., Nave, B. T., Soos, M. A., Orr, S. R., Hayward, A. C. and Siddle, K.** (1997). "Insulin receptor/IGF-I receptor hybrids are widely distributed in mammalian tissues: quantification of individual receptor species by selective immunoprecipitation and immunoblotting." *The Biochemical Journal* 327 ( Pt 1): 209-215.
- Barbato, C., Calissano, M., Pickford, A., Romano, N., Sandmann, G. and Macino, G.** (1996). "Mild RIP-an alternative method for in vivo mutagenesis of the albino-3 gene in *Neurospora crassa*." *Mol Gen Genet* 252(4): 353-361.
- Barja, B. C., Remorino, A., Roberti, M. J., Aramendia, P. F.** (2005). "Luminiscence quenching of europium (III) and terbium (III) carboxylates by transition metals in solution." *J. Argent. Chem. Soc.* 93(4-6): 81-96.
- Baserga, R., Peruzzi, F. and Reiss, K.** (2003). "The IGF-1 receptor in cancer biology." *Int J Cancer* 107(6): 873-877.
- Baxter, R. C.** (2000). "Insulin-like growth factor (IGF)-binding proteins: interactions with IGFs and intrinsic bioactivities." *Am J Physiol Endocrinol Metab* 278(6): E967-976.
- Baylis, J. R. J. D., A. G.** (1967). "Frequency of multinucleate conidia in microconidiating strains." *Neurospora Newsletter* 11: 9.
- Bayne, M. L., Applebaum, J., Chicchi, G. G., Miller, R. E. and Cascieri, M. A.** (1990). "The roles of tyrosines 24, 31, and 60 in the high affinity binding of insulin-like growth factor-I to the type 1 insulin-like growth factor receptor." *J Biol Chem* 265(26): 15648-15652.
- Bayne, M. L., Applebaum, J., Underwood, D., Chicchi, G. G., Green, B. G., Hayes, N. S. and Cascieri, M. A.** (1989). "The C region of human insulin-like growth factor (IGF) I is required for high affinity binding to the type 1 IGF receptor." *J Biol Chem* 264(19): 11004-11008.
- Belfiore, A., Frasca, F., Pandini, G., Sciacca, L. and Vigneri, R.** (2009). "Insulin receptor isoforms and insulin receptor/insulin-like growth factor receptor hybrids in physiology and disease." *Endocr Rev* 30(6): 586-623.
- Bell, G. I., Gerhard, D. S., Fong, N. M., Sanchez-Pescador, R. and Rall, L. B.** (1985). "Isolation of the human insulin-like growth factor genes: insulin-like growth factor II and insulin genes are contiguous." *Proc Natl Acad Sci U S A* 82(19): 6450-6454.
- Bell, G. I., Merryweather, J. P., Sanchez-Pescador, R., Stempien, M. M., Priestley, L., Scott, J. and Rall, L. B.** (1984). "Sequence of a cDNA clone encoding human preproinsulin-like growth factor II." *Nature* 310(5980): 775-777.
- Bell, G. I., Stempien, M. M., Fong, N. M. and Rall, L. B.** (1986). "Sequences of liver cDNAs encoding two different mouse insulin-like growth factor I precursors." *Nucleic Acids Res* 14(20): 7873-7882.
- Bennetzen, J. L. and Hall, B. D.** (1982). "Codon selection in yeast." *J Biol Chem* 257(6): 3026-3031.
- Benyoucef, S., Surinya, K. H., Hadaschik, D. and Siddle, K.** (2007). "Characterization of insulin/IGF hybrid receptors: contributions of the insulin receptor L2 and Fn1 domains and the alternatively spliced exon 11 sequence to ligand binding and receptor activation." *Biochem J* 403(3): 603-613.
- Bernath, K., Hai, M., Mastrobattista, E., Griffiths, A. D., Magdassi, S. and Tawfik, D. S.** (2004). "In vitro compartmentalization by double emulsions: sorting and gene enrichment by fluorescence activated cell sorting." *Anal Biochem* 325(1): 151-157.
- Bhaumick, B., Goren, H. J. and Bala, R. M.** (1981). "Further characterization of human basic-somatomedin: comparison with insulin-like growth factors I and II." *Horm Metab Res* 13(9): 515-518.



- Blagodatski, A. and Katanaev, V. L.** (2011). "Technologies of directed protein evolution in vivo." *Cell Mol Life Sci* 68(7): 1207-1214.
- Blanquart, C., Achi, J. and Issad, T.** (2008). "Characterization of IRA/IRB hybrid insulin receptors using bioluminescence resonance energy transfer." *Biochem Pharmacol* 76(7): 873-883.
- Bloom, J. D. and Arnold, F. H.** (2009). "In the light of directed evolution: pathways of adaptive protein evolution." *Proc Natl Acad Sci U S A* 106 Suppl 1: 9995-10000.
- Bloom, J. D., Labthavikul, S. T., Otey, C. R. and Arnold, F. H.** (2006). "Protein stability promotes evolvability." *Proc Natl Acad Sci U S A* 103(15): 5869-5874.
- Bloom, J. D., Meyer, M. M., Meinhold, P., Otey, C. R., MacMillan, D. and Arnold, F. H.** (2005). "Evolving strategies for enzyme engineering." *Curr Opin Struct Biol* 15(4): 447-452.
- Blundell, T. L., Bedarkar, S. and Humbel, R. E.** (1983). "Tertiary structures, receptor binding, and antigenicity of insulinlike growth factors." *Fed Proc* 42(9): 2592-2597.
- Blundell, T. L., Bedarkar, S., Rinderknecht, E. and Humbel, R. E.** (1978). "Insulin-like growth factor: a model for tertiary structure accounting for immunoreactivity and receptor binding." *Proc Natl Acad Sci U S A* 75(1): 180-184.
- Boder, E. T. and Wittrup, K. D.** (2000). "Yeast surface display for directed evolution of protein expression, affinity, and stability." *Methods Enzymol* 328: 430-444.
- Bohula, E. A., Salisbury, A. J., Sohail, M., Playford, M. P., Riedemann, J., Southern, E. M. and Macaulay, V. M.** (2003). "The efficacy of small interfering RNAs targeted to the type 1 insulin-like growth factor receptor (IGF1R) is influenced by secondary structure in the IGF1R transcript." *J Biol Chem* 278(18): 15991-15997.
- Bordoli, L., Kiefer, F., Arnold, K., Benkert, P., Battey, J. and Schwede, T.** (2009). "Protein structure homology modeling using SWISS-MODEL workspace." *Nat Protoc* 4(1): 1-13.
- Borkovich, K. A., Alex, L. A., Yarden, O., Freitag, M., Turner, G. E., Read, N. D., Seiler, S., Bell-Pedersen, D., Paietta, J., Plesofsky, N., Plamann, M., Goodrich-Tanrikulu, M., Schulte, U., Mannhaupt, G., Nargang, F. E., Radford, A., Selitrennikoff, C., Galagan, J. E., Dunlap, J. C., Loros, J. J., Catcheside, D., Inoue, H., Aramayo, R., Polymenis, M., Selker, E. U., Sachs, M. S., Marzluf, G. A., Paulsen, I., Davis, R., Ebbole, D. J., Zelter, A., Kalkman, E. R., O'Rourke, R., Bowring, F., Yeadon, J., Ishii, C., Suzuki, K., Sakai, W. and Pratt, R.** (2004). "Lessons from the genome sequence of *Neurospora crassa*: tracing the path from genomic blueprint to multicellular organism." *Microbiol Mol Biol Rev* 68(1): 1-108.
- Bornscheuer, U. T., Altenbuchner, J. and Meyer, H. H.** (1999). "Directed evolution of an esterase: screening of enzyme libraries based on pH-indicators and a growth assay." *Bioorg Med Chem* 7(10): 2169-2173.
- Bowring, F. J. and Catcheside, D. E.** (1991). "The initiation site for recombination cog is at the 3' end of the his-3 gene in *Neurospora crassa*." *Mol Gen Genet* 229(2): 273-277.
- Breuhahn, K., Longrich, T. and Schirmacher, P.** (2006). "Dysregulation of growth factor signaling in human hepatocellular carcinoma." *Oncogene* 25(27): 3787-3800.
- Brissenden, J. E., Ullrich, A. and Francke, U.** (1984). "Human chromosomal mapping of genes for insulin-like growth factors I and II and epidermal growth factor." *Nature* 310(5980): 781-784.
- Broekhuijsen, M. P., Mattern, I. E., Contreras, R., Kinghorn, J. R. and van den Hondel, C. A.** (1993). "Secretion of heterologous proteins by *Aspergillus niger*: production of active human interleukin-6 in a protease-deficient mutant by KEX2-like processing of a glucoamylase-hIL6 fusion protein." *J Biotechnol* 31(2): 135-145.

- Brown, J., Delaine, C., Zaccheo, O. J., Siebold, C., Gilbert, R. J., van Boxel, G., Denley, A., Wallace, J. C., Hassan, A. B., Forbes, B. E. and Jones, E. Y.** (2008). "Structure and functional analysis of the IGF-II/IGF2R interaction." *EMBO J* 27(1): 265-276.
- Brzozowski, A. M., Dodson, E. J., Dodson, G. G., Murshudov, G. N., Verma, C., Turkenburg, J. P., de Bree, F. M. and Dauter, Z.** (2002). "Structural origins of the functional divergence of human insulin-like growth factor-I and insulin." *Biochemistry* 41(30): 9389-9397.
- Buck, M. A., Olah, T. A., Weitzmann, C. J. and Cooperman, B. S.** (1989). "Protein estimation by the product of integrated peak area and flow rate." *Anal Biochem* 182(2): 295-299.
- Bulter, T., Alcalde, M., Sieber, V., Meinhold, P., Schlachtbauer, C. and Arnold, F. H.** (2003). "Functional expression of a fungal laccase in *Saccharomyces cerevisiae* by directed evolution." *Appl Environ Microbiol* 69(2): 987-995.
- Burgess, R. R.** (2009). "Refolding solubilized inclusion body proteins." *Methods Enzymol* 463: 259-282.
- Bustin, S. A. and Jenkins, P. J.** (2001). "The growth hormone-insulin-like growth factor-I axis and colorectal cancer." *Trends in Molecular Medicine* 7(10): 447-454.
- Callanan, M. J., Russell, W. M. and Klaenhammer, T. R.** (2007). "Modification of *Lactobacillus* beta-glucuronidase activity by random mutagenesis." *Gene* 389(2): 122-127.
- Cambareri, E.** (2008). "Methods for in vivo diversification of single genes."
- Cambareri, E. B., Jensen, B. C., Schabtach, E. and Selker, E. U.** (1989). "Repeat-induced G-C to A-T mutations in *Neurospora*." *Science* 244(4912): 1571-1575.
- Cambareri, E. B., Singer, M. J. and Selker, E. U.** (1991). "Recurrence of repeat-induced point mutation (RIP) in *Neurospora crassa*." *Genetics* 127(4): 699-710.
- Camps, M., Naukkarinen, J., Johnson, B. P. and Loeb, L. A.** (2003). "Targeted gene evolution in *Escherichia coli* using a highly error-prone DNA polymerase I." *Proc Natl Acad Sci U S A* 100(17): 9727-9732.
- Catcheside, D. E.** (1986). "A restriction and modification model for the initiation and control of recombination in *Neurospora*." *Genet Res* 47(3): 157-165.
- Catcheside, D. E., Rasmussen, J. P., Yeadon, P. J., Bowring, F. J., Cambareri, E. B., Kato, E., Gabe, J. and Stuart, W. D.** (2003). "Diversification of exogenous genes in vivo in *Neurospora*." *Appl Microbiol Biotechnol* 62(5-6): 544-549.
- Chan, S. J., Cao, Q. P. and Steiner, D. F.** (1990). "Evolution of the insulin superfamily: cloning of a hybrid insulin/insulin-like growth factor cDNA from amphioxus." *Proc Natl Acad Sci U S A* 87(23): 9319-9323.
- Chan, S. J., Cao, Q. P., Nagamatsu, S. & Steiner, D. F.** (1993). Insulin and insulin-like growth factor genes in fishes and other primitive chordates. *Biochemistry and Molecular Biology of Fishes*. H. a. Mommsen, Elsevier Science Publishers. 2: 407-417.
- Chan, S. J. S., D.F.** (2000). "Insulin through the ages: Phylogeny of a growth promoting and metabolic regulatory hormone." *Am Zool* 40: 213-222.
- Chao, W. and D'Amore, P. A.** (2008). "IGF2: epigenetic regulation and role in development and disease." *Cytokine Growth Factor Rev* 19(2): 111-120.
- Chen, J.-W., Ledet, T., Orskov, H., Jessen, N., Lund, S., Whittaker, J., De Meyts, P., Larsen, M. B., Christiansen, J. S. and Frystyk, J.** (2003). "A highly sensitive and specific assay for determination of IGF-I bioactivity in human serum." *Am J Physiol Endocrinol Metab* 284(6): E1149-1155.

