INVESTIGATION OF INSULIN-LIKE RECEPTOR SYSTEMS

THE UNIVERSITY
OF ADELAIDE
AUSTRALIA

THIS THESIS IS SUBMITTED IN PART OF THE REQUIREMENTS FOR THE DEGREE OF DOCTORATE OF PHILOSOPHY IN THE SCHOOL OF MOLECULAR AND BIOMEDICAL SCIENCE.

ERIC RICHARD BONYTHON
B.SC.(HON)

JUNE 2005

DECLARATION

To the best of my knowledge, this thesis does not contain any material previously submitted for a degree or diploma in any University or any material previously written or published by any other person, except where due reference is made in the text.

Eric Richard Bonython

June 16th 2005

NAME: **PROGRAM:** PhD

This work contains no material which has been accepted for the award of any other degree or diploma in any university or other tertiary institution and, to the best of my knowledge and belief, contains no material previously published or written by another person, except where due reference has been made in the text.

I give consent to this copy of my thesis, when deposited in the University Libraries, being available for photocopying and loan.

SIGNATURE: **DATE:** 3/11/05

"No regrets then?"
"A few. But just a few. You?"
"Oh, enough to fill a lifetime. So much has been lost, so much forgotten. So much pain, so much blood. And for what, I wonder. The past tempts us, the present confuses us, and the future frightens us. And our lives slip away, moment by moment, lost in that vast terrible inbetween. But there is still time to seize that one last fragile moment. To choose something better, to make a difference, as you say. And I intend to do just that."

BABYLON 5
Centauri Emperor and **Sheridan**, *The Coming of Shadows*
J. Michael Straczynski

AKNOWLEDGEMENTS

Firstly I would like to thank the then Department of Biochemistry and Professor Peter Rathjen for allowing me to begin a PhD, and of course my supervisors Dr. Grant Booker and Dr. Leah Cosgrove for accepting me. I thank them for helpful discussions on science and what to do with myself in the future. I should thank them for putting up with me for so long and being patient and understanding during my experimentally challenged moments (and months ☺). In the end I am just glad that my research has laid the groundwork for other peoples success with related projects.

The whole adventure would not have been as fun, or bearable without the support and kindness shown by the members of the Booker Lab, Cosgrove Lab, and Wallace Lab. I'm truly grateful to the help and friendship provided by everyone. Big thanks go to Cvetan for his friendship and technical expertise, Anita for amusing me (and sharing a bench), Sue and Carlie for their ears and interesting chats. Finally I must thank Wendy and my mum for all their love, support and encouragement over many years.

SUMMARY

The insulin and insulin-like growth factor receptor (IR and IGF-1R respectively) networks are ancient and fundamental systems that control growth and metabolism in multicellular organisms. This thesis has examined several aspects of this field focusing on mammalian receptor biology and a comparison of the similarities and differences between the insulin and IGF receptor signalling systems.

The insulin receptor family of proteins consist of eleven structural domains, of which the extracellular domains contain all the ligand binding and specificity determinants. The insert domain, within the extracellular region is the least understood of all the domains, and it has no similarity to any other protein sequence. It does however contain the cleavage site which separates the receptor into two subunits and also a small stretch of residues shown to directly contact bound ligand and which is absolutely required for ligand binding in short recombinant forms of the receptor. In addition, the human insulin receptor, expressed as one of two isoforms, A and B, results in the exclusion or inclusion of 12 amino acids directly adjacent to the ligand contacting amino acids in the insert domain. The A isoform lacking exon11 is expressed ubiquitously and the B isoform containing exon11 is co-expressed mainly in the traditional insulin responsive tissues of liver, muscle, adipocytes and kidney, where it is the dominant isoform.

In this thesis recombinant insert domain was expressed in a bacterial system in an attempt to purify folded protein suitable for NMR structural analysis. The results of the expression studies indicated that the insert domain was unstructured in isolation and was unable to be adequately refolded by all conditions tried, although hydrophobic conditions appeared to partially stabilize the structure. The overall conclusions of this project were that the Insert domain is likely to have limited structure, and probably buried within the receptor, and therefore requires the presence of the rest of the extracellular domains to adopt its correct structure.

A comparison of the ligand binding and phsophorylation potential between the two human isoforms of the insulin receptor was made. A competition binding assay using europium labelled insulin was developed, that found that both IGF-1 and IGF-2 had an increased affinity for the hIR-A, but insulin had a slightly reduced affinity. These results differ from the established literature in the raw values, however the relative ratios of binding strength are consistent. The

most likely reason for this is that the europium labelled insulin has a different mode of binding the receptors due to the location of the europium chelate. Interestingly, using europium labelled IGF-1 produced results nearly identical to those of conventional competition assays.

Phosphorylation assays indicated that the hIR-B isoform was more responsive than hIR-A. Even though IGF-2 and IGF-1 had improved affinity for hIR-A, the level of phosphorylation was not as high. The ability of each growth factor to promote cellular proliferation correlated well with the relative strength of binding and activation of the receptor.

The regions of the IR and IGF-1R involved in binding substrates and regulators are predominantly found in the juxtamembrane domain and the C-terminal domain, which contain several potential tyrosine and serine phosphorylation target sequences. In this study the effect of mutations in unique tyrosine residues and other residues in the C-terminal domain of the hIGF-1R was investigated. Results of time-course phosphorylation assays showed that mutation of Tyrosine^{1251} to phenylalanine caused hyperphosphorylation of the receptor and increased proliferation, which was caused by deregulation of a tyrosine phosphatase. A Tyrosine^{1250} to phenylalanine mutation had altered kinetics of phosphorylation, displaying an unchanging rate of phosphorylation over time after ligand stimulation. However, proliferation was unaltered, indicating that even under extended exposure to ligand, the initial strength of receptor activation is more critical to affecting the biological response.

The *Caenorhabditis elegans* insulin-like peptide family is a very large family consisting of possibly 38 peptides likely to be both agonists and antagonists of Daf-2 Receptor (IR homologue) signalling. Comparative modelling of all 38 peptides was performed based on the known structures of mammalian peptides. The overall results indicated that good quality models of ins peptides could be made despite the low sequence similarity with the templates. This suggested that it is the conformational shape of the molecule allowable by the individual residues that is most important when modelling and not having a perfect sequence match.

CHAPTER 1

INTRODUCTION

CHAPTER 1 INTRODUCTION

1.1 COMPONENTS OF THE INSULIN/IGF SYSTEM

The Insulin/Insulin-like Growth Factor (IGF) receptor networks are ancient and fundamental systems that control growth and metabolism in multicellular organisms.

This thesis has examined several aspects of this field including receptor structural biology, receptor-ligand interactions, receptor activation and the analysis of early representatives of insulin-like ligands. Due to the scope encompassed by these topics this main introduction will focus on mammalian receptor biology and a comparison of the similarities and differences between the Insulin and IGF receptor signalling systems.

In mammals and many other vertebrate species the Insulin/IGF system comprises multiple homologous receptors and ligands in addition to ligand specific 'Binding Proteins'. Via the intracellular protein kinase activity of the receptor this system feeds into a multitude of intracellular signalling pathways, through recruitment and phosphorylation of receptor substrates. In humans, three separate receptors were found that each bound insulin, IGF-1 and IGF-2 with preferential affinities (Jacobs *et al.,* 1977) (Massague *et al.*, 1982). Two of these receptors, hIR and hIGF-1R, were found to be homologous, while the third turned out to be the same protein as the unrelated mannose-6-phosphate receptor (M6PR/IGF-2R).

The occurrence of distinct IR and IGF-1R genes appears to coincide with the evolution of vertebrates from protochordates. The protochordate Amphioxus contains a single IR-like receptor cDNA whereas the most primitive extant vertebrate, the Hagfish, has two different IR-like cDNAs (Pashmforoush *et al.,* 1996). The number of insulin-like proteins found in a species appears to vary more dramatically ranging from 10 in humans, to an astonishing 38 in *Caenorhabditis elegans* (see section 6.1).

The general function of insulin in vertebrates is the regulation of glucose metabolism where it acts to increase the rate of glucose uptake in skeletal muscle and adipose tissue and inhibit the production of glucose in the liver. It also increases lipid formation in adipose tissue and at the same time prevents lipid breakdown. Deregulation of the Insulin system is the key area of malfunction in Type-1 and the highly prevalent Type-2 Diabetes Mellitus.

In contrast, IGF-1 and IGF-2 mainly function to stimulate cellular proliferation and inhibit apoptosis. They have additional context dependent effects including promoting cellular differentiation and increasing metastasis. Not surprisingly the upregulation of IGF system components including the ligands and the receptor have frequently been found to be associated with developing and mature cancers.

The Insulin and IGF systems are highly relevant areas of research that will lead to the eventual treatment of several major human diseases. Given the homology between both the ligands and the receptors and the overlapping activation of downstream signalling pathways, it is important to understand the way in which these two functionally distinct systems interact. This should allow the design of methods to manipulate one system and avoid a simultaneous detrimental effect to the other.

1.2 STRUCTURAL BIOLOGY OF INSULIN/IGF RECEPTORS

1.2.1 RECEPTOR GENE STRUCTURE

The human Insulin Receptor (hIR) and IGF-IR (hIGF-IR) cDNAs were cloned and sequenced in the mid 1980s (Ebina *et al.*, 1985) (Ullrich *et al.*, 1985) (Ullrich *et al.*, 1986) followed shortly by the sequencing of their respective genes (See Figure 1.1) (Seino *et al.*, 1989) (Abbott *et al.*, 1992).

Human insulin receptor is encoded by a single gene >130kb, located on chromosome 19, composing 22 exons and 21 introns (Seino, 1989). It was quickly recognized that the IR cDNAs sequenced earlier had been of an alternative isoform resulting from the exclusion of exon 11 and the loss of a 12 aa sequence near the end of the α–subunit in the insert domain (Seino *et al.*, 1989) (See sections 1.2.2, 1.2.3). The two isoforms were named hIR-A (exon11-) and hIR-B (exon11+). It was found that the alternative splicing is mainly controlled by two elements in intron 10. A 48-nucleotide purine rich sequence at the 5′ end acts as a splicing enhancer, increasing exon11 inclusion whereas a 43-nucleotide sequence at the 3′ end promotes skipping of exon11 (Kosaki *et al.*, 1998). The factors involved in controlling switching the splicing of exon11 have not been fully elucidated. However, it has been found that the aberrant phosphorylation and cellular localization of the CUG RNA Binding protein, CUG-BP, causes a switch in isoform expression in muscle from exon11+ to exon11-, and results in insulin resistance (Savkur *et al.*, 2001). Binding to the intronic sequence upstream of exon11 by CUG-BP was also demonstrated *in vitro*.

Similar to hIR, hIGF-IR has a ~100kb gene also on chromosome 19, comprising 21 exons and 20 introns (Abbott *et al.*, 1992). The hIGF-IR gene lacks an equivalent of the hIR exon 11 and

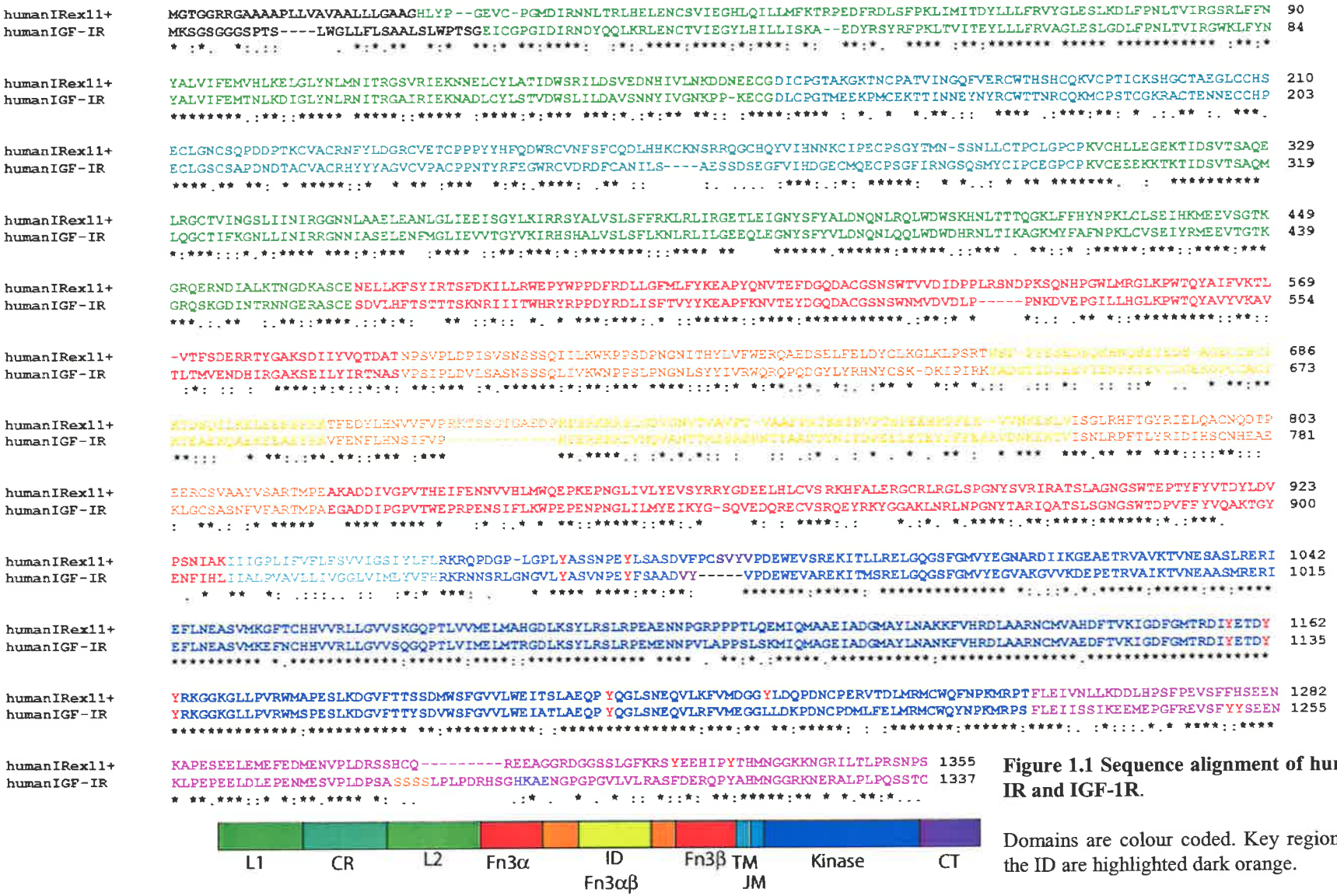
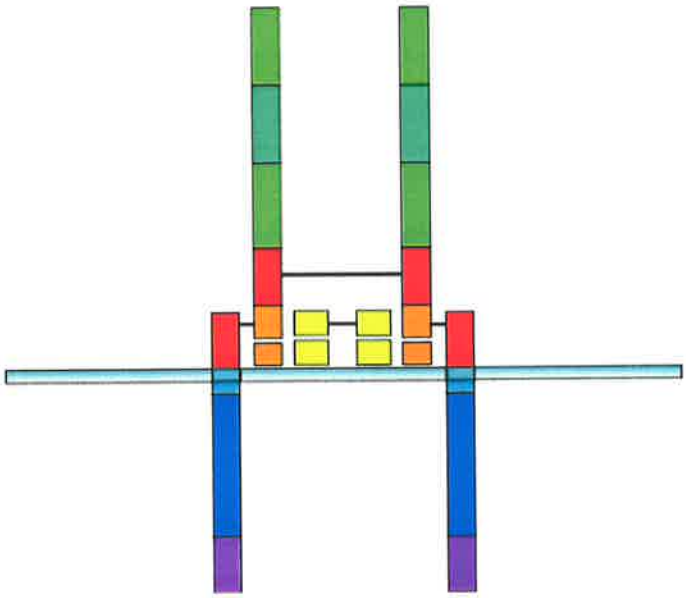


Figure 1.1 Sequence alignment of human IR and IGF-1R.

Domains are colour coded. Key regions in the ID are highlighted dark orange.



Domain	Number of residues (hIGF-IR/hIRex11+)	Identity between receptors	Known Function
L1	150/152 aa	69%	Ligand Binding
Cys Rich	149/152 aa	48%	Ligand Binding
L2	160/160 aa	64%	Ligand Binding
Fn3α	108/112 aa	57%	Scaffold
Fn3αβ1	63/64 aa	56%	Scaffold
ID	117/124 aa	35%	Ligand Binding
Fn3αβ2	39/39 aa	49%	Scaffold
Fn3β	108/109 aa	47%	Scaffold
Transmembrane	23/23 aa	22%	Membrane Placement
Juxtamembrane	26/25 aa	58%	Protein Docking
Kinase	273/278 aa	83%	Tyrosine Kinase
C-terminal	109/100 aa	45%	Protein Docking

Figure 1.2 Schematic diagram of the Insulin Receptor and the Type-1 IGF Receptor

1998) (Olson *et al.*, 1988). The two lectin chaperones, calnexin and calreticulin, together with the HSP70 homologue, BiP, also assist in the correct folding of the receptor and must dissociate from the receptor after folding has been completed to allow the final dimerisation to occur (Bass *et al.*, 1998). This is followed by transport from the endoplasmic reticulum to the cell surface membrane.

1.2.3.1 Hybrid Receptors

A consequence of the close similarity between the IGF-IR and the IR is the formation of hybrid receptors containing one of each receptor αβ monomer. The existence of two isoforms of the insulin receptor results in the potential occurrence of two types of hybrid receptor. Hybrids are naturally occurring where each receptor is expressed in a single cell. The wide distribution of both receptors results in a large amount of overlap in expression. In fact it was found in both rabbit and human tissues that hybrids occurred in every tissue examined ranging from skeletal muscle to brain tissue (Bailyes *et al.*, 1997). In the majority of these tissues over 50% of the available IGF-IR half-receptors were contained within hybrids, and hybrids appeared to be the dominant form in heart and muscle.

It has been shown that hybrid formation occurs by association of αβ monomers in the ER at the normal time of heterotetramer formation and that afterwards they are trafficked normally to the extracellular membrane. It is not entirely clear whether the formation of hybrids is random or slightly biased but the majority of reports provide evidence favouring random assembly (Belfiore *et al.*, 1999) (Bailyes *et al.*, 1997) (Pandini *et al.*, 1999). Receptor hybrids have been shown to have ligand binding profiles similar to those of IGF-IR, binding IGF-1 and IGF-2 better than insulin (Kasuya *et al.*, 1993). This suggests that insulin requires IR specific sequences from each monomer for high affinity binding. The two different hybrids have slightly different affinities for their three ligands. The hIR-A/IGF-IR hybrid was activated by IGF-I, IGF-II, and insulin, while the hIR-B/IGF-IR hybrid bound IGF-1 with high affinity, IGF-2 with low affinity and only very weakly to insulin (Pandini *et al.*, 2002).

Binding of ligand produces receptor autophosphorylation and activation of downstream signalling pathways similar to the IGF-IR, however the physiological significance of these receptors has yet to be properly assessed. Given their prevalence and seemingly wild-type activity, it is likely that they are responsible for a large part of the biological responses from IGF-1 and IGF-2, and depending on the hybrid, insulin as well, therefore care should be taken not to misinterpret data obtained from cell lines expressing significant numbers of both receptors. This also has a wider implication for the development of therapeutic agents against

the actions of IGF-1 and 2, particularly if the target site is in the intracellular regions of the receptor, as a hybrid will always retain a functionally active half.

1.2.4 THREE-DIMENSIONAL STRUCTURE OF THE RECEPTOR

A complete molecular structure of either the Insulin receptor or IGF-IR has yet to be solved. This has been mainly due to an inability to form appropriate crystals for X-ray crystallography analysis. However, the atomic structures of portions of each receptor have been solved using this method. The kinase domains of each receptor in both the activated and inactivated state have been determined (See section 1.3.1). They were found to be very similar to other tyrosine kinase domains and adopt the typical domain fold of this family (See section 1.3.1, figure 1.6). The only crystal structure of domains in the extracellular portion of the receptor was solved relatively recently, and consists of a monomeric structure of the first three domains of the human IGF-IR (Garrett *et al.*, 1998). Overall the three domains are roughly situated on three sides of a central cavity that would be large enough to accommodate an IGF ligand (See figure 1.3). Both the L1 and L2 domains are rectangular box-like structures formed mostly from perpendicular β sheets. The Cys-rich domain is a slightly extended rectangular structure with what appear to be flexible hinge regions connecting to both the L1 and L2 domains. Mapping insulin receptor mutations that effect insulin binding onto a homology model of the receptors first three domains suggests that there are distinct surfaces of the L1 and L2 domains involved in forming the ligand-binding pocket (See section 1.2.5.1, Table 1.2). The relative orientation of the 3 domains may not be representative of the native structure due to flexibility of the linking regions between each domain and the lack of the presence of the Fn3 domains and the insert domain. It should also be noted that this recombinant protein was unable to bind ligand as a C-terminal myc tag used in the purification was found to be sitting in the space bounded by the three domains.

Initial electron microscopy studies into the overall shape of the insulin receptor suggested that it adopted a 'V' or 'Y" shape in the cell membrane, presumably with the two arms representing each αβ half receptor (Schaefer *et al.*, 1992). More recently, imaging of single insulin receptor molecules bound to individual antibodies was examined using electron microscopy. The receptor resembled a U-shaped prism with a central cleft that is 30–40 Å wide, which is adequate to accommodate a molecule of insulin. The localization of the monoclonal antibodies indicated an antiparallel configuration of the two half-receptors (Tulloch *et al.*, 1999). The antibody 83-7 (binds the CR domain), bound the receptor in two places diametrically opposite each other half-way up each side-arm, compared to 83-14

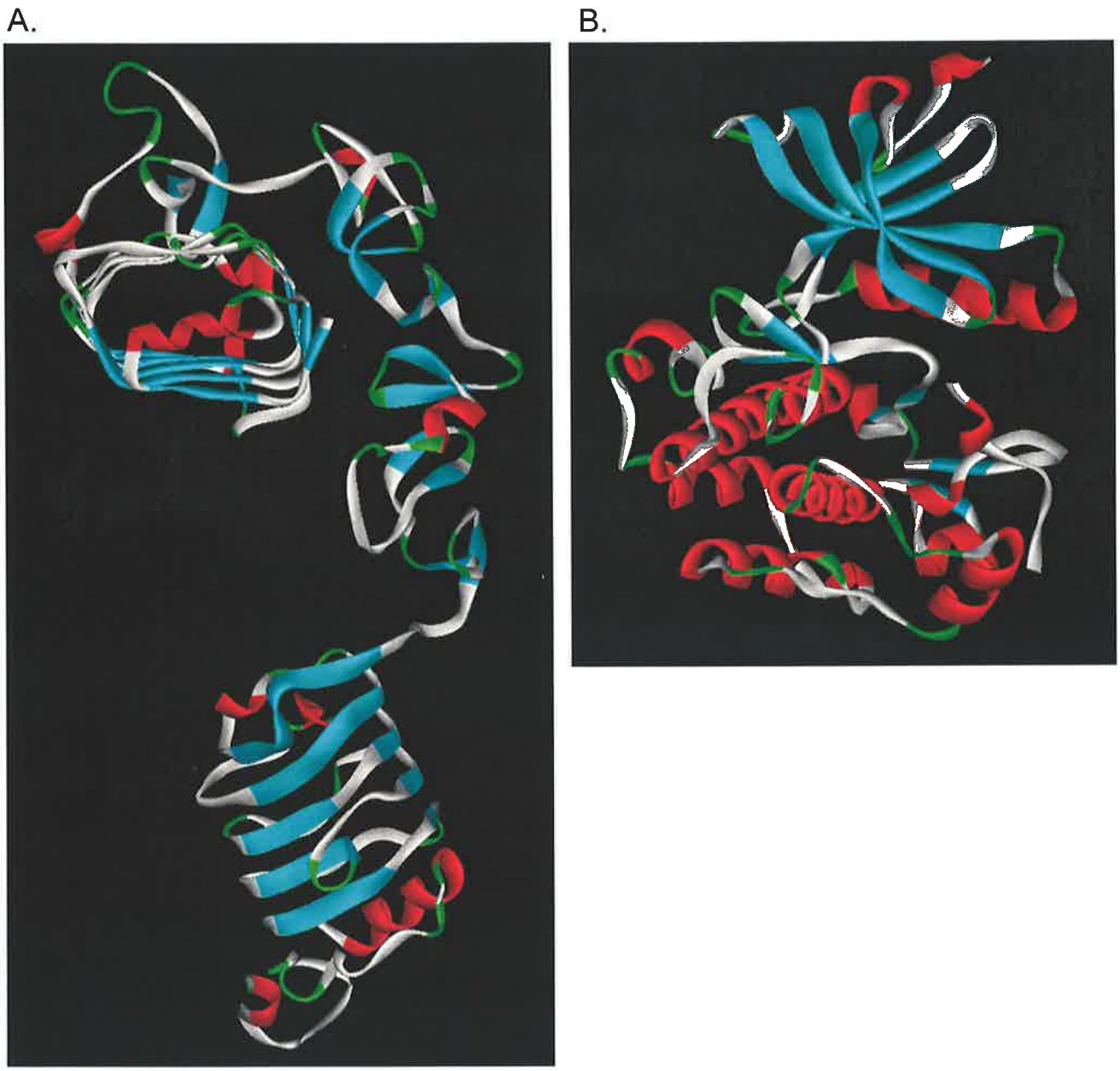


Figure 1.3 Solved Crystal Structures of the IGF-IR

A. L1/CR/L2 Domains (PDB: 1IGR), Garrett *et al.*, 1998
B. Tyrosine Kinase Domain (PDB: 1GAG, Perang *et al.*, 2001
Peptide Backbone is represented as a solid ribbon. β-Strands are coloured blue, α-helices are coloured red and β-turns are coloured green. (Garrett *et al.*, 1998)

(binds the Fn3α domain) and 18-44 (binds the ID), which both appeared to bind towards the base of the molecule. This points towards the L1, CR, and L2 domains occupying the top portion of the receptor and the Fn3 domains and the ID providing the base.

In theory this type of structure would allow the ligand simultaneous access to both halves of the receptor. More recent studies utilising electron cryomicroscopy and gold labelled insulin have proposed the relative orientation of the known domain structures in relation to the labelled insulin, based on space fill models and mutational data (Luo *et al.*, 1999) (Slade *et al.*, 2002). However, it must be noted that this structure was analysed without the presence of a plasma membrane and the overall structure is very compact and globular and quite unlike the extended structures seen in the earlier electron microscopy studies. Also, the relative positioning of the domains in these studies relied upon fitting homologous but dissimilar fibronectin domains, and also the unaltered crystal structure of the combined L1-CR-L2, which is unlikely to be in the correct orientation due to its inability to bind ligand and flexibility in the hinge regions between the domains. It therefore remains to be seen whether the receptor adopts a more compact structure upon binding ligand while membrane embedded.

1.2.4.1 Anti-receptor Antibodies

Antibodies raised against hIR and hIGF-1R have been used to purify and identify receptors, study ligand binding, the positional relationship between domains and for a clinical purpose as potent inhibitors of receptor function (Summarised in table 1.1). The vast majority of commonly used anti-receptor antibodies appear to affect ligand binding. As indicated in table 1.1, these effects can be inhibitory or enhancing, however in most cases triggering of receptor signalling can occur independently of ligand binding. Inhibitory antibodies block receptor access for the ligand but usually are themselves able to activate the receptor to varying degrees. These antibodies map to the extracellular regions of the receptor and individual antibodies have been found that bind to nearly all of the extracellular domains. Ligand binding enhancing antibodies are believed to function by altering the structure of the ligand-binding pocket, giving it higher affinity for the ligand. Interestingly, antibodies that enhance receptor phosphorylation independent of ligand binding have been found to bind to the receptor not only around the tyrosine kinase domain but also to some extracellular domains including the L1, CR, Fn3α and Fn3αβ (Nakae *et al.*, 1995) (Desbois-Mouthon *et al.*, 1996). This would suggest that the extracellular domain binding antibodies are affecting the conformational structure of the receptor, forcing it into an active form that would normally

Antibody	Epitope	Effect on ligand binding	Reference
IR		Insulin	
3D7	191-297 (CR)	no effect	Zhang and Roth, 1991
5D9	469-593 (Fn3α)	inhibits	Zhang and Roth, 1991
18-44	765-770 (ID)	inhibits	Taylor et al., 1987
25-49	469-593 (Fn3α)	inhibits	Taylor et al., 1987
47-9	469-593 (Fn3α)	inhibits	Taylor et al., 1987
83-7	191-297 (CR)	no effect	Taylor et al., 1987
83-14	469-593 (Fn3α)	inhibits	Taylor et al., 1987
MA-5	469-593 (Fn3α)	inhibits	Zhang and Roth, 1991
MA-10	469-593 (Fn3α)	inhibits	Zhang and Roth, 1991
MA-20	469-593 (Fn3α)	inhibits	Zhang and Roth, 1991
MC-51	469-593 (Fn3α)	inhibits	Zhang and Roth, 1991
IGF-1R		IGF-1	
IR-3	217-274 (CR)	inhibits	Gustafson and Roth,1990
4-52	62-184 (L1-CR)	enhances	Soos et al., 1992
16-13	62-184 (L1-CR)	enhances	Soos et al., 1992
17-69	514-586 (Fn3α)	inhibits	Soos et al., 1992, Schumacher et al., 1993
24-31	283-440 (CR-L2)	no effect	Soos et al., 1992, Schumacher et al., 1993
24-55	440-586 (L2-Fn3α)	inhibits	Soos et al., 1992, Schumacher et al., 1993
24-60	184-283 (CR)	inhibits	Soos et al., 1992, Schumacher et al., 1993
25-57	440-586 (L2-Fn3α)	inhibits	Soos et al., 1992, Schumacher et al., 1993
26-3	283-440 (CR-L2)	enhances	Soos et al., 1992, Schumacher et al., 1993

Table 1.1 Anti-receptor antibodies

only occur upon ligand binding, or it may also be allowing aggregation or cross-linking of receptors to occur.

1.2.5 MOLECULAR BASIS OF LIGAND BINDING

Given the similarity between the receptors and ligands, much research has been done looking into the regions and residues involved in directly binding ligand and in determining ligand specificity. The main methods employed have included using photo-affinity labelled ligands, receptor chimeras and mutants, and anti-receptor antibodies and cryoelectron microscopy. Unfortunately, as mentioned previously, a 3D structure of an insulin-like ligand bound to a receptor has not been able to be solved, so the specific molecular details remain unknown.

1.2.5.1 Role of the L1 Domain

Photoaffinity labelling human insulin receptor with the insulin derivative (B29-(2-nitro-4-azidophenyl)-biocytinyl-insulin) resulted in the isolation of a 14kDa receptor fragment beginning at Leu^{20} and ending ~80 residues later still within the L1 domain (Wedekind *et al.*, 1989). Further evidence to implicate the L1 domain as a binding determinant came from recombinant receptor chimera studies between the hIR and hIGF-1R. Exchanging the majority of the L1 domain between the receptors in either combination results in the loss of binding affinity for both Insulin and IGF-1 (Schumacher *et al.*, 1991)(Schumacher *et al.*, 1993). A number of mutations in hIR L1 domain have been found to affect Insulin binding (summarised in table 1.2). The residues exhibiting the most dramatic effects on binding were Arg^{14}, Asn^{15}, Phe^{64}, Arg^{86}, and Phe^{89}, of which only the latter two occur in hIR compared with hIGF-IR (Williams *et al.*, 1995) (De Meyts *et al.*, 1990). The majority of these residues map to a single surface on the crystal structure of the hIGF-1R L1 domain that faces towards the Insulin/IGF sized space between the domains. A recent cross-linking study using an insulin derivative has indicated that insulin's B chain alpha helix directly contacts the L1 beta-helix in the insulin receptor (Huang et al., 2004). Comparatively few mutations have been found in the hIGF-IR L1 domain that affect IGF-I binding. In fact only Asp^{8}, Asn^{11} and Phe^{58} (equivalent of hIR Asp^{12}, Asn^{15}, and Phe^{64}) had a moderate effect by decreasing IGF-1 affinity by 4-fold (Mynarcik *et al.*, 1997). It was further found that Tyr^{28}, His^{30}, Leu^{33}, Leu^{56}, Arg^{59}, and Phe^{90} also reduced ligand affinity and along with Asp^{8}, Asn^{11} and Phe^{58} map to the same surface as the IR mutations (Whittaker *et al.*, 2001). It therefore appears that the L1 domain functions primarily as a binding and specificity determinant for insulin and the IGFs. Recent alanine scanning mutagenesis of the IGF-1R looking for residues affecting IGF-2 binding has found that N-terminal L1 domain residues and C-terminal residues of the Insert Domain in the α-subunit are a critical part of the epitope (Sorensen *et al.*, 2004).

	Mutation	Effects	References
hIR			
L1 domain	$Asn^{15}Lys$ (Nat)	decreased affinity(5X),processing, surface expression	Kadowaki, 1995, JBC
	$Asn^{15}Ala$	decreased affinity (250X)	Williams, 1995,JBC
	$Asp^{12}Ala$	decreased affinity	Williams, 1995,JBC
	$Ile^{13}Ala$	decreased affinity	Williams, 1995,JBC
	Arg^{14} Ala	decreased affinity(700X)	Williams, 1995,JBC
	$Gln^{34}Ala$	decreased affinity	Williams, 1995,JBC
	$Leu^{36}Ala$	decreased affinity	Williams, 1995,JBC
	$Met^{38}Ala$	decreased affinity	Williams, 1995,JBC
	$Phe^{39}Ala$	decreased affinity	Williams, 1995,JBC
	$Glu^{44}Ala$	decreased affinity	Williams, 1995,JBC
	$Phe^{64}Ala$	decreased affinity (>700X)	Williams, 1995,JBC
	$phe^{89}Ala$	decreased affinity	Williams, 1995,JBC
	$Phe^{89}Leu$	no binding	Rouard, 1999, JBC
	$Asn^{90}Ala$	decreased affinity	Williams, 1995,JBC
	$Tyr^{91}Ala$	decreased affinity	Williams, 1995,JBC
	$Asp^{59}Gly$ (Nat)	decreased binding (X4)	Rouard, 1999, JBC
	$Leu^{62}Pro$ (Nat)	folding,oligomerisation,surface transport	Rouard, 1999, JBC
	$Arg^{86}Pro$ (Nat)	no binding, some constitutive receptor activity,leprechaunism	Nakae, 1995, JBC
	$Leu^{87}Pro$ (Nat)	leprechaunism	Nakae, 1995, JBC
	$Leu^{87}Ile$	increased affinity (X4)	Nakae, 1995, JBC
	$Leu^{87}Ala$	decreased affinity (X6)	Nakae, 1995, JBC
hIGF-IR			
L1 domain	$Asp^{8}Ala$	decreased affinity (X4)	Mynarcik,1997,JBC
	$Asn^{11}Ala$	decreased affinity (X4)	Mynarcik,1997,JBC
	$Tyr^{28}Ala$	decreased affinity	Whittaker,2001,JBC
	$His^{30}Ala$	decreased affinity	Whittaker,2001,JBC
	$Leu^{33}Ala$	decreased affinity	Whittaker,2001,JBC
	$Leu^{56}Ala$	decreased affinity	Whittaker,2001,JBC
	$Phe^{58}Ala$	decreased affinity (X4)	Mynarcik,1997,JBC
	$Arg^{59}Ala$	decreased affinity	Whittaker,2001,JBC
	$Phe^{90}Ala$	decreased affinity	Whittaker,2001,JBC
	$Trp^{79}Ala$	decreased affinity	Whittaker,2001,JBC

Table 1.2 Effects of Mutations in the L1 domains of the IR and IGF-IR

1.2.5.2 Role of the Cys-Rich Domain

The Cys-rich domain module has also been found to occur in the related EGF Receptor family of proteins (Bajaj *et al.*, 1987). In this family it is also located between the first two L domains however additional modules occur afterwards in place of the Fn3 domains. Iodinated Lys B29 Insulin, bound a 23kDa fragment, amino acids 203-316 from the CR domain to the start of L2, implicating the CR domain as forming part of the ligand binding pocket (Yip *et al.*, 1988). Few mutational studies have focused on the CR domain, however a $Pro^{243}Arg/Pro^{244}Arg/His^{246}Asp$ triple mutant was found to increase the affinity of the Insulin Receptor for insulin two to three fold (Rafaeloff *et al.*, 1989). A chimeric hIR with the whole CR domain and flanking regions of both L1 and L2 exchanged with hIGF-1R retained high affinity insulin binding while at the same time gaining high affinity IGF-1 binding. The corresponding hIGF-1R chimera displayed the opposite properties losing affinity for both ligands. The same results were achieved when an exchange of both the L1 and L2 domains was made. Similar results were obtained when slightly different exchange variations were used with soluble ectodomain receptors, thereby avoiding the formation of hybrid receptors to complicate experiments performed with membrane based full-length receptors (Reviewed in (Adams *et al.*, 2000)). Overall these experiments indicated that high affinity insulin binding and specificity required the presence of elements from L1 and L2 domains from the insulin receptor while IGF-1 required the CR domain from the IGF-1R. Disrupting the interplay between the L1 and L2 domains by only exchanging one domain severely affected insulin and IGF-1 binding, implying that part of the binding pocket is reliant on the correct relationship between L1 and L2 and another region by the CR domain. A deletion of Asn^{281} resulted in decreased ligand affinity but constitutive kinase activity, indicating that the CR domain plays an important role in the conformational switch between the active and inactive states (Desbois-Mouthon *et al.*, 1996). In the IGF-IR alanine mutagenesis of Arg^{240}, Phe^{241}, Glu^{242}, and Phe^{251}, together with Trp^{79}, of the L1 domain appear to form a patch at the interface of the L1 and CR domains that effect IGF binding (See table 1.3)(Whittaker *et al.*, 2001). Recent mutagenesis studies however, indicate that the CR domain does not play a major role for IGF-2 binding specifically (Sorensen *et al.*, 2004).

1.2.5.3 Role of the L2 Domain

A contact site for the ectodomain of the insulin receptor was found to lie within an 18kDa fragment in the L2 domain downstream of Gly^{390}, extending approximately to Arg^{488} near the

hIR			
CR domain	$Pro^{243}Arg/Pro^{244}Arg/His^{246}Asp$	increased affinity (X2-3)	Rafaeloff, 1989
	Asn^{281} deletion	decreased affinity, constitutive kinase domain	Desbois-Mouthon, 1996
hIGF-IR			
CR domain	$Arg^{240}Ala$	decreased affinity	Whittaker,2001
	$Phe^{241}Ala$	decreased affinity	Whittaker,2001
	$Glu^{242}Ala$	decreased affinity	Whittaker,2001
	$Phe^{251}Ala$	decreased affinity	Whittaker,2001
hIR			
L2 domain	$Ser^{323}Leu$	no binding	Roach, 1994
	$Gly^{366}Arg$ (Nat)	defective proreceptor processing, transport	Taylor, 1994
	$Phe^{382}Val$ (Nat)	defective proreceptor processing, transport, no insulin triggered autophosphorylation, conformationally restrained	Taylor, 1994
	$Trp^{412}Ser$ (Nat)	defective proreceptor processing, transport, no insulin triggered autophosphorylation	Taylor, 1994
	$Lys^{460}Glu$ (Nat)	increased receptor degradation	Kadowaki, 1995
	$Asn^{462}Ser$ (Nat)	increased receptor degradation	Kadowaki, 1995
	$Val^{335}del$ (Nat)	no binding	Johansen,2003
hIR			
Insert Domain	$Thr^{704}Ala$	decreased affinity	Mynarcik, 1996
	$Phe^{705}Ala$	decreased affinity	Mynarcik, 1996
	$Leu^{709}Ala$	decreased affinity	Mynarcik, 1996
	$His^{710}Ala$	decreased affinity	Mynarcik, 1996
	$Asn^{711}Ala$	decreased affinity	Mynarcik, 1996
	$Phe^{714}Ala$	decreased affinity	Mynarcik, 1996
hIGF-IR			
Insert Domain	$Phe^{701}Ala$	No Binding	Mynarcik,1997

Table 1.3 Effects of Mutations in the CR, L2, and Insert domains of the IR and IGF-IR

start of the FnIIIα domain (Fabry *et al.*, 1992).This was identified using the photoreactive insulin (4-azidosalicyloyl(B1-biocytinyl-B2-lysine)-insulin). It is therefore likely that the L2 domain participates in the formation of the ligand-binding pocket.

1.2.5.4 Role of the Fibronectin Type-III Domains

There are no reported mutations in Fn3α or Fn3αβ domains in either IR or IGF-1R that affect ligand binding, processing or display. The cysteines in this region involved in disulphide bonding, however, have been mapped (Sparrow *et al.*, 1997). These commonly found fibronectin domains function as a scaffold or stabilising structure, which in the case of IR and IGF-1R are used to build the ligand binding domains and allow activation of the tyrosine kinase domain upon ligand binding. Interestingly, all of the antibodies which bind to Fn3 domains listed in table 1.1 are able to inhibit ligand binding. This could suggest that their binding induces a conformational change in the binding site or that the position of the antibodies when bound inhibits access to the site.

1.2.5.5 Role of the Insert Domain

In a surprising result, an additional region in the insulin receptor involved in directly binding insulin was identified from an experiment involving the receptor cross-linking of another biotin-labelled photoreactive insulin derivative (despentapeptide-(B26-B30) (B25 p-azidophenylalanine-alpha-carboxamide) insulin). After Lys-C endoproteinase digestion of the receptor/ligand complex, the smallest fragment found bound to insulin comprised residues 704-718, which lie within the insert domain, directly adjacent to the residues encoded by exon 11 and the furin cleavage site (Kurose *et al.*, 1994)(See figure 1.4). This suggested a role for at least this region of the Insert Domain in forming part of the ligand binding pocket for insulin.

Confirming the results of the previous experiment, alanine scanning mutagenesis of this 13 aa stretch at the C-terminus of the α-subunit identified several residues which when mutated severely affected insulin binding affinity (Mynarcik *et al.*, 1996). These were Thr^{704}, Phe^{705}, Glu^{706}, Tyr^{708}, Leu^{709}, His^{710}, Asn^{711}, and Phe^{714} (See Figure 1.4). Since this experiment was carried out using the soluble ectodomain form of hIR, which only exhibits low affinity one-site, binding it can be concluded that these residues form part of the initial site for the receptor-ligand interaction. A comparable experiment in the IGF-1R found that only residue Phe^{701} (equivalent of hIR Phe^{714}) had a large effect on IGF-1 affinity, however this resulted in undetectable binding (Mynarcik *et al.*, 1997).

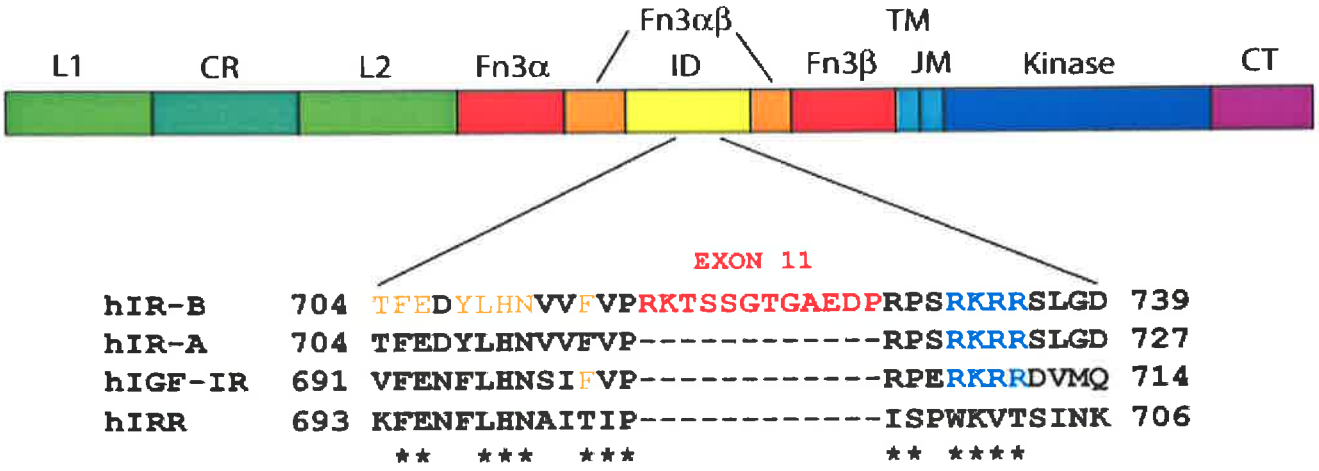


Figure 1.4 Comparison of the sequences surrounding exon 11.

hIR-B, hIR-A, hIGF-IR, and hIRR are shown. Exon 11 residues are coloured red. The residies constituting the furin cleavage site are coloured blue. The residues found to significantly affect ligand binding when mutated to alanine in hIR-B and hIGF-IR are coloured orange [Mynarcik *et al.*, 1996,1997].

Swapping the C-terminal 16 residues of the alpha subunit between hIR-A, hIGF-IR, and hIRR only abolished binding of the cognate ligand when the hIRR stretch was used, highlighting the importance of the Phe^{714}/Phe^{701} residue, which in hIRR is a Threonine (Kristensen *et al.*, 1999). It is therefore clear that this stretch of residues is involved in the binding of both insulin and IGF-1 though to a much greater extent in insulin binding.

After the initial identification of the two alternative isoforms of the Insulin Receptor, various studies were performed to ascertain whether there was any functional difference. Ligand binding experiments using only insulin and IGF-1 revealed a very little difference in insulin affinity and a moderate 10 fold increase in IGF-1 affinity for hIR-A (exon 11-) over hIR-B (exon 11+) (Yamaguchi *et al.*, 1993). However the increase in IGF-1 affinity was not considered sufficient enough to make hIR-A physiologically relevant for this ligand. Coupled with observations that hIR-A had a reduced signalling capability (Condorelli *et al.*, 1994) and that tissue expression of the isoform was limited to non-classical insulin responsive tissues, was a virtual end to attempts to assess its biological significance. It wasn't until evidence began mounting that both IGF-1 and IGF-2 were likely to have significant effects through the insulin receptor during embryogenesis, and that the residues directly upstream were directly involved in the receptor/ligand interaction that the binding properties were re-investigated, this time comparing insulin and IGF-2. It was found that in experiments conducted on mouse embryo fibroblasts derived from a mouse with a targeted disruption of the IGF-IR, that IGF-2 had an affinity for hIR-A comparable to that of hIGF-1R (Frasca *et al.*, 1999).

It is important to note that the chimeric receptor experiments discussed in section 1.2.5 were performed using the hIR-A isoform and therefore probably underplayed the importance of this additional region in ligand binding and also did not examine the binding of IGF-2.

1.2.5.6 Different Mechanisms of Binding between the Insulin Receptor and IGF-IR

It seems clear that, although the ligands rely on similar areas of each receptor to form the binding site, they are not coincident. Various studies analysing the kinetics of ligand binding have found that insulin binds the insulin receptor with a curvilinear scatchard plot, indicating a two-site binding mechanism. In addition, at insulin concentrations above 100nM, negative-cooperativity is apparent, resulting in competition experiments showing an increase in the amount of tracer binding (DeMeyts, 1994). This is due both to the binding of up to three molecules of insulin to the receptor and the formation of dimers and hexamers of insulin that

may also bind the receptor. The biological consequence of insulin dimer binding was tested *in vitro* with a covalently dimerised insulin derivative, which although binding to the receptor with a similar affinity to monomeric insulin, lead to reduced activation of glucose transporters and DNA synthesis (Joost *et al.*, 1989) (Weiland *et al.*, 1990). Insulin forms dimers only at micromolar concentration and will subsequently form hexamers in the presence of zinc ions however blood insulin concentrations remain sub-nanomolar even after a massive increase after eating. This however does not rule out having significance for instances of higher local tissue concentrations. In contrast IGF-1, exhibits a linear scatchard, signifying one-site binding, however it remains unclear as to whether the receptor can allow the simultaneous binding of two IGF molecules (Christoffersen *et al.*, 1994).

1.2.5.7 Minimized Ligand Binding Receptor

In an effort to further identify critical regions of the IR and IGF-IR required for high affinity ligand binding and to create a recombinant receptor suitable for X-ray crystallographic analysis, many studies have been performed to minimise the size of a ligand binding receptor fragment. Since only the extracellular portion of the receptor has been found to have a direct effect on ligand binding the obvious first step was to express a form of the receptor truncated before the start of the transmembrane domain. For both hIR and hIGF-IR this ectodomain form of the receptor was found to have a reduced affinity for their respective cognate ligands, and in the case of the hIR_{ED} also showed a linear scatchard plot, indicating the loss of a two-site interaction with insulin (Cosgrove *et al.*, 1995)(Jansson *et al.*, 1997).

The ligand binding affinity of the ectodomain form of the insulin receptor was found to be improved by the C-terminal addition of either a fragment of the self-associating constant domain of mouse γ-globulin (mFc) or a 33 residue leucine zipper from the *S. cerevisiae* transcriptional activator GCN4. Not only do these recombinant receptors bind insulin with an affinity comparable to wild-type membrane anchored receptor, but also unlike ectodomain receptors they exhibit a curvilinear scatchard (Hoyne *et al.*, 2000).

Interestingly, the smallest recombinant receptor fragment capable of binding ligand with near wild-type affinity consists of the first 255 residues, covering the L1, and the majority of the CR, in addition to 16 residues from the Insert Domain (aa 704-719, see 1.2.2.5) (Kristensen *et al.*, 2002). This receptor binds insulin with an apparent affinity of 11±4 nM, which is only 5-10 fold lower than wild-type, suggesting that the majority of the binding pocket is formed from the first two domains and the small section from the Insert domain. The major function of the other extracellular domains therefore appears to be to correctly orient the binding

pocket and to accommodate the conformational change required to activate the kinase domains.

1.2.6 Structure and Function of Insulin-like Proteins

Insulin-like proteins have been found in all species where an insulin-like receptor has been identified indicating the maintained co-evolution of this insulin system and the requirement for conservation of receptor and ligand structure in the binding interaction. Insulin-like proteins are generally characterised by a conserved pattern of six cysteine residues.

The human gene for insulin is situated on the short arm of chromosome 11, and directly adjacent to its 3' end lies the gene for IGF-2, while the gene for IGF-1 maps to the long arm of chromosome 12 (Werner *et al.*, 1994). Transcription of all the genes can occur from multiple promoters, and can be regulated in a developmental and tissue-specific manner. Insulin and the IGFs are highly conserved among vertebrate species, and there is even enough similarity with more distantly related species such as *Aedes aegypti* (mosquito) or *C. elegans* (nematode worm) for them to remain biologically active when expressed in these systems (Graf *et al.*, 1997) (Pierce *et al.*, 2001).

Insulin is synthesised with an N-terminal signal peptide and a C domain separating the B and A domains, which is subsequently removed during processing of the protein by carboxypeptidase H (Davidson *et al.*, 1987). IGFs also have a signal peptide and C domain in addition to a C-terminal D domain. Unlike insulin however, the C domain is not removed during processing of IGFs and from NMR structural analysis appears to be a relatively unstructured loop between the B and A chains (Cooke *et al.*, 1991) (Sato *et al.*, 1992) (Sato *et al,* 1993) (Torres *et al.*, 1995). The insulin-like domain consists of core 'B' and 'A' chains which fold together to form a conformationally restrained molecule consisting of two shorter, roughly anti-parallel alpha helices from the A chain, lying perpendicular to a third longer alpha helix in the B chain. A more detailed description of the involvement of residues in the function of insulin-like proteins is given in section 6.1 (See Figure 1.5). There also exist several alternate unprocessed extended versions of IGF-2 containing an extra C-terminal O-glycosylated E domain (Haselbacher and Humbel, 1982) (Hudgins *et al.*, 1992). These forms of IGF-2 are mitogenically active and do not bind IGFBPs, however their physiological role is unclear (Blahovec *et al.*, 2001). It has been suggested that they may be developmentally more important as this long form is more bioavailable during embryogenesis.

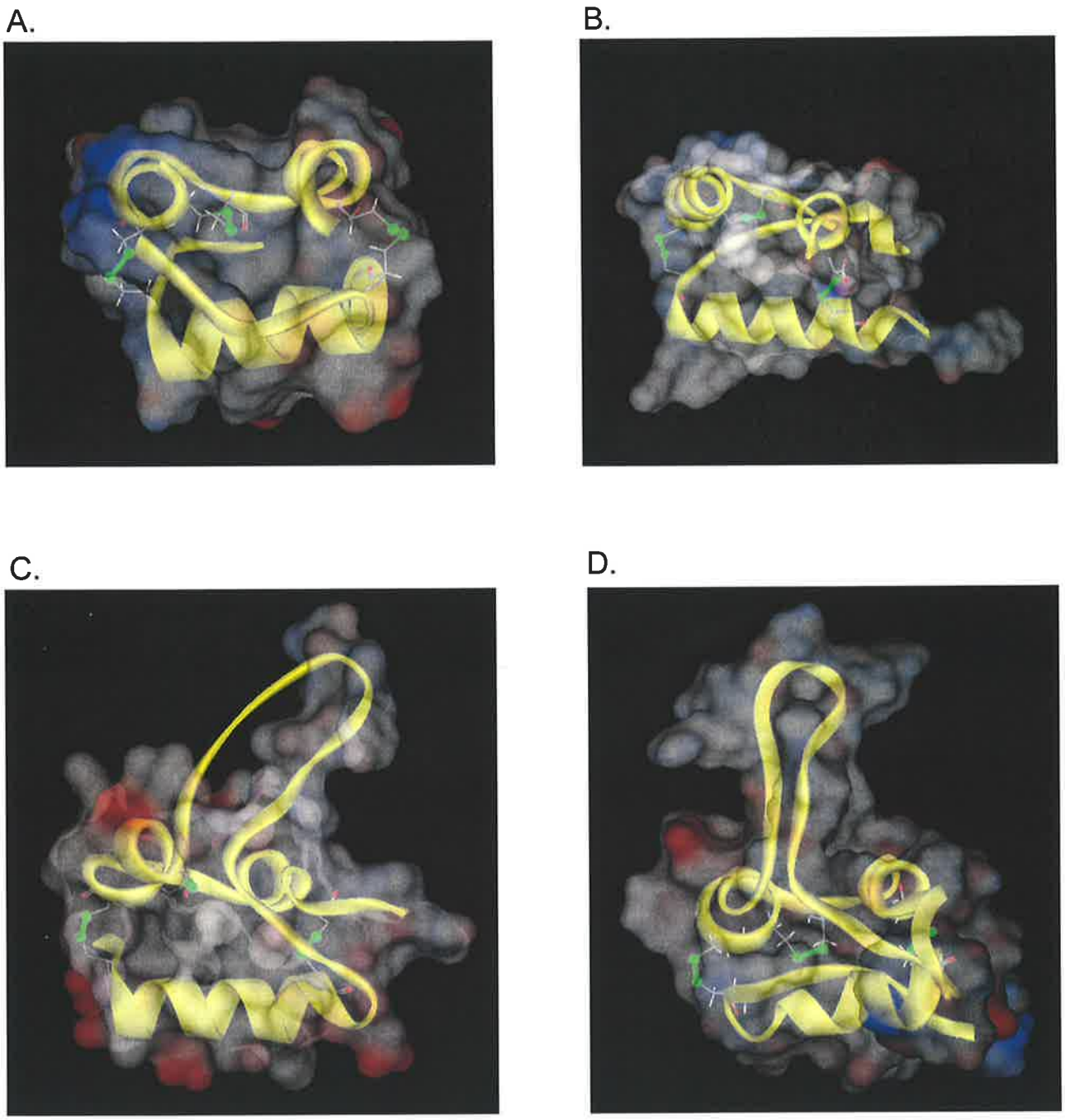


Figure 1.5 Three Dimensional Structures of human Insulin-like proteins

A. Human Insulin (PDB: 1MHJ), NMR Spectroscopy, Jorgensen *et al.*, 1996
B. Human Relaxin (PDB: 6RLX), X-Ray Diffraction, Eigenbrot *et al.*,1991
C. Human IGF-1 (PDB: 1IMX), X-Ray Diffraction, Vajdos *et al.*, 2001
D. Human IGF-2 (PDB: 1IGL), NMR Spectroscopy, Torres *et al.*, 1995
Peptide Backbone is represented as a yellow solid ribbon. Disulphide bonds are represented by a green ball and stick. A transparent solvent exposed surface was added using ViewerPro, with charged atoms coloured either red (+ve) or blue (-ve).

Receptor specificity is maintained between insulin, IGF-1 and IGF-2 mostly by the C and D domains, which when added to insulin increase IGF-IR affinity and decrease IR affinity and conversely, when removed from IGF-1 have the opposite effect (Bayne *et al.*, 1988) (Cara *et al.*, 1990).

1.2.7 Insulin-like Growth Factor II Receptor

The Insulin-like Growth Factor II Receptor (IGF-IIR) or cation independent mannose-6-phosphate receptor is quite distinct from the IR family. It is synthesised from a single chain polypeptide into a large monomeric glycoprotein and consists of 15 contiguous Cys-rich repeats, each with a conserved arrangement of the cysteine residues (Lobel *et al.*, 1988). Its well characterised function is to translocate mannose-6-phosphate containing lysosomal enzymes from the trans-golgi network into the endosome (Nielson *et al.*, 1992). Cell surface expressed IGF-IIR, however, can bind extracellular glycoproteins and also IGF-2, which upon binding are then endocytosed into endosomes and the ligand degraded (Oka *et al.*, 1985). Interestingly IGF-2 is only able to bind IGF-IIR from mammalian vertebrate species and not chicken or *xenopus* homologues indicating that any physiological role for the interaction came after the evolutionary divergence of the species (Clairmont *et al.*, 1989). IGF-2 has higher affinity for the IGF-IIR than it does for the IGF-1R and binds specifically to domain 11 (Schmidt *et al.*, 1995). IGF-1 however only has very weak binding affinity for this site (500-1000 fold lower affinity), while insulin is unable to bind at all. An IGF-IIR knockout mouse is embryonic lethal whereas a double IGF-IIR/IGF-2 knockout results in variable survival and an average birth weight similar to that of an IGF-2 knockout mouse (Barlow *et al.*, 1991) (Filson *et al.*, 1993). This suggests that the IGF-IIR may be required to prevent a toxic build-up of IGF-2 and provides further evidence that its role is to regulate circulating IGF-2 concentration. Recent studies have focused on the role of the IGF-IIR in modulating IGF-2 levels, and have found that increased expression of IGF-IIR reduces the rate of proliferation of several cancer types. Signal transduction emanating from the IGF-IIR has not been definitively proven, although binding of IGF-2 does stimulate association and GTPase activity of GTP-binding proteins that may initiate a cascade (Okamoto *et al.*, 1990).

1.2.8 Insulin-like Growth Factor Binding Proteins

In addition to binding the various receptors, IGFs in particular are further regulated *in vivo* by a family of soluble high affinity Insulin-like Growth Factor Binding Proteins (IGFBPs) of which seven members have been identified. Circulating IGF proteins are almost entirely bound to IGFBPs, whose interaction with the IGFs is of a higher affinity than IGFs with the receptors. Individual binding proteins exhibit preferential affinities for IGF-1 or IGF-2, and

have little affinity for insulin. It appears that IGFBPs function to transport IGFs and regulate their metabolic clearance by significantly increasing their half-life (Jones and Clemmons, 1995). The majority of circulating IGF-1 and IGF-2 is bound to IGFBP-3 and an acid-labile subunit (ALS), this complex serving to increase the stability of the growth factors and preventing their action until required. Tissue specific localisation and evidence that IGFBPs may both inhibit and enhance presentation of ligand to the receptor indicates that these proteins are key players in the *in vivo* control of IGF action.

1.3 RECEPTOR SIGNAL TRANSDUCTION: SIGNALLING AND SPECIFICITY

1.3.1 Tyrosine Kinase Activation

Upon ligand binding the receptor, a substantial conformational change occurs resulting in the activation of the intracellular tyrosine kinase domain. This leads to the autophosphorylation of both halves of the receptor, which immediately leads to the phosphorylation of receptor bound substrates. The tyrosine kinase domains of IR and IGF-1R are very highly conserved (84%), however the regulatory regions including the juxtamembrane domain and the C-terminal domain contain substantial differences in the position and number of potential phosphorylation sites and also protein docking sites. Activation of the kinase requires the sequential phosphorylation of three conserved tyrosines within the activation loop of the catalytic domain (hIR Tyr^{1158}, Tyr^{1162}, Tyr^{1163}/ hIGF-1R Tyr^{1131}, Tyr^{1135}, Tyr^{1136}), with functional deletion of these residues rendering the domain inactive (Li *et al.*, 1994).

Crystal structures of both the hIR and hIGF-IR kinase domains in the active and inactive states have been solved (Hubbard *et al.*, 1994) (Hubbard *et al.*, 1997) (Favelyukis *et al.*, 2001) (Pautsch *et al.*, 2001). It was seen that in the inactive state the activation loop behaves as a pseudosubstrate with Tyr^{1162}/Tyr^{1135} lying in the active site preventing access to other substrates while itself remaining unphosphorylated. The interaction is maintained by a hydrogen bond between Tyr^{1162}/Tyr^{1135} and the conserved residue Asp^{1132}/Asp^{1105}. Trans-phosphorylation of the three tyrosines forces the activation loop to leave the site, which is then stabilised by another hydrogen bond between Tyr^{1163}/Tyr^{1136} and Arg^{1155}/Arg^{1128}, allowing other substrates to access the site (See Figure 1.6).

1.3.2 Phosphorylation and Activation of Common Receptor Substrates

The primary role of the kinase domain of the IR and IGF-1R upon ligand induced activation appears to be the phosphorylation of residues in the adjacent juxtamembrane and C-terminal domains followed by the recruitment and activation of various receptor substrates. The

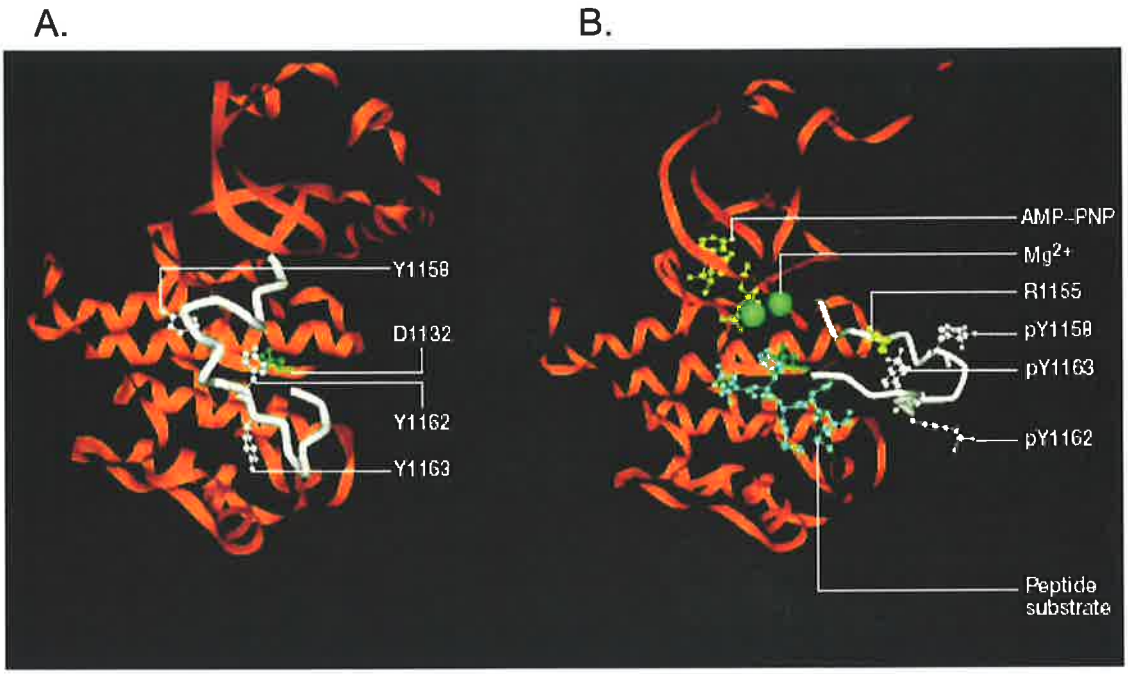


Figure 1.6 Structures of the inactive and activated tyrosine kinase domain of the Human Insulin Receptor

A. Inactivated Tyrosine Kinase Domain (PDB: 1IRK), X-Ray Diffraction, Hubbard *et al.*, 1994
B. Activated Tyrosine Kinase Domain and peptide substrate (PDB:1IR3), X-Ray Diffraction, Hubbard *et al.*,1997
Peptide Backbone is represented as a orange ribbon. The Activation loop is depicted as a white solid line with the critical tyrosines triplet also shown.
Adapted from De Meyts and Whittaker, 2002

phosphorylation of constitutively receptor bound substrates is also likely. Given the sequence similarity, many of the substrates that bind the IR are also bound and activated by the IGF-1R. Many of these substrates are also shared by other receptor tyrosine kinases and cytokine receptors. Numerous mutational studies have identified specific residues which when changed abrogate or reduce the capacity of the receptors to autoregulate, bind and activate specific substrates, as well as initiate normal biological responses (See Figure 1.7 for a summary of insulin signalling pathways).

The primary receptor substrates able to bind to both the IR and IGF-IR are large adapter proteins named Insulin Receptor Substrate (IRS) proteins and Shc (Src homology/collagen) proteins (p46/p52/p66), which trigger two major signalling pathways. The four members of the IRS family, IRS-1 to 4, are modular proteins containing PH (Plekstrin Homology) and PTB (Phospho-Tyrosine Binding) domains. Apart from the critical tyrosine triplet in the activation loop of the kinase domain, $Tyr^{960/950}$ in the juxtamembrane domain, occurring as part of the common PTB domain binding motif NPEY, has been shown to be very important for the binding of the insulin receptor family's primary adaptor proteins, the Insulin Receptor Substrates (IRS). The residue $Tyr^{960/950}$ is common to both hIR and hIGF-IR, and when mutated to phenylalanine appears to greatly reduce the proliferative and transforming potential of the IGF-1R, but does not have a significant effect on anti-apoptotic activity (Miura *et al.*, 1995). It has been shown to be important for the binding of IRS-1, IRS-2 and SHC (Gustafson *et al.*, 1995) (He *et al.*, 1995).

1.3.2.1 Activation of Pathways Controlling Biological Effects

It is well established in the literature that the IR and IGF-IR are both capable of increasing cellular proliferation or augmenting the proliferative ability of other factors.

1.3.2.1.1 Phosphatidylinositol 3-kinase Pathway

A major pathway activated by both the IR and the IGF-IR is the phosphatidylinositol 3-kinase (PI3-Kinase) pathway. The role of PI3-Kinases activated by growth factors is to convert phosphatidylinositol 4, 5-bisphosphate ($PI(4, 5)P_2$) into the active phosphatidylinositol 3, 4, 5 trisphosphate (PIP_3) molecule. PI3-kinases function as heterodimers consisting of a regulatory subunit associated with a catalytic subunit, in which several isoforms of each subunit exist (Shepard *et al.*, 1998). Insulin and IGFs primarily activate PI3-kinases containing the Class 1a p110 kinase subunit isoform. The relative importance of each regulatory subunit isoform (including p85α, p85β, p55α, and p50α) has not yet been resolved. The activation of the

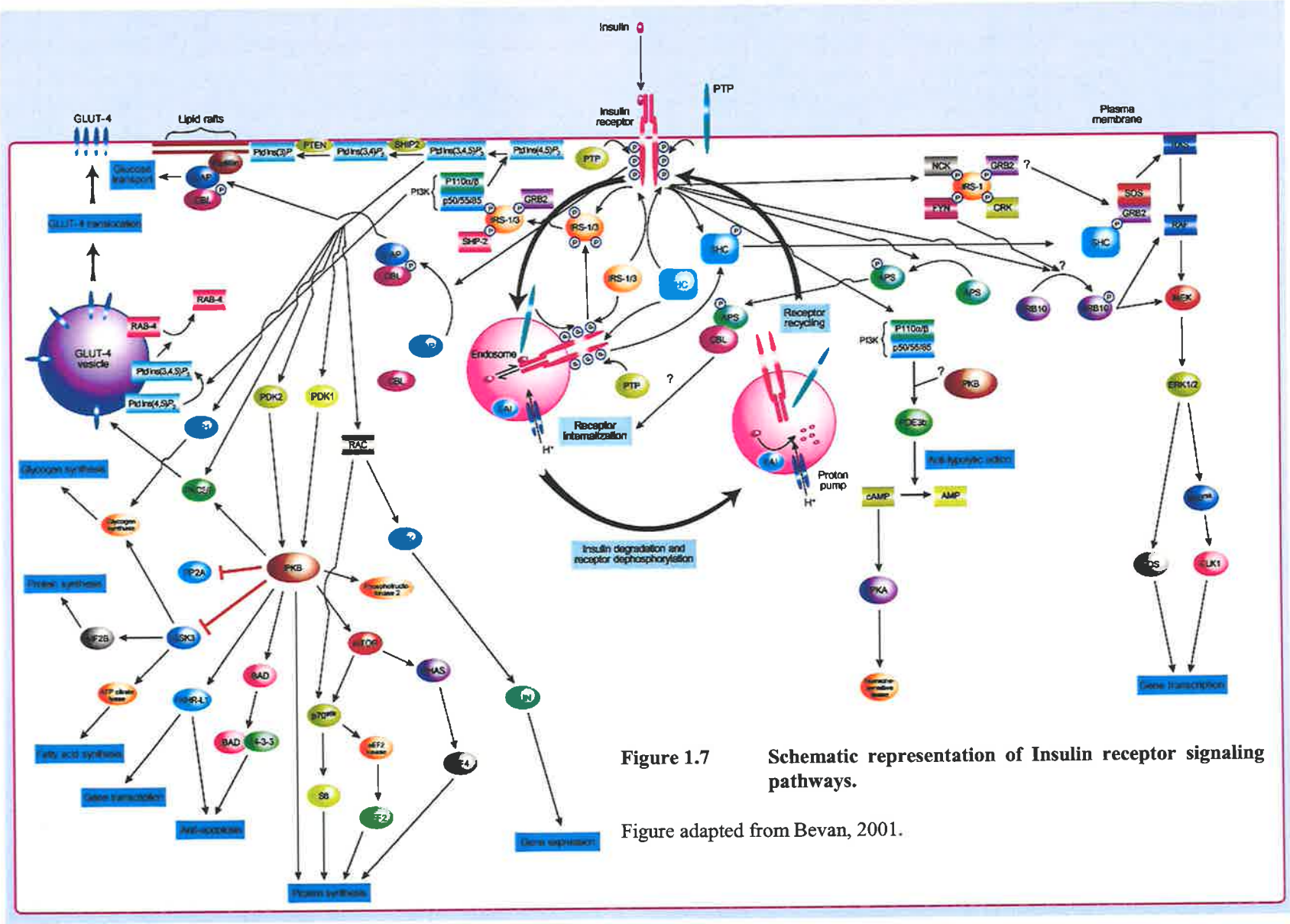


Figure 1.7 **Schematic representation of Insulin receptor signaling pathways.**

Figure adapted from Bevan, 2001.

complex results in a transient, acute increase in PIP_3 levels. IR and IGF-IR activation is linked to the activation of PI3-kinase both by direct interaction with the regulatory subunit of PI3-kinase and probably more importantly through the IRS proteins. Interactions are achieved through the binding of SH2 domains in the regulatory subunit preferentially for tyrosine phosphorylated residues in the YMXM motif, which is also a preferred substrate for the IR and IGF-IR kinase domain (Songyang *et al.*, 1993) (Songyang *et al.*, 1995).

The PI3-kinase pathway has been shown to be involved in initiating nearly every aspect of insulin and IGF action including stimulation of glucose transport, glycogen synthesis, protein synthesis, mitogenesis, protection from apoptosis, and regulation of gene transcription (Shepard *et al.*, 1998). The key step appears to be the production of PIP_3 which itself activates a large variety of molecules. Increased PIP_3 levels allow PDK1 and PDK2 to phosphorylate a major target of insulin signalling Protein Kinase B (PKB) (also known as Akt). PKB is a serine/threonine kinase and its activity is increased by the phosphorylation of residues Thr^{308} and Ser^{473} (PKBα isoform numbering) (Alessi *et al.*, 1996). PKB activation can be blocked by inhibition of PI3-kinase function.

Two of insulin's main functions, glucose uptake and glycogen synthesis, are mediated through PI3-kinase. The key step to upregulation of glucose uptake into the cell through translocation of GLUT4 glucose transporters to the outer membrane, is the activation of PI3-kinase and the production of PIP_3 (Kanai *et al.*, 1993) (Jiang *et al.*, 1998). However, it is as yet unclear as to which proteins provide the definitive link between GLUT4 and PI3K although there is some evidence to suggest the involvement of PKB, PKC, and Rab4 (Standaert *et al.*, 1997) (Mora *et al.*, 1997). The main method for increasing the storage of glucose as glycogen appears to be through inactivation of GSK3β by phosphorylation by PKB, which leads to glycogen synthase activity (reviewed in Lawrence *et al.*, 1997).

Insulin and IGFs signal through PI3-kinase to promote cell growth, which can be attenuated through blocking its activation (Cheatham *et al.*, 1994). Protection against apoptosis is also controlled through PI3-kinase, however the key step is the activation of PKB and its pathways (Kulik *et al.*, 1997). Cell survival in part can be controlled by the direct phosphorylation of BAD by PKB which enables sequestering of BAD by 14-3-3 protein, blocking it from dimerising with Bcl-2 and promoting apoptosis (Datta *et al.*, 1997). Pathways from PKB can also serve to block apoptosis induced cell detachment (Khwaja *et al.*, 1997). The involvement of further proteins is as yet undetermined.

1.3.2.1.2 Mitogen Activated Kinase Pathway

The Mitogen Activated Kinase Pathway (MAP kinase pathway) is activated downstream of many receptor kinases and is involved in the control of metabolism, mitogenesis and differentiation (Su and Karin, 1996). The principle method of activation of this pathway by the Insulin or IGF Receptor is through the adapter protein Shc, which upon binding to tyrosine 950/960 in the juxtamembrane domain is phosphorylated and able to bind another adapter protein, Grb2. Grb2 in turn facilitates the activation of Ras (a GTP-binding protein) via the guanine nucleotide exchange factor Sos. Activated Ras phosphorylates, two intermediary kinases, c-Raf and MEK, which themselves activate the serine-threonine MAP kinases (including ERK-1 and ERK-2). Contribution to activation of the MAP kinase pathway may also be through Grb-2 binding IRS-1 (Valverde *et al.*, 1995). The specific contribution of the MAP kinase pathway tends to vary with cell type, for example for stimulation of proliferation by IGF-1 in MCF7 breast cancer cells it contributes equally to the PI3-kinase cascade, however in foetal brown adipocytes it is essential (Jackson *et al.,* 1998) (Valverde *et al.,* 1995). In NIH3T3 cells the p38 MAP kinase pathway was exclusively found to control upregulation of T cell Death Associated Gene 51(TDAG51), whose presence was critical for IGF-1 mediated prevention of apoptosis (Toyoshima *et al.,* 2004). It is therefore clear that there is a certain amount of redundancy between the MAP kinase and PI3 kinase pathways, however in some contexts and cell types the roles are tightly segregated.

1.3.2.2 Separating the IR and IGF-IR Specific Pathways and Substrates

Site directed mutagenesis has been employed to identify specific regions and residues in the juxtamembrane domain, kinase domain, and CT domain that control substrate interactions and kinase regulation. The CT domain is the least conserved intracellular region in the receptor and contains the most number of obvious differences between the IR and IGF-IR. Despite this, deletion or conversely overexpression of the entire CT domain tethered to the membrane has little overall effect on receptor signaling (Myers *et al.,* 1991) (Paz *et al.,* 2000). Isolated CT domain from the IGF-1R was however able to bind to both the IGF-1R and IR. Strangely, this had the effect of activating tyrosine kinase activity from the IR but not the IGF-1R (Li *et al.,* 1998). An isolated IGF-1R C-terminal domain peptide consisting of amino acids 1282-1298 (unique to IGF-1R), linked to a stearic acid moiety, was able to promote apoptosis, interestingly, by activating the caspase cascade rather than by competition with endogenous IGF-1R (Reiss *et al.,* 1999). This overall effect in itself was surprising since the IGF-1R normally has potent anti-apoptotic effects. It may be that particular sequences are required to recruit the pro-apoptotic proteins enabling them to be inhibited by other proteins activated by

the IGF-IR, and that in an isolated context, these sequences serve only as activators. In this stretch of sequence, phosphorylated Ser^{1283} is believed to be the site of interaction of the IGF-1R with the 14-3-3 isoform ε, and a S1280-1283A mutant has been shown to lack transforming ability (Craparo *et al.,* 1997) (Li *et al.,* 1996). An H1293F/K1294R mutant remained mitogenically capable, but was unable to protect from apoptosis (O'Connor *et al.*, 1997). It is therefore likely that this region of the receptor modulates the more subtle differences between the receptors.

Tyr^{1250} and Tyr^{1251}, whose homologous residues in the insulin receptor are phenylalanine and histidine, are obvious targets to contribute to the biological differences between the two receptors. The Y1250F mutation exhibited a greatly reduced capacity to form colonies in soft agar (Blakesley *et al.*, 1998), and showed a 70% reduction in the rate of internalisation compared to both wild-type and a Y1251F mutant (Miura *et al.,* 1997). This most likely indicates that the proteins that bind this residue contribute to receptor internalisation and also that the signalling pathways initially activated upon the receptor being internalised are important for cellular transformation. NIH 3T3 fibroblast cells expressing the Y1251F mutant showed actin cytoskeleton disruption (Blakesley *et al.,* 1998), and like Y1250F receptors, a greatly reduced ability to make colonies in soft agar (Blakesley *et al.,* 1998). Cells also appeared to be mitogenically competent but with little anti-apoptotic activity (O'Connor *et al.,* 1997). Altogether these data demonstrate that phosphorylation of Shc, and assembly of the Shc complex necessary for activation of Ras and the MAPK pathways are deficient in cells expressing the Y1250/1251F mutant IGF-IR. This would explain the loss of IGF-I-mediated survival in FL5.12 cells expressing this mutant and may also explain why this mutant IGF-IR is deficient in functions associated with cellular transformation and cell migration in fibroblasts and epithelial tumour cells (Leahy *et al.,* 2004).

Recent two-hybrid screens for IGF-1R interacting proteins identified a new protein RACK1, a homologue of the β-subunit of heterotrimeric G proteins. It specifically interacts with IGF-1R and not the IR, and this interaction is independent of receptor activation, but did require Ser 1248 in the C-terminus. Its role in signalling is unclear as there are differing reports from overexpression studies are performed in NIH or MCF7 cells, however under diffferent conditions it may be involved in the modulation of proliferation, anti-apoptosis and cell spreading (Kiely *et al.*, 2002) (Hermanto *et al.*, 2002).

It is apparent that the types of downstream pathways capable of being activated by the IR and the IGF-IR are virtually indistinguishable in *in vitro* experiments. It is therefore in the precise

details of concentrations and binding affinities of substrates and the levels of activation that allow the systems to have different biological functions.

1.4 ROLE OF INSULIN-LIKE PROTEINS IN DISEASE

1.4.1 DIABETES

Diabetes is a multifactorial disorder that generally describes a condition where carbohydrate and lipid metabolism is no longer properly controlled by the Insulin system. Diabetes is split into two categories, Type-1/Insulin-Dependent Diabetes Mellitus (IDDM), and Type-2/Non-Insulin Dependent Diabetes Mellitis (NIDDM), which have common complications but a completely different pathophysiology.

1.4.1.1 Normal Regulation of Glucose Metabolism

The control of glucose metabolism is essential for the survival of all animals as it provides a ready source of energy in the form of ATP. Hypoglycemia causes cell death, and hyperglycemia can damage whole organs so the body maintains a plasma glucose level of around 5 mM. This homeostasis is regulated primarily by skeletal muscle controlling glucose disposal after eating, and by the liver when fasting, which produces glucose through gluconeogenesis. The balance is controlled by the two opposing hormones insulin and glucagon. Insulin is rapidly secreted into the bloodstream after increases in plasma glucose are sensed by the β cells of the islets of Langerhans in the pancreas. Insulin then acts to increase glucose uptake in skeletal muscle and in adipose tissue through increasing the level of cell surface GLUT 4 glucose transporter. The storage of excess glucose as glycogen in muscle is stimulated by activation of glycogen synthase and in adipose tissue as lipid by activation of lipid synthesizing enzymes, while at the same time inhibiting glycogen and lipid breakdown. Insulin also inhibits gluconeogenesis through phosphorylation and de-phosphorylation of metabolic enzymes and by reducing the expression of the rate-limiting enzyme in gluconeogenesis, phosphoenylpyruvate carboxylase (PEPCK). As plasma glucose levels fall glucagon is secreted by the neighbouring α cells, stimulating the release of stored (glycogen and lipid breakdown) and synthesized (liver) glucose, reversing insulin's effects.