- Chen, W. Y., Chen, N. Y., Yun, J., Wagner, T. E. and Kopchick, J. J.** (1994). "In vitro and in vivo studies of antagonistic effects of human growth hormone analogs." *J Biol Chem* 269(22): 15892-15897.
- Chernausek, S. D., Jacobs, S. and Van Wyk, J. J.** (1981). "Structural similarities between human receptors for somatomedin C and insulin: analysis by affinity labeling." *Biochemistry* 20(26): 7345-7350.
- Cherry, J. R., Lamsa, M. H., Schneider, P., Vind, J., Svendsen, A., Jones, A. and Pedersen, A. H.** (1999). "Directed evolution of a fungal peroxidase." *Nat Biotechnol* 17(4): 379-384.
- Chew, S. L., Lavender, P., Clark, A. J. and Ross, R. J.** (1995). "An alternatively spliced human insulin-like growth factor-I transcript with hepatic tissue expression that diverts away from the mitogenic IBE1 peptide." *Endocrinology* 136(5): 1939-1944.
- Chiou, C. H., Miller, M., Wilson, D. L., Trail, F. and Linz, J. E.** (2002). "Chromosomal location plays a role in regulation of aflatoxin gene expression in *Aspergillus parasiticus*." *Appl Environ Microbiol* 68(1): 306-315.
- Chitnis, M. M., Yuen, J. S., Protheroe, A. S., Pollak, M. and Macaulay, V. M.** (2008). "The type 1 insulin-like growth factor receptor pathway." *Clin Cancer Res* 14(20): 6364-6370.
- Choppin, G. R.** (1997). "Factors in Ln(III) complexation." *Journal of Alloys and Compounds* 249(1-2): 1-8.
- Christman, J. K., Schneiderman, N. and Acs, G.** (1985). "Formation of highly stable complexes between 5-azacytosine-substituted DNA and specific non-histone nuclear proteins. Implications for 5-azacytidine-mediated effects on DNA methylation and gene expression." *J Biol Chem* 260(7): 4059-4068.
- Christoffersen, C. T., Bornfeldt, K. E., Rotella, C. M., Gonzales, N., Vissing, H., Shymko, R. M., ten Hoeve, J., Groffen, J., Heisterkamp, N. and De Meyts, P.** (1994). "Negative cooperativity in the insulin-like growth factor-I receptor and a chimeric IGF-I/insulin receptor." *Endocrinology* 135(1): 472-475.
- Clemmons, D. R.** (2001). "Use of mutagenesis to probe IGF-binding protein structure/function relationships." *Endocr Rev* 22(6): 800-817.
- Clemmons, D. R., Dehoff, M. L., Busby, W. H., Bayne, M. L. and Cascieri, M. A.** (1992). "Competition for binding to insulin-like growth factor (IGF) binding protein-2, 3, 4, and 5 by the IGFs and IGF analogs." *Endocrinology* 131(2): 890-895.
- Clemmons, D. R. and Underwood, L. E.** (1991). "Nutritional regulation of IGF-I and IGF binding proteins." *Annu Rev Nutr* 11: 393-412.
- Coco, W. M., Encell, L. P., Levinson, W. E., Crist, M. J., Loomis, A. K., Licato, L. L., Arensdorf, J. J., Sica, N., Pienkos, P. T. and Monticello, D. J.** (2002). "Growth factor engineering by degenerate homoduplex gene family recombination." *Nat Biotechnol* 20(12): 1246-1250.
- Coco, W. M., Levinson, W. E., Crist, M. J., Hektor, H. J., Darzins, A., Pienkos, P. T., Squires, C. H. and Monticello, D. J.** (2001). "DNA shuffling method for generating highly recombined genes and evolved enzymes." *Nat Biotechnol* 19(4): 354-359.
- Cohen, P., Clemmons, D. R. and Rosenfeld, R. G.** (2000). "Does the GH-IGF axis play a role in cancer pathogenesis?" *Growth Horm IGF Res* 10(6): 297-305.
- Collet, C., Candy, J. & Sara, V. (1997). Evolutionary Aspects of the IGF System. Proceeds of the 4th International Symposium on Insulin-like Growth Factors, okyo International Forum, Tokyo, Japan, Elsevier.

- Collins, M. E., Briggs, G., Sawyer, C., Sheffield, P. and Connerton, I. F.** (1991). "An inducible gene expression system for *Neurospora crassa*." *Enzyme Microb Technol* 13(5): 400-403.
- Cooke, R. M., Harvey, T. S. and Campbell, I. D.** (1991). "Solution structure of human insulin-like growth factor 1: a nuclear magnetic resonance and restrained molecular dynamics study." *Biochemistry* 30(22): 5484-5491.
- Cramer, A., Raillard, S. A., Bermudez, E. and Stemmer, W. P.** (1998). "DNA shuffling of a family of genes from diverse species accelerates directed evolution." *Nature* 391(6664): 288-291.
- Crouch, S. P., Kozlowski, R., Slater, K. J. and Fletcher, J.** (1993). "The use of ATP bioluminescence as a measure of cell proliferation and cytotoxicity." *J Immunol Methods* 160(1): 81-88.
- Dagnaes-Hansen, F., Duan, H., Rasmussen, L. M., Friend, K. E. and Flyvbjerg, A.** (2004). "Growth hormone receptor antagonist administration inhibits growth of human colorectal carcinoma in nude mice." *Anticancer Res* 24(6): 3735-3742.
- Dalby, P. A.** (2011). "Strategy and success for the directed evolution of enzymes." *Curr Opin Struct Biol*.
- Daughaday, W. H. and Rotwein, P.** (1989). "Insulin-like growth factors I and II. Peptide, messenger ribonucleic acid and gene structures, serum, and tissue concentrations." *Endocrine Reviews* 10(1): 68-91.
- Davis, R. H.** (2000). *Neurospora Contributions of a Model Organism*. New York, Oxford University Press.
- Davis, R. H., and F. F. DeSerres.** (1970). "Genetic and microbiological research techniques for *Neurospora crassa*." *Methods Enzymol.* 17A: 79-143.
- de Andrade Leite, S. R., Couto dos Santos, M. A., Carubelli, C. R. and Galindo Massabni, A. M.** (1999). "L-Histidine-europium(III) complex: a spectroscopical study." *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy* 55(6): 1185-1191.
- De Meyts, P.** (1994). "The structural basis of insulin and insulin-like growth factor-I receptor binding and negative co-operativity, and its relevance to mitogenic versus metabolic signalling." *Diabetologia* 37 Suppl 2: S135-148.
- De Meyts, P.** (2008). "The insulin receptor: a prototype for dimeric, allosteric membrane receptors?" *Trends Biochem Sci* 33(8): 376-384.
- De Meyts, P., Gauguin, L., Svendsen, A. M., Sarhan, M., Knudsen, L., Nohr, J. and Kiselyov, V. V.** (2009). "Structural basis of allosteric ligand-receptor interactions in the insulin/relaxin peptide family: implications for other receptor tyrosine kinases and G-protein-coupled receptors." *Ann N Y Acad Sci* 1160: 45-53.
- De Meyts, P., Palsgaard, J., Sajid, W., Theede, A. M. and Aladdin, H.** (2004). "Structural biology of insulin and IGF-1 receptors." *Novartis Found Symp* 262: 160-171; discussion 171-166, 265-168.
- De Meyts, P. and Whittaker, J.** (2002). "Structural biology of insulin and IGF1 receptors: implications for drug design." *Nat Rev Drug Discov* 1(10): 769-783.
- de Pagter-Holthuis, P., Jansen, M., van Schaik, F. M., van der Kammen, R., Oosterwijk, C., Van den Brande, J. L. and Sussenbach, J. S.** (1987). "The human insulin-like growth factor II gene contains two development-specific promoters." *FEBS Lett* 214(2): 259-264.
- De Terra, N. and Tatum, E. L.** (1961). "Colonial growth of *Neurospora*. Sorbose and enzymes alter the composition of the cell wall and induce morphological changes." *Science* 134: 1066-1068.

- DeChiara, T. M., Robertson, E. J. and Efstratiadis, A.** (1991). "Parental imprinting of the mouse insulin-like growth factor II gene." *Cell* 64(4): 849-859.
- Delaine, C., Alvino, C. L., McNeil, K. A., Mulhern, T. D., Gauguin, L., De Meyts, P., Jones, E. Y., Brown, J., Wallace, J. C. and Forbes, B. E.** (2007). "A novel binding site for the human insulin-like growth factor-II (IGF-II)/mannose 6-phosphate receptor on IGF-II." *J Biol Chem* 282(26): 18886-18894.
- Denley, A., Bonython, E. R., Booker, G. W., Cosgrove, L. J., Forbes, B. E., Ward, C. W. and Wallace, J. C.** (2004). "Structural determinants for high-affinity binding of insulin-like growth factor II to insulin receptor (IR)-A, the exon 11 minus isoform of the IR." *Mol Endocrinol* 18(10): 2502-2512.
- Denley, A., Brierley, G. V., Carroll, J. M., Lindenberg, A., Booker, G. W., Cosgrove, L. J., Wallace, J. C., Forbes, B. E. and Roberts, C. T., Jr.** (2006). "Differential activation of insulin receptor isoforms by insulin-like growth factors is determined by the C domain." *Endocrinology* 147(2): 1029-1036.
- Denley, A., Carroll, J. M., Brierley, G. V., Cosgrove, L., Wallace, J., Forbes, B. and Roberts, C. T., Jr.** (2007). "Differential activation of insulin receptor substrates 1 and 2 by insulin-like growth factor-activated insulin receptors." *Mol Cell Biol* 27(10): 3569-3577.
- Denley, A., Cosgrove, L. J., Booker, G. W., Wallace, J. C. and Forbes, B. E.** (2005). "Molecular interactions of the IGF system." *Cytokine Growth Factor Rev* 16(4-5): 421-439.
- Derby, M. C. and Gleeson, P. A.** (2007). "New insights into membrane trafficking and protein sorting." *Int Rev Cytol* 261: 47-116.
- Di Cola, G., Cool, M. H. and Accili, D.** (1997). "Hypoglycemic effect of insulin-like growth factor-1 in mice lacking insulin receptors." *J Clin Invest* 99(10): 2538-2544.
- Divisova, J., Kuitse, I., Lazard, Z., Weiss, H., Vreeland, F., Hadsell, D. L., Schiff, R., Osborne, C. K. and Lee, A. V.** (2006). "The growth hormone receptor antagonist pegvisomant blocks both mammary gland development and MCF-7 breast cancer xenograft growth." *Breast Cancer Res Treat* 98(3): 315-327.
- Dransfield, D. T., Cohen, E. H., Chang, Q., Sparrow, L. G., Bentley, J. D., Dolezal, O., Xiao, X., Peat, T. S., Newman, J., Pilling, P. A., Phan, T., Priebe, I., Brierley, G. V., Kastrapeli, N., Kopacz, K., Martik, D., Wassaf, D., Rank, D., Conley, G., Huang, Y., Adams, T. E. and Cosgrove, L.** (2010). "A human monoclonal antibody against insulin-like growth factor-II blocks the growth of human hepatocellular carcinoma cell lines in vitro and in vivo." *Mol Cancer Ther* 9(6): 1809-1819.
- Drucker, H.** (1972). "Regulation of exocellular proteases in *Neurospora crassa*: induction and repression of enzyme synthesis." *J Bacteriol* 110(3): 1041-1049.
- Duguay, S. J., Chan, S. J., Mommsen, T. P. and Steiner, D. F.** (1995). "Divergence of insulin-like growth factors I and II in the elasmobranch, *Squalus acanthias*." *FEBS Lett* 371(1): 69-72.
- Duguay, S. J., Jin, Y., Stein, J., Duguay, A. N., Gardner, P. and Steiner, D. F.** (1998). "Post-translational processing of the insulin-like growth factor-2 precursor. Analysis of O-glycosylation and endoproteolysis." *J Biol Chem* 273(29): 18443-18451.
- Dunn, S. E., Ehrlich, M., Sharp, N. J., Reiss, K., Solomon, G., Hawkins, R., Baserga, R. and Barrett, J. C.** (1998). "A dominant negative mutant of the insulin-like growth factor-I receptor inhibits the adhesion, invasion, and metastasis of breast cancer." *Cancer Res* 58(15): 3353-3361.
- Dupont, J. and LeRoith, D.** (2001). "Insulin and insulin-like growth factor I receptors: similarities and differences in signal transduction." *Horm Res* 55 Suppl 2: 22-26.

- Duret, L. and Mouchiroud, D.** (1999). "Expression pattern and, surprisingly, gene length shape codon usage in *Caenorhabditis*, *Drosophila*, and *Arabidopsis*." *Proc Natl Acad Sci U S A* 96(8): 4482-4487.
- Ebbole, D. S., M. S.** (1990). "A rapid and simple method for isolation of *Neurospora crassa* homokaryons using microconidia." *Fungal Genet. Newslett.* 37.
- Edelmann, S. E. and Staben, C.** (1994). "A statistical analysis of sequence features within genes from *Neurospora crassa*." *Experimental Mycology* 18(1): 70-81.
- Edwards, R. M. O., GB2), Bawden, Lindsay (Oxford, GB2)** (1998). "IGF-II analogues."
- Emanuelsson, O., Brunak, S., von Heijne, G. and Nielsen, H.** (2007). "Locating proteins in the cell using TargetP, SignalP and related tools." *Nat Protoc* 2(4): 953-971.
- Farinas, E. T., Bulter, T. and Arnold, F. H.** (2001). "Directed enzyme evolution." *Curr Opin Biotechnol* 12(6): 545-551.
- Fernandez-Fuentes, N., Zhai, J. and Fiser, A.** (2006). "ArchPRED: a template based loop structure prediction server." *Nucleic Acids Res* 34(Web Server issue): W173-176.
- Fincham, J. R.** (1990). "Generation of new functional mutant alleles by premeiotic disruption of the *Neurospora crassa* *am* gene." *Curr Genet* 18(5): 441-445.
- Fisher, C. L. P., G. K.** (1997). "Modification of a PCR-based site-directed mutagenesis method." *Biotechniques* 23(4): 570-574.
- Foss, H. M., Roberts, C. J., Claeys, K. M. and Selker, E. U.** (1993). "Abnormal chromosome behavior in *Neurospora* mutants defective in DNA methylation." *Science* 262(5140): 1737-1741.
- Fottner, C., Engelhardt, D. and Weber, M. M.** (1998). "Regulation of steroidogenesis by insulin-like growth factors (IGFs) in adult human adrenocortical cells: IGF-I and, more potently, IGF-II preferentially enhance androgen biosynthesis through interaction with the IGF-I receptor and IGF-binding proteins." *J Endocrinol* 158(3): 409-417.
- Foulstone, E., Prince, S., Zaccheo, O., Burns, J. L., Harper, J., Jacobs, C., Church, D. and Hassan, A. B.** (2005). "Insulin-like growth factor ligands, receptors, and binding proteins in cancer." *J Pathol* 205(2): 145-153.
- Francis, G. L., Aplin, S. E., Milner, S. J., McNeil, K. A., Ballard, F. J. and Wallace, J. C.** (1993). "Insulin-like growth factor (IGF)-II binding to IGF-binding proteins and IGF receptors is modified by deletion of the N-terminal hexapeptide or substitution of arginine for glutamate-6 in IGF-II." *Biochem J* 293 ( Pt 3): 713-719.
- Francis, G. L., Ross, M., Ballard, F. J., Milner, S. J., Senn, C., McNeil, K. A., Wallace, J. C., King, R. and Wells, J. R.** (1992). "Novel recombinant fusion protein analogues of insulin-like growth factor (IGF)-I indicate the relative importance of IGF-binding protein and receptor binding for enhanced biological potency." *J Mol Endocrinol* 8(3): 213-223.
- Francisco, J. A., Campbell, R., Iverson, B. L. and Georgiou, G.** (1993). "Production and fluorescence-activated cell sorting of *Escherichia coli* expressing a functional antibody fragment on the external surface." *Proc Natl Acad Sci U S A* 90(22): 10444-10448.
- Francisco, J. A., Earhart, C. F. and Georgiou, G.** (1992). "Transport and anchoring of beta-lactamase to the external surface of *Escherichia coli*." *Proc Natl Acad Sci U S A* 89(7): 2713-2717.
- Frasca, F., Pandini, G., Scalia, P., Sciacca, L., Mineo, R., Costantino, A., Goldfine, I. D., Belfiore, A. and Vigneri, R.** (1999). "Insulin receptor isoform A, a newly

recognized, high-affinity insulin-like growth factor II receptor in fetal and cancer cells." *Molecular and Cellular Biology* 19(5): 3278-3288.

**Fuh, G., Cunningham, B. C., Fukunaga, R., Nagata, S., Goeddel, D. V. and Wells, J. A.** (1992). "Rational design of potent antagonists to the human growth hormone receptor." *Science* 256(5064): 1677-1680.

**Galagan, J. E., Calvo, S. E., Borkovich, K. A., Selker, E. U., Read, N. D., Jaffe, D., FitzHugh, W., Ma, L. J., Smirnov, S., Purcell, S., Rehman, B., Elkins, T., Engels, R., Wang, S., Nielsen, C. B., Butler, J., Endrizzi, M., Qui, D., Ianakiev, P., Bell-Pedersen, D., Nelson, M. A., Werner-Washburne, M., Selitrennikoff, C. P., Kinsey, J. A., Braun, E. L., Zelter, A., Schulte, U., Kothe, G. O., Jedd, G., Mewes, W., Staben, C., Marcotte, E., Greenberg, D., Roy, A., Foley, K., Naylor, J., Stange-Thomann, N., Barrett, R., Gnerre, S., Kamal, M., Kamvysselis, M., Mauceli, E., Bielke, C., Rudd, S., Frishman, D., Krystofova, S., Rasmussen, C., Metzzenberg, R. L., Perkins, D. D., Kroken, S., Cogoni, C., Macino, G., Catcheside, D., Li, W., Pratt, R. J., Osmani, S. A., DeSouza, C. P., Glass, L., Orbach, M. J., Berglund, J. A., Voelker, R., Yarden, O., Plamann, M., Seiler, S., Dunlap, J., Radford, A., Aramayo, R., Natvig, D. O., Alex, L. A., Mannhaupt, G., Ebbole, D. J., Freitag, M., Paulsen, I., Sachs, M. S., Lander, E. S., Nusbaum, C. and Birren, B.** (2003). "The genome sequence of the filamentous fungus *Neurospora crassa*." *Nature* 422(6934): 859-868.

**Garrett, T. P., McKern, N. M., Lou, M., Frenkel, M. J., Bentley, J. D., Lovrecz, G. O., Elleman, T. C., Cosgrove, L. J. and Ward, C. W.** (1998). "Crystal structure of the first three domains of the type-1 insulin-like growth factor receptor." *Nature* 394(6691): 395-399.

**Gauguin, L., Delaine, C., Alvino, C. L., McNeil, K. A., Wallace, J. C., Forbes, B. E. and De Meyts, P.** (2008a). "Alanine scanning of a putative receptor binding surface of insulin-like growth factor-I." *J Biol Chem* 283(30): 20821-20829.

**Gauguin, L., Klaproth, B., Sajid, W., Andersen, A. S., McNeil, K. A., Forbes, B. E. and De Meyts, P.** (2008b). "Structural basis for the lower affinity of the insulin-like growth factors for the insulin receptor." *J Biol Chem* 283(5): 2604-2613.

**Gavric, O., dos Santos, D. B. and Griffiths, A.** (2007). "Mutation and divergence of the phospholipase C gene in *Neurospora crassa*." *Fungal Genet Biol* 44(4): 242-249.

**Georgiou, G.** (2000). "Analysis of large libraries of protein mutants using flow cytometry." *Adv Protein Chem* 55: 293-315.

**Georgiou, G., Stathopoulos, C., Daugherty, P. S., Nayak, A. R., Iverson, B. L. and Curtiss, R., 3rd** (1997). "Display of heterologous proteins on the surface of microorganisms: from the screening of combinatorial libraries to live recombinant vaccines." *Nat Biotechnol* 15(1): 29-34.

**Gill, R., Wallach, B., Verma, C., Urso, B., De Wolf, E., Grotzinger, J., Murray-Rust, J., Pitts, J., Wollmer, A., De Meyts, P. and Wood, S.** (1996). "Engineering the C-region of human insulin-like growth factor-1: implications for receptor binding." *Protein Eng* 9(11): 1011-1019.

**Gouka, R. J., Punt, P. J., Hessing, J. G. and van den Hondel, C. A.** (1996). "Analysis of heterologous protein production in defined recombinant *Aspergillus awamori* strains." *Appl Environ Microbiol* 62(6): 1951-1957.

**Gouka, R. J., Punt, P. J. and van den Hondel, C. A.** (1997a). "Efficient production of secreted proteins by *Aspergillus*: progress, limitations and prospects." *Appl Microbiol Biotechnol* 47(1): 1-11.

- Gouka, R. J., Punt, P. J. and van den Hondel, C. A.** (1997b). "Glucoamylase gene fusions alleviate limitations for protein production in *Aspergillus awamori* at the transcriptional and (post) translational levels." *Appl Environ Microbiol* 63(2): 488-497.
- Goya, M., Miyamoto, S., Nagai, K., Ohki, Y., Nakamura, K., Shitara, K., Maeda, H., Sangai, T., Kodama, K., Endoh, Y., Ishii, G., Hasebe, T., Yonou, H., Hatano, T., Ogawa, Y. and Ochiai, A.** (2004). "Growth inhibition of human prostate cancer cells in human adult bone implanted into nonobese diabetic/severe combined immunodeficient mice by a ligand-specific antibody to human insulin-like growth factors." *Cancer Res* 64(17): 6252-6258.
- Grayburn, W. S. and Selker, E. U.** (1989). "A natural case of RIP: degeneration of the DNA sequence in an ancestral tandem duplication." *Mol Cell Biol* 9(10): 4416-4421.
- Greener, A., Callahan, M. and Jerpseth, B.** (1996). "An efficient random mutagenesis technique using an *E. coli* mutator strain." *Methods Mol Biol* 57: 375-385.
- Greener, A., Callahan, M. and Jerpseth, B.** (1997). "An efficient random mutagenesis technique using an *E. coli* mutator strain." *Mol Biotechnol* 7(2): 189-195.
- Griffiths, A. D. and Tawfik, D. S.** (2003). "Directed evolution of an extremely fast phosphotriesterase by in vitro compartmentalization." *EMBO J* 22(1): 24-35.
- Guan, K. L. and Dixon, J. E.** (1991). "Eukaryotic proteins expressed in *Escherichia coli*: an improved thrombin cleavage and purification procedure of fusion proteins with glutathione S-transferase." *Anal Biochem* 192(2): 262-267.
- Guo, H. H., Choe, J. and Loeb, L. A.** (2004). "Protein tolerance to random amino acid change." *Proc Natl Acad Sci U S A* 101(25): 9205-9210.
- Gupta, R. D. and Tawfik, D. S.** (2008). "Directed enzyme evolution via small and effective neutral drift libraries." *Nat Methods* 5(11): 939-942.
- Gustafsson, C., Govindarajan, S. and Minshull, J.** (2004). "Codon bias and heterologous protein expression." *Trends Biotechnol* 22(7): 346-353.
- Hameed, M., Orrell, R. W., Cobbold, M., Goldspink, G. and Harridge, S. D.** (2003). "Expression of IGF-I splice variants in young and old human skeletal muscle after high resistance exercise." *J Physiol* 547(Pt 1): 247-254.
- Hammarberg, B., Nygren, P. A., Holmgren, E., Elmlad, A., Tally, M., Hellman, U., Moks, T. and Uhlen, M.** (1989). "Dual affinity fusion approach and its use to express recombinant human insulin-like growth factor II." *Proc Natl Acad Sci U S A* 86(12): 4367-4371.
- Handl, H. L. and Gillies, R. J.** (2005). "Lanthanide-based luminescent assays for ligand-receptor interactions." *Life Sci* 77(4): 361-371.
- Hanes, J. and Pluckthun, A.** (1997). "In vitro selection and evolution of functional proteins by using ribosome display." *Proc Natl Acad Sci U S A* 94(10): 4937-4942.
- Hanson, M. A. and Marzluf, G. A.** (1975). "Control of the synthesis of a single enzyme by multiple regulatory circuits in *Neurospora crassa*." *Proc Natl Acad Sci U S A* 72(4): 1240-1244.
- Hassan, A. B. and Howell, J. A.** (2000). "Insulin-like growth factor II supply modifies growth of intestinal adenoma in *Apc(Min/+)* mice." *Cancer Res* 60(4): 1070-1076.
- He, M., Yang, Z. Y., Nie, Y. F., Wang, J. and Xu, P.** (2001). "A new type of class I bacterial 5-enopyruvylshikimate-3-phosphate synthase mutants with enhanced tolerance to glyphosate." *Biochim Biophys Acta* 1568(1): 1-6.
- Heiniger, U. and Matile, P.** (1974). "Protease secretion in *Neurospora crassa*." *Biochem Biophys Res Commun* 60(4): 1425-1432.
- Hellawell, G. O., Turner, G. D., Davies, D. R., Poulson, R., Brewster, S. F. and Macaulay, V. M.** (2002). "Expression of the type 1 insulin-like growth factor receptor



is up-regulated in primary prostate cancer and commonly persists in metastatic disease." *Cancer Res* 62(10): 2942-2950.