1.4.1.2 Molecular Basis of Insulin Dependent Diabetes Mellitis

IDDM is normally early-onset, manifesting itself sometime during childhood through to puberty and is characterized by loss of insulin production caused by autoimmune destruction

of the pancreatic β cells of the islets of Langerhans. This results in the inability of the body's muscle and fat cells to take up sugar efficiently and control fat storage. This disease can be controlled by the regular self-administration of insulin before each main meal, which allows patients to have a normal lifespan.

1.4.1.3 Molecular Basis of Non-Insulin Dependent Diabetes Mellitis

NIDDM is by far the most common form of diabetes and generally occurs in adults, resulting from defects in both insulin secretion and utilization. The key step in the development of NIDDM is the state of insulin resistance which is defined as a state in which a greater than normal amount of insulin is required to elicit a quantitatively normal response. Insulin resistance is commonly associated with obesity, aging, and a sedentary lifestyle, but there also appears to be some genetic predisposition. NIDDM overall is therefore thought to be mostly a disease caused by an inappropriate lifestyle. Skeletal muscle insulin resistance reduces glucose uptake, resulting in β cell compensation by increasing insulin levels, either through increased insulin production or increased β cell mass. Insulin resistance in adipose tissue increases lipolysis leading to fatty acid release, which has the additional effect of inhibiting insulin from suppressing gluconeogenesis in the liver. The ever-increasing concentration of plasma glucose causes the β cells to further increase insulin production leading to hyperinsulinemia and further exacerbating cellular insulin resistance. At some point the β cells can no longer compensate, and start producing ever-decreasing amounts of insulin resulting in an inability of the body to respond appropriately to glucose (Kahn, 2003).

At a cellular and molecular level the precise events that lead to the initial insulin resistance and subsequent loss of glucose homeostasis can be varied and are not totally understood. Naturally occurring mutations of the hIR affecting ligand binding or receptor activation are rare and usually result in extreme insulin resistance and leprachaunism, but at this level are not believed to contribute to the pathogenesis of NIDDM (Virkamaki *et al.,* 1999).

Insulin receptor inactivation in mice, using cre-lox recombination driven by a tissue specific promoter has given valuable insight into the relevance of receptor signaling in specific tissues. Muscle specific insulin receptor knockout (MIRKO) mice had increased fat stores and hypertriglyceridemia, but did not develop hyperinsulinemia and diabetes, due to the excess glucose being utilised by adipose tissue instead of muscle (Bruning *et al.,* 1998) (Kim *et al.,* 2000). In contrast ablation of the glucose transporter GLUT4 in muscle can lead to a diabetic state, indicating that there are compensatory mechanisms that bypass the insulin receptor to

activate GLUT4 (Zisman *et al.,* 2000). This probably includes the IGF-IR, as overexpression of a dominant negative IGF-IR, which inactivated the IR through hybrid formation, produced mice that also develop diabetes (Fernandez *et al.,* 2001). When insulin receptor is removed from adipose tissue, mice are resistant to obesity and do not develop insulin resistance, and also exhibit lower fasting insulin levels a vastly reduced ability for adipocytes to uptake glucose (Bluher, *et al.*, 2002). Liver-specific ablation of the insulin receptor had a profound effect with mice showing insulin resistance, glucose intolerance, hyperinsulinemia, and the expected lack of insulin induced suppression of hepatic glucose production (Michael *et al.,* 2000). This indicates that the liver plays an important role in glucose homeostasis and insulin clearance. Lack of the insulin receptor or IGF-IR in the pancreatic β cells impairs insulin secretion, leading to glucose intolerance and in some cases diabetes (Kulkarni *et al.,* 1999) (Kulkarni *et al.,* 2002). Double mutant mice of either IR/IRS-1 or IR/IRS-2 both develop diabetes at similar rate with the IR/IRS-1 mutants being insulin resistant mainly in the muscle and the IR/IRS-2 in the liver (Kido *et al.,* 2000). This highlights the polygenic nature of the disorder and the way in which the same disease can occur through slightly different initial mechanisms. Interestingly knocking out the p85 alpha regulatory isoform of PI3K sensitizes the mice to insulin and improves glucose tolerance through upregulation of the other two regulatory isoforms (Fruman *et al.,* 2000).

Higher frequencies of mutations in IRS-1 have been found in NIDDM populations over normal controls. The most well studied mutation in IRS-1 (G972R) reduces the binding affinity of IRS-1 for the p85 subunit of PI3K and results in an overall 35-40% decrease in insulin's activity (Almind *et al.,* 1996). Mice bearing a disrupted IRS-1 gene develop insulin resistance but not the diabetic state, whereas IRS-2 KO mice do develop diabetes associated with insulin resistance and impaired insulin secretion (Araki *et al.,* 1994) (Tamemoto *et al.,* 1994) (Withers *et al.,* 1998). Additionally, since IRS-2 appears to be unnecessary for skeletal muscle glucose uptake, this suggests that the key to the progression to diabetes is in the β cell function.

1.4.1.3.1 Is there a Role for Altered Expression of the IR Isoforms in NIDDM?

It was proposed that the alternative splicing and therefore the relative expression of the two hIR isoforms may be affected during the pathogenesis of type-2 diabetes, however there have been conflicting results between studies performed on rats, monkeys and humans. In human patients with diagnosed NIDDM or insulin resistance, there was found to be increased expression of hIR-B in muscle to a level similar to or greater than hIR-A (Sesti *et al.*, 1991) (Kellerer *et al.*, 1993) (Haring *et al.*, 1994) (Sbraccia *et al.*, 1996). In a diabetic rhesus

monkey model the opposite trend was seen with increases in hIR-A expression in muscle and liver (Huang *et al.*, 1994) (Huang *et al.*, 1996). Finally in a diabetic rat model no change was detected in the expression of the isoforms in muscle (Sbraccia *et al.*, 1998). The intrinsic effect of changing the isoform expression patterns has not been tested and it remains to be seen how this would effect *in vivo* insulin signaling. In light of recent evidence highlighting the importance of tissue specific expression of the insulin receptor and the preferential binding of IGF-2 to the IR-A isoform it would be prudent to re-evaluate the expression of the isoforms in the liver, adipose tissue and the pancreatic β cells for any changes throughout the development of NIDDM.

1.4.2 CANCER

The progression of a normal cell to a cancerous cell is a complicated process involving the stepwise degradation of controls regulating cellular proliferation, DNA integrity, matrix attachment and ability to undergo apoptosis. The details are beyond the scope of this review so instead I will focus on the involvement of the insulin/IGF system in the development and maintenance of human cancers.

Evidence to suggest that the IGF system plays an important role in cancer came from numerous knockout mouse experiments. Single gene knockouts of any of IGF-1, IGF-2, IGF-IR, or IR and in combination have shown that the IGFs in particular play an important role in cell growth and differentiation throughout development (Reviewed in Butler and LeRoith, 2001). IGF-2 is particularly important in mice and the regulation of its availability during development by the IGF-2R is critical for survival. Recently it has been found that increased expression of IGF-2R in cancer cells producing IGF-2 is able to inhibit their growth and correspondingly a loss of expression of IGF-2R promotes growth and increases intracellular signalling from the insulin receptor and IGF-1R (O'Gorman et al., 2002) (Ospio et al., 2001).

1.4.2.1 Evidence of a Role for the IGF-IR in Transformation and Maintenance of Tumours

Studies utilising the IGF-IR KO mouse derived fibroblasts (R-) found that these cells, unlike most other fibroblasts, were unable to be transformed by a number of common cancer causing agents, including the SV40 T antigen, Human papilloma virus E7, Ras, c-src, overexpression of Epidermal Growth Factor Receptor (EGF-R), PDGF-R, IR, and IRS-1.

Receptor number plays an important role in the ability of the IGF system to allow transformation. Cells with $3x10^3$ receptors/cell did not respond to IGF-1. With 1.5 X 10^4

receptors/cell IGF-1 induced DNA synthesis but not cell division, while at 2.2 X 10^4 receptors/cell full proliferation was achieved without the ability to grow colonies in soft agar (transformation). Only at receptor levels of $3x10^4$ receptors/cell were the cells fully transforming (Rubini *et al.*, 1997) (Reiss *et al.*, 1998). This suggests that there are critical thresholds for the activation of particular downstream pathways and that those needed for transformation have the highest requirement for a strong signal from the IGF-IR.

While the relevance of IGFs preventing apoptosis during development is unclear, it has been found in a large number of studies that the most important role of the IGF system in cancer progression and maintenance is that of an anti-apoptotic agent. IGFs help to protect the cancer cell from immune system targeted destruction and also from human mediated chemo and radiation therapy. Activation of the IGF-1R has been found *in vitro* to protect against hypoxia, growth factor withdrawal, chemotherapeautic agents, UV-B irradiation, and TNF-α (Reviewed in Adams, *et al.*, 2000). As discussed in 1.3.2.1.1 the major route through which this occurs is through the activation of PI3-kinase and in many cases PKB. Recent studies have also focused on the role of IGF signalling in metastasis and membrane remodelling where alteration of the expression or inhibiting the function of integrins reduces the ability of IGFs to promote cell motility (Zhang *et al.*, 2004).

1.4.2.2 Breast Cancer

As an example of a specific case, many forms of breast cancer have been found to have an altered IGF system. IGF-1 and IGF-2 are both potent mitogens for breast cancer, and act synergistically with estrogen to stimulate cell growth (Sachdev and Yee, 2001). Both oestradiol and IGF-1 are required to promote mammary gland growth (Ruan *et al.*, 1999). Overexpression of IGF-1 in the mammary gland causes hyperplasia and protects these cells from post-lactational apoptosis and, combined with the loss of the p53 tumour suppressor, accelerates breast tumourigenesis (Neuenschwander *et al.*, 1996) (Hadsell *et al.*, 2000). A positive correlation was found between circulating concentrations of IGF-1 and the risk of breast cancer in pre-menopausal but not post-menopausal women even after IGFBP-3 levels were taken into account (Hankinson *et al.*, 1998). It appears that in breast cancer the IGF system has several roles including the promotion of proliferation, anti-apoptosis, and even migration and invasion.

1.4.2.2.1 Is there a Role for Altered Expression of IR Isoforms in Cancer?

Recently, several studies have found that the expression of the hIR-A isoform has increased in cancerous cells, including breast, thyroid, ovarian, and muscle, relative to the untransformed

cells (Sciacca *et al.*, 2002) (Kalli *et al.*, 2002) (Vella *et al.*, 2002) (Sciacca *et al.*, 1999). It has been speculated that in these cases the cells are reverting to a more 'fetal' like nature, where IGF-2 signalling through the IR-A isoform is more relevant to cell growth. This occurrence together with an increase in IGF-2 concentration was found to contribute to the increased proliferative potential of the cells and resistance to apoptotic signals. Therefore in this instance the IR-A isoform is functioning more like an IGF-IR. The question of whether this is a unique result of IGF-2 binding or a normal property of the insulin receptor if overexpressed in the appropriate context remains to be determined.

1.4.3 TARGETING THE INSULIN/IGF SYSTEM

Researchers are faced with an interesting problem when considering how to develop strategies to treat diseases involving insulin and the IGFs. Treatments are required to enhance insulin signalling in diabetes, and conversely they are required to inhibit IGF signalling in cancer. Given the cross-reactive nature of the ligands and the overlapping signalling pathways emanating from the IR and IGF-IR it is easy to visualise situations where treatment of one disease would increase the risk and development of the other.

A variety of strategies have been successfully used both *in vitro* and *in vivo*, to inhibit cancer progression or growth by inhibiting the action or expression of the IGF-IR. Anti-sense RNA to IGF-IR has been successfully used to either inhibit the growth or increase sensitivity to apoptotic drug treatment, of Ewings sarcoma, breast, cervical, and brain tumour cells (Chernicky *et al.*, 2002) (Chernicky *et al.*, 2000) (Scotlandi *et al.*, 2002) (Nakamura *et al.*, 2000) (Resnicoff *et al.*, 1996). Neutralising antibodies that block IGF binding to the IGF-IR have also been shown to inhibit the progression of many different cancer cell lines (Maloney *et al.*, 2003) (Hailey *et al.*, 2002). It is now apparent however, that even if a method of specifically blocking the interaction of IGFs with the IGF-IR were achieved, that IGF-2 in particular would still be able to act through the IR-A isoform if present. IGF-IR kinase substrate inhibitors and kinase ATP mimetics, although effective, also blocked the highly similar IR kinase, which would be clinically undesirable (Blum *et al.*, 2000) (Parrizas *et al.*, 1997). Recently, more selective inihibitors have been developed by pharmaceutical companies including Roche and Merck.

The treatment of Type-2 diabetes is an ever-growing area of research that generally involves indirectly modulating the levels and sensitivity of insulin. Metformin, sulphonylureas, and thiazolidinediones are the most commonly used drugs in the treatment of type-2 diabetes.

Metformin appears to function through inhibiting gluconeogenesis in the liver thereby reducing hyperglycaemia and allowing eventual resensitization to insulin (Hundal and Inzucchi, 2003). Sulphonylurea compounds stimulate pancreatic beta cell production of insulin and the relatively new thiazolidinedione compounds bind to the nuclear peroxisome proliferator-activated receptor-gamma proteins (PPARgamma), which activates transcription of genes controlling adipogenesis, and glucose and lipid metabolism (Diamant and Heine, 2003). Overall these treatments tend to affect downstream targets of insulin that eventually leads to insulin sensitisation rather than directly targeting the cause.

1.5 SUMMARY AND PROJECT AIMS

The insulin/IGF system traverses control of both growth and metabolism through partially common signalling networks in addition to the cross-reactivity of the ligands. Signalling specificity is maintained through a combination of tissue specific expression of receptors and substrates, alternate modes of ligand binding utilising overlapping but distinct sites on the receptors, and by the presence of different phosphorylation and substrate binding sites within the regulatory regions adjacent to the intracellular tyrosine kinase domain. The substantial functional differences between the IGF and insulin systems, but use of common factors, requires a more detailed understanding of the way the systems overlap so that effective strategies to control them can be formulated.

Of all the insulin/IGF receptor domains, the insert domain is the most mysterious, having no homology to any structurally known domain. Other than its involvement as part of the ligand binding site and as the site for αβ cleavage, it is still unclear how this domain(s) is oriented in relation to the other domains. Therefore an initial aim of this project was to ascertain the 3D structure of the insert domain using NMR spectroscopy.

The insulin receptor exists as two isoforms that differ only by 12 amino acids situated at the end of the α-subunit in the Insert domain. Initial studies have suggested that the absence of the 12 amino acids (IR-A isoform) increases the affinity of the receptor for IGF-2, and to a lesser extent IGF-1, while retaining high affinity for insulin. They also found that IGF-2 induced greater proliferative effects from the IR-A receptor. A further aim of the project was to confirm these observations, and compare and analyse the binding, phosphorylation and proliferative potential of each ligand for each of the three receptors, IR-A, IR-B and IGF-IR.

In order to identify specific differences in signalling potential between the insulin receptor and IGF-IR intracellular regions, mutants of the IGF-IR in potential phosphorylation or binding sites that do not occur in the insulin receptor, were analysed for their effect on overall receptor phosphorylation and proliferation.

The *C. elegans* genome has been found to contain 38 separate insulin-like peptides but only one insulin-like receptor. The possibility exists that some of these peptides may be able to modulate activation of human insulin-like receptors both positively and negatively, as at least some of these peptides are inhibitory in *C. elegans*. An understanding of their structure may therefore lead to the development of clinically useful proteins to use in the treatment of both

diabetes and cancer. Each insulin-like peptide was homology modelled using a variety of different template structures to determine how structurally similar they were likely to be with each other and the templates. Attempts to clone several of the different types of insulin like peptides were made, for the purpose of expression and eventual NMR structural analysis, to examine in reality how closely related they were to insulin and the IGFs and to verify the accuracy of the homology modelling.

OVERALL AIMS:

A. Determine the structure of the insert domain by NMR spectroscopy.

B. Examine the biological differences between the hIR-A and hIR-B isoforms.

C. Investigate the role of unique residues in the C-terminal domain of the IGF-IR.

D. Use comparative modelling to predict the structures of the *C. elegans* ins peptides.

CHAPTER 2

Materials And Methods

CHAPTER 2 MATERIALS AND METHODS

2.1 ABBREVIATIONS

AA	amino acid
ABTS	2, 2' azinobis 3-ethylbenzthiazoline-6 sulfonic acid
Amp	ampicillin
APS	ammonium persulphate
ATCC	American Tissue Culture Collection
ATP	adenosine triphosphate
BCA	bicinchoninic acid
BME	β-mercaptoethanol
bp	base pair
BSA	bovine serum albumin
DEPC	diethyl pyrocarbonate
DMEM	Dulbecco's modified eagles medium
DMSO	dimethyl sulfoxide
DNA	deoxyribonucleic acid
dNTP	deoxynucleotide triphosphate
DTT	dithiothreitol
ECL	enhanced chemiluminescence
EDTA	ethylene diamine tetra-acetic acid
ELISA	enzyme linked immunosorbent assay
EtBr	ethidium bromide
FCS	fetal calf serum
GLB	gel load buffer
HEPES	N-2-hydroxyethylpiperazine-N'-2-ethanesulfonic acid
HRP	horseradish peroxidase
IGF	Insulin-Like Growth Factor
IGF-1R	Type-1 Insulin-Like Growth Factor Receptor
IP	immunoprecipitation
IPTG	isopropyl-β-Δ-thiogalactopyranoside
IR	Insulin Receptor
kb	kilobase pair
LA100	LB plus Amp (100 μg /mL)
LF2000	Lipofectamine 2000™
MQ	Milli-Q
$OD_{x\ nm}$	optical density at x nm wavelength
PBS	phosphate buffered saline
PBST	phosphate buffered saline and tween-20
PMSF	phenylmethylsulfonyl fluoride
RNA	ribonucleic acid
RNase	ribonuclease
Rpm	revolutions per minute
SD	super duper buffer
SDS	sodium dodecyl sulphate
SDS-PAGE	sodium dodecyl sulphate polyacrylamide gel electrophoresis
TBS	tris buffered saline
TBST	tris buffered saline and tween-20
TK	thymidine kinase
Tm	melting temperature at which half maximal binding occurs
tris	tris(hydroxymethyl) aminomethane

U	units (active)
UV	ultra violet
V	volt

2.2 Materials

The materials that were used and their main supplier/manufacturer are listed.

2.2.1 General Materials

3MM chromatography paper	Whatman Ltd
Centricon concentrators, 30 and 100 kDa MWCO	Millipore
Filter, 100 kDa MWCO	Amicon
Fluro Nunc 96-well	Nunc
Hybond™-C (nitrocellulose membrane)	Amersham
Nunclon Delta white 96 microwell	Nunc
LumiTrac™ 600	Greiner
Minisart 0.2 or 0.45 µM filters, syringe top	Sartorius
Needles, various gauges	Terumo
PD-10 Desalting Column	Amersham
Photographic slide film	Kodak
Sealed plastic container	Supermarket
Syringes, various volumes	Becton Dickinson
Tissue Culture Vessels, various	Falcon
X-ray film	Konica, AGFA
500ml bottletop filter 0.45 µM	Corning

2.2.2 Chemicals and Reagents

All chemicals and reagents were of analytical grade, or of the highest purity available. The majority of common laboratory chemicals were obtained from either Sigma or BDH Chemicals Ltd. Specialized chemicals are listed below.

ABTS	Boehringer Mannheim
acrylamide (Acryl/bis 29:1)	Astral Scientific
Acti-rapid™ Insulin	Novo Nordisk
Amp	Sigma
APS	BDH Chemicals Ltd.
ATP	Sigma
bacto-agar	Difco Labs Ltd.
Bacto-tryptone	Difco Labs Ltd.
Benchmark Size markers	Invitrogen
BCA assay reagent	Pierce
β-mercapto-ethanol	Sigma
Bradford reagent	Biorad
bromophenol blue	Sigma
BSA	Sigma
BSA (TC quality) (#A8806)	Sigma
Coomassie stain	Sigma
Colour Markers (protein) (#C3437)	Sigma

DELFIA Enhancement Solution	Perkin Elmer Life Sciences
DEPC	Sigma
DMEM (#12430-054)	Gibco/BRL
DMSO	Sigma
dNTP	Sigma
DTT	Diagnostic Chemicals
EDTA	Sigma
Ethidium bromide	Sigma
FCS	Gibco/BRL
Gentamycin (G418)	Gibco/BRL
Guanidium	Sigma
Horse Serum	CSL
Immidazole	Sigma
Insulin (media grade)	Gropep Pty Ltd
IGF1 (media grade #IU100)	Gropep Pty Ltd
IGF1 (receptor grade #CM001)	Gropep Pty Ltd
IGF2 (media grade #OU100)	Gropep Pty Ltd
IGF2 (receptor grade #OM001)	Gropep Pty Ltd
IgG (G-4386)	Sigma
IPTG	Sigma
Lipofectamine™ Plus	Gibco/BRL
Lipofectamine™ 2000	Gibco/BRL
Luminol	Sigma
Methylene Blue	Sigma
$NaVO_3$	Sigma
NaF	Sigma
nickel-IDA agarose	Zymatrobe
Nonidet P-40	BDH Chemicals Ltd.
Oxaloacetic acid	Sigma
Paraformaldehyde	Sigma
Pfu buffer	Stratagene
PMSF	Sigma
Protease inhibitor cocktail (#P8340)	Sigma
PUC19 DNA markers	Geneworks
Puromycin	Sigma
RNAlater	Ambion
RNaseZAP	Ambion
Sepharose-A (#P3391)	Sigma
SDS	Sigma
skim milk powder	Diploma
sodium azide	Sigma
SPP1 DNA markers	Geneworks
temed	BIORAD
Trypan Blue	Sigma
Trypsin/EDTA	Gibco/BRL
triton X-100	Sigma
TRIzol™	Gibco/BRL
tween-20 (polyoxyethylene-sorbitan monolaurate)	Sigma
yeast extract	Difco Labs Ltd.

2.2.3 Enzymes

Enzymes were obtained from the following listed suppliers:

DNAase, RNAase free	Boehringer Mannheim
Lysozyme	Sigma
RNaseA	Sigma
RT-mMLV (#M6125H)	Epicenter
Turbo Pfu DNA polymerase	Stratagene

2.2.4 Antibodies

All antibodies were diluted according to the manufacturer's recommendation. The following antibodies were obtained from these sources:

a) Primary Antibodies:

anti-phosphotyrosine monoclonal (4G10)	1:2000	Upstate Biotech
europium anti-phosphotyrosine monoclonal (PY20)		Santa Cruz
europium anti-phosphotyrosine monoclonal (PT100)		Wallac (Perkin Elmer Life Sciences)
anti-insulin monoclonal (CT-1)	1:1000	(gift from Ken Siddle, Cambridge)
anti-insulin monoclonal (83-7)	N/A	(gift from Ken Siddle, Cambridge)
anti-insulin monoclonal (83-14)	N/A	(gift from Ken Siddle, Cambridge)
anti-IGF-IR monoclonal (24-31)	N/A	(gift from Ken Siddle, Cambridge)
anti-IGF-IR monoclonal (16-13)	N/A	(gift from Ken Siddle, Cambridge)
anti-IGF-IR monoclonal (26-3)	N/A	(gift from Ken Siddle, Cambridge)
anti-IGF-IR monoclonal (24-55)	N/A	(gift from Ken Siddle, Cambridge)

b) Secondary Antibodies:

anti-mouse-HRP conjugate, sheep polyclonal	1:2500	Silenus (AMRAD)
anti-rabbit-HRP conjugate, goat polyclonal	1:2000	DAKO
anti-mouse HRP conjugate, goat polyclonal	N/A	Silenus (AMRAD)

2.2.5 Bacterial Strains

DH5α: The *E. coli* DH5α strain was used in transformations and was a host for all recombinant plasmids. {supE44, DlacU169 (phi80, lacZ, DM15), hsdR17, recA1, endA1, gyrrA96, thi1, relA1}.

BL21(DE3): The *E. coli* BL21(DE3) strain was used in transformations and was a host for recombinant protein expression plasmids. {hsdS, gal (λcIts857, ind1, Sam7, nin5, lacUV5-T7gene1)}

ADA294: The *E. coli* ADA294 strain was used in transformations and was a host for recombinant protein expression plasmids.

2.2.6 Tissue Culture Cell Lines

The following cell lines have been used throughout the course of this study:

R- : embryionic fibroblast from IGF-IR KO mouse (Renato Baserga, Kimmel Cancer Center, Philadelphia)
P6 : IGF-IR overexpressing fibroblasts (Colin Ward, CSIRO, Melbourne)
NIH 3T3 hIR : hIR-B overexpressing fibroblasts (Colin Ward, CSIRO, Melbourne)
Lec8 hIRex11-EDZIP : carbohydrate conjugating deficient fibroblasts secreting hIR-A ED ZIP (Colin Ward, CSIRO, Melbourne)
CT5 : hIR-B ED secreting fibrolblasts (Colin Ward, CSIRO Melbourne)

2.2.7 Bacterial Cloning and Protein Expression Vectors

pET32a+ :(Novagen) cloning and bacterial expression vector was used to generate fusion proteins with a thioredoxin (Trx) (109aa) and 6-His tag.
pBluescriptII-KS+ :(Stratagene) general purpose cloning vector
pEE14 :(Celltech) Mammalian expression vector
pGEX4T-2 :(Amrad) bacterial expression vector
pSP72 :general purpose cloning vector
pGEM-Teasy :(Invitrogen) general purpose cloning vector
pEZZ18 :(CSIRO, Melbourne) bacterial expression vector
pET32a :(Invitrogen) bacterial expression vector

2.2.8 Mammalian Cell Culture Expression and Reporter Vectors

Stable mammalian cell lines were selected for using the antibiotic resistance plasmids to neomycin and puromycin using the pEF-IRES vector as a backbone (Invitrogen).

2.2.9 PCR Primers

The following primers were used in PCR. Their nucleotide sequence is as shown:

Oligonucleotides

DNA primers were synthesised by Geneworks

PCR primers:
Y950F lower 5′ AGCAGCGCTGAAGACTCCGGTTTTA 3′
5'IGF-IR/ID 5′ TAC GGA TCC AAA GAC AAA ATC CCC 3′
3'IGF-IR/ID 5′ AGT GAA TTC TTA GTT ATC CAC TCT GCT CTC 3′
5'hIR/mFc 5′ CGCGGATCCAATATTGCAAAAGGTTGGTAAGCCTTGCATATG 3′
5'IGF-IR/FN3-ID 5′ CGCGGATCCGTTCCTTCCATTCCC 3′

3'IGF-IR/FN3-ID	5' GCGGAATTCTTATGCGGGCATAGTCC 3'
5'Irex11	5' CTGGGAGAGGCAGGCGGAAGACAGTGAGCTGTTCG 3'
3'IRex11	5' CCTGGTTGCAAGCCTGCAGCTCGATGCGATAGCCC 3'
IRseq950	5' CACCACACCCTTGGC 3'
mFc-delhinge-upper	5' CCGTCAAATATTGCAAAAGTCCCAGAAGTATCATC 3'
mFc-delhinge-lower	5' GATGATACTTCTGGGACTTTTGCAATATTTGACGG 3'
5'Ins6Amp	5' ATGAACTCTGTCTTTACTATC 3'
3'Ins6Amp	5' TCATGGACAACAAGCAGATC 3'
3'Ins6RT	5' GACTGATATCGGAGTAAAGG 3'
5'Ins11Amp	5' ATGTCTAGTTACCGTCAAAC 3'
3'Ins11Amp	5' TTAAATGTCCCGAGTTGAC 3'
5'Ins17Amp	5' ATGTTCTCGACCAGAGGGG 3'
3'Ins17Amp	5' TTATCCAAAAGGATTGCAGATTG 3'
5'Ins23Amp	5' ATGTTCGTTCTTCTTATTATTC 3'
3'Ins23Amp	5' TCACTGAGAAGTTGGACAG 3'
5'Ins30Amp	5' ATGAGTTCTCACGCCCTG 3'
3'Ins30Amp	5' TCAGAAGACGAATGGTGTC 3'
5'Ins37Amp	5' ATGGCTGCTTTCCTGCCAATTG 3'
3'Ins37Amp	5' TTACTTTATTTTTACACAATTAATACAG 3'
5'Ins-6ExpZ	5' CGGGGATCCGCTGGTTCCGCGTGTTCCAGCACCAGG 3'
3'Ins-6Exp 3'	5' CGGGTCGACTCATGGACAACAAGCAGAGCGTATGTAGTC
5'Ins-6ExpE	5' CGGGGATCCCTGGTTCCGCGTGTTCCAGCACCAGG 3'
WT950	5' GTAAAACCGGAGT**A**CTTCAGCGCTGCT 3'
Y950F upper	5' GTAAAACCGGAGT**T**CTTCAGCGCTGCT 3'
WT Y1250	5' GAGGTCTCCTTCT**A**CTACAGCGAGGAG 3'
Y1250F upper	5' GAGGTCTCCTTCT**T**CTACAGCGAGGAG 3'
Y1250F lower	5' CTCCTCGCTGCAGAAGAAGGAGACCTC 3'
WTY1251	5' GTCTCCTTCTACT**A**CAGCGAGGAGAAC 3'
Y1251F upper	5' GTCTCCTTCTACT**T**CAGCGAGGAGAAC 3'
Y1251F lower	5' GTTCTCCTCGCTGAAGTAGAAGGAGAC 3'
WT S1283	5' GCCTCCTCGTCC**A**CACTGCCACTGCCC 3'
S1283A upper	5' GCCTCCTCGTCC**G**CACTGCCACTGCCC 3'
S1283A lower	5' GGGCAGTGGCAGTG**C**GGACGAGGAGGC 3'
WT H1293/K1294	5' AGACACTCAGGA**CA**CAAGGCCGAGAACGGC 3'
H1293F/K1294R u	5' AGACACTCAGGA**TT**CA**G**GGCCGAGAACGGC 3'
H1293F/K1294R l	5' GCCGTTCTCGGCC**C**TG**AA**TCCTGAGTGTCT 3'

2.2.10 Commercial Kits

DELFIA Eu-Labelling Kit	Wallac
ToTALLY RNA™ Extraction Kit (#1910)	Ambion
Ultraclean™ DNA purification	Geneworks
Cell Titre Glo™	Promega

2.2.11 Molecular Weight Standards

2.2.11.1 DNA Markers

Hpa II digested pUC19 markers were purchased from Geneworks.
Band sizes (bp): 501, 489, 404, 331, 242, 190, 147, 111, 110, 67, 34, 34, and 26.

EcoR1 digested SPP-1 markers were purchased from Geneworks.
Band sizes (bp): 8557, 7427, 6106, 4899, 3639, 2799, 1953, 1882, 1515, 1412, 1164, 992, 710, 492, 359, 81.

2.2.11.2 Protein Markers

SDS7 markers were purchased from Sigma. Protein constituents: Albumin (Bovine) 66kDa, Albumin (Egg) 45kDa, Glyceraldehyde-3-P-Dehydrogenase 36kDa, Carbonic Anhydrase (Bovine) 29kDa, Trypsinogen (Bovine Pancreas) 24kDa, Trypsin inhibitor (Soybean) 20kDa, α-Lactalbumin (Bovine milk) 14.2kDa

The following protein markers were purchased from Gibco:
Benchmark™ Prestained Protein Ladder molecular weight markers:
Contains a mixture of 10 proteins stained blue except 67 kDa stained pink. Approximate MW (kDa): 221, 133, 93, 67, 56, 42, 28, 23, 17, and 11.

2.2.12 Solutions

The Milli-Q Reagent Water System was employed to filter the water for these experiments. All solutions were sterilised by autoclaving or filtering with a 0.45 μM filter. Tissue culture solutions were filtered *in vacuo* using Corning disposable bottle top 0.22 μM filters. The following solutions were used.

General Laboratory Solutions

ABTS reagent: 60nM citric acid buffer (pH 9), 0.12% H_2O_2, 1 ABTS tablet per 10mls
Anode Buffer: 0.2 M tris (pH 8.9).
Ammonium acetate DNA buffer 1: 25mM tris (pH 7.5), 10mM EDTA (pH 8), 0.9% (w/v) glucose and 2 mg/ml Lysozyme.
Ammonium acetate DNA buffer 2: 1% (w/v) SDS and 0.2M NaOH.
Bicarbonate Buffer: 150mM Na_2CO_3 (pH 9.5)
Blocking Solution: 5% skim milk powder, PBST
Cathode Buffer: 0.1M tris, 1%SDS, 1 M tricine.
Citric Acid Buffer: 60mM citric acid and 81mM Na_2HPO_4 with a final pH of 4.
Coomassie blue stain: 0.1%(w/v) Coomassie brilliant blue, 30%(v/v) methanol, 10%(v/v) acetic acid.
Destain: 50%(v/v) methanol, 5%(v/v) glacial acetic acid.
ECL solution 1: 2.5mM luminol, 0.4mM courmaric acid, 0.1M tris (pH 8.5).
ECL solution 2: 0.0192% (v/v) H_2O_2, 0.1M tris (pH 8.5).
Europium elution buffer: 50mM tris (pH 7.8), 0.9% (w/v) NaCl and 0.05% (w/v) sodium azide.
Europium labelling buffer: 100mM Na_2Co_3 (pH 9.3).
Europium lysis buffer: 150mM NaCl, 1.5mM $MgCl_2$, 1mM EGTA, 20mM hepes, 10% (v/v) glycerol, 0.1% (v/v) triton-X-100, 1mM PMSF, 1mM Na_3VO_4, 10mM NaF and 1x Protease inhibitor cocktail.
Gel Buffer {3X}: 3M tris base, 0.3%(w/v) SDS.
GLB {10X}: 50% (v/v) glycerol, 0.05% (w/v) bromophenol blue and 0.05% (w/v) xylene cyanol.
Hepes Buffer (^{125}I-GF): 100mM HEPES, 120mM NaCl, 5mM KCl, 1.2mM $MgSO_4.7H_2O$, 8mM Glucose, 0.5% BSA

IP binding buffer: 10mM tris (pH 7.5), 0.1% triton-X-100 (or NP-40), 2mM EDTA, 1mM PMSF, 1mM Na_3VO_4, 10mM NaF and 1x Protease inhibitor cocktail.
IP wash buffer: 10mM tris (pH 7.5), 0.1% triton-X-100 (or NP-40), 0.15M KCl, 2mM EDTA, 1mM PMSF, 1mM Na_3VO_4, 10mM NaF and 1X Protease inhibitor cocktail.
Ligand binding buffer: 100mM Hepes (pH 8), 100mM NaCl, 2μM 4, 4'Dithio-bis(phenyl azide) and 0.05% (v/v) tween-20.
Ni-IDA binding buffer: 0.5M NaCl, 5mM imidazole, 20mM tris (pH 7.9).
Ni-IDA wash buffer: 0.5M NaCl, 60mM imidazole, 20mM tris (pH 7.9).
Ni-IDA elute buffer: 0.5M NaCl, 1M imidazole, 20mM tris (pH 7.9).
Ni-IDA strip buffer: 0.1M EDTA, 0.25M NaCl, 20mM tris (pH 7.9).
Ni-IDA charge buffer: 50mM $NiSO_4$
PBS: 8% (w/v) NaCl, 0.02% (w/v) KCl, 0.02% (w/v) KH_2PO_4, 0.115% (w/v) Na_2HPO_4. PBST is PBS supplemented with 0.1% (v/v) Tween-20.
Resolving Gel: 8% (v/v) acrylamide:bis 28:1, 10.5% (v/v) glycerol, 1 x gel buffer, 0.0014% (v/v) temed and 0.04% (w/v) APS.
SD buffer {10X}: 330mM tris (pH 8.3), 625mM KAc, 100mM MgAc, 40mM spermidine, 5mM DTE.
SDS load buffer {5X}: 0.25M tris (pH 6.8), 5% (v/v) β-mercaptoethanol, 10% (w/v) SDS, 0.5% (w/v) bromophenol blue and 50% (v/v) glycerol.
Stacking Gel: 4.5% (v/v) acrylamide:bis 28:1, 1 x gel buffer, 0.0009% (v/v) temed and .078% (w/v) APS.
Stripping Buffer: 0.1M β-mercaptoethanol, 2% SDS, 62.5mM tris (pH 6.7).
TAE: 40mM tris, 20mM NaAc, 10mM EDTA (pH 8.2).
TBS: 25mM tris (pH 7.4), 0.137M NaCl, 2.7mM KCl. TBST is TBS supplemented with 0.1% (v/v) Tween-20.
Western Transfer Buffer: 25mM tris (pH 8.4), 0.25M glycine and 20% methanol.
WCE buffer: 20mM HEPES, 420mM NaCl, 0.5% NP-40 (or 1% triton-X-100), 25% Glycerol, 0.2mM EDTA, 1.5mM $MgCl_2$ and freshly added 1mM PMSF, 1mM Na_3VO_4, 10mM NaF and 1x Protease inhibitor cocktail.

2.2.13 Online and Computing Resources

Data Analysis and Representation:

Excel 2000, Microsoft, (http://www.office.microsoft.com/)

Prism v3.03, GraphPad Software Inc. (http://www.graphpad.com/prism/)

DNA and Protein sequence visualisation, analysis and manipulation:

MacDNASIS v2.0, Hitachi software Engineering America Ltd. (http://www.hitachi-soft.com)

ABI Prism™ EditView v1.01, Applied Biosystems, (http://www.appliedbiosystems.com)

Chroma v1.0, (http://www.lg.ndirect.co.uk/chroma)

DNA and protein sequence alignment:

ClustalW v1.7, European Molecular Biology Laboratory, Heidelberg, Germany.(http://www.ebi.ac.uk/clustalw/) (Thompson, *et al.*, 1994)

T-Coffee v1.37, (http://www.ch.embnet.org/software/TCoffee.html) (Notredame *et al.*, 2000)

Protein structure file visualisation:

VMD v1.8a21, University of Illinois Theoretical and Computational Biophysics Group, (http://www.ks.uiuc.edu/Research/vmd/) (Humphrey *et al.*, 1996)

DS ViewerPro^{TM} v5.0, Accelrys Inc. (http://www.accelrys.com/viewer/)

Comparative Modelling:
MODELLER 6v2, (http://structbio.vanderbilt.edu/comp/soft/modeller/).

Protein structure refinement and evaluation:

NAMD v2.5, University of Illinois, Theoretical and Computational Biophysics Group, (http://www.ks.uiuc.edu/Research/namd/).

Procheck v3.5.4 (http://www.biochem.ucl.ac.uk/~roman/procheck/procheck.html).

ProSa 2003 v4.0, Center of Applied Molecular Engineering, ProCeryon Biosciences (http://www.proceryon.com/solutions/prosa.html)

STAMP Structural Alignment of Multiple Proteins v4.2, University of Oxford, Molecular Biophysics (http://www.compbio.dundee.ac.uk/Software/Stamp/stamp.html). (Russel and Barton, 1992)

2.3 METHODS

2.3.1 Bacterial Methods

2.3.1.1 Making Glycerol Stocks

A small-scale bacterial culture was set up in selective LB medium from a single transformed *Escherichia coli* colony and grown overnight at 37°C, shaking. The following morning, triplicate 500µL samples were diluted 1:1 with 80% glycerol, mixed and stored at –80°C.

2.3.1.2 Calcium Chloride Competent Cells

A small-scale bacterial culture was set up in LB medium from a single *E. coli* colony (DH5α, BL21(DE3) strain) and grown overnight at 37°C with shaking. The following morning, a 1% (v/v) subculture was made in 500 mL of LB medium and the culture was expanded to an optical density at 600nm wavelength (OD_{600nm}) of 0.4-0.5. The culture was chilled on ice for 10 min and bacterial cells were collected by centrifugation at 2990 x g (4°C) for 15 min. Pellets were then washed in 100 mL ice cold $CaCl_2$ solution (0.06 M $CaCl_2$, 15% v/v glycerol), centrifuged at 2990 x g (4°C) for 15 min and resuspended in 20 mL ice cold $CaCl_2$ solution. Approximately 200µL aliquots in Eppendorf tubes were snap frozen in a dry ice/ethanol bath and stored at -80°C until used.

2.3.1.3 Transformation of Competent Cells by Heat Shock

Competent cells were transformed with circular plasmids by the heat shock method as follows. Plasmid DNA or 20. µL ligation reaction was mixed with freshly thawed 200. µL $CaCl_2$ competent cells and incubated on ice for ~10-20 min before heat shock at 42°C for 2 min and then incubation on ice for a further 10 min. Cells were then incubated at 37°C for 10min before being plated onto LB agar plates with appropriate selection and incubated overnight at 37°C.

2.3.1.4 Large Scale Midiprep Kit

100 mL of overnight culture was processed using the QIAGEN Maxi prep kit according to the manufacturers recommendation. Briefly, cells were pelleted at 2990 x g for 10 min and resuspended in QIAGEN kit Buffer P1 and lysed by the addition of buffer P2. After addition of buffer P3 contaminants were removed by filtration through a disposable column into a QIAGEN-tip prepared by washing with buffer QBT. Plasmid DNA attached to the tip was washed with buffer QC and eluted in buffer QF, precipitated with isopropanol, washed with 70% ethanol and resuspended in TE. Approximately 500 µg of plasmid DNA suitable for mammalian cell transfection was obtained from each preparation.

2.3.1.5 Medium Scale Plasmid Preparation

Single bacterial colonies or frozen glycerol samples were inoculated into 50 mL of LB containing 100 µg/mL ampicillin, and incubated overnight at 37°C in an orbital shaker. ~40-45 mL of bacterial culture was aliquoted into oakridge tubes, and spun at 2990 x g at 4°C for 5 minutes. The supernatant was removed, and the pellet resuspended in 3 mL of TEG buffer (50mM Glucose, 25mM Tris-Cl pH 8.0, 10mM EDTA pH 8.0). 6 mL of fresh Lysis buffer (0.2M NaOH, 1% (w/v) SDS) was then added and the solution incubated on ice for 5 minutes. 4.5 mL of potassium acetate solution (5M KOAc pH 5.2, glacial acetic acid) was added and incubated for a further 20 minutes on ice. Following centrifugation at 30,000 x g for 15 minutes at 4°C, the supernatant was transferred to oakridge tubes containing 8mL isopropanol. These were centrifuged at 30,000 x g at 4°C for 5 minutes and the pellet resuspended in 400 µL H_2O. RNA was digested by adding 2 µL RNase A (10mg/mL) and incubating at 37° for 30 minutes. Proteins were removed by phenol/chloroform extraction in which an equal volume of phenol/chloroform was added, the mixture vortexed and centrifuged at 30,000 x g for 2 minutes and the aqueous layer removed. Residual phenol was removed by adding an equal volume of chloroform, vortexing and spinning. The aqueous layer was again removed and the DNA precipitated by adding $^1/_{10}$ volume of 3M NaAc pH 5.2, 1mL 95% (v/v) EtOH and incubating at -20° for 10-20 minutes. The DNA was pelleted by spinning at 30,000 x g for 10 minutes. The pellet was washed in 400 µL 70% (v/v) EtOH and vacuum dried, before being resuspended in 100-200 µL of 1xTE.

2.3.1.6 Small scale kit

Up to 5 mL of overnight culture was processed using the Ultraclean™ miniprep kit according to the manufacturers recommendation. Briefly, cells were pelleted and all traces of media were removed before resuspension in Ultraclean™ miniprep kit Solution 1. Cells were lysed by the addition of kit solution 2. After addition of kit solution 3 contaminants were removed by centrifugation at 20,800 x g for 1 min and filtration through a disposable column. Plasmid DNA attached to the column was washed with kit solution 4 by centrifugation and eluted in TE with centrifugation. Approximately 10 µg of plasmid DNA suitable for DNA sequencing or mammalian cell transfection was obtained from each preparation.

2.3.1.7 Small Scale Plasmid Preparation

Single bacterial colonies were inoculated from an agar plate or from a frozen glycerol sample into 2 mL of LB containing the appropriate antibiotic (100 µg/mL amp), and incubated overnight at 37°C in a rolling drum. Plasmid DNA was extracted by the rapid miniprep method. 1.5 mL of bacterial culture was aliquoted into an eppendorf and centrifuged at 14 000 x g for 30 seconds, then resuspended in ~50 µL of supernatant. This solution was then vortexed with 300 µL of Megadeath solution (0.1M NaOH, 0.5% (w/v) SDS, 10mM Tris pH 8, 1mM EDTA). 150 µL of 3M NaAc pH 5.2 was added, and the cell debri pelleted at 14 000 x g for 3minutes. Plasmid DNA in the supernatant was transferred to a fresh eppendorf, 1mL of cold 95% (v/v) ethanol added, and the mixture vortexed and spun at 14 000 x g for 4 minutes to pellet the plasmid DNA. The supernatant was aspirated, and cleaned with 400 µL of 70% (v/v) ethanol. After spinning for 4 minutes, the supernatant was removed and the pellet vacuum dried and resuspended in 20 µL of TE. Samples were treated with RNase A (~0.5µg/µL) and incubated at 37°C for 30 minutes.

2.3.2 Molecular Methods

2.3.2.1 Agarose Gel Electrophoresis

Agarose minigels were prepared by pouring 10-12mLs of gel solution (1-3% (w/v) Agarose (Sigma) in 1 x TAE) onto a 5.0cm x 7.5cm glass microscope slide. Gels were submerged in 1 x TAE and samples containing 1 x GLB were electrophoresed at 70-100 Volts. DNA or RNA was stained by a 1 µg/mL solution of EtBr and visualised by exposure to medium length UV light.

2.3.2.2 Purification of Linear DNA fragments

DNA samples were excised from agarose minigels using a sterile scalpel blade and purified using the ULTRAclean Nucleic Acid Purification Kit, from Geneworks, according to the instructions.

2.3.2.3 Restriction Endonuclease Digestion of Plasmid DNA

Plasmid DNA was digested with 2-5 units of enzyme/µg of DNA. Restrictions were done in SD buffer and incubated at 37°C for 45-60 minutes. Digestions were examined by TAE agarose gel electrophoresis.

2.3.2.4 Ligation Reactions

Ligation reactions contained ~50 ng of restricted, purified vector, a 3 fold molar excess of purified insert, 1 x Ligation Buffer with 1mM ATP and 1U of T4 DNA ligase in a 20 µL volume. Reactions were incubated for 2-4hrs or overnight at 16°C before transformation.

2.3.2.5 RNA purification from cells in tissue culture

The TRIzol™ RNA extraction protocol was employed to obtain RNA from tissue-culture cell pellets (5 x 10^6 cells) (Chomczynski and Sacchi, 1987). All solutions, including the electrophoresis gel and buffers were prepared with DEPC treated water. Briefly outlined, the procedure involved addition of 1 mL of reagent to these sample amounts in an eppendorf and tituration performed to ensure thorough homogenisation. After a 5 min RT incubation to allow complete disassociation of all complexes, 0.2 mL of chloroform was mixed vigorously by shaking and the phases allowed to seperate during 3 mins of incubation at room temperature. Micro-centrifugation at 12, 000 x g for 15 mins was followed by removal of the aqueous layer to a fresh eppendorf. The RNA was then extracted from the solution by addition of 0.5mL of isopropanol and a 10 min RT incubation. RNA was precipitated by centrifugation at 12, 000 x g for 10 mins, and the pellet washed with 75% (v/v) ethanol before being air-dried and resuspended in 50 µL RNase-free water.

RNA samples from either extraction method were electrophoresed (Section 2.3.2.1) to visualise the 18S (1.8kb) and 28S (5kb) ribosomal subunits as well as any contamination by DNA or protein. Spectroscopy was performed to obtain RNA concentrations (OD_{260nm} of 1 = 40mg/mL) and to ensure purity of samples (the 230:260 ratio required to be between 1.7 and 2.1 and a 260:280 ratio of 2 required). Only clean and intact samples were used for reverse transcription into cDNA (Section 2.3.2.7), if necessary samples underwent re-extraction with a phenol:chloroform mixture followed by salt precipitation to remove contaminating protein or were digested with RNase-free DNase 1 (0.5mg/mL) at 37°C(#458).

2.3.2.6 mRNA Purification from whole *c. elegans* RNA

mRNA was isolated from total RNA using the Micro-Fast Track™ Kit (Invitrogen). Briefly, *c. elegans* RNA stored in ethanol was pelleted and vacuum dried to remove the ethanol. Pellets were resuspended in 1 mL Micro-Fast TrackTM Lysis buffer and 63 uL 5M NaCl . An Oligo dT cellulose tablet was added to each tube and allowed to swell for 2 min before dispersion with a sterile tip. After a 20 min incubation at RT, the oligo dT cellulose was pelleted by a 4000 x g spin for 6 min and the supernatant removed. The cellulose pellet was

then washed three times with 1.3 mL Micro-Fast TrackTM Binding buffer with a 4000 x g spin for 6 min between each wash. The pellet was finally resuspended in 300 µL Binding buffer and added to a spin column, followed by a 10 sec sin at 4000 x g. A further three wash steps were performed with 1 mL Binding buffer on the column with intervening 10 sec spins. Non-polyadenylated RNA was removed with two 200 µL Micro-Fast TrackTM Low Salt Wash Buffer wash steps. mRNA was obtained from the column by adding two lots of 100 µL Elution Buffer with subsequent 10 sec spins at 4000 x g and collecting the flow through. mRNA was precipitated and stored for future use by the addition of 10 µL glycogen (2 mg/ml), 30 µL 2M NaAc and 600 µL 95% EtOH to the 200 µL of elution buffer and -70 C storage. mRNA for use was recovered by centrifugation and resuspension in 20 µL 10mM Tris pH 7.5.

2.3.2.7 Reverse Transcription

Reverse transcription was performed on various extracted RNA samples generated in Sections 2.3.2.5 and 2.3.2.6 to create cDNA to be used as a template in PCR (Section 2.3.2.9). In general, 1-2 µg of RNA was added to a 20 µL final solution containing 1 x RT buffer, 0.5 U/µL m-MLV Reverse Transcriptase, 0.625mM random hexamers, 10mM DTT, 1mM dNTPs and MQ water. This solution was incubated at 37°C for 1 hour to allow reaction to occur. Often to encourage extension of full-length cDNA containing the 3' end, an additional 3' or 3'UTR primer was 'spiked' into reaction at a final concentration of 2.5ng/µL. The reaction was then either stored at 4°C until required or immediately added as template (2-3µL) to a PCR (Section 2.3.2.9).

2.3.2.8 Primer Design

Primers for polymerase chain reaction (PCR) reactions were designed with the aid of Oligo^{TM} software. Generally, primers were >17 bp long, terminated in a G/C residue, had >40% G/C content, unique 3' end of 7 nucleotides (compared to the template) and did not self anneal. Primer pairs had similar melting temperature at which half maximal binding occurs. Primers designed to amplify sequences for cloning incorporated a restriction site at the beginning of the coding sequence and a different restriction site and stop codon at the end of the coding sequence.

2.3.2.9 Polymerase Chain Reaction (PCR)

PCR was used to produce insert domain constructs, check the exon 11 status of hIR constructs, and to amplify *C. elegans* genes from cDNA. Amplification reactions contained

PCR buffer, deoxynucleotides, oligonucleotides, template DNA and Taq polymerase. The amplification program used had the following basic conditions for denaturation and annealing, with variations as stated in the text to improve yield or purity : Cycle 1, 94°C 5 minutes / 40°C 1 minute, Cycle 2, 94°C 1 minute / 40°C 1 minute, Cycles 3-30, 94°C 1 minutes / 56°C 1 minute. The program was carried out by the PTC-100 Progammable Thermal Controller.

2.3.2.10 Cycle Sequencing of Plasmid DNA

DNA sequencing was performed with reactions containing 0.5-1.0 µg of plasmid DNA, 100ng sequencing primer and 6 µL DYETERM™ or BIGDYE™ ready reaction mix (AP Biotech) in a total volume of 15 µL. 26 rounds of PCR were performed with denaturation at 96°C for 30 sec, annealing at 50°C for 15 sec and extension at 60°C for 4 min. The PCR reaction was precipitated with ethanol (70% final) and sodium acetate (85mM final) at -20°C for 15 min followed by centrifugation at 30,000 x g for 10 min. Pellets were washed with 200 µL 70% (v/v) ethanol and dried before being processed at the IMVS DNA sequencing facility. Alternatively, the PCR reaction was precipitated with isopropanol (60% (v/v) final) at room temperature for 15 min followed by centrifugation at 30,000 x g for 20 min. Pellets were washed with 250 µL 75% v/v isopropanol and dried before being processed at the IMVS DNA sequencing facility which uses an ABI sequencer (Applied Biosystems).

2.3.3 In vitro Methods

2.3.3.1 Culture of Mammalian Cells

Standard tissue culture techniques were used in the culture of mammalian cells. These included thawing, subculturing, and freezing which are briefly described as follows. Cells in a vial stored in liquid N_2 were thawed rapidly in a 37°C water bath before dilution in 10 mL media, centrifugation (290 x g, 2 min) and resuspension in fresh medium and cultured. To passage cells, adherent cell monolayers were washed once with PBS and treated with trypsin (1 mL per 75 cm^2) for 5 min at room temperature. Detached cells were resuspended in medium containing 10% fetal bovine serum (FCS), centrifuged (290 x g, 3 min) and resuspended in fresh medium. Where appropriate a haemocytometer grid was used to estimate cell concentration prior to subculturing. Generally, cells were cultured in medium consisting of DMEM supplemented with 10% FBS. Freezing cells for long term storage was done by growing the cells to ~80-90% confluence, trypsinising them as above, followed by a pelleting at 290 x g for 2min, The trypsin/media was aspirated and the cells resuspended in freeze mix

(95%FCS, 5% DMSO) and aliquoted into freezing vials and placed at –80°C. Vials were then moved to storage under liquid nitrogen.

2.3.3.2 Cell counting using cytometer

Cell counts using a cytometer were performed on cells trypsinised to a single cell suspension. Depending on a visual approximation of the cell density either 5 µL or 10 µL of cells were added to 10% (w/v) Trypan blue in 1xPBS. Five large squares on the cytometer were counted and the cell concentration in cells/mL calculated according to the formula:

(Total cells from 5 squares/5) x Dilution factor x 10^4

Dead or dying cells were distinguished from viable cells by their ability to take up trypan blue stain.

2.3.3.3 Transfection of cells using Lipofectamine +™

The day before transfection the cells were counted and plated to a cell density to allow growth overnight to 50-80% confluence in a 10cm dish. For each transfection 4 µg DNA was mixed with 750 µL of serum free medium (OPTIMEM™) and 20 µL of PLUS Reagent™ and incubated at room temperature for 15 min. 30 µL of Lipofectamine + Reagent was diluted in 750 µL serum free medium and mixed with the DNA solution and incubated for a further 15 min. Cells to be transfected were washed once in 1xPBS and replaced with 5 mL serum free medium. The combined DNA/Lipofectamine mix was then added evenly to the top of the cells and mixed gently by rocking the dishes. After a 3 hr incubation at 37°C, the transfection media was replaced with complete medium. To select for stable transfectants antibiotics were added 24 hrs after the transfection to allow the cells time to recover form the process.

2.3.3.4 Methylene Blue cell viability assay

Cells to be assayed for viability by methylene blue staining were grown in clear Falcon 96 well tissue culture plates. Media was removed by aspiration and each well washed x1 with 0.15M NaCl. The cells were then fixed by the addition of 100 µL of methanol to each well and incubation at room temperature for 30 min. The methanol was subsequently aspirated and cells stained by adding 70 µL Methylene Blue Stain (1% (w/v) Methylene Blue/0.01M Borate) to each well followed by a further 30 min incubation at room temperature. The stain was aspirated and each well washed x2 with 0.01M Borate Buffer. The dye was eluted from the cells by adding 100 µL/well of 1:1 (v/v) 0.1M HCl/Ethanol. Plates were left for at least 10

min and shook to ensure even distribution of the dye. Absorbance at 580 nm was then read using the PolarStar Galaxy (BMG Technologies).

2.3.3.5 Cell titre Glo assay

Cells to be assayed for viability by the CellTitre Glo™ Assay were grown in clear Falcon 96 well tissue culture plates in a volume of 100 µL of media. Assays were essentially developed according to the manufacturers instructions. Briefly, CellTitre Glo™ buffer and substrate were thawed at room temperature, and mixed together to dissolve the substrate. Tissue culture plates were allowed to equilibrate to room temperature for 30 min prior to the addition of 100 µL/well of the combined CellTitre Glo™ reagent. Contents were mixed for 2 min on an orbital shaker to induce cell lysis and then the plate left for 10 min at RT to stabilize the luminescent signal. Luminescence was then recorded with an integration time of 0.5 sec/well on the PolarStar Galaxy (BMG Technologies).

2.3.3.6 Cell Lysis

Adherent cells to be lysed had their media aspirated and were washed once in 1xPBS. Ice cold Lysis buffer (20mM HEPES, 150mM NaCl, 1.5mM MgCl2, 10% (v/v) Glycerol, 1% (v/v) Triton X-100, 1mM EGTA, pH 7.5) containing 1/100 dilution of Protease Inhibitor mix (Sigma) was then added to the plate or well and incubated at 4°C for at least 30 min. If intact tyrosine phosphorylation was to be assayed for then 5mM Na Orthovanadate and 10mM NaF was also included in the lysis buffer. The cell lysate was then transferred to a new tube and spun at 2,500 x g for 5 min at 4°C before being used or stored in eppendorf tubes at –80°C.

2.3.3.7 Basic FACS analysis

Media was aspirated and cells washed x1 in 1xPBS. Cells were harvested with the addition of trypsin/EDTA (1.5ml/10cm dish, 1ml T25 Flask). Cells were removed in pre-warmed media into a 20ml blue cap.

Harvested cells checked for being a single cell suspension were centrifuged at 290 x g for 2 min. The media/trypsin mix was aspirated and the cell pellets were washed in1xPBS and spun again at 290 x g for 3min. The supernatant was aspirated and the cells resuspended in 1xPBS. Aliquots of ~20 µL cells for each sample were put into Falcon 2008 tubes. 50 µL of 1/50 diluted 1° Antibody (either 83-14 for hIR or 16-13 for hIGF-IR) and incubated on ice for ~40

min. Cells were washed x1 with 3 mL 1xPBS and pelleted at 290 x g, 3 min. The supernatant was again aspirated and the cells resuspended in 50 µL 1/50 diluted FITC labelled goat anti-mouse antibody and incubated on ice for a further 40 min. Cells were washed x2 in 3 mL 1xPBS, after each wash pelleting the cells by centrifugation at 290 x g for 3 min. Labelled cells were finally resusupended and fixed in 500 µL 1% paraformaldehyde in 1xPBS and stored in the dark on ice until FACS acquisition.

FACS machine Instrument settings for R- cells
P1 FSC E00 1.00 Lin
P2 SSC 350 1.00 Lin
P3 FL1 510 1.00 Log
P4 FL2 560 1.00 Log
P5 FL3 667 1.00 Log

2.3.3.8 Phosphorylation Assay

Stable cell lines expressing either IR or IGF-1R were plated in clear Falcon 96 well tissue culture plates at a cell density of 25, 000 cells/well and grown overnight at 37°C. The media was aspirated and 50 µL serum-free DMEM added to each well and incubated at 37°C for 4 hrs. Following this period, cells were stimulated with 50 µL pre-warmed serum-free media containing growth factor and incubated for 10-15 min at 37°C. The media was then quickly aspirated and 120 µL ice cold lysis buffer added to each well (see section 2.3.3.6). 100 µL of cell lysate was then transferred to a white lumintrac 600 96 well plate, which had been pre-coated with the appropriate anti-receptor antibody (See section 2.3.4.3) and incubated overnight at 4°C. The amount of tyrosine phosphorylation present in each well was then assayed (see section 2.3.4.7).

2.3.3.9 Time-course Phosphorylation Assay

Stable cell lines expressing either IR or IGF-1R were plated in clear Falcon 96 well tissue culture plates at a cell density of 25, 000 cells/well and grown overnight at 37°C. The media was aspirated and 50 µL serum-free DMEM added to each well and incubated at 37°C for 4hrs. In the case of phosphatase inhibition, sodium othovanadate was added to the media to a concentration of 30 µM for 30 min prior to growth factor stimulation (Tan *et al.*, 2001). Following this period, on a bench at room temperature, cells were stimulated with 50 µL pre-warmed serum-free media containing growth factor and time points taken at 0, 15, 30, 60, 120, 300, and 600 sec. At each time point the media was quickly aspirated and 120 µL ice cold lysis buffer added to each well (see section 2.3.3.6). 100 µL of cell lysate was then

transferred to a white lumintrac 600 96 well plate which had been pre-coated with the appropriate anti-receptor antibody (See section 2.3.4.3) and incubated overnight at 4°C. The amount of tyrosine phosphorylation present in each well was then assayed (see section 2.3.4.7).

2.3.4 Protein Methods

2.3.4.1 Western Blot

Briefly, proteins were separated using 8%, 10% or 12.5% (v/v) Tris/Tricine/SDS/acrylamide gels with a Hoefer gel cast system and transferred to Hybond-C (Amersham) membrane at 40mA/gel for 90 min in 1× WTB containing 20% (v/v) ethanol using a Hoefer SemiPhor Western Transfer apparatus (Pharmacia Biotech). A Hoefer 500 V/400 mA/200 W DC power supply was used. Membranes were blocked overnight in blocking solution washed where appropriate in wash solution and probed with antibody diluted in blocking solution. Secondary antibodies: anti-rabbit-HRP, anti-mouse-HRP and anti-goat-HRP were used at 1:4000. Enhanced chemiluminescence (ECL) detection was used as follows. Membranes were developed for 1 min in a freshly combined mixture of equal volumes of ECL solutions 1 and 2 and exposed to X-ray film that was developed using a CURIX 60 X-Ray developer. Digital images were recorded using a Canon laser scanner and UMAX MagicscanII software. To strip Western blots for reuse, they were submerged in Stripping buffer at 50°C for 30 min with occasional agitation. After washing twice in wash solution, blots were blocked overnight in blocking solution.

2.3.4.2 Europium-labelling proteins

The DELFIA Eu-labelling kit was used to europium label either antibodies or growth factors on the N-terminal and side chain amine groups. Proteins to be labelled were either resuspended from solid or buffer exchanged using a PD10 column (See 2.3.4.10), into labelling buffer (100mM Na_2CO_3 pH 8.5). In each case ~1mg of protein was mixed with Eu-labelling Reagent (used at a molar ratio of ~ 3:1 Eu:protein) and incubated in the dark at 4°C, overnight for 24hrs. Reactions were quenched with L-Lysine for 15min and the protein separated from unreacted eu-labelling reagent using a PD10 column and eluting from the column with 50mM Tris-HCl pH 7.5, 0.9% (w/v) NaCl, 0.05% (v/v) NaN_3. 500 µL fractions were collected in siliconised tubes and subsequently tested for protein content using a BCA assay (See 2.3.4.8) and the presence of europium using time-resolved fluorescence (See

2.3.4.5). Fractions containing both protein and europium were assumed to be labelled and stored at 4°C in the elution buffer.

2.3.4.3 Preparation of 96-well Plates with anti-receptor antibody

Receptor capturing plates were prepared using white 96 well Lumintrac 600 plates and anti-receptor antibodies 24-31 (α-IGF-IR) or 83-7 (α-IR). Antibodies were diluted in 1xPBS at a concentration of 2-2.5 µg/mL and 100 µL added to each well. Coating was allowed to occur at 4°C overnight, after which wells were aspirated and blocked by the addition of 125 µL/well of 0.5% (w/v) receptor grade BSA (SIGMA) in 1xPBS. This was incubated for up to 3 hrs at 37°C. Wells were again aspirated and washed x2 with TBST (1XTBS and 0.05% Tween 20). Plates were either used immediately or stored at -20°C in 200 µL/well TBST.

2.3.4.4 Insulin Receptor ELISA

Receptor capturing plates were prepared using clear 96 well TC grade plates and coated with the anti-insulin receptor antibody 83-7. Antibodies were diluted in 1xPBS at a concentration of 2-2.5 µg/mL and 100 µL added to each well. Coating was allowed to occur at 4°C overnight, after which wells were aspirated and blocked by the addition of 125 µL/well of 0.5% receptor grade BSA (SIGMA) in 1xPBS. This was incubated for up to 3 hrs at 37°C. Wells were again aspirated and washed x2 with 1xPBS. Plates were either used immediately or stored at -20°C in 200 µL/well PBS. 100 µL samples were then added to the wells for 2-4 hr at RT or overnight at 4°C. Following a further x2 washes with 1xPBS, 100 µL of biotinylated 83-14 diluted 1/1000, was added to each well. Wells were washed again with 1xPBS and BCIP substrate added and absorbance measured at 405 nm.

2.3.4.5 Competition Assay with eu-labeled growth factors

Prepared and washed plates were tapped dry and 100 µL/well of soluble or solubilised receptor was added (if purified then at 0.5 µg/mL, otherwise undiluted). Having incubated overnight at 4°C, wells were aspirated and washed twice with 200 µL/well TBST and patted dry. Eu-labeled growth factor was diluted in Ligand Binding Buffer (LBB: 100mM HEPES pH 8, 100mM NaCl, 0.05% (v/v) tween 20) and 50 µL added to each well (See section 4.2.5 for a discussion on the amount of eu-growth factor to use). 50 µL of unlabelled ligand diluted in LBB to double the required concentration was added to each well in triplicate and the plate incubated overnight at 4°C. The mix was aspirated and the wells washed x3 with 200 µL MQ

H_20 and finally patted dry. 100 µL/well of Wallac Enhancement solution was added to release the europium and left for ~10 min before measuring the time-resolved fluorescence from a PolarStar Galaxy (BMG lab technologies).
Instrument settings- Excitation Filter: 340-12, Emission Filter: 612-12, No of Flashes: 50, Integration Delay (µs): 0400, Integration Time (µs): 0800

2.3.4.6 ^{125}I growth factor competition assay

R- cells, wild-type or stably transfected, were plated in 24 well trays to yield 80-95% confluence overnight. Following at least 2 hrs growth in serum free medium, cells were washed x1 with ice cold 1xPBS. ^{125}I-Growth Factor (IGF-1, 2 or Insulin(Y19)) was diluted in ice cold HEPES buffer (100mM HEPES, 120mM NaCl, 5mM KCl, 1.2mM $MgSO_4.7H_2O$, 8mM Glucose, 0.5% BSA) to give ~10-15, 000 cpm/400 µL and 400 µL added to each well. Unlabeled growth was serially diluted (at x5 required concentration) in ice cold HEPES buffer and 100 µL added to each well. Having been left overnight at 4°C, competition mixes were aspirated and washed x1 with 2.5 mLs 1xPBS. Following a further aspiration, 1 mL of 0.5M NaOH/0.1% Triton X-100 was added to each well, triturated and transferred to eppendorfs for counting on a multigamma counter (CSIRO).

2.3.4.7 Detection of Tyrosine Phosphorylation using eu-labeled anti-phosphotyrosine Antibody

Prepared and washed plates were tapped dry and 100 µL/well of solubilised receptor was added (See section 2.3.3.6 for lysis method). Having incubated overnight at 4°C, wells were aspirated and washed x2 with 200 µL/well TBST and patted dry. Eu-labelled anti-phosphotyrosine antibody was diluted in LBB and 100 µL added to each well and incubated in the dark for at least 2 hrs at room temperature. The eu-antibody was aspirated and the wells washed x3 with 200 µL MQ H_20 and finally patted dry. 100 µL/well of Wallac Enhancement solution was added to release the europium and left for ~10 min before measuring the time-resolved fluorescence from a PolarStar Galaxy (BMG lab technologies) (Okada *et al.*, 1998).
Instrument settings- Excitation Filter: 340-12, Emission Filter: 612-12, No of Flashes: 50, Integration Delay (µs): 0400, Integration Time (µs): 0800

2.3.4.8 Quantitation of Protein concentration

Protein concentration was determined by either Bradford assay or the bicinchoninic acid (BCA) protein assay. 10 µL of various BSA standards (0 – 1.0 mg/mL) and sample (neat or diluted) were added to a 96 well tray along with either 200 µL of 25% Bradford reagent or BCA mixture. The OD_{600nm} was measured on a 96 well Emax microplate reader (Molecular Devices) or a PolarStar Galaxy (BMG lab technologies). A standard curve was plotted and the sample concentration was measured from the graph. The BCA assay was used more routinely as it has a greater tolerance for the Tween-20 and Triton-X-100 detergents used in protein preparations.

2.3.4.9 SDS-PAGE Electrophoresis

Denaturing SDS-PAGE gels (0.75–1 mm thick, 5-12.5%) were prepared using a Hoefer™ gel pouring apparatus. The resolving-gel was allowed to polymerise with an overlay of water to form a smooth interface. Once the gel had set and the water removed, a stacking-gel was added and allowed to polymerise with 10 or 15 well comb inserted. Samples were mixed to a final concentration of 1x SDS load buffer and heated to 100°C for 5 min before loading (up to 20 µL) onto the gel. Gels were electrophoresed at 50 mA per gel using a continuous buffer system with Anode Buffer and Cathode Buffer. The gels were elecrophoresed with cooling, until the ion front reached the bottom of the gel. The gels were then either electroblotted (See Section 2.3.4.1) or stained with Coomassie blue stain (0.1%(w/v) Coomassie brilliant blue, 30%(v/v) methanol, 10%(v/v) acetic acid) to visualise protein content for 30 mins at 37°C. Stain was removed by extensive washing in Destain solution (50%(v/v) methanol, 5%(v/v) glacial acetic acid). A digital image of the gel was recorded using a Canon laser scanner and UMAX MagicscanII software.

2.3.4.10 Ni-Affinity Chromatography

Ni-Affinity Chromatography was used to purify bacterially expressed fusion proteins from pET32a that contain a Histidine Tag. 2mLs of Ni-NTA agarose resin (Zymatrobe) was added to a PD10 column and allowed to settle before washing with 3 bed volumes of MQ H_2O, 5 bed volumes of Charge buffer (50mM NiSO4), and 3 bed volumes of Binding Buffer (0.5M NaCl, 5mM imidazole, 20mM tris pH 7.9). Resin was transferred in Binding Buffer as a slurry into a tube containing the protein sample and incubated for 1-1.5 hrs at 4°C on a shaking platform. The combined sample and resin were loaded back into the column and the flow through collected and repoured. The column was washed with 10 bed volumes Binding Buffer, followed by 6 bed volumes of Wash Buffer (0.5M NaCl, 20-60mM imidazole, 20mM

tris pH 7.9), and remaining protein eluted with 6 bed volumes of Elute Buffer (0.5M NaCl, 200-1000mM imidazole, 20mM tris pH 7.9). 1mL fractions were routinely collected of both the flow through from the Wash and Elute buffers and analysed by Bradford assay (See section 2.3.4.8) and SDS page gel electrophoresis (See Section 2.3.4.9) for protein content. Samples to be kept had 1/100 dilution of Sodium Azide added to them and were stored at 4°C. Ni-NTA resin was stripped of the Ni and any remaining bound protein by adding 3 bed volumes of Strip Buffer (0.1M EDTA, 0.25M NaCl, 20mM tris pH 7.9), followed by 6 bed volumes of MQ H_2O.

2.3.4.11 High Performance Liquid Chromatography Analysis

Reverse phase HPLC was used to separate peptide fragments and to check the purity after speration. A C4 analytical column (Brownlee Aquapore BU300, 7µm particle size, 300-A pore size 4.6X100 mm) and a C18 analytical column (Brownlee Analytical Rp-18, Spheri-5, 5µm particle size, 4.6X100 mm) were used in various studies. Samples were acidified to ~pH 3.0 with 0.1%-2% (v/v) TFA prior to loading onto the column. Columns were equilibrated with 0.1% (v/v) TFA between runs. Various linear gradients using 0.1% (v/v) TFA and 0.088% (v/v) TFA/80% acetonitrile were used to elute peptides from the column. A flow rate of 0.5 mL/min was used and the eluted proteins were detected by absorbance at 215nm (peptide bonds) and 280nm (aromatic rings) using a Flow-through Ultraviolet Detector (Waters, Model 490). Due to the small quantities of proteins being analysed if fractions were collected then siliconised eppendorf tubes were used. Reverse Phase HPLC was controlled using, MilleniumTM chromatography software.

2.3.4.12 Thrombin Cleavage of fusion proteins

Fusion protein purified by glutathione sepharose affinity chromatography was cleaved with bovine thrombin to remove the thioredoxin. Digests were performed in the presence of the chromatography elution buffer, in addition to 2.5mM $CaCl_2$ and between 1-5U/mg fusion protein of Thrombin. Reactions were incubated at 37°C from 1 hr to overnight to ensure complete cleavage occurs. Enterokinase digestions were performed in a similar manner only $CaCl_2$ was added to a final concentration of 10mM.

2.3.4.13 Proteinase K Digestion

Fusion protein purified by glutathione sepharose chromatography was assayed for structural compactness and stability by treatment with Proteinase K (Boehringer Mannheim) to a final

concentration of 1 mg/mL for up to 10 min at room temperature. Aliqouts of 15 µL were taken at several time points for analysis by gel electrophoresis.

2.3.4.14 NMR Spectroscopy

NMR experiments were performed on a Varian Inova 600 mHz spectrometer. Data sets were recorded at 25°C using a 5mm inverse triple resonance $^{1}H/^{13}C/^{15}N$ pfg probe. Insert Domain sample used for NMR spectroscopy were obtained from liophilised samples purified by HPLC. Thioredoxin sample was obtained from liopholised samples purified by Ni-affinity chromatography. These were dissolved in 90% H_2O/10% D_2O. The pH was adjusted using 1M NaOH or HCl.

2.3.5 Computational Methods for Comparative Modelling

2.3.5.1 Protein Sequence Alignments

Ins peptide sequences were aligned for use in comparative modelling using v1.37 of the T-Coffee algorithim (Notredame, Higgins *et al.*, 2000) using the 'slow pair' method, based on the Myers-Miller alignment algorithim (Myers, Selznick *et al.*, 1996).

2.3.5.2 Building the Comparative Models

Template PDB files were edited to remove headers, hetero-atoms, and water molecules. Aligned sequences from T-Coffee were modelled individually on each template in MODELLER (Andrej Sali, Rockefeller University), using the default parameters in addition to restraints specifying which cysteines would be disulphide bonded.

2.3.5.3 Simulated Annealing

Simulated annealing was carried out using a version of the Linge *et al.*, (2003) simulated annealing script that was modified for use in XPLOR and NAMD. Proteins for refinement were prepared by solvation in a 4Å wide shell of TIP3 water, addition of hydrogen atoms, setting the correct cysteines to be disulphide bonded, for final production of protein structure and parameter files using XPLOR and the CHARMM 22 forcefield for proteins (MacKerell, Bashford *et al.*, 1998). Proteins were first minimised by conjugate gradient minimisation for 5000 steps, before 1000 steps at each temperature during ramping from 200K to 500K in 100K increments. Simulation was carried out for a further 5000 steps before the temperature was lowered in 25K increments to 25K, with 500 steps of simulation at each temperature. A final 3000 step conjugate gradient minimisation step was then performed.

2.3.5.4 Model Evaluation

Models were evaluated by determining the overall energy of the structure and the quality of the structure/fold. Free energy was calculated in NAMD using the CHARMM22 forcefield (Kale *et al.*, 1999) (MacKerell *et al.*, 1998). Model energy potentials for structure/fold quality were evaluated using ProSa II v3.0, to determine a combined z-score (zp-comb) of the pair and surface energies, lower energy indicating a more stable structure ($z_p = (E_p - E_{avg.}) / \sigma$). Z-scores were then converted to q-scores to obtain a value independent of peptide length (Sippl, 1993) (q-score = z_p/ ln (peptide length)) (Sanchez and Sali, 1998). 3D structural alignments were performed using STAMP v4.0 to obtain RMSD's between backbone carbons (Structural Alignment of Multiple Proteins) (Russel and Barton, 1992).

CHAPTER 3

Expression Of Insert Domain Proteins

CHAPTER 3 EXPRESSION OF INSERT DOMAIN PROTEINS

3.1 INTRODUCTION

The Insulin Receptor (IR) and the Type-1 Insulin-Like Growth Factor Receptor (IGF-IR) are homologous receptors, each able to bind a sub-set of a family of Insulin-like ligands that includes Insulin, and Insulin-like Growth factors 1 and 2 (IGF-1 and IGF-2). The IR is primarily involved in controlling metabolic processes, particularly glucose uptake, while the IGF-IR predominantly signals to increase proliferation and protect cells from apoptosis.

An important step in the development of therapeutic agents is the mechanistic and molecular understanding of the relevant protein-protein interactions in the target system. In the IR and IGF-IR fields the Holy Grail has been to determine the structure of one of the ligands bound to either receptor. Due to the large size of a receptor/ligand complex, X-ray crystallography is the only method currently available to study this problem at atomic resolution. Unfortunately both the IR and IGF-IR have not been amenable to the formation of crystals of the required quality.

The Insulin receptor family of proteins consist of eleven structural domains (See Figure 1.2), of which the extracellular domains contain all the ligand binding and specificity determinants. The N-terminal region consists of two homologous domains, L1 and L2, separated by a cysteine rich (CR) domain (Bajaj *et al.*, 1987). The remainder of the extracellular domains were found to consist of three fibronectin type 3 (Fn3) domains with the central Fn3 domain interrupted by a ~120-130 Insert domain (ID) (Marino-Buslje *et al.*, 1998) (Mulhern *et al.*, 1998) (Ward *et al.*, 1999), which also contains the cleavage site between the α and β subunits. A transmembrane domain follows the final Fn3 domain marking the end of the extracellular domains of the receptor.

The 3D structure of the first three domains (L1/CR/L2) of the IGF-IR has been solved using X-ray crystallography (Garrett *et al.*, 1998). It confirms the existence of three discreet domains as predicted by earlier sequence analysis. As shown in figure 1.3, the overall structure is bilobal. The two L domains form a rectangular structure consisting of three beta sheet sides with each end capped by short alpha helices and a disulphide bond. Unfortunately, this fragment is unable to bind ligands as a C-terminal myc tag used in the purification was found to sit in the gap between the lobes and therefore block the likely sites of interaction. In addition you require a 12 amino acid insert peptide from the insert domain that has been shown to contribute to high affinity binding (Kristensen *et al.*, 1999). The relative orientation

of both L domains to the CR domain is probably incorrect in the crystal structure and likely requires the Fn3 domains and the insert domain to form an appropriate scaffold.

Fn3 domains are a commonly occurring scaffold domain, and although a structure has not been produced of a Fn3 domain from either the IR or the IGF-IR, several groups have used comparative modelling to predict the structures based on the relatively high sequence similarity with other known Fn3 domains (Marino-Buslje, 1998). These domains are predicted to form a seven stranded beta sandwich. The relative orientation of the three domains with each other is unknown, as is their orientation relative to the insert domain and the first three domains. Currently it is thought that their main function in extracellular proteins is to act as a spacer to correctly position the other functionally important domains.

There has been no evidence to date to suggest that the intracellular region has a direct role in ligand binding, but it appears to be required for a high affinity conformation, presumably by altering the orientation of the extracellular domains. This has been shown by the fact that a receptor truncated at the transmembrane domain has several-fold lower affinity for ligands (Jansson *et al.,* 1997) and that the addition of dimerisation domains (i.e. a mouse IgG Fc region or a yeast leucine zipper) at the C-terminus of a truncated receptor nearly restores full-length receptor affinity (Hoyne *et al.,* 2000). In order to reduce the size of a ligand binding receptor fragment even further, many studies have focused on minimizing the receptor size such that it contains only those domains and/or sequences necessary for wild-type binding affinity, and at the same time, hoping to produce a protein able to properly crystallize. However, suitable crystals for collecting x-ray diffraction data, from recombinant receptors able to bind ligand, have not yet been reported.

The insert domain is the least understood of all the IR/IGF-IR domains, as sequence analysis shows it is not similar to any other protein sequence. It also has the least amount of sequence similarity between any of the major hIR and hIGF-1R domains (35%), suggesting that there may be a significant number of loosely structured regions in this domain. A study in which bound insulin was cross-linked to the IR, followed by protease digestion of the complex and peptide sequencing, found that a 12 amino acid stretch in the insert domain was closely linked to insulin sequences (Kurose *et al.,* 1994). Interestingly, these 12 amino acids lie directly before the furin cleavage site, and represent the only well conserved region in the entire insert domain. In support of the previous study, alanine scanning mutagenesis of these residues produced receptors with reduced affinity for their ligands, and for one residue in particular (Phe 701) completely abolished ligand binding (Mynarcik *et al.*, 1997). This data suggests

that at least this part of the insert domain is involved in forming part of the ligand binding pocket.

3.1.1 Project Summary and Aims

The insert domain has an important role in the structure and function of the IR and IGF-IR. It is directly involved in ligand binding, it is the site for protease cleavage during receptor biosynthesis, and it contains a triplet of cysteine's from which a single inter-subunit disulphide bond forms. The general aim of this project was to elucidate the 3D structure of the Insert Domain using NMR spectroscopy and determine how its structure enables its functional role.

AIMS:

a) Purify and characterise recombinant insert domain Protein

b) Determine the structure of the insert domain by NMR spectroscopy

3.2 RESULTS

3.2.1 Construction of recombinant insert domain expressing vectors

In the first instance the insert domain was expressed as a fusion protein with bacterial thioredoxin. In many cases a thioredoxin fusion partner has been used to successfully express proteins that contain disulphide bonds (Carrick *et al.*, 2001). Since the insert domain is known to have several inter-domain disulphide bonds present between two insert domains, it was thought that if the insert domain dimerised naturally that the thioredoxin may be able to stabilize its structure or at least prevent inclusion body formation. The pET32 vector system (Novagen) is a commonly and successfully used bacterial expression system that allows the expression of a protein of interest as a C-terminal fusion partner of thioredoxin. The linker peptide region between the thioredoxin and the inserted protein contains a thrombin cleavage site and His-Tag to enable release and purification respectively.

Since the insert domain is cleaved at a furin cleavage site almost in the centre of the domain, there are at least two ways in which the domain could be processed. The cleavage may occur before any secondary structure has formed and therefore the two halves form independent structures in separate parts of the tertiary structure of the receptor. Alternatively, if the cleavage occurs after the domain forms, this may lead to a local change in conformation to allow the critical 12 amino acid C-terminus of the alpha subunit to become part of the ligand binding pocket. This presents a problem when it comes to deciding how best to attempt to get a structure of this domain. To cover both eventualities it was decided to construct both a single whole Insert Domain protein in addition to a double domain protein consisting of the interrupted Fn3αβ domain and the insert domain. In the latter case it was reasoned that the Fn3αβ domain would be able to fold correctly and bring the two halves of the domain together. This should allow the insert domain to fold by itself or, if required, use the Fn3αβ domain as a scaffold.

Primers were designed to amplify the whole insert domain, containing an additional 5' *BamH1*, a 3' *EcoRI* restriction site and a 3' translation stop codon, to allow insertion into the pET32a bacterial expression vector multiple cloning site. Insertion was designed to be in frame with the N-terminal fusion protein thioredoxin (See Figure 3.1). A further two primers were designed to amplify Fn3αβ/ID in an analogous way to the insert domain alone (See Figure 3.1). PCR amplification using *Pfu* polymerase from the plasmid pKS+ containing IGF-IR cDNA using each primer set yielded two products of the expected sizes (See Figure 3.2). Products were gel purified, digested with *BamH1* and *EcoRI*, and ligated into similarly cut

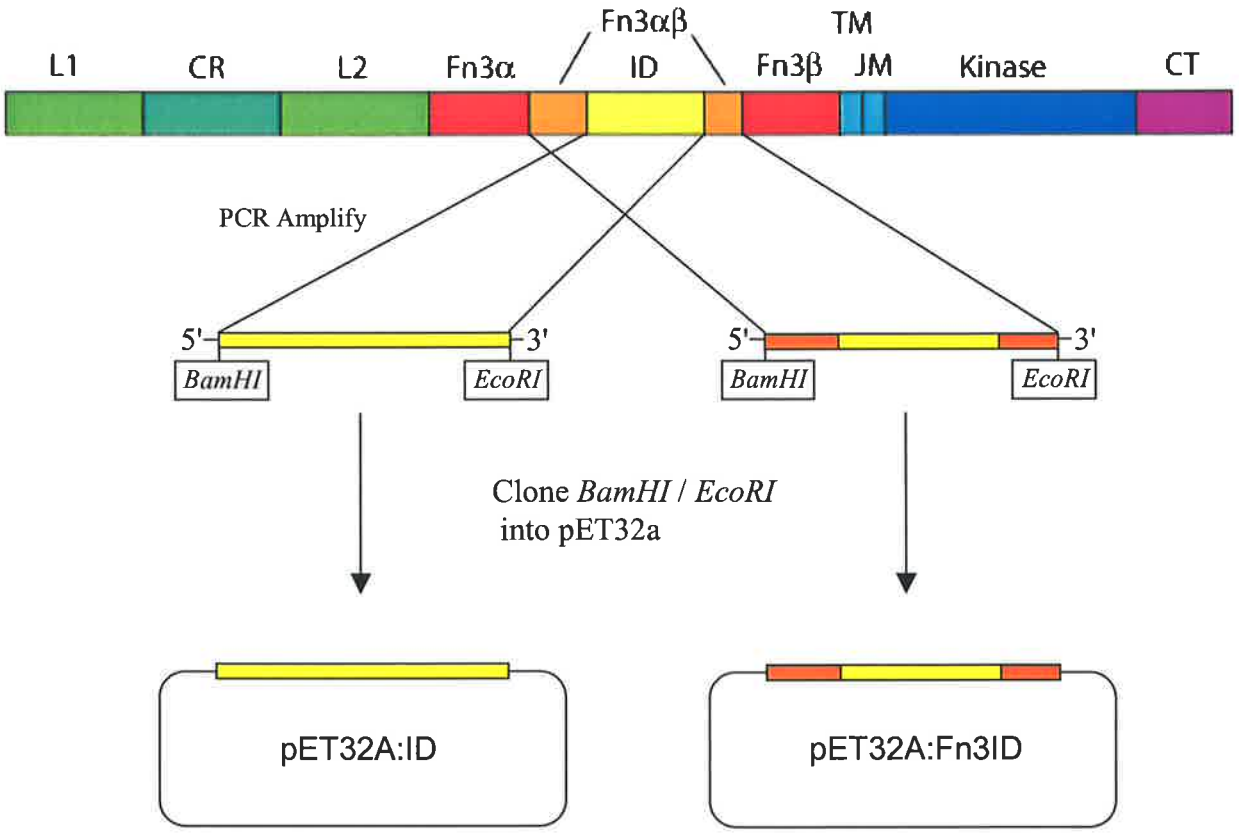


Figure 3.1 **Schematic representation of the cloning strategy used to make Insert Domain Constructs.**

Primers were designed incorporating a 5′ BamH1 and 3′ Stop codon and EcoR1 sites to amplify either the Insert Domain (ID) or both the Fn3αβ Domain and the Insert Domain (Fn3ID). PCR products were subsequently cloned in frame into the MCS of pET32a bacterial expression through BamH1/EcoR1 restriction digests.

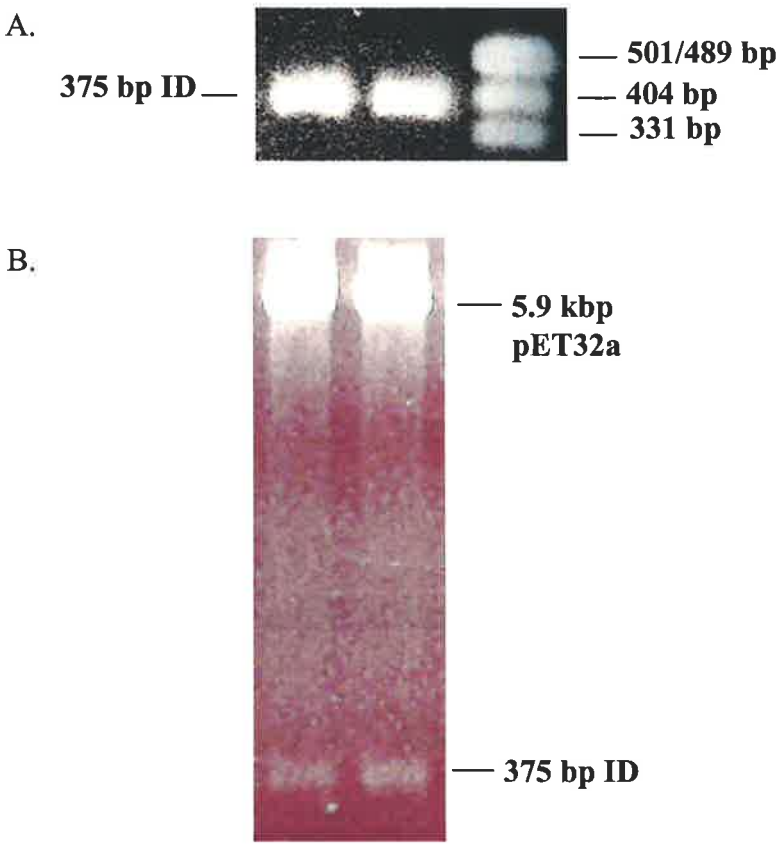


Figure 3.2 PCR amplification of ID and cloning into pET32a

A. Insert Domain was amplified from pKS+:IGF-IR. PCR products were run on 2% agarose gel and bands visualised by ethidium bromide staining.

B. The PCR product was digested with BamH1 and EcoR1 and ligated into similarly cut pET32a vector. Transformants were digested with BamHI and EcoR1 to drop insert, run on a 1% agarose gel, and visualised by ethidium bromide staining. Clones with the correct sized insert were subsequently sequenced for confirmation.

pET32a vector. Transformants containing the correct insert were selected on the basis of a further *BamHI* and *EcoRI* restriction digest dropping the right size DNA fragment (See Figure 3.2) and finally sequence verified to check for any random mutations from the PCR or cloning procedures. The results of the sequencing indicated that there were no changes from the wild-type receptor. Verified clones were finally transformed into BL21 (DE3) or ADA294 bacterial cells for induction of protein expression.

3.2.2 PROTEIN EXPRESSION

3.2.2.1 Expression of Thioredoxin-Insert Domain fusion protein

Initially, the solubility of the Thioredoxin-Insert Domain (trx-ID) fusion protein expressed at 37°C was assessed. A protein expressed as an insoluble inclusion body requires solubilization and refolding which may be difficult or impossible to do. The production of a foreign soluble protein indicates that the bacterium can accommodate its presence in high concentrations and in many cases has been found to at least have some secondary structure.

For trial expression experiments pET32a:ID transformed BL21 (DE3) or ADA294 cells (cell line that allows disulphide bonds to form more easily in the bacterial cytosol) were grown overnight, then inoculated 1/100 into 50 mL LB + ampicillin cultures and grown to OD_{600} = 0.6, and expression induced with 0.2mM IPTG. The amount of fusion protein produced in a 3 hr time course was assessed by PAGE analysis and, was maximal at the 3 hr time point (see figure 3.3). There appeared to be some fusion protein present in the 0 time point indicative of some leakiness from the theoretically repressed lacI promoter even before induction. The next step was to determine whether the fusion protein being produced was soluble, insoluble or a mixture of both. Bacteria induced for 3 hrs were sonicated and centrifuged to separate the insoluble cell debris from the soluble cell contents. A 12.5% polyacrylamide gel was used to analyse bacterial samples from before induction and after induction, in addition to the insoluble cell pellet and supernatant for expression of the trx-ID fusion protein. The results indicated that ~50% of the fusion protein was expressed in the insoluble fraction from the sonication (See Figure 3.4), probably transported into inclusion bodies immediately after expression.

The observation that 50% of the fusion protein was soluble at 37°C suggested that the production of the fusion protein was compromised under maximum bacterial proliferative conditions. Therefore, a trial expression was repeated at the lower temperate of 30°C, allowing the bacteria more time between division cycles, and therefore more time for the

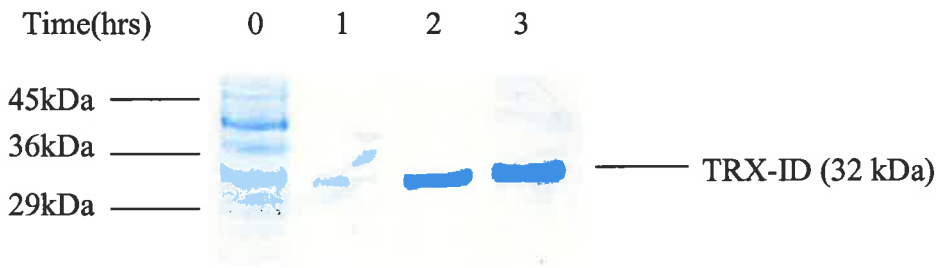


Figure 3.3 Induction Timecourse of TRX-ID

pET32a:ID transformed into BL21 *E.coli* were grown overnight and following 1/100 dilution into fresh media grown to OD_{600} 0.4 before induction with 0.2mM IPTG for up to 3hrs. Bacterial samples were collected before and after induction and run on a 12.5% polyacrylamide gel and protein stained with coomassie blue.

Figure 3.4 Comparison of 37 and 30 degree inductions of trx-ID

pET32a:ID transformed into BL21 E.coli were grown overnight and following 1/100 dilution into fresh media grown to OD_{600} 0.4 before induction with 0.2mM IPTG for 3 hrs. Bacterial samples were collected before and after induction in addition to sonicating the remaining bacteria and separating the soluble and insoluble fractions by centrifugation. The insoluble cell pellet was resuspended in the same volume as the supernatant to allow relative amounts to be observed. Samples were run on a 12.5% Tris-tricine polyacrylamide gel stained with coomassie blue protein stain. BI = BL21 before induction. AI = BL21 after induction. CP = Insoluble cell pellet after centrifugation. SN = Soluble supernatant after centrifugation

fusion protein to be processed. In this experiment, after the overnight incubation, cultures were either grown at 37°C or 30°C. The results shown in a 12.5% polyacrylamide gel, comparing two separate cultures of BL21 transformed with pET32a:ID induced at the two temperatures (see figure 3.4). When grown at 30°C the amount of fusion protein in the soluble fraction after sonication increases from ~40% up to ~90%. It was decided that a further drop in temperature to produce 100% soluble fusion protein was unnecessary. An additional trial was performed at 37°C in pET32a:ID transformed ADA294 cells to investigate whether the increased capacity of this cell line to handle disulphide bond formation would allow soluble fusion protein production, however, the results proved negative. The trials outcomes are summarised below (See Table 3.1). As only one set of conditions resulted in soluble protein production, all subsequent inductions of trx-ID were performed at 30°C in BL21 (DE3) cells.

Table 3.1 Summary of trial inductions using pET32a:ID
pET32a:ID was transformed into either BL21 or ADA294 cells and induced at either 37°C or 30°C, for 3 hrs. Samples were taken during the induction and run on a denaturing polyacrylamide gel and subsequently stained with coomassie blue to detect protein.

Domain	Vector	Cell line	Induction temperature	Majority of Soluble fusion protein
ID	pET32a	BL21	37°C	No
ID	pET32a	BL21	30°C	Yes
ID	pET32a	ADA294	37°C	No

3.2.2.2 Expression of Thioredoxin-Fibronectin 3-Insert Domain fusion protein

Since the trx-ID fusion protein was only fully soluble when expressed at 30°C, trial expressions at both 37°C and 30°C were performed with the Thioredoxin-Fibronectin 3-Insert Domain (trx-Fn3/ID fusion protein expression. Experiments were performed in the same way as in 3.3.2.1 above. Nearly 100% of the trx-Fn3/ID fusion protein was found in the insoluble pellet at both temperatures indicating that this protein was likely being shunted into inclusion bodies. The trial expression in ADA294 cells yielded the same result of insoluble fusion protein. In an effort to improve the ability of the cells to accommodate the fusion protein an additional plasmid constitutively expressing either the GroEL/ES (pACYC:GroE (5212bp)) or dnaK (poB45:dnaK(9544bp)) was co-transformed with pET32a:Fn3/ID (Lee *et al.*, 1991). Further trial expressions yielded the same result of insoluble fusion protein regardless of which cell line was used or what was co-expressed. All trial expression results are summarised below (see table 3.2).

Table 3.2 Summary of trial inductions using pET32a:Fn3/ID or pGEX4T-2:Fn3/ID

pET32a:ID was transformed into either BL21 or ADA294 cells. Plasmids expressing either GroEL/ES or dnaK were also co-transformed as indicated. Inductions were performed at either 37°C or 30°C, for 3 hrs. Samples were taken during the induction and run on a denaturing polyacrylamide gel and subsequently stained with coomassie blue to detect protein.

Domain	Vector	Co-transformant	Cell line	Induction temperature	Soluble fusion protein
Fn3/ID	pET32a	-	BL21	37°C	No
Fn3/ID	pET32a	-	BL21	30°C	No
Fn3/ID	pGEX4T-2	-	BL21	37°C	No
Fn3/ID	pGEX4T-2	-	BL21	30°C	No
Fn3/ID	pET32a	-	ADA294	37°C	No
Fn3/ID	pET32a	GroEL/ES	BL21	37°C	No
Fn3/ID	pET32a	dnaK	BL21	37°C	No
Fn3/ID	pET32a	GroEL/ES	ADA294	37°C	No
Fn3/ID	pET32a	dnaK	ADA294	37°C	No

Given the inability to obtain soluble fusion protein, there was no other option other than to proceed by using the insoluble pellet as the fusion protein source. All subsequent protein expressions of TRX-Fn3/ID were performed in BL21 (DE3) at 37°C as there appeared to be no difference between any of the trial expressions. Soluble TRX-Fn3/ID fusion protein was obtained by dissolving the insoluble pellet in solubilization buffer (8M urea, 100mM Tris-HCl pH 8.0, 50mM DTT, 1mM PMSF).

3.2.3 PROTEIN PURIFICATION

3.2.3.1 TRX-ID Purification by Ni-affinity chromatography

The 6-His tag in the linker peptide between thioredoxin and the Insert Domain was utilised as a first step in the purification process by performing Ni-affinity chromatography. This was a convenient first step to take, however Ni-affinity chromatography is known to pull down endogenous proteins that contain poly-histidine tracts in addition to the induced fusion protein, therefore further purification steps were anticipated.

Soluble TRX-ID fusion protein obtained from sonication of BL21 (DE3) cells grown at 30°C was added in binding buffer to the Ni-NTA agarose beads to form a slurry that was incubated for 1 hr at 4°C. Unbound proteins were washed through with several column volumes of binding buffer. A stringent wash step was then performed using 60mM imidazole wash

buffer, followed by an elution step using 1000mM Imidazole. Fractions were collected from each stage and analysed on 12.5% polyacrylamide gel (See Figure 3.5). The gel analysis shows that there was a larger proportion of the fusion protein in the wash rather than the elute fractions. There was very little fusion protein present in either the insoluble cell pellet or the flow through from the soluble fraction after sonication, indicating that the vast majority of the protein was being expressed in a soluble form and was binding well to the Ni-NTA agarose beads. Since the wash fraction contained the largest amount of the fusion protein, subsequent Ni-affinity chromatography performed used a reduced imidazole concentration of 20mM in the wash buffer and 200mM in the elute buffer. Figure 3.5B shows an example of the improvement in the yields from the elution buffer step using the altered imidazole concentrations. As expected, there were still a significant number of additional proteins co-eluting with the fusion protein therefore further purification steps were required.

3.2.3.2 TRX-ID Purification by Resource Q chromatography

As a further purification step ion-exchange chromatography was tested for its ability to separate and purify the trx-ID fusion protein. A Resource Q column (Amersham Biosciences), was equilibrated with 20mM Tris-HCl pH8.0. A 50 μL sample containing fusion protein eluted from the Ni-affinity chromatography column was added to the Resource Q column and initially eluted with a 0-100% gradient of 0.5M NaCl, 20mM Tris-HCl pH8.0 over 20 min. The trace recording showed two major peaks however these were poorly separated. The same experiment was repeated with a 30 minute elution gradient to try and separate the major peaks. This resulted in better separation of the peaks (See figure 3.6). Fractions of 0.5 mL were collected from the peaks, and samples run on a 12.5% polyacrylamide gel to assess protein content. The gel indicated that both major peaks contained mostly fusion protein (See figure 3.6). All fractions analysed contained some higher molecular weight proteins while fractions 6 and 8 each had one additional band ~3kDa smaller than the fusion protein. The purest and most concentrated fraction for the fusion protein (No. 8) correlated with the first third of the second major peak.

3.2.3.3 Thrombin Kinase digestion

In order for the insert domain to be analysed by NMR spectroscopy the thioredoxin fusion partner needed to be removed and separated. A thrombin cleavage site just upstream of the N-terminus of the Insert domain, if utilised left only a Glycine/serine leader sequence at the N-terminus. trx-ID fusion protein partially purified by Ni-affinity chromatography was digested with 1U/mg fusion protein of bovine Thrombin for up to 3 hrs at 37°C to assess the rate and

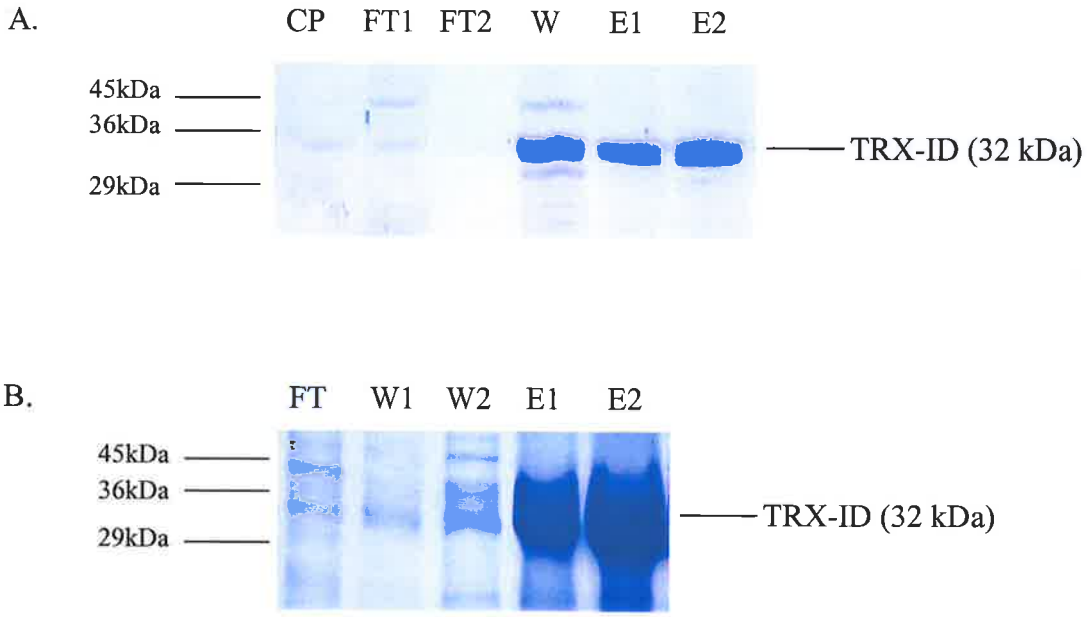


Figure 3.5 Optimisation of wash and elution buffers for Ni-affinity Chromatography

pET32a:ID transformed into BL21 E.coli were grown overnight and following 1/100 dilution into fresh media grown to OD_{600} 0.4 at 30°C before induction with 0.2mM IPTG for 3hrs. Bacteria were collected, sonicated and centrifuged to separate the soluble and insoluble fraction. The soluble fraction was added to Ni-NTA agarose beads and affinity chromatography performed using different concentrations of imidazole in the Wash and Elute buffers. Fractions collected were run on a 12.5% polyacrylamide gel and protein stained with coomassie blue.

A. Wash Buffer (60mM Imidazole), Elute Buffer (1000mM Imidazole), CP = Solubilised cell pellet, FT1 = Flow through 1, FT2 = Flow through 2, W = Wash fraction, E1 + E2 = Elute fractions.

B. Wash Buffer (20mM Imidazole), Elute Buffer (200mM Imidazole), CP = Solubilised cell pellet, FT = Flow through 1, W1 + W2 = Wash fractions, E1 + E2 = Elute fractions.

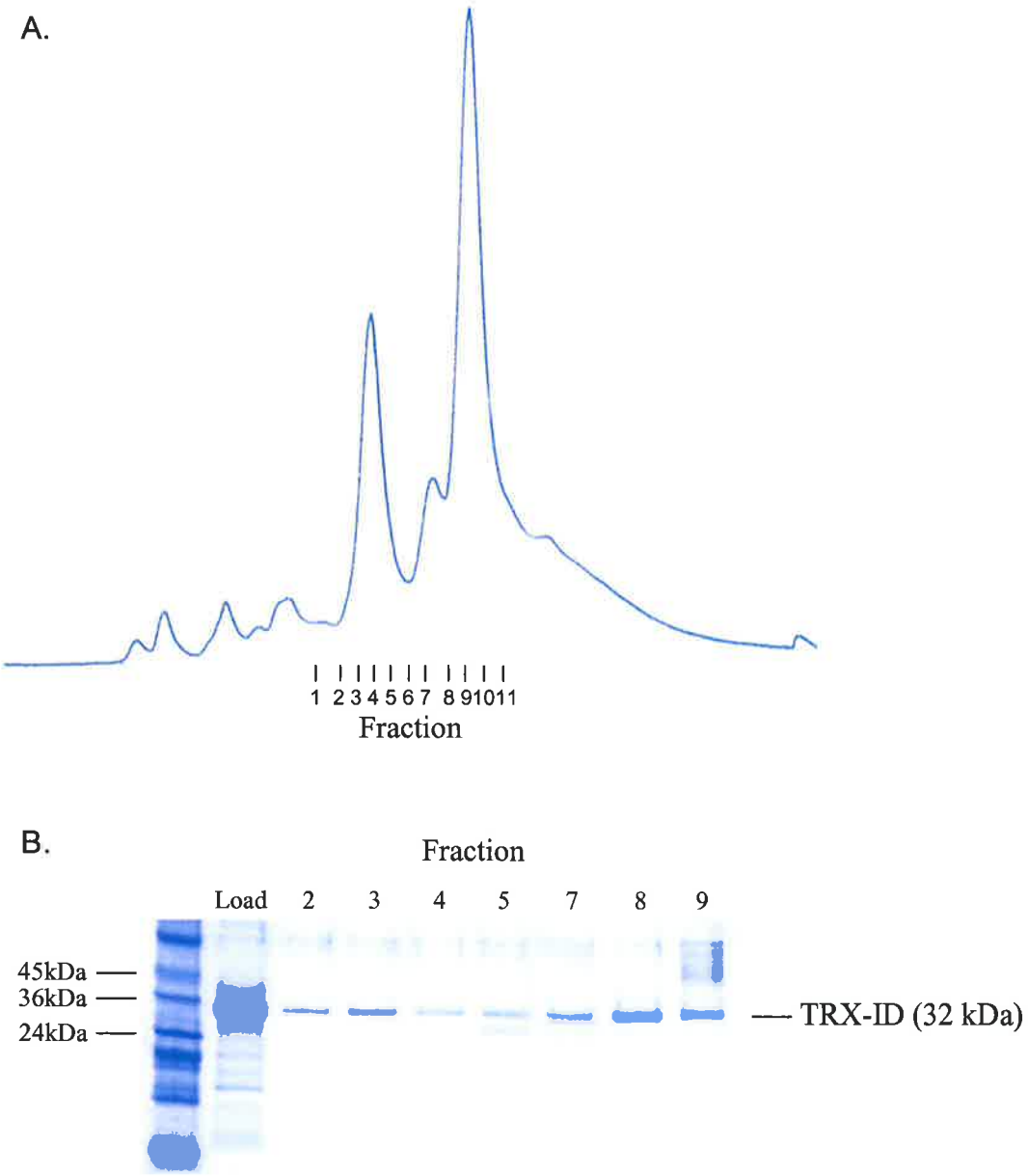


Figure 3.6 Ion-exchange chromatography

A. TRX-ID fusion protein semi-purified by Ni-affinity chromatography was loaded onto a 20mM Tris-HCl pH8.0 equilibrated Resource Q column for ion-exchange chromatography. Fusion protein was eluted with a 0-100% gradient of 0.5M NaCl, 20mM Tris-HCl pH8.0 over 30 min with a flow rate of 0.5 mL/min. The recorded trace of the absorbance at 280nm and the relative positions of fractions collected as the peaks were eluting off is shown.

B. 10 µL of the collected fractions and the initial loaded sample were run on a reduced 12.5% Tris-tricine polyacrylamide gel stained with coomassie blue protein stain.

extent of cleavage. The 12.5% polyacrylamide gel of the results shows that the fusion protein is rapidly cleaved and by 15 min the majority of the protein was running at the apparent molecular weights of ~22kDa, ~18kDa (expected size of ID= 17.4kDa), and ~14kDa (expected size of thioredoxin) (See Figure 3.7). It was decided that subsequent thrombin digests would be performed for at least 4 hrs to minimise the amount of the ~22kDa protein from being produced.

A further Ni-affinity purification was subsequently performed to separate the products of the reaction. The Ni charged column was able to bind to the 6-His tag between the Thioredoxin and the inserted protein and uncleaved full-length fusion protein as the His Tag is located between the thrombin cleavage site and the multiple cloning site.

3.2.3.4 HPLC Purification of ID

Insert domain protein cleaved from the thioredoxin was purified by High Performance Liquid chromatography (HPLC). Fractions containing proteins from the second Ni-affinity chromatography step after the thrombin digestion were compared with the semi-purified fusion protein. Samples were acidified with TFA and analysed on a C18 analytical column with a 20 min 0-80% acetonitrile gradient. The traces in figure 3.8 show that the fusion protein eluted at 28 min while the insert domain protein was eluted at 22 min. The insert domain peak was collected and a sample of the fraction rerun to check for purity (See figure 3.8).

3.2.3.5 NMR of Purified ID

The purified fraction of insert domain was freeze dried and prepared for NMR analysis by dissolving in 90% MQ H_2O/10% D_2O. As a control, the thioredoxin purified from the flow through from the second Ni-affinity chromatography step was buffer exchanged into MQ H_2O to remove the salt from the binding buffer before being freeze dried and prepared in the same way as for the insert domain. A 1D 1H NMR spectrum was recorded to assess the presence of secondary structure in the insert domain. The insert domain trace overall indicated that the protein was likely to be unstructured. The peaks in the amide and aromatic hydrogen region were not well dispersed and the peaks generally were sharp, which is typically indicative of a lack of secondary structure (See Figure 3.9). The insert domain trace was in stark contrast to that of the thioredoxin sample, which appeared to have well-dispersed, sharp peaks.

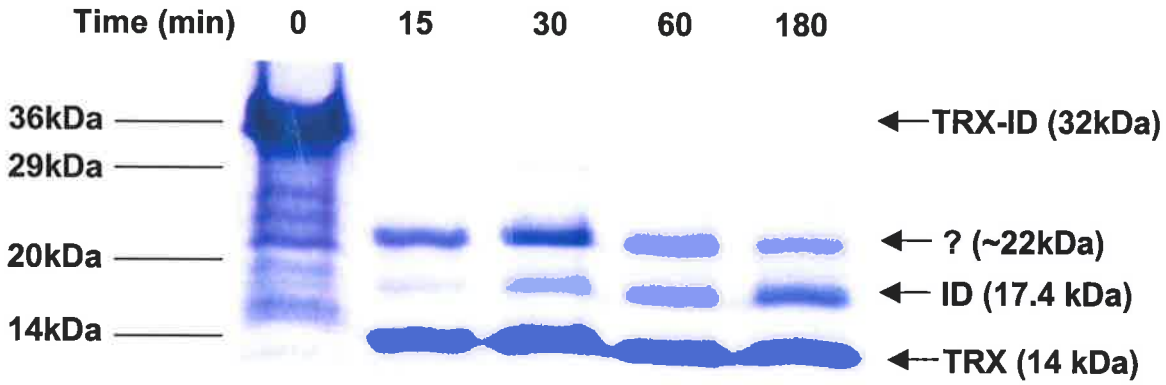


Figure 3.7 Thrombin Digest

TRX-ID fusion protein semi-purified by Ni-affinity chromatography was digested with 1U thrombin/ mg protein for up to 3 hrs at 37°C. Digestions were stopped by the addition of 1mM PMSF. 10 µL samples were run on a 12.5% Tris-tricine polyacrylamide gel stained with coomassie blue protein stain.

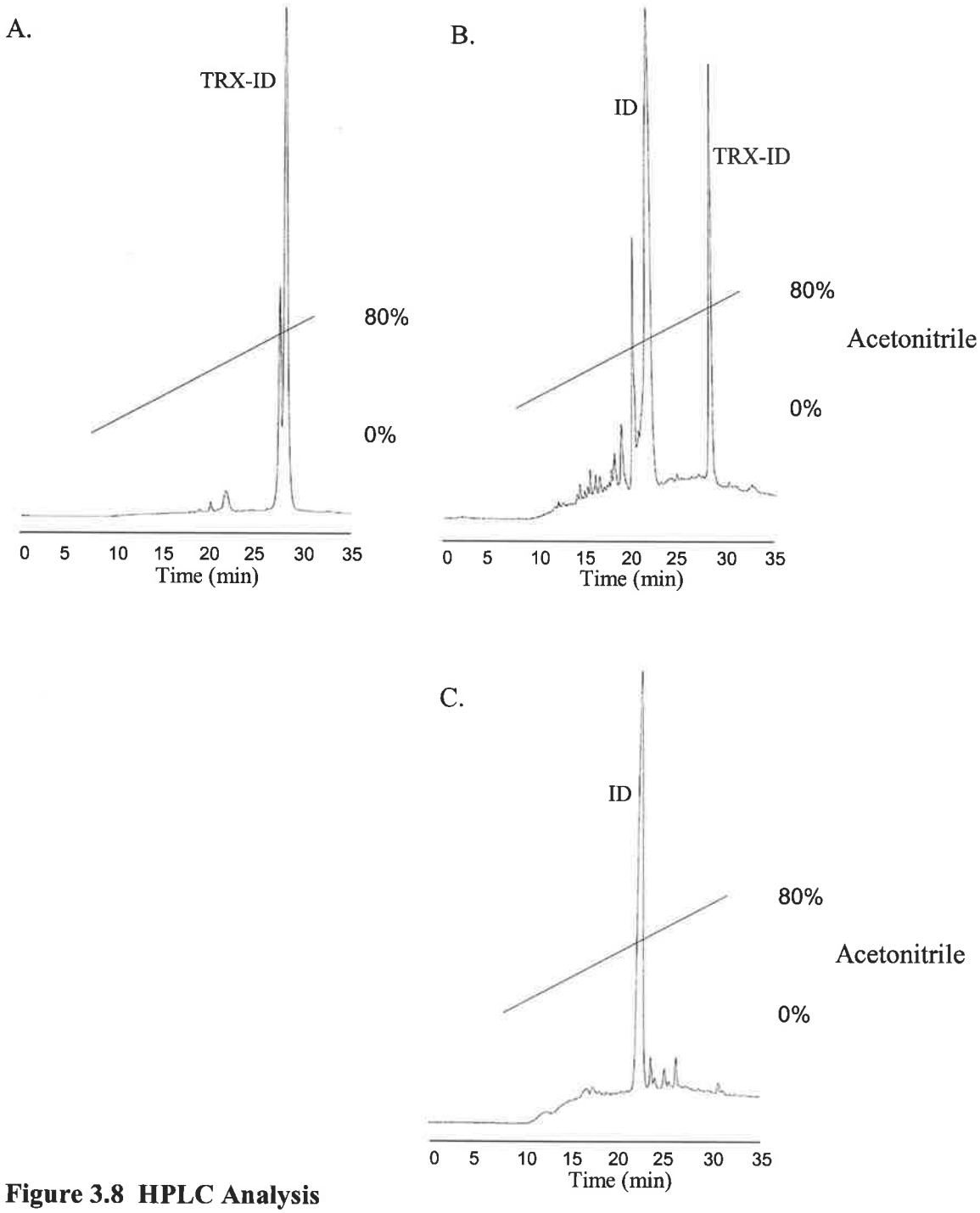


Figure 3.8 HPLC Analysis

Protein samples were acidified with 0.1%-2% (v/v) TFA prior to loading onto a C18 analytical column equilibrated with 0.1% (v/v) TFA. A 0-80% acetonitrile linear gradient using 0.1% (v/v) TFA and 0.088% (v/v) TFA/80% acetonitrile was used to elute peptides from the column at a flow rate of 0.5 mL/min. Traces show the detected absorbance at 215nm.

A. TRX-ID fusion protein from an elution fraction from Ni-affinity chromatography.
B. Semi-purifed ID from second Ni-affinity chromatography purification after thrombin cleavage of the fusion protein.
C. Purifed ID separated by a previous HPLC run.

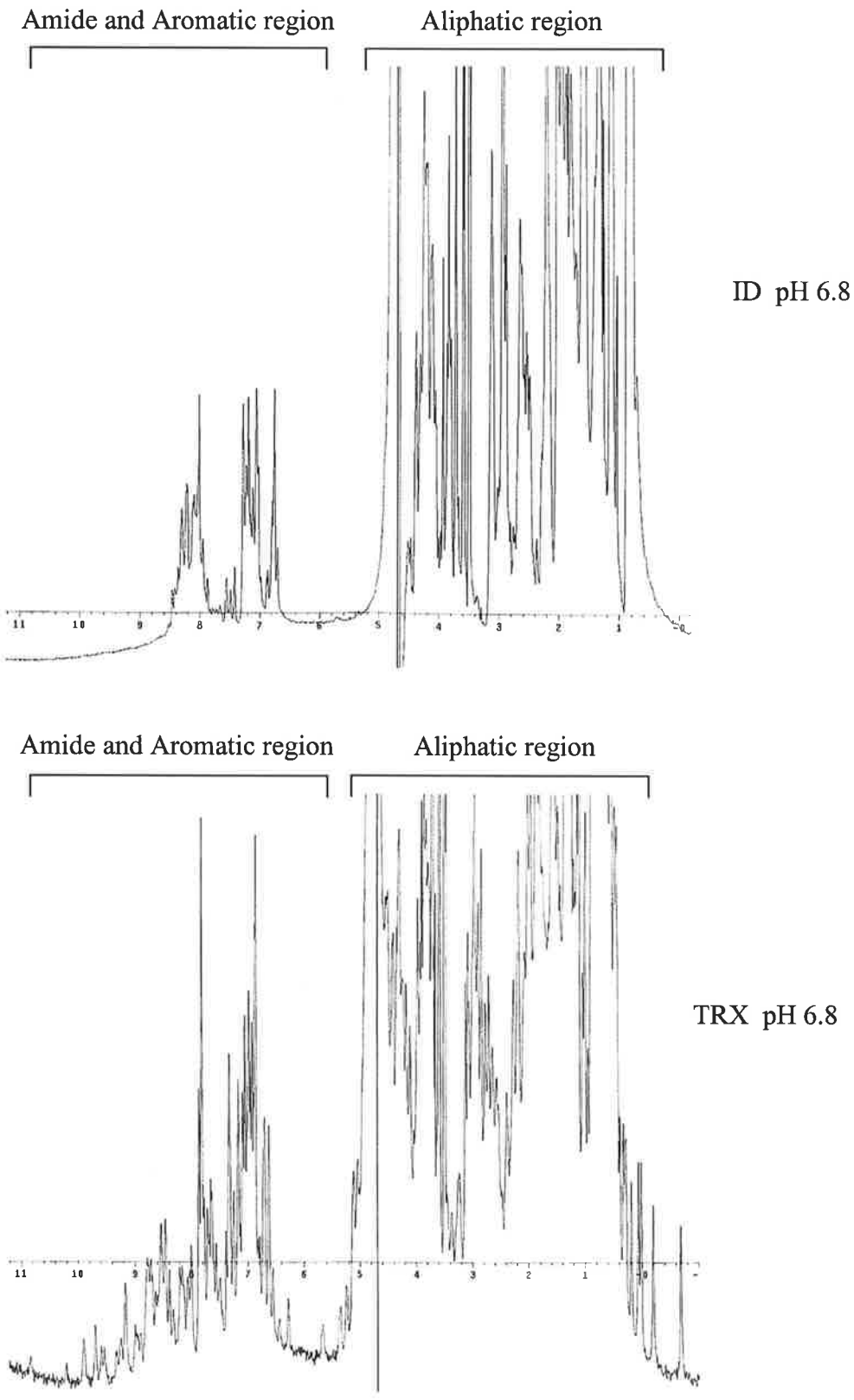


Figure 3.9 Comparison of 1D ^{1}H NMR spectra of Insert Domain and Thioredoxin

1D plot of ^{1}H NMR spectra at pH 6.8 of HPLC purified Insert Domain and Ni-affinity purified TRX. Protein samples were lypholised and dissolved in 90% MQ H_2O/10% D_2O.

3.2.4 REFOLDING STUDIES

3.2.4.1 Analysis of the State of ID by Gel Filtration Chromatography

Since the NMR analysis indicated that the Insert Domain was likely to be mostly unstructured it was necessary to further assess the state of the protein. It was possible that the ID protein was unfolded because of incorrect formation of disulphide bonds within or between two IDs. Samples were analysed by gel filtration using a superdex 75 column in order to assess the presence of multimers. The chromatography trace of untreated ID directly after resuspension indicated the presence of mainly dimers, however after several hours the main peak had shifted to indicate the presence of trimers (See figure 3.10). A control sample of ID in reduced conditions containing 10mM DTT was monomeric, however both the samples of ID incubated with the oxidised and reduced glutathione appeared to contain trimers and larger multimeric complexes (See figure 3.10). In contrast, the protein fraction from the Ni-affinity chromatography assumed to be thioredoxin showed a single peak of a size appropriate for monomeric thioredoxin when analysed by this form of chromatography.

3.2.4.2 Proteinase K Assay

Although gel filtration chromatography was able to identify variant multimeric forms of the ID, an alternative method was required to analyse the extent of structure in the protein. In order to assess the presence of more structured regions, limited proteolysis was performed using proteinase K. This enzyme is a non-specific protease and has long been used as a probe for the presence of compact globular domains. Therefore the susceptibility of ID protein to cleavage by proteinase K was indicative of the level of secondary structure. In order to provide a control for the reaction the fusion protein, trx-ID was used instead of purified ID and as can be seen in figure 3.11, after a 10min digestion the fusion protein was completely cleaved leaving only a protein corresponding to the size of thioredoxin. This confirms that the insert domain portion of the fusion protein is likely to be unstructured or possibly, in an extended, easily accessible conformation, whereas the thioredoxin portion of the fusion protein appeared to be compact.

3.2.4.3 Alteration of Refolding Conditions

As both the NMR results for the purified ID and the susceptibility of both the trx-ID and trx-Fn3/ID fusion proteins to proteinase K digestion, suggested that the insert domain proteins were unstructured, a variety of refolding conditions were tested. Since the conditions favourable for refolding were unknown, a series of different refolding conditions were tested on their ability to alter the results of a proteinase K assay. After researching the literature it was decided that the standard conditions for the refolding buffer should contain a low

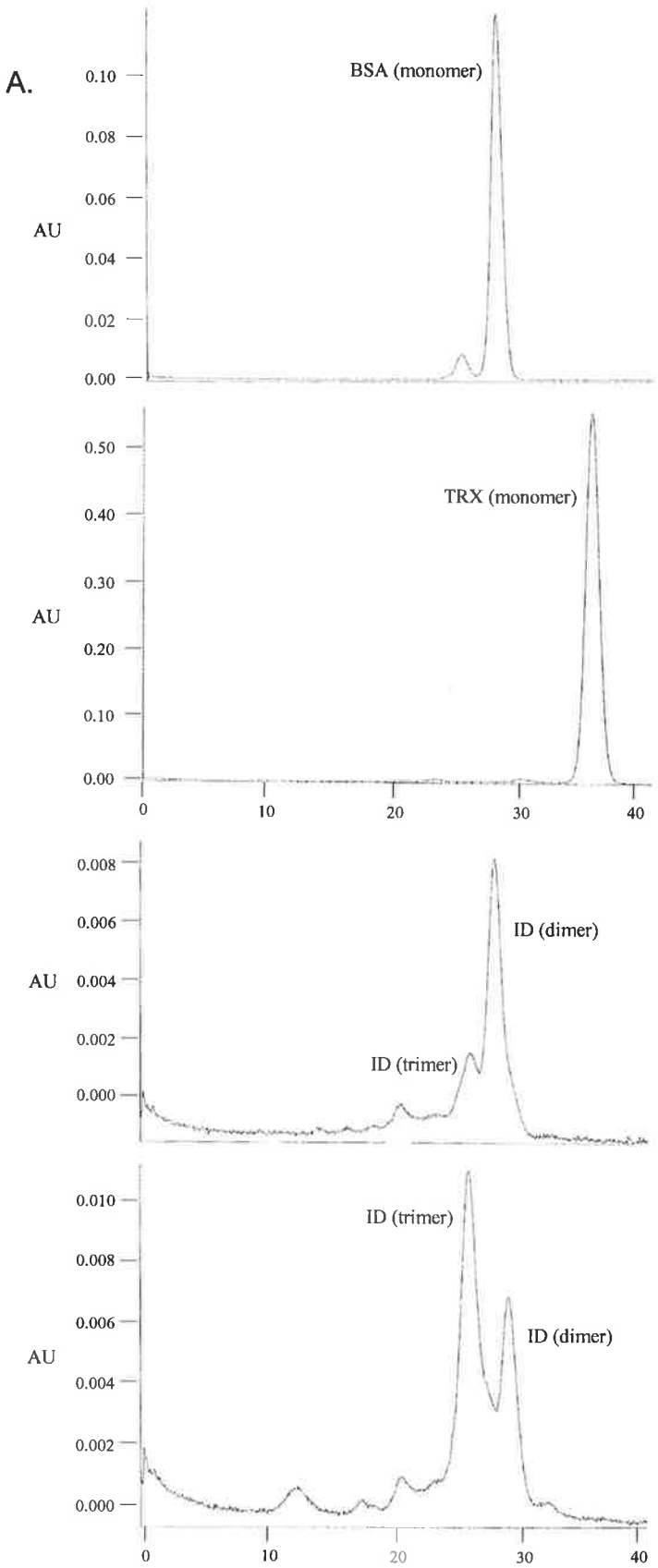


Figure 3.10 Gel Filtration Chromatography

HPLC Purified ID, TRX or BSA protein were loaded onto a 20mM Tris-HCl pH8.0 equilibrated Resource Q column for ion-exchange chromatography. Protein was eluted with a 0-100% gradient of 0.5M NaCl, 20mM Tris-HCl pH8.0 over 30 min with a flow rate of 0.5 mL/min. The recorded trace of the absorbance at 280nm and the relative positions of fractions collected as the peaks were eluting off.

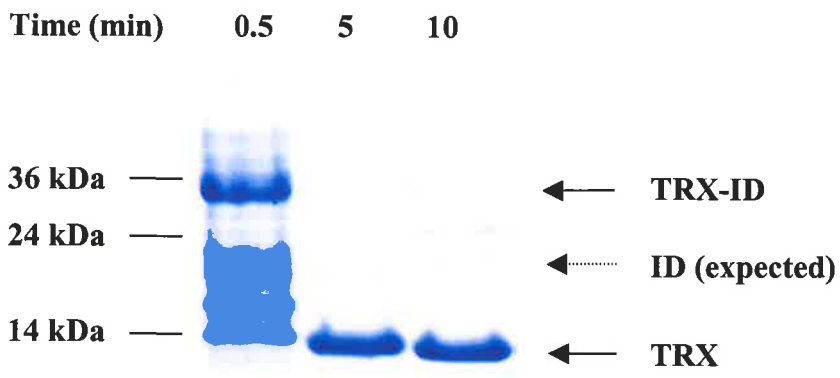


Figure 3.11 Limited proteolysis of TRX-ID by Proteinase K

TRX-ID fusion protein semi-purified by Ni-affinity chromatography was digested with 5U Proteinase K/ mg protein for up to 10 min at 37°C. Digestions were stopped by the addition of 1mM PMSF. 10 μL samples were run on a 12.5% Tris-tricine polyacrylamide gel stained with coomassie blue protein stain.

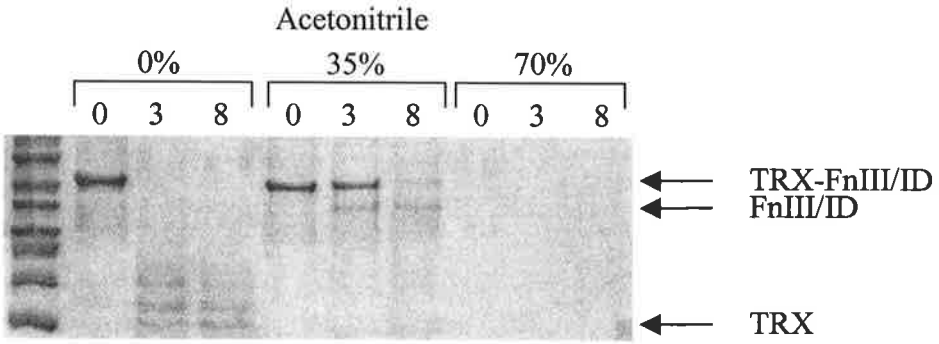
concentration of denaturant (2M urea) in order to reduce the likelihood of unfavourable folding states, 0.25M NaCl and 100mM Tris HCl (a standard salt concentration), at a pH of 8.0, and the samples should be incubated the sample at 4°C to reduce protein breakdown. Either trx-ID or trx-Fn3/ID at a 1 mg/mL concentration was added to 1 mL of refolding buffer and incubated for 48 hrs at 4°C, prior to proteinase K digestion on 200 µL fractions, and analysis of 15µL samples of PMSF stopped digestions on a 12.5% polyacrylamide gel (See figure 3.12 and table 3.3).

Protein	Conditions	Result
TRX-ID	2M urea, 0.25M NaCl, 100mM Tris HCl pH 8.0, Incubate 4°C	Rapid digestion of fusion protein leaving mostly only TRX
TRX-Fn3/ID	2M urea, 0.25M NaCl, 100mM Tris HCl pH 8.0, Incubate 4°C	Rapid digestion of fusion protein leaving mostly only TRX
TRX-ID	+ 25°C incubation	Enhanced degradation
TRX-ID	+ 37°C incubation	Enhanced degradation
TRX-Fn3/ID	+ 25°C incubation	Enhanced degradation
TRX-Fn3/ID	+ 37°C incubation	Enhanced degradation
TRX-ID	+ pH 5.0	No change from initial conditions
TRX-ID	+ pH 11.0	No change from initial conditions
TRX-Fn3/ID	+ pH 5.0	No change from initial conditions
TRX-Fn3/ID	+ pH 11.0	No change from initial conditions
TRX-ID	+ 0.5 M NaCl	No change from initial conditions
TRX-ID	+ 1 M NaCl	No change from initial conditions
TRX-Fn3/ID	+ 0.5 M NaCl	No change from initial conditions
TRX-Fn3/ID	+ 1 M NaCl	No change from initial conditions
TRX-ID	+ 10% Glycerol	Slight delay in cleavage
TRX-ID	+ 30% Glycerol	Slight delay in cleavage
TRX-Fn3/ID	+ 10% Glycerol	Slight delay in cleavage
TRX-Fn3/ID	+ 30% Glycerol	Slight delay in cleavage and partial stability of Fn3/ID product
TRX-ID	+ 35% Acetonitrile	Slight delay in cleavage
TRX-Fn3/ID	+ 35% Acetonitrile	Slight delay in cleavage and partial stability of Fn3/ID product

Table 3.3 Summary of the results of the alterations to the re-folding conditions

TRX-ID or TRX-Fn3/ID was mixed with various refolding buffers for 48 hrs after which proteinase K digestion was performed. Samples were run on a denaturing polyacrylamide gel and protein stained with coomassie blue.

A.



B.

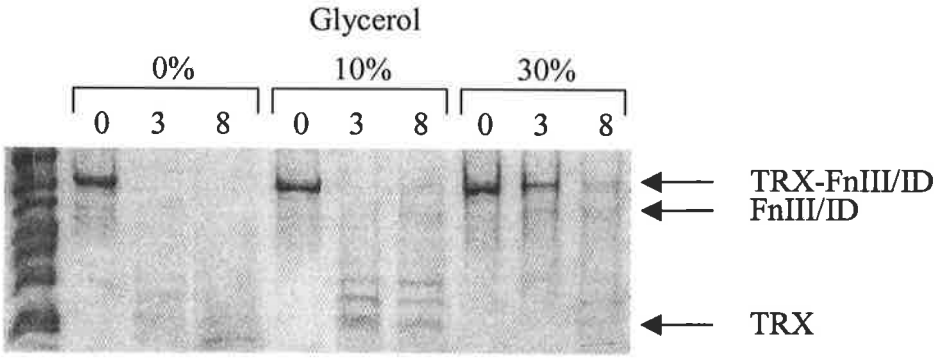


Figure 3.12 Effect of altered conditions on the Limited proteolysis of TRX-Fn3/ID by Proteinase K

Solubilised TRX-Fn3/ID was added to refolding buffer containing 2M urea, 0.25M NaCl, 100mM Tris HCl, pH of 8.0, and either acetonitrile or glycerol and incubated overnight at 4°C. Samples were subsequently digested with 5U Proteinase K/ mg protein for up to 8 min at 37°C. Digestions were stopped by the addition of 1mM PMSF. 10 μL samples were run on a 12.5% Tris-tricine polyacrylamide gel stained with coomassie blue protein stain.

A. Acetonitrile
B. Glycerol

3.3 DISCUSSION

The aim of this project was ultimately to determine the structure of the insert domain of the IGF-IR, either in isolation or together with the Fn3 within which it is inserted. Due to the inability to obtain a fully folded protein sample this was unable to be attempted. Both insert domain fusion proteins, when expressed at 37°C, were insoluble suggesting the bacterial cells have difficulty in folding these proteins to allow it to remain in a soluble form. Since both the NMR and limited proteolysis experiments both indicated that the thioredoxin portion of the fusion protein contained some structure it was clear that the Insert domain and the Fn3 domains were the reason for the problem. Since changing the expression cell line from BL21 to ADA294 had no effect on the amount of soluble protein produced, the potential presence of disulphides may not be the major problem. In addition it appears that it is also not the overextending of the bacterial refolding apparatus, as overexpression of either GroEL/ES, or DnaK had no effect. This suggests, that if able to be folded in a mammalian system, the Insert Domain requires mammalian folding apparatus.

The partial purification by Ni-affinity chromatography, successfully enriched the collected samples with fusion protein, although there was a significant amount of additional proteins collected by this technique. Ion-exchange chromatography was unable to completely separate the fusion protein from some other high molecular weight proteins. It is also unclear why the fusion protein was eluting in two separate peaks that appeared to be the same protein, unless one of the peaks represented an alternatively structured version that altered the number of accessible charged residues. There is a possibility that the two peaks could represent monomeric and dimeric forms which appear the same on a reduced gel.

The thrombin kinase digestion of trx-ID although resulting in the production of two proteins of the expected sizes, 17.8kDa (ID+linker) and 11.8 kDa (thioredoxin), also made an intermediate sized protein of ~22kDa. It is unclear as to what the ~22 kDa band represents as analysis of the full length fusion protein sequence did not reveal any obvious cleavage sites that would produce a peptide of this size. As the cleavage continued there was a decrease in the amount of the ~22kDa band and a concurrent increase in the ~18kDa band suggesting that the ~22kDa band contained the ID sequence. At this time mass spectral analysis was not available to confirm this.

1D ^{1}H NMR analysis of the HPLC purified ID and limited proteolysis strongly suggested a lack of secondary structure, compared to the thioredoxin protein. It is possible that the

additional N-terminal linker peptide derived from pET32a MCS could interfere with the correct folding of the ID. It was expected that the ID may form multimers due to its cysteine triplet, and the gel filtration chromatography provides some evidence of this, however this may also be interfering with the ability of the domain to fold. The addition of the Fn3αβ domain surrounding the ID appeared to have no effect in helping either domain form stable secondary or tertiary structures. It therefore seems likely that either, a large portion of the extracellular domains of the receptor is required to allow correct folding, or that the ID in particular is relatively unstructured even in the full-length receptor. Clearly, only 12 aa from the ID are needed to be added to truncated receptor constructs to produce a high affinity receptor.

Individual alterations to the conditions to promote refolding of the ID had mixed results in creating a more stable protein. Raising the incubation temperature from 4°C to 25°C, or 37°C in order to increase the overall energy of the system, only succeeded in causing degradation of the starting material. Changing the pH to 5.0 or 11.0 and increasing the NaCl concentration to 0.5M and 1M all had no effect on the stability of the fusion protein. The addition of either 10% or 30% glycerol resulted in a delay in the rate of cleavage by proteinase K, therefore this may only be inhibiting the function or affecting the structure of the proteinase. In the case of TRX-Fn3/ID, the 30% glycerol test had a much more pronounced delay and a band of the expected size of Fn3/ID appeared. However this still only appeared to be a small proportion of the starting material. Changing the conditions to become more hydrophobic with the addition of 35% acetonitrile had a similar effect to 30% glycerol in that the protease cleavage was delayed for both TRX-ID and TRX-Fn3/ID. Again a Fn3/ID band appeared but a similar band in the same TRX-ID experiment did not. However unlike the 30% glycerol condition there were no lower molecular weight bands indicating that the TRX had been degraded. This suggests that in the presence of more hydrophobic conditions the structure of TRX is destabilised, whereas for Fn3/ID there may be some secondary structure forming allowing it some protection from the proteinase K cleavage. This experiment suggests that simulating an environment with a reduced water exposure increases the stability of the Fn3/ID protein. Taken together the glycerol and acetonitrile experiments also suggest that the Fn3 domain is required for the stabilising effect to occur. Since it is unclear exactly where the insert domain is located within the overall structure of the receptor this evidence points to the domain being buried somewhere within the receptor.

It is clear that the insert domain was not able to be folded correctly by production from a bacterial system under the conditions presented here. A next logical step may have been to

move the construct into a mammalian expression vector although NMR spectroscopic analysis would require expensive protein labelling procedures. However, the expression from a construct of the Insert domain and the three adjacent Fn3 domains in a mammalian cell line, unlike in bacteria produced a soluble protein however, based on the protein sizes on a reducing polyacrylamide gel, it was processed in to two separate units due to furin cleavage within the insert domain (Personal communication from Peter Hoyne, CSIRO Melbourne). This protein was also used in a crystal formation trial however proving unsuitable for this process. The refolding experiments suggest that the insert domain becomes more stable in a hydrophobic environment suggesting that it is probably buried within the receptor and without the presence of the domains to mask it from water, it is unlikely to be able to be folded properly. Producing an even larger fragment more likely to adopt a correct structure was not feasible since the available NMR spectrometer is unable to be used to analyse proteins >25-35kDa.

The overall conclusions of this project were that the insert domain is likely to have limited structure, is buried within the receptor, and requires the presence of the rest of the extracellular domains to adopt its correct structure. It seems we will have to wait for the crystal structure of the entire receptor to resolve the molecular basis for the involvement of the insert domain in binding ligand.

CHAPTER 4

Cross-Reactivity Of Growth Factors In The Insulin System

CHAPTER 4 CROSS-REACTIVITY OF GROWTH FACTORS IN THE INSULIN SYSTEM

4.1 INTRODUCTION

Human insulin receptor is encoded by a single gene located on chromosome 19, composing 22 exons. Human insulin receptor may be expressed as either of two isoforms, A and B, resulting from the alternative splicing of the primary transcript to either exclude or include exon 11 respectively. The A isoform lacking exon 11 appears to be expressed ubiquitously and the B isoform containing exon 11 is co-expressed mainly in the traditional insulin responsive tissues of liver, muscle, adipocytes and kidney, where it is the dominant isoform (Seino and Bell, 1989). Even though the hIR-A isoform is more widely and in many cases more highly expressed overall, studies into the biology of the insulin receptor have focused almost entirely on the hIR-B form, due to it being the major isoform expressed in the classical insulin responsive cells.

Initial investigations found that IR-A bound insulin ~1.5 fold better than IR-B (Yamaguchi *et al.,* 1993) but that IR-B was more efficient at auto-phosphorylating itself, recombinant IRS-1 and a peptide substrate poly-Glu_4:Tyr_1 (Kosaki *et al.,* 1995). This fact, together with the expression pattern of hIR-A being in classical insulin non-responsive tissues, led researchers to presume that the IR-A isoform was essentially unimportant to the major functions that insulin was known to control.

Murine knockout studies of members of the insulin/IGF system have suggested that the insulin receptor has an additional role in fetal growth, most likely through IGF-2 (discussed in section 1.2.6). Expression of the hIR-A isoform is also increased in fetal tissue compared with the same adult tissue pointing to this isoform being more important in embryonic development than in the adult. Recent studies identified the hIR-A but not the hIR-B isoform as having a high affinity for IGF-2 and also suggested that the hIR-A isoform was able to give a stronger mitogenic signal when activated by IGF-2 than it normally would by insulin (Frasca *et al.,* 1999). Additional studies into whether the IR-A isoform has a role in cancer, have found that it is upregulated in several different tumour types, together with IGF-2 (See 1.4.2.2), however these cells also appeared mitogenically responsive to insulin as well, so whether IGF-2 is the only relevant ligand is yet to be established.

4.1.1 Project Summary

The hIR-A isoform appears to have a role in promoting cellular growth through binding IGF-2. The overall aims of this project were to evaluate the high affinity binding of hIR-A isoform by IGF-2 and to further investigate the binding by using recombinant forms of each receptor. An assessment of the ability of the growth factors to trigger phosphorylation of the receptor and the subsequent effect of growth factor stimulation on cellular growth and protection from apoptosis was also made.

AIMS:

a) Assess the role of exon 11 in affecting the binding insulin, IGF-1, and IGF-2 using both full-length and recombinant receptors.

b) Analyze the difference in phosphorylation of hIR-A and hIR-B in response to ligand stimulation.

c) Assess the proliferative response of hIR-A and hIR-B to extended ligand stimulation.

d) Test the ability of hIR-A and hIR-B activation to protect cells from drug induced apoptosis.

e) Produce an antibody able to differentiate between hIR-A and hIR-B.

4.2 RESULTS AND DISCUSSION

4.2.1 Construction of Full Length hIR Isoforms in Mammalian Expression Vectors

The full-length hIR-A isoform was constructed from pECE:hIR-B, containing the full-length receptor sequence, and pHIR/P12.1 (Ullrich *et al.*, 1985) containing the entire cytoplasmic domain of the insulin receptor lacking exon11. pECE:hIR-B was digested with AatII to remove a 1.8kb fragment spanning exon 11. The corresponding 1.8kb fragment from AatII cut p12.1 lacking exon 11 was exchanged into the pECE:hIR fragment to make pECE:hIR-A. The full-length receptors were subsequently cloned into the mammalian expression vector pEFIRESneo. This vector expresses a bicistronic message of both the 5' inserted protein and the neomycin resistance gene expressed from an internal ribosome entry site (IRES) under the control of the strong constitutive Elongation Factor 2 promoter (Hobbs *et al.,* 1998). pECE:hIR-A and hIR-B plasmids were restricted with Sal1 and Xba1 to release a 2.9kb fragment containing the insulin receptor and ligated to Xho1/Xba1 cut pEFIRESneo through their compatible overhangs. Transformants containing inserts were identified by EcoRI/XbaI restriction digest and the exon 11 status confirmed by PCR analysis. PCR reactions of parent and constructed plasmids were performed using primers to amplify a region spanning the insert domain of the insulin receptor. The resulting PCR products were of two distinguishable sizes on a 2% agarose gel and the size difference correlates with the presence or absence of exon 11, yielding the predicted fragment sizes of 508bp and 472bp for exon 11+ and exon 11- containing products respectively (See Figure 4.1).

4.2.2 Construction of Recombinant hIR-B Ectodomain with a C-terminal Leucine Zipper

The recombinant hIR-B ectodomain with a C-terminal leucine zipper (hIR-B EDZIP) was constructed from pEE14:hIR-A EDZIP (a kind gift from Colin Ward, CSIRO Melbourne). Due to the presence of numerous AatII sites in pEE14 a small fragment containing the exon11 was unable to be swapped directly in the vector as performed in 4.2.1. Instead the entire recombinant fragment of hIR-A EDZIP was moved into pBluescriptII KS+, through a BamH1/Xba1 restriction digest. An exchange from exon11- to exon11+ was then performed as described in 4.2.1, to make pBluescriptII KS+:hIR-B EDZIP. The exon11 status of the construct was verified by PCR as described previously (data not shown). hIR-B EDZIP fragment was removed from pBluescriptII KS+ using Sal1 and Xba1 and cloned instead into Xho1/Xba1 cut pEFIRESneo (NB. hIR-B EDZIP was attempted to be moved back into pEE14 through BamH1/Xba1 digestion, however was unable to be cloned).

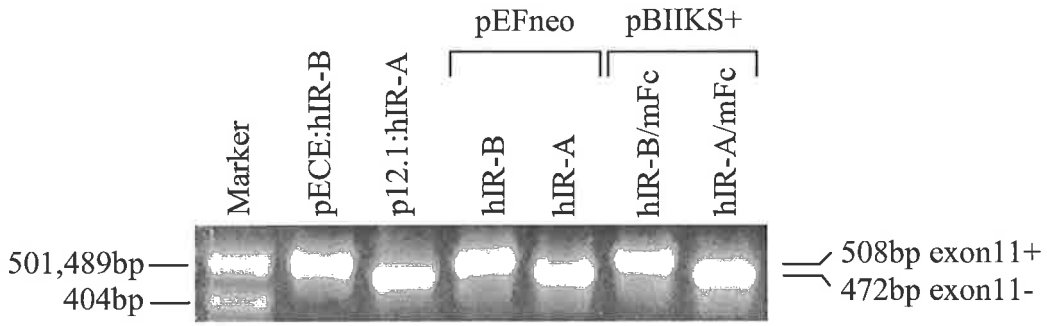


Figure 4.1 PCR analysis of hIR constructs

PCR amplification of a short region spanning exon 11 was performed and products visualized by ethidium bromide staining.

4.2.3 Creation of Cell Lines Expressing Full Length hIR Isoforms

It was decided that stable cell lines expressing the full length receptor isoforms would be made instead of doing transient transfections to avoid having to perform repeated transfections that would yield variable populations of receptor expressing cells. The cell line chosen to study the effects of the hIR isoforms were R- cells (A kind gift from Renato Baserga, Philadelphia), a cell line derived from the embryonic fibroblasts of an IGF-IR KO mouse. Since studies involving IGFs were to be a key part of the project it was thought imperative that there should be no background receptor cross-reactivity that would make it difficult to determine through which receptor system a growth factor was acting. Additionally, R- cells have also been found to express very low levels of insulin receptor (~5 X 10^3 receptors/cell) and are virtually unresponsive to insulin (Sell *et al.,* 1994) and grow slowly even in the presence of serum (See Figure 4.16).

Plasmid constructs containing either hIR-A or hIR-B were prepared using the Ultraclean™ miniprep kit in order to obtain a purity suitable for tissue culture transfections. Subsequently, plasmid purity was tested and the concentration quantitated by UV spectroscopy. Transfections were then performed on R- cells grown overnight to a confluency of ~50% using Lipofectamine+™ and 4μg of plasmid DNA for each construct. After a 3hr transfection, cells were allowed to recover overnight in DMEM + 10% FCS before selection with G418 (Genetecin: Gibco/BRL) at a final concentration of 500μg/ml. Media was replaced every 3-4 days for 2 weeks and drug resistant cells screened for IR expression by FACS analysis using the monoclonal anti-IR antibody 83-7 (For the detailed method see 2.3.3.7). Briefly, a single 10 cm dish of each clonal line was grown to ~80% confluency and then a single cell suspension obtained by treating the cells with a trypsin/EDTA solution. Aliquots were incubated on ice with 83-14, and again after washing with FITC labelled goat anti-mouse antibody. After final washes cells were resuspended in 1% paraformaldehyde until FACS acquisition. The initial FACS analysis indicated that less than 50% of the cell population was expressing human insulin receptor.

In order to obtain a more homogenous population, cells from each transfected line were re-cultured and IR expressing cells subsequently isolated using a FACS cell sorter (IMVS/Hanson centre). Cell populations from the low, middle and high end of the receptor expressing peak were sorted separately, in order to obtain cell lines expressing differing numbers of receptors. After, several weeks of propagation in DMEM+ 500μg/mL G418 it became apparent that only those cells sorted from the middle portion of the receptor

expressing peak had survived, owing mostly to the greatly increased number of cells isolated from this region. A final FACS analysis of several clones of each of hIR-A and hIR-B indicated that the sorting process was successful and that homogenous hIR receptor expressing cells had been obtained (See Figure 4.2). All the clones examined appeared to have similar numbers of receptors both within and between the isoforms. This was an important control for later assessment of the biological effect of a specific amount of ligand, as previous studies at least with the IGF-1R have shown that receptor numbers can affect the strength of the cellular response (Rubini *et al.,* 1997).

Testing the fidelity of each cell line for expressing specifically one isoform could not be determined by using an antibody based method due to the lack of an antibody able to distinguish between each isoform. Therefore, an RT-PCR approach was used. RNA from 10cm plates of each isoform and an untransfected control, grown to 40% confluence, was prepared using the TRIzol™ RNA extraction protocol (Chomczynski and Sacchi, 1987). The quality of the RNA was assessed by visualisation of the 18S (1.8kb) and 28S (5kb) ribosomal subunits. Subsequently, the RNA was reverse-transcribed and amplified using primers spanning exon11. The results of the RT-PCR indicated that each cell line was expressing the one expected isoform, and the control line was not expressing any detectable human insulin receptor RNA (See Figure 4.3).

4.2.4 Creation of a Cell Line Expressing Recombinant hIR-B EDZIP Isoform

The pEFIRESneo:hIR-B EDZIP construct was prepared using the Ultraclean™ miniprep kit and transfected into R- cells grown overnight to a confluency of ~50% using Lipofectamine+™ (as described above in 4.2.3). After a 3hr transfection, cells were allowed to recover overnight in DMEM + 10% FCS before selection with G418 (Genetecin: Gibco/BRL) at a final concentration of 500µg/ml. Media was replaced every 3-4 days for 2 weeks and the presence of receptor was tested in a 96 well plate assay using eu-labelled 83-14 (See figure 4.4). hIR-A EDZIP was collected from the Lec8 hIRex11-EDZIP cell line and tested in the same way as indicated for hIR-B EDZIP. The concentration of receptor secreted into the media was determined based on a standard curve of a known concentration of hIR-A ED, and was found to be ~10ng/mL. In order to obtain a concentration of receptor needed for the following experiments the media was concentrated to ~20ng/mL, using a stirred cell and a 100kDa cut-off nitrocellulose membrane.

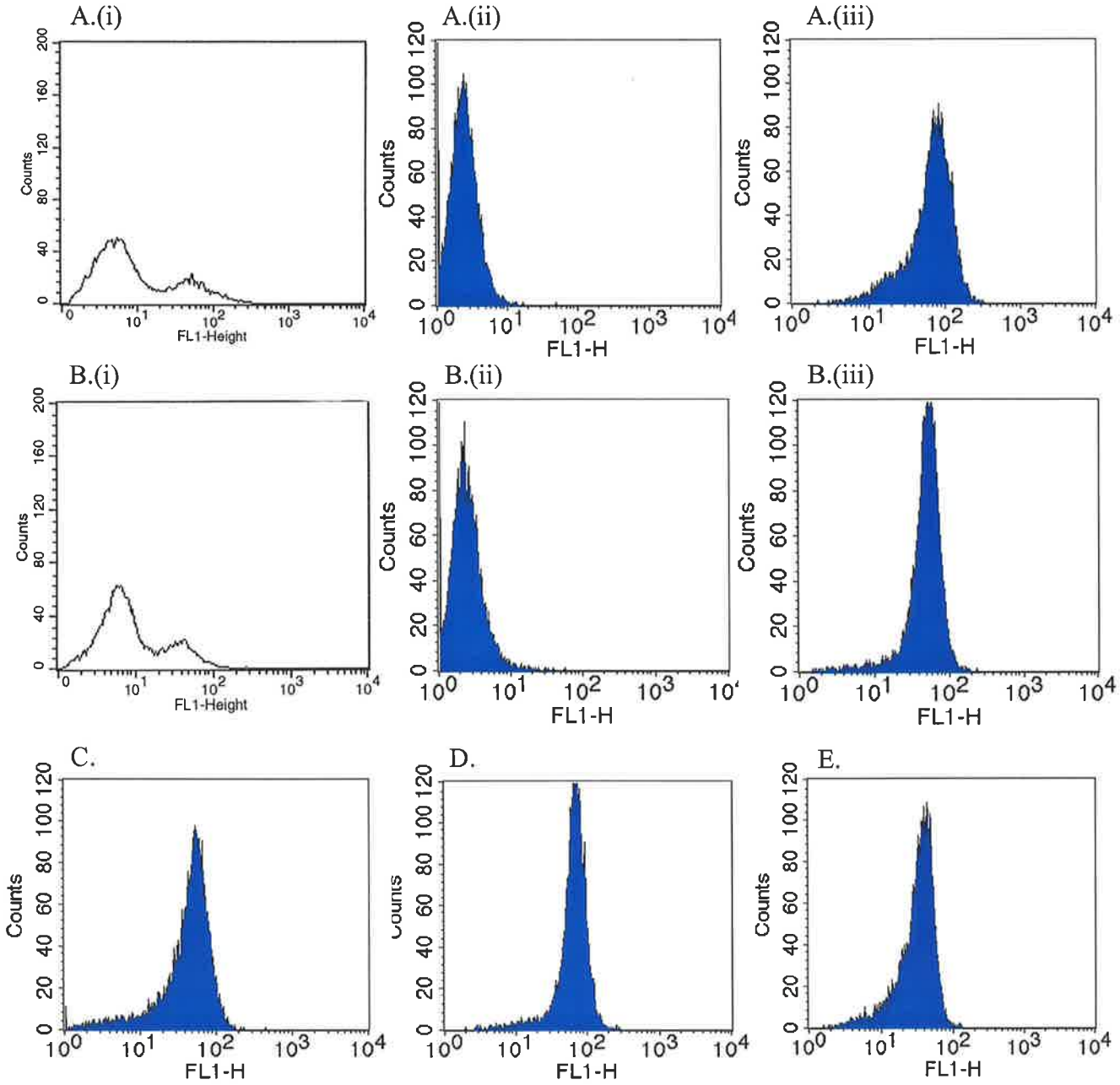


Figure 4.2 Comparison of R-hIR-A and R-hIR-B cells before and after single cell sorting.

Cells were grown to ≤90% confluency, trypsinised and incubated with the mouse monoclonal anti-hIR antibody 83-7 or a mouse IgG control. Cells were subsequently labeled with a FITC conjugated anti-mouse monoclonal antibody, fixed with paraformaldehyde and fluoresencing cells counted on a FACS cell sorter.

A. R-hIR-A, **B**. R-hIR-B
(i) Before single cell sorting: 83-7
(ii) After single cell sorting: IgG
(iii) After single cell sorting: 83-7

C. After single cell sorting hIR-A (Clone 2) : 83-7
D. After single cell sorting R-hIR-B (Clone 2) : 83-7
E. After single cell sorting R-hIR-B (Clone 3) : 83-7

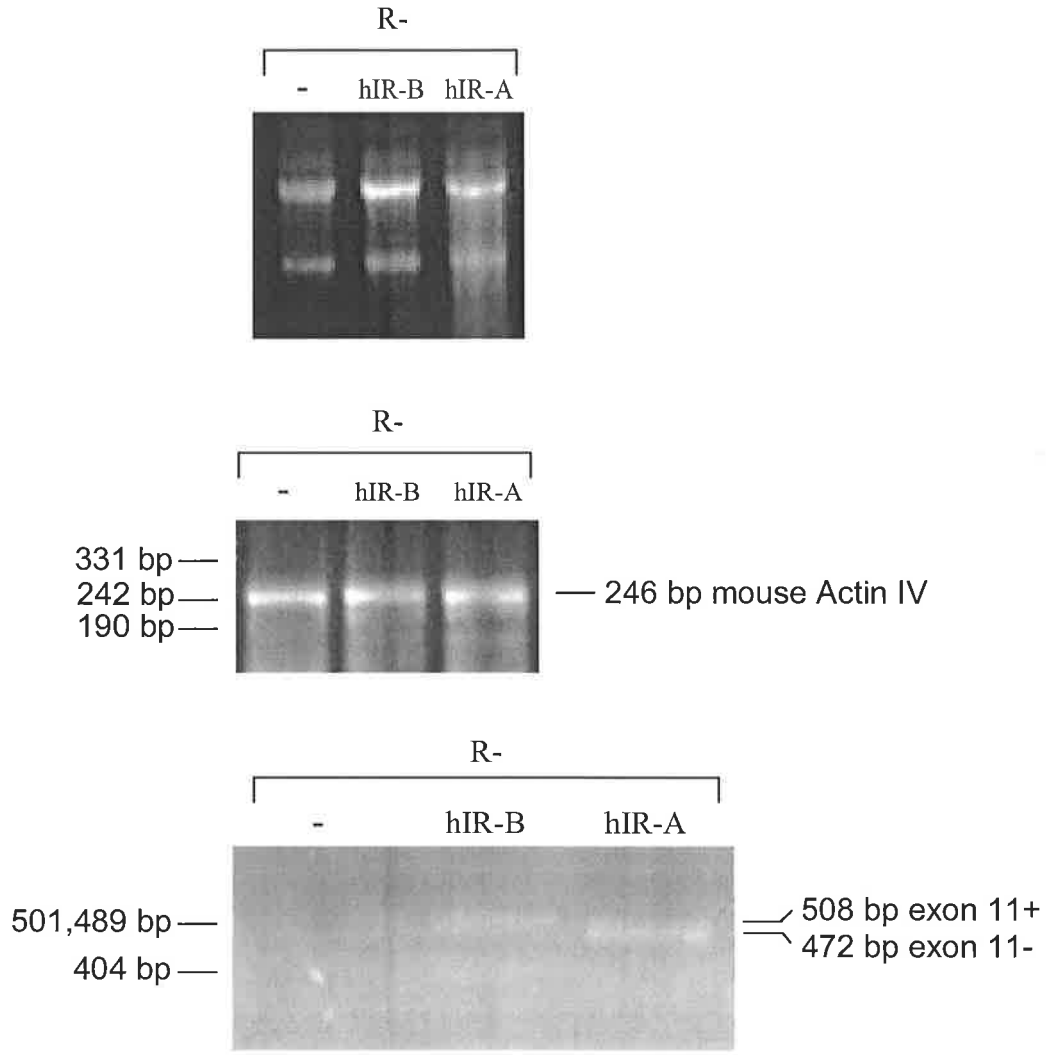
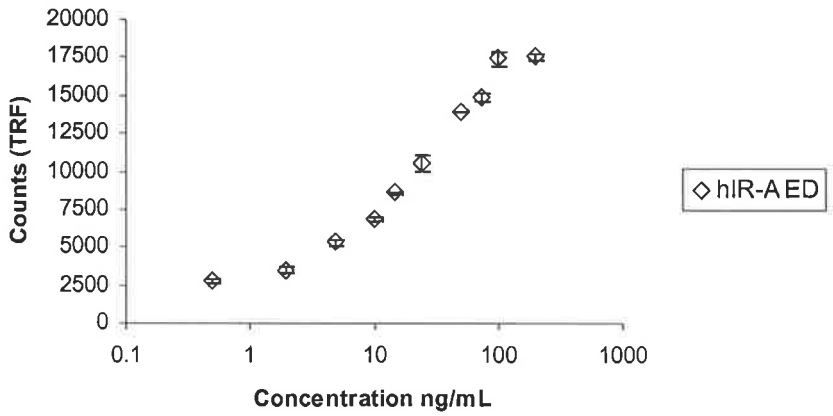


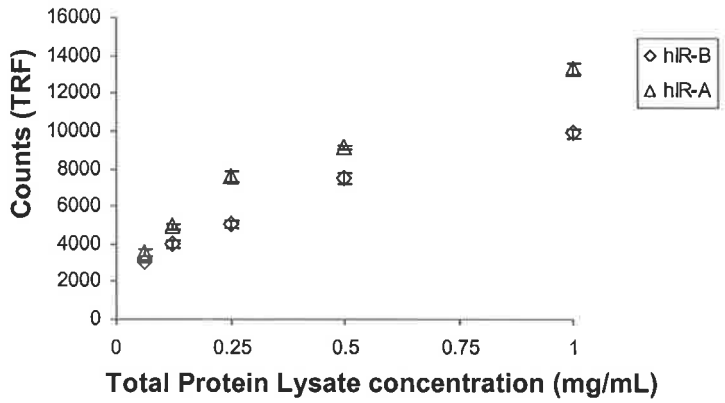
Figure 4.3 RT-PCR analysis of RNA isolated from stable R- cell lines hIR constructs.

A. RNA isolated from untransfected R-, R-hIR-B, and R-hIR-A.
B. Control RT-PCR amplifying mouse Actin IV.
C. RT-PCR amplifying a short region spanning exon 11.