**Henderson, S. T., Eariss, G. A. and Catcheside, D. E. A.** (2005). "Reliable PCR amplification from *Neurospora crassa* genomic DNA obtained from conidia." *Fungal Genet. Newsl.* 52.

**Henikoff, S. and Henikoff, J. G.** (1992). "Amino acid substitution matrices from protein blocks." *Proc Natl Acad Sci U S A* 89(22): 10915-10919.

**Hohenblum, H., Gasser, B., Maurer, M., Borth, N. and Mattanovich, D.** (2004). "Effects of gene dosage, promoters, and substrates on unfolded protein stress of recombinant *Pichia pastoris*." *Biotechnol Bioeng* 85(4): 367-375.

**Hoyne, P. A., Elleman, T. C., Adams, T. E., Richards, K. M. and Ward, C. W.** (2000). "Properties of an insulin receptor with an IGF-1 receptor loop exchange in the cysteine-rich region." *FEBS Lett* 469(1): 57-60.

**Hudgins, W. R., Hampton, B., Burgess, W. H. and Perdue, J. F.** (1992). "The identification of O-glycosylated precursors of insulin-like growth factor II." *J Biol Chem* 267(12): 8153-8160.

**Ikemura, T.** (1981). "Correlation between the abundance of *Escherichia coli* transfer RNAs and the occurrence of the respective codons in its protein genes: a proposal for a synonymous codon choice that is optimal for the *E. coli* translational system." *J Mol Biol* 151(3): 389-409.

**Iwashita, K.** (2002). "Recent studies of protein secretion by filamentous fungi." *J Biosci Bioeng* 94(6): 530-535.

**J. Irelan, V. M. a. E. U. S.** (1993). "Small scale DNA preps for *Neurospora crassa*." *Fungal Genet. Newsl.* 40: 24.

**Jackel, C., Kast, P. and Hilvert, D.** (2008). "Protein design by directed evolution." *Annu Rev Biophys* 37: 153-173.

**Jansen, M., van Schaik, F. M., Ricker, A. T., Bullock, B., Woods, D. E., Gabbay, K. H., Nussbaum, A. L., Sussenbach, J. S. and Van den Brande, J. L.** (1983). "Sequence of cDNA encoding human insulin-like growth factor I precursor." *Nature* 306(5943): 609-611.

**Jansen, M., van Schaik, F. M., van Tol, H., Van den Brande, J. L. and Sussenbach, J. S.** (1985). "Nucleotide sequences of cDNAs encoding precursors of human insulin-like growth factor II (IGF-II) and an IGF-II variant." *FEBS Lett* 179(2): 243-246.

**Jeenes, D. J., Marczinke, B., MacKenzie, D. A. and Archer, D. B.** (1993). "A truncated glucoamylase gene fusion for heterologous protein secretion from *Aspergillus niger*." *FEMS Microbiol Lett* 107(2-3): 267-271.

**Jones, D. S., Silverman, A. P. and Cochran, J. R.** (2008). "Developing therapeutic proteins by engineering ligand-receptor interactions." *Trends Biotechnol* 26(9): 498-505.

**Jones, J. I. and Clemmons, D. R.** (1995). "Insulin-like growth factors and their binding proteins: biological actions." *Endocrine Reviews* 16(1): 3-34.

**Jones, J. I., D'Ercole, A. J., Camacho-Hubner, C. and Clemmons, D. R.** (1991). "Phosphorylation of insulin-like growth factor (IGF)-binding protein 1 in cell culture and in vivo: effects on affinity for IGF-I." *Proc Natl Acad Sci U S A* 88(17): 7481-7485.

**Juneau, K., Miranda, M., Hillenmeyer, M. E., Nislow, C. and Davis, R. W.** (2006). "Introns regulate RNA and protein abundance in yeast." *Genetics* 174(1): 511-518.

**Kalebic, T., Tsokos, M. and Helman, L. J.** (1994). "In vivo treatment with antibody against IGF-1 receptor suppresses growth of human rhabdomyosarcoma and down-regulates p34cdc2." *Cancer Res* 54(21): 5531-5534.

- Kaleko, M., Rutter, W. J. and Miller, A. D.** (1990). "Overexpression of the human insulinlike growth factor I receptor promotes ligand-dependent neoplastic transformation." *Mol Cell Biol* 10(2): 464-473.
- Kane, J. F.** (1995). "Effects of rare codon clusters on high-level expression of heterologous proteins in Escherichia coli." *Curr Opin Biotechnol* 6(5): 494-500.
- Kavoosi, M., Creagh, A. L., Kilburn, D. G. and Haynes, C. A.** (2007). "Strategy for selecting and characterizing linker peptides for CBM9-tagged fusion proteins expressed in Escherichia coli." *Biotechnol Bioeng* 98(3): 599-610.
- Kawarasaki, Y., Griswold, K. E., Stevenson, J. D., Selzer, T., Benkovic, S. J., Iverson, B. L. and Georgiou, G.** (2003). "Enhanced crossover SCRATCHY: construction and high-throughput screening of a combinatorial library containing multiple non-homologous crossovers." *Nucleic Acids Res* 31(21): e126.
- Keefe, A. D. and Szostak, J. W.** (2001). "Functional proteins from a random-sequence library." *Nature* 410(6829): 715-718.
- Keyhanfar, M., Booker, G. W., Whittaker, J., Wallace, J. C. and Forbes, B. E.** (2007). "Precise mapping of an IGF-I-binding site on the IGF-1R." *Biochem J* 401(1): 269-277.
- Khandwala, H. M., McCutcheon, I. E., Flyvbjerg, A. and Friend, K. E.** (2000). "The effects of insulin-like growth factors on tumorigenesis and neoplastic growth." *Endocrine Reviews* 21(3): 215-244.
- Khosravi, J., Diamandi, A., Bodani, U., Khaja, N. and Krishna, R. G.** (2005). "Pitfalls of immunoassay and sample for IGF-I: comparison of different assay methodologies using various fresh and stored serum samples." *Clin Biochem* 38(7): 659-666.
- Kim, H. S., Nagalla, S. R., Oh, Y., Wilson, E., Roberts, C. T., Jr. and Rosenfeld, R. G.** (1997). "Identification of a family of low-affinity insulin-like growth factor binding proteins (IGFBPs): characterization of connective tissue growth factor as a member of the IGFBP superfamily." *Proc Natl Acad Sci U S A* 94(24): 12981-12986.
- Kim, J. J. and Accili, D.** (2002). "Signalling through IGF-I and insulin receptors: where is the specificity?" *Growth Horm IGF Res* 12(2): 84-90.
- Kimura, M.** (1968a). "Evolutionary rate at the molecular level." *Nature* 217(5129): 624-626.
- Kimura, M.** (1968b). "Genetic variability maintained in a finite population due to mutational production of neutral and nearly neutral isoalleles." *Genet Res* 11(3): 247-269.
- Kimura, M.** (1983). "Rare variant alleles in the light of the neutral theory." *Mol Biol Evol* 1(1): 84-93.
- Kimura, M. and Ogiwara, M.** (1997). "Proliferation of adult rat hepatocytes in primary culture induced by insulin is potentiated by cAMP-elevating agents." *Eur J Pharmacol* 327(1): 87-95.
- King, R., Wells, J. R., Krieg, P., Snoswell, M., Brazier, J., Bagley, C. J., Wallace, J. C., Ballard, F. J., Ross, M. and Francis, G. L.** (1992). "Production and characterization of recombinant insulin-like growth factor-I (IGF-I) and potent analogues of IGF-I, with Gly or Arg substituted for Glu3, following their expression in Escherichia coli as fusion proteins." *J Mol Endocrinol* 8(1): 29-41.
- Kinnaird, J. H., Burns, P. A. and Fincham, J. R.** (1991). "An apparent rare-codon effect on the rate of translation of a Neurospora gene." *J Mol Biol* 221(3): 733-736.
- Kiselyov, V. V., Versteyhe, S., Gauguin, L. and De Meyts, P.** (2009). "Harmonic oscillator model of the insulin and IGF1 receptors' allosteric binding and activation." *Mol Syst Biol* 5: 243.

- Koh, L. Y. and Catchside, D. E.** (2007). "Mutation of *msh-2* in *Neurospora crassa* does not reduce the incidence of recombinants with multiple patches of donor chromosome sequence." *Fungal Genet Biol* 44(7): 575-584.
- Kohek, M., Leme, C., Nakamura, I. T., De Oliveira, S. A., Lando, V. and Mendonca, B. B.** (2002). "Effects of EDTA and Sodium Citrate on hormone measurements by fluorometric (FIA) and immunofluorometric (IFMA) methods." *BMC Clin Pathol* 2(1): 2.
- Kosaki, A., Pillay, T. S., Xu, L. and Webster, N. J.** (1995). "The B isoform of the insulin receptor signals more efficiently than the A isoform in HepG2 cells." *The Journal of Biological Chemistry* 270(35): 20816-20823.
- Kouzminova, E. and Selker, E. U.** (2001). "dim-2 encodes a DNA methyltransferase responsible for all known cytosine methylation in *Neurospora*." *Embo J* 20(15): 4309-4323.
- Kristensen, C., Andersen, A. S., Hach, M., Wiberg, F. C., Schaffer, L. and Kjeldsen, T.** (1995). "A single-chain insulin-like growth factor I/insulin hybrid binds with high affinity to the insulin receptor." *Biochem J* 305 ( Pt 3): 981-986.
- Kristensen, C., Wiberg, F. C. and Andersen, A. S.** (1999). "Specificity of insulin and insulin-like growth factor I receptors investigated using chimeric mini-receptors. Role of C-terminal of receptor alpha subunit." *The Journal of Biological Chemistry* 274(52): 37351-37356.
- Krywicki, R. F. and Yee, D.** (1992). "The insulin-like growth factor family of ligands, receptors, and binding proteins." *Breast Cancer Res Treat* 22(1): 7-19.
- Kuang, Z., Yao, S., McNeil, K. A., Forbes, B. E., Wallace, J. C. and Norton, R. S.** (2009). "Insulin-like growth factor-I (IGF-I): solution properties and NMR chemical shift assignments near physiological pH." *Growth Horm IGF Res* 19(3): 226-231.
- Lahm, H., Amstad, P., Wyniger, J., Yilmaz, A., Fischer, J. R., Schreyer, M. and Givel, J. C.** (1994). "Blockade of the insulin-like growth-factor-I receptor inhibits growth of human colorectal cancer cells: evidence of a functional IGF-II-mediated autocrine loop." *Int J Cancer* 58(3): 452-459.
- Lawrence, M. C., McKern, N. M. and Ward, C. W.** (2007). "Insulin receptor structure and its implications for the IGF-1 receptor." *Curr Opin Struct Biol* 17(6): 699-705.
- Lee, C. T., Park, K. H., Adachi, Y., Seol, J. Y., Yoo, C. G., Kim, Y. W., Han, S. K., Shim, Y. S., Coffee, K., Dikov, M. M. and Carbone, D. P.** (2003). "Recombinant adenoviruses expressing dominant negative insulin-like growth factor-I receptor demonstrate antitumor effects on lung cancer." *Cancer Gene Ther* 10(1): 57-63.
- Lee, C. Y., Yu, K. O., Kim, S. W. and Han, S. O.** (2010). "Enhancement of the thermostability and activity of mesophilic *Clostridium cellulovorans* EngD by in vitro DNA recombination with *Clostridium thermocellum* CelE." *J Biosci Bioeng* 109(4): 331-336.
- Lee, I. H., Walline, R. G. and Plamann, M.** (1998). "Apolar growth of *Neurospora crassa* leads to increased secretion of extracellular proteins." *Mol Microbiol* 29(1): 209-218.
- Legerton, T. L. and Yanofsky, C.** (1985). "Cloning and characterization of the multifunctional *his-3* gene of *Neurospora crassa*." *Gene* 39(2-3): 129-140.
- LeRoith, D., Kavsan, V. M., Koval, A. P. and Roberts, C. T., Jr.** (1993). "Phylogeny of the insulin-like growth factors (IGFs) and receptors: a molecular approach." *Mol Reprod Dev* 35(4): 332-336; discussion 337-338.
- LeRoith, D. and Roberts, C. T., Jr.** (2003). "The insulin-like growth factor system and cancer." *Cancer Lett* 195(2): 127-137.

- Levin, A. M. and Weiss, G. A.** (2006). "Optimizing the affinity and specificity of proteins with molecular display." *Mol Biosyst* 2(1): 49-57.
- Li, X., Cui, H., Sandstedt, B., Nordlinder, H., Larsson, E. and Ekstrom, T. J.** (1996). "Expression levels of the insulin-like growth factor-II gene (IGF2) in the human liver: developmental relationships of the four promoters." *J Endocrinol* 149(1): 117-124.
- Li, X., Nong, Z., Ekstrom, C., Larsson, E., Nordlinder, H., Hofmann, W. J., Trautwein, C., Odenthal, M., Dienes, H. P., Ekstrom, T. J. and Schirmacher, P.** (1997). "Disrupted IGF2 promoter control by silencing of promoter P1 in human hepatocellular carcinoma." *Cancer Res* 57(10): 2048-2054.
- Lien, S., Milner, S. J., Graham, L. D., Wallace, J. C. and Francis, G. L.** (2001). "Linkers for improved cleavage of fusion proteins with an engineered alpha-lytic protease." *Biotechnol Bioeng* 74(4): 335-343.
- Lindberg, R. A., Rhodes, W. G., Eirich, L. D. and Drucker, H.** (1982). "Extracellular acid proteases from *Neurospora crassa*." *J Bacteriol* 150(3): 1103-1108.
- Linnell, J., Groeger, G. and Hassan, A. B.** (2001). "Real time kinetics of insulin-like growth factor II (IGF-II) interaction with the IGF-II/mannose 6-phosphate receptor: the effects of domain 13 and pH." *J Biol Chem* 276(26): 23986-23991.
- Liu, J. P., Baker, J., Perkins, A. S., Robertson, E. J. and Efstratiadis, A.** (1993). "Mice carrying null mutations of the genes encoding insulin-like growth factor I (Igf-1) and type 1 IGF receptor (Igf1r)." *Cell* 75(1): 59-72.
- Liu, K., Sandgren, E. P., Palmiter, R. D. and Stein, A.** (1995). "Rat growth hormone gene introns stimulate nucleosome alignment in vitro and in transgenic mice." *Proc Natl Acad Sci U S A* 92(17): 7724-7728.
- Long-McGie, J., Liu, A. D. and Schellenberger, V.** (2000). "Rapid in vivo evolution of a beta-lactamase using phagemids." *Biotechnol Bioeng* 68(1): 121-125.
- Lopez, T. and Hanahan, D.** (2002). "Elevated levels of IGF-1 receptor convey invasive and metastatic capability in a mouse model of pancreatic islet tumorigenesis." *Cancer Cell* 1(4): 339-353.
- Lutz, S. and Ostermeier, M.** (2003). "Preparation of SCRATCHY hybrid protein libraries: size- and in-frame selection of nucleic acid sequences." *Methods Mol Biol* 231: 143-151.
- Lutz, S., Ostermeier, M. and Benkovic, S. J.** (2001a). "Rapid generation of incremental truncation libraries for protein engineering using alpha-phosphothioate nucleotides." *Nucleic Acids Res* 29(4): E16.
- Lutz, S., Ostermeier, M., Moore, G. L., Maranas, C. D. and Benkovic, S. J.** (2001b). "Creating multiple-crossover DNA libraries independent of sequence identity." *Proc Natl Acad Sci U S A* 98(20): 11248-11253.
- Magee, B. A., Shooter, G. K., Wallace, J. C. and Francis, G. L.** (1999). "Insulin-like growth factor I and its binding proteins: a study of the binding interface using B-domain analogues." *Biochemistry* 38(48): 15863-15870.
- Maheshwari, R.** (1999). "Microconidia of *Neurospora crassa*." *Fungal Genet Biol* 26(1): 1-18.
- Massague, J. and Czech, M. P.** (1982). "Role of disulfides in the subunit structure of the insulin receptor. Reduction of class I disulfides does not impair transmembrane signalling." *J Biol Chem* 257(12): 6729-6738.
- Mathews, L. S., Norstedt, G. and Palmiter, R. D.** (1986). "Regulation of insulin-like growth factor I gene expression by growth hormone." *Proc Natl Acad Sci U S A* 83(24): 9343-9347.

- Matsuura, T. and Yomo, T.** (2006). "In vitro evolution of proteins." *J Biosci Bioeng* 101(6): 449-456.
- McCutcheon, I. E., Flyvbjerg, A., Hill, H., Li, J., Bennett, W. F., Scarlett, J. A. and Friend, K. E.** (2001). "Antitumor activity of the growth hormone receptor antagonist pegvisomant against human meningiomas in nude mice." *J Neurosurg* 94(3): 487-492.
- McKern, N. M., Lawrence, M. C., Streltsov, V. A., Lou, M. Z., Adams, T. E., Lovrecz, G. O., Elleman, T. C., Richards, K. M., Bentley, J. D., Pilling, P. A., Hoyne, P. A., Cartledge, K. A., Pham, T. M., Lewis, J. L., Sankovich, S. E., Stoichevska, V., Da Silva, E., Robinson, C. P., Frenkel, M. J., Sparrow, L. G., Fernley, R. T., Epa, V. C. and Ward, C. W.** (2006). "Structure of the insulin receptor ectodomain reveals a folded-over conformation." *Nature* 443(7108): 218-221.
- McRory, J. E. and Sherwood, N. M.** (1997). "Ancient divergence of insulin and insulin-like growth factor." *DNA Cell Biol* 16(8): 939-949.
- Melcher, K.** (2000). "A modular set of prokaryotic and eukaryotic expression vectors." *Anal Biochem* 277(1): 109-120.
- Merimee, T. J., Zapf, J., Froesch, E. R.** (1982). "Insulin-like growth factor(s) (IGFs) in pygmies and subjects with the pygmy trait: Characterization of the metabolic actions of IGF I and IGF II in man. ." *J Clin. Endocr. Metab.* 55: 1081-1088.
- Michalsky, E., Goede, A. and Preissner, R.** (2003). "Loops In Proteins (LIP)--a comprehensive loop database for homology modelling." *Protein Eng* 16(12): 979-985.
- Michnick, S. W. and Arnold, F. H.** (1999). ""Itching" for new strategies in protein engineering." *Nat Biotechnol* 17(12): 1159-1160.
- Min, Y., Adachi, Y., Yamamoto, H., Ito, H., Itoh, F., Lee, C. T., Nadaf, S., Carbone, D. P. and Imai, K.** (2003). "Genetic blockade of the insulin-like growth factor-I receptor: a promising strategy for human pancreatic cancer." *Cancer Res* 63(19): 6432-6441.
- Minson, A. C. and Creaser, E. H.** (1969). "Purification of a trifunctional enzyme, catalysing three steps of the histidine pathway, from *Neurospora crassa*." *Biochem J* 114(1): 49-56.
- Miyamoto, S., Nakamura, M., Shitara, K., Nakamura, K., Ohki, Y., Ishii, G., Goya, M., Kodama, K., Sangai, T., Maeda, H., Shi-Chuang, Z., Chiba, T. and Ochiai, A.** (2005). "Blockade of paracrine supply of insulin-like growth factors using neutralizing antibodies suppresses the liver metastasis of human colorectal cancers." *Clin Cancer Res* 11(9): 3494-3502.
- Moks, T., Abrahmsen, L., Holmgren, E., Bilich, M., Olsson, A., Uhlen, M., Pohl, G., Sterky, C., Hultberg, H., Josephson, S. and et al.** (1987). "Expression of human insulin-like growth factor I in bacteria: use of optimized gene fusion vectors to facilitate protein purification." *Biochemistry* 26(17): 5239-5244.
- Moller, D. E., Yokota, A., Caro, J. F. and Flier, J. S.** (1989). "Tissue-specific expression of two alternatively spliced insulin receptor mRNAs in man." *Mol Endocrinol* 3(8): 1263-1269.
- Moore, T., Constancia, M., Zubair, M., Bailleul, B., Feil, R., Sasaki, H. and Reik, W.** (1997). "Multiple imprinted sense and antisense transcripts, differential methylation and tandem repeats in a putative imprinting control region upstream of mouse *Igf2*." *Proc Natl Acad Sci U S A* 94(23): 12509-12514.
- Mosthaf, L., Grako, K., Dull, T. J., Coussens, L., Ullrich, A. and McClain, D. A.** (1990). "Functionally distinct insulin receptors generated by tissue-specific alternative splicing." *The Embo Journal* 9(8): 2409-2413.
- Mukkala, V. M., Mikola, H. and Hemmila, I.** (1989). "The synthesis and use of activated N-benzyl derivatives of diethylenetriaminetetraacetic acids: Alternative

- reagents for labeling of antibodies with metal ions." *Analytical Biochemistry* 176(2): 319-325.
- Myal, Y., Shiu, R. P., Bhaumick, B. and Bala, M.** (1984). "Receptor binding and growth-promoting activity of insulin-like growth factors in human breast cancer cells (T-47D) in culture." *Cancer Res* 44(12 Pt 1): 5486-5490.
- Mynarcik, D. C., Williams, P. F., Schaffer, L., Yu, G. Q. and Whittaker, J.** (1997). "Identification of common ligand binding determinants of the insulin and insulin-like growth factor 1 receptors. Insights into mechanisms of ligand binding." *J Biol Chem* 272(30): 18650-18655.
- Nagamatsu, S., Chan, S. J., Falkmer, S. and Steiner, D. F.** (1991). "Evolution of the insulin gene superfamily. Sequence of a preproinsulin-like growth factor cDNA from the Atlantic hagfish." *J Biol Chem* 266(4): 2397-2402.
- Nakano, E. T., R. D. Fox, D.E. Clements, K. Koo, W. D. Stuart and J. M. Ivy** (1993). "Expression vectors for *Neurospora crassa* and expression of a bovine preprochymosin cDNA." *Fungal genet. Newslett.* 40: 54-56.
- Neuenschwander, S., Roberts, C. T., Jr. and LeRoith, D.** (1995). "Growth inhibition of MCF-7 breast cancer cells by stable expression of an insulin-like growth factor I receptor antisense ribonucleic acid." *Endocrinology* 136(10): 4298-4303.
- Ng, P. C. and Henikoff, S.** (2001). "Predicting deleterious amino acid substitutions." *Genome Res* 11(5): 863-874.
- Ng, P. C. and Henikoff, S.** (2003). "SIFT: Predicting amino acid changes that affect protein function." *Nucleic Acids Res* 31(13): 3812-3814.
- Ng, P. C. and Henikoff, S.** (2006). "Predicting the effects of amino acid substitutions on protein function." *Annu Rev Genomics Hum Genet* 7: 61-80.
- Nguyen, A. W. and Daugherty, P. S.** (2003). "Production of randomly mutated plasmid libraries using mutator strains." *Methods Mol Biol* 231: 39-44.
- Ninomiya, Y., Suzuki, K., Ishii, C. and Inoue, H.** (2004). "Highly efficient gene replacements in *Neurospora* strains deficient for nonhomologous end-joining." *Proc Natl Acad Sci U S A* 101(33): 12248-12253.
- Nussbaum, T., Samarin, J., Ehemann, V., Bissinger, M., Ryschich, E., Khamidjanov, A., Yu, X., Gretz, N., Schirmacher, P. and Breuhahn, K.** (2008). "Autocrine insulin-like growth factor-II stimulation of tumor cell migration is a progression step in human hepatocarcinogenesis." *Hepatology* 48(1): 146-156.
- O'Brien, J., Wilson, I., Orton, T. and Pognan, F.** (2000). "Investigation of the Alamar Blue (resazurin) fluorescent dye for the assessment of mammalian cell cytotoxicity." *Eur J Biochem* 267(17): 5421-5426.
- O'Sullivan, D. C., Szostak, T. A. and Pell, J. M.** (2002). "Regulation of IGF-I mRNA by GH: putative functions for class 1 and 2 message." *Am J Physiol Endocrinol Metab* 283(2): E251-258.
- Orr-Weaver, T. L. and Szostak, J. W.** (1985). "Fungal recombination." *Microbiol Rev* 49(1): 33-58.
- Ostermeier, M., Shim, J. H. and Benkovic, S. J.** (1999). "A combinatorial approach to hybrid enzymes independent of DNA homology." *Nat Biotechnol* 17(12): 1205-1209.
- Otten, L. G., Sio, C. F., Vrielink, J., Cool, R. H. and Quax, W. J.** (2002). "Altering the substrate specificity of cephalosporin acylase by directed evolution of the Beta - subunit." *J Biol Chem* 277(44): 42121-42127.
- Paisley, A. N., Trainer, P. and Drake, W.** (2004). "Pegvisomant: a novel pharmacotherapy for the treatment of acromegaly." *Expert Opin Biol Ther* 4(3): 421-425.