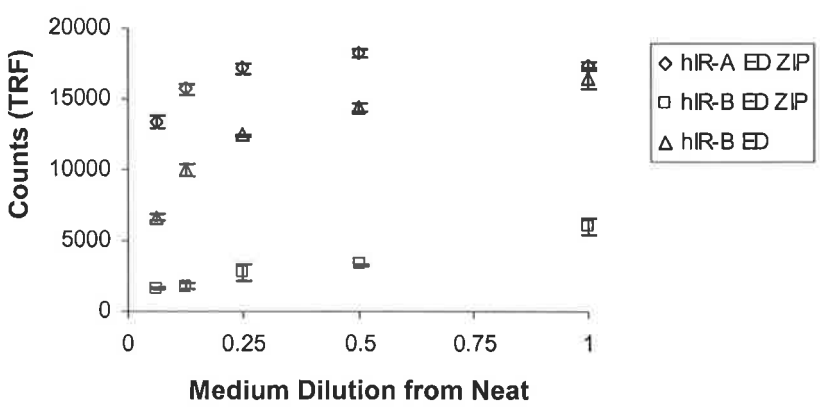
A.



B.



C.



Receptor	hIR-A	hIR-B	hIR-A EDZIP	hIR-B EDZIP	hIR-B ED
Concentration	40 ng/mL	25 ng/mL	640 ng/mL	10 ng/mL	160 ng/mL

Figure 4.4 Quantitation of the amount of receptor produced from each cell line

A standard curve of known receptor concentration using hIR-A ED, capturing receptor with 83-7 and detection of receptor done using a europium assay by eu-labeled 83-14 . Receptor concentrations were then determined for unknown samples. Cells expressing full-length receptor were harvested and the lysate diluted to give a final total protein concentration of 1 mg/mL. Recombinant receptor secreted into the media was taken from cells plated at equal numbers and allowed to grow for 3 days.

A. hIR-A ED Standard Curve.
B. Full-Length Receptors.
C. Recombinant Receptors.

4.2.5 Development of a Receptor Competition Assay using Europium Labeled Ligands

Traditionally studies to determine the binding affinity of a ligand for its receptor have been achieved by performing experiments that compete an I^{125} labeled tracer with an unlabelled ligand for receptor binding. Although I^{125} is very sensitive it has three major drawbacks. Researchers must handle the tracer carefully to avoid absorbing it through the skin. Prolonged exposure could potentially lead to a build up of I^{125} in the thyroid and eventual thyroid dysfunction. The second drawback is that during the experiment each competition mix must be handled and assayed in separate test tubes. Finally I^{125} tracer has a half-life of ~2 months and is only functionally useable for about a month after labeling, even when stored at -80°C. Combined, these drawbacks result in quite a slow throughput assay that must be performed in a relatively short period of time once the tracer is available.

Using fluorophore labeled tracer has not been a practical alternative as high backgrounds and quenching compromises the sensitivity. Recently technologies have been developed that enable the use of the lanthanide elements as fluorophores in a time-resolved fluorescence assay that is potentially very amenable to receptor competition assays. Europium is the most commonly used, and like the other lanthanides has several attractive properties. Compared to I^{125} it is safe to the researcher although care must be taken not to contaminate pipettes. It overcomes the problems of normal fluorophores by having a long fluorescence decay time and a relatively short background fluorescence time resulting in one being able to take measurements after background fluorescence has dissipated. A large Stokes Shift (the difference between excitation and emission wavelength) of almost 300nm means that the measurement is taken where the influence of non-specific signal is minimal. The emission peak is also particularly sharp which raises the possibility of using several different easily distinguishable lanthanides in the one assay and measuring their emission simultaneously. Europium also has a high fluorescence intensity. Finally, europium-labeled proteins are stable for up to 1 year being stored at 4°C. Measurements of time-resolved fluorescence are now typically performed in 96 well trays allowing for high throughput readings, better consistency between samples, and less use of reagents.

4.2.5.1 Europium Labelling Insulin, IGF-1 and IGF-2

In order to label the growth factors, Europium, chelated to N^1-(p-isothiocyanatobenzyl)-DTTA (diethylenetriamine-N^1, N^2, N^3, N^3-tetraacetic acid) was reacted with purified human insulin, IGF-1, and IGF-2 which had been buffer exchanged into labeling buffer (see 2.3.4.2). Reactions were incubated overnight at room temperature and stopped by the addition of poly-

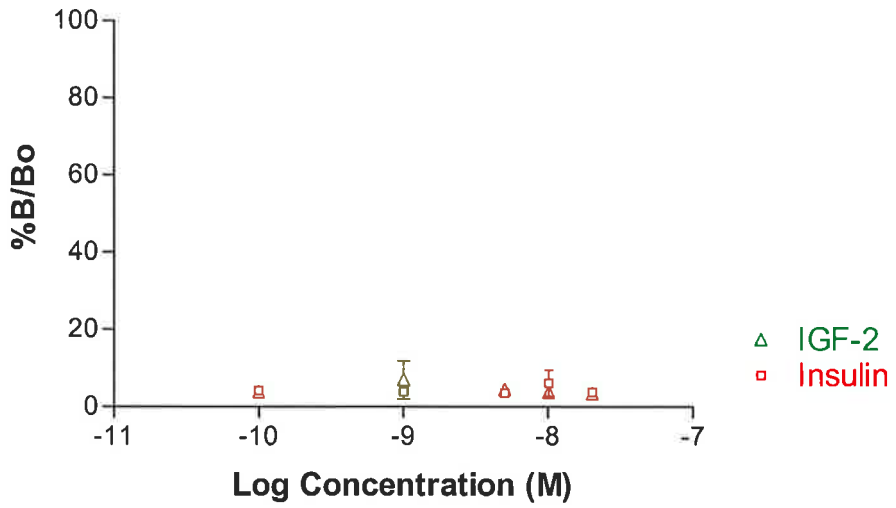
L-Lysine, which quenches any unreacted europium chelate. Since the reaction was non-specific and would occur at any free amine the resultant labeled growth factors were produced as a heterogenous mix. Eu-labeled insulin may be labeled with any combination of the B chain N-terminus, Lysine B29, or the A chain N-terminus. Structural studies had determined that the C-terminal end of the B-chain including Lysine B29 is involved in a conformational change upon binding the insulin receptor, and flexibility in this region has been shown to be very important for high affinity binding (Derewenda *et al.*, 1991) (Ludvigsen *et al.*, 1998). It was unclear how much having a europium chelate attached to Lys B29 would affect this regions ability to allow a native binding affinity. HPLC was performed to determine the number of labeled and unlabeled species in the competition mix and indicated that there were several species in each reaction mix although dominated by one major species (Data not shown). Individual species were unable to be isolated due to complications with the eu-labeled peptides sticking to the HPLC column and only eluting off after the chelate had been removed. Lablled peptides were used despite their heterogenous nature.

4.2.5.2 Optimization of the Europium Competition Assay

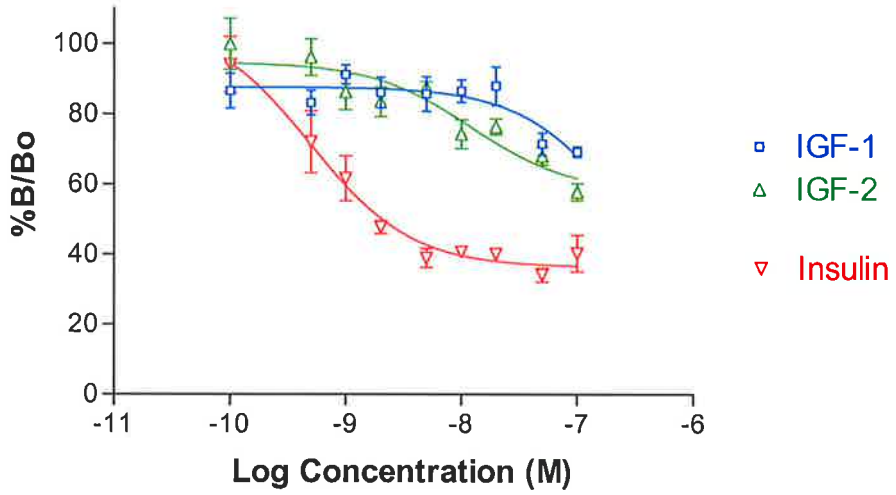
The functionality of the europium labeled growth factors were tested for their ability to act as tracers in competition assays using insulin receptor solubilisd from NIH 3T3 cells overexpressing human insulin receptor or IGF-IR solubilised from stably transfected R- cells (See 5.2.2). For this and all subsequent competition experiments cells were grown in 10cm dishes to ~90% confluency before being serum deprived for 3 hrs (to allow any bound growth factors to be internalized and degraded), followed by cell lysis. Lysates were then aliquoted into anti-receptor antibody coated white (Lumintrac 400) 96 well plates and incubated overnight at 4°C. The europium labeled tracer was added to each well separately then unlabelled growth factors added, which were serially diluted in Binding Buffer. Competition mixes were incubated overnight at 4°C and the amount of labeled tracer remaining after several washes quantitated using time resolved fluorescence.

Figure 4.5 shows an experiment using hIR solubilised from NIH 3T3 cells, captured with the anti-insulin receptor antibody 83-7, and Eu-insulin used as the tracer. As expected, insulin competes well with itself for binding the insulin receptor with an EC_{50} of ~1nM and IGF-1 and 2 have much poorer affinity. In a further control assay, figure 4.6 shows competition assay instead using hIGF-IR solubilised from a hIGF-IR transfected R- cell line captured with 24-31 and eu-IGF-1 used as the tracer (See section 5.2.2). These results clearly indicate high affinity binding for IGF-1 and 2 and much weaker binding for insulin.

A.



B.

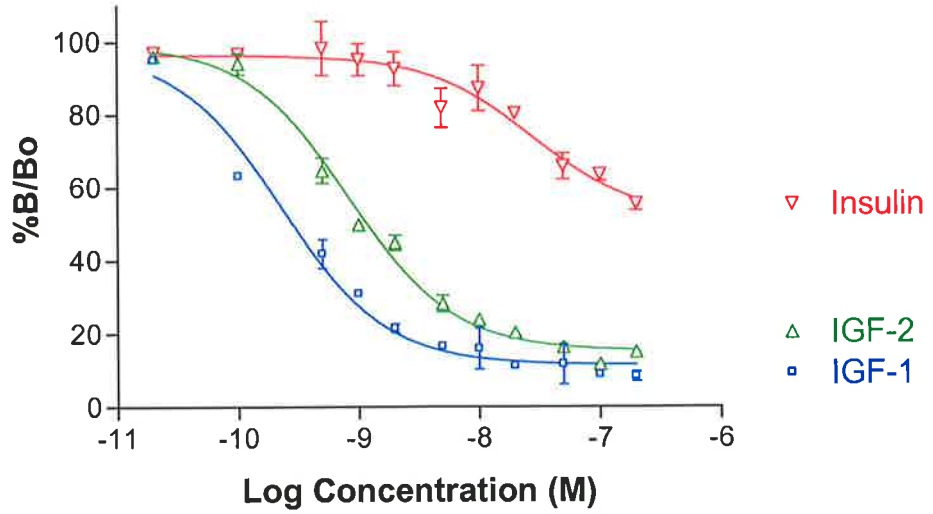


	IGF-1	IGF-2	Insulin
hIR-B	249 ± 16.2 nM	12.2 ± 2.0 nM	0.49±0.24 nM

Figure 4.5 Europium-labelled Ligand Competition Binding Assay

Stable cell lines expressing each receptor were grown in 10cm dishes to ~90% confluency before being harvested in lysis buffer and hIR captured on 83-7 coated white 96 well plates. A competition mix between Eu-insulin and serial dilutions of either unlabelled IGF-1, IGF-2, or insulin was incubated with the receptor overnight at 4°C before unbound growth factor being washed away and the amount of remaining Eu-insulin measured using time-resolved fluorescence. Data was converted to a percentage of maximum counts bound and fitted to a one-site competition binding curve using Prism™.

A. R- fibroblasts
B. Full-length hIR-B solubilised from NIH3T3 fibroblasts overexpressing hIR-B



	IGF-1	IGF-2	Insulin
hIGF-IR	0.23 ± 0.1 nM	0.83 ± 0.06 nM	28 ± 5.3 nM

Figure 4.6 IGF-IR Competition Binding Assay using Eu-IGF-1

Stable cell lines expressing each receptor were grown in 10cm dishes to ~90% confluency before being harvested in lysis buffer and IGF-IR captured on 24-31 coated white 96 well plates. A competition mix between Eu-IGF-1 and serial dilutions of either unlabelled IGF-1 ,IGF-2, or insulin was incubated with the receptor overnight at 4°C before unbound growth factor being washed away and the amount of remaining Eu-IGF-1 measured using time-resolved fluorescence. Data was converted to a percentage of maximum counts bound and fitted to a one-site competition binding curve using Prism™.

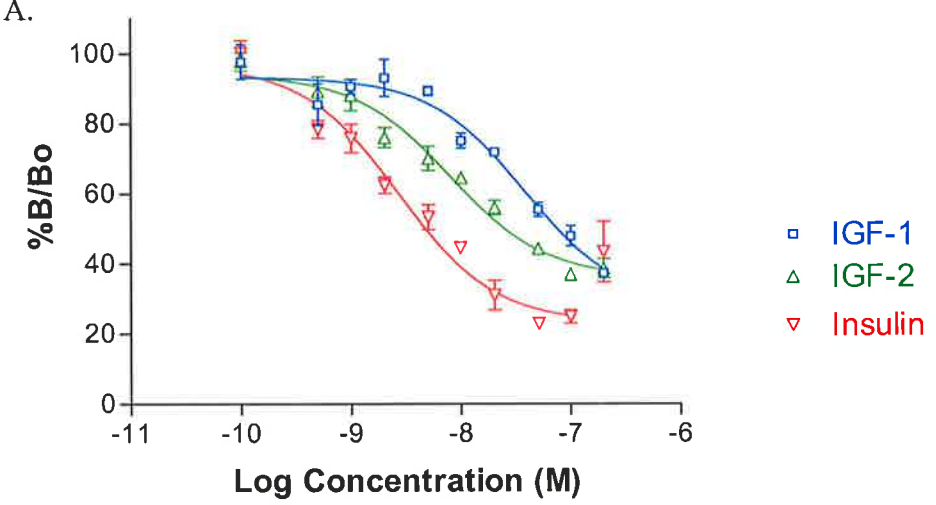
As mentioned previously mouse R- cells express a very level of insulin receptor, so a control experiment was performed on untransfected R-s to determine if any mouse insulin receptors were captured using the monoclonal anti-human insulin-receptor antibody 83-7 (See Figure 4.5). The results indicated that the cell line does not produce any background signal therefore any data gained must be due to the exogenous expression of human insulin receptor from the stably transfected vector.

4.2.6 Comparison of Ligand Binding between Full-length hIR Isoforms in the Europium Competition Assay

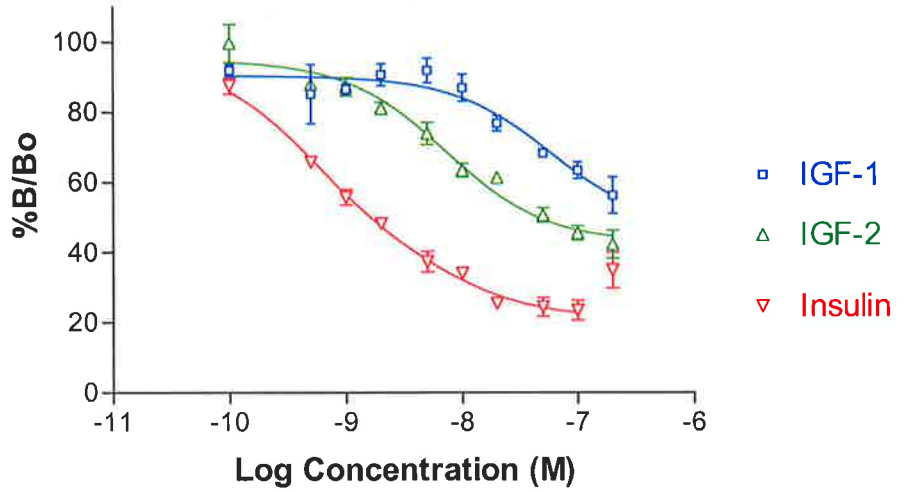
Europium competition assays were performed on the two R- cell lines expressing each of the human insulin receptor isoforms. Cells from each line were grown in 10cm dishes to a confluency of ~90%, and further incubated for 2 hrs in serum free media to allow any internalized receptors to be recycled back to the membrane and remove any insulin, IGF-1 or IGF-2 that may be in the serum. Cells were subsequently lysed in ice cold lysis buffer and the insoluble fraction pelleted by centrifugation. Soluble lysate from two 10 cm dishes was added to a single 83-7 coated white 96 well tray and incubated overnight at 4°C to allow receptor capture. Wells were washed to remove unbound receptor and serially diluted unlabelled growth factor was then added with Eu-labeled insulin, and the competition mix incubated overnight at 4°C. Unbound growth factor was then removed and the amount of bound Eu-insulin quantitated by time-resolved fluorescence.

The graphs of the results shown in figure 4.7, show a representative of four separate experiments. In this assay the hIR-B isoform appears to have a clear preference for insulin with an EC_{50} of 0.93 nM, with IGF-2 and IGF-1 having 7.7 and 65 fold lower affinity for the receptor relative to insulin's EC_{50}. The hIR-A isoform does not have as distinct a ligand preference, shown by a decreased EC_{50} of 2.7 nM for insulin and increased affinity for both IGF-1 and IGF-2 at only 3 and 14 fold less than insulin respectively. Insulin in each isoform also exhibits the classical bell shaped dose response competition curve that begins to curve upwards at around 100nM insulin concentration. The absolute EC_{50} values when compared between each isoform suggest that hIR-B has a higher affinity for insulin than hIR-A and that there is only a two fold improvement for IGF-2 between hIR-A and hIR-B and a similar improvement for IGF-1. Previous analyses of the relative affinities of insulin for the two IR isoforms have lacked a consensus. It has been reported that IR-B has higher affinity

A.



B.



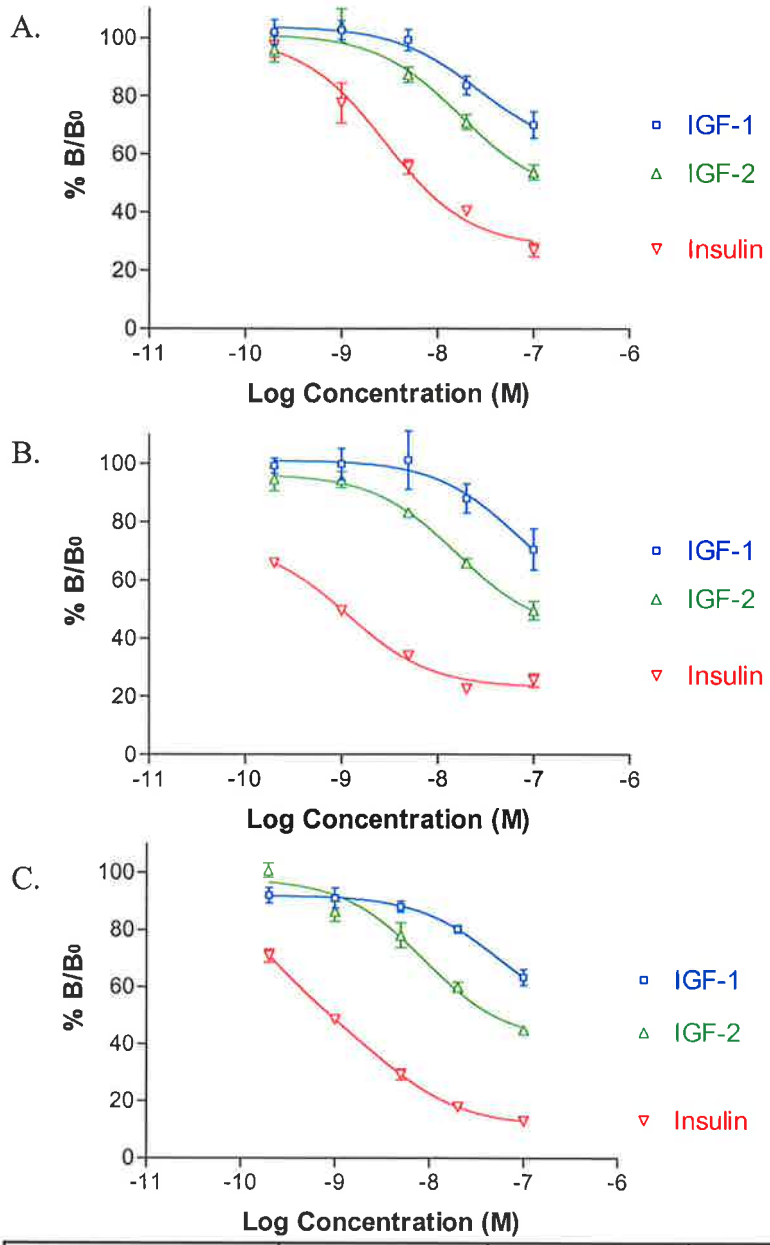
EC50 relative to Insulin

	IGF-1	IGF-2	Insulin
hIR-A	37.8 ± 16 nM	7.9 ± 2.3 nM	2.7 ± 0.95 nM
hIR-B	59.7 ± 28 nM	7.2 ± 2.2 nM	0.93 ± 0.18 nM

Figure 4.7 Comparison of full-length hIR Isoforms in a Competition Binding Assay

Stable cell lines expressing each receptor were grown in 10cm dishes to ~90% confluency before being harvested in lysis buffer and hIR captured on 83-7 coated white 96 well plates. A competition mix between Eu-insulin and serial dilutions of either unlabelled IGF-1, IGF-2, or insulin was incubated with the receptor overnight at 4°C before unbound growth factor being washed away and the amount of remaining Eu-insulin measured using time-resolved fluorescence. Data was converted to a percentage of maximum counts bound and fitted to a one or two-site competition binding curve using Prism™.

A. Full-length hIR-A solubilised from R- stable cell line expressing hIR-A.
B. Full-length hIR-B solubilised from R- stable cell line expressing hIR-B.



	IGF-1	IGF-2	Insulin
hIR-A (clone 2)	26 nM	17 nM	2.8 nM
hIR-B (clone 2)	68 nM	15.7 nM	1.2 nM
hIR-B (clone 3)	54 nM	8.8 nM	1.3 nM

Figure 4.8 **Comparison of different cell lines expressing full-length hIR Isoforms in a Competition Binding Assay.**

Stable cell lines expressing each receptor were grown in 10cm dishes to ~90% confluency before being harvested in lysis buffer and hIR captured on 83-7 coated white 96 well plates. A competition mix between Eu-insulin and serial dilutions of either unlabelled IGF-1, IGF-2, or insulin was incubated with the receptor overnight at 4°C before unbound growth factor being washed away and the amount of remaining Eu-insulin measured using time-resolved fluorescence. Data was converted to a percentage of maximum counts bound and fitted to a one or two-site competition binding curve using Prism™.

A. Full-length hIR-A (clone 2) solubilised from R- stable cell line expressing hIR-A.
B. Full-length hIR-B (clone 2) solubilised from R- stable cell line expressing hIR-B.
C. Full-length hIR-B (clone 3) solubilised from R- stable cell line expressing hIR-B

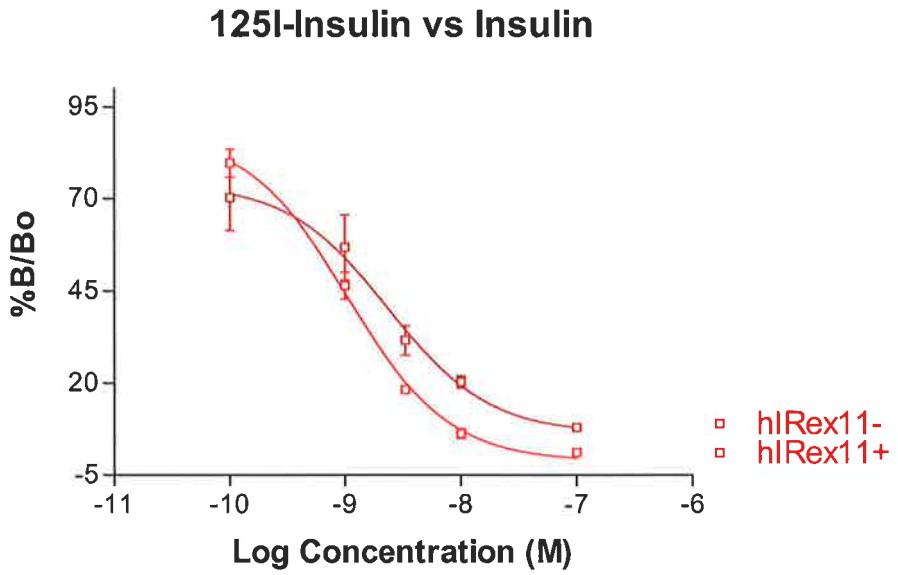


Figure 4.9 ^{125}I-insulin Competition Assay

R- cells, wild-type or stably transfected, were plated in 24 well trays to yield 80-95% confluence overnight. Following at least 2 hrs growth in serum free medium, ^{125}I-insulin(Y19)) was added to each well in ice cold buffer in addition to serially diluted unlabelled insulin. Plates were left overnight at 4°C, competition mixes were then aspirated and washed and cells harvested and transferred to eppendorfs for counting on a multigamma counter (CSIRO). Data was converted to a percentage of maximum counts bound and fitted to a one-site competition binding curve using Prism™.

binding curve. It has been shown that both hIR-B and hIR-B isoforms have curvilinear scatchard plots and therefore should have two-site competition binding curves (DeMeyts, 1994). These assays however did not suggest that IGF-1 or IGF-2 had a two-site interaction.

It has been reported that tracer concentration influences obtained ligand affinities. A low tracer concentration is able to increase the observed affinity of a ligand. This may be affecting this assay since the concentration of eu-Insulin added is based on loading a specific number of counts to each well. The exact concentration of the labeled insulin could not be obtained due to the labeling procedure resulting in a heterogeneous mix, and the usual method of HPLC quantitation being unachievable (See 4.2.5.1).

The lower affinities obtained in this assay may also be attributed to the use of solubilised receptors, that lacking a membrane may not allow the full length receptor to adopt the exact same tertiary confirmation of a membrane bound receptor. This change in overall receptor conformation has been seen experimentally, as there is a significant difference between electron microscopy images obtained from insulin receptors embedded in a plasma membrane and from isolated receptors (Tulloch *et al.,* 1999) (Luo *et al.,* 1999).

Although the exact EC_{50} values obtained from this experiment do not agree completely with published values, the ratios between the affinities of insulin, IGF-1 and IGF-2 are generally in agreement and therefore overall, this assay is suitable for internally comparing the relative affinity of different ligands for the insulin receptor (Denley *et al.*, 2004).

4.2.6.1 The C and D domains of IGF-1 and IGF-2 Determine Receptor Specificity

In order to determine whether the C and D domains of IGF-2 are responsible for enabling higher affinity binding to the hIR and in particular hIR-A, domain swaps were made between IGF-2 and IGF-1. The C and D domains were exchanged singly and in tandem between the two growth factors, expressed in *E. coli*, refolded, and purified to homogeneity. Quantitated amounts of these recombinant proteins were provided by Adam Denley (Wallace laboratory, Adelaide University) for use in competition assays with R-IGF-IR, R-hIR-A, and R-hIR-B. Assays were performed as described in section 4.2.6, and a representative of at least three separate experiments for each cell line are shown in figure 4.10. In each receptor system, exchanging either the C or D domain results in a shift in affinity towards the exchanged parent, with the C domain appearing to have a greater effect. The double domain swaps

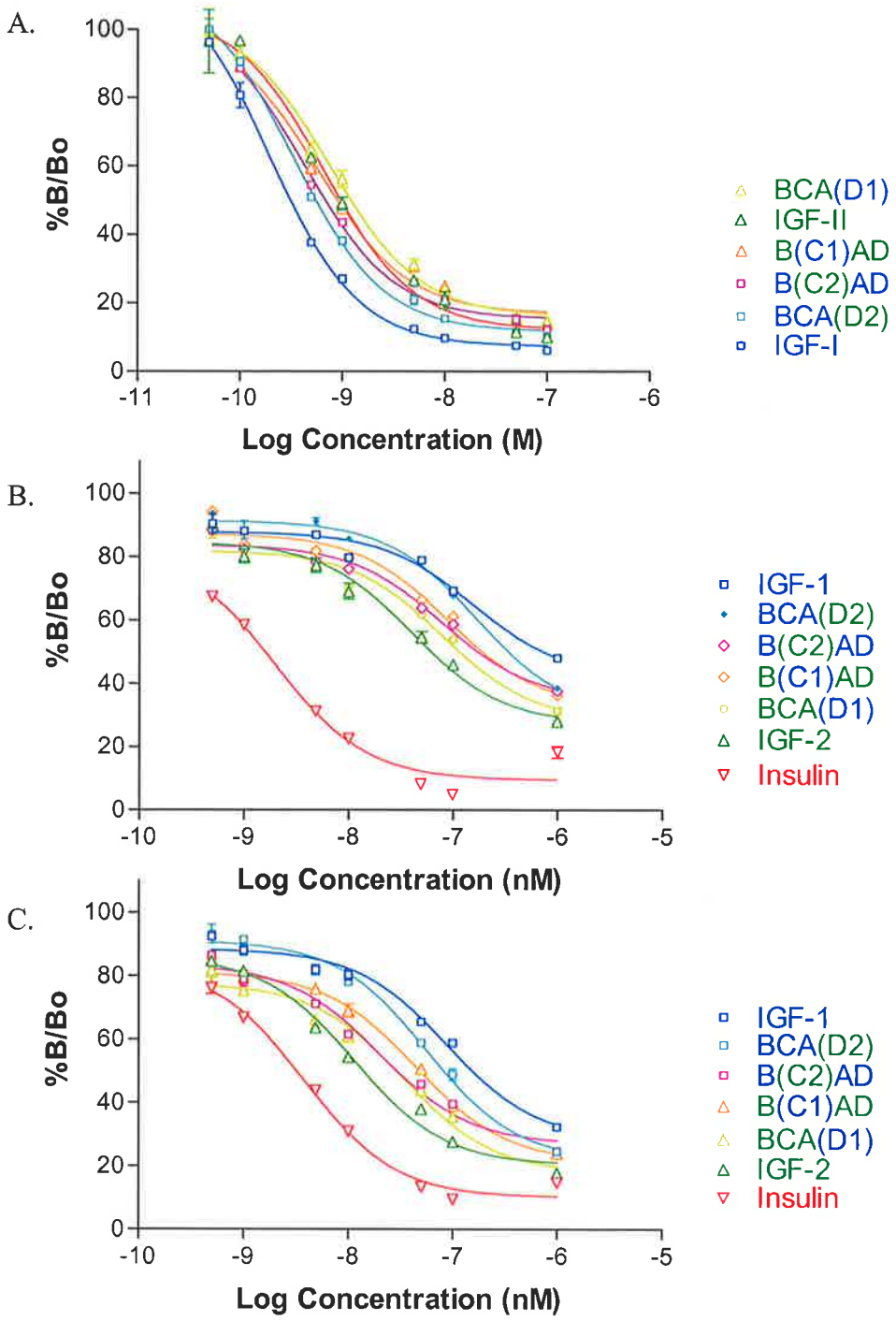


Figure 4.10 Effect on receptor binding of swapping domains between IGF-1 and IGF-2.

Stable cell lines expressing each receptor were grown in 10cm dishes to ~90% confluency before being harvested in lysis buffer and IGF-1R or IR captured on 24-31 or 83-7 coated white 96 well plates respectively. A competition mix between eu-IGF-1 or Eu-insulin and serial dilutions of either unlabelled IGF-1 ,IGF-2, insulin, or recombinant IGFs was incubated with the receptor overnight at 4°C before unbound growth factor being washed away and the amount of remaining Eu-IGF-1 measured using time-resolved fluorescence. Data was converted to a percentage of maximum counts bound and fitted to a one-site competition binding curve using Prism™.

A. hIGF-1R, eu-IGF-1
B. hIR-B, eu-Insulin
C. hIR-A, eu-Insulin

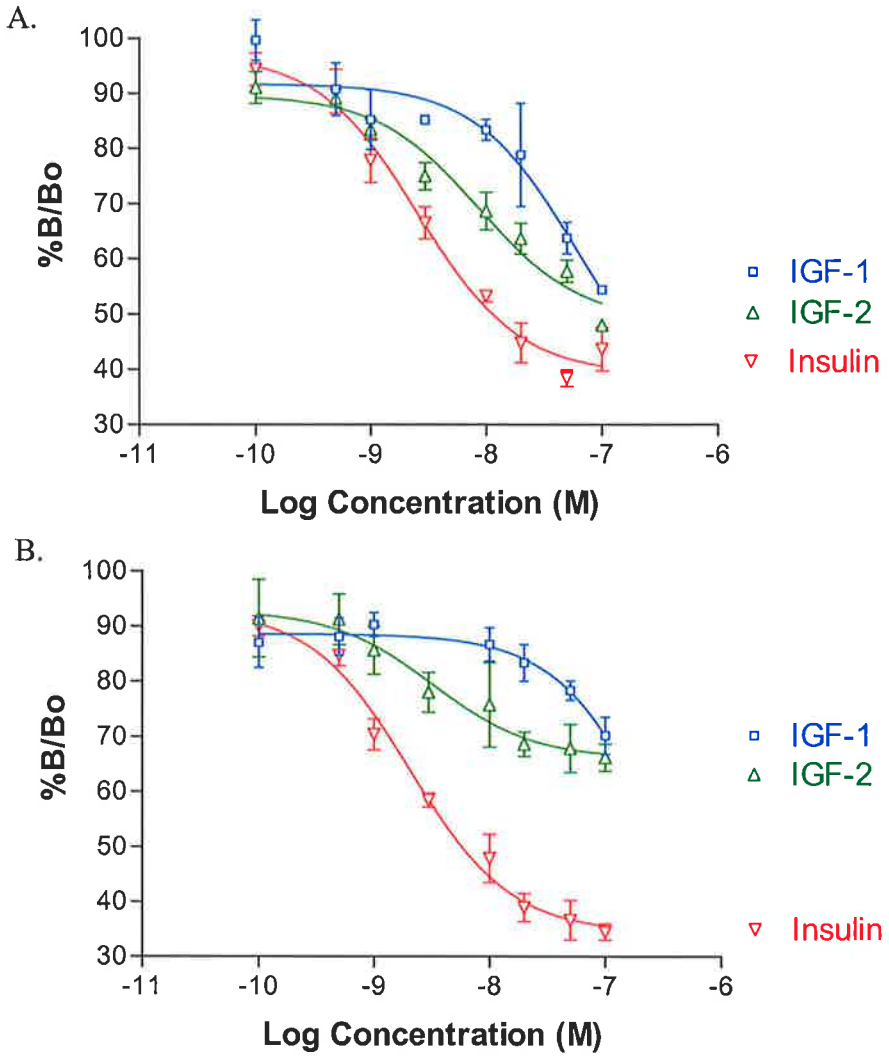
effectively changes the affinity of the recombinant IGF to that of the C and D domains parent, such that IGF-1 with the IGF-2 C and D domains has the same affinity as IGF-2 and vice versa. Interestingly this holds true for the IGF-IR and both isoforms of the IR (Denley *et al.*, 2004). Therefore it appears that the C and D domains of the IGF proteins are mostly responsible for the differences in specificity. The molecular basis for the change cannot be determined from these experiments, however for the C domain in particular, the two most likely reasons are, that either the shortened length of the IGF-2 C domain reduces steric hindrance when binding the insulin receptor, or that the increased basic charge, from two additional arginines, improves the electrostatic fit.

4.2.7 Comparison of Ligand Binding between Recombinant Soluble Ectodomain Isoforms in the Europium Competition Assay.

Recombinant ectodomain forms of each isoform of the insulin receptor were received as purified hIR-A ED protein, and a cell line expressing hIR-B ED called CT5 (kind gifts from Colin Ward, CSIRO Melbourne). For the europium competition assay hIR-A ED protein was added at a concentration of 0.5µg/100µL in each well and the rest of the assay performed as described above. The concentration of hIR-B ED secreted into the media of CT5 cells was determined to be ~0.2 µg/100µL in an ELISA using quantitated hIR-A ED as the standard. It was therefore added neat to the wells as the amount of protein produced was low and the experiment only yielded analyzable results when this concentration was used. The results as seen in figure 4.11, are a representative of two separate successful experiments, and show a similar relationship between the growth factors and between receptors as seen with the full length receptor, only the EC_{50} values appear to be lower, especially for IGF-1 and IGF-2. This may be due to IGF-1 and IGF-2 only being able to bind well with a native receptor conformation. Overall the reduced affinities are consistent with previous observations (Cosgrove *et al.*, 1995)(Jansson *et al.*, 1997).

4.2.8 Comparison of Ligand Binding between Recombinant Soluble Ectodomain ZIP Isoforms in the Europium Competition Assay.

Medium was collected from Lec8hIR-A EDZIP (a kind gift from Colin Ward CSIRO Melbourne), and R-hIR-B EDZIP cells, and used in a europium competition assay as described above only, medium from the Lec8hIR-A EDZIP cells was diluted ¼, whereas medium from the R-hIR-B EDZIP cells was used neat, reflecting the reduced levels of



EC50 relative to Insulin

	IGF-1	IGF-2	Insulin
hIR-A ED	45.5 ± 13 nM	19.8 ± 5.1 nM	2.6 ± 0.72 nM
hIR-B ED	198 ± 21 nM	96.8 ± 48 nM	2.1 ± 0.5 nM

Figure 4.11 Comparison of recombinant hIR ED Isoforms in a Competition Binding Assay

Stable cell lines secreting each receptor were grown in 10cm dishes and media collected after 3 days. hIR was captured on 83-7 coated white 96 well plates. A competition mix between Eu-insulin and serial dilutions of either unlabelled IGF-1, IGF-2, or insulin was incubated with the receptor overnight at 4°C before unbound growth factor being washed away and the amount of remaining Eu-insulin measured using time-resolved fluorescence. Data was converted to a percentage of maximum counts bound and fitted to a one or two-site competition binding curve using Prism™.

A. Purified hIR-A ED secreted from CT5 stable cell line expressing hIR-A.
B. hIR-B ED secreted from CT5 stable cell line expressing hIR-B.

receptor secretion. Assays were performed three times in triplicate and a representative graph of one of the experiments is shown (See Figure 4.12). The graphs show a similar trend to those of the full-length receptors, however the pronounced change in shape of the curve of insulin is not seen beyond 100nM. The relative affinities of the ligands are more similar to the ectodomain constructs and slightly higher than full-length receptors. The consistent observation that insulin has a higher affinity for the hIR-B isoform would continue to suggest either a difference between europium and iodine based assays or the differently labeled ligands themselves. In this assay that the hIR EDZIP receptors unexpectedly produced lower EC50 values than the hIR ED receptors. This may again be indicative of the Eu-labelled tracer having a slightly different mode of receptor binding than iodinated tracer.

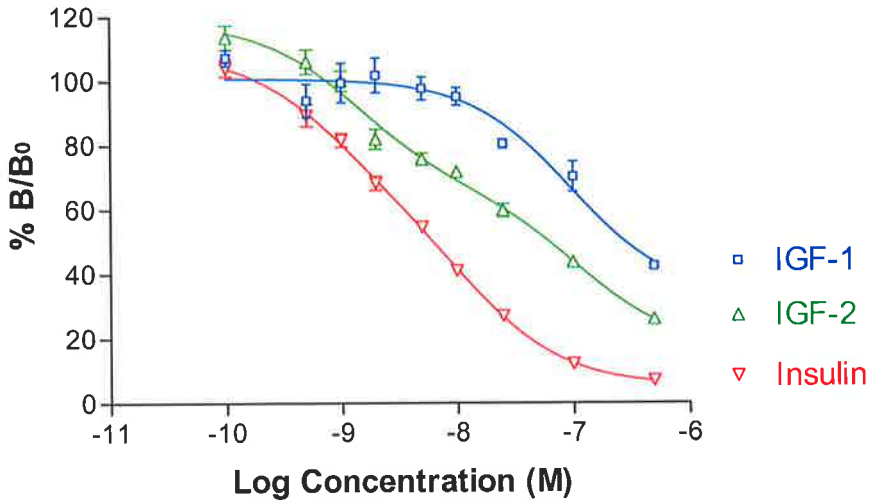
4.2.9 Phosphorylation Assays

A study by Kosaki *et al.*, 1995 found that IR-B was more efficient at auto-phosphorylating itself and recombinant IRS-1. The competition assays performed above indicated that insulin bound IR-B better than IR-A and that IGF-1 and 2 bound IR-A better than IR-B. The next step was to investigate whether receptor and substrate phosphorylation in response to growth factor correlated with the relative binding affinity.

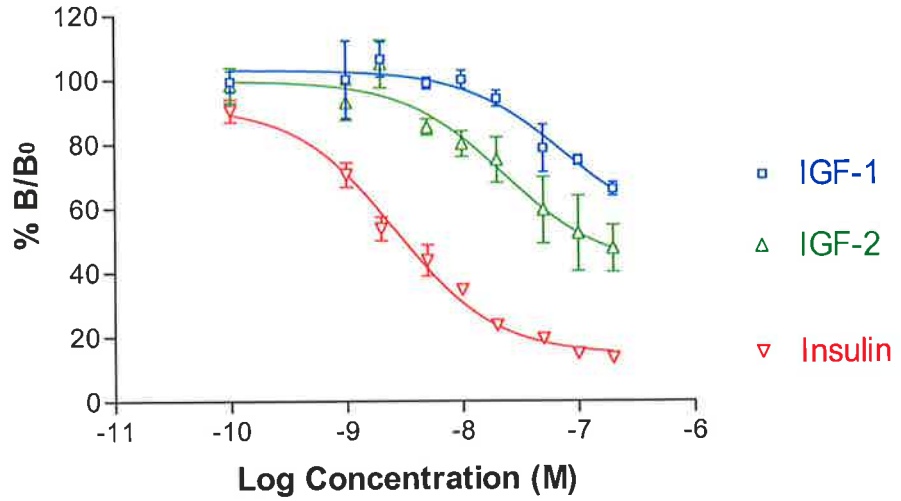
Stable cell lines expressing each of the receptors were seeded into clear 96 well plates to give ~90% confluence with overnight growth. The cells were placed into serum free medium for 3hrs prior to stimulation with serially diluted growth factor for 12 min at 37°C. The media was aspirated and lysis buffer containing protease and phosphatase inhibitors added to each well. The cell lysate was subsequently added to white 96 well plates pre-coated with the anti-insulin receptor antibody 83-7. After an overnight incubation at 4°C, wells were washed and incubated with a europium labeled anti-phosphotyrosine antibody (PT-100). Time-resolved fluorescence was then used to quantitate the amount of antibody bound in each well.

The results (See Figure 4.13) show, as expected, that with increasing amounts of growth factor the amount of tyrosine phosphorylation from the immunopreciptated receptor and any receptor bound proteins also increases. For hIR-B Insulin is ~10 fold better at inducing tyrosine phsophorylation than IGF-2 which is itself ~2.5 fold better than IGF-1. This correlates fairly well with the competition data in terms of relative potency. It is interesting to note that at the higher concentrations of IGF phosphorylation levels approach that of insulin, unlike with the IGF-IR (See Figure 4.13) where insulin has a minimal effect at all concentrations tested. The hIR-A data shows that insulin has slightly less potency indicated

A.



B.



EC50 relative to Insulin

	IGF-1	IGF-2	Insulin
hIR-A EDZIP	229 nM	30.7 nM	4.6 nM
hIR-B EDZIP	219 nM	70 nM	2.6 nM

Figure 4.12 Comparison of recombinant hIR ED ZIP Isoforms in a Competition Binding Assay

Stable cell lines secreting each receptor were grown in 10cm dishes and media collected after 3 days. hIR was captured on 83-7 coated white 96 well plates. A competition mix between Eu-insulin and serial dilutions of either unlabelled IGF-1, IGF-2, or insulin was incubated with the receptor overnight at 4°C before unbound growth factor being washed away and the amount of remaining Eu-insulin measured using time-resolved fluorescence. Data was converted to a percentage of maximum counts bound and fitted to a one or two-site competition binding curve using Prism™.

A. hIR-A ED ZIP secreted from lec8 stable cell line expressing hIR-A.
B. hIR-B ED ZIP secreted from R- stable cell line expressing hIR-B.

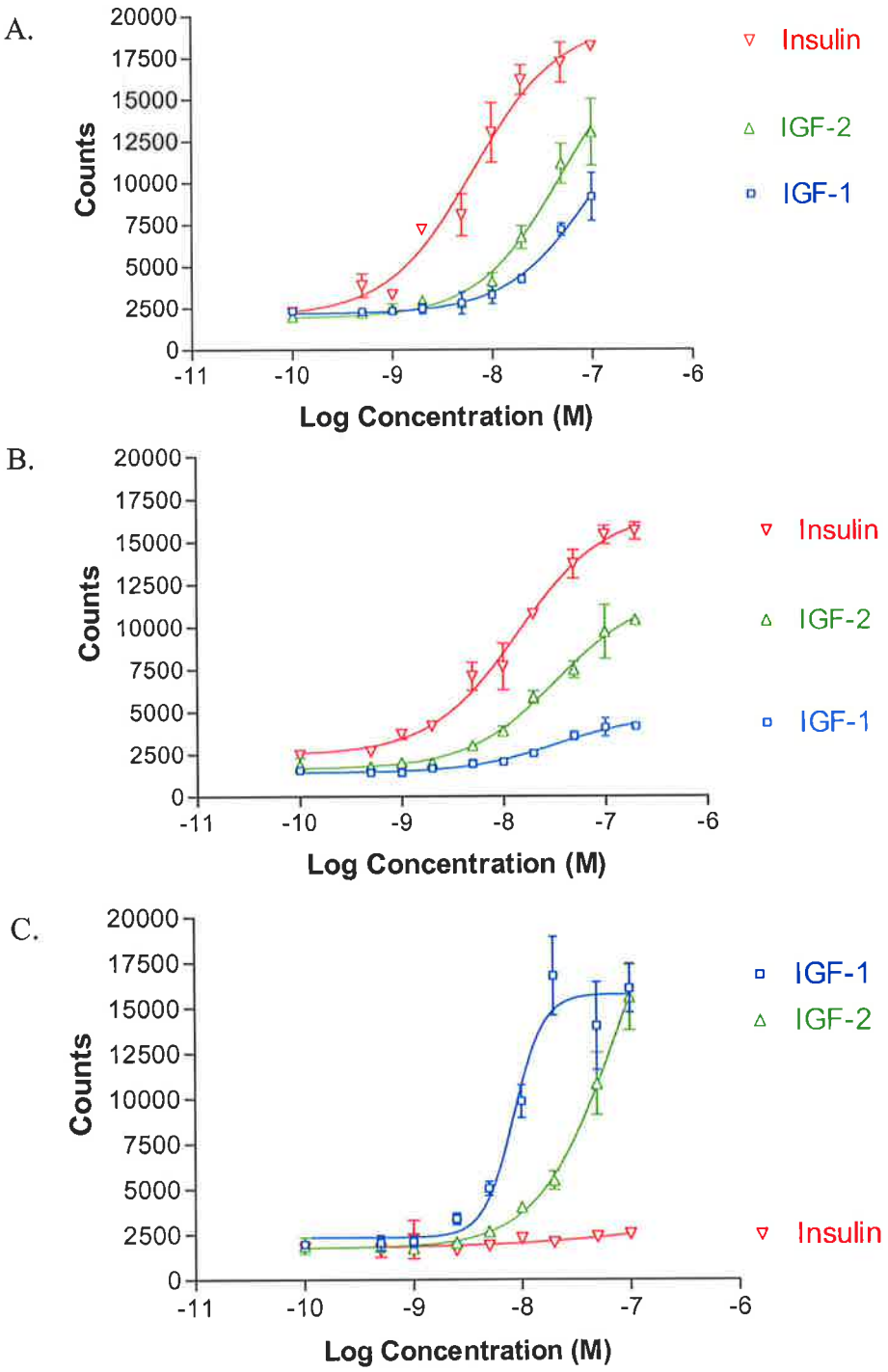


Figure 4.13 Receptor tyrosine phosphorylation assay.

Stable cell lines expressing insulin receptor were grown in 96 well trays to ~90% confluency and allowed to go serum free for 4hrs before stimulation with a range of serially diluted growth factors for 15 min. The medium was aspirated, and ice cold lysis buffer added and solubilised receptor captured on 83-7 coated white 96 well plates. The amount of phosphotyrosine phosphorylation within the captured receptor/substrate complex was obtained using eu-PT100 antibody and measuring bound antibody by time-resolved fluorescence. Phosphorylation data was fitted to an uphill dose response curve using Prism™ software.

A. R-hIR-B.
B. R-hIR-A.
C. R-hIGF-1R.

by it reaching its maximal phosphorylation level at a higher concentration, keeping consistent with its reduced affinity for hIR-A compared to hIR-B. IGF-2 has a slightly improved potency on hIR-A and relative to insulin is only ~3-5 fold worse. IGF-1 however although binding to hIR-A with an improved affinity over hIR-B appears to induce less tyrosine phosphorylation. Taken overall the tyrosine phosphorylation correlates well with the binding strengths of the ligands, although there may be difference in the capacity for each receptor to be activated, as found originally by Kosaki *et al.*, 1995.

Additionally, a time course phosphorylation assay was performed in order to determine if there were any differences in the kinetics of the way the isoforms were activated by insulin or the IGFs. In this assay R- cells expressing hIR-A, hIR-B, or hIGF-1R were plated at the same cell density overnight into 96 well plates. Cells were subsequently serum starved for 3 hrs and then treated with either 10nM or 50nM growth factor included in serum-free medium. At various time points, the medium was aspirated and ice-cold lysis buffer including protease and phosphatase inhibitors was added to stop the reaction. Lysate was added to pre-coated white 96 well plates as with the europium competition assay, and the amount of tyrosine phosphorylation of the receptor and bound proteins detected using the europium labeled anti-phosphotyrosine antibody (PT-100). The results using 10 nM growth factor show clearly that insulin is far more potent than IGF-1 and IGF-2 on both insulin receptors, and IGF-1 is significantly stronger than IGF-2 on the IGF-1R (See figure 4.14). These results did not clearly show a difference between IGF-1 and IGF-2 on hIR-A, although it did appear that IGF-2 was more potent in each case. In order to differentiate between them better, a further assay was performed using 50 nM growth factor (See figure 4.15).

When the phosphorylation results from this assay are compared to results obtained by other groups it can be seen that these results exhibit a higher concentration of ligand required for 50% maximal phosphorylation. This may be due to the time taken for the cells to be serum free. This assay was performed with a 3hr serum free period before stimulation whereas other studies have used 24hr serum free incubations. The extra time may sensitize the cells more to ligand stimulation and therefore result in lower concentrations giving a greater phosphorylation response. The serum free conditions are an artificially created state and probably do not mimic the kind of environment seen by most cells under normal physiological conditions. A cell is unlikely to experience a prolonged absence from serum and other growth factors and therefore an extended serum free incubation is not a realistic scenario.

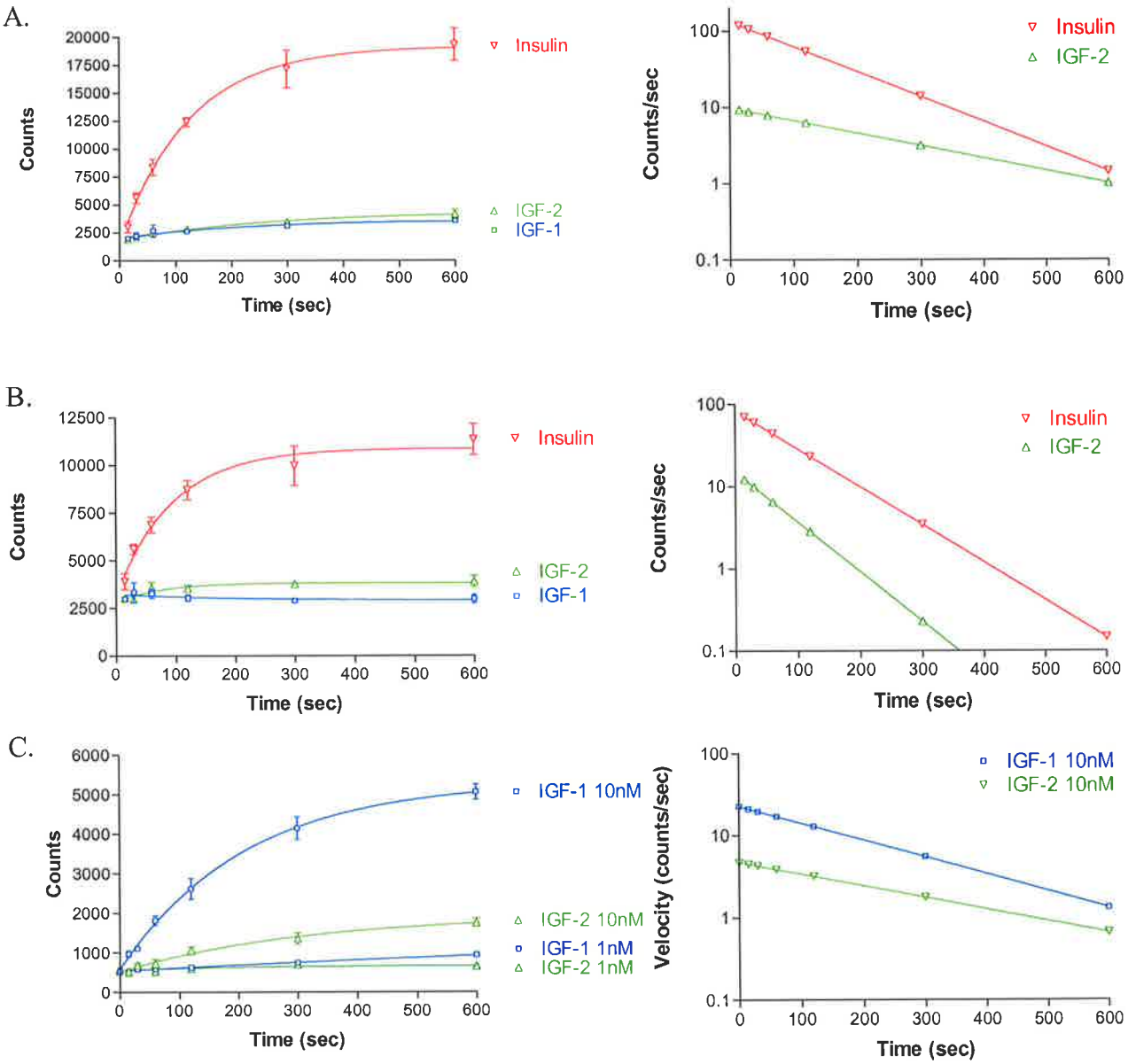


Figure 4.14 Time Course Phosphorylation Assay using 10nM Growth Factor.

Stable cell lines expressing each receptor were grown in 96 well trays to ~90% confluency and allowed to go serum free for 4hrs before stimulation with 10nM or 1nM growth factor. At various time points media was aspirated, ice cold lysis buffer added and solubilised receptor captured on 83-7 coated white 96 well plates. The amount of phosphotyrosine phosphorylation within the captured receptor/substrate complex was obtained using eu-PT100 antibody and measuring bound antibody by time-resolved fluorescence. Phosphorylation data was fitted to an exponential association curve using Prism™ software. The velocity at each point was calculated using the 'k' value and 'MAX' value obtained from the equation of the phosphorylation curve and graphed in Prism™.

A. R-hIR-B.
B. R-hIR-A.
C. R-hIGF-1R.

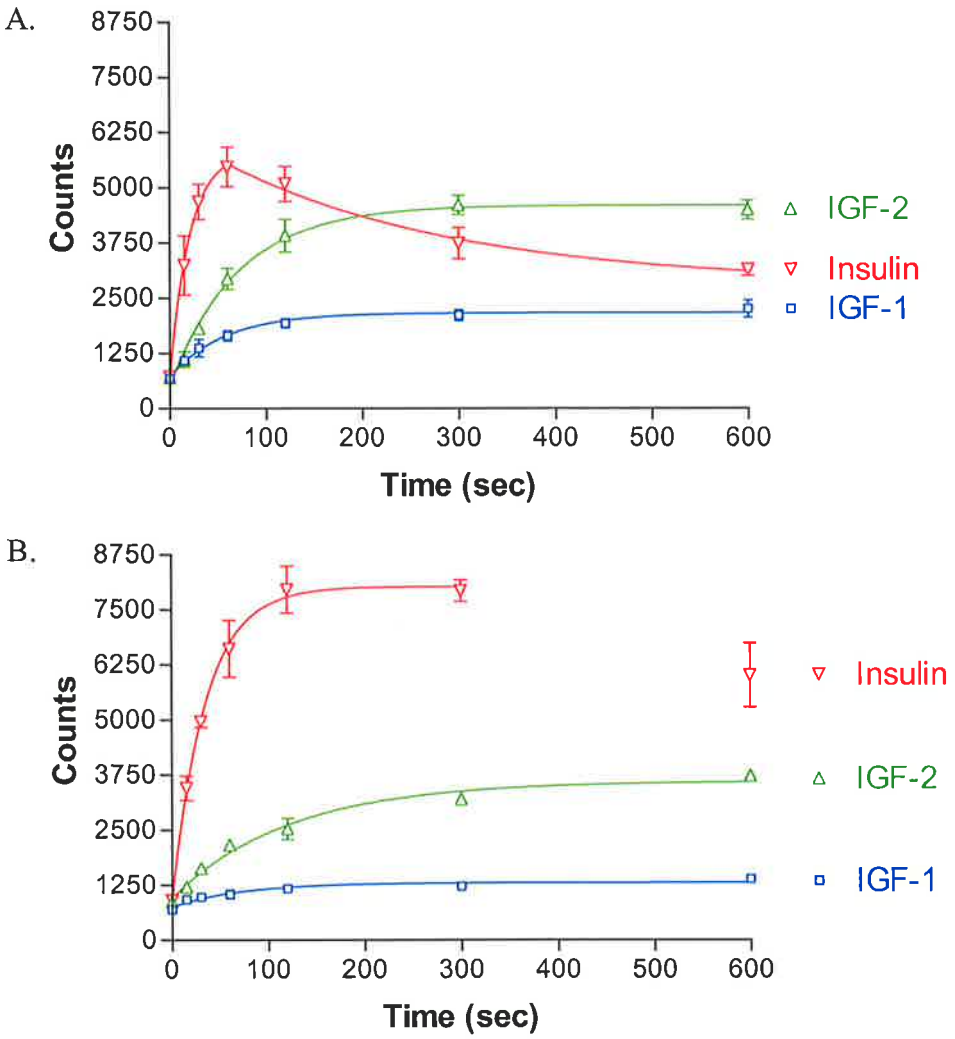


Figure 4.15 Time Course Phosphorylation Assay using 50nM Growth Factor.

Stable cell lines expressing each receptor were grown in 96 well trays to ~90% confluency and allowed to go serum free for 4hrs before stimulation with 50nM growth factor. At various time points media was aspirated, ice cold lysis buffer added and solubilised receptor captured on 83-7 coated white 96 well plates. The amount of phosphotyrosine phosphorylation within the captured receptor/substrate complex was obtained using eu-PT100 antibody and measuring bound antibody by time-resolved fluorescence. Phosphorylation data was fitted to an exponential association curve using Prism™ software. The velocity at each point was calculated using the 'k' value and 'MAX' value obtained from the equation of the phosphorylation curve and graphed in Prism™.

A. R-hIR-B.
B. R-hIR-A.

4.2.10 Proliferation Assays

The competition binding and phosphorylation assays gave a clear indication as to the relative potency of each ligand at a receptor level, so the next step was to assess whether this translated into a similar relationship with a biological end point effect. It is well established that both the IR and IGF-IR can elicit mitogenic responses from a wide variety of cell types and several reports have indicated that IGF-2 was acting through IR-A in this manner.

As an initial characterization to optimize conditions for a proliferation assay, cell lines expressing hIR-A and hIR-B were grown for up to 72 hrs in either 10% FCS or serum free conditions. This would determine whether the absence of serum caused cellular apoptosis and if there was a large enough difference between the two conditions to reliably detect as it was thought to be unlikely that a single growth factor would be able to substitute for full serum conditions. The results indicate that the presence of 10% FCS significantly improves the proliferation of both R-hIR-A and R-hIR-B at a similar rate for each cell line (See Figure 4.16). After 48hrs there was a significant difference between serum and serum free conditions, therefore subsequent assays were performed for at least 48 hrs.

An assessment of the proliferative effect of continual growth factor stimulation on R-hIR-A and R-hIR-B cells grown in serum free medium was made. Cells were seeded into 96 well plates at a concentration of 12, 000 cells/well and allowed to recover overnight in normal medium. After 2.5 hrs of serum deprivation the media was replaced with treatments in triplicate of serum free medium, containing serially diluted growth factor and further incubated for 48 hrs. The number of viable cells was determined by methylene blue staining, with higher absorption at OD_{580} being indicative of more viable cells. The results show that cells expressing either hIR-A or hIR-B can respond proliferatively to growth factor stimulation (See Figure 4.17). In both cell lines insulin elicited the largest increase in proliferation and IGF-1 and IGF-2 displayed similar potencies except at 30nM in hIR-A where IGF-2 has a significantly stronger effect. This data is different to that seen in the literature which suggests that IGF-2 is similarly, if not more potent than insulin acting through hIR-A. An analogous experiment performed using untransfected cells was unable to respond to growth factor stimulation (Data not shown), while another control experiment using IGF-1 stimulation of R-IGF-1R displayed a moderate growth response with an EC50 of 1.63 nM and maximal proliferative response at 10 nM IGF-1 (See Figure 4.18). However in

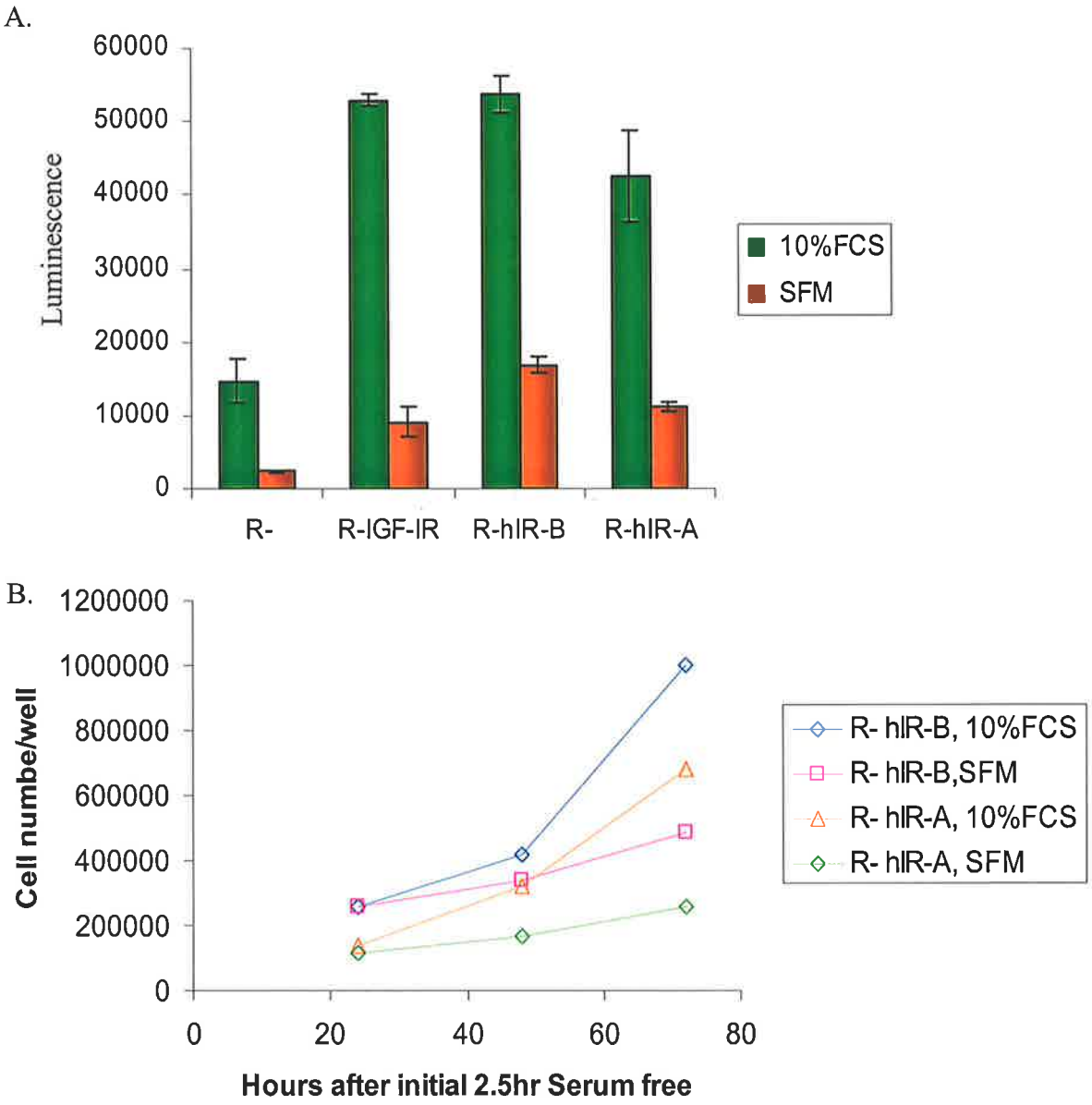
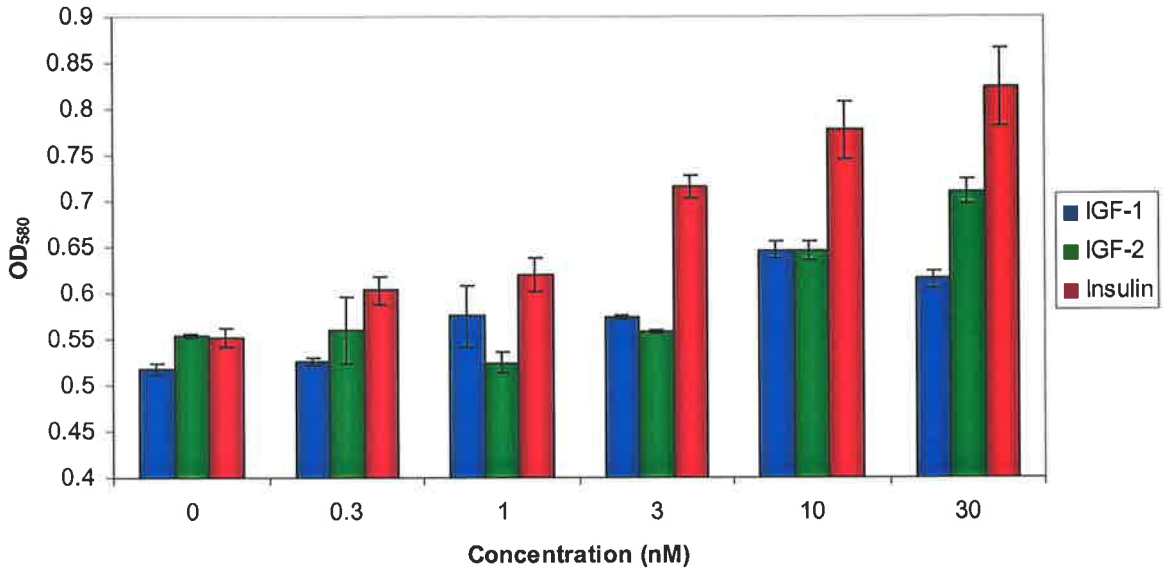


Figure 4.16 Proliferation of hIR-A and hIR-B up to 48 hrs under serum or serum-free conditions.

Cell lines were plated in 96 well plates at a density of 15,000 cells/well and grown overnight before going serum free for 2.5 hrs. Medium was then replaced by serum free medium supplemented with serially diluted growth factor and the cells incubated at 37°C for a further 48 hrs. Proliferation was assessed using the Cell Titre Glo™ kit, measuring ATP levels. The averages and standard deviations of the results from the triplicate treatments were calculated in Microsoft Excel.

A.



B.

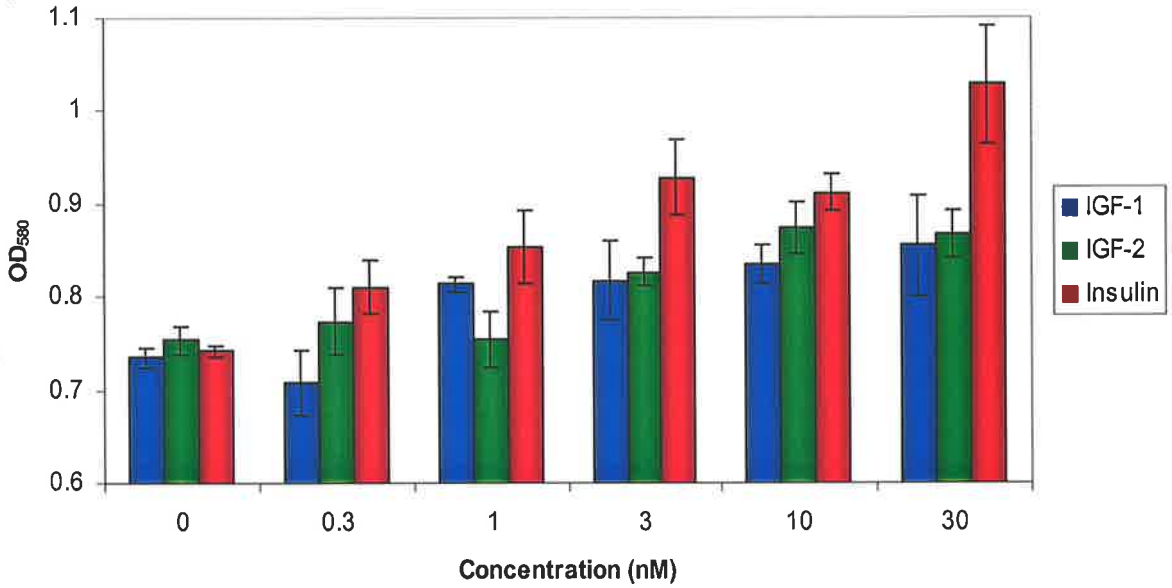


Figure 4.17 Proliferation of hIR-B and hIR-A after 48 hrs in response to continual growth factor stimulation.

Cell lines were plated in 96 well plates at a density of 12,000 cells/well and grown overnight before going serum free for 2.5 hrs. Medium was then replaced by serum free medium supplemented with serially diluted growth factor and the cells incubated at 37°C for a further 48 hrs. Proliferation was assessed by methylene blue staining. The averages and standard deviations of the results from the triplicate treatments were calculated in Microsoft Excel.

A. R-hIR-A
B. R-hIR-B

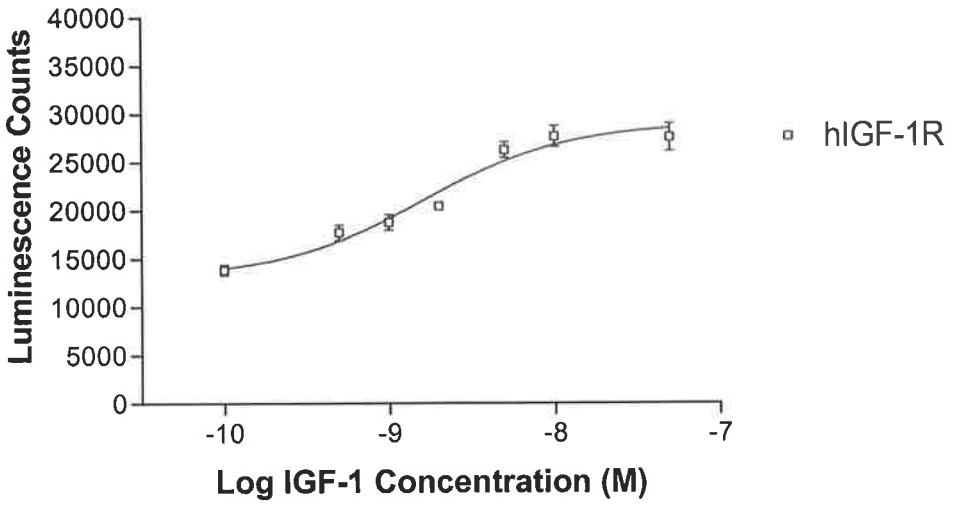


Figure 4.18 Proliferation of R- hIGF-1R 48 hrs in response to IGF-1.

Cell lines were plated in 96 well plates at a density of 15,000 cells/well and grown overnight before going serum free for 2.5 hrs. Medium was then replaced by serum free medium supplemented with serially diluted growth factor and the cells incubated at 37°C for a further 48 hrs. Proliferation was assessed using the Cell Titre Glo™ kit, measuring ATP levels. Data of the results from the triplicate treatments were fitted to an exponential association curve in Prism™.

these experiments, both the competition and phosphorylation assays predict that IGF-2 should not have as great an effect as insulin on either receptor.

4.2.11 Anti-Apoptosis Assays

Activated IGF-1R is well established at being able to inhibit apoptosis triggered by a variety of conditions including growth factor withdrawal and chemotherapeautic agents. It was therefore necessary to examine under which conditions R- cells would undergo apoptosis. Unlike, in previous studies using R- cells, it was seen that this batch did not appear to go into apoptosis when serum was withdrawn, rather only have a slow rate of growth (See Figure 4.16). This lack of observable apoptosis may have been due to the short 48-72hr time points examined. Performing experiments over much longer time periods was considered impractical, as the assays were to be done in 96 well plates. It was observed in preliminary experiments that seeding wells in 96 well plates with numbers less than 8000, did not grow consistently and gave variable results between triplicates. A long experiment would have necessitated a low seeding density and likely would have produced unreliable results. Therefore several drugs, including methotrexate (inhibitor of folate-dependent enzyme dihydrofolate reductase (DHFR)), etoposide (inhibits the ability of topoisomerase II to ligate cleaved DNA molecules), and butyrate (inhibitor of histone deacetylase) were tested for their ability to kill R- cells transfected with IR or IGF-1R.

In numerous preliminary experiments, the Cell titre glo^{TM} assay was used to assess cell numbers after treatment with the apoptotic drugs, however it became clear that under these conditions that the assay gave inconsistent and highly variable results, which did not correlate with the observable number of cells seen under a microscope (Data not shown). This assay quantitates the amount of ATP in a well, as a measure of cell numbers. Since the apoptotic drugs are also known to affect metabolism, the amount of ATP produced by the cell may not correlate normally in the presence of the drugs.

Therefore further experiments were carried out instead by counting viable cells through methylene blue staining. The ability of IGF-2 to protect R- cells transfected with hIR-B, hIR-A, or IGF-1R against pro-apoptotic drugs was tested. Cells were plated overnight in 96 well plates at a density of 12,000 cells/well. After serum starvation for 3.5hrs, medium was replaced with serially diluted methotrexate, etoposide, or butyrate alone or together with 20nM IGF-2. 20nM IGF-2 was chosen due to it being ~3 fold higher than its EC_{50} in the competition assays with both hIR receptors and therefore if it is going to have an effect it should be visible at this concentration. Cells were incubated at 37°C for 48hrs and the viable

cells stained with methylene blue. The results of treatment with methotrexate indicated that there was a dose dependent effect on the cells using this drug, shown by a decreased effect at lower concentrations, however IGF-2 as unable to rescue this effect in fact, it appeared to increase the effect resulting in higher numbers of cell death (See figure 4.19). Etoposide was highly efficient at killing the R- cells and in the range tested did not show a dose-dependent effect. In addition IGF-2 was unable to protect them from apoptosis (See figure 4.20). Butyrate induced apoptosis was seen to be dose-dependent and was partially rescued by 20nM IGF-2 in all the cell lines (See figure 4.21). There did not appear to be a significant difference between the effectiveness of IGF-2 between hIR-B and hIR-A, however IGF-2 had a more potent effect in the IGF-1R. This remains consistent with the previous binding, phosphorylation and proliferation experiments. Due to time constraints further experiments using IGF-1 and insulin were unable to be performed.

Methylene blue staining is unable to give a perfectly accurate picture of the amount of apoptosis occurring, as it cannot detect cells in the initial stages of apoptosis. To follow on from this initial work, future experiments involving the direct measurement of the number of apoptotic cells would be necessary, such as using Annexin-V staining and FACS analysis.

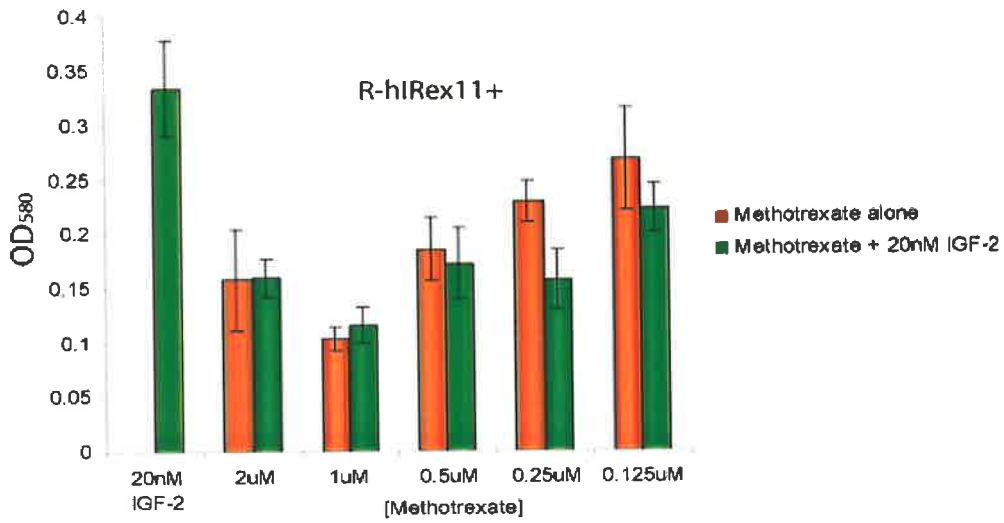
4.2.12 Creation of antibody that can differentiate between the insulin receptor isoforms

An attempt was made to produce a polyclonal antibody that would be able to differentiate between the hIR-A and hIR-B, as previously the only method available to distinguish between the two was by PCR. The method chosen was to inject rabbits with a peptide consisting of the residues encoded by exon11 (RKTSSGTGAEDPC), which was acetylated on the C-terminus and N-terminally conjugated to a diphtheria toxoid moiety to improve the immune response. The injections were performed by Joe Wrin (Biochemistry Department, University of Adelaide) on two rabbits and serum bleeds were taken before and 6 weeks and 12 weeks post injection. A further injection of peptide was made 3 months after the first bleed and another serum sample taken after 8 weeks. The rabbits were terminally bled 2 months after the second sample was taken.

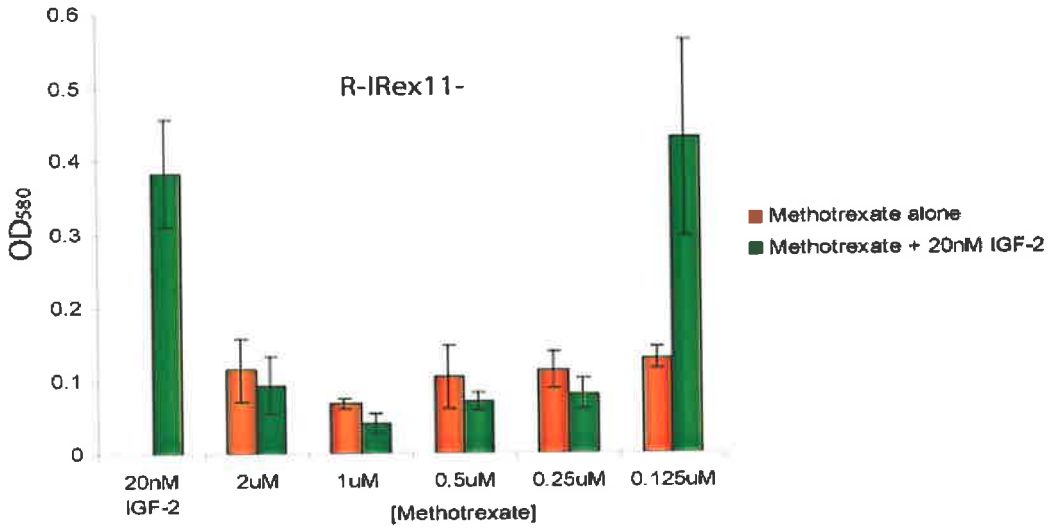
4.2.12.1 Characterisation of the polyclonal anti-exon11 serum

The initial characterisation was to determine whether the antibodies in the serum from the injected rabbits and not the pre-immune serum were able to recognise the exon11 peptide. Exon11 peptide without the diphtheria toxoid conjugate was coated on a clear 96 well plate at a concentration of 50 ng/well. Various dilutions of the serum samples from 6 weeks after the

A.



B.



C.

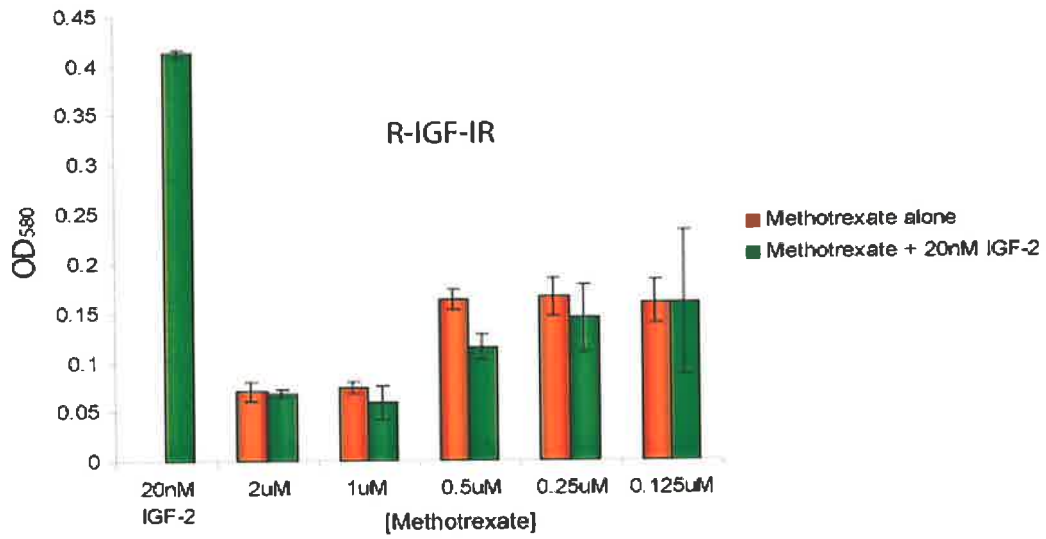
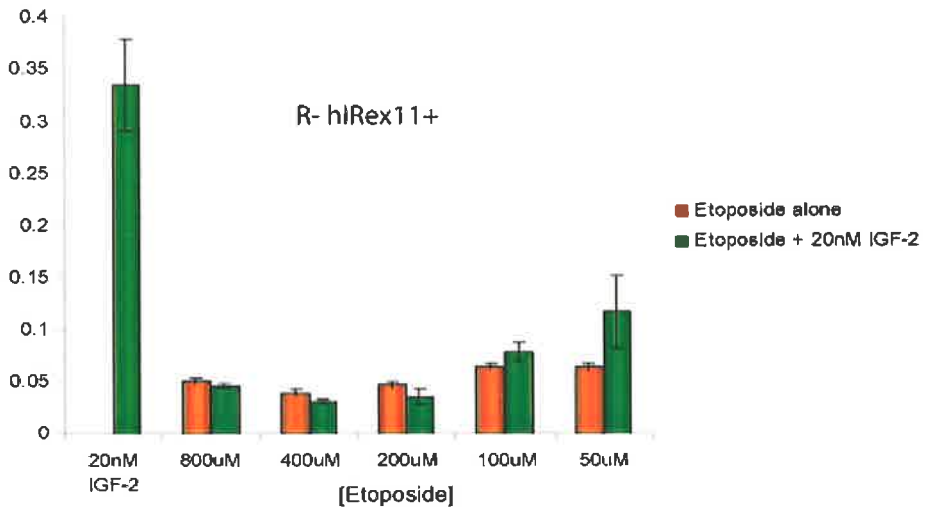


Figure 4.19 Anti-apoptotic effect of IGF-2 co-treated with methotrexate.

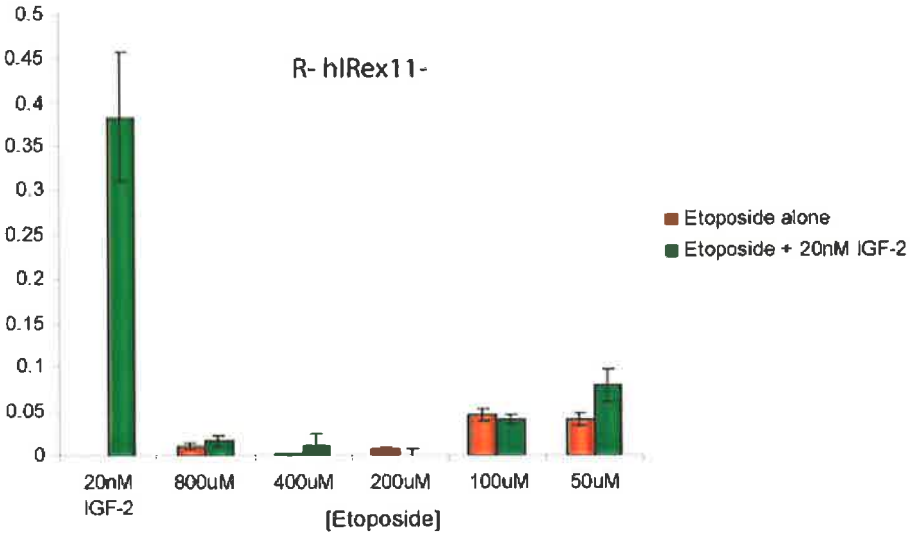
Cell lines were plated in 96 well plates at a density of 12,000 cells/well and grown overnight before going serum free for 2.5 hrs. Medium was then replaced by serum free medium supplemented with 20nM IGF-2 with or without methotrexate and the cells incubated at 37°C for a further 48 hrs. Proliferation was assessed by methylene blue staining. The averages and standard deviations of the results from the triplicate treatments were calculated in Microsoft Excel.

A. R-hIR-B B. R-hIR-A C. R-hIGF-1R

A.



B.



C.

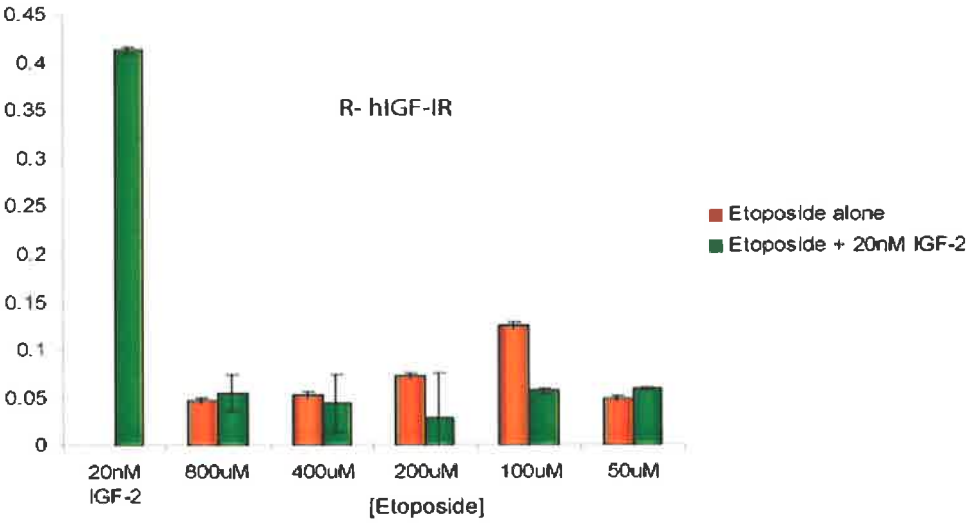
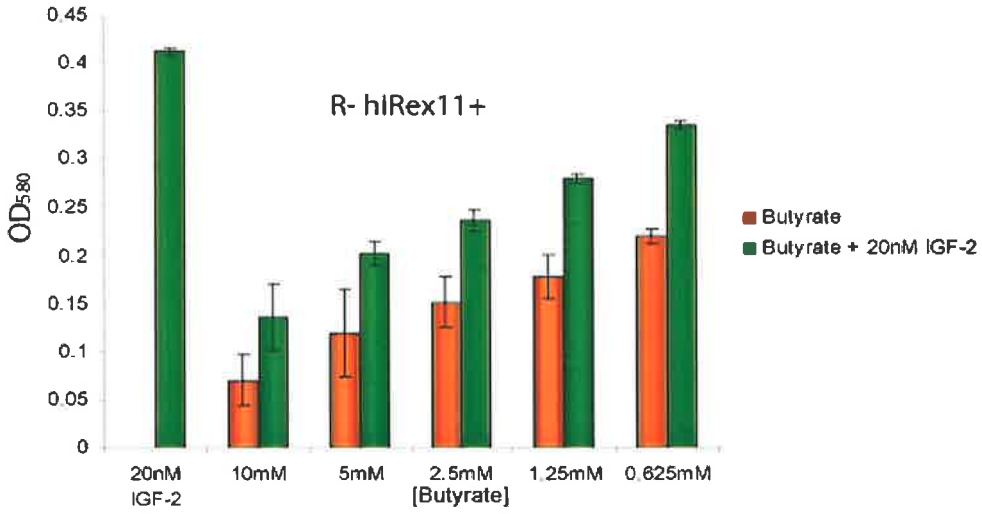


Figure 4.20 Anti-apoptotic effect of IGF-2 co-treated with Etoposide.

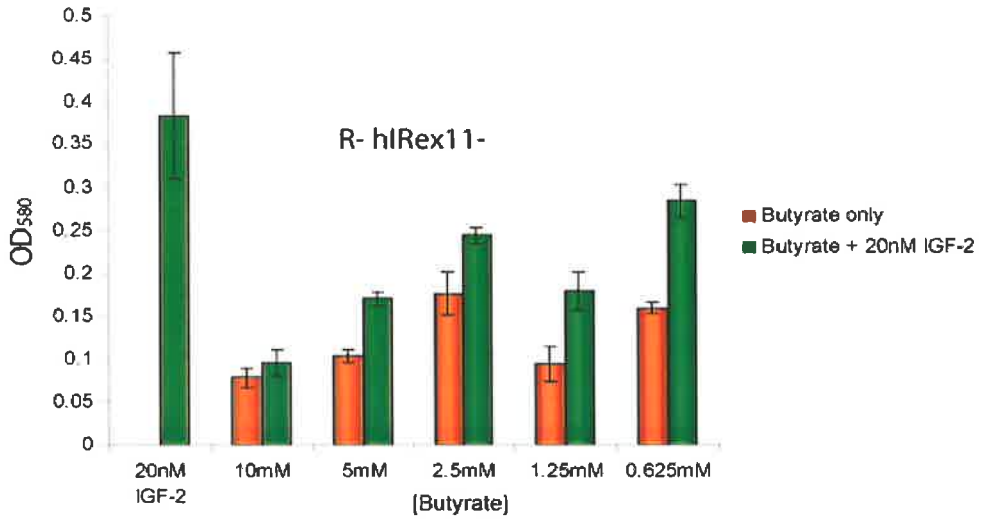
Cell lines were plated in 96 well plates at a density of 12,000 cells/well and grown overnight before going serum free for 2.5 hrs. Medium was then replaced by serum free medium supplemented with 20nM IGF-2 with or without etoposide and the cells incubated at 37°C for a further 48 hrs. Proliferation was assessed by methylene blue staining. The averages and standard deviations of the results from the triplicate treatments were calculated in Microsoft Excel.

A. R-hIR-B B. R-hIR-A C. R-hIGF-1R

A.



B.



C.

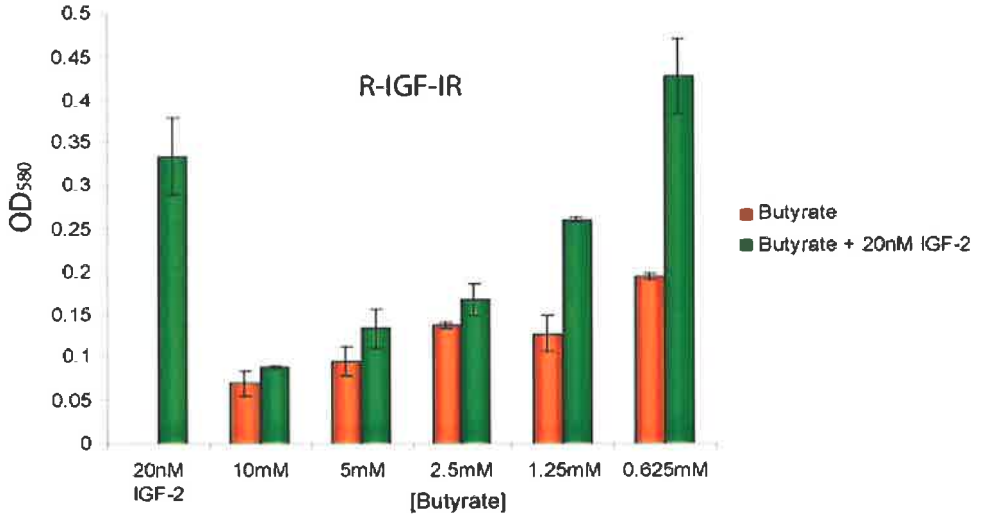


Figure 4.21 Anti-apoptotic effect of IGF-2 co-treated with Butyrate.

Cell lines were plated in 96 well plates at a density of 12,000 cells/well and grown overnight before going serum free for 2.5 hrs. Medium was then replaced by serum free medium supplemented with 20nM IGF-2 with or without Butyrate and the cells incubated at 37°C for a further 48 hrs. Proliferation was assessed by methylene blue staining. The averages and standard deviations of the results from the triplicate treatments were calculated in Microsoft Excel.

A. R-hIR-B B. R-hIR-A C. R-hIGF-1R

first injection were made from 1/10 to 1/50, 000, and were incubated in each well for 2hrs at room temperature. An alkaline phosphatase (AP) conjugated goat anti-rabbit secondary antibody was used to detect any bound antibody from the serum and was detected after an AP substrate was added by measuring the optical density at 405 nm. The graph of the results show that the serum from both rabbits post injection contained antibodies that bound the exon 11 peptide and that these antibodies were not detectable in the pre-immune serum (See Figure 4.22). It appeared that rabbit A contained more exon11 specific antibodies as it reached the maximum distinguishable range by 1/100 dilution whereas the maximum for rabbit B occurred at around 1/10 dilution.

A comparison of the exon11 binding capabilities of several serum bleeds from each rabbit was tested in a similar assay, using a standard concentration 1/20. These results indicated that there was little difference between each bleed and between the two animals (See Figure 4.22). There was no consistent trend between subsequent samples and although a further injection of the peptide was made between serum bleeds 2 and 3, there was not the expected increase in the amount of exon11 binding antibodies. This may suggest that the exon11 peptide sequence is not particularly immunogenic even with the presence of the diphtheria toxoid amplifying the immune response.

The specificity of the antibodies in the serum to recognize the exon11 sequence and not another peptide sequence of similar length was tested by comparing binding to the exon11 peptide and an unrelated peptide from the SH3 linker region of mouse Tec Kinase. It is clear from the graph that the antibodies are specific for the exon11 peptide and only show a low level of reactivity with the linker peptide at a 1/20 dilution (See Figure 4.22).

The next step was to determine whether the serum could differentiate between the two isoforms of the insulin receptor. Soluble ectodomain forms of each isoform were captured using the antibody 83-7 and incubated with either a 1/50 dilution of serum or the anti-receptor antibody 83-14. The graph using the hIR-B ectodomain (exon11+) shows a concurrent parallel increase in absorbance with increasing amounts of receptor for each treatment, indicating that the antibodies in the serum are recognizing the receptor and also suggests that the stoicheometry of both the antibody-receptor interactions are the same (See Figure 4.23). This could mean that two anti-exon11 antibodies are able to bind to a single receptor on each half of the receptor in the same way that 83-14 can bind on opposite sides of the heterotetrameric receptor (Tulloch *et al.*, 1999). The results using the hIR-A ectodomain show that the serum antibodies cannot bind this receptor as there is no increase in absorbance

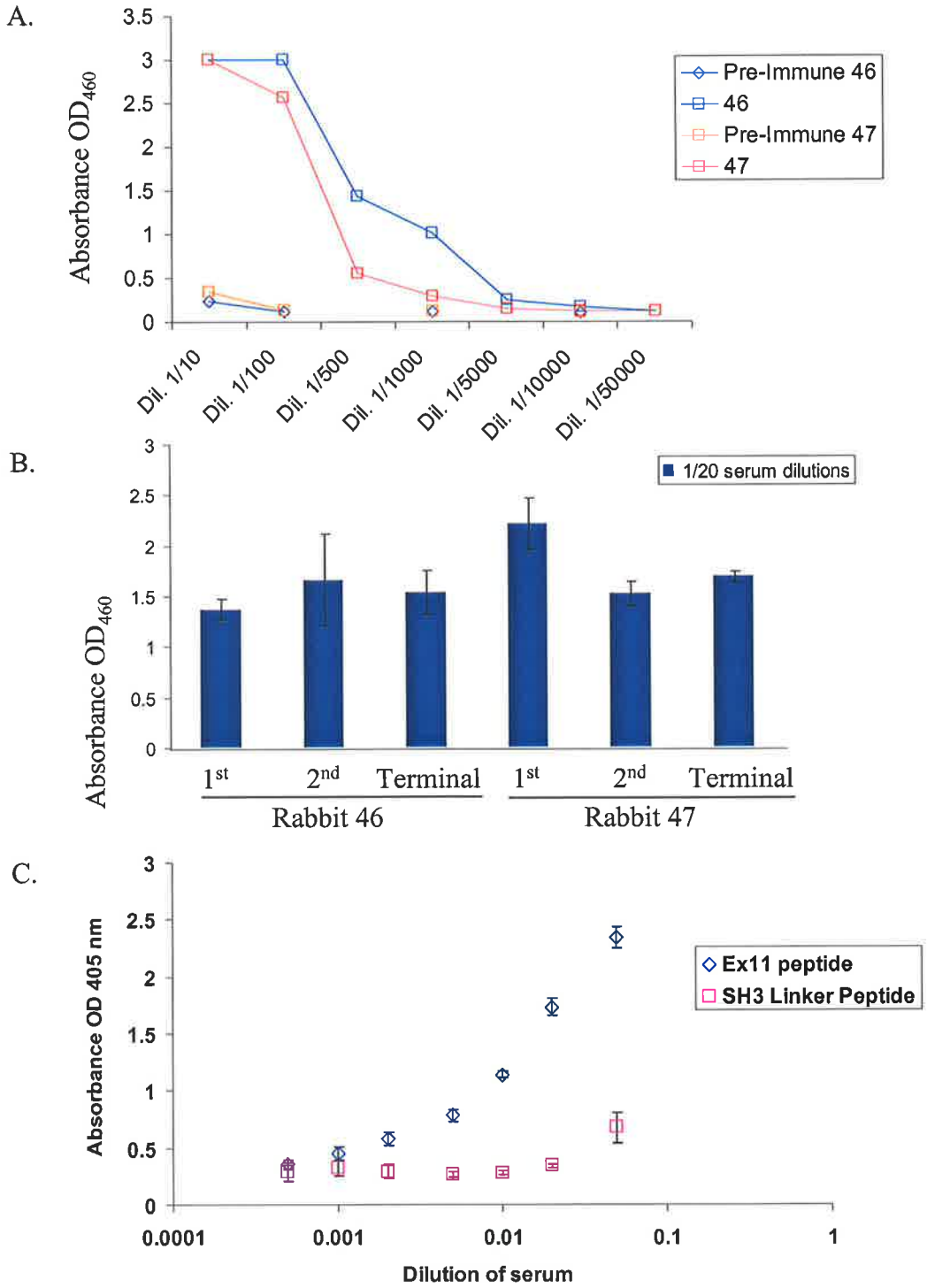


Figure 4.22 Testing the rabbit serum for containing antibodies capable of binding exon11 peptide using Enzyme-Linked Immuno Assays.

Clear 96 well plates were coated with 50ng/well Exon11 Peptide or the unrelated Tec SH3 domain linker peptide and serum diluted in 1xPBS used as the primary antibody. Bound antibody was detected using AP conjugated Goat α-rabbit Antibody.

A. Comparison of pre-immune serum to serum after injection by the diptheria toxoid conjugated exon 11 peptide.
B. Comparison of different serum bleeds, after 1st injection, after 2nd injection, and terminal bleed.
C. Comparison of the specificity of the serum for binding Exon 11 peptide and a Tec SH3 linker domain peptide.

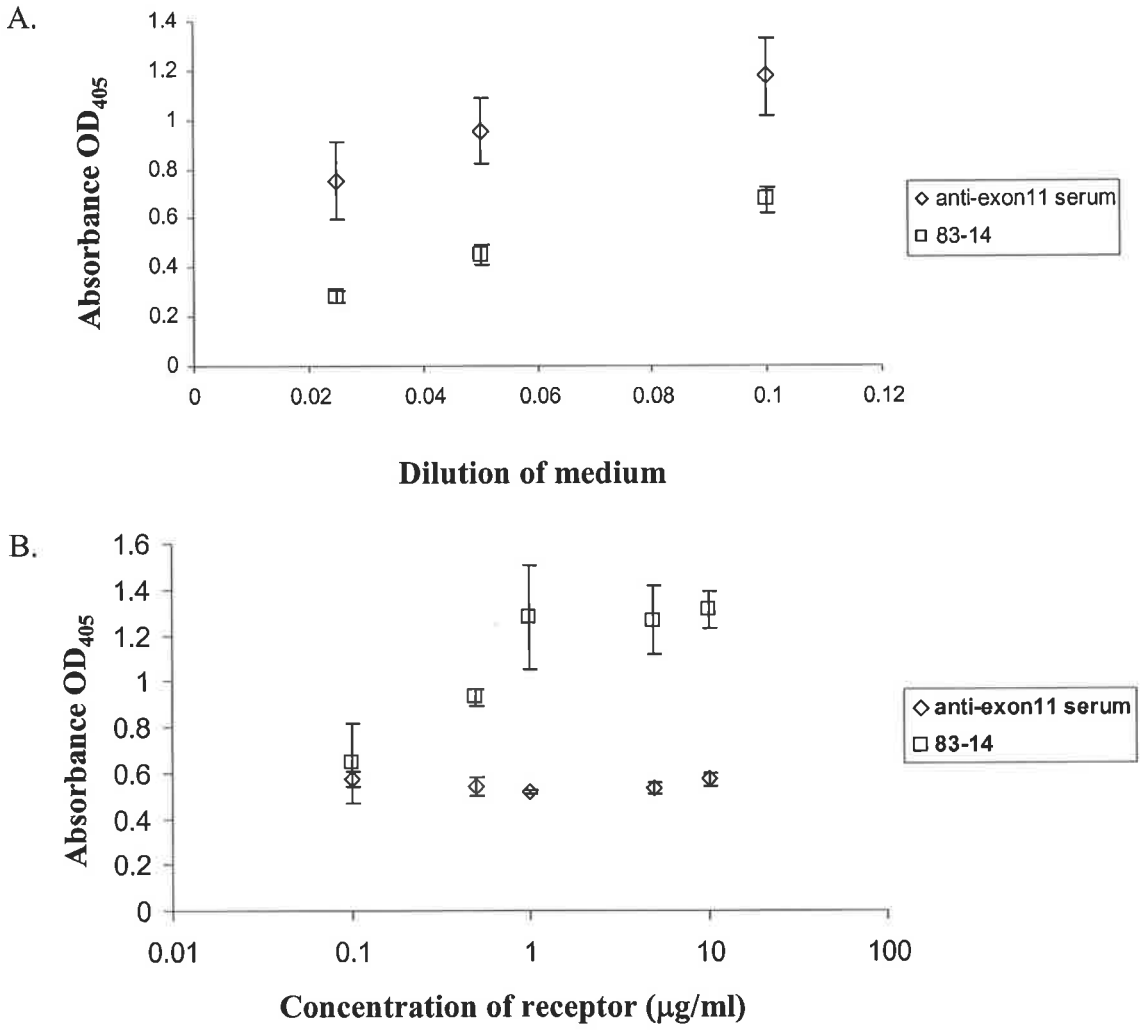


Figure 4.23 Enzyme-Linked Immuno Assay using recombinant hIR.

ELISA assays were performed using either purifed hIR-A ED or media from CT5 cells expressing hIR-B ED. Insulin Receptor was captured using 83-7 and either a specific concentration of hIRA ED or serially diluted media of hIR-B ED used. Primary antibodies used were either those in the exon11 serum or 83-14. Bound antibody was detected using an AP-conjugated anti-rabbit monoclonal antibody.

A. hIR-B ED
B. hIR-A ED

regardless of the concentration of the receptor (See Figure 4.23). In a further experiment the ability of the serum to bind to full length hIR-B receptor was tested by using triton X-100 detergent solubilised receptor from the two R- cell lines expressing each isoform. A receptor null control was also used containing only lysis buffer as it had been previously seen that some detergents (including triton X-100) would give positive readings above background in this assay. The graphs show that hIR-B has significantly higher absorbance readings at every serum dilution over hIR-A which itself does not stray from the background null receptor control samples (See Figure 4.24).

It can therefore be concluded that the rabbit polyclonal serum contains antibodies that specifically recognize the exon11 sequence and can distinguish between the two insulin receptor isoforms presumably through binding to the exon11 sequence in hIR-B. The fact that antibodies are able to bind to exon11 suggests that this sequence lies on the exterior of the receptor even though this may be within the ligand-binding pocket.

The polyclonal serum was tested for use as a primary antibody in FACS analysis but even the profiles obtained at the lowest dilution of serum were indistinguishable from the control antibody, suggesting that the antibodies are unable to bind to membrane bound receptor or that they are at a far too low concentration for this experiment to work (Data not shown).

4.3 SUMMARY AND CONCLUSION

In this project, comparisons of the effectiveness of insulin, IGF-1, and IGF-2 were made, focusing on their ability to bind, activate, and trigger biological responses from the two human isoforms of the Insulin receptor, hIR-A, and hIR-B. Full length and recombinant hIR isoforms were examined in competition binding assays using europium labeled insulin. The overall results confirmed that IGF-2 and insulin had closer affinities in hIR-A than hIR-B. The ratios of binding affinities between the different growth factors was consistent with the published literature using iodine labeled insulin, however the raw values were different. This was most likely due to the use of the europium-labeled insulin where the position of the europium chelate may be affecting mode of binding to the receptor, through disrupting the conformational change at the B-chain C-terminus and A-chain N-terminus. These results were in contrast to those using europium labeled IGF-1 on the IGF-1R, which yielded very similar affinities to those already published. Within a single receptor the phosphorylation assays were consistent with the binding data, however the hIR-B isoform appeared to be more strongly phoshorylated in response to growth factor. The cellular response to growth factor stimulation

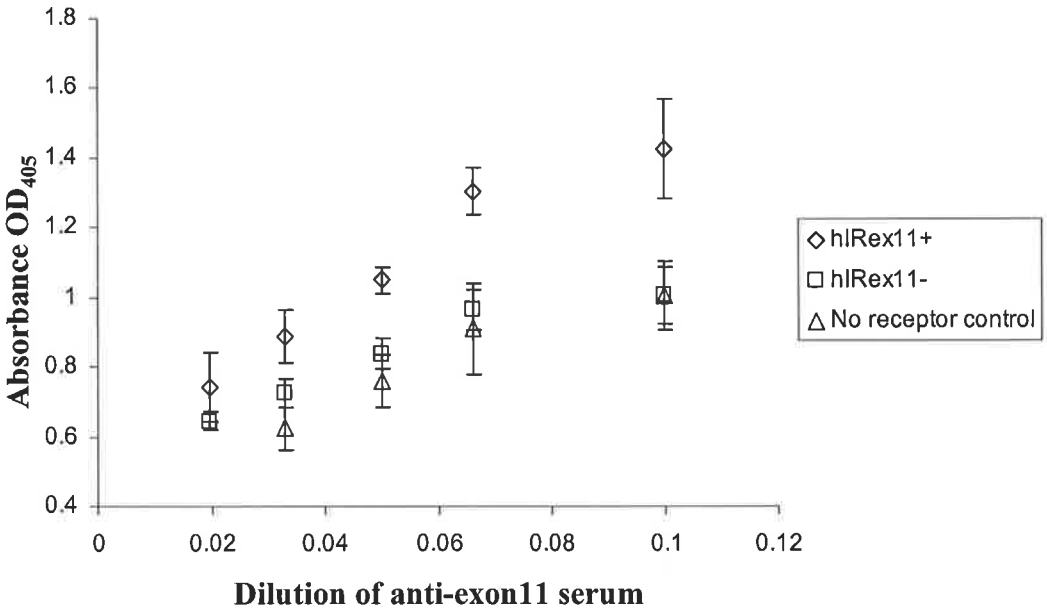


Figure 4.24 Enzyme-Linked Immuno Assay using full length hIR.

ELISA assays were performed using full length receptor isolated from either R-hIR-A or R-hIR-B, in addition to a lysis buffer control. Insulin Receptor was captured using 83-7, and serum used as the primary antibody. Bound antibody was detected using an AP-conjugated anti-rabbit monoclonal antibody.

correlated well with binding and phosphorylation data and did not suggest that IGF-2 was able to stimulate proliferation better than insulin. In order to examine any differences in signalling invidual proteins activated by the IR would need to be examined, such as PI3K and PKB. These studies have shown that the cell lines will be a useful reagent for future studies of this system.

CHAPTER 5

C-Terminal Mutants Of The IGF-1R

CHAPTER 5 C-TERMINAL MUTANTS OF THE IGF-1R

5.1 INTRODUCTION

The IGF-1R is able to mediate a diverse range of biological responses upon binding its cognate ligands IGF-1 and IGF-2. These effects appear to be context dependent and vary considerably between different cell lines. Stimulation with ligand may result in cellular proliferation, differentiation, protection from apoptosis and transformation. The effects of IGF-1R are potentiated by the activation of its intracellular tyrosine kinase domain, which results in cross-phosphorylation of each receptor half, and further recruitment and activation of assorted receptor substrates. The regions of the receptor involved in binding substrates and regulators are predominantly found in the juxtamembrane domain and the C-terminal domain, which contain several potential tyrosine and serine phosphorylation target sequences. The structure of these two domains is unknown, although they are likely to be flexible to allow movement of the phosphorylation sites in and out of the catalytic pocket of the kinase domain and also to accommodate the large number and size of the receptor substrates.

Due to the great similarity between the hIGF-1R and the hIR researchers have been interested in identifying the mechanistic way in which the different biological responses are elicited. It seems likely that different ligand affinities, modes of binding, extent and duration of receptor phosphorylation, and receptor substrate specificity, all play a role in separating the biological effect of co-expressed receptors. Studies to identify unique pathways emanating from the hIGF-IR have focused on the unique tyrosine residues (potential phosphorylation targets) and other unique residues in the C-terminal domain of the hIGF-1R. These include Tyr^{1250}, Tyr^{1251}, $Ser^{1280\text{-}1283}$, His^{1293}, and Lys^{1294}(discussed in section 1.3.2.2).With the recent advances in DNA microarray technologies several studies have looked at the differences in gene expression between hIR expressing cells stimulated with insulin and hIGF-IR expressing cells stimulated with IGF-1. The results indicated that quite distinct changes occur, and although the mechanism of regulation of the expression of these target genes can be correlated to the receptor, it still does not resolve how the two receptors can achieve separate effects (Mulligan *et al.*, 2002). The majority of evidence to date indicates that the vast majority of substrates are able to bind and be activated by either the IGF-1R or IR in *in vitro* systems.

5.1.1 Project Summary and Aims

The aim of this study was to create stable transfected R- cell lines expressing single and multiple mutations of residues involved in mediating signaling downstream of the hIGF-1R and 'double' mutants including a common Y950F mutant. Since it is clear that there is some redundancy and overlap in the origin of major biological responses from the receptor it was reasoned that inhibition of binding of IRS proteins and Shc, combined with mutations in unique residues would produce receptors that were functionally incapable of activating specific pathways.

AIMS:
a) Compare the amount of ligand triggered receptor phosphorylation between mutants.
b) Compare the ligand activated proliferative potential of the mutants.

5.2 RESULTS AND DISCUSSION

5.2.1 Construction of hIGF-IR Cytoplasmic Mutant Plasmids

The initial construction of the single amino acid mutations within the cytoplasmic region of the IGF-IR was made by Sven Delaney as part of a separate Honours project (See Figure 5.1). A 635bp EcoRV to BamH1 IGF-IR fragment was cloned into pBluescript and site-directed mutagenesis (Stratagene) used to introduce the Y1250F, Y1251F, S1283A, and H1293F/K1294R mutations (See section 2.2.9 for primer sequences). These mutant fragments were sub-cloned into a 961bp BamH1/HindIII IGF-IR fragment containing the wildtype 635bp sequence. This was done to avoid an additional EcoRV site in the full-length sequence that would complicate the cloning procedure. Mutant 961bp fragments were then swapped by BamH1/HindIII digestion into either full length wild-type or Y950F constructs in pSP72-S/P (Sma-PvuII fragment removed) to create hIGF-1R (Y1250F), (Y1250F/Y950F), (Y1251F), (Y1251F/Y950F), (S1283A), (S1283A/Y950F), (H1293F/K1294R), and (H1293F/K1294R/Y950F). Confirmation of the mutations and all subsequent work was undertaken as part of this thesis. The mutations were confirmed by cycle sequencing from pSP72-S/P transformants (See Figure 5.2) and correct constructs were cloned into the mammalian expression vector pEFIRES-neo by EcoR1 digestion and correctly oriented clones identified by restriction endonuclease digestion using Xho1 (See Figure 5.3). Correct clones showed two bands, while wrongly oriented clones displayed 3 bands.

5.2.2 Creation of Stable Cell Lines Expressing Mutants

Stable cell lines expressing each of the mutant receptors were made based on the need to have a population of cells that all expressed a high, reproducible number of receptors. Transient transfection of receptor constructs, although faster to produce, yielded a much lower level of receptor expression in terms of both the percentage of receptor expressing cells and the relative amount of receptor expressed by individual cells. Since it was possible that some of the effects of the mutants would be subtle it was also thought that having every cell expressing the mutant receptor would increase the chances of being able to detect variation from the wild-type receptor expressing cells and the untransfected cell line. R- cells (A kind gift from Renato Baserga, Philadelphia), a cell line derived from the embryonic fibroblasts of a IGF-IR KO mouse, were chosen as the model cell line due to their lack of wild-type receptors and also because they were characterised in the literature as being unresponsive to IGFs and insulin unless exogenous receptor was introduced.

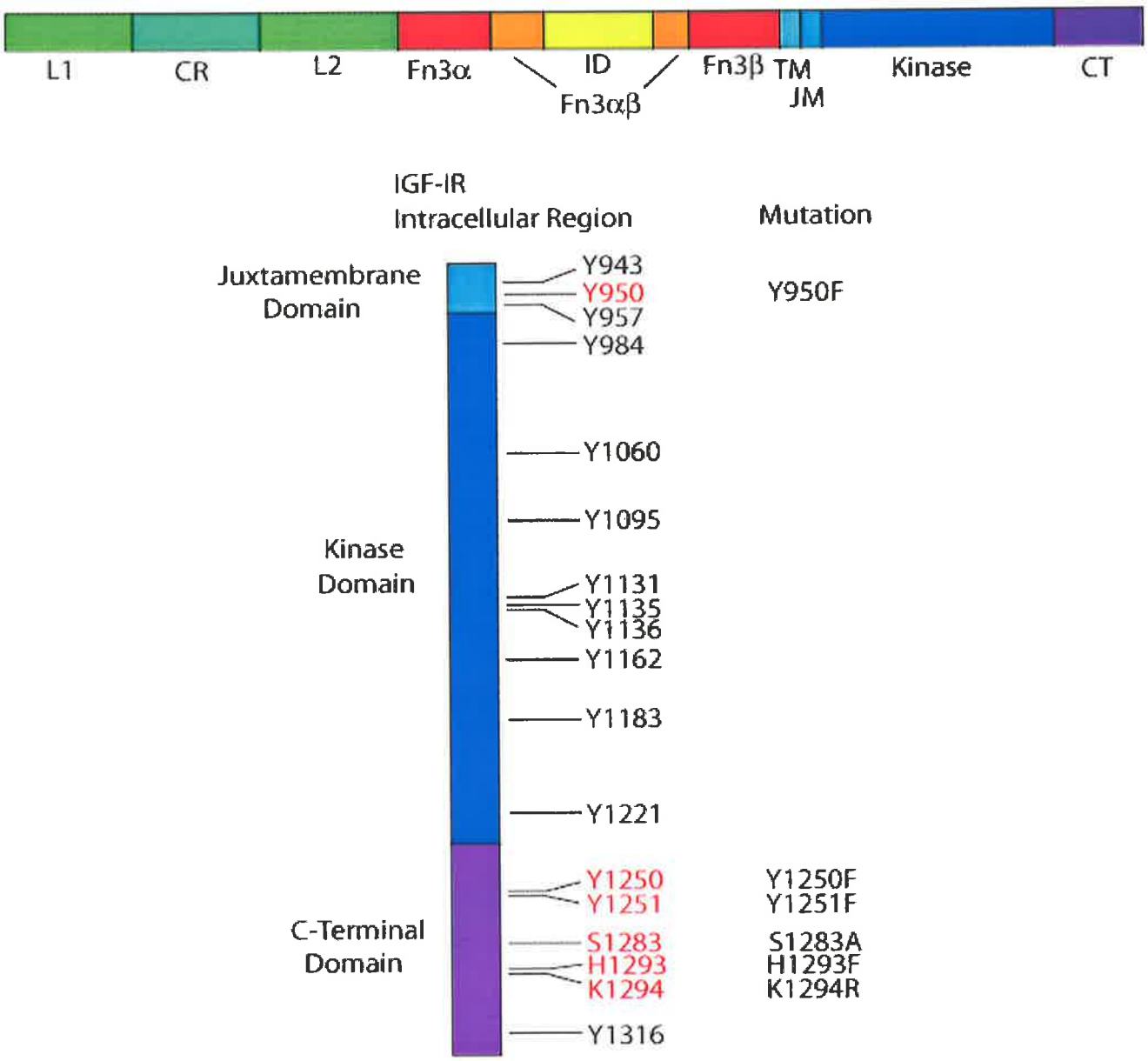


Figure 5.1 Schematic representation of the domains of IGF-1R and the relative position of residues mutated in this study

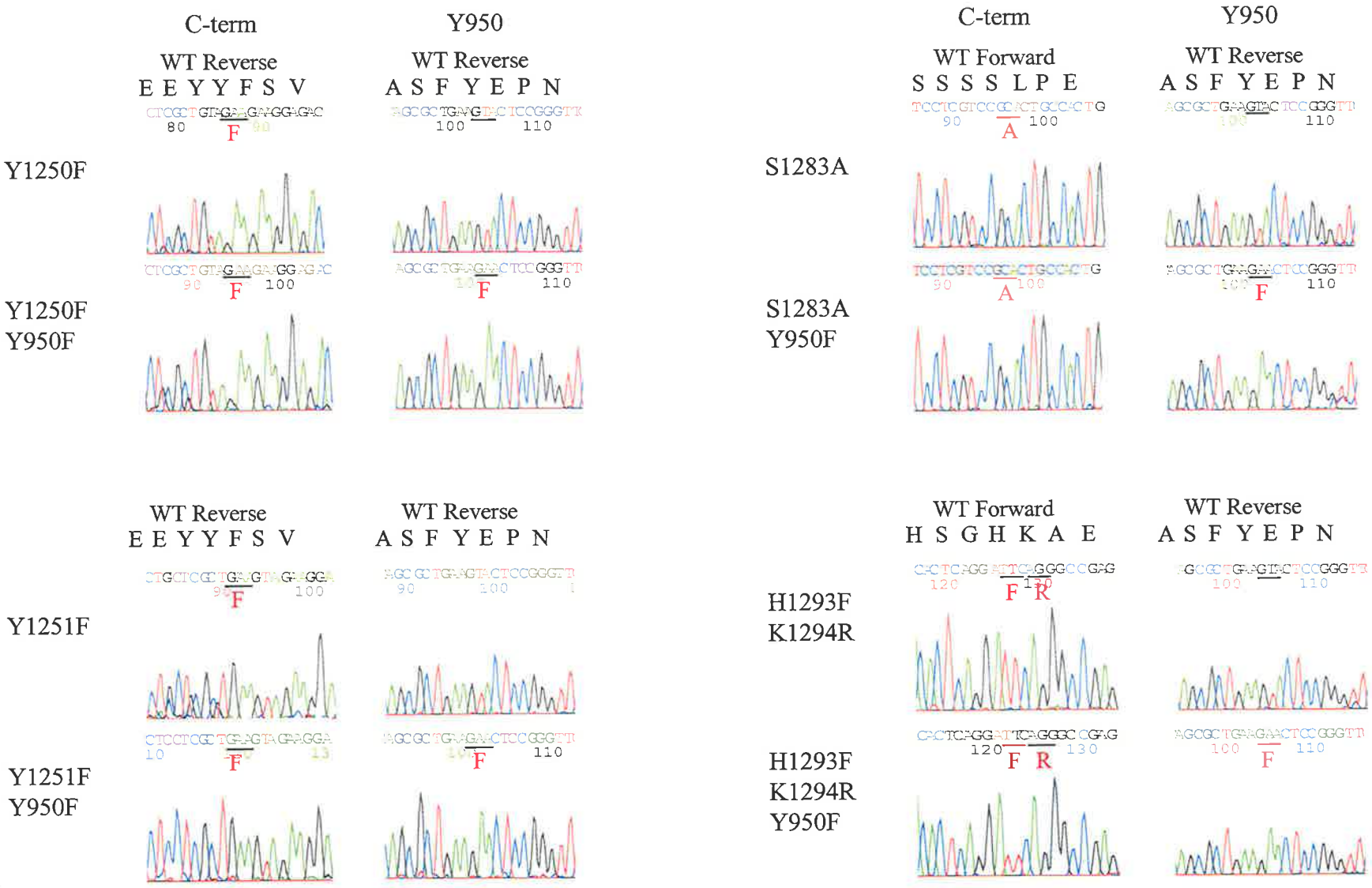


Figure 5.2 Sequencing of IGF-IR mutants

Plasmid DNA from putative IGF-IR mutants was produced using the UltracleanTM miniprep kit and used in a cycle sequencing PCR reaction with the IGF-IRcterm primer and the seq950 primer. Sequencing traces are shown and the position of the mutated residue underlined and compared to the wild-type amino-acid sequence.

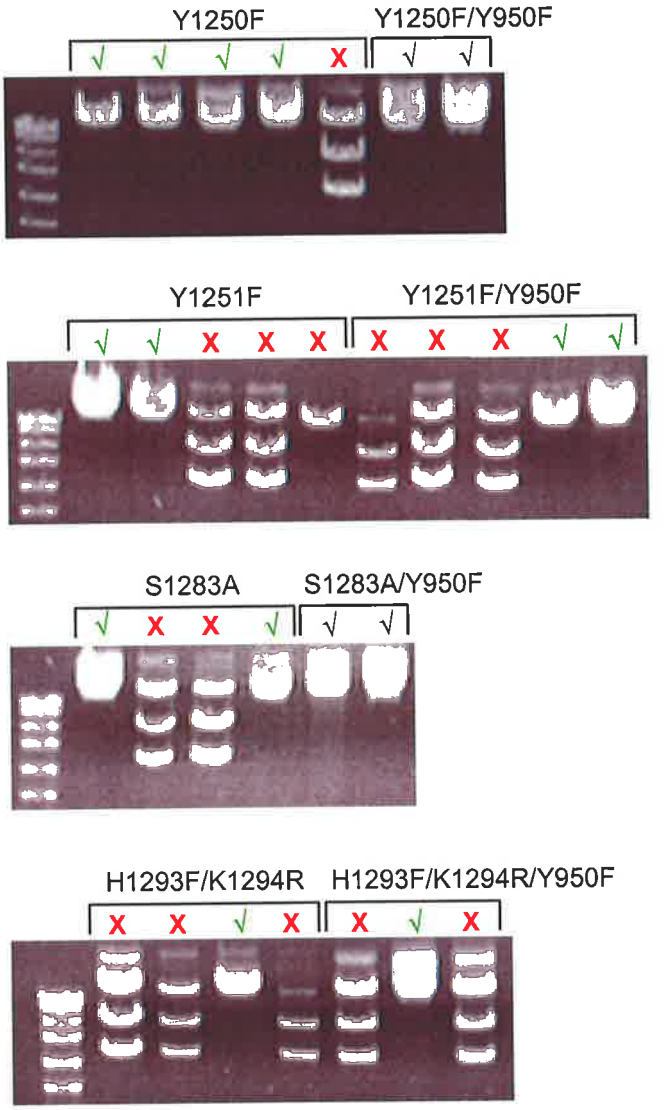


Figure 5.3 Identification of correctly oriented IGF-IR mutant clones by restriction digestion analysis

Plasmid DNA from transformants containing pEFneo:IGF-IR mutants were digested with XhoI for 2hrs at 37°C and run on a 1% agarose gel.

Plasmid constructs containing the mutant IGF-1Rs were prepared using the Ultraclean™ miniprep kit in order to obtain a purity suitable for tissue culture transfections. Subsequently, plasmid purity was tested and the concentration quantitated by UV spectroscopy. Transfections were then performed on R- cells grown overnight to a confluency of ~50% using Lipofectamine+™ and 4μg of plasmid DNA for each construct (See section 2.3.3.3). After a 3hr transfection, cells were allowed to recover overnight in DMEM + 10% FCS before selection with G418 (Genetecin: Gibco/BRL) at a final concentration of 500μg/ml. Media was replaced every 3-4 days for 2 weeks and drug resistant cells screened for IGF-IR expression by FACS analysis using the monoclonal anti-IGF-IR antibody 16-13 (See Tables 5.1 and Figure 5.4). The FACS profiles of cells fluorescing in the FITC range indicated that an average of 60% of the various populations were fluorescing above 10^1 units in the 16-13 labelled samples compared to an average of ~2 units in the control IgG samples. This indicated that a majority of the cells were expressing hIGF-1R. It also indicated that the receptors were likely to be in a correct structural confirmation, at least in the L1 and CR domains as the detecting antibody 16-13 is conformationally dependent. In addition, in later assays another conformationally dependent antibody 24-31 gave a positive result.

Having established that populations from each transfected cell line were expressing receptor, cells were re-cultured and IGF-IR expressing cells subsequently isolated using a FACS cell sorter (IMVS/Hanson centre). As shown in figure 5.4 cells fluorescing in the middle range of the receptor expressing population were selected. This population was chosen because when sorting and selecting for hIR expressing cells, those taken from the highly fluorescent portion of the population, had poor viability after plating. The poor viability may have been due to the smaller numbers of cells isolated, which was previously shown to reduce plating recovery (See 4.2.10). In addition a greater number of cells were selected to further increase the chances of a successful recovery from the FACS procedure. Selection from 500μg/ml G418 was maintained throughout the subsequent passaging process. Due to the increased number of cells selected, plates reached confluency 1 week post-sorting.

A final FACS analysis was then performed on each cell line to confirm that the sorting procedure was a success (See Tables 5.2-5.5 and Figures 5.5-5.8). The results showed that in all cell lines except R-IGF-1R (S1283A) there was an increase in the percentage of the population fluorescing above 10^1 units. The average percentage of fluorescing cells across all lines increased from 60% to 79% and in the majority of cases the shape of the profiles became sharper indicating a more homogenous population of cells expressing a similar number of receptors. The mean FL1-H value of the receptor expressing cells from the sorted cell lines

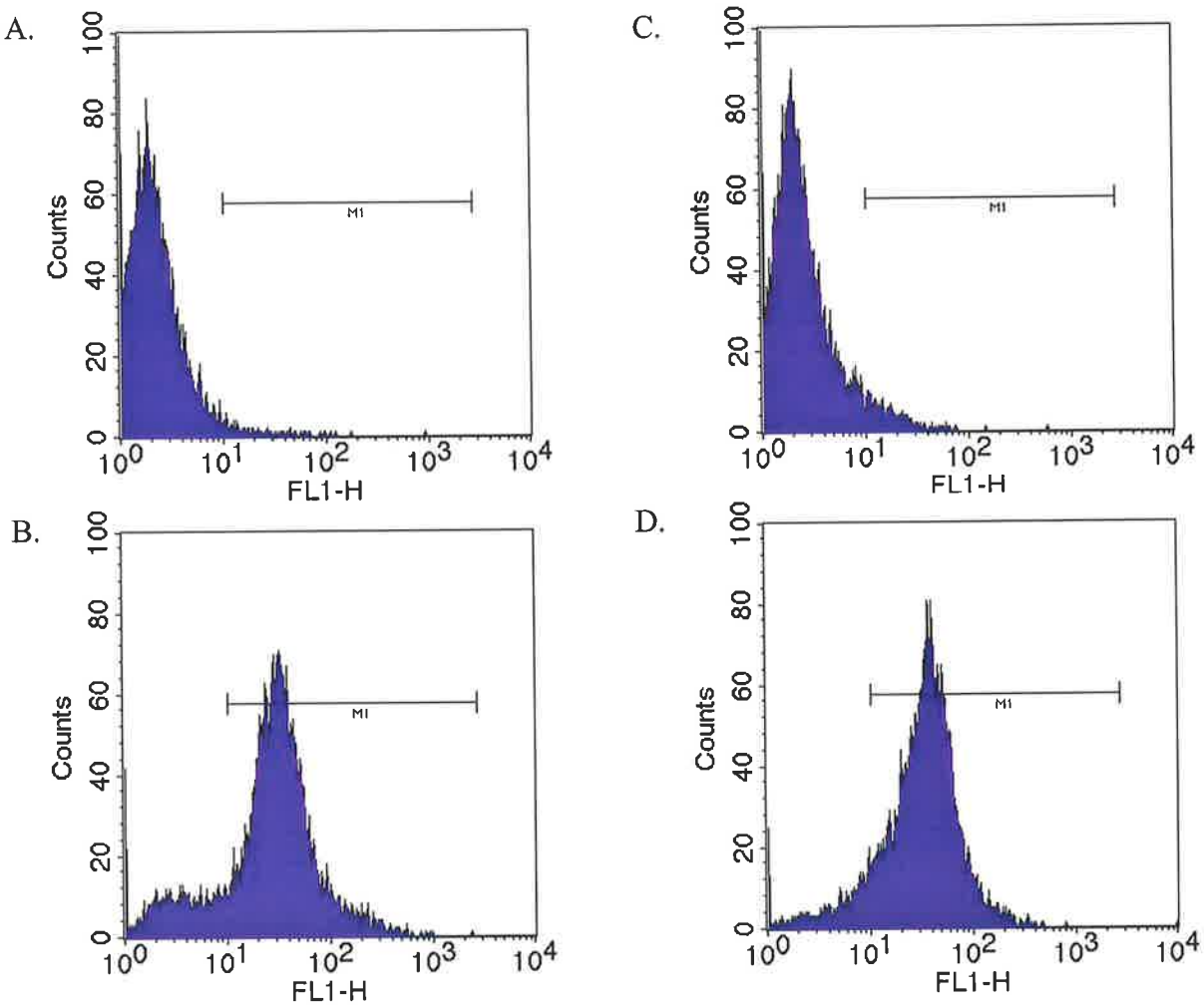


Figure 5.4 Comparison of R-IGF-1R (WT) cells before and after single cell sorting

Cells were grown to ≤90% confluency, trypsinised and incubated with the mouse monoclonal anti-IGF-1R antibody 16-13 or a mouse IgG control. Cells were subsequently labeled with a FITC conjugated anti-mouse monoclonal antibody, fixed with paraformaldehyde and fluoresencing cells analysed on a FACS cell sorter.

A. Before single cell sorting R-IGF-1R (WT) : IgG
B. Before single cell sorting R-IGF-1R (WT) : 16-13
C. After single cell sorting R-IGF-1R (WT) : IgG
D. After single cell sorting R-IGF-1R (WT) : 16-13

Cell Line	% Expressing Cells (M1)	Mean FL1-H
A. R-:IGF-IR (WT) IgG Control	1.48%	2
B. R-:IGF-IR (WT) 16-13	83.77%	30
C. R-:IGF-IR (WT) IgG Control	4.84%	2
D. R-:IGF-IR (WT) 16-13	90.31%	40

Table 5.1 Comparison of the % expressing cells and the mean FL1-H value of the major peak

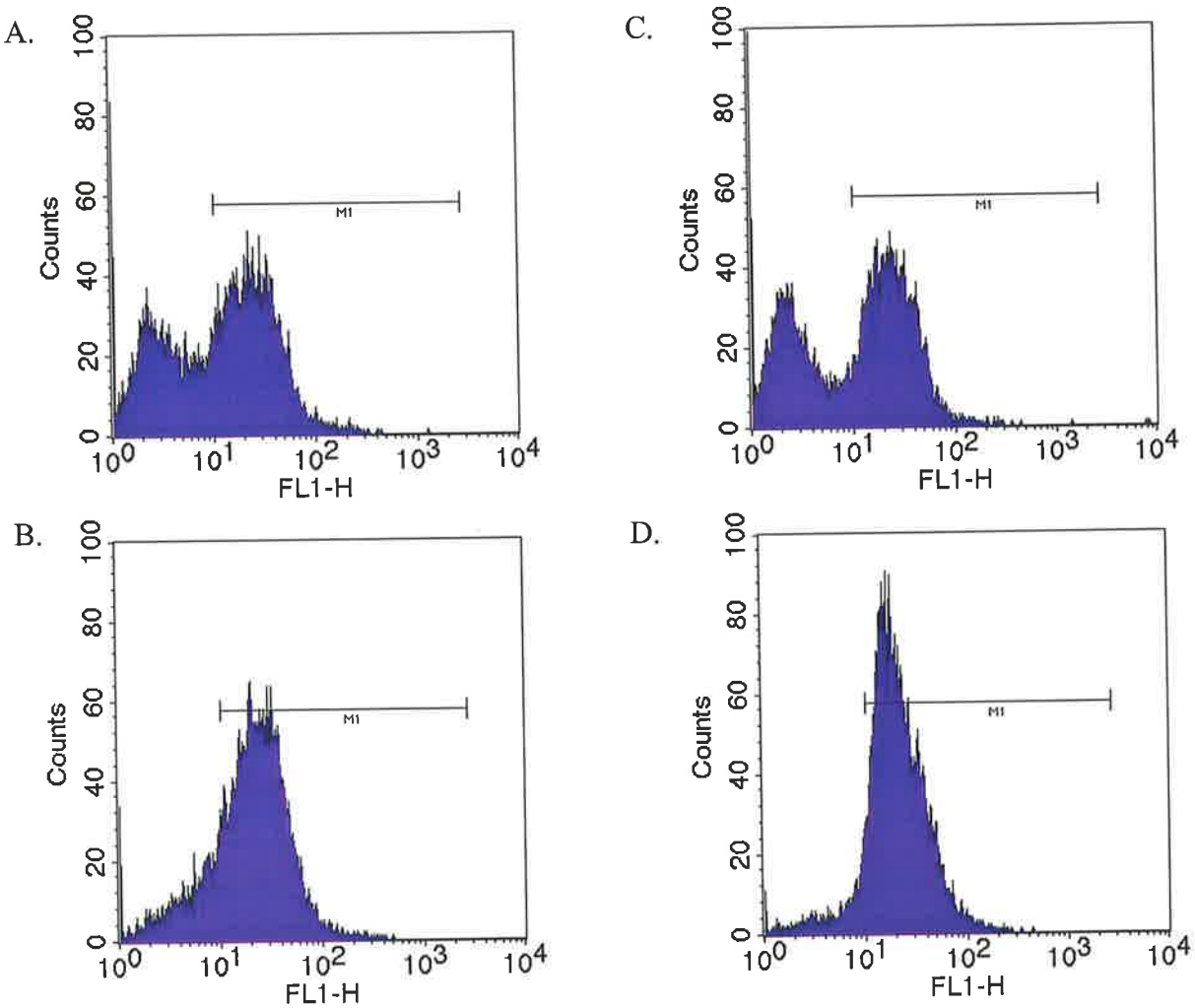


Figure 5.5 Comparison of R-IGF-1R mutant cells before and after single cell sorting

Cells were grown to ≤90% confluency, trypsinised and incubated with the mouse monoclonal anti-IGF-1R antibody 16-13 or a mouse IgG control. Cells were subsequently labeled with a FITC conjugated anti-mouse monoclonal antibody, fixed with paraformaldehyde and fluoresencing cells analysed on a FACS cell sorter.

A. Before single cell sorting R-IGF-1R (Y1250F) : 16-13
B. After single cell sorting R-IGF-1R (Y1250F) : 16-13
C. Before single cell sorting R-IGF-1R (Y1250F/Y950F) : 16-13
D. After single cell sorting R-IGF-1R (Y1250F/Y950F) : 16-13

Cell Line	% Expressing Cells (M1)	Mean FL1-H
A. R-:IGF-IR (Y1250F) 16-13	58.12%	20
B. R-:IGF-IR (Y1250F) 16-13	80.44%	25
C. R-:IGF-IR (Y1250F/Y950F) 16-13	55.97%	25
D. R-:IGF-IR (Y1250F/Y950F) 16-13	90.47%	20

Table 5.2 Comparison of the % expressing cells and the mean FL1-H value of the major peak

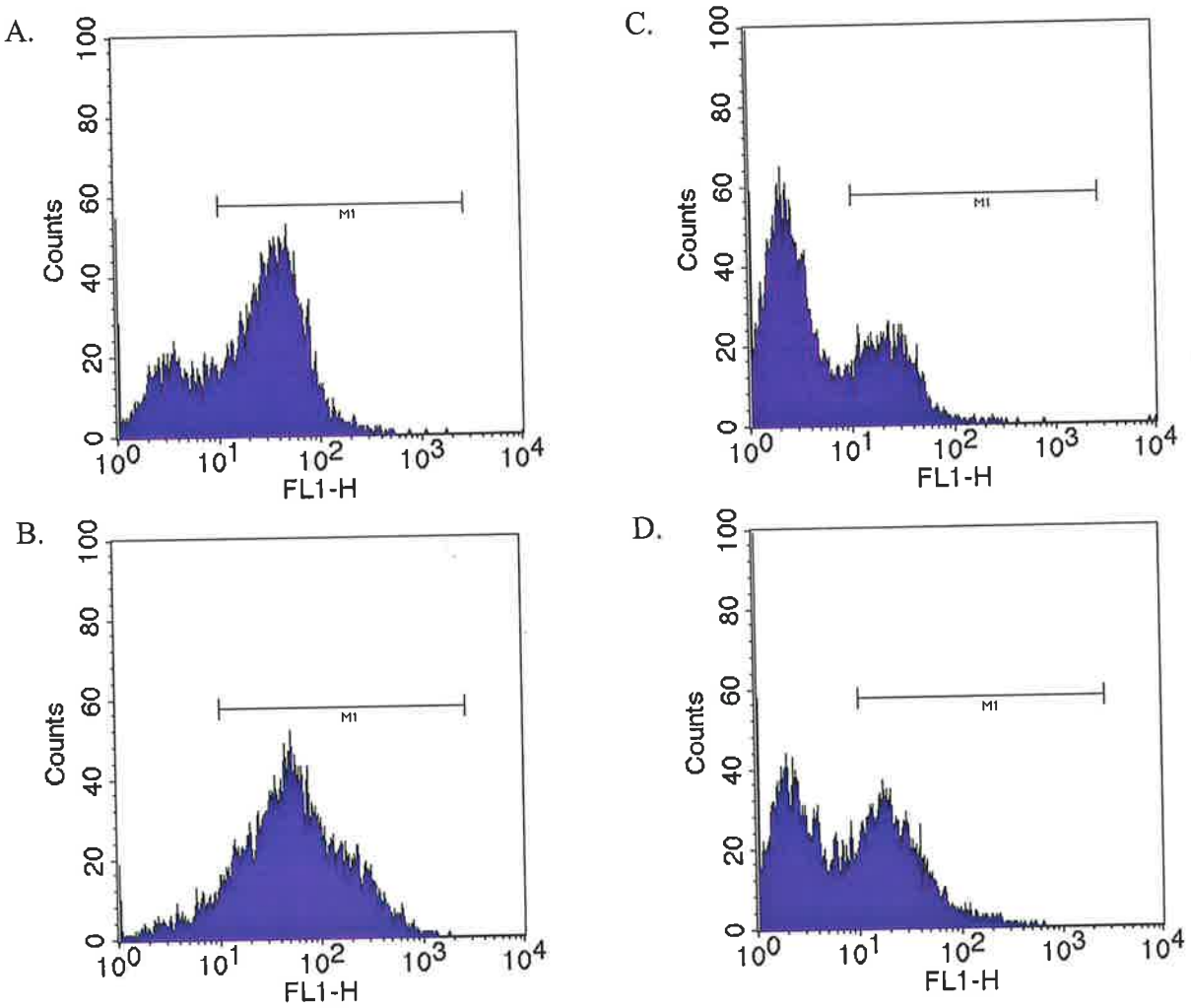


Figure 5.6 Comparison of R-IGF-1R mutant cells before and after single cell sorting

Cells were grown to ≤90% confluency, trypsinised and incubated with the mouse monoclonal anti-IGF-1R antibody 16-13 or a mouse IgG control. Cells were subsequently labeled with a FITC conjugated anti-mouse monoclonal antibody, fixed with paraformaldehyde and fluoresencing cells analysed on a FACS cell sorter.

A. Before single cell sorting R-IGF-1R (Y1251F) : 16-13
B. After single cell sorting R-IGF-1R (Y1251F) : 16-13
C. Before single cell sorting R-IGF-1R (Y1251F/Y950F) : 16-13
D. After single cell sorting R-IGF-1R (Y1251F/Y950F) : 16-13

Cell Line	% Expressing Cells (M1)	Mean FL1-H
A. R-:IGF-IR (Y1251F) 16-13	72.70%	40
B. R-:IGF-IR (Y1251F) 16-13	91.55%	50
C. R-:IGF-IR (Y1251F/Y950F) 16-13	28.01%	25
D. R-:IGF-IR (Y1251F/Y950F) 16-13	42.08%	20

Table 5.3 Comparison of the % expressing cells and the mean FL1-H value of the major peak

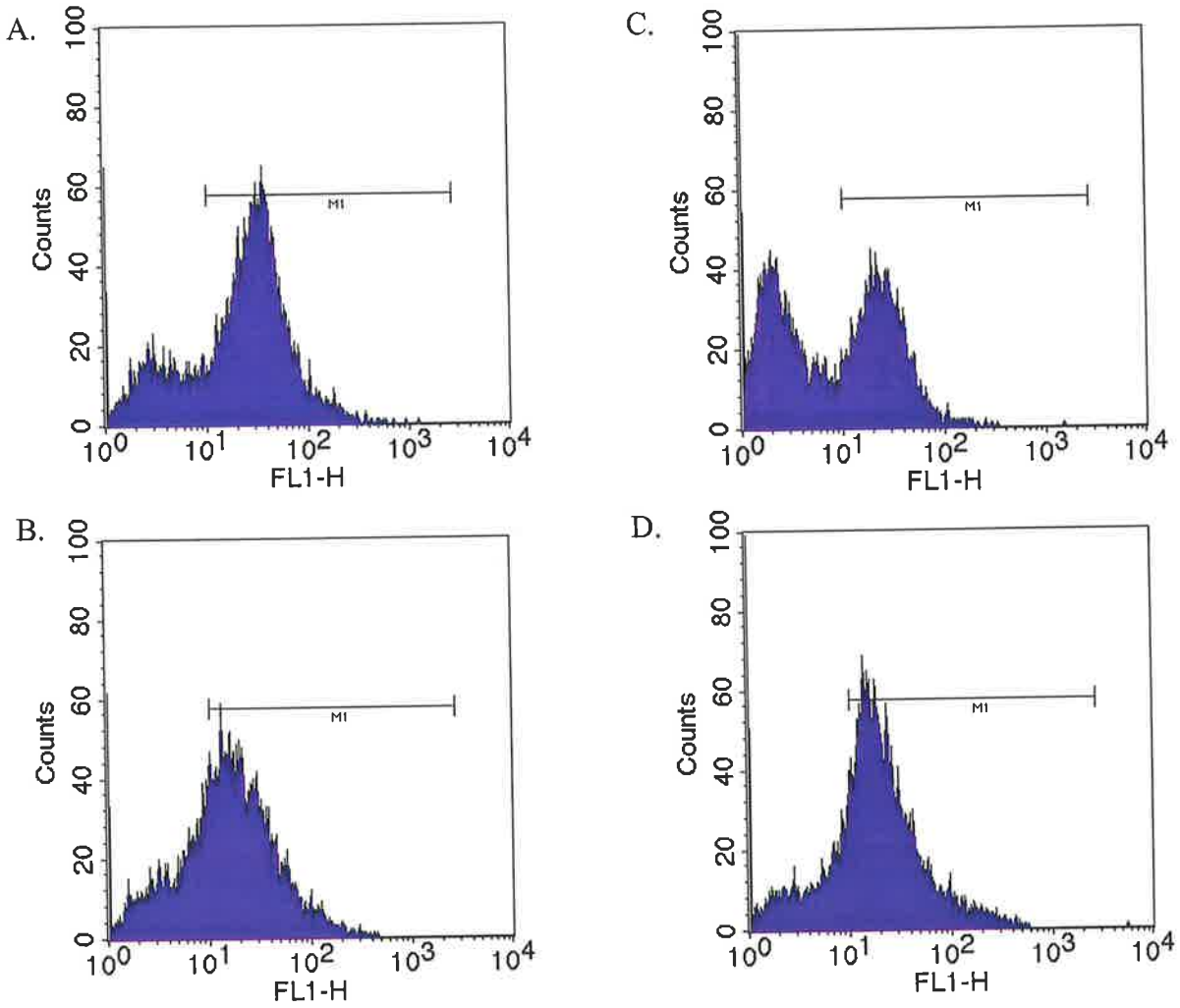


Figure 5.7 Comparison of R-IGF-1R mutant cells before and after single cell sorting

Cells were grown to ≤90% confluency, trypsinised and incubated with the mouse monoclonal anti-IGF-1R antibody 16-13 or a mouse IgG control. Cells were subsequently labeled with a FITC conjugated anti-mouse monoclonal antibody, fixed with paraformaldehyde and fluoresencing cells analysed on a FACS cell sorter.

A. Before single cell sorting R-IGF-1R (S1283A) : 16-13
B. After single cell sorting R-IGF-1R (S1283A) : 16-13
C. Before single cell sorting R-IGF-1R (S1283A/Y950F) : 16-13
D. After single cell sorting R-IGF-1R (S1283A/Y950F) : 16-13

Cell Line	% Expressing Cells (M1)	Mean FL1-H
A. R-:IGF-IR (S1283A) 16-13	76.42%	35
B. R-:IGF-IR (S1283A) 16-13	67.63%	15
C. R-:IGF-IR (S1283A/Y950F) 16-13	47.79%	25
D. R-:IGF-IR (S1283A/Y950F) 16-13	75.07%	15

Table 5.4 Comparison of the % expressing cells and the mean FL1-H value of the major peak

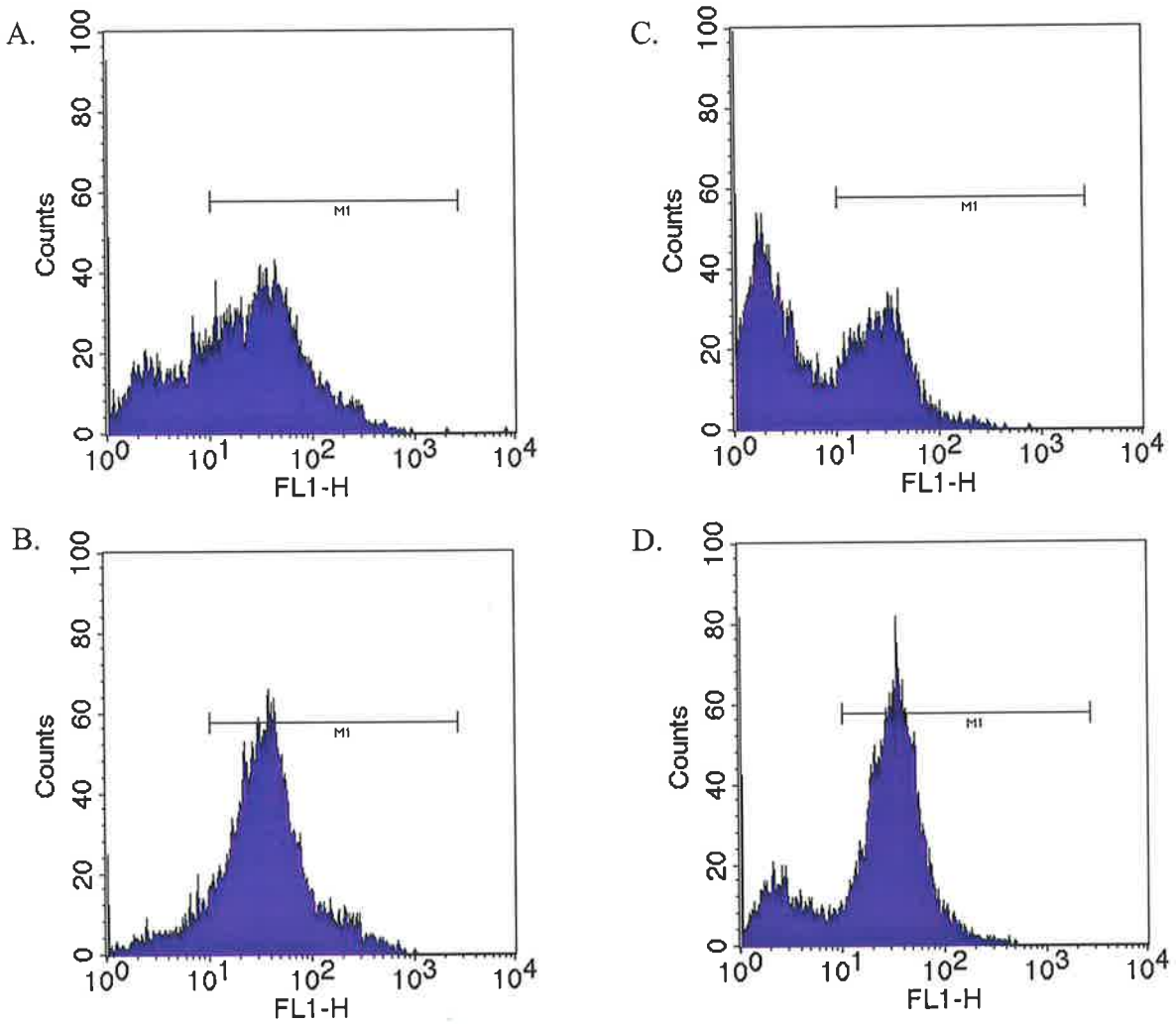


Figure 5.8 Comparison of R-IGF-1R mutant cells before and after single cell sorting

Cells were grown to ≤90% confluency, trypsinised and incubated with the mouse monoclonal anti-IGF-1R antibody 16-13 or a mouse IgG control. Cells were subsequently labeled with a FITC conjugated anti-mouse monoclonal antibody, fixed with paraformaldehyde and fluoresencing cells analysed on a FACS cell sorter.
A. Before single cell sorting R-IGF-1R (H1293F/K1294R) : 16-13
B. After single cell sorting R-IGF-1R (H1293F/K1294R) : 16-13
C. Before single cell sorting R-IGF-1R (H1293F/K1294R/Y950F) : 16-13
D. After single cell sorting R-IGF-1R (H1293F/K1294R/Y950F) : 16-13

Cell Line	% Expressing Cells (M1)	Mean FL1-H
A. R-:IGF-IR (H1293F/K1294R) 16-13	70.48%	40
B. R-:IGF-IR (H1293F/K1294R) 16-13	90.51%	40
C. R-:IGF-IR (H1293F/K1294R /Y950F) 16-13	40.55%	30
D. R-:IGF-IR (H1293F/K1294R /Y950F) 16-13	79.98%	30

Table 5.5 Comparison of the % expressing cells and the mean FL1-H value of the major peak

ranged between 15 and 50 with average of 30 units indicating that there is approximately a 3-fold range in the average number of receptors per cell between the lines. Based on these results only sorted cell lines were used in subsequent experiments except for R-IGF-1R (S1283A). In this case the FACS profile indicated a drop in the percentage of expressing cells in the sorted cell line relative to the original line. Therefore the better original line was used for the subsequent experiments.

Despite treating each cell line in an identical manner there was some variation in the percentage of receptor expressing cells, ranging from 28-84%, after the initial selection procedure. This may have been due to the initial transfection resulting in varying amounts of plasmid absorption between the lines. The cell sorting procedure was successful, resulting in an increase in the percentage of receptor expressing cells except for the S1283A mutant line. In general, the profiles also displayed a sharper receptor expressing peak, which was to be expected as a result of the sorting procedure.

Even though only receptor expressing cells were selected all of the cell lines still had non-expressing cells after further selection as shown by FACS analysis. This is probably due to the untransfected R- cells having a certain amount of innate resistance to the G418 selection, although it could be reasoned that the transfected cells should out compete them in terms of growth and proliferation. Alternatively the untransfected cells may be able to remain alive so long as a reasonable proportion of resistant cells were present, which would indicate either the transfer of resistance or another cell to cell survival signal. It is also possible that there may be a certain amount of downregulation of receptor expression while maintaining the ability to survive the antibiotic selection.

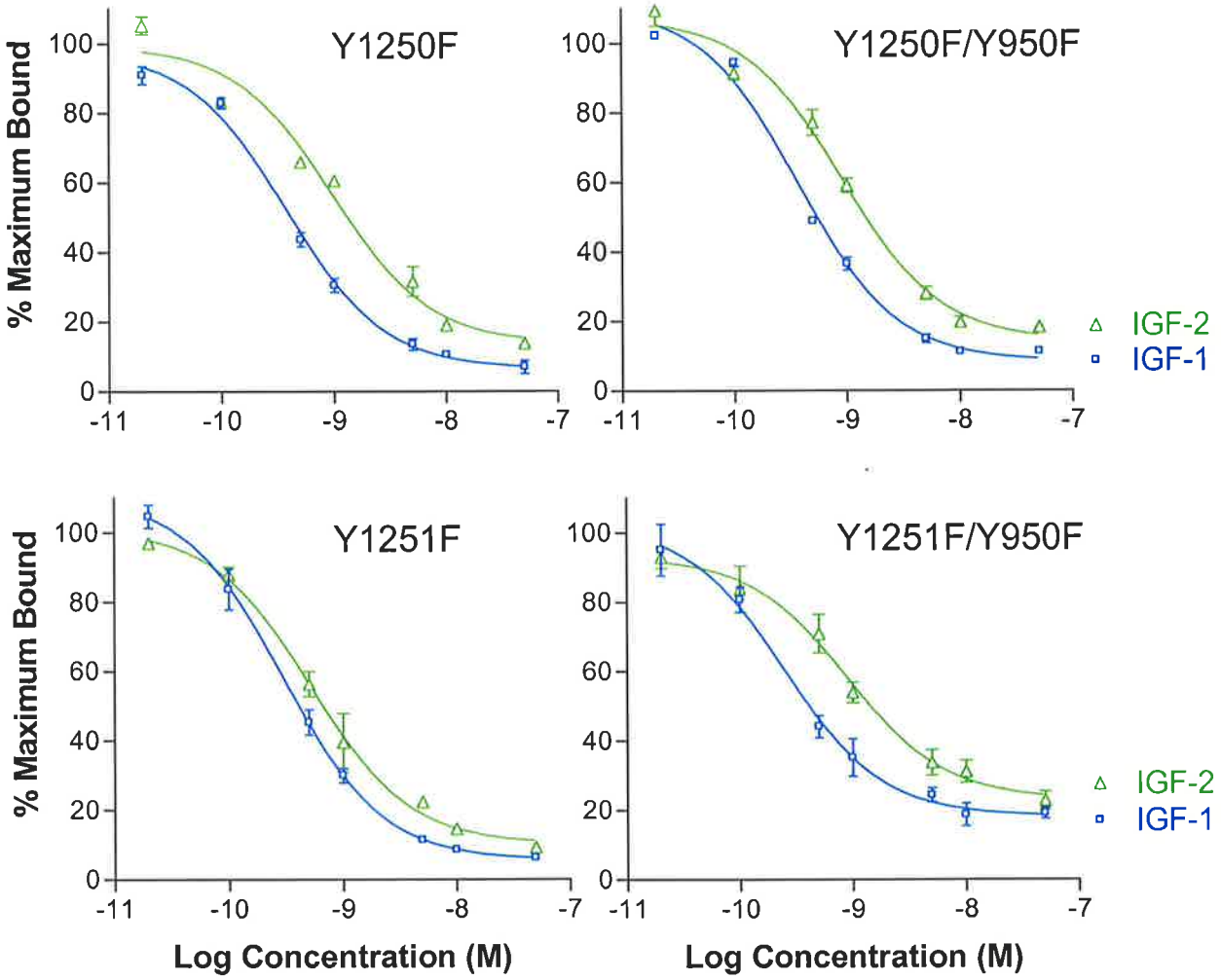
Interestingly both the mutant lines involving the S1283A mutation showed a significant decrease in the mean FL1-H value of their receptor expressing peak after cell sorting, even though the cells were taken from the middle of the original peak. This may point to the cells not tolerating higher numbers of these mutant receptors over extended periods of time. The Y1251F mutant displayed both the broadest profile and the highest mean FL1-H of its receptor expressing peak. It was later shown to give the highest phosphorylation in response to IGF-1, proportionally larger than expected given the difference in receptor numbers with the wild-type cell line, indicating that there this was a real effect. It may therefore be reasoned that this cell line contained a greater number of highly expressing cells because they gained some kind of advantage.

5.2.3 Competition Assays on Solubilised Receptors with IGF-1, 2

Having verified that the stable cell lines were expressing IGF-1R it was then necessary to confirm that they were able to bind to their cognate ligands IGF-1 and IGF-2 with the same affinity as wild-type receptor. A competition assay was therefore performed on solubilised receptor using europium-labeled IGF-1 and unlabelled IGF-1 or IGF-2 (See section 4.2.5.2 for a discussion of this assay). Each cell line was grown in 10cm dishes to approximately 90% confluency before being lysed and the solubilised receptor captured on 24-31 coated white 96 well plates. Europium labeled IGF-1 (Eu-IGF-1) was mixed with serially diluted IGF-1 or IGF-2 and the competition performed overnight at 4°C. After removing the competition mix and washing the plates, the amount of Eu-IGF-1 still bound to the receptor was measured using time-resolved fluorescence. The competition curves all fit a one-site binding curve and show clearly that all the mutants have a similar affinity for each ligand as the wild-type receptor with the mutants having a collective average EC_{50}s for IGF-1 and IGF-2 of 0.33±0.06 nM and 0.88±0.23 nM, compared with the wild-type EC_{50}s of 0.23±0.1 nM and 0.83±0.06 nM respectively (See Figure's 5.9, 5.10, and 4.6). Of the mutants Y1251F and S1283A had IGF-2 EC50 values that were significantly lower than the average (0.52±0.5 nM, and 0.67±0.4 nM respectively) or the H1293F/K1294R/Y950F mutant that was higher (1.31±0.07 nM). This may indicate a subtle change in the conformation of the intracellular region of the receptor that resulted in a slight change to the binding site. However since neither the Y1251F/Y950F, the S1283A/Y950F, nor the H1293F/K1294R receptors showed this change it is more likely to be experimental variation. Additionally, all receptors were recognized by the available antibodies used in these and subsequent assays. Overall, the results were to be expected due to the mutations occurring in the cytoplasmic region of the receptor and therefore unlikely to contribute to ligand binding.

The ligand binding competition assays with IGF-1 and IGF-2 showed that all the mutant receptors had similar affinities to the wild-type receptor at least for IGF-1. The obtained EC_{50}s are comparable with those values obtained in the established literature, which indicate a 2-5 fold difference in affinity between IGF-1 and IGF-2 (Germain-Lee *et al.*, 1992) (Forbes *et al.*, 2002). Therefore in the phosphorylation assays IGF-1 was used to stimulate cells based on its higher potency and more similar binding profiles.