- Pandini, G., Conte, E., Medico, E., Sciacca, L., Vigneri, R. and Belfiore, A.** (2004). "IGF-II binding to insulin receptor isoform A induces a partially different gene expression profile from insulin binding." *Ann N Y Acad Sci* 1028: 450-456.
- Pandini, G., Frasca, F., Mineo, R., Sciacca, L., Vigneri, R. and Belfiore, A.** (2002). "Insulin/insulin-like growth factor I hybrid receptors have different biological characteristics depending on the insulin receptor isoform involved." *J Biol Chem* 277(42): 39684-39695.
- Pashmforoush, M., Yoshimasa, Y. and Steiner, D. F.** (1994). "Exon 11 enhances insulin binding affinity and tyrosine kinase activity of the human insulin proreceptor." *The Journal of Biological Chemistry* 269(51): 32639-32648.
- Pelletier, J. N.** (2001). "A RACHITT for our toolbox." *Nat Biotechnol* 19(4): 314-315.
- Pollak, M.** (2008a). "Insulin and insulin-like growth factor signalling in neoplasia." *Nat Rev Cancer* 8(12): 915-928.
- Pollak, M.** (2008b). "Insulin, insulin-like growth factors and neoplasia." *Best Pract Res Clin Endocrinol Metab* 22(4): 625-638.
- Pollak, M.** (2008c). "Targeting insulin and insulin-like growth factor signalling in oncology." *Curr Opin Pharmacol* 8(4): 384-392.
- Pompon, D. and Nicolas, A.** (1989). "Protein engineering by cDNA recombination in yeasts: shuffling of mammalian cytochrome P-450 functions." *Gene* 83(1): 15-24.
- Rajewsky, K., Forster, I. and Cumano, A.** (1987). "Evolutionary and somatic selection of the antibody repertoire in the mouse." *Science* 238(4830): 1088-1094.
- Rall, L. B., Scott, J. and Bell, G. I.** (1987). "Human insulin-like growth factor I and II messenger RNA: isolation of complementary DNA and analysis of expression." *Methods Enzymol* 146: 239-248.
- Rasmussen-Wilson, S. J., Palas, J. S., Wolf, V. J., Taft, C. S. and Selitrennikoff, C. P.** (1997). "Expression of a plant protein by *Neurospora crassa*." *Appl Environ Microbiol* 63(9): 3488-3493.
- Rasmussen, J. P., Bowring, F. J., Yeadon, P. J. and Catcheside, D. E.** (2002). "Targeting vectors for gene diversification by meiotic recombination in *Neurospora crassa*." *Plasmid* 47(1): 18-25.
- Reinecke, M. and Collet, C.** (1998). "The phylogeny of the insulin-like growth factors." *Int Rev Cytol* 183: 1-94.
- Renehan, A. G., Zwahlen, M., Minder, C., O'Dwyer, S. T., Shalet, S. M. and Egger, M.** (2004). "Insulin-like growth factor (IGF)-I, IGF binding protein-3, and cancer risk: systematic review and meta-regression analysis." *Lancet* 363(9418): 1346-1353.
- Rinderknecht, E. and Humbel, R. E.** (1978a). "The amino acid sequence of human insulin-like growth factor I and its structural homology with proinsulin." *J Biol Chem* 253(8): 2769-2776.
- Rinderknecht, E. and Humbel, R. E.** (1978b). "Primary structure of human insulin-like growth factor II." *FEBS Lett* 89(2): 283-286.
- Ripoche, M. A., Kress, C., Poirier, F. and Dandolo, L.** (1997). "Deletion of the H19 transcription unit reveals the existence of a putative imprinting control element." *Genes Dev* 11(12): 1596-1604.
- Roberts, C. T., Jr., Lasky, S. R., Lowe, W. L., Jr., Seaman, W. T. and LeRoith, D.** (1987). "Molecular cloning of rat insulin-like growth factor I complementary deoxyribonucleic acids: differential messenger ribonucleic acid processing and regulation by growth hormone in extrahepatic tissues." *Mol Endocrinol* 1(3): 243-248.
- Roberts, R. W. and Szostak, J. W.** (1997). "RNA-peptide fusions for the in vitro selection of peptides and proteins." *Proc Natl Acad Sci U S A* 94(23): 12297-12302.

- Rochester, M. A., Riedemann, J., Hellawell, G. O., Brewster, S. F. and Macaulay, V. M.** (2005). "Silencing of the IGF1R gene enhances sensitivity to DNA-damaging agents in both PTEN wild-type and mutant human prostate cancer." *Cancer Gene Ther* 12(1): 90-100.
- Rohrer, J., Schweizer, A., Johnson, K. F. and Kornfeld, S.** (1995). "A determinant in the cytoplasmic tail of the cation-dependent mannose 6-phosphate receptor prevents trafficking to lysosomes." *J Cell Biol* 130(6): 1297-1306.
- Romero, P. A. and Arnold, F. H.** (2009). "Exploring protein fitness landscapes by directed evolution." *Nat Rev Mol Cell Biol* 10(12): 866-876.
- Rosa, A. L., Folco, H. D. and Mautino, M. R.** (2004). "In vivo levels of S-adenosylmethionine modulate C:G to T:A mutations associated with repeat-induced point mutation in *Neurospora crassa*." *Mutat Res* 548(1-2): 85-95.
- Rosenzweig, S. A. and Atreya, H. S.** (2010). "Defining the pathway to insulin-like growth factor system targeting in cancer." *Biochem Pharmacol* 80(8): 1115-1124.
- Ross, R. J., Leung, K. C., Maamra, M., Bennett, W., Doyle, N., Waters, M. J. and Ho, K. K.** (2001). "Binding and functional studies with the growth hormone receptor antagonist, B2036-PEG (pegvisomant), reveal effects of pegylation and evidence that it binds to a receptor dimer." *J Clin Endocrinol Metab* 86(4): 1716-1723.
- Rossier, C., Oulevey, N. and Turian, G.** (1973). "Electron microscopy of selectively stimulated microconidiogenesis in wild type *Neurospora crassa*." *Arch Mikrobiol* 91(4): 345-353.
- Roth, B. V., Burgisser, D. M., Luthi, C. and Humbel, R. E.** (1991). "Mutants of human insulin-like growth factor II: expression and characterization of analogs with a substitution of TYR27 and/or a deletion of residues 62-67." *Biochem Biophys Res Commun* 181(2): 907-914.
- Rotwein, P.** (1991). "Structure, evolution, expression and regulation of insulin-like growth factors I and II." *Growth Factors* 5(1): 3-18.
- Rotwein, P. and Hall, L. J.** (1990). "Evolution of insulin-like growth factor II: characterization of the mouse IGF-II gene and identification of two pseudo-exons." *DNA Cell Biol* 9(10): 725-735.
- Rotwein, P., Pollock, K. M., Didier, D. K. and Krivi, G. G.** (1986). "Organization and sequence of the human insulin-like growth factor I gene. Alternative RNA processing produces two insulin-like growth factor I precursor peptides." *J Biol Chem* 261(11): 4828-4832.
- Rountree, M. R. and Selker, E. U.** (1997). "DNA methylation inhibits elongation but not initiation of transcription in *Neurospora crassa*." *Genes Dev* 11(18): 2383-2395.
- Sachdev, D. and Yee, D.** (2007). "Disrupting insulin-like growth factor signaling as a potential cancer therapy." *Mol Cancer Ther* 6(1): 1-12.
- Sakano, K., Enjoh, T., Numata, F., Fujiwara, H., Marumoto, Y., Higashihashi, N., Sato, Y., Perdue, J. F. and Fujita-Yamaguchi, Y.** (1991). "The design, expression, and characterization of human insulin-like growth factor II (IGF-II) mutants specific for either the IGF-II/cation-independent mannose 6-phosphate receptor or IGF-I receptor." *J Biol Chem* 266(31): 20626-20635.
- Saltiel, A. R. and Pessin, J. E.** (2002). "Insulin signaling pathways in time and space." *Trends Cell Biol* 12(2): 65-71.
- Samani, A. A., Yakar, S., LeRoith, D. and Brodt, P.** (2007). "The role of the IGF system in cancer growth and metastasis: overview and recent insights." *Endocr Rev* 28(1): 20-47.



- Sandhu, M. S., Dunger, D. B. and Giovannucci, E. L.** (2002). "Insulin, insulin-like growth factor-I (IGF-I), IGF binding proteins, their biologic interactions, and colorectal cancer." *Journal of the National Cancer Institute* 94(13): 972-980.
- Sato, A., Nishimura, S., Ohkubo, T., Kyogoku, Y., Koyama, S., Kobayashi, M., Yasuda, T. and Kobayashi, Y.** (1993). "Three-dimensional structure of human insulin-like growth factor-I (IGF-I) determined by 1H-NMR and distance geometry." *Int J Pept Protein Res* 41(5): 433-440.
- Saupe, S. J.** (2000). "Molecular genetics of heterokaryon incompatibility in filamentous ascomycetes." *Microbiol Mol Biol Rev* 64(3): 489-502.
- Schaffer, L.** (1994). "A model for insulin binding to the insulin receptor." *Eur J Biochem* 221(3): 1127-1132.
- Schaffer, M. L., Deshayes, K., Nakamura, G., Sidhu, S. and Skelton, N. J.** (2003). "Complex with a phage display-derived peptide provides insight into the function of insulin-like growth factor I." *Biochemistry* 42(31): 9324-9334.
- Schmidt, B., Kiecke-Siensen, C., Waheed, A., Brulke, T. and von Figura, K.** (1995). "Localization of the insulin-like growth factor II binding site to amino acids 1508-1566 in repeat 11 of the mannose 6-phosphate/insulin-like growth factor II receptor." *J Biol Chem* 270(25): 14975-14982.
- Schreuder, M. P., Brekelmans, S., van den Ende, H. and Klis, F. M.** (1993). "Targeting of a heterologous protein to the cell wall of *Saccharomyces cerevisiae*." *Yeast* 9(4): 399-409.
- Sciacca, L., Costantino, A., Pandini, G., Mineo, R., Frasca, F., Scalia, P., Sbraccia, P., Goldfine, I. D., Vigneri, R., Belfiore, A., Vella, V. and Mazzon, E.** (1999). "Insulin receptor activation by IGF-II in breast cancers: evidence for a new autocrine/paracrine mechanism." *Oncogene* 18(15): 2471-2479.
- Sciacca, L., Mineo, R., Pandini, G., Murabito, A., Vigneri, R. and Belfiore, A.** (2002). "In IGF-I receptor-deficient leiomyosarcoma cells autocrine IGF-II induces cell invasion and protection from apoptosis via the insulin receptor isoform A." *Oncogene* 21(54): 8240-8250.
- Scotlandi, K., Benini, S., Nanni, P., Lollini, P. L., Nicoletti, G., Landuzzi, L., Serra, M., Manara, M. C., Picci, P. and Baldini, N.** (1998). "Blockage of insulin-like growth factor-I receptor inhibits the growth of Ewing's sarcoma in athymic mice." *Cancer Res* 58(18): 4127-4131.
- Seino, S. and Bell, G. I.** (1989). "Alternative splicing of human insulin receptor messenger RNA." *Biochemical and Biophysical Research Communications* 159(1): 312-316.
- Seino, S., Seino, M., Nishi, S. and Bell, G. I.** (1989). "Structure of the human insulin receptor gene and characterization of its promoter." *Proceedings of the National Academy of Sciences of the United States of America* 86(1): 114-118.
- Selker, E. U.** (1997). "Epigenetic phenomena in filamentous fungi: useful paradigms or repeat-induced confusion?" *Trends Genet* 13(8): 296-301.
- Selker, E. U., Cambareri, E. B., Jensen, B. C. and Haack, K. R.** (1987). "Rearrangement of duplicated DNA in specialized cells of *Neurospora*." *Cell* 51(5): 741-752.
- Selker, E. U. and Stevens, J. N.** (1985). "DNA methylation at asymmetric sites is associated with numerous transition mutations." *Proc Natl Acad Sci U S A* 82(23): 8114-8118.
- Sell, C., Dumenil, G., Deveaud, C., Miura, M., Coppola, D., DeAngelis, T., Rubin, R., Efstratiadis, A. and Baserga, R.** (1994). "Effect of a null mutation of the insulin-