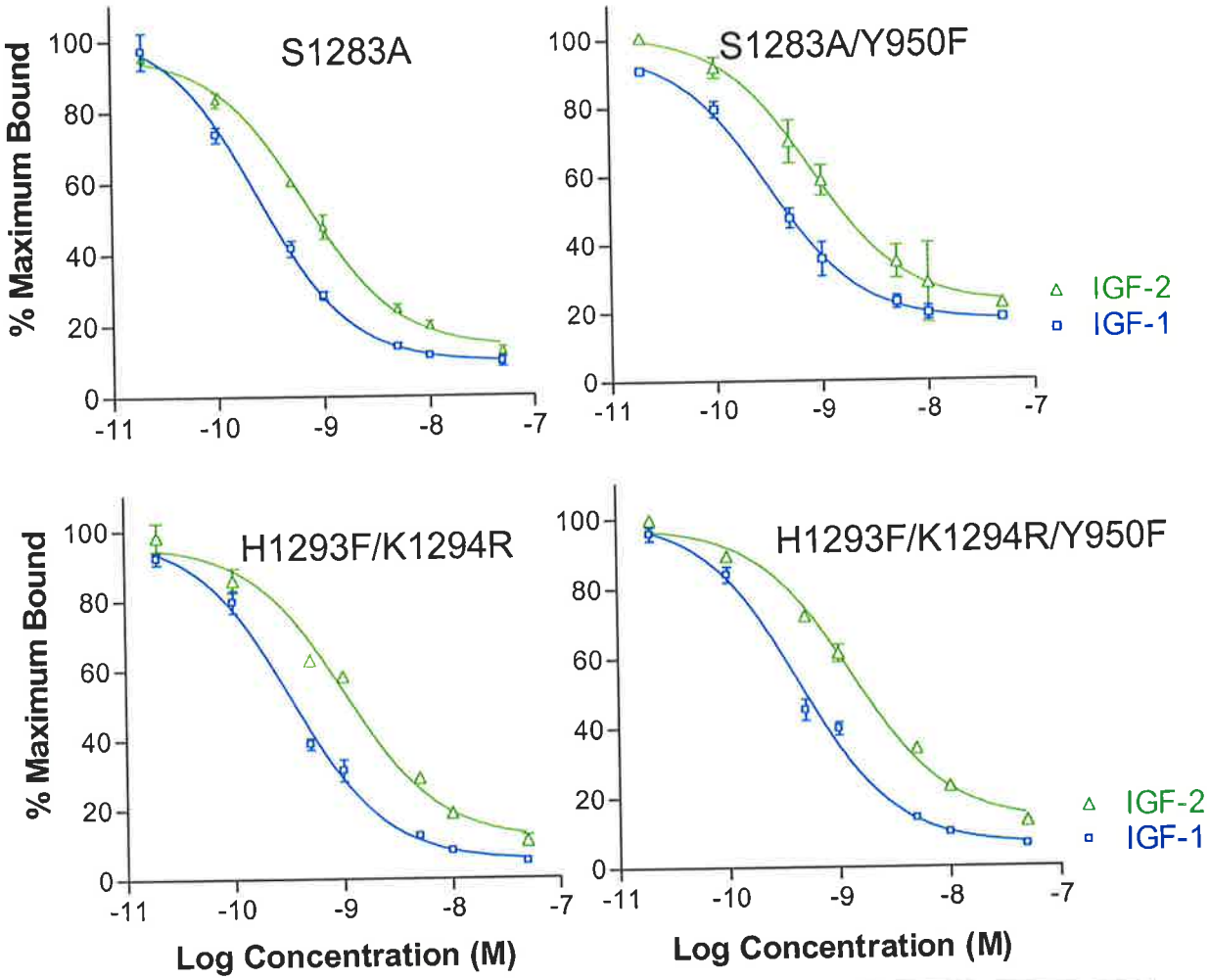
Eu-IGF-1 is a heterogeneous mixture based on HPLC analysis (data not shown), and the different species would have possible labeling sites at the amino terminal Gly^1, and the amino



	EC50 VALUES			
Growth Factor	Y1250F	Y1250F Y950F	Y1251F	Y1251F Y950F
IGF-1	0.37±0.06 nM	0.37±0.07 nM	0.30±0.01 nM	0.25±0.06 nM
P value	0.8 (NS)	0.97 (NS)	0.8 (NS)	0.97 (NS)
IGF-2	0.96±0.15 nM	0.91±0.09 nM	0.52±0.05 nM	0.90±0.07 nM
P value	0.54 (NS)	0.8 (NS)	1.0 (NS)	0.97 (NS)

Figure 5.9 IGF-IR Competition Binding Assay using IGF-1 and IGF-2

Stable cell lines expressing each receptor were grown in 10cm dishes to ~90% confluency before being harvested in lysis buffer and IGF-IR captured on 24-31 coated white 96 well plates. A competition mix between Eu-IGF-1 and serial dilutions of either unlabelled IGF-1 and IGF-2 was incubated with the receptor overnight at 4°C before unbound growth factor being washed away and the amount of remaining Eu-IGF-1 measured using time-resolved fluorescence. Data was converted to a percentage of maximum counts bound and fitted to a one-site competition binding curve using PrismTM (P value (Runs Test), NS = Not significant >0,05).



	EC50 VALUES			
Growth Factor	S1283A	S1283A Y950F	H1293F K1294R	H1293F K1294R Y950F
IGF-1	0.24±0.04 nM	0.31±0.03 nM	0.33±0.06 nM	0.43±0.07 nM
P value	0.54 (NS)	0.97 (NS)	0.8 (NS)	0.8 (NS)
IGF-2	0.67±0.04 nM	0.81±0.03 nM	1.02±0.1 nM	1.31±0.07 nM
P value	0.54 (NS)	0.54 (NS)	0.54 (NS)	0.54 (NS)

Figure 5.10 IGF-IR Competition Binding Assay using IGF-1 and IGF-2

Stable cell lines expressing each receptor were grown in 10cm dishes to ~90% confluency before being harvested in lysis buffer and IGF-IR captured on 24-31 coated white 96 well plates. A competition mix between Eu-IGF-1 and serial dilutions of either unlabelled IGF-1 and IGF-2 was incubated with the receptor overnight at 4°C before unbound growth factor being washed away and the amount of remaining Eu-IGF-1 measured using time-resolved fluorescence. Data was converted to a percentage of maximum counts bound and fitted to a one-site competition binding curve using PrismTM (P value (Runs Test), NS = Not significant > 0.05).

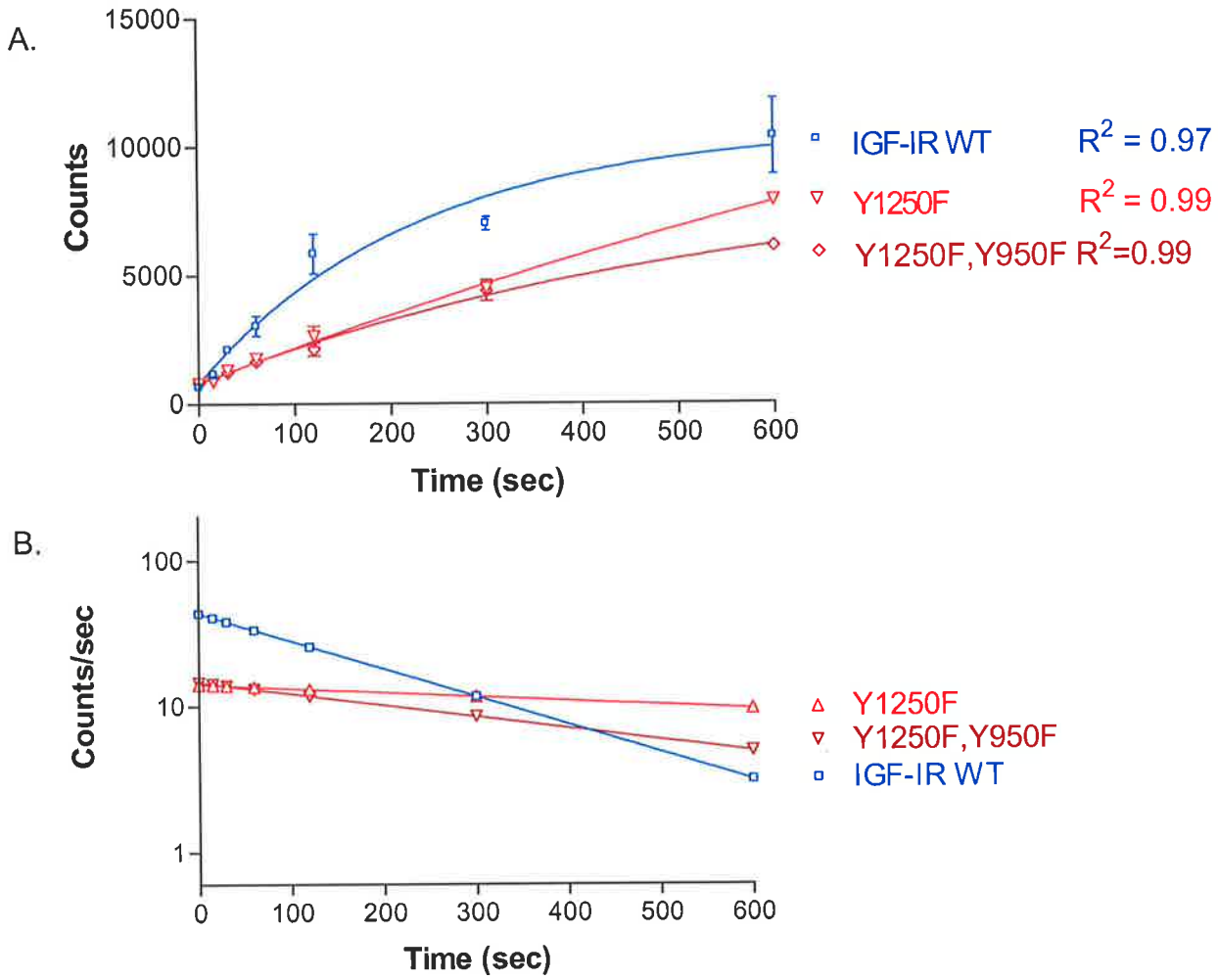
group side chains of Lys^{26}, Lys^{64}, and Lys^{67}. Gly^{1}, Lys^{64}, and Lys^{67}, lie in the flexible termini of IGF-1, which have not been found to greatly affect receptor binding affinity. Lys^{26} sits at the end of the B chain near the start of the C domain. Several residues near Lys^{26} including Phe^{23}, Tyr^{24}, and Tyr^{31}, have been implicated in receptor binding (See section 6.1). Even though a bulky europium chelate may disrupt this surface in a rigid molecule, the highly flexible nature of the C domain is likely to be able to compensate for its presence. Overall, since using either I^{125}-IGF-1 or eu-IGF-1 results in a very similar competition profile it is likely that they do not have significantly different modes or affinities for binding the IGF-1R, unlike the Eu-insulin tracer (see section 4.2.5.2).

5.2.4 Mutant Receptor Complexes Show Altered Amounts of Tyrosine Phosphorylation in Response to IGF-1.

The various mutations, particularly Y950F, are believed to disrupt the binding of important substrates to the receptor (including Shc, IRS-2, and IRS-10) and therefore it was thought that their absence would reduce the overall tyrosine phosphorylation of the ligand stimulated receptor/substrate complex. A time-course phosphorylation assay was developed that utilized a europium-labeled anti-phosphotyrosine antibody PT100 (as described in 4.2.9).

Each cell line was plated at a density of 25,000 cells/well in a 96 well plate and grown overnight. After growth in the absence of serum for at least 4hrs, the cells were stimulated with 10nM IGF-1 and time points taken by quickly aspirating the media and adding lysis buffer. Solubilised receptor was captured on 24-31 coated white 96 well plates left overnight at 4°C. Plates were then washed and developed and the amount of anti-phosphotyrosine antibody bound measured by time resolved fluorescence.

The data of the results were analyzed in Prism™ and each time-course fitted to an exponential association curve. The derived 'k' value was then taken and used to plot the velocity at each time point as described in 4.2.9 (See figures 5.11-5.14). Overall the results showed that the mutants had about the same or reduced amount of phosphorylation except for the Y1251F mutant, which was greatly enhanced to, about double the amount of phosphorylation of wild-type. Compared with wild-type receptor the Y1251F mutant showed more rapid and enhanced phosphorylation overall in the 10 min time course. The increase could have been due to the deregulation of intrinsic receptor kinase activity, lack of binding a phosphorylation regulator such as a phosphatase and therefore attenuated signalling, or an increased number of substrates binding to the receptor.



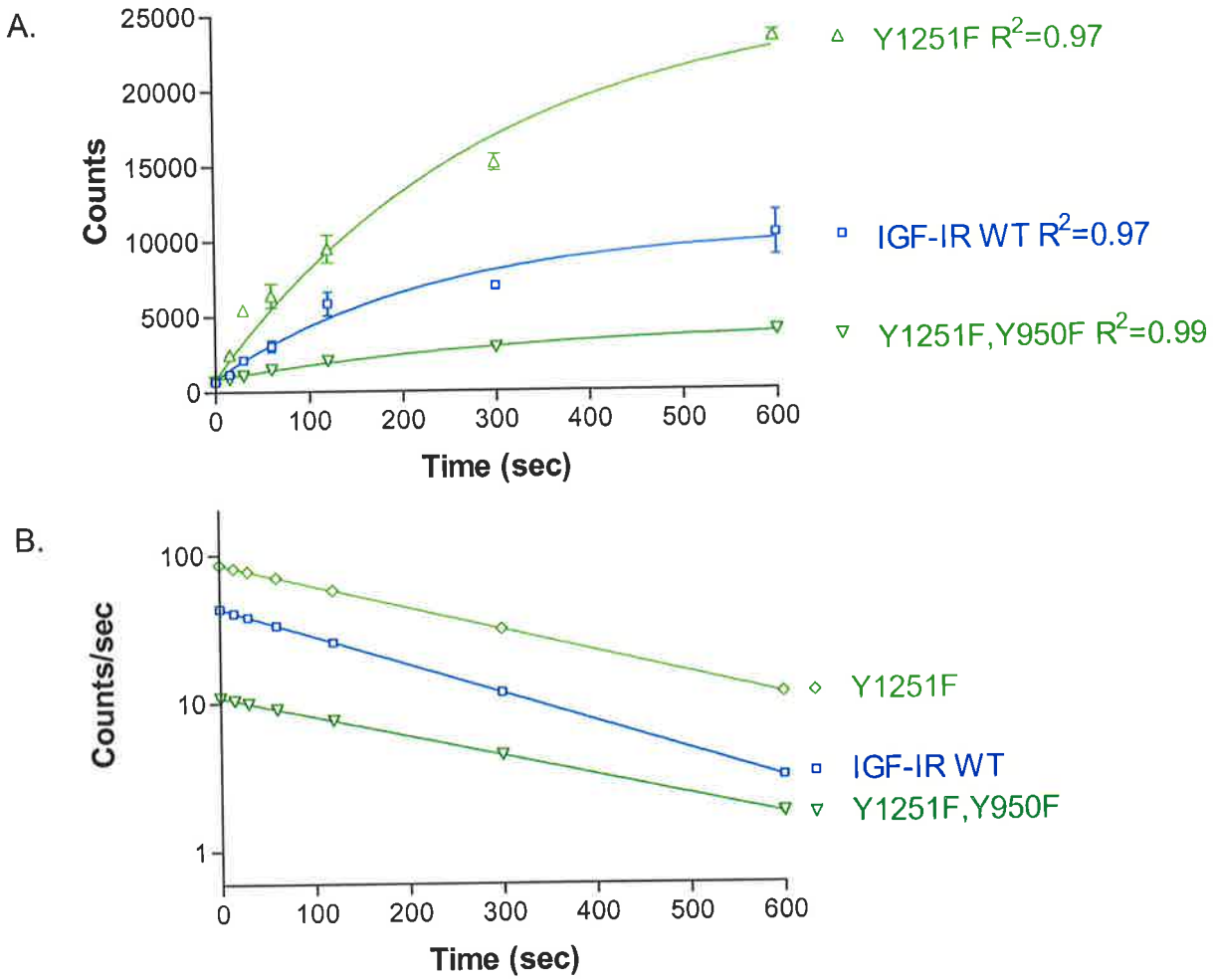
Receptor	WT	Y1250F	Y1250F/Y950F
Initial Rate	43.2 counts/sec	14.1 counts/sec	14.6 counts/sec
P value (Runs Test)	0.97 (NS)	0.54 (NS)	0.71 (NS)

Figure 5.11 Time Course Phosphorylation Assay

Stable cell lines expressing each receptor were grown in 96 well trays to ~90% confluency and allowed to go serum free for 4hrs before stimulation with 10nM IGF-1. At various time points media was aspirated, ice cold lysis buffer added and solubilised receptor captured on 24-31 coated white 96 well plates.

A. The amount of phosphotyrosine phosphorylation within the captured receptor/substrate complex was obtained using eu-PT100 antibody and measuring bound antibody by time-resolved fluorescence. Phosphorylation data was fitted to an exponential association curve using Prism™ software (P value (Runs Test), NS = Not Significant > 0.05).

B. The velocity at each point was calculated using the 'k' value and 'MAX' value obtained from the equation of the phosphorylation curve and graphed in Prism™.



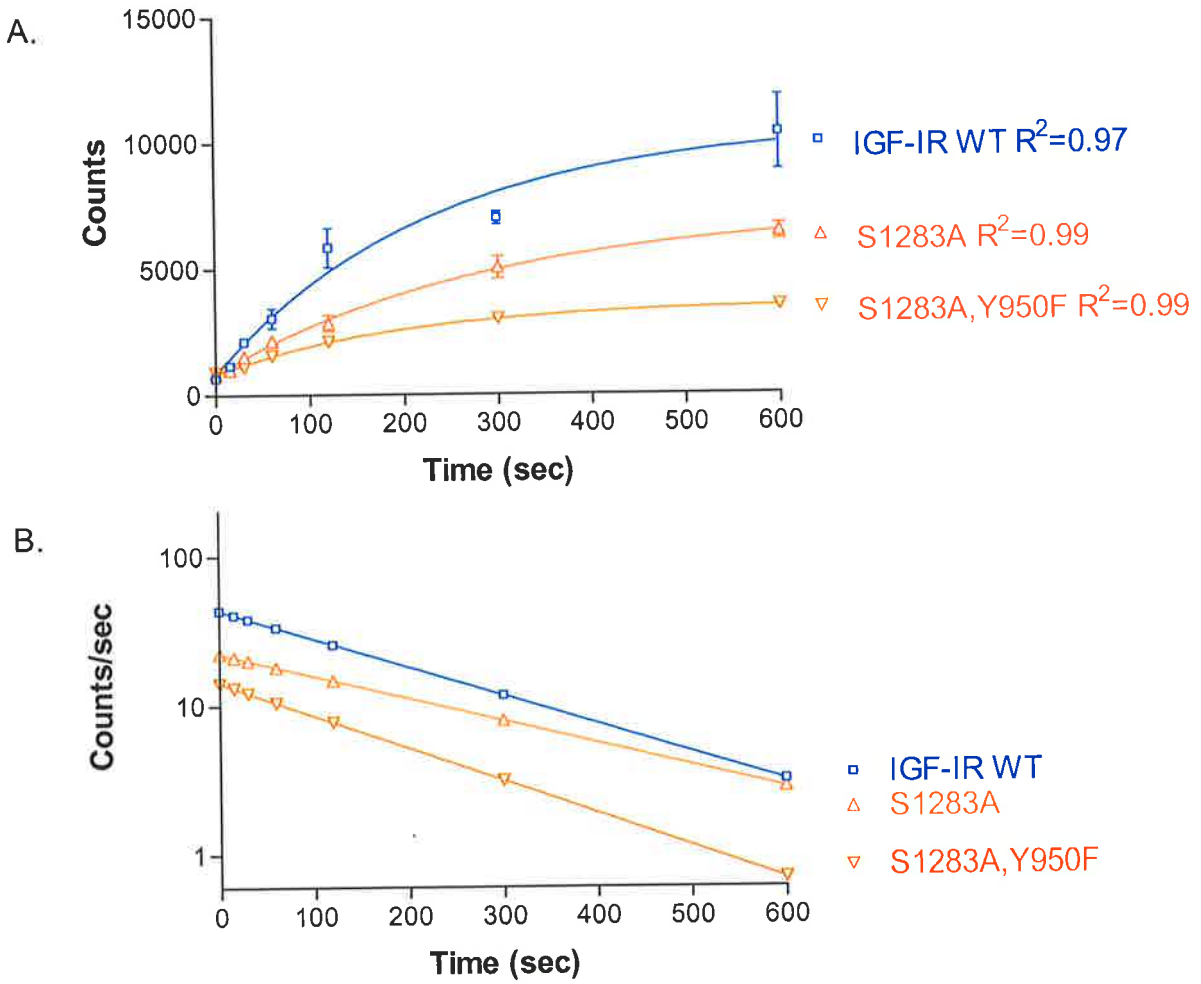
Receptor	WT	Y1251F	Y1251F/Y950F
Initial Rate	43.2 counts/sec	85.0 counts/sec	10.8 counts/sec
P value (Runs Test)	0.97 (NS)	0.54 (NS)	0.54 (NS)

Figure 5.12 Time Course Phosphorylation Assay

Stable cell lines expressing each receptor were grown in 96 well trays to ~90% confluency and allowed to go serum free for 4hrs before stimulation with 10nM IGF-1. At various time points media was aspirated, ice cold lysis buffer added and solubilised receptor captured on 24-31 coated white 96 well plates.

A. The amount of phosphotyrosine phosphorylation within the captured receptor/substrate complex was obtained using eu-PT100 antibody and measuring bound antibody by time-resolved fluorescence. Phosphorylation data was fitted to an exponential association curve using Prism™ software (P value (Runs Test), NS = Not Significant > 0.05).

B. The velocity at each point was calculated using the 'k' value and 'MAX' value obtained from the equation of the phosphorylation curve and graphed in Prism™.



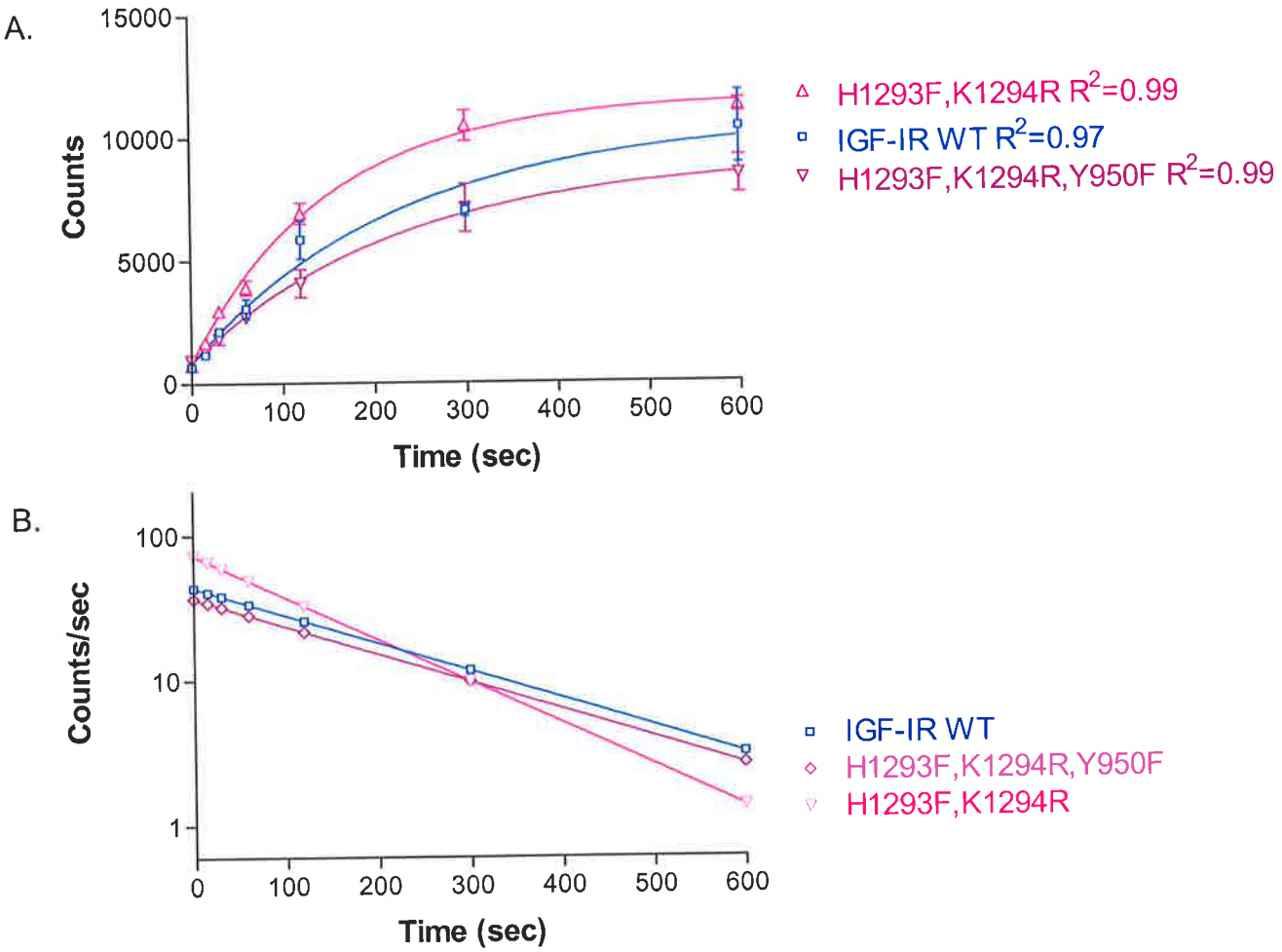
Receptor	WT	S1283A	S1283A/Y950F
Initial Rate	43.2 counts/sec	22.3 counts/sec	14.1 counts/sec
P value (Runs Test)	0.97 (NS)	0.97 (NS)	0.54 (NS)

Figure 5.13 Time Course Phosphorylation Assay

Stable cell lines expressing each receptor were grown in 96 well trays to ~90% confluency and allowed to go serum free for 4hrs before stimulation with 10nM IGF-1. At various time points media was aspirated, ice cold lysis buffer added and solubilised receptor captured on 24-31 coated white 96 well plates.

A. The amount of phosphotyrosine phosphorylation within the captured receptor/substrate complex was obtained using eu-PT100 antibody and measuring bound antibody by time-resolved fluorescence. Phosphorylation data was fitted to an exponential association curve using Prism™ software (P value (Runs Test), NS = Not Significant).

B. The velocity at each point was calculated using the 'k' value and 'MAX' value obtained from the equation of the phosphorylation curve and graphed in Prism™.



Receptor	WT	H1293F/K1294R	H1293F/K1294R/ Y950F
Initial Rate	43.2 counts/sec	71.0 counts/sec	36.4 counts/sec
P value (Runs Test)	0.97 (NS)	0.97 (NS)	0.71 (NS)

Figure 5.14 Time Course Phosphorylation Assay

Stable cell lines expressing each receptor were grown in 96 well trays to ~90% confluency and allowed to go serum free for 4hrs before stimulation with 10nM IGF-1. At various time points media was aspirated, ice cold lysis buffer added and solubilised receptor captured on 24-31 coated white 96 well plates.

A. The amount of phosphotyrosine phosphorylation within the captured receptor/substrate complex was obtained using eu-PT100 antibody and measuring bound antibody by time-resolved fluorescence. Phosphorylation data was fitted to an exponential association curve using Prism™ software (P value (Runs Test), NS = Not Significant).

B. The velocity at each point was calculated using the 'k' value and 'MAX' value obtained from the equation of the phosphorylation curve and graphed in Prism™.

The phosphorylation velocity graphs showed that all the mutants except those involving Y1250F had almost identical slopes to wild-type, suggesting that this aspect of the regulation had not been affected. The Y1250F and, to a lesser extent Y1250F/Y950F mutants, although having a slightly reduced amount of phosphorylation after 10min relative to wild-type, did not exhibit the typical decrease in velocity of phosphorylation in this time period, instead remaining almost flat. This indicates that that there is a loss in the regulation of phosphorylation different to that of Y1251F.

The anti-phosphotyrosine antibody PT-100 was used to assess the level of tyrosine phosphorylation in IGF-1 activated receptor/substrate complexes. The majority of the mutants except Y1251F had a similar or slightly reduced level of tyrosine phosphorylation than wild-type. The small reductions are probably due to lower levels of receptor expression as seen by the FACS analysis, resulting in less receptor being captured.

The Y1250F and Y1250F/Y950F mutants displayed a level of phosphorylation compared to wild-type that correlated well with the level of receptor expression, however unlike all the other mutants, had a flattened phosphorylation curve that over the ten minute time course of the experiment did not flatten out as quickly. The graph of the velocity of phosphorylation at each time point indicated that the rate of phosphorylation was not reducing much over time. It therefore appeared that the Y1250F mutants were lacking a level of regulation that would normally begin inhibiting the initial fast rate of phosphorylation as soon as the receptor was activated. The Y1250F mutant has been previously shown to be internalized at 30% the rate of wild-type receptor (Miura *et al.*, 1997). This may mean that the process of internalization is required for the rapid decrease in the rate of phosphorylation and that the binding of a protein to phosphorylated Y^{1250} is involved in enhancing the speed of this process. The increased rate of de-phosphorylation, or reduced rate of phosphorylation by the receptor, could be due to the internalized receptor being more accessible to particular cytosolic phosphatases or to factors that regulate the receptor kinase activity respectively. Interestingly, the Y1250F/Y950F mutant does not exhibit quite as steady a rate of phosphorylation as the Y1250F mutant, suggesting that the binding of factors reliant on phosphorylated Y^{950} slightly inhibits the reduction in the rate of phosphorylation. Whether this is through inhibiting internalization or the action of some other protein remains unclear.

The Y1251F mutant exhibited a hyperphosphorylation phenotype even after receptor numbers were taken into account. However it displayed the same rate of reduction in phosphorylation

as the wild-type receptor. This conflicts with published data also from Y1251F mutants expressed in R- cells that found that there was no gross alteration of tyrosine phosphorylation as seen on a western blot of a whole cell lysate after a 5 min stimulation with 20ng/ml IGF-1 (~28nM) (Miura, 1995). The data presented here differs from this in that there is a receptor capturing step which will magnify the localized changes in tyrosine phosphorylation that may not be visible in the whole cell lysate. From these results it appears that the Y1251F mutant has lost a level of regulation controlling the initial rate of phosphorylation of the ligand activated kinase while retaining the slower acting level of phosphorylation inhibition perturbed by the Y1250F mutants There are also several possibilities for the cause of the hyperphosphorylation. Phosphorylated Tyr^{1251} is most likely a docking site for particular receptor substrates. Since the lack of a functional sidechain at this position results in hyperphosphorylation the normal action of protein(s) bound here would appear to be as a repressor of phosphorylation. This protein could therefore have been a phosphatase, acting generally to de-phosphorylate residues in the receptor and its substrates, a kinase inhibitor, acting directly on the receptors tyrosine kinase domain or even as a protein binding inhibitor, preventing key receptor substrates from being phosphorylated. General inhibition of phosphatases by sodium orthovanadate did not greatly change the amount of phosphorylation of the Y1251F mutant, however it did bring the phosphorylation profile of the wild-type receptor up to a comparable level. This suggests that either phosphorylated Y^{1251} normally binds a phosphatase, or that even deregulating the removal of phosphate groups cannot make the receptor phosphorylate substrates any quicker.

Interestingly the Y1251F/Y950F mutant did not have increased phosphorylation and was significantly lower than wild-type, however the FACS profile also suggested it had much lower average receptor expression. The phosphorylation result suggests that the derepression seen in the Y1251F mutant requires the presence of a protein that normally requires phosphorylated Tyr^{950}. This could include the adapter proteins IRS-1/2 or even Shc. Alternatively, this could suggest that the increase in phosphorylation is due to enhanced phosphorylation of the substrates reliant on a functional Tyr^{950} and not on a general de-regulation of the receptor kinase activity. Somewhat strangely, sodium orthovanadate mediated phosphatase inhibition resulted in an increase in phosphorylation for the Y1251F/Y950F mutant, providing evidence that would reject both previous proposals. A possible explanation could be that the combined absence of both phosphorylated Tyr^{950} and Tyr^{1251} allows another phosphatase to bind to the receptor complex that would not be able to in WT and Y1251F, whose inhibition by sodium orthovanadate results in the observed increase in phosphorylation.

All of the mutants involving either S1283A or H1293F/K1294R did not differ significantly from wild-type once the receptor numbers were taken into account, indicating that proteins that interact with these residues do not affect phosphorylation regulation or the binding of major receptor substrates. Taken overall, the time-course phosphorylation experiment was able to show differences in the tyrosine phosphorylation of receptor/substrate complexes, and the analysis of the mutant data seems to suggest that there are at least two levels of tyrosine phosphorylation repression that occur after kinase activation. The first affects the initial rate of phosphorylation as shown in the Y1251F mutant and presumably requires factors that are already very near or bound to the receptor complex so that they can act almost instantaneously. The second level of repression inhibited by the Y1250F mutant effects the gradual reduction in the rate of phosphorylation over the time-course, and may be related to the rate of receptor internalization.

5.2.5 Effect of Phosphatase Inhibition on Tyrosine Phosphorylation

The cause for the increase in tyrosine phosphorylation in the Y1251F mutant was likely due to either inhibition or absence of binding of a phosphatase to the receptor. In order to test this hypothesis time-course phosphorylation was performed with an additional incubation before stimulation in which 30μM sodium orthovanadate (NaOV) was added to the media. This compound inhibits both tyrosine and serine de-phosphorylation by inhibiting phosphatases (Tan *et al.*, 2001). In this assay WT, Y1251F and Y1251F/Y950F expressing cell lines were tested for the effect of NaOV pre-treatment on IGF-1 stimulated tyrosine phosphorylation. The assay was performed twice and a representative is shown (See figure 5.15). In the first instance, the results indicate that there is very little effect from adding NaOV to the Y1251F mutant, shown by the overlapping phosphorylation curve. Both WT and Y1251F/Y950F constructs show an increase in phosphorylation, the WT to a level comparable to that of Y1251F. This result therefore provides some evidence that the reason for elevated phosphorylation levels in Y1251F is because of the inhibition or lack of the ability of a phosphatase to bind to the receptor complex. Overall the results had lowered phosphorylation counts even though the number of cells plated for the assay was the same. The change in sensitivity may have been due to the use of a new batch of Eu-labeled anti-phosphotyrosine antibody, or experimental variation between data sets. However, even though the raw counts were lowered, the relationships between the control treatments were similar to those in the initial results.

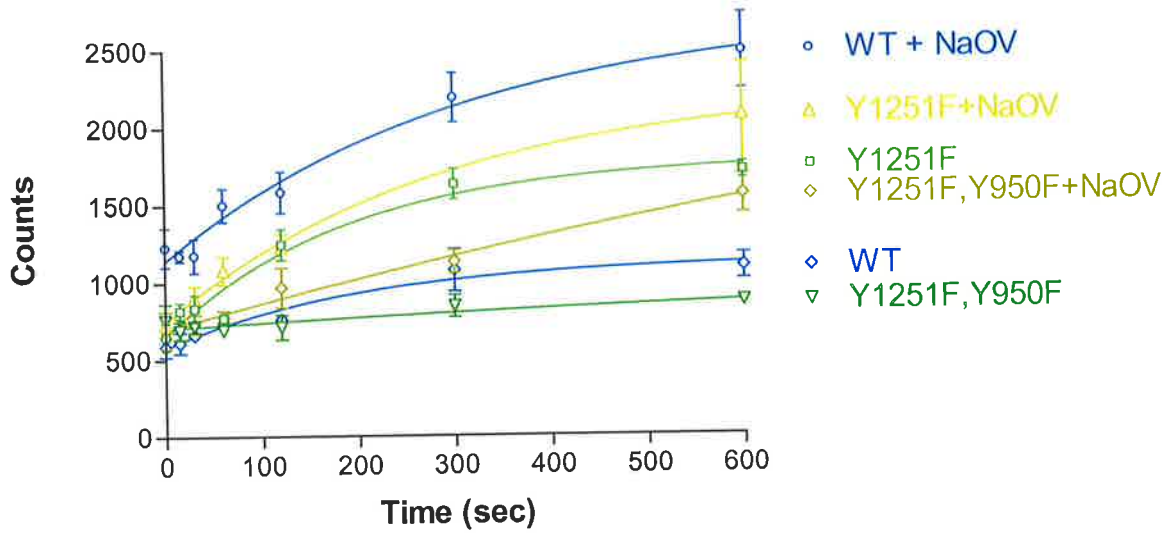


Figure 5.15 Effect of phosphatase inhibition on the Time Course Phosphorylation Assay

Stable cell lines expressing each receptor were grown in 96 well trays to ~90% confluency and allowed to go serum free for 4hrs. 300μM Sodium orthovanadate (NaOV) was added 30min before stimulation with 10nM IGF-1. At various time points media was aspirated, ice cold lysis buffer added and solubilised receptor captured on 24-31 coated white 96 well plates. A. The amount of phosphotyrosine phosphorylation within the captured receptor/substrate complex was obtained using eu-PT100 antibody and measuring bound antibody by time-resolved fluorescence. Phosphorylation data was fitted to an exponential association curve using Prism™ software (P value (Runs Test), NS = Not Significant).

5.2.6 Effect of Mutations on the Proliferative Potential of the Receptor

The proliferative potential of IGF-1 and IGF-2 acting through the IGF-IR has been long established, however the exact combination of residues within the cytoplasmic region of the receptor required to trigger the activation of the appropriate pathways has not yet been fully elucidated. There appears to be some redundancy of several pathways that can originate from the IGF-IR, so it was thought that the Y950F mutation in combination with mutations in some of the unique IGF-IR residues would be able to abolish the receptors ability to stimulate proliferation. In this experiment the number of viable cells present after incubation with varying concentrations of IGF-1 was measured after 48hrs as described previously (See Section 4.2.10).

The results as presented in figure 5.16 show that wild-type receptor was able to increase proliferation above serum free levels to approximately 40% of the 10%FCS serum control. The maximal increase was achieved at a concentration of ~10nM. The untransfected R- control cell line was unresponsive to IGF-1, and showed only a small increase in growth in the presence of serum. All the mutants were able to respond to IGF-1 stimulation, resulting in an increase in cell number over serum-free control wells, and also showed better proliferation than the untransfected control cells.

Both the Y1250F and the Y1250F/Y950F mutants had a slightly reduced but similar response to IGF-I to the wild-type receptor. This is consistent with the time-course phosphorylation data that also demonstrated this relationship. The data shows that the Y1251F receptor has increased proliferation, up to ~90% of the serum control, an increase of 50% compared with wild-type. The double Y1251F/Y950F mutant is also consistent with its phosphorylation data showing a reduced proliferative capacity. Again, similar to the phosphorylation assays the S1283A, S1283A/Y950F, H1293F/K1294R, and H1293F/K1294R/Y950F mutants did not have greatly altered responses compared with the wild-type receptor. The results from the Y950F containing mutants were unexpected due to this single mutant being previously shown to affect the receptor's ability to cause proliferation.

The functional capacity of the mutants to increase the rate of proliferation in response to IGF-1 was tested in a proliferation assay measuring the number of viable cells after 48hr stimulation. All cell lines except the untransfected control cells showed increased proliferation above a serum free control. The 10nM maximum effect for the wild-type receptor line is consistent with previous findings in the literature.

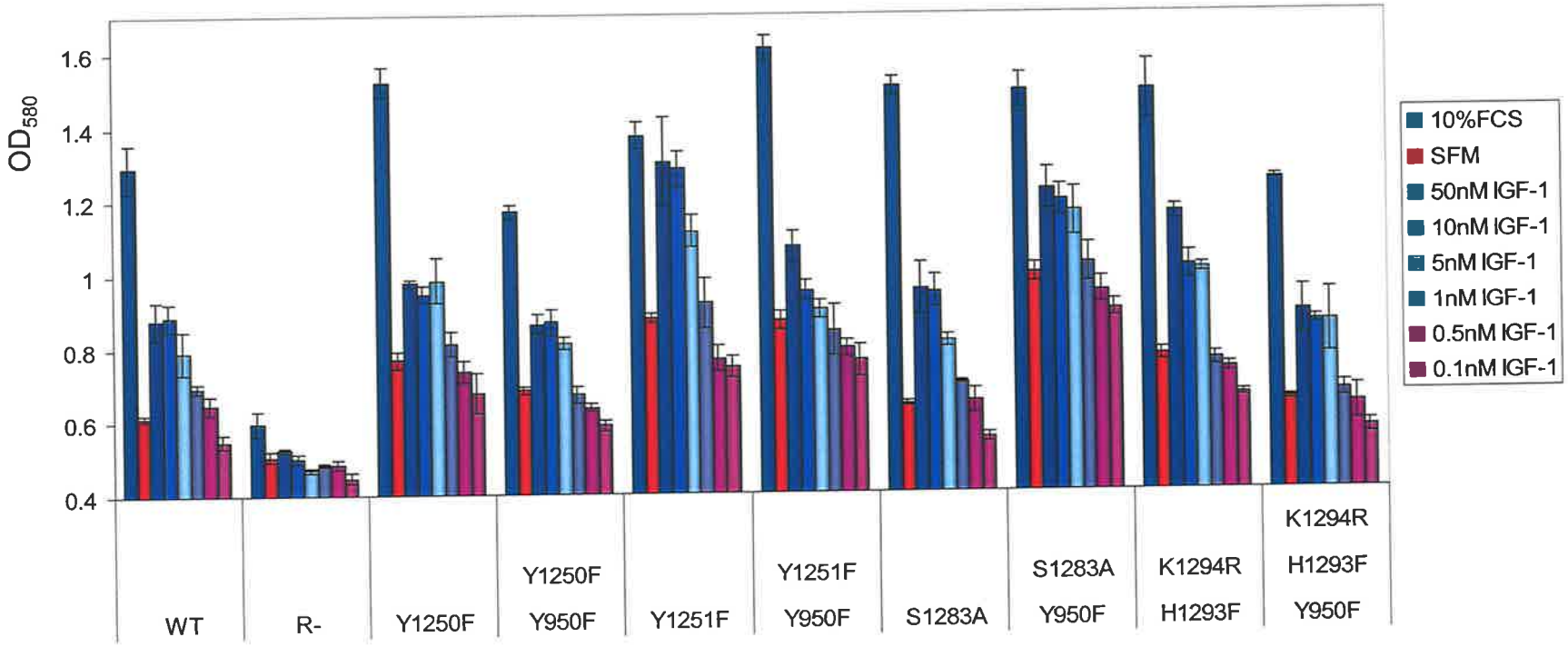


Figure 5.16 Proliferation of IGF-IR mutants in response to IGF-1

R- cells and stable cell lines expressing each receptor were plated in clear 96 well trays overnight before going serum free for 4 hrs and being stimulated with serially diluted IGF-1 for 48hrs. Viable cells were then stained with methylene blue and the optical density at 580nM measured.

The Y1250F mutant may have been expected to give a higher proliferative response than the results indicated since, unlike the wild-type receptor, the phosphorylation curve for Y1250F appeared as if it would increase in a linear fashion beyond the 10min time point. One could therefore predict that this would result in an extended activation of downstream signaling pathways, and subsequently a stronger proliferative response. Since this clearly has not occurred, one may conclude that either the initial amplitude of the phosphorylation, or the initial rate of internalization, determines how large the final response is. From this experiment it is impossible to know when or if the steady rate of phosphorylation would have reached a plateau after the 10min time-point or continued indefinitely.

The results of the time-course phosphorylation experiments predict that the Y1251F mutant would have a greater proliferative capacity based on its hyperphosphorylation, provided another protein important in the signaling cascade was not also perturbed. The increased level of proliferation of this cell line relative to the 10% FCS serum control over the wild-type receptor cell line suggests that the hyperphosphorylation is probably the cause. Each other mutant line showed a phosphorylation response that correlated well with its ability to be tyrosine phosphorylated. It was not expected that the double mutants including the Y950F mutation would be able to increase cellular proliferation to a similar level to wild-type, since it had been previously established that the single Y950F IGF-1R mutant has a vastly reduced mitogenic activity in R- cells and 32D cells (a cell line that lacks endogenous IRS proteins) (Jiang, 1996) (Yam, 2001). It is unclear why this would be the case unless the assay was not performed for a sufficiently long time for any differences in cell proliferation. However recently, it was found that the IGF-IR is able to activate the PI3-K pathway and induce DNA synthesis in a normal fashion when an analagous loss of function mutation was introduced into the NPXY motif (Kataoka, 2004).

Overall it is clear that all of the mutant receptors are able to increase proliferation in response to IGF-1 even if only slightly. It also seems to be the case that the number of receptors expressed in the cell lines contributes to the level of proliferation via the extent of the phosphorylation which reinforces the observation that increases in IGF-1R expression is associated with developing malignancy. It is interesting to note that all transfected cell lines were able to grow much more proficiently in serum than untransfected R-s, even though some had a vastly reduced ability to respond to IGF-1. The reason for this is unclear, although it does point to the presence of the receptor being important for proliferation in a non-ligand dependent manner. This could potentially be from recruitment of intracellular signaling

molecules to the membrane by inactivated receptor, participating in other growth factor initiated signaling. This idea has been tested in the case of the Epidermal Growth Factor Receptor (EGFR) co-transfected with IGF-1R in R- cells. These results showed that EGFR required the presence of IGF-1R in order to cause proliferation or colony formation in soft agar, and that this only required a functional IGF-1R kinase domain (Burgaud, 1996).

5.2.7 PTP-1D Localization is not Affected in the Y1251F or Y1250F Mutants

The association of the protein phosphatase PTP-1D with the IGF-1R was tested by a western blot analysis. Wild-type, Y1250F, Y1251F, and Y1251F/Y950F cell lines were grown in 10cm dishes to ~80-90% confluency, and incubated in serum free conditions for 2hrs before treatment with 10nM IGF-1 for up to 12 min. Stimulation was stopped by the addition of ice cold lysis buffer, before IGF-1R was immunoprecipitated from the soluble lysate with the antibody 24-31. Immunoprecipitates and a sample of whole cell extract were run on a 10% polyacrylamide gel, transferred to a nitrocellulose membrane and probed with an α-PTP-1D antibody (1/2500 dilution) (See Figure 5.17). The results show the presence of higher molecular weight bands appearing in the IGF-1R immunoprecipitates in addition to a band of the expected size of PTP-1D. There does not appear to be a large change in the concentration of any of the bands either upon ligand stimulation or between the different cell lines, which may suggest that the localization of this protein is not perturbed in any of the mutants and is therefore unlikely to be responsible for the other observed effects.

In a preliminary attempt to identify the protein responsible for causing the hyperphosphorylation seen in the Y1251F mutant, western blots probing for the phosphatase PTP-1D in IGF-1R immunoprecipitates were performed on IGF-1 stimulated cells. PTP-1D (also called SHP-2 or SH-PTP2) is known to mediate the activation of PI3K and Akt by growth factors including IGF (Wu, 2001). The results showed no significant difference between wild-type cells, Y1250F, Y1251F, or Y1251F/Y950F cells. Detectable PTP-1D in the whole cell extract had the expected apparent molecular weight of 72 kDa, however in the anti-receptor immunoprecipitates the major form appeared in two bands at ~80 and ~95 kDa, although the expected sized form was also visible. The higher molecular weight bands probably represent activated phosphorylated forms of the protein. Interestingly there do not appear to be major changes in the amount of any of the forms upon ligand stimulation, suggesting that this protein is in a permanently active state. This may be to ensure that the receptor or its substrates are maintained in an inactivated dephosphorylated state while receptor ligand is not present and that there is not persistent phosphorylation after ligand stimulation has occurred. Alternative candidate proteins could include PTP-1B, SOCS, and

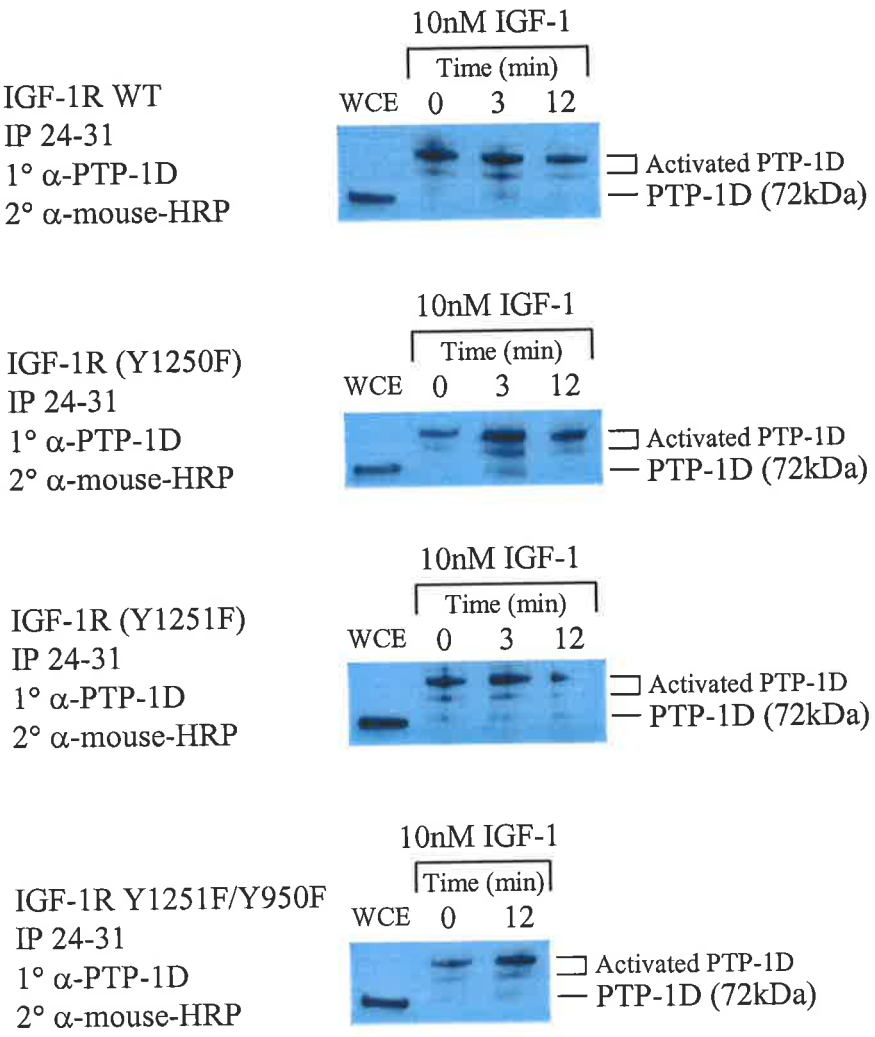


Figure 5.17 Detection of PTP-1D in IGF-IR immunopreciptates

Stable cell lines expressing each receptor were grown in 10cm dishes to ~90% confluency before being incubated in serum free media for 2hrs. Cells were stimulated with 10nM IGF-1 for 12 min before being harvested in lysis buffer and IGF-IR immunoprecipitated with 24-31 overnight at 4°C. Immunocomplexes were captured with protein A sepharose and run on a 10% polyacrylamide gel. Proteins were transferred to a nitrocellulose membrane and probed with 1/2500 diluted mouse α-PTP-1D antibody and subsequently 1/2500 diluted α-mouse-HRP antibody. Enhanced chemiluminescence was used to detect antibody bound protein bands which were visualised by a 30 sec exposure on X-Ray film.

FAK. Due to time constraints these proteins could not be tested for their ability to bind to the mutant receptors or be affected by them.

5.3 SUMMARY AND CONCLUSION

The aim of this project was to assess the effect of mutations in C-terminal domain residues alone and in combination with a Y950F mutation on the regulation of signaling from the hIGF-1R. Site-directed mutagenesis was used to make the changes Y1250F, Y1251F, S1283A, and H1293F/K1294R in both wild-type and Y950F hIGF-1R. Control competition assays proved that the mutants had no effect on either IGF-1 or IGF-2 binding. Ligand activated receptor complexes showed differing amounts of tyrosine phosphorylation, with the Y1251F mutant having a enhanced degree of phosphorylation. This was not affected by tyrosine phosphatase inhibition, indicating that the regulation of a phosphatase had been perturbed. Proliferation assays correlated well with the phosphorylation data and showed that the Y1251F mutant had an increased growth stimulus by IGF-1. The phosphatase affected by the Y1251F mutation was not identified, although PTP-1D has been ruled out, as a candidate. The next step would have been to test other phosphatases or inhibitors of phosphorylation including PTP-1B, and the SOCS family of proteins. Alternatively, a yeast 2-hybrid screen using the C-terminal domain as the ligand trap, could be used to determine which proteins aren't binding to the mutant receptors. A larger scale approach employing proteomic and mass spectral analysis or using microarrays would also address this work.

CHAPTER 6

COMPARATIVE MODELLING OF *C. ELEGANS* INSULIN-LIKE PEPTIDES

CHAPTER 6 COMPARATIVE MODELLING OF *C. ELEGANS* INSULIN-LIKE PEPTIDES

6.1 INTRODUCTION

The insulin/IGF system has been found to be well conserved in all vertebrates studied. It therefore represents an early system controlling the growth and metabolism of higher organisms. *Caenorhabditis Elegans* (*C. elegans*), the nematode worm, has been a widely used animal model for genetic and developmental studies. The fate of each and every cell has been mapped during its development and most of its genome has also been sequenced.

This species represents a present form of one of the last common ancestor of mammals and nematodes and establishes that the genetic circuitry controlling glucose or fat metabolism was already in place before the split.

The first evidence of an insulin-like signalling pathway existing in more primitive invertebrates was identified in *C. elegans* as a result of analysing mutants of genes that effected dauer formation ('daf' mutants). Dauer Larvae are in a stress resistant, metabolically inactive state that is normally caused by overcrowding or nutrient starvation and allows the larvae to remain alive for up to eight times the length of a normal two-week lifespan (See Figure 6.1). Upon the resumption of favourable conditions the dauer larvae undergoes further developmental changes and re-enters the normal lifecycle at the larval L4 stage, subsequently becoming an adult worm. Interestingly, larvae regardless of whether they enter the dauer state, still have the same lifespan as an adult. In effect, dauer formation increases lifespan albeit due to a form of hibernation. However, 'daf' mutants, had significantly extended lifespan as an adult even after leaving the dauer stage, probably because they retained some dauer-like characteristics. Sequencing the gene encoded by *daf-2*, whose loss caused all larvae to enter into the dauer state, identified a homolog of the mammalian insulin/IGF receptors (Kimura *et al.*, 1997). Similar to mammals, the insulin-like signalling pathway in *C. elegans* controls both glucose and fat metabolism, in addition to ageing.

It was later found that daf-2 operated upstream in a signalling pathway with several components common to the mammalian pathway (See Figure 1.7). These included age-1, a PI3K homolog (Morris *et al.*, 1996), akt-1, akt-2, pdk-1 (Paradis *et al.*, 1999; Paradis and Ruvkun, 1998), and daf-16, a forkhead transcription factor (Lin *et al.*, 1997; Ogg *et al.*, 1997). It is assumed that an insulin-like ligand binds the daf-2 receptor initiating a signalling cascade

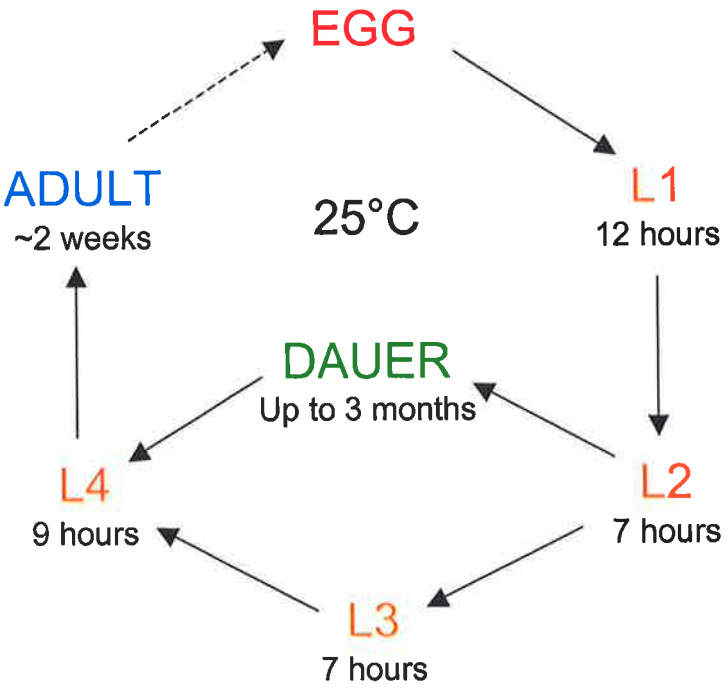


Figure 6.1 **Schematic illustration of the life-cycle of *C. elegans* genes. L= Larval stage**

that activates age-1, that in turn activates akt, which acts to block daf-16 from translocating to the nucleus and upregulating transcription of environmental stress resistance, development and dauer formation genes.

A search of the entire *C. elegans* genome by several groups for insulin-related peptides produced the surprising result of 37 candidate genes (Duret *et al.*, 1998; Gregoire *et al.*, 1998; Pierce *et al.*, 2001), designated *ins-1-37*. One further insulin-like gene has been recently discovered through sequencing the gene encoded by *daf-28* (Hua *et al.*, 2003), bringing the total to 38. This ins gene was not identified in the original search because the genomic sequences corresponding to the *daf-28* locus, were not contained in the database. All these genes contain an extended N-terminal signal sequence. Most importantly they contained either two or three predicted disulphide bond pairs formed from conserved cysteine residues in the B and A domains. Of the 38 genes, 34 had intron/exon organization identical to previously described insulin family members in mammals. Due to these sequences diverging from mammalian insulin's at conserved sites required for biological activity in mammals their classification as insulin-like has been presumptive.

The first insulin-like peptide to be cloned came from the gene *ins-18* which was predicted to contain a cleavable C-peptide. However, when expression was detected using a polyclonal serum raised against the peptide, in both reducing and non-reducing conditions, the peptide appeared as a 8.1kDa band consistent with having lost its N-terminal leader sequence but retaining contiguous B, C, and A domains (Kawano *et al.*, 2000). In this way it appeared to be more like the IGFs rather than insulin. Its expression was limited to embryonic and larval stages (except L1) and it was not detected in dauer larvae or in the adult worm. RNA-mediated interference (RNAi) of *ins-18*, resulted in a long-lived phenotype similar to some *daf-2* mutants, however unlike *daf-2,* had no effect on dauer larvae formation.

In a separate study it was found that an increase in gene dosage of *ins-1* results in promotion of dauer arrest, suggesting that Ins-1 is an antagonist of daf-2 signalling. Therefore, a high gene dosage of either *ins-1* or *ins-18* resulted in antagonism of daf-2 signalling and dauer arrest (Pierce *et al.*, 2001). In addition human insulin under the expression of the promoter from the 5′ *ins-1* flanking DNA also had similar antagonizing effects raising the possibility that *C. elegans* ins peptides may act as both antagonists and agonists of the daf-2 receptor. daf-28 was recently found to regulate daf-16 nuclear localization and therefore, likely functions as a ligand of DAF-2. A daf-28_GFP fusion protein was seen to be expressed in two

pairs of sensory neurons, and this expression was down-regulated in response to dauer-inducing environmental and sensory signals.

Insulin-like proteins have been found to be well conserved throughout vertebrate evolution and various studies have identified the specific residues and surfaces involved in receptor binding, and in the case of insulin, dimerisation and hexamerisation. The consensus surface residues responsible for receptor binding include GlyA1, Gln A5, Tyr A19, Asn A21, Val B12, Tyr B16, Gly B23, Phe B24, Phe B25, and Tyr B26, which also overlaps the dimerisation interface (Pullen *et al.,* 1976) (Baker *et al.,* 1988). There is additional evidence suggesting that Ile A2 and Val A3, which are not on the surface, become exposed and interact with the receptor after displacement of insulin's B chain during the binding process (Hua *et al.,* 1991) (Ludvigsen *et al.,* 1998). However it also clear that other residues outside of these surfaces contribute to receptor binding. In addition to many of the residues in Insulin, IGF-1 and IGF-2 require Ala 8, Asp 12, Phe 23, and Tyr 24 in the B domain, Tyr 31, Arg 36, Arg 37, in the C domain, and Met 59, and Tyr 60 in the A domain (Adams *et al.,* 2000). The extra C and D domains in IGF-1 and IGF-2 confer mainly IGF-1R specificity and reduce their affinity for the IR. As mentioned above the *C. elegans* Ins peptides lack almost every residue found to be important for receptor binding with a low overall similarity to their mammalian homologues. This is surprising considering the much higher similarity between the Daf-2 and hIR and hIGF-1R, unless the overall 3D structure is retained in the ligands.

6.1.1 Project Summary and Aims

The *C. elegans* insulin-like peptide family is a very large family consisting of peptides likely to be both agonists and antagonists of daf-2 Receptor signalling. An analysis and understanding of the 3D structure of Ins peptides and why some are agonists and others antagonists may enable future researchers to design mammalian insulin-like proteins that act as antagonists of the IGF-IR that would be clinically useful proteins to treat IGF-dependent cancer tissues.

AIMS:
a) Determine the 3D structure of all the ins peptides by homology modelling.
b) Choose several ins gene targets to RT-PCR and clone.

6.2 RESULTS AND DISCUSSION

6.2.1 Analysis of Ins Peptide Sequences

Initially, the 38 theoretical ins proteins were aligned using CLUSTAL W 1.82. For the purposes of the alignment the three insulin-like motifs from ins-31 were treated as separate sequences. The results show that the ins proteins can cluster into several classes according to the number and arrangement of cysteine residues they contain. Generally, those *ins* genes that are found next to each other on the chromosome were seen to be the most closely related suggesting that they were likely formed from recent gene duplications. Duret *et al.*, 1998 and Gregoire *et al.*, 1998, have suggested that there are 3 classes formed based on their alignments (α, β, γ). The α class peptides have an additional disulphide pair, B23/A23 (Human Insulin numbering), that is not seen in mammalian insulin-like proteins, although they are lacking the cysteine pair at positions A6/A11. In place of this pair are usually either non-polar, or large aromatic residues. The β class peptides have the additional cysteine pair as well retaining the A6/A11 pair. γ class peptides follow the three canonical cysteine pair structure conserved in mammalian insulin and the IGFs. The additional disulphide present in many of the ins peptides tethering cysteines at position 84 and 111, corressponds in human insulin to residues Gly B23 and Asn A21 which are also in close proximity.

Interestingly, only two peptides (ins-1(beta class) and ins-18(alpha class)) are predicted to contain a cleavable C-peptide between the B and A domains although as mentioned above ins-18 was expressed as a single uncleaved peptide (Kawano *et al.*, 2000). Ins-31 is perhaps the most unusual gamma-class peptide having three insulin chains linked together that if capable of correctly folding would likely form a trimeric insulin-like molecule. One could predict that if expressed this protein would probably act as an antagonist similar to dimeric or hexameric insulin in mammals.

6.2.2 Comparative Modelling of Ins Peptides

Given the large number of *C. elegans* ins proteins and the wealth of detailed information on the structure of insulin-like protein domain fold, the quickest and most practical way of analysing the structure these sequences would likely adopt, was to apply comparative modelling methods. Comparative (or Homology) modelling of a protein sequence based on the known 3D structure of a related protein can provide useful information about the

functionality of a protein. Comparative modelling is the only method available that can reliably predict the 3D structure of a protein with an accuracy similar to that of a low-resolution experimentally determined structure (Marti-Renom *et al.*, 2000). The four stages of the comparative modelling process are template selection, target-template alignment, model building and model evaluation (See Figure 6.2). Of these stages template selection and alignment accuracy have the greatest impact on model quality for cases where there is <40% sequence identity between target and template. As a result, these two stages were given the most attention.

6.2.2.1 Selection of Modelling Templates

The first stage of the comparative modelling process is to select a suitable class of proteins with a particular fold from which to select a template to base the theoretical structure. In the case of ins proteins this step is unnecessary as the template must belong to the insulin family of proteins. Initially, potential templates in the Protein Data Bank (PDB) (contains proteins whose structures have been solved by either X-ray crystallography or NMR spectroscopy) were searched for with WU-BLAST 2.0 using each Ins protein sequence as the query. The relatively low sequence similarity across the entire Ins peptide sequences resulted in some cases of the search program bringing up non-insulin related proteins as good matches. Such proteins included phospholipase A2, Ribosomal Protein S4. The absolute scores for these proteins however were low and also only short 10-30aa stretches of the query peptide were lined up. Changing the parameters of WU-BLAST 2.0 (ie. Gap penalties, search matrices) did very little to change the results. Therefore instead of allowing a program to pull out sequences from the PDB, the different insulin-like proteins were manually extracted from the ~150 existing structures. Analysis of these structures showed that only eight different insulin-like peptide structures had been solved. Representatives to be used in the homology modelling, were chosen based on the quality of the structure (highest resolution of X-ray derived structures, largest number of restraints/residue for NMR derived structures) and in the case of insulin, to minimise the number of mutations (unmutated insulin forms dimers and hexamers at the concentrations used in structural analysis).

The sequences of these potential templates were obtained from the Protein Data Bank and compared to each ins protein for the best possible match using CLUSTAL W 1.83. The multiple sequence alignment and percentage identity table (see Appendix table A.) show that the ins peptides cluster together and the majority are most similar to either human relaxin or silkmoth bombyxin-II. The highest maximum full-length percentage identity was 34% (ins-6

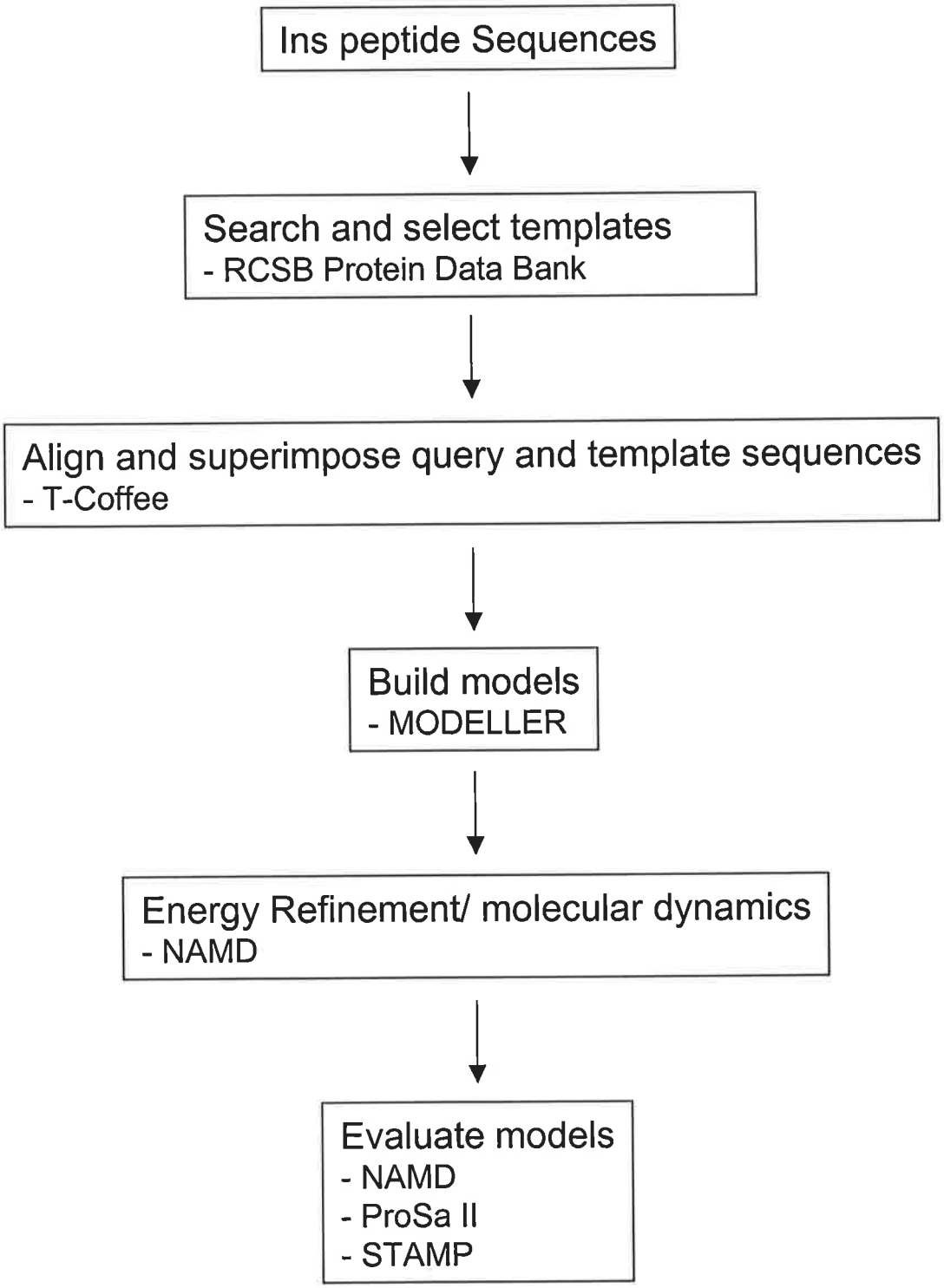


Figure 6.2 Flowchart of the Comparative Modelling process

with Bombyxin-II) and the lowest 13% (Ins-10 with insulin (human)) with an average maximum of 22.9%. This range of sequence identity puts these proteins firmly in the 'twilight' zone for protein modelling, a description given to modelling done with sequence similarities of less than 40% (Marti-Renom *et al.*, 2000).

During the analysis of the PDB searches it was observed that in most cases the stretch of amino acids that was aligned from the query corresponded only to the A chain. This raised the possibility that the A and B chains may have different 'best' template sequences than the full-length peptide. An analysis of edited ins peptide sequences looking individually at either the A or B chain, using CLUSTAL W revealed that the average maximum sequence similarity with the best template was significantly increased in the A chain (28.9%) over the B chain (22.4%) and the full length peptide (22.9%) (See Appendix A.). The other striking difference was that the vast majority of the domain A sequences were most similar to either Relaxin or Bombyxin-II. In addition the best templates for the majority of the Ins proteins differed between full length, A and B domains. Generally, the greater the overall sequence similarity the better the final model therefore it was decided that, when appropriate, two templates would be used for a single ins peptide. Comparative modelling can be successfully performed using two templates with a realistic outcome for separate parts of a target sequence, provided the structured sequences do not overlap and one has some way of aligning them. For insulin-like proteins the disulphide pattern is a defining feature and these inter-residue bonds bridge the A and B domains, which would allow the modelling program to orient the two separate structures.

However in the case of the ins proteins it was also likely to be important that templates closely matched the length of the conserved B and A domains so as to reduce the amount of protein sequence that has no equivalent in the template.

6.2.2.2 Target-Template Alignment

Following the template selection the target sequences were aligned with the chosen template. This is an extremely important step as there is currently no modelling procedure or program that can produce an accurate model from an initial misalignment. In an initial attempt the multiple sequence alignment program ALIGN2D (within MODELLER, Andrej Sali, Rockefeller University) was used since MODELLER was going to be used to make the models. The ALIGN2D method takes into account structural information from the template when making the alignment. This is achieved through use of a variable gap penalty that places gaps in solvent exposed and curved regions, which lie outside secondary structure segments

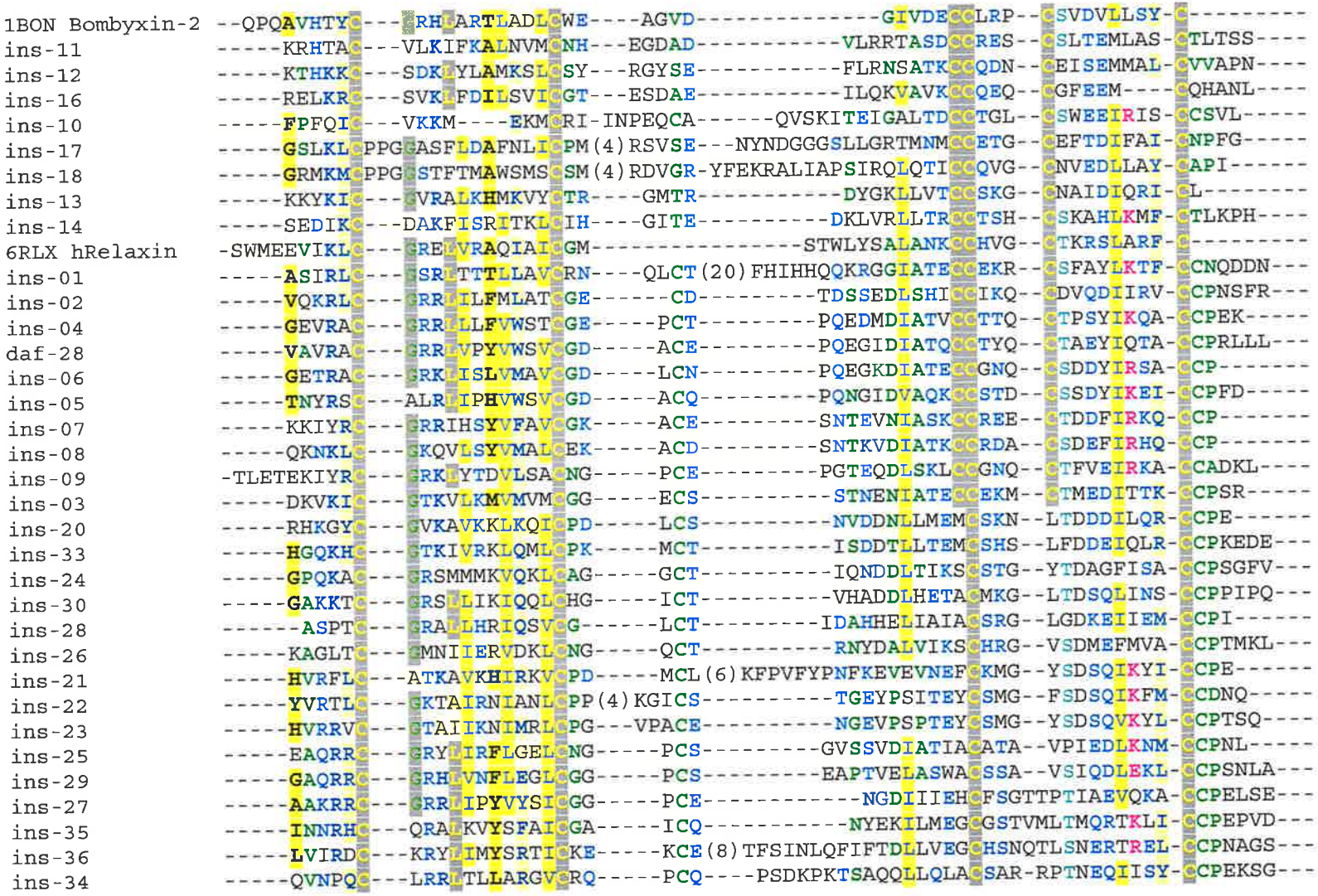
and between two Cα atoms of individual amino acids that are close in space. As a result alignment errors should be reduced by ~1/3 compared to standard alignments (Marti-Renom *et al.*, 2000). Despite this ALIGN2D did not align the cysteine's correctly and so another method was attempted. Clustal W although better than ALIGN2D still had errors in the alignment of many cysteine's and also introduced large gaps in the alignment. The overall results of the ALIGN2D alignment however, when examined, contained a large number of gaps that did not correspond well to the predicted separation of the B and A chains. A second program T-Coffee was then tested. This relatively new program uses a two-stage method that performs some pre-processing on all the pair-wise alignments to aide in the second stage progressive alignment. Intermediate alignments were therefore based not only on the next sequence to be added to the alignment but also on how all sequences align with each other. T-Coffee produced the best alignment with the fewest number of gaps, with only the last cysteine in some cases requiring a manual adjustment (Notredame *et al.*, 2000). The final target-template alignments are shown in figure 6.3, and consensus residues coloured using the program Chroma (Goodstadt and Ponting, 2001).

6.2.2.3 Building the Models

Protein structures for the 38 predicted insulin-related peptides were modelled using the program MODELLER (Andrej Sali, Rockefeller University). MODELLER constructs the 3D structure of the target protein based on the satisfaction of spatial restraints by comparing the known structure of the template protein to the aligned sequence of the target. The restraints, including distance, angles, dihedral angles and pairs of dihedral angles, are used to produce a structure containing all mainchain and sidechain non-hydrogen atoms. In the case of Ins peptides that were predicted to have an extra disulphide bond, an additional restraint was added to the modelling procedure to force the formation of the disulphide bond.

It has been observed that in some cases the use of multiple templates equidistant from a target increases the accuracy of the model. An assessment of this technique was made using ins-6 as the target and bombyxin-II and hIGF-1 as the two simultaneous target templates and comparing the single model obtained to the models made separately from Bombyxin-II and hIGF-1. The objective function of the model based on the two templates was far higher than either of the two single template models and it was therefore decided that this would not be a suitable method for modelling these peptides (data not shown).

Since it was clear from the analysis of the alignments of the individual B and A domains that most ins peptides had different proteins they were most similar to for each domain, a hybrid


1BON Bombyxin-2 --QPQAVHTYC---RHCARTLADLCWE----AGVD-------------GIVDECCLRP--CSVDVLLSY-C---------
ins-11 -----KRHTAC---VLKIFKALNVMCNH---EGDAD-----------VLRRTASDCCRES--CSLTEMLAS-CTLTSS----
ins-12 -----KTHKKC---SDKLYLAMKSLCSY---RGYSE-----------FLRNSATKCCQDN--CEISEMMAL-CVVAPN----
ins-16 -----RELKRC---SVKIFDILSVICGT---ESDAE-----------ILQKVAVKCCQEQ--CGFEEM----CQHANL----
ins-10 -----FPFQIC---VKKM----EKMCRI-INPEQCA------QVSKITEIGALTDCCTGL--CSWEEIRIS-CCSVL-----
ins-17 -----GSLKLCPPGGASFLDAFNLICPM(4)RSVSE---NYNDGGGSLLGRTMNMCCETG--CEFTDIFAI-CNPFG-----
ins-18 -----GRMKMCPPGCSTFTMAWSMSCSM(4)RDVGR-YFEKRALIAPSIRQLQTICCQVG--CNVEDLLAY-CAPI------
ins-13 -----KKYKICC---GVRALKHMKVYCTR----GMTR-----------DYGKLLVTCCSKG--CNAIDIQRI-CL--------
ins-14 -----SEDIKC---DAKFISRITKLCIH----GITE----------DKLVRLLTRCCTSH--CSKAHLKMF-CTLKPH----
6RLX hRelaxin -SWMEEVIKLC---GRELVRAQIAICGM----------STWLYSALANKCCHVG--CTKRSLARF-C---------
ins-01 -----ASIRLC---GSRLTTTLLAVCRN----QLCT(20)FHIHHQQKRGGIATECCEKR--CSFAYLKTF-CCNQDDN---
ins-02 -----VQKRLC---GRRLILFMLATCGE------CD---------TDSSEDLSHICCIKQ--CDVQDIIRV-CCPNSFR---
ins-04 -----GEVRAC---GRRLLLFVWSTCGE-----PCT---------PQEDMDIATVCCTTQ--CTPSYIKQA-CCPEK-----
daf-28 -----VAVRAC---GRRLVPYVWSVCGD-----ACE---------PQEGIDIATQCCTYQ--CTAEYIQTA-CCPRLLL---
ins-06 -----GETRAC---GRKLISLVMAVCGD-----LCN---------PQEGKDIATECCGNQ--CSDDYIRSA-CCP-------
ins-05 -----TNYRSC---ALRLIPHVWSVCGD-----ACQ---------PQNGIDVAQKCCSTD--CSSDYIKEI-CCPFD-----
ins-07 -----KKIYRC---GRRIHSYVFAVCGK-----ACE---------SNTEVNIASKCCREE--CTDDFIRKQ-CCP-------
ins-08 -----QKNKLC---GKQVLSYVMALCEK-----ACD---------SNTKVDIATKCCRDA--CSDEFIRHQ-CCP-------
ins-09 -TLETEKIYRC---GRKLYTDVLSACNG-----PCE---------PGTEQDLSKLCCGNQ--CTFVEIRKA-CCADKL----
ins-03 -----DKVKIC---GTKVLKMVMVMCGG-----ECS----------STNENIATECCEKM--CTMEDITTK-CCPSR-----
ins-20 -----RHKGYC---GVKAVKKLKQICPD-----LCS----------NVDDNLLMEMCSKN--LTDDDILQR-CCPE------
ins-33 -----HGQKHC---GTKIVRKLQMLCPK-----MCT----------ISDDTLLTEMCSHS--LFDDEIQLR-CCPKEDE---
ins-24 -----GPQKAC---GRSMMMKVQKLCAG-----GCT----------IQNDDLTIKSCSTG--YTDAGFISA-CCPSGFV---
ins-30 -----GAKKTC---GRSLLIKIQQLCHG-----ICT----------VHADDLHETACMKG--LTDSQLINS-CCPPIPQ---
ins-28 ------ASPTC---GRALLHRIQSVCG------LCT---------IDAHHELIAIACSRG--LGDKEIIEM-CCPI------
ins-26 -----KAGLTC---GMNIIERVDKLCNG-----QCT----------RNYDALVIKSCHRG--VSDMEFMVA-CCPTMKL---
ins-21 -----HVRFLC---ATKAVKHIRKVCPD-----MCL(6)KFPVFYPNFKEVEVNEFCKMG--YSDSQIKYI-CCPE------
ins-22 -----YVRTLC---GKTAIRNIANLCPP(4)KGICS----------TGEYPSITEYCSMG--FSDSQIKFM-CCDNQ-----
ins-23 -----HVRRVC---GTAIIKNIMRLCPG---VPACE----------NGEVPSPTEYCSMG--YSDSQVKYL-CCPTSQ----
ins-25 -----EAQRRC---GRYLIRFLGELCNG-----PCS---------GVSSVDIATIACATA--VPIEDLKNM-CCPNL-----
ins-29 -----GAQRRC---GRHLVNFLEGLCGG-----PCS---------EAPTVELASWACSSA--VSIQDLEKL-CCPSNLA---
ins-27 -----AAKRRC---GRRLIPYVYSICGG-----PCE----------NGDIIIEHCFSGTTPTIAEVQKA-CCPELSE---
ins-35 -----INNRHC---QRALKVYSFAICGA-----ICQ----------NYEKILMEGCGSTVMLTMQRTKLI-CCPEPVD---
ins-36 -----LVIRDC---KRYLIMYSRTICKE-----KCE(8)TFSINLQFIFTDLLVEGCHSNQTLSNERTREL-CCPNAGS---
ins-34 -----QVNPQC---LRRLTLLARGVCRQ-----PCQ----PSDKPKTSAQQLLQLACSAR-RPTNEQIISY-CCPEKSG---

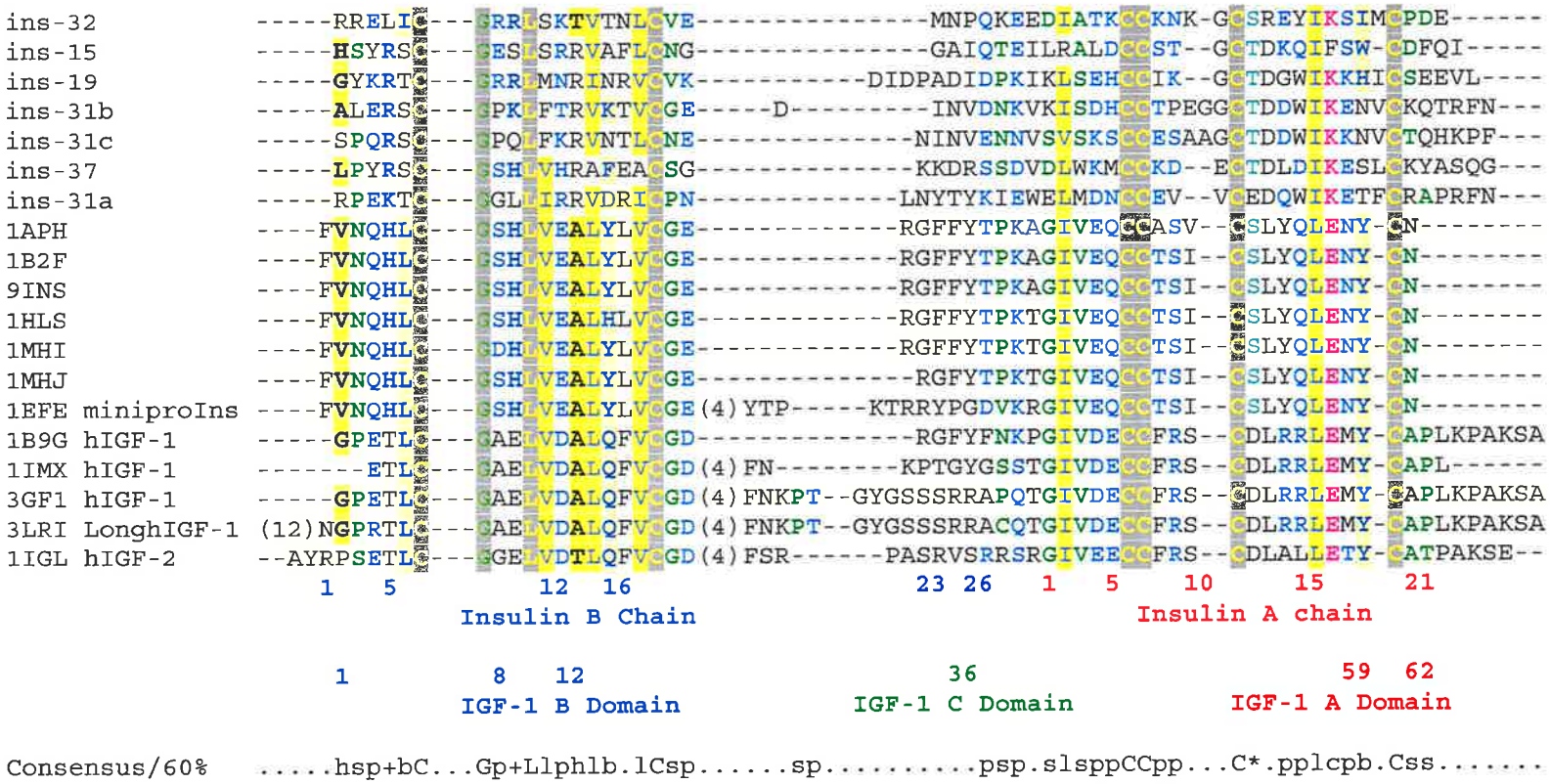


Figure 6.3 **Multiple Sequence Alignment of all 38 ins peptides and the templates extracted from the PDB database.**

The alignment was performed by the program T-Coffee and residues coloured using Chroma. Long N-terminal sequences in the Ins peptides were deleted before the final model construction.

model was made, again using ins-6 as the target. hIGF-1 B domain and bombyxin-II A domain were aligned separately to ins-6 ensuring that there was no overlap in the template sequences. Cysteine pairs were also explicitly made to cross the templates where appropriate. The resultant model however did not have a recognisable fold and the A and B domains were not aligned in an insulin-like fold. It was therefore concluded that this method would also not be suitable for the modelling process.

Models of each ins peptide were then produced based individually on each of the selected templates, rather than choosing an individual template for each peptide. The rationale being that there was no clear method of choosing a 'best' template that didn't rely on sequence alignment. It was unclear as to how important sequence identity would be considering how low the overall percentage similarities were and it was also unclear how much template quality would affect the final model.

The models produced in MODELLER were scored on their objective function, which is related to how well the model agrees with a calculated set of restraints, with the lower the score meaning a model which satisfies the restraint criteria better. According to the MODELLER programmers a score of <300 should be an accurate model equivalent to a low resolution structure, if the alignment is perfect.

The average lowest objective function obtained for each peptide is 508 (See Appendix B.). The best objective function score for any peptide was 232 for ins-11 modelled on relaxin. The template that produced the highest number of 'best' objective function scores was the crystal structure of IGF-1 (1IMX). The stereochemical quality of each model and the templates was analysed by Procheck (Laskowski, 1993)(See Appendix C.). The quality of the models correlated well with the quality of the template it was modelled on, which would in part explain why the 1IMX IGF-1 template generally produced higher quality models. The objective function values also loosely correlated with the Procheck percentage (core) values, as those models with a higher objective value tended to have better Procheck scores as shown by the trend lines sloping down as the objective function increased.

Several observations can be made of the overall results. Firstly, on an individual basis the percentage identity between the target and the template did not generally correlate with the quality of the model produced. It therefore appears that the alignment of the key residues, in this case the cysteines and chemically similar residues is the most important to produce an accurate model. The presence of a longer linker region between B and A domains did not

particular affect the result as shown by ins-1 having the second lowest average obective function, indicating it could be modelled well on a number of the templates. Interestingly, most ins peptides had clear preferences for individual templates and had minimum objective functions much lower than their average. Generally, ins peptides from the same genomic location had the same best template, correlating with them sharing the most sequence identity with other members of their cluster.

Since the sterochemical quality of the template had a strong effect on the quality of the model, in order to relax the model structures and allow the models to adopt a more favourable structure for each sequence, refinement was performed by conjugate gradient minimisation and simulated annealing.

6.2.2.4 Refinement of the Models

The models were refined by simulated annealing. Proteins for refinement were prepared in XPLOR by solvation in a 4Å wide shell of TIP3 water, addition of hydrogen atoms, and setting the correct cysteines to be disulphide bonded and the structure and parameter files calculated using a CHARMM 22 forcefield (MacKerell, *et al.*, 1998). Minimisation was performed in NAMD, and consisted of an initial conjugate gradient minimisation, followed by simulated annealing and a final conjugate gradient minimisation step. An attempt at minimisation setting the parameters to using point harmonic restraints and removing disulphide bond restraints resulted in too much backbone relaxation and in the majority of cases secondary structure was destroyed and significant reductions in model quality were observed (Data not shown). It was therefore reasoned that the disulphide bond linkages should be enfocorced to maintain overall structure and instead remove point harmonic restraints to allow some movement in the backbone (See results below).

6.2.2.5 Evaluation of the Models

The models both before and after refinement were then evaluated on their potential energy, both empirically (in NAMD) and in a 'knowledge based' manner (in ProSa II). Empirical energies calculated in NAMD using the CHARM22 forcefield, include covalent bond energies, dihedral angles, van der waals forces and electrostatic energies, which are a measure of how energetically stable a structure is and whether there are any unfavourable clashes in the different fields. Output from ProSa II is in the form of a z-score which approximates the difference in free energy of an evaluated model and the mean free energy of the same sequence threaded through unrelated folds, expressed in units of SD. Z-scores were converted to q-scores to obtain an energy value independent of peptide length (See 2.3.5.4, and

Appendix) (Sanchez and Sali, 1998). A lower value for both empirical energy and q-scores indicates a more stable and favourable structure.

The results are given in Appendix D. and E., and a summary of the best structures according to consensus scoring of the empirical energy and q-scores is shown in figure 6.4. In the majority of cases refinement of the models succeeded in improving both empirical energies and q-scores.

Ins-10 was predicted to be highly unlikely to be able to form an insulin-like structure based on a much shorter than unusual amino acid spacing between the first two cysteine residues which would normally form the B chain α-helix.. None of the models made for ins-10 had three well defined helices and its best model based on relaxin, even though having reasonable q and empirical values, structurally only had 2 of the 3 insulin fold helices and in conclusion this peptide is probably non-functional. For the remaining models, even though the percentage similarities between the ins peptides and the templates were consistently less than 30% and in some cases as low as 5%, reasonable models were able to be made containing the 3 alpha helices and disulphide bonds arranged in a consistent way.

6.2.2.6 Model Comparisons

Due to space constraints in this thesis, rather than examine every peptide model individually, I will discuss only selected ones. During this work, the NMR structure of ins-6 was solved but the coordinates have yet to be released. An email to the author requesting them was unsuccessful. Therefore, based on a visual comparison with figures from the journal article, it appears that the models most closely resembling the NMR structure were based on insulin templates. Therefore the two best representatives of the previously predicted peptide families based on an insulin template will be discussed.

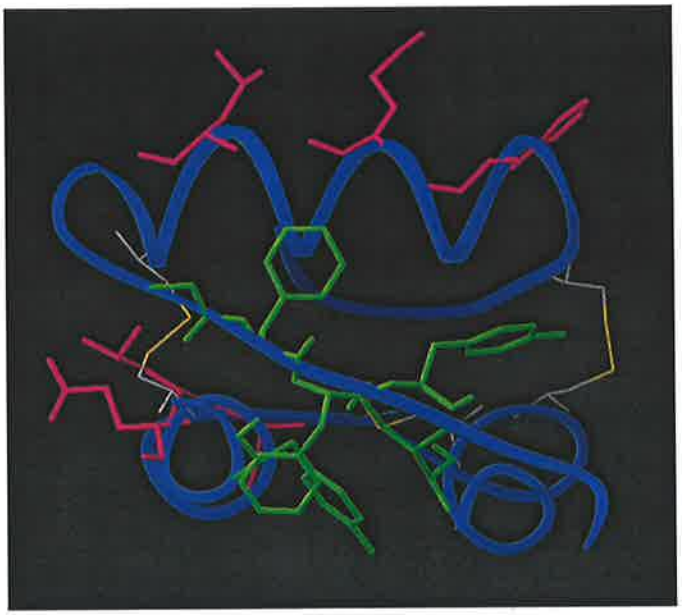
Ins-25 and ins-28 are the two highest scoring representatives of the α-class peptides based on an insulin template. Their best models were both based on the hInsulin structure 1MHJ, with empirical scores of –823 and –401, and very good q-scores of –1.79 and –2.05 respectively. Compared to insulin, the overall topology of the ins-25 and ins-28 models is more globular with a lack of defined flatter surfaces, which in insulin's case are used to both bind the receptor and to allow dimerisation and hexamerisation (See figure 6.5 and 6.6). The charge distribution on the molcules is also different. The key residues in insulin found to most greatly influence receptor binding are not well conserved in either ins-25 or ins-28. Based on the

					Modeling			
	Residues	Chromosome	Locus	Class	% Identity in aligned residues	Free Energy	q-score	Best Template
DAF-28	54	V	23.52	β	24	-642	-1.48	hIGF-1 (3GF1)
INS-1	80	IV	4.6	β	31	-1418	-1.05	pInsulin (9INS)
INS-2	53	II	-0.74	β	22	-1040	-1.49	hIGF-1 (1B9G)
INS-3	51	II	-0.74	β	19	-316	-1.62	bInsulin (1APH)
INS-4	52	II	-0.72	β	23	-499	-1.65	hIGF-1 (3GF1)
INS-5	52	II	-0.72	β	21	-420	-1.65	bInsulin (1APH)
INS-6	50	II	-0.72	β	23	-796	-1.85	hRelaxin (6RLX)
INS-7	50	IV	4.37	β	19	-936	-1.79	hInsulin (1HLS)
INS-8	50	IV	4.37	β	18	-641	-1.66	hIGF-1 (1B9G)
INS-9	57	X	0.43	β	7	-905	-1.27	hInsulin (1MHI)
INS-10	55	V	5.96	β	7	-249	-1.31	hRelaxin (6RLX)
INS-11	53	II	-1.16	γ	22	-814	-1.29	hRelaxin (6RLX)
INS-12	53	II	-1.17	γ	19	-452	-1.31	hInsulin (1MHI)
INS-13	48	II	-1.19	γ	15	-643	-1.39	bInsulin (1APH)
INS-14	53	II	0.08	γ	19	-438	-1.17	bInsulin (1APH)
INS-15	51	II	0.08	γ	23	-866	-1.53	hRelaxin (6RLX)
INS-16	50	III	-16.58	γ	18	-716	-1.42	hIGF-1 (1B9G)
INS-17	67	III	-3.17	γ	16	-1297	-1.49	hIGF-2 (1IGL)
INS-18	68	I	2.61	γ	21	-1539	-1.13	hIGF-1 (3GF1)
INS-19	57	II	-8.39	γ	15	-1092	-1.26	hInsulin (1MHI)
INS-20	50	II	-0.7	α	19	-861	-1.55	hRelaxin (6RLX)
INS-21	63	III	1.42	α	9	-309	-1.13	hInsulin (1MHI)
INS-22	57	III	1.43	α	9	-435	-1.43	hIGF-1 (3GF1)
INS-23	54	III	1.44	α	14	-373	-1.52	pInsulin (9INS)
INS-24	53	I	22.54	α	17	-96	-1.62	bInsulin (1APH)
INS-25	52	I	22.51	α	21	-823	-1.79	hInsulin (1MHJ)
INS-26	53	I	22.56	α	10	-708	-1.52	hminiproins (1EFE)
INS-27	53	I	22.53	α	12	-491	-1.43	hIGF-1 (1B9G)
INS-28	49	I	22.49	α	12	-401	-2.05	hInsulin (1MHJ)
INS-29	54	I	22.5	α	26	-473	-1.99	hIGF-1 (3GF1)
INS-30	53	I	22.55	α	13	-175	-1.56	hInsulin (1MHI)
INS-31a	56	II	-8.38	γ	15	-1523	-1.40	hInsulin (1MHI)
INS-31b	57	II	-8.38	γ	23	-1128	-1.63	pInsulin (9INS)
INS-31c	57	II	-8.38	γ	17	-773	-1.63	bInsulin (1APH)
INS-32	52	II	-6.28	γ	16	-1138	-1.56	hIGF-1 (3GF1)
INS-33	53	I	21.44	α	20	-565	-1.50	hIGF-1 (1IMX)
INS-34	60	IV	11.34	α	13	-1219	-1.32	hIGF-1 (3GF1)
INS-35	54	V	24.98	α	15	-563	-1.19	pInsulin (9INS)
INS-36	70	I	7.44	α	15	-1692	-1.04	hInsulin (1HLS)
INS-37	55	II	20.56	γ	18	-805	-1.50	hminiproins (1EFE)

Figure 6.4 Summary of the results of simulated annealing

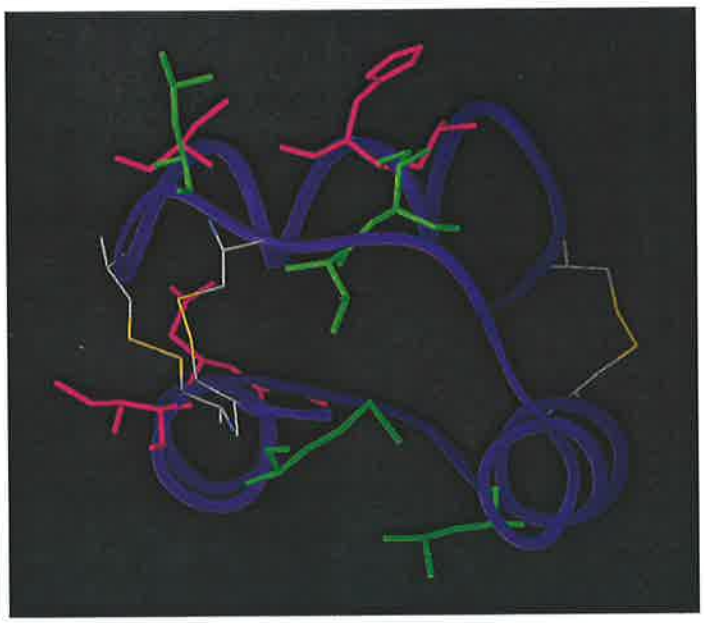
Chromosome locations of Ins genes obtained from [illegible]. Class Types: α: 3 cysteine pairs including B23/A23 but lacking A6/A11 (hInsulin numbering) β: 4 cysteine pairs including the additional B23/A23 pair. γ: 3 cysteine pairs conserved in mammalian insulin. Percentage Identity was obtained from ClustalW alignments.The best templates were determined by consensus scoring both Free Energy and q-scores. Free Energy values were calculated in NAMD using a CHARM22 forcefield. Q-scores were derived from z-scores calculated in ProSa 2 v3.0.

A.



B.

C.



α-helix α-helix α-helix

```
ins-25    -EAQRRCGRYLIRFLGELCNG-PCSGVSSVDIATIACATAVPIEDLKNMCCPNL
ins-28    --ASPTCGRALLHRIQSVCG--LCTIDAHHELIAIACSRGLGDKEIIEMCCPI-
pInsulin FVNQHLCGSHLVEALYLVCGERGFFYTPKAGIVEQCCTSICSLYQLENYCN---
          1   5       12  16     23 26   1   5    10   15    21
                insulin B Chain             insulin A chain
```

Figure 6.5 Comparison of Models of ins-25 and ins-28 with porcine Insulin

Ins models were made in MODELLER, and refined by simulated annealing. Residues highlighted correspond to those found to affect binding to the insulin receptor.
A. pInsulin : 9INS
B. ins-25 : Modelled on 1MHJ (hInsulin)
C. ins-28 : Modelled on 1MHJ (hInsulin)

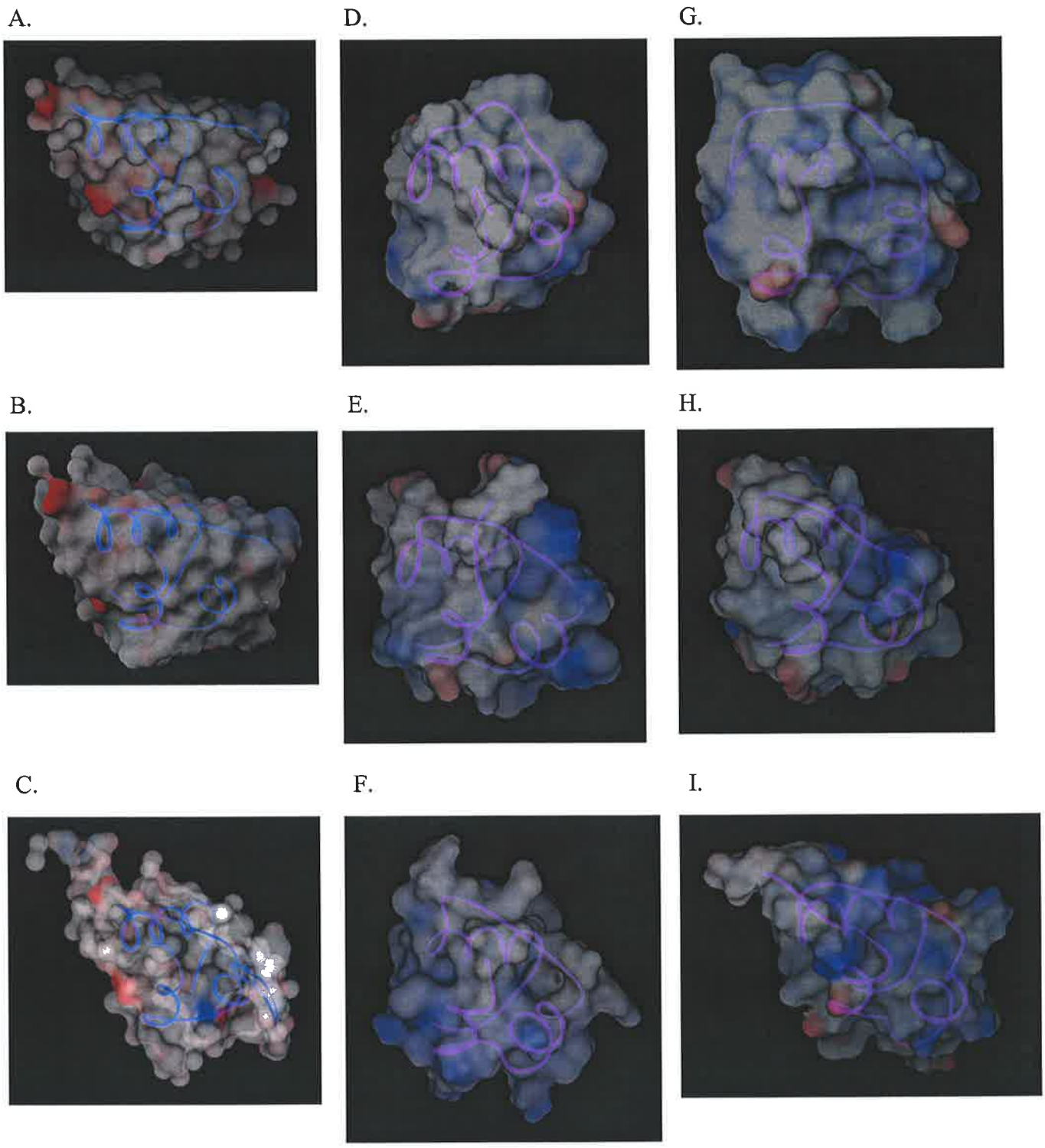


Figure 6.6 Structures of selected templates and ins peptide models with solvent accessible surface showing.

Backbone is displayed as a ribbon structure overlayed with a solvent accessible surface, calculated using a 1.4 A probe and indicating electrostatic potential.

A. bInsulin (1APH)	D. ins-25	G. ins-28
B. pInsulin (9INS)	E. ins-5	H. ins-6
C. IGF-1 (1IMX)	F. ins-13	I. ins-31b

orientation of the helices, one would predict that ins-28 more closely approximates insulin than ins-25 and may be more likely to bind the insulin receptor.

The first structure of a *C. elegans* ins peptide has been recently solved (Hua *et al.,* 2003). The target was ins-6 and the structure was determined by standard NMR techniques. Unlike the strategy hoped to be adopted in this thesis (See section 6.2.3) their ins-6 peptide was synthesized as a 50 residue polypeptide derived from two fragments, joined by native ligation. The secondary structure of the protein consisted of three insulin-like helices and the predicted four disulphide bridges.

Ins-5 and ins-6 are β-class peptides, and chromosomally located adjacent to each other. They are likely to be the result of a recent gene duplication and have quite high sequence identity and would therefore be predicted to have similar structures. Although the top ranking ins-6 structure was based on relaxin, a visual comparison with the published NMR structure indicated that the models based on insulin templates were more similar (See Figures 6.7 and 6.8). A structural alignment could not be performed as the atomic coordinates have not been released. The largest difference between the model and the NMR structure are in the A-chain helices, that in the model, are oriented differently compared the B-chain helix and are of different shape. The position of the disulphide bonds are however quite close. When compared to insulin, the ins-6 model appears more similar in overall topology and in the shape of the long B-chain helix, therefore one might predict that ins-6 would bind better to the insulin receptor than ins-5.

The biological functionality of Ins-6 was tested in a competition assay with iodinated human insulin for binding to Insulin Receptor derived from human placental membranes (mostly hIR-B isoform). The experiment yielded an EC_{50} value ~100 fold lower than insulin, similar to that of hIGF-1 (Hua *et al.,* 2003). This was unexpected as the ins-6 sequence diverges considerably from the human insulin residues Phe B24, Phe B25, Val A3, and Tyr A19 which have been shown to be critical to its activity (Baker, 1988) (De Meyts, 2002). In addition substitution of Phe B24 or Phe B25 by non-aromatic amino acids impairs receptor binding, and the respective substitutions of Tyr A19 by Phe and Ala impair activity by 8- and 1000-fold (Mirmira et al., 1991) (Du and Tang, 1998) (Kristensen *et al.*, 1997). Also, substitution of Leu B6 and Val A3 by Ala also hinders activity by 40- and 70-fold (Nakagawa and Tager, 1991) (Nakagawa and Tager, 1992). The divergence at these positions of ins-6, if considered one at a time, would be expected to block binding to the insulin receptor, yet its overall binding is only reduced 100-fold. This means that it is the multiple amino acid substitutions

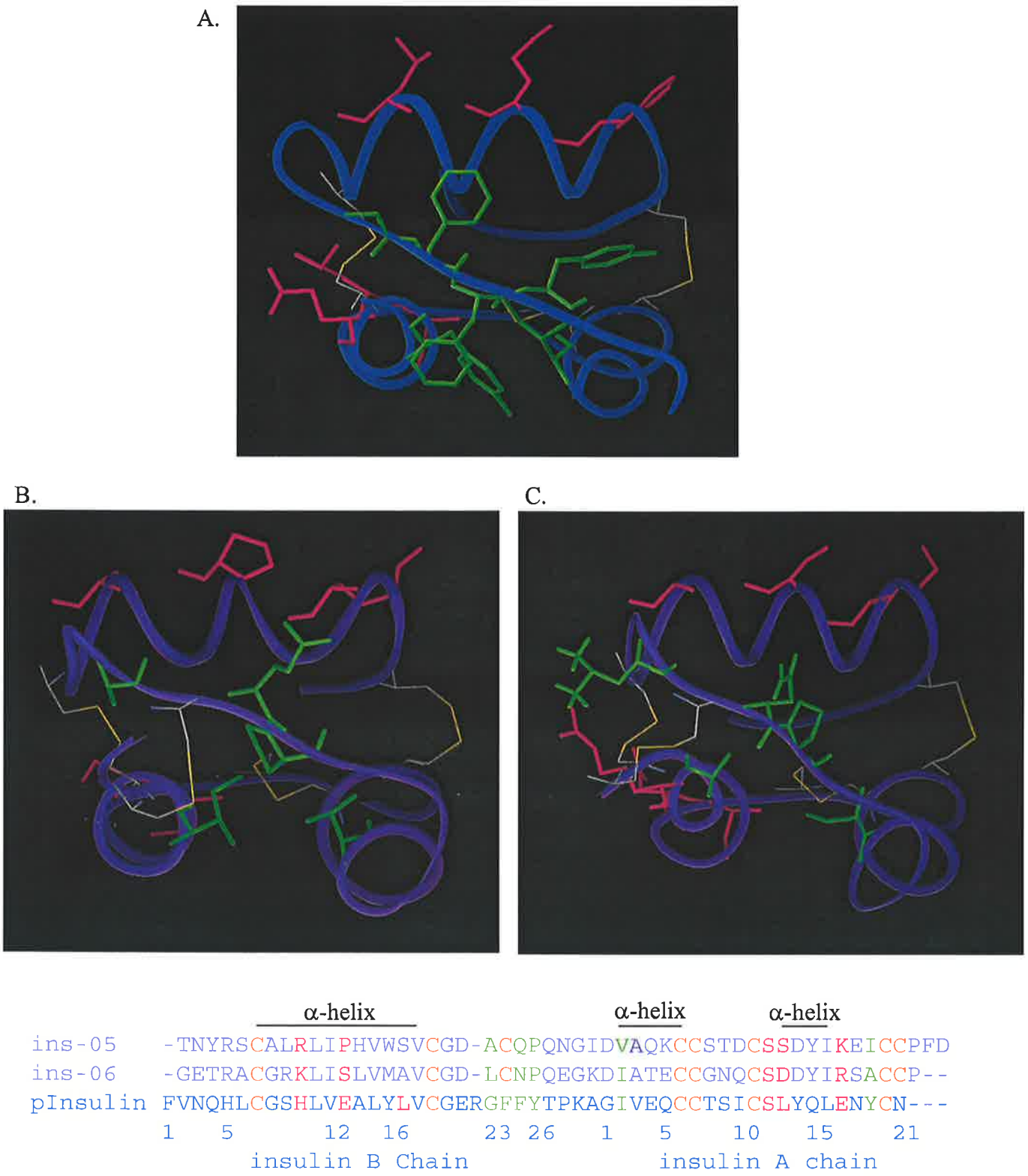
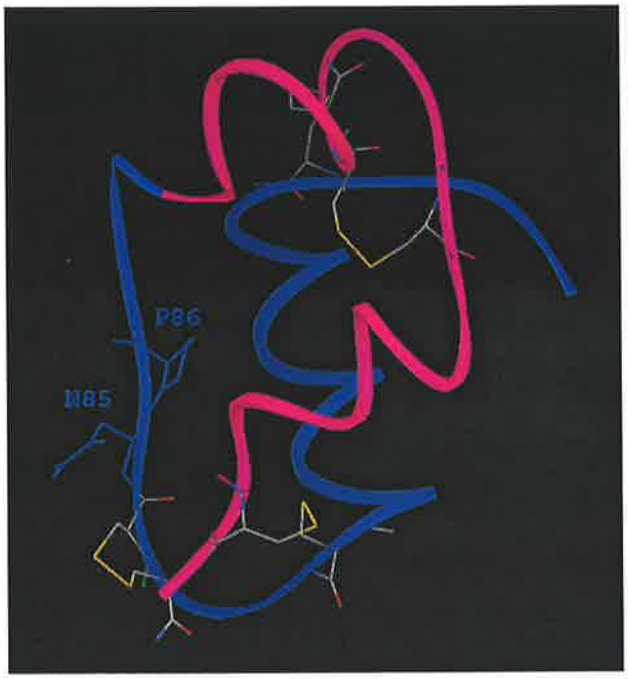


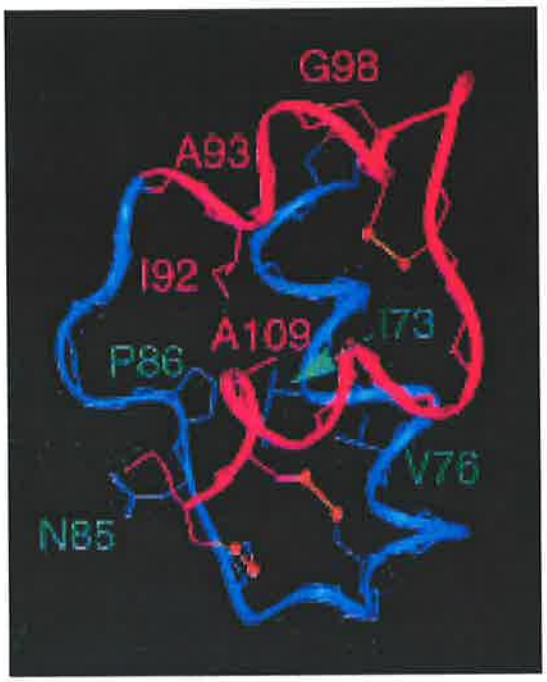
Figure 6.7 Comparison of Models of Ins-5 and Ins-6 with porcine Insulin

Ins models were made in MODELLER, and refined by simulated annealing. Residues highlighted correspond to those found to affect binding to the insulin receptor.
A. pInsulin : 9INS
B. ins-5 : Modelled on 1APH (bInsulin)
C. ins-6 : Modelled on 9INS (pInsulin)

A.



C.



B.

Figure 6.8 Comparison of ins-6 models with the solved NMR structure.

A. ins-6 (9INS)
B. Alignment of ins-6 models based on insulin templates.
C. ins-6 solved by NMR analysis (Hua *et al.*, 2003)

that together compensate for the lack of any one particular residue and allow a functional insulin-like fold to be maintained. At the concentration used in the NMR experiment INS-6 is monomeric and it is clear from the structure that it differs substantially from insulin on the surface responsible for multimeric insulin complex formation.

The representatives of the γ-class, ins-13, and ins-31b modelled well on an insulin template and they have the same cysteine pattern. The ins-31b model has the most similar overall topology to insulin of all 6 scrutinized models, even though it contains an additional two glycine residues between the A-chain helices (See Figures 6.6 and 6.9). The sequence similarity with important residues in insulin, but not identity, is quite good except in the C-terminus of the B-chain. On topology alone it may have a better chance to effectively bind the hIR than Ins-13.

The comparative modelling analysis suggests that the α, β, and γ classes of cysteine patterns do not have distinct preferences to be modelled on a particularly insulin protein, and that the inclusion or absence of disulphide bonds does not impact significantly on its ability to adopt an insulin fold. A recent folding and unfolding experiment using amphioxus insulin-like peptide, identified that an intial folding intermediate always included a formed Cys B19 to Cys A20 disulphide bond and the Cys B7 to Val B18 alpha helix (Chen *et al.*, 2004). This disulphide pairing and the Cys B7 to Cys A7 pairing are the only common pairs between all ins peptides and mammalian homologues. In addition the 'B' chain helix was consistently formed in all models except for Ins-10, and even though there is little sequence identity, the residues in this helix region are similar. This suggests that these two disulphide bonds and the helix are the essential scaffold for the insulin family upon which the other two helices and disulphide pairings form.

6.2.3 RT-PCR of Ins Peptides

Simultaneous to comparative modelling, several of the ins peptides were cloned, so as to eventually perform structural analyses. Since their sequences are so divergent, particularly at the N and C termini, it was obvious that it would be impossible to design a specific pair of oligonucleotides that would amplify multiple *ins* genes, therefore it was necessary to choose *ins* genes that represented each of the different sub-classes. Ins-6 was chosen as a representative of the β class peptides and also because it has a potential signal peptidase cleavage site near the first cysteine and a very short C-terminus after the final cysteine similar

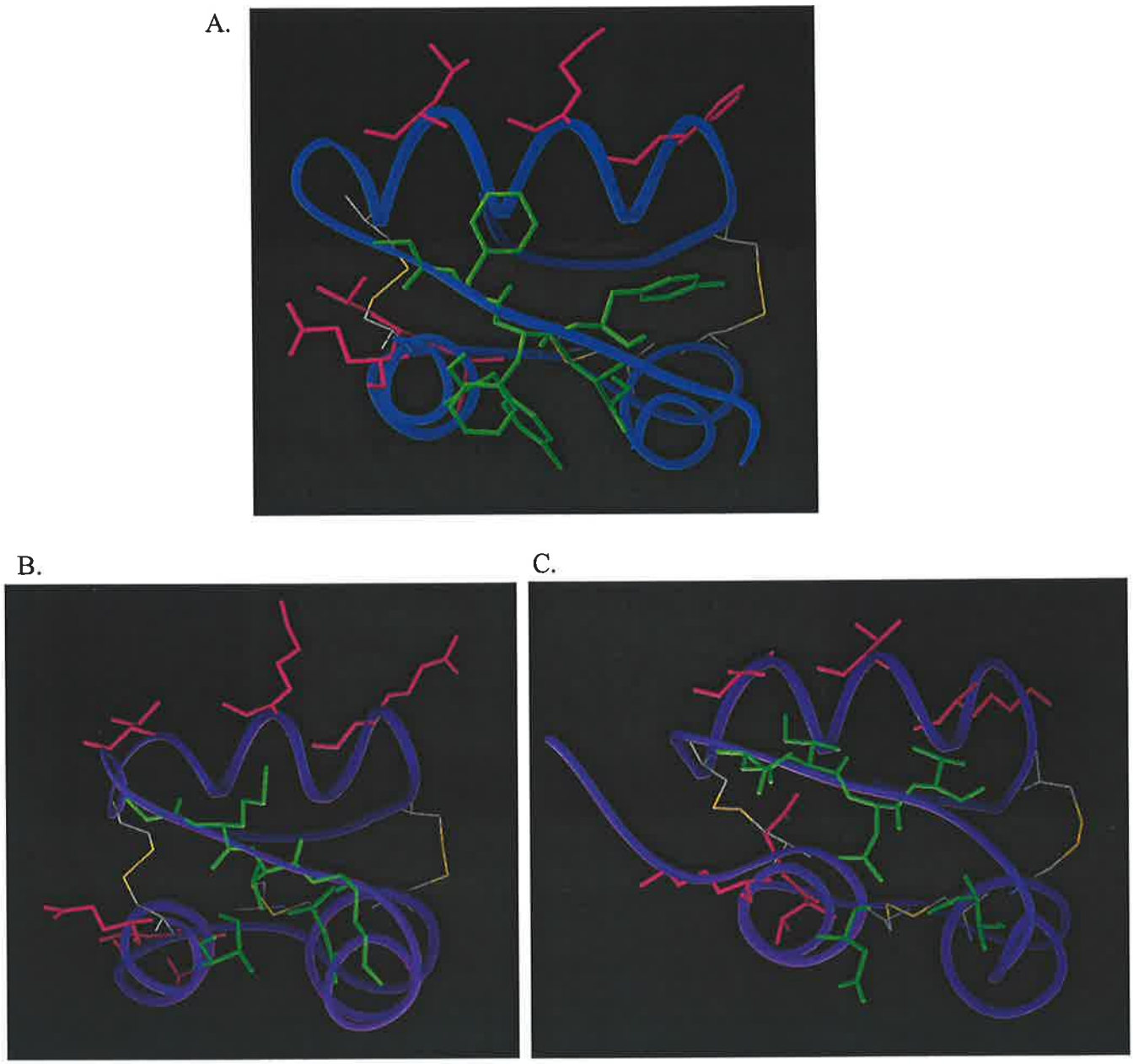


Figure 6.9 Comparison of Models of ins-13 and ins-31b with porcine insulin

Ins models were made in MODELLER, and refined by simulated annealing. Residues highlighted correspond to those found to affect binding to the insulin receptor.
A. pInsulin : 9INS
B. ins-13 : Modelled on 1APH (bInsulin)
C. ins-31b : Modelled on 9INS (pInsulin)

to mammalian insulin. Ins-11 was chosen as the member of the gamma class peptides, containing no C domain, like processed insulin, but also having a C-terminus of similar length to mammalian IGF-1. Ins-23, an α class peptide was chosen because it had two tyrosine's in place of the normal A6/A11 cysteine pair, which is predicted to stack via the aromatic rings of each residue and compensate for the lack of the cysteines (see figure 6.3). Ins-30, another member of the α class, unlike ins-23 has two small non-polar residues in place of A6/A11 cysteine pair and therefore is likely to adopt a slightly different conformation. Ins-17 is another γ class peptide however it has a C domain of similar length to IGF-1 and a C-terminus of intermediate length between insulin and IGF-1. Finally, ins-37 was chosen as the only ins gene to have an extremely long C-terminus similar in length to IGF-2s E domain in its extended form. The cloning strategy is shown in Figure 6.10.

Ins-6 was the first target for cloning. This protein belongs to the β subclass of ins proteins and therefore contains an additional pair of cysteines and has no C-peptide equivalent compared with mammalian insulin. However, similar to insulin it has a short C-terminus after its final cysteine and also contains dual basic residues near to its first cysteine that is predicted to be cleaved to release the mature peptide.

Total RNA obtained from whole adult *C. elegans* was provided by Dr. Warwick Grant (Flinders University). mRNA was isolated from total RNA using the Micro-Fast TrackTM Kit (Invitrogen) (See 2.3.2.6). cDNA was synthesized from the mRNA immediately after purification using MMLV Reverse transcriptase and either Random Hexamer (RH) primers or oligo-dT primers with a 1 hr incubation at 37°C. A control PCR reaction amplifying Actin IV was subsequently performed to assess the quality of the cDNA synthesis and the presence of genomic DNA contamination in the RNA (See Figure 6.11). The results show ~50 bp size difference between using genomic DNA and either oligo-dT or RH derived cDNA, which was expected since Actin IV in *C. elegans* has a 50bp intron in the amplified region.

A primer specifically from the 3′ untranslated sequence from ins-6 was used to 'spike' random hexamers to make more cDNA and a PCR amplification reaction performed using Ampli-Taq Gold polymerase, and primers designed to amplify the coding sequence of ins-6 (see figure 6.12), on the three different cDNA sources. Figure 6.12 shows a 2.5% agarose gel of the PCR reaction. A single band corresponding to the expected siize of ins-6 was detected from each type of cDNA. The spiked RH produced the most of this band size however also showed a significant number of extra bands running lower than the main one. The RH cDNA

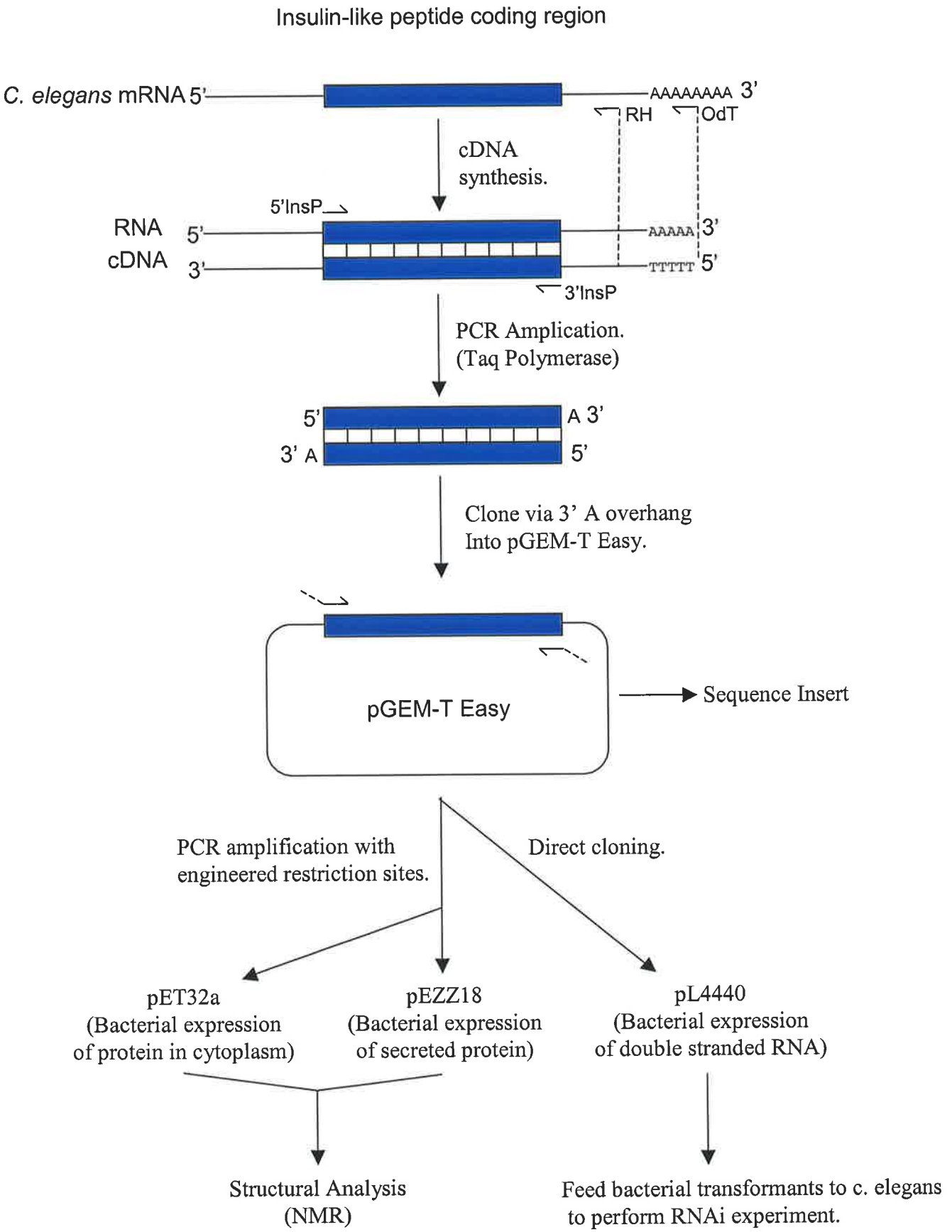


Figure 6.10 **Schematic illustration of the cloning strategy used for isolating the coding region of *C. elegans* genes**

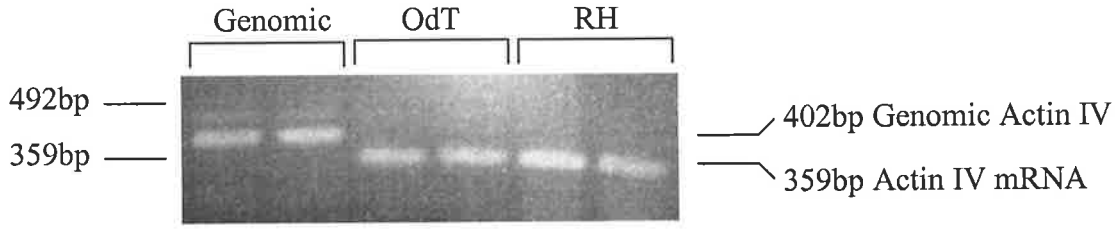


Figure 6.11 Actin IV PCR control of cDNA synthesis

PCR reactions were performed using Actin IV primers and samples of *c. elegans* genomic DNA, cDNA synthesised using oligo dT primers (OdT), or cDNA synthesised using random hexamer primers (RH). PCR proucts were run on a 2% agarose gel and the bands visualised by ethidium bromide staining.

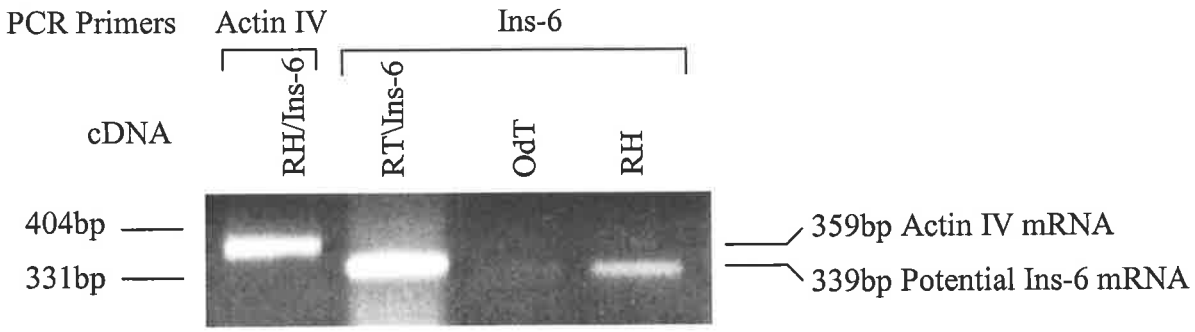


Figure 6.12 PCR amplification of Ins-6 from cDNA

PCR reactions were performed using either Actin IV or Ins-6 primers and samples of cDNA synthesised using oligo dT primers (OdT), and cDNA synthesised using random hexamer primers (RH), or cDNA synthesised using Ins-6 specific primer spiked random hexamers. PCR proucts were run on a 2% agarose gel and the bands visualised by ethidium bromide staining.

produced a clean band as did the oligo-dT cDNA albeit at lower concentration. Since using random hexamers and oligo-dT as the cDNA synthesis primers yielded the cleanest results it was decided that subsequent ins genes to be cloned would use cDNA created from these methods. Therefore an additional RT primer for each gene would not be required.

Primers to amplify ins-11, ins-17, ins-23, ins-30, and ins-37 were designed in an analogous way to ins-6 and used in a PCR reaction on RH cDNA. The results (See Figure 6.13) show that the same band for ins-6 (Lane 5) was amplified in RH cDNA in addition to bands corresponding to the expected sizes for ins-11 (Lane 7) and Ins-17 (Lane 9). It is unclear why the oligo-dT cDNA did not produce any PCR bands although the ins-6 band concentration appeared to be fairly low in the initial reaction compared to the ins-6 specific and RH cDNA reactions. No bands were detected from the PCR reactions for ins-23, 30 or 37. This result could suggest one of several possibilities. *C. elegans* may not express ins-23, 30, or 37 under the optimal growth conditions the worms were grown and harvested in, or at all (ie. pseudo-genes). Alternatively, these proteins may be expressed at particular times during development and therefore not show up in the adult. It is also possible that these genes require a specific environmental signal in order to be expressed. Finally, they could be expressed at such low levels as to be undetectable by this RNA/cDNA amplification method.

6.2.4 Cloning and Sequencing of Ins-6, 11, 17

In order to clone and sequence the bands made from the RT-PCR, products from the ins-6, 11, and 17 reactions were ligated into pGEM-TEasy through the 3′ 'A' tail left behind by the AmpliTaq Gold polymerase. Transformants containing the correct sized insert were identified by EcoRI restriction digests (See Figure 6.14 and 6.15). Even after several attempts the band corresponding to ins-17 was unable to be cloned into pGEMT-Easy. This may have been due to some contaminating DNA visible below the main band. However the ins-17 band looked far more concentrated than the DNA around it. Putative clones of ins-6 and ins-11 were sequenced from pGEMT-Easy using 5′USP primer. Clones were identified containing ins-6 and ins-11, with sequences that exactly matched the predicted coding sequence from the database.

Both ins-6 and ins-11 were sub-cloned into the bacterial vector pL4440 which transcribes RNA from each end, into the multiple cloning site, which results in the formation of double stranded RNA for RNAi experiments. Inserts from EcoR1 digested pGEM-TEasy were ligated into EcoR1 cut pL4440 and the presence of insert determined by EcoR1 digests of

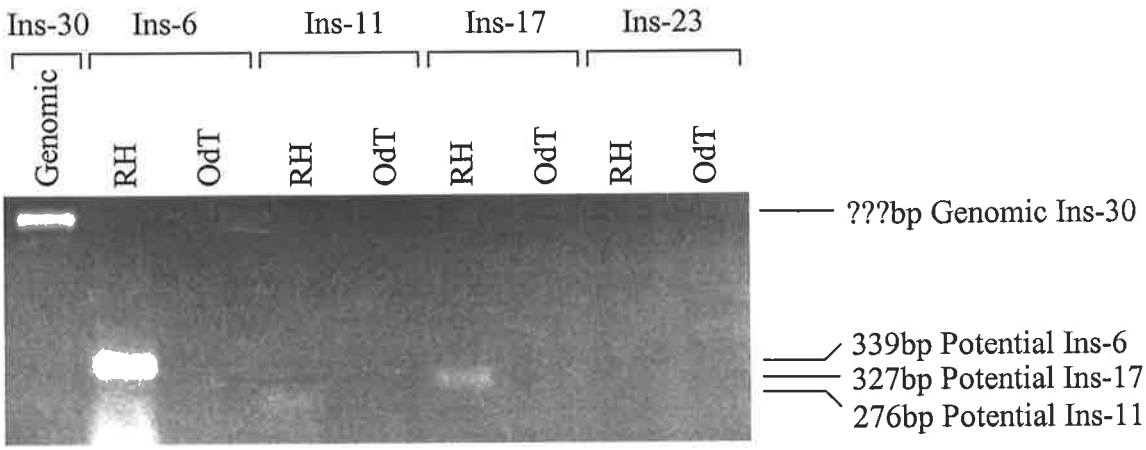


Figure 6.13 PCR amplification of Ins-6, Ins-11 and Ins-17 from cDNA

Taq polymerase PCR reactions were performed using Ins-6, Ins-11, Ins-17, Ins-23 or Ins-30 primers and samples of *c. elegans* genomic DNA, cDNA synthesised using oligo dT primers (OdT), or cDNA synthesised using random hexamer primers (RH). PCR products were run on a 2% agarose gel and the bands visualised by ethidium bromide staining.

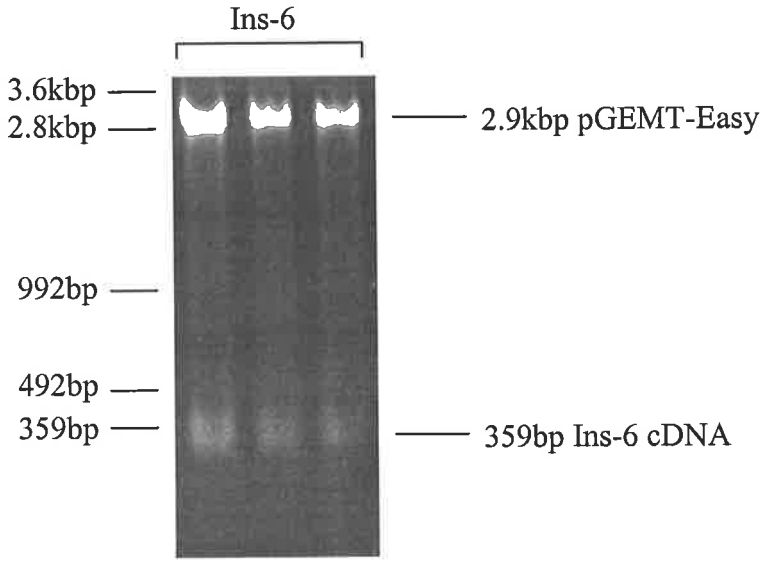


Figure 6.14 Cloning of Ins-6 cDNA into pGEMT-Easy

Ins-6 PCR product was ligated into the pGEMT-Easy via the 5' 'A' overhangs added by the taq polymerase. Transformants containing insert were picked by the use of blue/white colour selection. Plasmid from picked colonies was prepared and digested with EcoRI to release the insert. Restriction digests were run on a 1% agarose gel and the bands visualised by ethidium bromide staining.

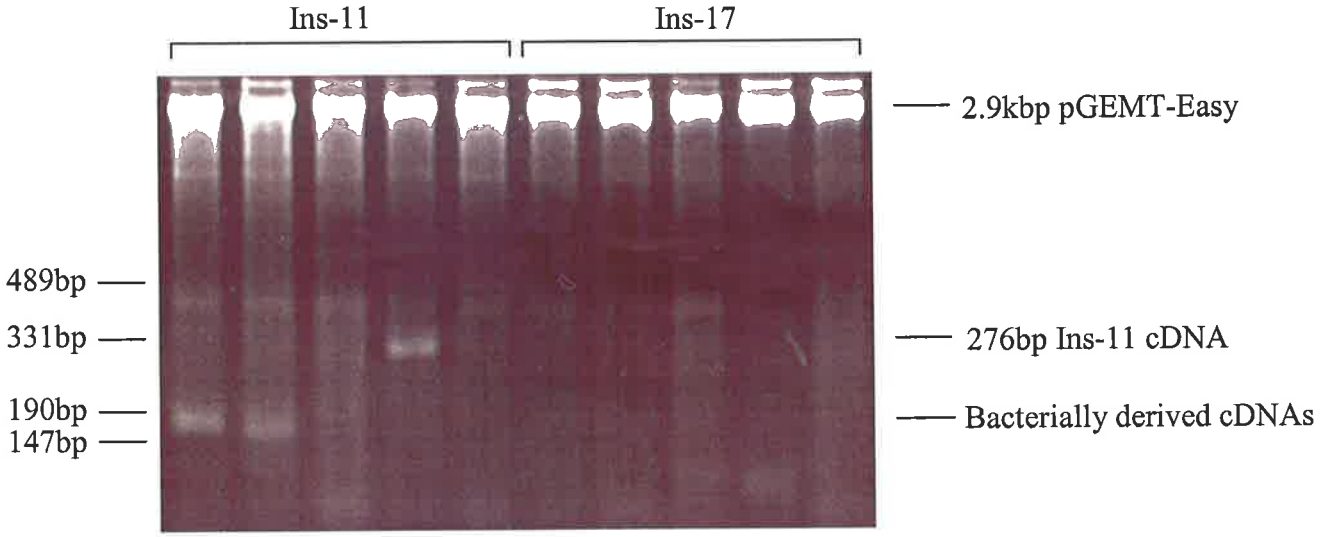


Figure 6.15 Cloning of Ins-11 and Ins-17 cDNA into pGEMT-Easy

Ins-11 and Ins-17 PCR products was ligated into the pGEMT-Easy via the 5' 'A' overhangs added by the taq polymerase. Transformants containing insert were picked by the use of blue/white colour selection. Plasmid from picked colonies was prepared and digested with EcoRI to release the insert. Restriction digests were run on a 1% agarose gel and the bands visualised by ethidium bromide staining.

transformants. RNAi experiments were performed by Warwick Grant on adult worms, however there was no effect on dauer formation, or life expectancy (Data not shown). This may indicate redundancy for ins-6, possibly by the most similar ins peptides ins-4 and ins-5, which also have a similar expression pattern in the adult worm (Pierce *et al.*, 2001). It may also mean that ins-6 is not involved in dauer formation or ageing. In order to further assess ins-6 function, future experiments could focus on different metabolic processes.

A shorter version of the Ins-6 coding region was amplified for introduction into the bacterial expression vectors pEZZ18 and pET32a. PCR products included 5′ BamHI and 3′ EcoRI restriction sites for directional cloning and deleted 58 N-terminal residues up to and including a potential cleavage site in order for the expressed peptide to be similar to processed insulin. The results of the PCR amplification indicated a product of the correct length was made (Data not shown). The PCR products were digested with BamHI and EcoRI and ligated into their respective target vectors.

At this time, a key collaborator (Warwick Grant) moved overseas so it was decided that this project would not be able to be completed through lack of *C. elegans* handling expertise and loss of the daf-2 receptor clone, which was being made in that lab. Other collaborators were looked for in Australia but there were no groups currently doing the types of experiments required using *C. elegans*. Although, expression of ins-6 and possible structure determination might have been achievable, by itself was not deemed to have been as valuable as other aspects of this thesis, particularly in light of the structure being published by another group.

It is possible that ins-6 transcripts are in higher abundance in adult worms than either ins-11 or ins-17, due to the higher band concentration produced from the RT-PCR, however since each primer pair for the PCR was not individually optimized, one cannot be certain. The fact that ins-6, ins-11 and ins-17 are expressed in adult worms growing under optimal conditions may suggest that they are required for normal regulation of growth and metabolism. However, it has been seen that repressing the expression of individual ins genes often had no effect on the phenotype, presumably because of redundancy amongst the peptides, yet some have shown in dauer screens, such as daf-28. The sequences of both ins-6 and ins-11 exactly matched those from the sequenced *c. elegans* genome database, indicating a lack of variation in these genes between strains.

6.3 SUMMARY AND CONCLUSION

C. elegans represents a common ancestor between vertebrate animals and the invertebrate worm. Similar to mammals, it has a fully functional Insulin-like system that appears to control both growth and metabolism, and even uses at least part of the same intracellular circuitry. It was surprising to discover the existence of 38 insulin-like peptides that although containing similar cysteine residue patterning had very little sequence homology to the highly conserved vertebrate insulin and IGFs. Interestingly, the single insulin receptor-like homologue is far more similar to the vertebrate gene counterparts, suggesting that the structural fold induced by the disulphide bonding in the insulin-like peptides is of more importance than the specific residues.

The aims of this project were to use comparative modelling to assess which of the current insulin-like proteins, whose structures have been solved, are most similar to the *C. elegans* peptides. The results of the comparative modelling indicated that there was no obvious, single best template structure, although the known ins-6 structure was similar to those models based on insulin templates, perhaps suggesting that the ins peptides are structurally more closely related to insulin than to either relaxin or bombyxin-II. The overall topology of the models was more globular without flat defined surfaces and of the models ins-31b, and ins-6 were most similar in structure and may be able to interact best with the human receptor homologues. There did not appear to be any conserved charged regions between the models or with insulin or IGF-1. However with the exception of ins-10, every ins peptide sequence was able to be modelled successfully on at least one insulin-like template.

CHAPTER 7

FINAL DISCUSSION

CHAPTER 7 FINAL DISCUSSION

7.1 ROLE OF THE INSERT DOMAIN IN RECEPTOR FUNCTION

The insulin family of receptors are all transmembrane receptor tyrosine kinases with a common modular structure. The most unconserved regions between the receptors lie in the insert domain in the extracellular portion of the receptor and in the C-terminal tail of the intracellular portion of the receptor. The insert domain does not share significant homology with any other known protein domain and between the insulin family of receptors, the only conserved region lies at the end of the alpha subunit directly adjacent to the furin cleavage site.

In this thesis the properties and function of the insert domain were examined in two ways, through protein expression of the isolated domain, and through an investigation of the functional differences between the two isoforms of the insulin receptor, which differ by 12 amino acids within the insert domain.

Two recombinant insert domain proteins were expressed from the pET32a vector, containing either the isolated insert domain or, both the insert domain and the flanking Fn3 domain in which it is inserted. The results of the expression studies indicated that bacteria had difficulty in producing these proteins under all conditions tested, including varying the temperature, cell line used, or co-expressing additional bacterial folding proteins. Additionally, NMR analysis and limited proteolysis suggested that the insert domain protein produced in bacteria is unstructured. Refolding studies were unable to produce a stable insert domain protein, however it appeared that increasing the hydrophobicity, was able to partially stabilise the insert domain structure, as indicated by an increased resistance to proteolysis. Taken together this evidence suggests that the insert domain is only loosely structured and is likely to be buried in the core of the receptor based on its preference for more hydrophobic conditions. It is known that only a small part of the insert domain participates directly in ligand binding and, based on the above expression studies it seems most likely that the insert domain does not form a discrete globular domain, instead forming extended loops that hug the other domains enabling it to position the critical C-terminal tail of the alpha-subunit into the ligand binding pocket and also position the cysteines required for linking the $\alpha\beta$ monomer.

The key to confirming the position of the insert domain within the whole receptor remains solving at least the structure of the extracellular portion of the receptor. It appears unlikely that the insert domain structure will be able to be resolved without the presence of the other domains to act as a scaffold for it. A preliminary assessment of the stability and globularity of the insert domain would be possible by making various truncated forms of the extracellular region of the receptor including the insert domain, and performing limited proteolysis of the recombinant protein, followed by protein sequencing of the remaining stable fragments. This experiment would be able to determine what if any domains are required to stabilise the insert domain structure.

The function of the insert domain was assessed through analysing the differences in ligand binding and receptor phosphorylation of the two isoforms of the insulin receptor, using both recombinant and full-length forms. The affinity of both insulin and the IGFs for binding the isoforms was calculated using a competition assay with europium-labelled insulin. The overall results indicated that increasing the length of the C-terminal tail by the 12 amino acids, found in IR-B, improved insulin binding but decreased IGF-1 and IGF-2 binding. Given that this region is directly involved in ligand binding, the results suggest that the extended C-terminal tail in IR-B sterically interferes with IGF-1 and IGF-2 binding, while slightly improving the insulin fit. This agrees well with the situation in the IGF-1R, which lacks the equivalent additional 12 amino acids found in IR-B. In order to further examine this theory one could make a similarly extended insulin receptor to IR-B using unrelated amino acids and by performing the same competition experiments and seeing if it is only the presence of additional volume that affects ligand binding. If this is not the case then individual mutants within the exon 11 encoded amino acids would have to be analysed.

Within each isoform the amount of tyrosine phosphorylation achieved upon ligand stimulation correlated well with the relative affinities for the receptor. However, it appeared that the IR-A isoform had an overall decreased level of phosphorylation which was particularly surprising since the relative affinity of IGFs was increased. This result points to this region of the insert domain having not only an effect on ligand binding but also on transmitting the signal through to the intracellular kinase domain. This may be a direct effect or indirect through conformational inhibition of one of the other extracellular domains like the cys-rich domain which is known to contribute to switching the receptor from the inactive to active state (Desbois-Mouthon *et al.*, 1996).

7.2 CROSS-REACTIVITY OF IGF-2 WITH THE INSULIN RECEPTOR

Knockout mouse studies have long suggested that IGF-2 required a receptor other than the IGF-1R for promoting embryonic growth. It is now clear that IGF-2 also functions through the Insulin receptor, and probably primarily through the alternate exon 11 minus isoform IR-A. The comparative strength of IGF-2 binding to each isoform has not been as well characterised in addition to whether there is any biological difference between each isoform, activated by either insulin or IGF-2.

In order to investigate these issues, an IGF-1R null cell line, R-, was stably transfected with full-length receptor constructs, expressing IGF-1R, IR-A, or IR-B. In addition, recombinant receptor constructs of IR-A and IR-B containing, either the entire extracellular region of the receptor truncated at the start of the transmembrane domain, or the same truncated receptor with a C-terminal extension consisting of a *S. cereviseae* leucine zipper domain. Competition assays were performed on each receptor type using either europium-labelled insulin or europium-labelled IGF-1 as the tracer. In this assay IGF-2 had an EC_{50} for binding the IGF-1R of ~1nM which dropped 10 and 20 fold when binding to IR-A and IR-B respectively. The value obtained for binding the IGF-1R is similar to that found in the literature using iodine-labelled IGF-1, suggesting that the mechanics of the tracer binding in the competition assay are the same in each case. However, the values obtained when using insulin appear to be lower than what has been published in the literature, indicating a change in the mechanics of the heterogenous europium-labelled insulin binding compared to mono-iodine-labelled insulin. Interestingly the relative strength of binding either IR-A or IR-B between insulin, IGF-1, and IGF-2 is the same between the two competition assays. A way to overcome the difference would be to produce mono-Eu-labelled insulin.

The experiments performed in this thesis, although supporting the idea that IGF-2 is a stronger ligand for hIR-A than hIR-B, do not suggest that it is as biologically potent as insulin, even though it does have a higher affinity for hIR-A. However, the higher affinity appears to be offset by the fact that hIR-A is not phosphorylated to the same extent as hIR-B by any of the growth factors, which had been observed earlier (Kosaki *et al.,* 1995). It would be interesting to determine to what extent the absolute levels of insulin receptor phosphorylation influenced the type of downstream pathways activated. This may explain how the same ligand could illicit alternate outcomes.

The proliferation experiments were consistent with the relative affinities of the ligands and the anti-apoptosis assays suggested that IGF-2 had only a minimal effect. It is however clear that insulin and IGF-2 can trigger hIGF-1R like responses although not as strongly, possibly due to the significantly different ligand affinities between hIGF-1R and hIR. DNA microarray studies performed on hIR-A cells stimulated by either the same concentration of insulin or IGF-2, and therefore different levels of receptor activation, revealed differences in gene expression, however is not clear whether this is due to a ligand specific effect or a due to the different strength of receptor activation (Mulligan *et al.*, 2002). The duration of receptor occupancy and rate of dissociation has also been seen to influence the type of biological response seen from activated receptor and may account for some of the differences observed between insulin and IGF-2 activation of hIR-A. Future microarray experiments using different growth factors should attempt to use ligand concentrations that stimulate equivalent total receptor phosphorylation or initial rates of phosphorylation to address these issues.

7.3 ROLE OF TYROSINE 1250 AND 1251 IN IGF-1R PHOSPHORYLATION

The IGF-IR and the IR have long been known to activate many of the same intracellular signalling pathways, however biologically functioning with quite distinct roles. In *in vitro* experiments it has been difficult to identify specific pathways that are completely unique to one or other receptor under all conditions and cell types including receptor and ligand concentrations. Recent microarray experiments have suggested that at the level of gene regulation there are significant differences both between receptors and also depending on which ligand is used to activate the receptor (Pandini *et al.*, 2003). The intracellular domains of the IGF-IR and IR are very similar and contain many of the same potential phosphorylation and docking sites for substrates. However the IGF-IR contains several altered and additional stretches of sequences that may account for some of the biological differences between it and the IR. The identification of individual residues, that when mutated have specific effects on intracellular signalling pathways, can then be used to examine the role of receptor substrates in terms of both the individual substrate and also the overall contribution to pathways or biological effects.

In this thesis several IGF-IR mutants in the intracellular region of the receptor were compared with wild-type receptor for their ability to undergo IGF-I stimulated tyrosine phosphorylation and trigger cellular proliferation in the IGF-IR null cell line, R-. The results of the analysis indicated that, in particular, the Y1250F and Y1251F single amino acid mutations displayed unexpected phenotypes.

The Y1251F but not the Y1251F/Y950F mutant exhibited hyperphosphorylation in response to IGF-1, which was independent of receptor number. An increased proliferative response to prolonged exposure to IGF-1 was also seen. The reason for the hyperphosphorylation appears to be due to the inability of this receptor to be regulated by one or several tyrosine phosphatases at the initial stage of ligand stimulation. A possible candidate may be PTP-1B as in a recent study, IGF-I induced higher levels of IGF-IR autophosphorylation and kinase activity from embryonic fibroblast cell lines derived from PTP-1B knockout mice, than control lines (Buckley *et al.*, 2002). The phosphorylation response does however retain the same overall shape as the wild-type receptor indicating that the mechanism preventing a continual increase in receptor phosphorylation remains in effect. This may be dependent on receptor internalisation. The Y1250F mutant was deficient in the secondary mechanism as its phosphorylative response did not appear to be approaching a plateau. A previous study had found that the Y1250F mutant exhibited a reduced rate of internalisation (Miura *et al.*, 1997). A comparative study of these mutant receptors examining the activation of downstream signalling pathways, or end point analysis using microarrays would be able to identify the full extent of their effect and provide a better understanding of the mechanisms controlling the ability of the receptor to trigger intracellular signalling.

7.4 THE RELATIONSHIP OF INSULIN-LIKE PEPTIDES IN *C. ELEGANS* WITH MAMMALIAN INSULIN-LIKE PROTEINS

The discovery of a single insulin/IGF receptor (daf-2) in *C. elegans* controlling metabolism and ageing prompted the search for insulin-like peptides. Analysis of the genome revealed the presence of 38 genes containing insulin-like cysteine patterns, however also having very low sequence identity with mammalian insulin proteins. Several of these genes were found to act upstream of daf-2 and were therefore most likely ligands for the receptor. The huge number of genes, many of which appear to be recent gene duplications, coupled with the findings that some of the genes appear to have developmentally and spatially restricted expression patterns, suggests that these peptides may have specialised roles which are not dissimilar to those of insulin, IGF, and relaxins in mammals. In order to see whether the peptides form any structurally related sub-groups, the 3D structure of each peptide was predicted using comparative modelling, based on the existing known structures.

The results of the modelling and refinement revealed that all peptides except ins-10 were able to produce a reasonable model. Generally, the quality of the template had a profound effect on

the quality of the model produced as the hIGF-1 crystal structure 1IMX after the initial modelling had the largest number of best models created from it. The percentage sequence identity generally had little effect on the quality of the model, rather it was the position of the cysteines, and the similarity between the other aligned amino acids that was important. In the case of ins-6, where a picture of the solved structure was known but the coordinates not publically available, the most similar models were derived from the templates with a C-domain of the nearest length. It can therefore be concluded that for target and template sequences that differ greatly in sequence, it is the similarity and not the identity that will allow a good model to be produced. If the ins peptides are all similar to ins-6 in structure then the results of the modelling would predict that they are more similar to insulin than to relaxin or bombyxin.

In the future, given that ins-6 has been shown to bind the insulin receptor (Hua *et al.,* 2003), showing that the residues thought to be previously unreplacable can be changed in the right combination, a functional analysis of the *C. elegans* peptides using site-directed mutagenesis may reveal novel ways to design insulin-receptor family ligands. When expressed in *C. elegans* insulin behaves as an inhibitor of daf-2 signalling. This suggests that the mechanism for binding the insulin receptor has been conserved amongst the ligands but the mode of activation of the receptor has changed. This is intriguing since unlike the ligands there has been greater sequence conservation between the daf-2 receptor and mammalian receptors. An analysis of the daf-2 receptor to determine how its activation differs from mammalian receptors may identify regions that could be targeted by antibodies or other drugs to inhibit or enhance human insulin/IGF receptors.

This thesis examined several aspects of the insulin and IGF systems including receptor-ligand interactions, receptor autophosphorylation, and the conservation of structure of insulin-like peptides from the nematode worm. It may be that only through the combined analysis of the insulin/IGF systems throughout the animal kingdom that we will be able to understand and control one of the most important growth and metabolism systems in biology.

CHAPTER 8 References

Abbott, A.M., Bueno, R., Pedrini, M.T., Murray, J.M. and Smith, R.J. (1992) Insulin-like growth factor I receptor gene structure. *J Biol Chem*, **267**, 10759-63.

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. *Cell Mol Life Sci*, **57**, 1050-93.

Alessi, D.R., Andjelkovic, M., Caudwell, B., Cron, P., Morrice, N., Cohen, P. and Hemmings, B.A. (1996) Mechanism of activation of protein kinase B by insulin and IGF-1. *Embo J*, **15**, 6541-51.

Almind, K., Inoue, G., Pedersen, O. and Kahn, C.R. (1996) A common amino acid polymorphism in insulin receptor substrate-1 causes impaired insulin signaling. Evidence from transfection studies. *J Clin Invest*, **97**, 2569-75.

Andersen, A.S., Kjeldsen, T., Wiberg, F.C., Christensen, P.M., Rasmussen, J.S., Norris, K., Moller, K.B. and Moller, N.P. (1990) Changing the insulin receptor to possess insulin-like growth factor I ligand specificity. *Biochemistry*, **29**, 7363-6.

Andersen, A.S., Kjeldsen, T., Wiberg, F.C., Vissing, H., Schaffer, L., Rasmussen, J.S., De Meyts, P. and Moller, N.P. (1992) Identification of determinants that confer ligand specificity on the insulin receptor. *J Biol Chem*, **267**, 13681-6.

Araki, E., Lipes, M.A., Patti, M.E., Bruning, J.C., Haag, B., 3rd, Johnson, R.S. and Kahn, C.R. (1994) Alternative pathway of insulin signalling in mice with targeted disruption of the IRS-1 gene. *Nature*, **372**, 186-90.

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. *Biochem J*, **327 (Pt 1)**, 209-15.

Bajaj, M., Waterfield, M.D., Schlessinger, J., Taylor, W.R. and Blundell, T. (1987) On the tertiary structure of the extracellular domains of the epidermal growth factor and insulin receptors. *Biochim Biophys Acta*, **916**, 220-6.

Baker, E.N., Blundell, T.L., Cutfield, J.F., Cutfield, S.M., Dodson, E.J., Dodson, G.G., Hodgkin, D.M., Hubbard, R.E., Isaacs, N.W., Reynolds, C.D. and et al. (1988) The structure of 2Zn pig insulin crystals at 1.5 A resolution. *Philos Trans R Soc Lond B Biol Sci*, **319**, 369-456.

Barlow, D.P., Stoger, R., Herrmann, B.G., Saito, K. and Schweifer, N. (1991) The mouse insulin-like growth factor type-2 receptor is imprinted and closely linked to the Tme locus. *Nature*, **349**, 84-7.

Bass, J., Chiu, G., Argon, Y. and Steiner, D.F. (1998) Folding of insulin receptor monomers is facilitated by the molecular chaperones calnexin and calreticulin and impaired by rapid dimerization. *J Cell Biol*, **141**, 637-46.

Bass, J., Kurose, T., Pashmforoush, M. and Steiner, D.F. (1996) Fusion of insulin receptor ectodomains to immunoglobulin constant domains reproduces high-affinity insulin binding in vitro. *J Biol Chem*, **271**, 19367-75.

Bayne, M.L., Applebaum, J., Chicchi, G.G., Hayes, N.S., Green, B.G. and Cascieri, M.A. (1988) Structural analogs of human insulin-like growth factor I with reduced affinity for serum binding proteins and the type 2 insulin-like growth factor receptor. *J Biol Chem*, **263**, 6233-9.

Belfiore, A., Pandini, G., Vella, V., Squatrito, S. and Vigneri, R. (1999) Insulin/IGF-I hybrid receptors play a major role in IGF-I signaling in thyroid cancer. *Biochimie*, **81**, 403-7.

Bevan, P. (2001) Insulin signalling. *J Cell Sci*, **114**, 1429-30.

Blahovec, J., Kostecka, Z., Lacroix, M.C., Cabanie, L., Godeau, F., Mester, J. and Cavaille, F. (2001) Mitogenic activity of high molecular weight forms of insulin-like growth factor-II in amniotic fluid. *J*

Endocrinol, **169**, 563-72.

Blakesley, V.A., Koval, A.P., Stannard, B.S., Scrimgeour, A. and LeRoith, D. (1998) Replacement of tyrosine 1251 in the carboxyl terminus of the insulin-like growth factor-I receptor disrupts the actin cytoskeleton and inhibits proliferation and anchorage-independent growth. *J Biol Chem*, **273**, 18411-22.

Bluher, M., Michael, M.D., Peroni, O.D., Ueki, K., Carter, N., Kahn, B.B. and Kahn, C.R. (2002) Adipose tissue selective insulin receptor knockout protects against obesity and obesity-related glucose intolerance. *Dev Cell*, **3**, 25-38.

Blum, G., Gazit, A. and Levitzki, A. (2000) Substrate competitive inhibitors of IGF-1 receptor kinase. *Biochemistry*, **39**, 15705-12.

Bravo, D.A., Gleason, J.B., Sanchez, R.I., Roth, R.A. and Fuller, R.S. (1994) Accurate and efficient cleavage of the human insulin proreceptor by the human proprotein-processing protease furin. Characterization and kinetic parameters using the purified, secreted soluble protease expressed by a recombinant baculovirus. *J Biol Chem*, **269**, 25830-7.

Bruning, J.C., Gautam, D., Burks, D.J., Gillette, J., Schubert, M., Orban, P.C., Klein, R., Krone, W., Muller-Wieland, D. and Kahn, C.R. (2000) Role of brain insulin receptor in control of body weight and reproduction. *Science*, **289**, 2122-5.

Bruning, J.C., Michael, M.D., Winnay, J.N., Hayashi, T., Horsch, D., Accili, D., Goodyear, L.J. and Kahn, C.R. (1998) A muscle-specific insulin receptor knockout exhibits features of the metabolic syndrome of NIDDM without altering glucose tolerance. *Mol Cell*, **2**, 559-69.

Bruning, J.C., Winnay, J., Bonner-Weir, S., Taylor, S.I., Accili, D. and Kahn, C.R. (1997) Development of a novel polygenic model of NIDDM in mice heterozygous for IR and IRS-1 null alleles. *Cell*, **88**, 561-72.

Buckley, D.A., Cheng, A., Kiely, P.A., Tremblay, M.L., O'Connor, R. (2002) Regulation of insulin-like growth factor type I (IGF-I) receptor kinase activity by protein tyrosine phosphatase 1B (PTP-1B) and enhanced IGF-I-mediated suppression of apoptosis and motility in PTP-1B-deficient fibroblasts. *Mol. Cell. Biol.*, 22(7), 1998-2010.

Burgaud, J.L. and Baserga, R. (1996) Intracellular transactivation of the insulin-like growth factor I receptor by an epidermal growth factor receptor. *Exp Cell Res*, **223**, 412-9.

Butler, A.A. and LeRoith, D. (2001) Minireview: tissue-specific versus generalized gene targeting of the igf1 and igf1r genes and their roles in insulin-like growth factor physiology. *Endocrinology*, **142**, 1685-8.

Cara, J.F., Mirmira, R.G., Nakagawa, S.H. and Tager, H.S. (1990) An insulin-like growth factor I/insulin hybrid exhibiting high potency for interaction with the type I insulin-like growth factor and insulin receptors of placental plasma membranes. *J Biol Chem*, **265**, 17820-5.

Carrick, F.E., Forbes, B.E. and Wallace, J.C. (2001) BIAcore analysis of bovine insulin-like growth factor (IGF)-binding protein-2 identifies major IGF binding site determinants in both the amino- and carboxyl-terminal domains. *J Biol Chem*, **276**, 27120-8. Epub 2001 May 16.

Cascieri, M.A., Chicchi, G.G., Applebaum, J., Hayes, N.S., Green, B.G. and Bayne, M.L. (1988) Mutants of human insulin-like growth factor I with reduced affinity for the type 1 insulin-like growth factor receptor. *Biochemistry*, **27**, 3229-33.

Cheatham, B., Vlahos, C.J., Cheatham, L., Wang, L., Blenis, J. and Kahn, C.R. (1994) Phosphatidylinositol 3-kinase activation is required for insulin stimulation of pp70 S6 kinase, DNA synthesis, and glucose transporter translocation. *Mol Cell Biol*, **14**, 4902-11.

Chen, y., Jin, R., Dong, H., and Feng, Y. (2004) In Vitro Refolding/Unfolding Pathways of Amphioxus Insulin-like Peptide. *J. Biol. Chem.* **279** (53), 55224-55233.

Chernicky, C.L., Tan, H., Yi, L., Loret de Mola, J.R. and Ilan, J. (2002) Treatment of murine breast cancer cells with antisense RNA to the type I insulin-like growth factor receptor decreases the level of plasminogen activator transcripts, inhibits cell growth in vitro, and reduces tumorigenesis in vivo. *Mol Pathol*, **55**, 102-9.