- like growth factor I receptor gene on growth and transformation of mouse embryo fibroblasts." *Mol Cell Biol* 14(6): 3604-3612.
- Sell, C., Rubini, M., Rubin, R., Liu, J. P., Efstratiadis, A. and Baserga, R.** (1993). "Simian virus 40 large tumor antigen is unable to transform mouse embryonic fibroblasts lacking type 1 insulin-like growth factor receptor." *Proc Natl Acad Sci U S A* 90(23): 11217-11221.
- Sen, S., Venkata Dasu, V. and Mandal, B.** (2007). "Developments in directed evolution for improving enzyme functions." *Appl Biochem Biotechnol* 143(3): 212-223.
- Shao, Y., Gao, Z., Marks, P. A. and Jiang, X.** (2004). "Apoptotic and autophagic cell death induced by histone deacetylase inhibitors." *Proc Natl Acad Sci U S A* 101(52): 18030-18035.
- Shimatsu, A. and Rotwein, P.** (1987a). "Mosaic evolution of the insulin-like growth factors. Organization, sequence, and expression of the rat insulin-like growth factor I gene." *J Biol Chem* 262(16): 7894-7900.
- Shimatsu, A. and Rotwein, P.** (1987b). "Sequence of two rat insulin-like growth factor I mRNAs differing within the 5' untranslated region." *Nucleic Acids Res* 15(17): 7196.
- Shimotohno, A., Oue, S., Yano, T., Kuramitsu, S. and Kagamiyama, H.** (2001). "Demonstration of the importance and usefulness of manipulating non-active-site residues in protein design." *J Biochem* 129(6): 943-948.
- Shoichet, B. K., Baase, W. A., Kuroki, R. and Matthews, B. W.** (1995). "A relationship between protein stability and protein function." *Proc Natl Acad Sci U S A* 92(2): 452-456.
- Siddle, K., Urso, B., Niesler, C. A., Cope, D. L., Molina, L., Surinya, K. H. and Soos, M. A.** (2001). "Specificity in ligand binding and intracellular signalling by insulin and insulin-like growth factor receptors." *Biochem Soc Trans* 29(Pt 4): 513-525.
- Singer, M. J., Kuzminova, E. A., Tharp, A., Margolin, B. S. & Selker, E. U.** (1995). "Different frequencies of rip among early vs late ascospores of *Neurospora crassa*." *Fungal Genet. Newslett.* 42: 74-75.
- Singer, M. J., Marcotte, B. A. and Selker, E. U.** (1995). "DNA methylation associated with repeat-induced point mutation in *Neurospora crassa*." *Mol Cell Biol* 15(10): 5586-5597.
- Singh, B., Halestrap, A. P. and Paraskeva, C.** (1997). "Butyrate can act as a stimulator of growth or inducer of apoptosis in human colonic epithelial cell lines depending on the presence of alternative energy sources." *Carcinogenesis* 18(6): 1265-1270.
- Slaaby, R., Schaffer, L., Lautrup-Larsen, I., Andersen, A. S., Shaw, A. C., Mathiasen, I. S. and Brandt, J.** (2006). "Hybrid receptors formed by insulin receptor (IR) and insulin-like growth factor I receptor (IGF-IR) have low insulin and high IGF-1 affinity irrespective of the IR splice variant." *J Biol Chem* 281(36): 25869-25874.
- Sleckman, B. P., Gorman, J. R. and Alt, F. W.** (1996). "Accessibility control of antigen-receptor variable-region gene assembly: role of cis-acting elements." *Annu Rev Immunol* 14: 459-481.
- Smith, B. J., Huang, K., Kong, G., Chan, S. J., Nakagawa, S., Menting, J. G., Hu, S. Q., Whittaker, J., Steiner, D. F., Katsoyannis, P. G., Ward, C. W., Weiss, M. A. and Lawrence, M. C.** (2010). "Structural resolution of a tandem hormone-binding element in the insulin receptor and its implications for design of peptide agonists." *Proc Natl Acad Sci U S A* 107(15): 6771-6776.
- Smith, G. P.** (1985). "Filamentous fusion phage: novel expression vectors that display cloned antigens on the virion surface." *Science* 228(4705): 1315-1317.

- Smith, J. M.** (1970). "Natural selection and the concept of a protein space." *Nature* 225(5232): 563-564.
- Soares, M. B., Turken, A., Ishii, D., Mills, L., Episkopou, V., Cotter, S., Zeitlin, S. and Efstratiadis, A.** (1986). "Rat insulin-like growth factor II gene. A single gene with two promoters expressing a multitranscript family." *J Mol Biol* 192(4): 737-752.
- Soos, M. A., Field, C. E. and Siddle, K.** (1993). "Purified hybrid insulin/insulin-like growth factor-I receptors bind insulin-like growth factor-I, but not insulin, with high affinity." *The Biochemical Journal* 290 ( Pt 2): 419-426.
- Soos, M. A. and Siddle, K.** (1989). "Immunological relationships between receptors for insulin and insulin-like growth factor I. Evidence for structural heterogeneity of insulin-like growth factor I receptors involving hybrids with insulin receptors." *Biochem J* 263(2): 553-563.
- Sorensen, H., Whittaker, L., Hinrichsen, J., Groth, A. and Whittaker, J.** (2004). "Mapping of the insulin-like growth factor II binding site of the Type I insulin-like growth factor receptor by alanine scanning mutagenesis." *FEBS Lett* 565(1-3): 19-22.
- Stemmer, W. P.** (1994a). "DNA shuffling by random fragmentation and reassembly: in vitro recombination for molecular evolution." *Proc Natl Acad Sci U S A* 91(22): 10747-10751.
- Stemmer, W. P.** (1994b). "Rapid evolution of a protein in vitro by DNA shuffling." *Nature* 370(6488): 389-391.
- Stone, P. J., Makoff, A. J., Parish, J. H. and Radford, A.** (1993). "Cloning and sequence analysis of the glucoamylase gene of *Neurospora crassa*." *Curr Genet* 24(3): 205-211.
- Surinya, K. H., Forbes, B. E., Occhiodoro, F., Booker, G. W., Francis, G. L., Siddle, K., Wallace, J. C. and Cosgrove, L. J.** (2008). "An investigation of the ligand binding properties and negative cooperativity of soluble insulin-like growth factor receptors." *J Biol Chem* 283(9): 5355-5363.
- Tanori, M., Santone, M., Mancuso, M., Pasquali, E., Leonardi, S., Di Majo, V., Rebessi, S., Saran, A. and Pazzaglia, S.** (2010). "Developmental and oncogenic effects of insulin-like growth factor-I in Ptc1+/- mouse cerebellum." *Mol Cancer* 9: 53.
- Taverna, D. M. and Goldstein, R. A.** (2002). "Why are proteins marginally stable?" *Proteins* 46(1): 105-109.
- Tawfik, D. S. and Griffiths, A. D.** (1998). "Man-made cell-like compartments for molecular evolution." *Nat Biotechnol* 16(7): 652-656.
- Terao, Y., Miyamoto, K. and Ohta, H.** (2006). "Improvement of the activity of arylmalonate decarboxylase by random mutagenesis." *Appl Microbiol Biotechnol* 73(3): 647-653.
- Terasawa, H., Kohda, D., Hatanaka, H., Nagata, K., Higashihashi, N., Fujiwara, H., Sakano, K. and Inagaki, F.** (1994). "Solution structure of human insulin-like growth factor II; recognition sites for receptors and binding proteins." *Embo J* 13(23): 5590-5597.
- Thorvaldsen, J. L., Duran, K. L. and Bartolomei, M. S.** (1998). "Deletion of the H19 differentially methylated domain results in loss of imprinted expression of H19 and Igf2." *Genes Dev* 12(23): 3693-3702.
- Tiekink, E. R. T., Hall, V. J., Buntine, M. A. & Hook, J.** (2000). "Examination of the effect of crystal packing forces on geometric parameters: a combined crystallographic and theoretical study of 2,2'-bipyridyl adducts of R2SnCl2." *Zeitschrift für Kristallographie* 215: 23-33.

- Torres, A. M., Forbes, B. E., Aplin, S. E., Wallace, J. C., Francis, G. L. and Norton, R. S.** (1995). "Solution structure of human insulin-like growth factor II. Relationship to receptor and binding protein interactions." *J Mol Biol* 248(2): 385-401.
- Tracewell, C. A. and Arnold, F. H.** (2009). "Directed enzyme evolution: climbing fitness peaks one amino acid at a time." *Curr Opin Chem Biol* 13(1): 3-9.
- Travers, K. J., Patil, C. K., Wodicka, L., Lockhart, D. J., Weissman, J. S. and Walter, P.** (2000). "Functional and genomic analyses reveal an essential coordination between the unfolded protein response and ER-associated degradation." *Cell* 101(3): 249-258.
- Troll, C., Alexander, D., Allen, J., Marquette, J. and Camps, M.** (2011). "Mutagenesis and functional selection protocols for directed evolution of proteins in *E. coli*." *J Vis Exp*(49).
- Tsuchiya, Y., Morioka, K., Shirai, J., Yokomizo, Y. and Yoshida, K.** (2003). "Gene design of signal sequence for effective secretion of protein." *Nucleic Acids Res Suppl*(3): 261-262.
- Tsuchiya, Y., Morioka, K., Yoshida, K., Shirai, J., Kokuho, T. and Inumaru, S.** (2007). "Effect of N-terminal mutation of human lysozyme on enzymatic activity." *Nucleic Acids Symp Ser (Oxf)*(51): 465-466.
- Ullrich, A., Gray, A., Tam, A. W., Yang-Feng, T., Tsubokawa, M., Collins, C., Henzel, W., Le Bon, T., Kathuria, S., Chen, E. and et al.** (1986). "Insulin-like growth factor I receptor primary structure: comparison with insulin receptor suggests structural determinants that define functional specificity." *EMBO J* 5(10): 2503-2512.
- Underwood, L. E., Thissen, J. P., Lemozy, S., Ketelslegers, J. M. and Clemmons, D. R.** (1994). "Hormonal and nutritional regulation of IGF-I and its binding proteins." *Horm Res* 42(4-5): 145-151.
- Vajdos, F. F., Ultsch, M., Schaffer, M. L., Deshayes, K. D., Liu, J., Skelton, N. J. and de Vos, A. M.** (2001). "Crystal structure of human insulin-like growth factor-1: detergent binding inhibits binding protein interactions." *Biochemistry* 40(37): 11022-11029.
- Valkonen, M., Penttila, M. and Saloheimo, M.** (2003a). "Effects of inactivation and constitutive expression of the unfolded- protein response pathway on protein production in the yeast *Saccharomyces cerevisiae*." *Appl Environ Microbiol* 69(4): 2065-2072.
- Valkonen, M., Ward, M., Wang, H., Penttila, M. and Saloheimo, M.** (2003b). "Improvement of foreign-protein production in *Aspergillus niger* var. *awamori* by constitutive induction of the unfolded-protein response." *Appl Environ Microbiol* 69(12): 6979-6986.
- Van den Berg, C. L., Cox, G. N., Stroh, C. A., Hilsenbeck, S. G., Weng, C. N., McDermott, M. J., Pratt, D., Osborne, C. K., Coronado-Heinsohn, E. B. and Yee, D.** (1997). "Polyethylene glycol conjugated insulin-like growth factor binding protein-1 (IGFBP-1) inhibits growth of breast cancer in athymic mice." *Eur J Cancer* 33(7): 1108-1113.
- van den Hombergh, J. P., van de Vondervoort, P. J., Fraissinet-Tachet, L. and Visser, J.** (1997). "Aspergillus as a host for heterologous protein production: the problem of proteases." *Trends Biotechnol* 15(7): 256-263.
- Vann, D. C.** (1995). "Electroporation-based transformation of freshly harvested conidia of *Neurospora crassa*." *Fungal Genet. Newsl.* 42A: 53.
- Veldhuis, J. D., Bidlingmaier, M., Anderson, S. M., Wu, Z. and Strasburger, C. J.** (2001). "Lowering total plasma insulin-like growth factor I concentrations by way of a novel, potent, and selective growth hormone (GH) receptor antagonist, pegvisomant (B2036-peg), augments the amplitude of GH secretory bursts and elevates

- basal/nonpulsatile GH release in healthy women and men." *J Clin Endocrinol Metab* 86(7): 3304-3310.
- Vella, V., Pandini, G., Sciacca, L., Mineo, R., Vigneri, R., Pezzino, V. and Belfiore, A.** (2002). "A novel autocrine loop involving IGF-II and the insulin receptor isoform-A stimulates growth of thyroid cancer." *J Clin Endocrinol Metab* 87(1): 245-254.
- Verdoes, J. C., Punt, P. J., Stouthamer, A. H. and van den Hondel, C. A.** (1994). "The effect of multiple copies of the upstream region on expression of the *Aspergillus niger* glucoamylase-encoding gene." *Gene* 145(2): 179-187.
- Vogt, B., Carrascosa, J. M., Ermel, B., Ullrich, A. and Haring, H. U.** (1991). "The two isotypes of the human insulin receptor (HIR-A and HIR-B) follow different internalization kinetics." *Biochemical and Biophysical Research Communications* 177(3): 1013-1018.
- Voigt, C. A., Kauffman, S. and Wang, Z. G.** (2000). "Rational evolutionary design: the theory of in vitro protein evolution." *Adv Protein Chem* 55: 79-160.
- Wang, B. Q., Lei, L. and Burton, Z. F.** (1994). "Importance of codon preference for production of human RAP74 and reconstitution of the RAP30/74 complex." *Protein Expr Purif* 5(5): 476-485.
- Wang, C. L., Yang, D. C. and Wabl, M.** (2004a). "Directed molecular evolution by somatic hypermutation." *Protein Eng Des Sel* 17(9): 659-664.
- Wang, L., Jackson, W. C., Steinbach, P. A. and Tsien, R. Y.** (2004b). "Evolution of new nonantibody proteins via iterative somatic hypermutation." *Proc Natl Acad Sci U S A* 101(48): 16745-16749.
- Wang, N., Yoshida, Y. and Hasunuma, K.** (2007). "Loss of Catalase-1 (Cat-1) results in decreased conidial viability enhanced by exposure to light in *Neurospora crassa*." *Mol Genet Genomics* 277(1): 13-22.
- Wang, X., Minasov, G. and Shoichet, B. K.** (2002). "Evolution of an antibiotic resistance enzyme constrained by stability and activity trade-offs." *J Mol Biol* 320(1): 85-95.
- Ward, C., Lawrence, M., Streltsov, V., Garrett, T., McKern, N., Lou, M. Z., Lovrecz, G. and Adams, T.** (2008). "Structural insights into ligand-induced activation of the insulin receptor." *Acta Physiol (Oxf)* 192(1): 3-9.
- Ward, C. W. and Lawrence, M. C.** (2009). "Ligand-induced activation of the insulin receptor: a multi-step process involving structural changes in both the ligand and the receptor." *Bioessays* 31(4): 422-434.
- Watters, M. K., Randall, T. A., Margolin, B. S., Selker, E. U. and Stadler, D. R.** (1999). "Action of repeat-induced point mutation on both strands of a duplex and on tandem duplications of various sizes in *Neurospora*." *Genetics* 153(2): 705-714.
- Werner, H. and Bruchim, I.** (2009). "The insulin-like growth factor-I receptor as an oncogene." *Arch Physiol Biochem* 115(2): 58-71.
- Werner, H. and LeRoith, D.** (1996). "The role of the insulin-like growth factor system in human cancer." *Adv Cancer Res* 68: 183-223.
- Whittaker, J., Groth, A. V., Mynarcik, D. C., Pluzek, L., Gadsboll, V. L. and Whittaker, L. J.** (2001). "Alanine scanning mutagenesis of a type 1 insulin-like growth factor receptor ligand binding site." *J Biol Chem* 276(47): 43980-43986.
- Whittaker, J., Sorensen, H., Gadsboll, V. L. and Hinrichsen, J.** (2002). "Comparison of the functional insulin binding epitopes of the A and B isoforms of the insulin receptor." *The Journal of Biological Chemistry* 277(49): 47380-47384.
- Whittaker, L., Hao, C., Fu, W. and Whittaker, J.** (2008). "High-affinity insulin binding: insulin interacts with two receptor ligand binding sites." *Biochemistry* 47(48): 12900-12909.

- Whitten, A. E., Smith, B. J., Menting, J. G., Margetts, M. B., McKern, N. M., Lovrecz, G. O., Adams, T. E., Richards, K., Bentley, J. D., Trehwella, J., Ward, C. W. and Lawrence, M. C.** (2009). "Solution structure of ectodomains of the insulin receptor family: the ectodomain of the type 1 insulin-like growth factor receptor displays asymmetry of ligand binding accompanied by limited conformational change." *J Mol Biol* 394(5): 878-892.
- Wilson, D. S. and Keefe, A. D.** (2001). "Random mutagenesis by PCR." *Curr Protoc Mol Biol* Chapter 8: Unit8 3.
- Windhofer, F., Catcheside, D. E. and Kempken, F.** (2000). "Methylation of the foreign transposon Restless in vegetative mycelia of *Neurospora crassa*." *Curr Genet* 37(3): 194-199.
- Wittrup, K. D.** (2001). "Protein engineering by cell-surface display." *Curr Opin Biotechnol* 12(4): 395-399.
- Wosten, H. A., Moukha, S. M., Sietsma, J. H. and Wessels, J. G.** (1991). "Localization of growth and secretion of proteins in *Aspergillus niger*." *J Gen Microbiol* 137(8): 2017-2023.
- Yamaguchi, Y., Flier, J. S., Benecke, H., Ransil, B. J. and Moller, D. E.** (1993). "Ligand-binding properties of the two isoforms of the human insulin receptor." *Endocrinology* 132(3): 1132-1138.
- Yandell, C. A., Francis, G. L., Wheldrake, J. F. and Upton, Z.** (1998). "Purification, amino acid sequence and characterisation of kangaroo IGF-I." *J Endocrinol* 156(1): 195-204.
- Yeadon, P. J., Bowring, F. J. and Catcheside, D. E.** (2004a). "Alleles of the hotspot cog are codominant in effect on recombination in the his-3 region of *Neurospora*." *Genetics* 167(3): 1143-1153.
- Yeadon, P. J., Bowring, F. J. and Catcheside, D. E.** (2004b). "Sequence heterology and gene conversion at his-3 of *Neurospora crassa*." *Curr Genet* 45(5): 289-301.
- Yeadon, P. J., Bowring, F. J. and Catcheside, D. E.** (2010). "High density analysis of randomly selected *Neurospora* octads reveals conversion associated with crossovers located between cog and his-3." *Fungal Genet Biol* 47(10): 847-854.
- Yeadon, P. J. and Catcheside, D. E.** (1995). "The chromosomal region which includes the recombinator cog in *Neurospora crassa* is highly polymorphic." *Curr Genet* 28(2): 155-163.
- Yeadon, P. J. and Catcheside, D. E.** (1998). "Long, interrupted conversion tracts initiated by cog in *Neurospora crassa*." *Genetics* 148(1): 113-122.
- Yeadon, P. J. and Catcheside, D. E.** (1999). "Polymorphism around cog extends into adjacent structural genes." *Curr Genet* 35(6): 631-637.
- Yeadon, P. J., Koh, L. Y., Bowring, F. J., Rasmussen, J. P. and Catcheside, D. E.** (2002). "Recombination at his-3 in *Neurospora* declines exponentially with distance from the initiator, cog." *Genetics* 162(2): 747-753.
- Yeadon, P. J., Rasmussen, J. P. and Catcheside, D. E.** (2001). "Recombination events in *Neurospora crassa* may cross a translocation breakpoint by a template-switching mechanism." *Genetics* 159(2): 571-579.
- Yee, D., Jackson, J. G., Kozelsky, T. W. and Figueroa, J. A.** (1994). "Insulin-like growth factor binding protein 1 expression inhibits insulin-like growth factor I action in MCF-7 breast cancer cells." *Cell Growth Differ* 5(1): 73-77.
- Yin, J., Li, G., Ren, X. and Herrler, G.** (2007). "Select what you need: a comparative evaluation of the advantages and limitations of frequently used expression systems for foreign genes." *J Biotechnol* 127(3): 335-347.

- Yuan, L., Kurek, I., English, J. and Keenan, R.** (2005). "Laboratory-directed protein evolution." *Microbiol Mol Biol Rev* 69(3): 373-392.
- Yuen, C. M. and Liu, D. R.** (2007). "Dissecting protein structure and function using directed evolution." *Nat Methods* 4(12): 995-997.
- Zeslawski, W., Beisel, H. G., Kamionka, M., Kalus, W., Engh, R. A., Huber, R., Lang, K. and Holak, T. A.** (2001). "The interaction of insulin-like growth factor-I with the N-terminal domain of IGFBP-5." *Embo J* 20(14): 3638-3644.
- Zhang, J. H., Dawes, G. and Stemmer, W. P.** (1997a). "Directed evolution of a fucosidase from a galactosidase by DNA shuffling and screening." *Proc Natl Acad Sci U S A* 94(9): 4504-4509.
- Zhang, L., Zhou, W., Velculescu, V. E., Kern, S. E., Hruban, R. H., Hamilton, S. R., Vogelstein, B. and Kinzler, K. W.** (1997b). "Gene expression profiles in normal and cancer cells." *Science* 276(5316): 1268-1272.
- Zhang, W., Gustafson, T. A., Rutter, W. J. and Johnson, J. D.** (1994). "Positively charged side chains in the insulin-like growth factor-1 C- and D-regions determine receptor binding specificity." *J Biol Chem* 269(14): 10609-10613.
- Zhao, H.** (2007). "Directed evolution of novel protein functions." *Biotechnol Bioeng* 98(2): 313-317.
- Zhao, H. and Arnold, F. H.** (1999). "Directed evolution converts subtilisin E into a functional equivalent of thermitase." *Protein Eng* 12(1): 47-53.
- Zhao, H., Giver, L., Shao, Z., Affholter, J. A. and Arnold, F. H.** (1998). "Molecular evolution by staggered extension process (StEP) in vitro recombination." *Nat Biotechnol* 16(3): 258-261.
- Zhao, H. and Zha, W.** (2006). "In vitro 'sexual' evolution through the PCR-based staggered extension process (StEP)." *Nat Protoc* 1(4): 1865-1871.
- Zia, F., Jacobs, S., Kull, F., Jr., Cuttitta, F., Mulshine, J. L. and Moody, T. W.** (1996). "Monoclonal antibody alpha IR-3 inhibits non-small cell lung cancer growth in vitro and in vivo." *J Cell Biochem Suppl* 24: 269-275.
- Zumarraga, M., Camarero, S., Shleev, S., Martinez-Arias, A., Ballesteros, A., Plou, F. J. and Alcalde, M.** (2008). "Altering the laccase functionality by in vivo assembly of mutant libraries with different mutational spectra." *Proteins* 71(1): 250-260.
- Zumstein, P. P., Luthi, C. and Humbel, R. E.** (1985). "Amino acid sequence of a variant pro-form of insulin-like growth factor II." *Proc Natl Acad Sci U S A* 82(10): 3169-3172.