Chernicky, C.L., Yi, L., Tan, H., Gan, S.U. and Ilan, J. (2000) Treatment of human breast cancer cells with antisense RNA to the type I insulin-like growth factor receptor inhibits cell growth, suppresses tumorigenesis, alters the metastatic potential, and prolongs survival in vivo. *Cancer Gene Ther*, **7**, 384-95.

Chomczynski, P. and Sacchi, N. (1987) Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction. *Anal Biochem*, **162**, 156-9.

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**, 472-5.

Clairmont, K.B. and Czech, M.P. (1989) Chicken and Xenopus mannose 6-phosphate receptors fail to bind insulin-like growth factor II. *J Biol Chem*, **264**, 16390-2.

Condorelli, G., Bueno, R. and Smith, R.J. (1994) Two alternatively spliced forms of the human insulin-like growth factor I receptor have distinct biological activities and internalization kinetics. *J Biol Chem*, **269**, 8510-6.

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**, 5484-91.

Cosgrove, L., Lovrecz, G.O., Verkuylen, A., Cavaleri, L., Black, L.A., Bentley, J.D., Howlett, G.J., Gray, P.P., Ward, C.W. and McKern, N.M. (1995) Purification and properties of insulin receptor ectodomain from large-scale mammalian cell culture. *Protein Expr Purif*, **6**, 789-98.

Craparo, A., Freund, R. and Gustafson, T.A. (1997) 14-3-3 (epsilon) interacts with the insulin-like growth factor I receptor and insulin receptor substrate I in a phosphoserine-dependent manner. *J Biol Chem*, **272**, 11663-9.

Craparo, A., O'Neill, T.J. and Gustafson, T.A. (1995) Non-SH2 domains within insulin receptor substrate-1 and SHC mediate their phosphotyrosine-dependent interaction with the NPEY motif of the insulin-like growth factor I receptor. *J Biol Chem*, **270**, 15639-43.

Datta, S.R., Dudek, H., Tao, X., Masters, S., Fu, H., Gotoh, Y. and Greenberg, M.E. (1997) Akt phosphorylation of BAD couples survival signals to the cell-intrinsic death machinery. *Cell*, **91**, 231-41.

Davidson, H.W., Peshavaria, M. and Hutton, J.C. (1987) Proteolytic conversion of proinsulin into insulin. Identification of a Ca2+-dependent acidic endopeptidase in isolated insulin-secretory granules. *Biochem J*, **246**, 279-86.

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**, S135-48.

De Meyts, P., Gu, J.L., Shymko, R.M., Kaplan, B.E., Bell, G.I. and Whittaker, J. (1990) Identification of a ligand-binding region of the human insulin receptor encoded by the second exon of the gene. *Mol Endocrinol*, **4**, 409-16.

De Meyts, P. and Whittaker, J. (2002) Structural biology of insulin and IGF1 receptors: implications for drug design. *Nat Rev Drug Discov*, **1**, 769-83.

Denley A, Bonython ER, Booker GW, Cosgrove LJ, Forbes BE, Ward CW, Wallace JC. (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), 2502-12.

Derewenda U, Derewenda Z, Dodson EJ, Dodson GG, Bing X, Markussen J. (1991) X-ray analysis of the single chain B29-A1 peptide-linked insulin molecule. A completely inactive analogue. *J Mol Biol.* **220**(2):425-33.

Desbois-Mouthon, C., Sert-Langeron, C., Magre, J., Oreal, E., Blivet, M.J., Flori, E., Besmond, C., Capeau, J. and Caron, M. (1996) Deletion of Asn281 in the alpha-subunit of the human insulin receptor causes constitutive activation of the receptor and insulin desensitization. *J Clin Endocrinol Metab*, **81**, 719-27.

Diamant, M. and Heine, R.J. (2003) Thiazolidinediones in type 2 diabetes mellitus: current clinical evidence. *Drugs*, **63**, 1373-405.

Du, X. and Tang, J.G. (1998) Hydroxyl group of insulin A19Tyr is essential for receptor binding: studies on (A19Phe)insulin. *Biochem Mol Biol Int*, **45**, 255-60.

Duret, L., Guex, N., Peitsch, M.C. and Bairoch, A. (1998) New insulin-like proteins with atypical disulfide bond pattern characterized in Caenorhabditis elegans by comparative sequence analysis and homology modeling. *Genome Res*, **8**, 348-53.

Ebina, Y., Ellis, L., Jarnagin, K., Edery, M., Graf, L., Clauser, E., Ou, J.H., Masiarz, F., Kan, Y.W., Goldfine, I.D. and et al. (1985) The human insulin receptor cDNA: the structural basis for hormone-activated transmembrane signalling. *Cell*, **40**, 747-58.

Fabry, M., Schaefer, E., Ellis, L., Kojro, E., Fahrenholz, F. and Brandenburg, D. (1992) Detection of a new hormone contact site within the insulin receptor ectodomain by the use of a novel photoreactive insulin. *J Biol Chem*, **267**, 8950-6.

Favelyukis, S., Till, J.H., Hubbard, S.R. and Miller, W.T. (2001) Structure and autoregulation of the insulin-like growth factor 1 receptor kinase. *Nat Struct Biol*, **8**, 1058-63.

Fernandez, A.M., Kim, J.K., Yakar, S., Dupont, J., Hernandez-Sanchez, C., Castle, A.L., Filmore, J., Shulman, G.I. and Le Roith, D. (2001) Functional inactivation of the IGF-I and insulin receptors in skeletal muscle causes type 2 diabetes. *Genes Dev*, **15**, 1926-34.

Filson, A.J., Louvi, A., Efstratiadis, A. and Robertson, E.J. (1993) Rescue of the T-associated maternal effect in mice carrying null mutations in Igf-2 and Igf2r, two reciprocally imprinted genes. *Development*, **118**, 731-6.

Forbes BE, Hartfield PJ, McNeil KA, Surinya KH, Milner SJ, Cosgrove LJ, Wallace JC. (2002) Characteristics of binding of insulin-like growth factor (IGF)-I and IGF-II analogues to the type 1 IGF receptor determined by BIAcore analysis. *Eur J Biochem*. **269**(3):961-8.

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. *Mol Cell Biol*, **19**, 3278-88.

Fruman, D.A., Mauvais-Jarvis, F., Pollard, D.A., Yballe, C.M., Brazil, D., Bronson, R.T., Kahn, C.R. and Cantley, L.C. (2000) Hypoglycaemia, liver necrosis and perinatal death in mice lacking all isoforms of phosphoinositide 3-kinase p85 alpha. *Nat Genet*, **26**, 379-82.

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**, 395-9.

Germain-Lee EL, Janicot M, Lammers R, Ullrich A, Casella SJ. (1992) Expression of a type I insulin-like growth factor receptor with low affinity for insulin-like growth factor II. *Biochem J.* **281** (Pt 2):413-7.

Goodstadt, L. and Ponting, C.P. (2001) CHROMA: consensus-based colouring of multiple alignments for publication. *Bioinformatics*, **17**, 845-6.

Graf, R., Neuenschwander, S., Brown, M.R. and Ackermann, U. (1997) Insulin-mediated secretion of ecdysteroids from mosquito ovaries and molecular cloning of the insulin receptor homologue from ovaries of bloodfed Aedes aegypti. *Insect Mol Biol*, **6**, 151-63.

Gregoire, F.M., Chomiki, N., Kachinskas, D. and Warden, C.H. (1998) Cloning and developmental regulation of a novel member of the insulin-like gene family in Caenorhabditis elegans. *Biochem Biophys Res Commun*, **249**, 385-90.

Guerra, C., Navarro, P., Valverde, A.M., Arribas, M., Bruning, J., Kozak, L.P., Kahn, C.R. and Benito, M. (2001) Brown adipose tissue-specific insulin receptor knockout shows diabetic phenotype without insulin resistance. *J Clin Invest*, **108**, 1205-13.

Gustafson, T.A., He, W., Craparo, A., Schaub, C.D. and O'Neill, T.J. (1995) Phosphotyrosine-dependent interaction of SHC and insulin receptor substrate 1 with the NPEY motif of the insulin receptor via a novel non-SH2 domain. *Mol Cell Biol*, **15**, 2500-8.

Gustafson, T.A. and Rutter, W.J. (1990) The cysteine-rich domains of the insulin and insulin-like growth factor I receptors are primary determinants of hormone binding specificity. Evidence from receptor chimeras. *J Biol Chem*, **265**, 18663-7.

Hadsell, D.L., Murphy, K.L., Bonnette, S.G., Reece, N., Laucirica, R. and Rosen, J.M. (2000) Cooperative interaction between mutant p53 and des(1-3)IGF-I accelerates mammary tumorigenesis. *Oncogene*, **19**, 889-98.

Hailey, J., Maxwell, E., Koukouras, K., Bishop, W.R., Pachter, J.A. and Wang, Y. (2002) Neutralizing anti-insulin-like growth factor receptor 1 antibodies inhibit receptor function and induce receptor degradation in tumor cells. *Mol Cancer Ther*, **1**, 1349-53.

Hankinson, S.E., Willett, W.C., Colditz, G.A., Hunter, D.J., Michaud, D.S., Deroo, B., Rosner, B., Speizer, F.E. and Pollak, M. (1998) Circulating concentrations of insulin-like growth factor-I and risk of breast cancer. *Lancet*, **351**, 1393-6.

Haring, H.U., Kellerer, M. and Mosthaf, L. (1994) Modulation of insulin receptor signalling: significance of altered receptor isoform patterns and mechanism of hyperglycaemia-induced receptor modulation. *Diabetologia*, **37 Suppl 2**, S149-54.

Haring, H.U., Kellerer, M. and Mosthaf, L. (1994) Modulation of insulin signalling in non-insulin-dependent diabetes mellitus: significance of altered receptor isoform patterns and mechanisms of glucose-induced receptor modulation. *Horm Res*, **41 Suppl 2**, 87-91; discussion 92.

Haselbacher, G. and Humbel, R. (1982) Evidence for two species of insulin-like growth factor II (IGF II and "big" IGF II) in human spinal fluid. *Endocrinology*, **110**, 1822-4.

He, W., Craparo, A., Zhu, Y., O'Neill, T.J., Wang, L.M., Pierce, J.H. and Gustafson, T.A. (1996) Interaction of insulin receptor substrate-2 (IRS-2) with the insulin and insulin-like growth factor I receptors. Evidence for two distinct phosphotyrosine-dependent interaction domains within IRS-2. *J Biol Chem*, **271**, 11641-5.

He, W., O'Neill, T.J. and Gustafson, T.A. (1995) Distinct modes of interaction of SHC and insulin receptor substrate-1 with the insulin receptor NPEY region via non-SH2 domains. *J Biol Chem*, **270**, 23258-62.

Hermanto, U., Zong, C., Li, W., and Wang, L. (2002) RACK1, an Insulin-Like Growth Factor 1 (IGF-1) receptor-interacting protein, modulates IGF-1-dependent integrin signaling and promotes cell spreading and contact with extracellular matrix. *Mol. Cell. Biol.* **22**(7), 2345-2365.

Hobbs, S., Jitrapakdee, S. and Wallace, J.C. (1998) Development of a bicistronic vector driven by the human polypeptide chain elongation factor 1alpha promoter for creation of stable mammalian cell lines that express very high levels of recombinant proteins. *Biochem Biophys Res Commun*, **252**, 368-72.

Holzenberger, M., Dupont, J., Ducos, B., Leneuve, P., Geloen, A., Even, P.C., Cervera, P. and Le Bouc, Y. (2003) IGF-1 receptor regulates lifespan and resistance to oxidative stress in mice. *Nature*, **421**, 182-7.

Hoyne, P.A., Cosgrove, L.J., McKern, N.M., Bentley, J.D., Ivancic, N., Elleman, T.C. and Ward, C.W. (2000) High affinity insulin binding by soluble insulin receptor extracellular domain fused to a leucine zipper. *FEBS Lett*, **479**, 15-8.

Hua, Q.X., Nakagawa, S.H., Wilken, J., Ramos, R.R., Jia, W., Bass, J. and Weiss, M.A. (2003) A divergent INS protein in Caenorhabditis elegans structurally resembles human insulin and activates the human insulin receptor. *Genes Dev*, **17**, 826-31.

Hua, Q.X., Shoelson, S.E., Kochoyan, M. and Weiss, M.A. (1991) Receptor binding redefined by a structural switch in a mutant human insulin. *Nature*, **354**, 238-41.

Huang, K., Xu, B., Hu, S.Q., Chu, Y.C., Hua, Q.X., Qu, Y., Li, B., Wang, S., Wang, R.Y., Nakagawa, S.H., Theede, A.M., Whittaker, J., De Meyts, P., Katsoyannis, P.G. and Weiss, M.A. (2004) How insulin binds: the B-chain alpha-helix contacts the L1 beta-helix of the insulin receptor. *J Mol Biol*, **341**, 529-50.

Huang, Z., Bodkin, N.L., Ortmeyer, H.K., Hansen, B.C. and Shuldiner, A.R. (1994) Hyperinsulinemia is associated with altered insulin receptor mRNA splicing in muscle of the spontaneously obese diabetic rhesus monkey. *J Clin Invest*, **94**, 1289-96.

Huang, Z., Bodkin, N.L., Ortmeyer, H.K., Zenilman, M.E., Webster, N.J., Hansen, B.C. and Shuldiner, A.R. (1996) Altered insulin receptor messenger ribonucleic acid splicing in liver is associated with deterioration of glucose tolerance in the spontaneously obese and diabetic rhesus monkey: analysis of controversy between monkey and human studies. *J Clin Endocrinol Metab*, **81**, 1552-6.

Hubbard, S.R. (1997) Crystal structure of the activated insulin receptor tyrosine kinase in complex with peptide substrate and ATP analog. *Embo J*, **16**, 5572-81.

Hubbard, S.R., Wei, L., Ellis, L. and Hendrickson, W.A. (1994) Crystal structure of the tyrosine kinase domain of the human insulin receptor. *Nature*, **372**, 746-54.

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**, 8153-60.

Humphrey, W., Dalke, A., and Schulten, K. (1996) VMD - Visual Molecular Dynamics. *J. Mol. Graph.* **14,** 33-38, 27-28.

Hundal, R.S. and Inzucchi, S.E. (2003) Metformin: new understandings, new uses. *Drugs*, **63**, 1879-94.
Jackson, J.G., White, M.F. and Yee, D. (1998) Insulin receptor substrate-1 is the predominant signaling molecule activated by insulin-like growth factor-I, insulin, and interleukin-4 in estrogen receptor-positive human breast cancer cells. *J Biol Chem*, **273**, 9994-10003.

Jacobs, S., Shechter, Y., Bissell, K. and Cuatrecasas, P. (1977) Purification and properties of insulin receptors from rat liver membranes. *Biochem Biophys Res Commun*, **77**, 981-8.

Jansson, M., Hallen, D., Koho, H., Andersson, G., Berghard, L., Heidrich, J., Nyberg, E., Uhlen, M., Kordel, J. and
Nilsson, B. (1997) Characterization of ligand binding of a soluble human insulin-like growth factor I receptor variant suggests a ligand-induced conformational change. *J Biol Chem*, **272**, 8189-97.

Jiang, T., Sweeney, G., Rudolf, M.T., Klip, A., Traynor-Kaplan, A. and Tsien, R.Y. (1998) Membrane-permeant esters of phosphatidylinositol 3, 4, 5-trisphosphate. *J Biol Chem*, **273**, 11017-24.

Jiang, Y., Chan, J.L., Zong, C.S. and Wang, L.H. (1996) Effect of tyrosine mutations on the kinase activity and transforming potential of an oncogenic human insulin-like growth factor I receptor. *J Biol Chem*, **271**, 160-7.

Jones, J.I. and Clemmons, D.R. (1995) Insulin-like growth factors and their binding proteins: biological actions. *Endocr Rev*, **16**, 3-34.

Kale, L., Skeel, R., Bhandarkar, M., Brunner, R., Gursoy, A., Krawetz, N., Phillips, J., Shinozaki, A., Varadarajan, K., and Schulten, K. (1999) NAMD2: greater scalability for parallel molecular dynamics. *J. Comput. Phys.* **151,** 283-312.

Kahn, S.E. (2003) The relative contributions of insulin resistance and beta-cell dysfunction to the pathophysiology of Type 2 diabetes. *Diabetologia*, **46**, 3-19

Kalli, K.R., Falowo, O.I., Bale, L.K., Zschunke, M.A., Roche, P.C. and Conover, C.A. (2002) Functional insulin receptors on human epithelial ovarian carcinoma cells: implications for IGF-II mitogenic signaling. *Endocrinology*, **143**, 3259-67.

Kanai, F., Ito, K., Todaka, M., Hayashi, H., Kamohara, S., Ishii, K., Okada, T., Hazeki, O., Ui, M. and Ebina, Y. (1993) Insulin-stimulated GLUT4 translocation is relevant to the phosphorylation of IRS-1 and the activity of PI3-kinase. *Biochem Biophys Res Commun*, **195**, 762-8.

Kasuya, J., Paz, I.B., Maddux, B.A., Goldfine, I.D., Hefta, S.A. and Fujita-Yamaguchi, Y. (1993) Characterization of human placental insulin-like growth factor-I/insulin hybrid receptors by protein microsequencing and purification. *Biochemistry*, **32**, 13531-6.

Kataoka, K., Yu, D. and Miura, M. (2004) Activation of the phosphatidylinositol-3' kinase pathway and DNA synthesis by a mutant insulin-like growth factor I receptor lacking the NPXY motif. *J Endocrinol*, **181**, 139-46.

Kawano, T., Ito, Y., Ishiguro, M., Takuwa, K., Nakajima, T. and Kimura, Y. (2000) Molecular cloning and characterization of a new insulin/IGF-like peptide of the nematode Caenorhabditis elegans. *Biochem*

Biophys Res Commun, **273**, 431-6.

Kellerer, M., Sesti, G., Seffer, E., Obermaier-Kusser, B., Pongratz, D.E., Mosthaf, L. and Haring, H.U. (1993) Altered pattern of insulin receptor isotypes in skeletal muscle membranes of type 2 (non-insulin-dependent) diabetic subjects. *Diabetologia*, **36**, 628-32.

Khwaja, A., Rodriguez-Viciana, P., Wennstrom, S., Warne, P.H. and Downward, J. (1997) Matrix adhesion and Ras transformation both activate a phosphoinositide 3-OH kinase and protein kinase B/Akt cellular survival pathway. *Embo J*, **16**, 2783-93.

Kido, Y., Burks, D.J., Withers, D., Bruning, J.C., Kahn, C.R., White, M.F. and Accili, D. (2000) Tissue-specific insulin resistance in mice with mutations in the insulin receptor, IRS-1, and IRS-2. *J Clin Invest*, **105**, 199-205.

Kiely, P.A., Sant, A., O'Connor, R. (2002) RACK1 is an Insulin-like Growth Factor 1 (IGF-1) Receptor-interacting protein that can regulate IGF-1-mediated Akt activation and protection from cell death. *J. Biol. Chem.* **277**(25), 22581-22589.

Kim, J.K., Michael, M.D., Previs, S.F., Peroni, O.D., Mauvais-Jarvis, F., Neschen, S., Kahn, B.B., Kahn, C.R. and Shulman, G.I. (2000) Redistribution of substrates to adipose tissue promotes obesity in mice with selective insulin resistance in muscle. *J Clin Invest*, **105**, 1791-7.

Kimura, K.D., Tissenbaum, H.A., Liu, Y. and Ruvkun, G. (1997) daf-2, an insulin receptor-like gene that regulates longevity and diapause in Caenorhabditis elegans. *Science*, **277**, 942-6.

Kjeldsen, T., Andersen, A.S., Wiberg, F.C., Rasmussen, J.S., Schaffer, L., Balschmidt, P., Moller, K.B. and Moller, N.P. (1991) The ligand specificities of the insulin receptor and the insulin-like growth factor I receptor reside in different regions of a common binding site. *Proc Natl Acad Sci U S A*, **88**, 4404-8.

Kjeldsen, T., Wiberg, F.C. and Andersen, A.S. (1994) Chimeric receptors indicate that phenylalanine 39 is a major contributor to insulin specificity of the insulin receptor. *J Biol Chem*, **269**, 32942-6.

Kosaki, A., Nelson, J. and Webster, N.J. (1998) Identification of intron and exon sequences involved in alternative splicing of insulin receptor pre-mRNA. *J Biol Chem*, **273**, 10331-7.

Kristensen, C., Andersen, A.S., Ostergaard, S., Hansen, P.H., Brandt, J., Surinya, K.H., Molina, L., Soos, M.A. and Siddle, K. (2002) Functional reconstitution of insulin receptor binding site from non-binding receptor fragments
Role of insulin receptor dimerization domains in ligand binding, cooperativity, and modulation by anti-receptor antibodies. *J Biol Chem*, **277**, 18340-5.

Kristensen, C., Kjeldsen, T., Wiberg, F.C., Schaffer, L., Hach, M., Havelund, S., Bass, J., Steiner, D.F. and Andersen, A.S. (1997) Alanine scanning mutagenesis of insulin. *J Biol Chem*, **272**, 12978-83.

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. *J Biol Chem*, **274**, 37351-6.

Kulik, G., Klippel, A. and Weber, M.J. (1997) Antiapoptotic signalling by the insulin-like growth factor I receptor, phosphatidylinositol 3-kinase, and Akt. *Mol Cell Biol*, **17**, 1595-606.

Kulkarni, R.N., Bruning, J.C., Winnay, J.N., Postic, C., Magnuson, M.A. and Kahn, C.R. (1999) Tissue-specific knockout of the insulin receptor in pancreatic beta cells creates an insulin secretory defect similar to that in type 2 diabetes. *Cell*, **96**, 329-39.

Kulkarni, R.N., Holzenberger, M., Shih, D.Q., Ozcan, U., Stoffel, M., Magnuson, M.A. and Kahn, C.R. (2002) beta-cell-specific deletion of the Igf1 receptor leads to hyperinsulinemia and glucose intolerance but does not alter beta-cell mass. *Nat Genet*, **31**, 111-5.

Kurose, T., Pashmforoush, M., Yoshimasa, Y., Carroll, R., Schwartz, G.P., Burke, G.T., Katsoyannis, P.G. and Steiner, D.F. (1994) Cross-linking of a B25 azidophenylalanine insulin derivative to the carboxyl-terminal region of the alpha-subunit of the insulin receptor. Identification of a new insulin-binding domain in the insulin receptor. *J Biol Chem*, **269**, 29190-7.

Lawrence, J.C., Jr. and Roach, P.J. (1997) New insights into the role and mechanism of glycogen synthase activation by insulin. *Diabetes*, **46**, 541-7.

Leahy, M., Lyons, A., Krause, D. and O'Connor, R. (2004) Impaired Shc, Ras, and MAPK activation, but normal Akt activation in FL5.12 cells expressing an IGF-IR mutated at tyrosines 1250 and 1251. *J Biol Chem*, **12**, 12.

Li, S., Ferber, A., Miura, M. and Baserga, R. (1994) Mitogenicity and transforming activity of the insulin-like growth factor-I receptor with mutations in the tyrosine kinase domain. *J Biol Chem*, **269**, 32558-64.

Li, S., Resnicoff, M. and Baserga, R. (1996) Effect of mutations at serines 1280-1283 on the mitogenic and transforming activities of the insulin-like growth factor I receptor. *J Biol Chem*, **271**, 12254-60.

Li, S., Termini, J., Hayward, A., Siddle, K., Zick, Y., Koval, A., LeRoith, D. and Fujita-Yamaguchi, Y. (1998) The carboxyl-terminal domain of insulin-like growth factor-I receptor interacts with the insulin receptor and activates its protein tyrosine kinase. *FEBS Lett*, **421**, 45-9.

Li, W., Kennedy, S.G. and Ruvkun, G. (2003) daf-28 encodes a C. elegans insulin superfamily member that is regulated by environmental cues and acts in the DAF-2 signaling pathway. *Genes Dev*, **17**, 844-58.

Lin, K., Dorman, J.B., Rodan, A. and Kenyon, C. (1997) daf-16: An HNF-3/forkhead family member that can function to double the life-span of Caenorhabditis elegans. *Science*, **278**, 1319-22.

Linge, J.P., Williams, M.A. (2003) Refinement of protein structures in explicit solvents. *Proteins*, **50**(3), 496-506.

Lobel, P., Dahms, N.M. and Kornfeld, S. (1988) Cloning and sequence analysis of the cation-independent mannose 6-phosphate receptor. *J Biol Chem*, **263**, 2563-70.

Ludvigsen, S., Olsen, H.B. and Kaarsholm, N.C. (1998) A structural switch in a mutant insulin exposes key residues for receptor binding. *J Mol Biol*, **279**, 1-7.

Luo, R.Z., Beniac, D.R., Fernandes, A., Yip, C.C. and Ottensmeyer, F.P. (1999) Quaternary structure of the insulin-insulin receptor complex. *Science*, **285**, 1077-80.

MacKerell, A. D., Bashford, D., Bellott, M., Dunbrack, R. L., Evanseck, J. D., Field, M. J., Fischer, S., Gao, J., Guo, H., Ha, S., Joseph-McCarthy, D., Kuchnir, L., Kuczera, K., Lau, F. T. K., Mattos, C., Michnick, S., Ngo, T., Nguyen, D. T., Prodhom, B., Reiher, W. E., Roux, B., Schlenkrich, M., Smith, J. C., Stote, R., Straub, J., Watanabe, M., Wiórkiewicz-Kuczera, J., Yin, D., and Karplus, M. (1998) All-atom empirical potential for molecular modeling and dynamics Studies of proteins. *J. Phys. Chem. B* **102 (18),** 3586-3616.

Maloney, E.K., McLaughlin, J.L., Dagdigian, N.E., Garrett, L.M., Connors, K.M., Zhou, X.M., Blattler, W.A., Chittenden, T. and Singh, R. (2003) An anti-insulin-like growth factor I receptor antibody that is a potent inhibitor of cancer cell proliferation. *Cancer Res*, **63**, 5073-83.

Marino-Buslje, C., Mizuguchi, K., Siddle, K. and Blundell, T.L. (1998) A third fibronectin type III domain in the extracellular region of the insulin receptor family. *FEBS Lett*, **441**, 331-6.

Marti-Renom, M.A., Stuart, A.C., Fiser, A., Sanchez, R., Melo, F. and Sali, A. (2000) Comparative protein structure modeling of genes and genomes. *Annu Rev Biophys Biomol Struct*, **29**, 291-325.

Massague, J. and Czech, M.P. (1982) The subunit structures of two distinct receptors for insulin-like growth factors I and II and their relationship to the insulin receptor. *J Biol Chem*, **257**, 5038-45.

Michael, M.D., Kulkarni, R.N., Postic, C., Previs, S.F., Shulman, G.I., Magnuson, M.A. and Kahn, C.R. (2000) Loss of insulin signaling in hepatocytes leads to severe insulin resistance and progressive hepatic dysfunction. *Mol Cell*, **6**, 87-97.

Mirmira, R.G., Nakagawa, S.H. and Tager, H.S. (1991) Importance of the character and configuration of residues B24, B25, and B26 in insulin-receptor interactions. *J Biol Chem*, **266**, 1428-36.

Miura, M. and Baserga, R. (1997) The tyrosine residue at 1250 of the insulin-like growth factor I receptor is required for ligand-mediated internalization. *Biochem Biophys Res Commun*, **239**, 182-5.

Miura, M., Li, S. and Baserga, R. (1995) Effect of a mutation at tyrosine 950 of the insulin-like growth factor I receptor on the growth and transformation of cells. *Cancer Res*, **55**, 663-7.

Miura, M., Surmacz, E., Burgaud, J.L. and Baserga, R. (1995) Different effects on mitogenesis and transformation of a mutation at tyrosine 1251 of the insulin-like growth factor I receptor. *J Biol Chem*, **270**, 22639-44.

Mora, S., Monden, I., Zorzano, A. and Keller, K. (1997) Heterologous expression of rab4 reduces glucose transport and GLUT4 abundance at the cell surface in oocytes. *Biochem J*, **324**, 455-9.

Morris, J.Z., Tissenbaum, H.A. and Ruvkun, G. (1996) A phosphatidylinositol-3-OH kinase family member regulating longevity and diapause in Caenorhabditis elegans. *Nature*, **382**, 536-9.

Mosthaf, L., Eriksson, J., Haring, H.U., Groop, L., Widen, E. and Ullrich, A. (1993) Insulin receptor isotype expression correlates with risk of non-insulin-dependent diabetes. *Proc Natl Acad Sci U S A*, **90**, 2633-5.

Mosthaf, L., Grako, K., Dull, T.J., Coussens, L., Ullrich, A., McClain, D.A. (1990) Functionally distinct insulin receptors generated by tissue-specific alternative splicing. *EMBO J.* **9**(8), 2409-13.

Mosthaf, L., Vogt, B., Haring, H.U. and Ullrich, A. (1991) Altered expression of insulin receptor types A and B in the skeletal muscle of non-insulin-dependent diabetes mellitus patients. *Proc Natl Acad Sci U S A*, **88**, 4728-30.

Mulhern, T.D., Booker, G.W. and Cosgrove, L. (1998) A third fibronectin-type-III domain in the insulin-family receptors. *Trends Biochem Sci*, **23**, 465-6.

Mulligan C, Rochford J, Denyer G, Stephens R, Yeo G, Freeman T, Siddle K, O'Rahilly S. (2002) Microarray analysis of insulin and insulin-like growth factor-1 (IGF-1) receptor signaling reveals the selective up-regulation of the mitogen heparin-binding EGF-like growth factor by IGF-1. *J Biol Chem.* **277**(45):42480-7.

Myers, G., Selznick, S. (1996) Progressive multiple alignment with constraints. *J Comput Biol*, **3**(4), 563-572.

Myers, M.G., Backer, J.M., Siddle, K. and White, M.F. (1991) The insulin receptor functions normally in Chinese hamster ovary cells after truncation of the C terminus. *J Biol Chem*, **266**, 10616-23.

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**, 18650-5.

Mynarcik, D.C., Yu, G.Q. and Whittaker, J. (1996) Alanine-scanning mutagenesis of a C-terminal ligand binding domain of the insulin receptor alpha subunit. *J Biol Chem*, **271**, 2439-42.

Nakagawa, S.H. and Tager, H.S. (1991) Implications of invariant residue LeuB6 in insulin-receptor interactions. *J Biol Chem*, **266**, 11502-9.

Nakagawa, S.H. and Tager, H.S. (1992) Importance of aliphatic side-chain structure at positions 2 and 3 of the insulin A chain in insulin-receptor interactions. *Biochemistry*, **31**, 3204-14.

Nakamura, K., Hongo, A., Kodama, J., Miyagi, Y., Yoshinouchi, M. and Kudo, T. (2000) Down-regulation of the insulin-like growth factor I receptor by antisense RNA can reverse the transformed phenotype of human cervical cancer cell lines. *Cancer Res*, **60**, 760-5.

Neuenschwander, S., Schwartz, A., Wood, T.L., Roberts, C.T., Jr., Hennighausen, L. and LeRoith, D. (1996) Involution of the lactating mammary gland is inhibited by the IGF system in a transgenic mouse model. *J Clin Invest*, **97**, 2225-32.

Nielsen, F.C. (1992) The molecular and cellular biology of insulin-like growth factor II. *Prog Growth Factor Res*, **4**, 257-90.

Notredame, C., Higgins, D.G. and Heringa, J. (2000) T-Coffee: A novel method for fast and accurate multiple sequence alignment. *J Mol Biol*, **302**, 205-17.

O'Connor, R., Kauffmann-Zeh, A., Liu, Y., Lehar, S., Evan, G.I., Baserga, R. and Blattler, W.A. (1997) Identification of domains of the insulin-like growth factor I receptor that are required for protection from apoptosis. *Mol Cell Biol*, **17**, 427-35.

Ogg, S., Paradis, S., Gottlieb, S., Patterson, G.I., Lee, L., Tissenbaum, H.A. and Ruvkun, G. (1997) The Fork head transcription factor DAF-16 transduces insulin-like metabolic and longevity signals in C. elegans. *Nature*, **389**, 994-9.

O'Gorman DB, Weiss J, Hettiaratchi A, Firth SM, Scott CD (2002) Insulin-like growth factor-II/mannose 6-phosphate receptor overexpression reduces growth of choriocarcinoma cells in vitro and in vivo. *Endocrinology*. **143**(11):4287-94.

Oka, Y. and Czech, M.P. (1986) The type II insulin-like growth factor receptor is internalized and recycles in the absence of ligand. *J Biol Chem*, **261**, 9090-3.

Okada, Y., Yokono, K., Katsuta, A., Yoshida, M., Morita, S., Irino, H., Goto, T., Baba, S., Roth, R.A., Shii, K. (1998) Development of an assay for bioactive insulin. *Anal Biochem*. **257**(2), 134-8.

Okamoto, T., Nishimoto, I., Murayama, Y., Ohkuni, Y. and Ogata, E. (1990) Insulin-like growth factor-II/mannose 6-phosphate receptor is incapable of activating GTP-binding proteins in response to mannose 6-phosphate, but capable in response to insulin-like growth factor-II. *Biochem Biophys Res Commun*, **168**, 1201-10.

Olson, T.S., Bamberger, M.J. and Lane, M.D. (1988) Post-translational changes in tertiary and quaternary structure of the insulin proreceptor. Correlation with acquisition of function. *J Biol Chem*, **263**, 7342-51.

O'Neill, T.J., Craparo, A. and Gustafson, T.A. (1994) Characterization of an interaction between insulin receptor substrate 1 and the insulin receptor by using the two-hybrid system. *Mol Cell Biol*, **14**, 6433-42.

Osipo C, Dorman S, Frankfater A. (2001) Loss of insulin-like growth factor II receptor expression promotes growth in cancer by increasing intracellular signaling from both IGF-I and insulin receptors. *Exp Cell Res*. **264**(2):388-96.

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**, 39684-95.

Pandini, G., Medico, E., Conte, E., Sciacca, L., Vigneri, R. and Belfiore, A. (2003) Differential gene expression induced by insulin and insulin-like growth factor-II through the insulin receptor isoform A. *J Biol Chem*, **278**, 42178-89. Epub 2003 Jul 24.

Pandini, G., Vigneri, R., Costantino, A., Frasca, F., Ippolito, A., Fujita-Yamaguchi, Y., Siddle, K., Goldfine, I.D. and Belfiore, A. (1999) Insulin and insulin-like growth factor-I (IGF-I) receptor overexpression in breast cancers leads to insulin/IGF-I hybrid receptor overexpression: evidence for a second mechanism of IGF-I signaling. *Clin Cancer Res*, **5**, 1935-44.

Paradis, S., Ailion, M., Toker, A., Thomas, J.H. and Ruvkun, G. (1999) A PDK1 homolog is necessary and sufficient to transduce AGE-1 PI3 kinase signals that regulate diapause in Caenorhabditis elegans. *Genes Dev*, **13**, 1438-52.

Paradis, S. and Ruvkun, G. (1998) Caenorhabditis elegans Akt/PKB transduces insulin receptor-like signals from AGE-1 PI3 kinase to the DAF-16 transcription factor. *Genes Dev*, **12**, 2488-98.

Parrizas, M., Gazit, A., Levitzki, A., Wertheimer, E. and LeRoith, D. (1997) Specific inhibition of insulin-like growth factor-1 and insulin receptor tyrosine kinase activity and biological function by tyrphostins. *Endocrinology*, **138**, 1427-33.

Pashmforoush, M., Chan, S.J. and Steiner, D.F. (1996) Structure and expression of the insulin-like peptide receptor from amphioxus. *Mol Endocrinol*, **10**, 857-66.

Pautsch, A., Zoephel, A., Ahorn, H., Spevak, W., Hauptmann, R. and Nar, H. (2001) Crystal structure of bisphosphorylated IGF-1 receptor kinase: insight into domain movements upon kinase activation. *Structure (Camb)*, **9**, 955-65.

Paz, K., Boura-Halfon, S., Wyatt, L.S., LeRoith, D. and Zick, Y. (2000) The juxtamembrane but not the carboxyl-terminal domain of the insulin receptor mediates insulin's metabolic functions in primary adipocytes and cultured hepatoma cells. *J Mol Endocrinol*, **24**, 419-32.

Pierce, S.B., Costa, M., Wisotzkey, R., Devadhar, S., Homburger, S.A., Buchman, A.R., Ferguson, K.C., Heller, J., Platt, D.M., Pasquinelli, A.A., Liu, L.X., Doberstein, S.K. and Ruvkun, G. (2001) Regulation of DAF-2 receptor signaling by human insulin and ins-1, a member of the unusually large and diverse C. elegans insulin gene family.
Genes Dev, **15**, 672-86.

Pullen, R.A., Lindsay, D.G., Wood, S.P., Tickle, I.J., Blundell, T.L., Wollmer, A., Krail, G., Brandenburg, D., Zahn, H., Gliemann, J. and Gammeltoft, S. (1976) Receptor-binding region of insulin. *Nature*, **259**, 369-73.

Rafaeloff, R., Patel, R., Yip, C., Goldfine, I.D. and Hawley, D.M. (1989) Mutation of the high cysteine region of the human insulin receptor alpha-subunit increases insulin receptor binding affinity and transmembrane signaling. *J Biol Chem*, **264**, 15900-4.

Reiss, K., Valentinis, B., Tu, X., Xu, S.Q. and Baserga, R. (1998) Molecular markers of IGF-I-mediated mitogenesis. *Exp Cell Res*, **242**, 361-72.

Reiss, K., Yumet, G., Shan, S., Huang, Z., Alnemri, E., Srinivasula, S.M., Wang, J.Y., Morrione, A. and Baserga, R. (1999) Synthetic peptide sequence from the C-terminus of the insulin-like growth factor-I receptor that induces apoptosis and inhibition of tumor growth. *J Cell Physiol*, **181**, 124-35.

Resnicoff, M., Tjuvajev, J., Rotman, H.L., Abraham, D., Curtis, M., Aiken, R. and Baserga, R. (1996) Regression of C6 rat brain tumors by cells expressing an antisense insulin-like growth factor I receptor RNA. *J Exp Ther Oncol*, **1**, 385-9.

Ruan, W. and Kleinberg, D.L. (1999) Insulin-like growth factor I is essential for terminal end bud formation and ductal morphogenesis during mammary development. *Endocrinology*, **140**, 5075-81.

Rubini, M., Hongo, A., D'Ambrosio, C. and Baserga, R. (1997) The IGF-I receptor in mitogenesis and transformation of mouse embryo cells: role of receptor number. *Exp Cell Res*, **230**, 284-92.

Russell, R.B., Barton, G.J. (1992) Multiple protein sequence alignment from tertiary structure comparison. *Proteins: Struct. Funct. Genet.* **14**, 309-323.

Sanchez, R., Sali, A. (1998) Large-scale protein structure modelling of the Saccharomyces cerevisiae genome. *Proc Natl. Acad. Sci.*, **95**, 13597-13602.

Sasaoka, T., Ishiki, M., Wada, T., Hori, H., Hirai, H., Haruta, T., Ishihara, H. and Kobayashi, M. (2001) Tyrosine phosphorylation-dependent and -independent role of Shc in the regulation of IGF-1-induced mitogenesis and glycogen synthesis. *Endocrinology*, **142**, 5226-35.

Sato, A., Nishimura, S., Ohkubo, T., Kyogoku, Y., Koyama, S., Kobayashi, M., Yasuda, T. and Kobayashi, Y. (1992) 1H-NMR assignment and secondary structure of human insulin-like growth factor-I (IGF-I) in solution. *J Biochem (Tokyo)*, **111**, 529-36.

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**, 433-40.

Savkur, R.S., Philips, A.V. and Cooper, T.A. (2001) Aberrant regulation of insulin receptor alternative splicing is associated with insulin resistance in myotonic dystrophy. *Nat Genet*, **29**, 40-7.

Sbraccia, P., D'Adamo, M., Leonetti, F., Caiola, S., Iozzo, P., Giaccari, A., Buongiorno, A. and Tamburrano, G. (1996) Chronic primary hyperinsulinaemia is associated with altered insulin receptor mRNA splicing in muscle of patients with insulinoma. *Diabetologia*, **39**, 220-5.

Sbraccia, P., Giaccari, A., D'Adamo, M., Caiola, S., Morviducci, L., Zorretta, D., Maroccia, E., Buongiorno, A. and Tamburrano, G. (1998) Expression of the two insulin receptor isoforms is not altered in the skeletal muscle and liver of diabetic rats. *Metabolism*, **47**, 129-32.

Schaefer, E.M., Erickson, H.P., Federwisch, M., Wollmer, A. and Ellis, L. (1992) Structural organization of the human insulin receptor ectodomain. *J Biol Chem*, **267**, 23393-402.

Schaffer, L. and Ljungqvist, L. (1992) Identification of a disulfide bridge connecting the alpha-subunits of the extracellular domain of the insulin receptor. *Biochem Biophys Res Commun*, **189**, 650-3.

Schmidt, B., Kiecke-Siemsen, C., Waheed, A., Braulke, 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**, 14975-82.

Schumacher, R., Mosthaf, L., Schlessinger, J., Brandenburg, D. and Ullrich, A. (1991) Insulin and insulin-like growth factor-1 binding specificity is determined by distinct regions of their cognate receptors. *J Biol Chem*, **266**, 19288-95.

Schumacher, R., Soos, M.A., Schlessinger, J., Brandenburg, D., Siddle, K. and Ullrich, A. (1993) Signaling-competent receptor chimeras allow mapping of major insulin receptor binding domain determinants. *J Biol Chem*, **268**, 1087-94.

Sciacca, L., Costantino, A., Pandini, G., Mineo, R., Frasca, F., Scalia, P., Sbraccia, P., Goldfine, I.D., Vigneri, R. and Belfiore, A. (1999) Insulin receptor activation by IGF-II in breast cancers: evidence for a new autocrine/paracrine mechanism. *Oncogene*, **18**, 2471-9.

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**, 8240-50.

Scotlandi, K., Maini, C., Manara, M.C., Benini, S., Serra, M., Cerisano, V., Strammiello, R., Baldini, N., Lollini, P.L., Nanni, P., Nicoletti, G. and Picci, P. (2002) Effectiveness of insulin-like growth factor I receptor antisense strategy against Ewing's sarcoma cells. *Cancer Gene Ther*, **9**, 296-307.

Seino, S. and Bell, G.I. (1989) Alternative splicing of human insulin receptor messenger RNA. *Biochem Biophys Res Commun*, **159**, 312-6.

Seino, S., Seino, M., Nishi, S. and Bell, G.I. (1989) Structure of the human insulin receptor gene and characterization of its promoter. *Proc Natl Acad Sci U S A*, **86**, 114-8.

Sell, C., Dumenil, G., Deveaud, C., Miura, M., Coppola, D., DeAngelis, T., Rubin, R., Efstratiadis, A., Baserga, R., Rubini, M. and Liu, J.P. (1994) Effect of a null mutation of the insulin-like growth factor I receptor gene on growth and transformation of mouse embryo fibroblasts Simian virus 40 large tumor antigen is unable to transform mouse embryonic fibroblasts lacking type 1 insulin-like growth factor receptor. *Mol Cell Biol*, **14**, 3604-12.

Sesti, G., D'Alfonso, R., Vargas Punti, M.D., Frittitta, L., Trischitta, V., Liu, Y.Y., Borboni, P., Longhi, R., Montemurro, A. and Lauro, R. (1995) Peptide-based radioimmunoassay for the two isoforms of the human insulin receptor. *Diabetologia*, **38**, 445-53.

Sesti, G., Marini, M.A., Tullio, A.N., Montemurro, A., Borboni, P., Fusco, A., Accili, D. and Lauro, R. (1991) Altered expression of the two naturally occurring human insulin receptor variants in isolated adipocytes of non-insulin-dependent diabetes mellitus patients. *Biochem Biophys Res Commun*, **181**, 1419-24.

Shepherd, P.R., Withers, D.J. and Siddle, K. (1998) Phosphoinositide 3-kinase: the key switch mechanism in insulin signalling. *Biochem J*, **333**, 471-90.

Slade, A., Luh, J., Ho, S. and Yip, C.M. (2002) Single molecule imaging of supported planar lipid bilayer--reconstituted human insulin receptors by in situ scanning probe microscopy. *J Struct Biol*, **137**, 283-91.

Songyang, Z., Carraway, K.L., 3rd, Eck, M.J., Harrison, S.C., Feldman, R.A., Mohammadi, M., Schlessinger, J., Hubbard, S.R., Smith, D.P., Eng, C. and et al. (1995) Catalytic specificity of protein-tyrosine kinases is critical for selective signalling. *Nature*, **373**, 536-9.

Songyang, Z., Shoelson, S.E., Chaudhuri, M., Gish, G., Pawson, T., Haser, W.G., King, F., Roberts, T., Ratnofsky, S., Lechleider, R.J. and et al. (1993) SH2 domains recognize specific phosphopeptide sequences. *Cell*, **72**, 767-78.

Sorensen, H., Whittaker, L., Hinrichsen, J., Groth, A., Whittaker, J. and Gadsboll, V.L. (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**, 19-22.

Sparrow, L.G., McKern, N.M., Gorman, J.J., Strike, P.M., Robinson, C.P., Bentley, J.D. and Ward, C.W. (1997) The disulfide bonds in the C-terminal domains of the human insulin receptor ectodomain. *J Biol Chem*, **272**, 29460-7.

Standaert, M.L., Galloway, L., Karnam, P., Bandyopadhyay, G., Moscat, J. and Farese, R.V. (1997) Protein kinase C-zeta as a downstream effector of phosphatidylinositol 3-kinase during insulin stimulation in rat adipocytes. Potential role in glucose transport. *J Biol Chem*, **272**, 30075-82.

Su, B. and Karin, M. (1996) Mitogen-activated protein kinase cascades and regulation of gene expression. *Curr Opin Immunol*, **8**, 402-11.

Sun, H., Tu, X., Prisco, M., Wu, A., Casiburi, I. and Baserga, R. (2003) Insulin-like growth factor I receptor signaling and nuclear translocation of insulin receptor substrates 1 and 2. *Mol Endocrinol*, **17**, 472-86.

Surinya, K.H., Molina, L., Soos, M.A., Brandt, J., Kristensen, C. and Siddle, K. (2002) Role of insulin receptor dimerization domains in ligand binding, cooperativity, and modulation by anti-receptor antibodies. *J Biol Chem*, **277**, 16718-25.

Tamemoto, H., Kadowaki, T., Tobe, K., Yagi, T., Sakura, H., Hayakawa, T., Terauchi, Y., Ueki, K., Kaburagi, Y., Satoh, S. and et al. (1994) Insulin resistance and growth retardation in mice lacking insulin receptor substrate-1. *Nature*, **372**, 182-6.

Tan, C.M., Kelvin, D.J., Litchfield, D.W., Ferguson, S.S. and Feldman, R.D. (2001) Tyrosine kinase-mediated serine phosphorylation of adenylyl cyclase. *Biochemistry*, **40**, 1702-9.

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**, 385-401.

Toyoshima, Y., Karas, M., Yakar, S., Dupont, J., Helman, L. and LeRoith, D. (2004) TDAG51 mediates the effects of insulin-like growth factor I (IGF-I) on cell survival. *J Biol Chem*, **22**, 22.

Trombetta, E.S. and Helenius, A. (1998) Lectins as chaperones in glycoprotein folding. *Curr Opin Struct Biol*, **8**, 587-92.

Tulloch, P.A., Lawrence, L.J., McKern, N.M., Robinson, C.P., Bentley, J.D., Cosgrove, L., Ivancic, N., Lovrecz, G.O., Siddle, K. and Ward, C.W. (1999) Single-molecule imaging of human insulin receptor ectodomain and its Fab complexes. *J Struct Biol*, **125**, 11-8.

Ullrich, A., Bell, J.R., Chen, E.Y., Herrera, R., Petruzzelli, L.M., Dull, T.J., Gray, A., Coussens, L., Liao, Y.C., Tsubokawa, M. and et al. (1985) Human insulin receptor and its relationship to the tyrosine kinase family of oncogenes. *Nature*, **313**, 756-61.

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**, 2503-12.

Valverde, A.M., Lorenzo, M., Teruel, T. and Benito, M. (1995) cAMP inhibits IGF-I-induced mitogenesis in fetal rat brown adipocytes: role of p21 ras. *Exp Cell Res*, **218**, 305-9.

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**, 245-54.

Vidal, H., Auboeuf, D., Beylot, M. and Riou, J.P. (1995) Regulation of insulin receptor mRNA splicing in rat tissues. Effect of fasting, aging, and diabetes. *Diabetes*, **44**, 1196-201.

Virkamaki, A., Ueki, K. and Kahn, C.R. (1999) Protein-protein interaction in insulin signaling and the molecular mechanisms of insulin resistance. *J Clin Invest*, **103**, 931-43.

Ward, C.W. (1999) Members of the insulin receptor family contain three fibronectin type III domains. *Growth Factors*, **16**, 315-22.

Wedekind, F., Baer-Pontzen, K., Bala-Mohan, S., Choli, D., Zahn, H. and Brandenburg, D. (1989) Hormone binding site of the insulin receptor: analysis using photoaffinity-mediated avidin complexing. *Biol Chem Hoppe Seyler*, **370**, 251-8.

Werner, H., Adamo, M., Roberts, C.T., Jr. and LeRoith, D. (1994) Molecular and cellular aspects of insulin-like growth factor action. *Vitam Horm*, **48**, 1-58.

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**, 43980-6.

Whittaker, J., Sorensen, H., Gadsboll, V.L., Hinrichsen, J. (2002) Comparison of the functional insulin binding epitopes of the A and B isoforms of the insulin receptor. *JBC* **277**(49), 47380-4

Williams, P.F., Mynarcik, D.C., Yu, G.Q. and Whittaker, J. (1995) Mapping of an NH2-terminal ligand binding site of the insulin receptor by alanine scanning mutagenesis. *J Biol Chem*, **270**, 3012-6.

Withers, D.J., Gutierrez, J.S., Towery, H., Burks, D.J., Ren, J.M., Previs, S., Zhang, Y., Bernal, D., Pons, S., Shulman, G.I., Bonner-Weir, S. and White, M.F. (1998) Disruption of IRS-2 causes type 2 diabetes in mice. *Nature*, **391**, 900-4.

Wu, C.J., O'Rourke, D.M., Feng, G.S., Johnson, G.R., Wang, Q. and Greene, M.I. (2001) The tyrosine phosphatase SHP-2 is required for mediating phosphatidylinositol 3-kinase/Akt activation by growth factors. *Oncogene*, **20**, 6018-25.

Yam, A., Hyun, T. and Li, W. (2001) Characterization of insulin-like growth factor I (IGF-I) receptor mutants for their effects on IGF-I- and interleukin 4-mediated DNA synthesis of 32D cells. *J Biol Chem*, **276**, 24409-13.

Yamaguchi, Y., Flier, J.S., Yokota, A., Benecke, H., Backer, J.M., Moller, D.E. (1991) Functional properties of two naturally occurring isoforms of the human insulin receptor in Chinese hamster ovary cells. *Endocrinology*. **129**(4), 2058-66.

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**, 1132-8.

Yee, D., Lebovic, G.S., Marcus, R.R. and Rosen, N. (1989) Identification of an alternate type I insulin-like growth factor receptor beta subunit mRNA transcript. *J Biol Chem*, **264**, 21439-41.

Yip, C.C., Hsu, H., Patel, R.G., Hawley, D.M., Maddux, B.A. and Goldfine, I.D. (1988) Localization of the insulin-binding site to the cysteine-rich region of the insulin receptor alpha-subunit. *Biochem Biophys Res Commun*, **157**, 321-9.

Yu, D., Watanabe, H., Shibuya, H. and Miura, M. (2003) Redundancy of radioresistant signaling pathways originating from insulin-like growth factor I receptor. *J Biol Chem*, **278**, 6702-9.

Zhang, X., Kamaraju, S., Hakuno, F., Kabuta, T., Takahashi, S., Sachdev, D. and Yee, D. (2004) Motility response to insulin-like growth factor-I (IGF-I) in MCF-7 cells is associated with IRS-2 activation and integrin expression. *Breast Cancer Res Treat*, **83**, 161-70.

Zisman, A., Peroni, O.D., Abel, E.D., Michael, M.D., Mauvais-Jarvis, F., Lowell, B.B., Wojtaszewski, J.F., Hirshman, M.F., Virkamaki, A., Goodyear, L.J., Kahn, C.R. and Kahn, B.B. (2000) Targeted disruption of the glucose transporter 4 selectively in muscle causes insulin resistance and glucose intolerance. *Nat Med*, **6**, 924-8.

APPENDIX

A. Percentage similarity of ins peptide and template amino acid sequences

	IGF-1	IGF-2	Bombyxin-II	Insulin (bovine)	Insulin (human)	Insulin (Porcine)	Mini-proinsulin	Relaxin	Max%
INS-1	25	28	**31**	**31**	**31**	**31**	30	29	31
INS-2	22	22	25	23	23	23	24	**27**	27
INS-3	21	**23**	19	19	19	19	17	**23**	23
INS-4	23	18	**25**	21	23	23	21	23	25
INS-5	21	19	23	21	**27**	21	21	19	27
INS-6	27	24	**34**	27	27	27	22	23	34
INS-7	20	20	23	19	19	19	18	**31**	31
INS-8	18	18	**25**	23	23	23	20	**25**	25
INS-9	10	14	21	7	7	7	14	**27**	27
INS-10	8	8	12	11	**13**	**13**	12	7	13
INS-11	22	22	**25**	21	21	21	24	9	25
INS-12	17	19	19	19	19	19	19	**21**	21
INS-13	15	18	20	15	15	15	20	**25**	25
INS-14	14	14	**21**	19	**21**	**21**	18	**21**	21
INS-15	17	17	**25**	17	17	17	21	23	25
INS-16	18	16	17	17	17	17	16	**23**	23
INS-17	17	16	8	**19**	**19**	**19**	18	15	19
INS-18	**21**	11	17	13	13	13	11	15	21
INS-19	15	13	17	**23**	**23**	**23**	20	21	23
INS-20	21	**23**	**23**	10	10	10	10	19	23
INS-21	**12**	9	8	9	9	9	8	7	12
INS-22	9	9	14	**19**	**19**	**19**	7	**19**	19
INS-23	16	6	8	14	14	14	**18**	12	18
INS-24	**21**	14	19	17	15	17	8	**21**	21
INS-25	18	18	21	**23**	21	21	22	**23**	23
INS-26	12	16	17	7	7	7	10	**21**	21
INS-27	12	15	**19**	11	11	11	16	13	19
INS-28	24	**26**	19	12	12	12	12	22	26
INS-29	26	19	25	**33**	**33**	**33**	28	25	33
INS-30	12	18	**21**	11	13	13	16	**21**	21
INS-31a	13	20	**23**	17	15	15	13	**23**	23
INS-31b	17	10	21	21	**23**	**23**	21	9	23
INS-31c	7	6	10	**17**	**17**	**17**	15	11	17
INS-32	16	18	**23**	16	16	16	16	20	23
INS-33	**20**	11	12	11	11	11	11	17	20
INS-34	13	6	10	5	5	5	10	**15**	15
INS-35	4	6	10	9	15	15	8	**17**	17
INS-36	5	4	10	17	15	15	6	**19**	19
INS-37	16	20	19	21	21	21	18	**23**	23
DAF-28	24	24	25	31	**33**	**33**	26	25	33
AVG	**16.725**	**15.95**	**19.1**	**17.4**	**17.8**	**17.7**	**16.625**	**19.725**	**22.875**
No of Max	**5**	**3**	**12**	**7**	**10**	**10**	**1**	**19**	

B. Objective Function output from MODELLER (Lower number rated a better structure)

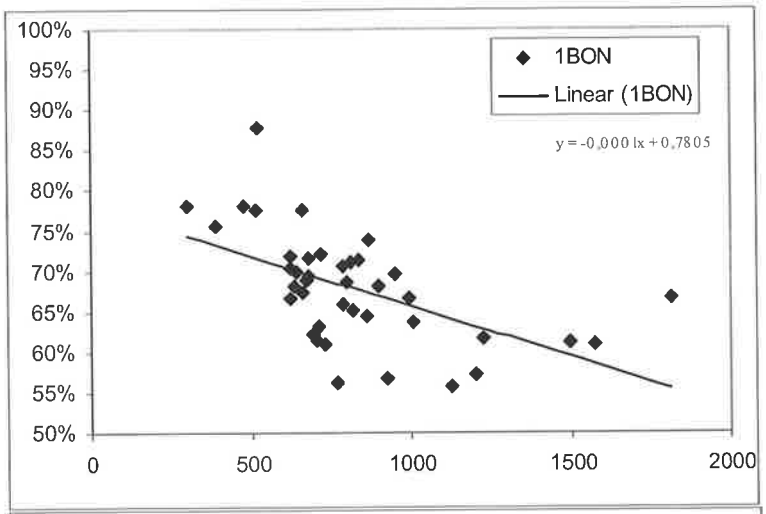
	1APH	1B2F	1B9G	1BON	1EFE	1HLS	1IGL	1IMX	1MHI	1MHJ	3LRI	6RLX	9INS		
ID	bInsulin	pInsulin	hIGF-1	Bombyxin-II	hminiproins	hInsulin	hIGF-2	hIGF-1	hInsulin	hInsulin	hLR3IGF-1	hRelaxin	pInsulin	Minimum	Average
Daf-28	527	529	866	727	1527	1190	2983	355	1508	605	3788	2515	547	355	1359
Ins-01	410	396	736	521	449	450	746	494	1009	555	1113	518	402	396	600
Ins-02	724	762	781	767	775	1430	901	359	1597	563	3919	2338	740	359	1204
Ins-03	687	714	570	814	1531	1337	2964	334	1497	517	3424	2169	685	334	1326
Ins-04	500	573	872	708	1591	1436	2980	361	1623	548	3697	2426	532	361	1373
Ins-05	574	534	736	693	1529	1207	2890	296	1479	617	3646	2553	598	296	1335
Ins-06	493	473	664	644	1490	1153	2797	415	1463	512	3602	2269	503	415	1268
Ins-07	475	488	671	660	1526	1257	2818	271	1419	528	3541	2704	515	271	1298
Ins-08	534	491	697	633	1524	1195	2862	290	1504	518	3596	2330	557	290	1287
Ins-09	615	610	875	677	1588	1396	3202	388	1526	689	3010	3051	658	388	1407
Ins-10	1613	1670	2123	1499	1979	2378	2373	2256	2701	1980	4390	3406	1681	1499	2311
Ins-11	244	316	421	305	254	269	380	292	416	305	449	232	249	232	318
Ins-12	465	452	695	1203	1308	957	2674	2388	1293	477	1966	445	509	445	1141
Ins-13	531	539	540	992	1264	965	2418	2171	1276	428	2315	791	541	428	1136
Ins-14	547	542	752	712	1289	1234	2822	2376	1471	539	2634	1854	541	539	1332
Ins-15	1251	1233	732	393	1062	1998	2119	1193	2097	897	2617	2976	1318	393	1530
Ins-16	1048	1000	1027	1574	1405	1521	2652	2401	1589	773	2226	917	1009	773	1472
Ins-17	1058	949	1282	625	718	1710	934	946	1911	853	1427	3686	898	625	1307
Ins-18	818	965	1190	658	653	1618	727	942	2063	852	1051	3548	971	653	1235
Ins-19	712	919	1109	622	1546	1564	2160	1223	2016	841	3138	2932	914	622	1515
Ins-20	719	704	650	924	1552	1326	2879	472	1448	543	3434	1855	738	472	1327
Ins-21	773	769	1025	623	1756	1596	3196	659	1922	874	4134	3322	784	623	1649
Ins-22	528	501	791	1006	919	1372	917	507	1462	749	2260	1625	528	501	1013
Ins-23	485	564	818	1126	1636	1247	2883	2454	1378	670	3516	2014	509	485	1485
Ins-24	627	667	642	817	1535	1276	2936	306	1403	513	3616	1609	674	306	1278
Ins-25	509	541	656	717	1560	1199	2965	285	1538	496	3655	2346	560	285	1310
Ins-26	639	636	748	949	1527	1300	3202	315	1395	608	3475	2157	668	315	1355
Ins-27	1181	1169	800	1809	1784	1470	3037	784	1555	815	3662	1180	1056	784	1562
Ins-28	501	445	756	788	1679	1163	3022	249	1434	517	3775	1798	472	249	1277
Ins-29	537	507	708	674	1545	1350	2987	506	1622	549	3713	2415	545	506	1358
Ins-30	655	649	654	863	1561	1293	3040	341	1454	496	3519	1872	671	341	1313
Ins-31a	944	817	1249	513	1386	1580	2574	2632	2054	872	2330	3054	900	513	1608
Ins-31b	1128	792	1002	838	1853	1625	3943	1078	2020	756	3625	3102	977	756	1749
Ins-31c	1129	1319	1082	786	1092	1563	2810	1661	2118	955	3013	3374	1161	786	1697
Ins-32	1655	1706	1332	678	1344	2460	2240	1493	2612	1082	2882	3239	1730	678	1881
Ins-33	716	728	741	902	1583	1375	3050	400	1480	589	3563	2151	763	400	1388
Ins-34	836	941	1000	798	1607	1651	3593	820	1931	1127	5687	3618	942	798	1888
Ins-35	1080	1131	817	1228	1827	1542	3214	614	1617	949	3814	1729	1387	614	1611
Ins-36	1158	1221	934	869	2003	1605	3427	775	1942	1211	4612	3429	904	775	1853
Ins-37	775	790	1071	479	997	1420	2373	1592	1995	702	2675	2919	822	479	1432
Minimum	244	316	421	305	254	269	380	249	416	305	449	232	249	232	318
Average	760	769	870	820	1394	1392	2567	942	1646	717	3163	2312	779	509	1395

C. Comparison of Objective Function and Procheck results.

ID	Objective Function			Procheck Value (core)	
	Minimum	Average	Best Template	Maximum	Average
Daf-28	354.89	1359.04	hIGF-1 (1IMX)	91.30%	73.08%
Ins-01	395.76	599.87	pInsulin (1B2F)	89.00%	79.97%
Ins-02	359.15	1204.36	hIGF-1 (1IMX)	91.70%	70.82%
Ins-03	334.21	1326.28	hIGF-1 (1IMX)	84.40%	71.62%
Ins-04	360.94	1372.78	hIGF-1 (1IMX)	88.60%	72.21%
Ins-05	296.38	1334.86	hIGF-1 (1IMX)	84.40%	69.57%
Ins-06	414.96	1267.53	hIGF-1 (1IMX)	88.40%	74.05%
Ins-07	270.93	1297.95	hIGF-1 (1IMX)	89.10%	70.05%
Ins-08	289.90	1287.00	hIGF-1 (1IMX)	89.40%	68.76%
Ins-09	388.18	1406.65	hIGF-1 (1IMX)	89.80%	73.00%
Ins-10	1499.05	2311.45	Bombyxin-II (1BON)	83.70%	63.11%
Ins-11	232.05	317.79	hRelaxin (6RLX)	96.00%	81.54%
Ins-12	445.22	1140.99	hRelaxin (6RLX)	83.70%	70.34%
Ins-13	427.74	1136.26	hInsulin (1MHJ)	90.50%	71.79%
Ins-14	538.70	1331.82	hInsulin (1MHJ)	85.70%	71.91%
Ins-15	393.03	1529.69	Bombyxin-II (1BON)	88.90%	72.32%
Ins-16	772.88	1472.48	hInsulin (1MHJ)	80.40%	68.22%
Ins-17	624.65	1307.46	Bombyxin-II (1BON)	87.00%	70.52%
Ins-18	652.53	1234.99	hMini-proins(1EFE)	79.30%	69.25%
Ins-19	622.14	1515.06	Bombyxin-II (1BON)	78.00%	68.15%
Ins-20	471.97	1326.50	hIGF-1 (1IMX)	84.10%	69.56%
Ins-21	623.46	1648.65	Bombyxin-II (1BON)	81.50%	66.25%
Ins-22	500.99	1012.85	pInsulin (1B2F)	93.60%	76.60%
Ins-23	485.31	1484.70	bInsulin (1APH)	86.00%	72.96%
Ins-24	305.66	1278.43	hIGF-1 (1IMX)	86.00%	72.44%
Ins-25	285.21	1309.75	hIGF-1 (1IMX)	86.00%	72.98%
Ins-26	315.11	1355.40	hIGF-1 (1IMX)	84.80%	72.74%
Ins-27	784.49	1561.63	hIGF-1 (1IMX)	83.30%	72.15%
Ins-28	249.25	1276.83	hIGF-1 (1IMX)	90.20%	75.41%
Ins-29	506.12	1358.28	hIGF-1 (1IMX)	84.40%	72.98%
Ins-30	341.01	1312.78	hIGF-1 (1IMX)	86.70%	72.14%
Ins-31a	512.93	1608.17	Bombyxin-II (1BON)	85.70%	71.27%
Ins-31b	755.56	1749.12	hInsulin (1MHJ)	81.60%	68.75%
Ins-31c	786.32	1697.19	Bombyxin-II (1BON)	88.00%	69.38%
Ins-32	677.57	1881.00	Bombyxin-II (1BON)	80.40%	68.72%
Ins-33	400.37	1387.67	hIGF-1 (1IMX)	85.10%	72.18%
Ins-34	798.07	1888.46	Bombyxin-II (1BON)	76.50%	65.46%
Ins-35	613.50	1611.47	hIGF-1 (1IMX)	80.90%	65.63%
Ins-36	775.15	1853.04	hIGF-1 (1IMX)	81.50%	65.08%
Ins-37	479.25	1431.62	Bombyxin-II (1BON)	82.00%	71.08%
AVG	508.52	1394.70	hIGF-1 (1IMX)	85.69%	71.10%

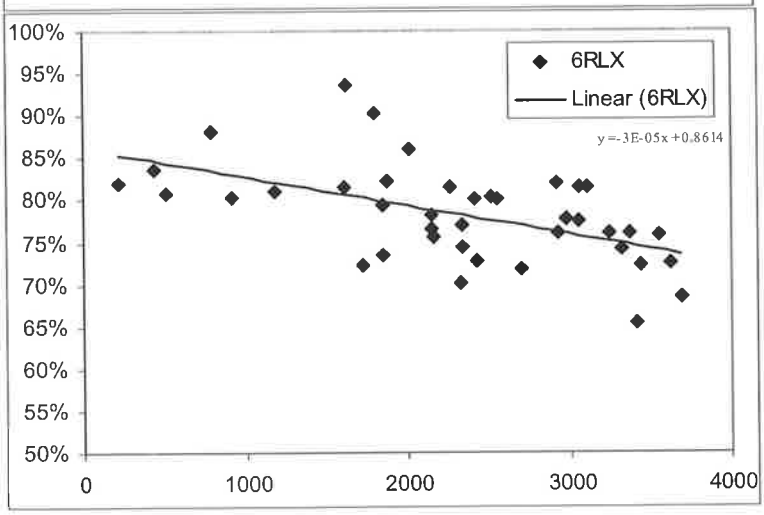
Bombyxin-II

Average Model OF value = 820
Number of ins peptide minimums = 10
Number with OF value <500 = 3
Average Model Procheck % (core) = 68



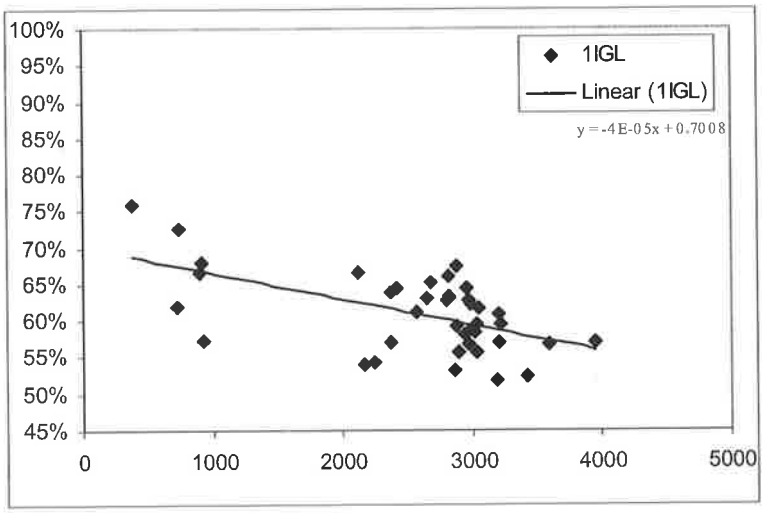
Human Relaxin

Average Model OF value = 2311
Number of ins peptide minimums = 2
Number with OF value <500 = 2
Average Model Procheck % (core) = 78



Human IGF-2

Average Model OF value = 2567
Number of ins peptide minimums = 0
Number with OF value <500 = 1
Average Model Procheck % (core) = 61



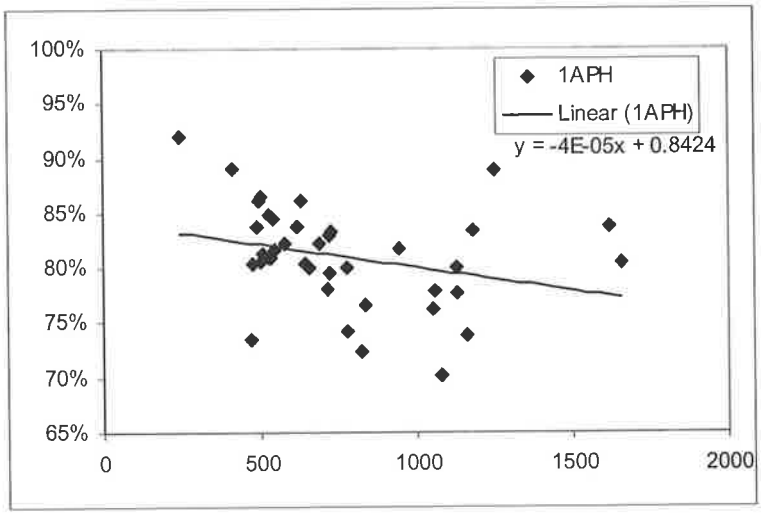
Comparison of templates used in the comparative modelling.

Scatter plots of the Objective Function (OF) on the x-axis vs the Procheck % score on the y-axis for each Ins peptide based on the templates.

A. 1BON
B. 6RLX
C. 1IGL

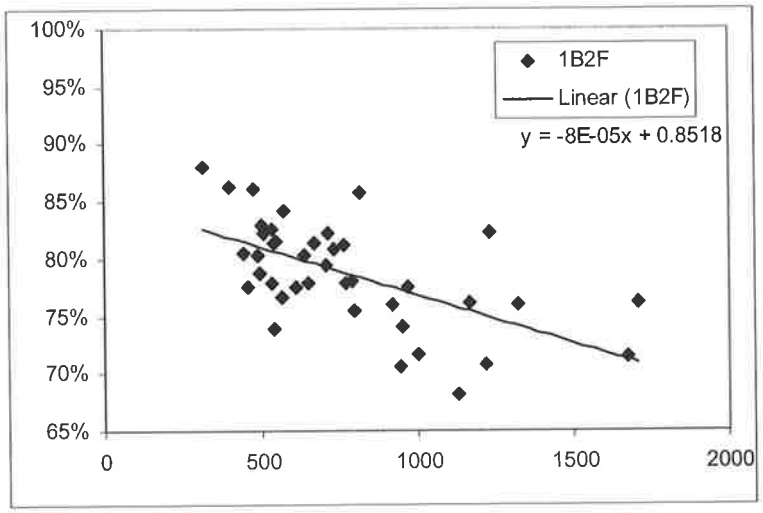
Bovine Insulin

Average Model OF value = 760
Number of ins peptide minimums = 1
Number with OF value <500 = 6
Average Model Procheck % (core) = 81



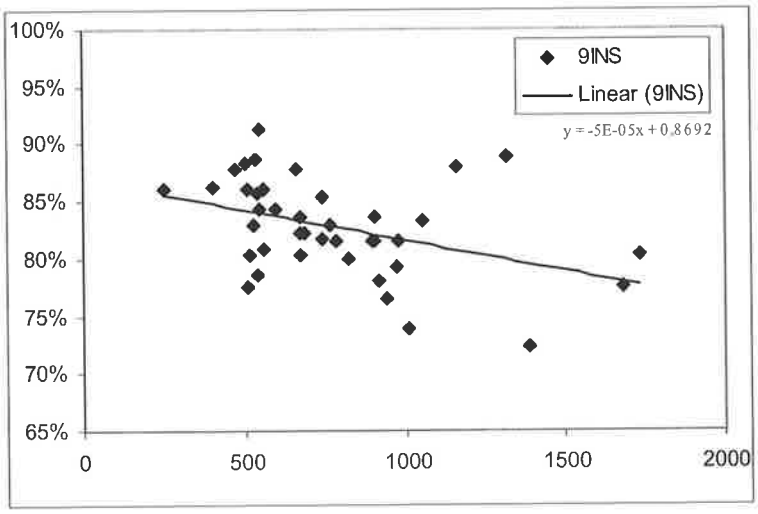
Pig Insulin

Average Model OF value = 769
Number of ins peptide minimums = 2
Number with OF value <500 = 7
Average Model Procheck % (core) = 79



Pig Insulin

Average Model OF value = 779
Number of ins peptide minimums = 0
Number with OF value <500 = 3
Average Model Procheck % (core) = 83



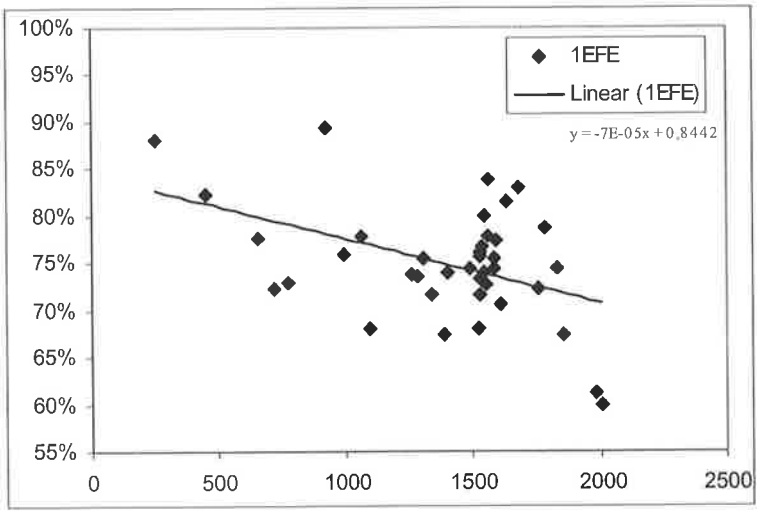
Comparison of templates used in the comparative modelling.

Scatter plots of the Objective Function (OF) on the x-axis vs the Procheck % score on the y-axis for each Ins peptide based on the templates.

A. 1APH
B. 1B2F
C. 9INS

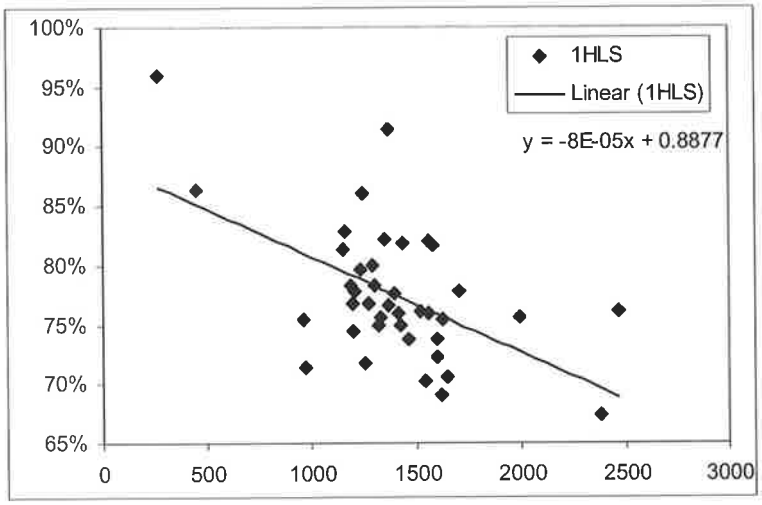
Human mini-proinsulin 1EFE

Average Model OF value = 1394
Number of ins peptide minimums = 1
Number with OF value <500 = 2
Average Model Procheck % (core) = 75



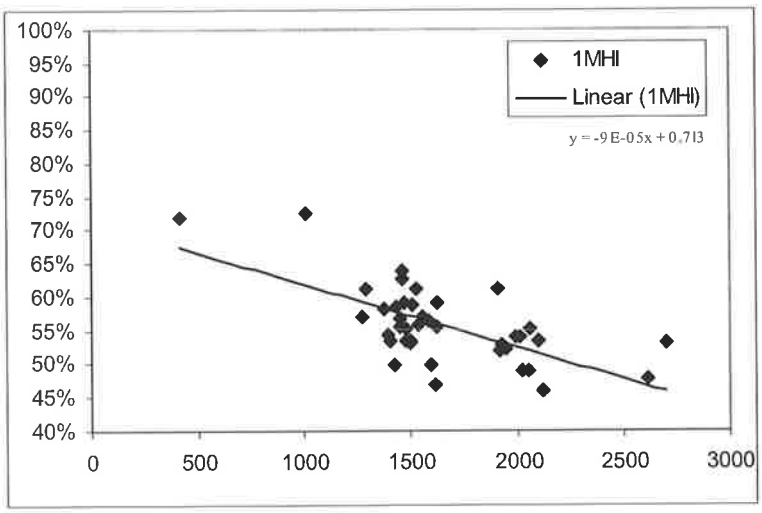
Human Insulin 1HLS

Average Model OF value = 1392
Number of ins peptide minimums = 0
Number with OF value <500 = 2
Average Model Procheck % (core) = 78



Human Insulin 1MHI

Average Model OF value = 1646
Number of ins peptide minimums = 0
Number with OF value <500 = 1
Average Model Procheck % (core) = 56



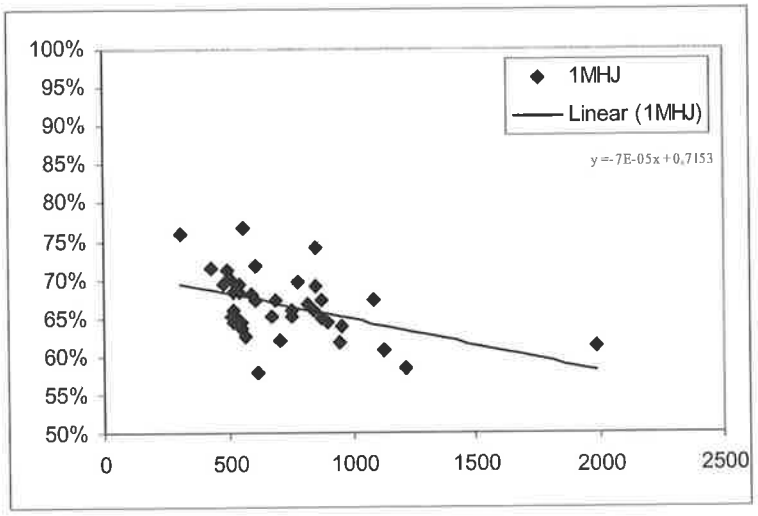
Comparison of templates used in the comparative modelling.

Scatter plots of the Objective Function (OF) on the x-axis vs the Procheck % score on the y-axis for each Ins peptide based on the templates.

A. 1EFE
B. 1HLS
C. 1MHI.

Human Insulin 1MHJ

Average Model OF value = 717
Number of ins peptide minimums = 4
Number with OF value <500 = 5
Average Model Procheck % (core) = 67



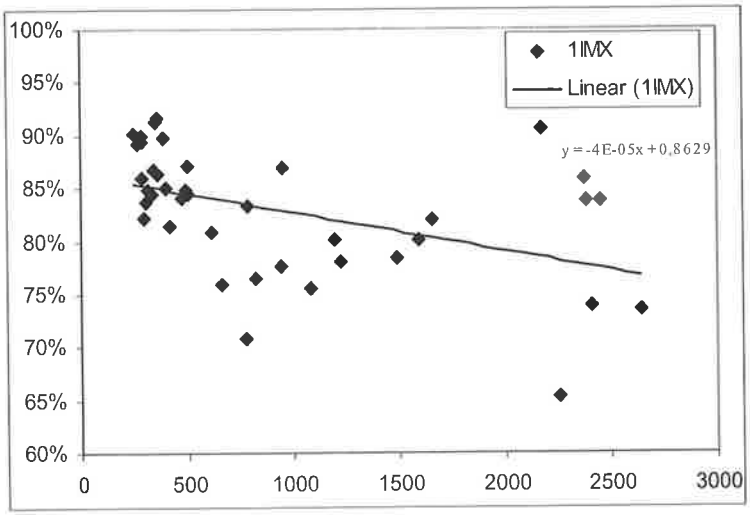
Comparison of templates used in the comparative modelling.

Scatter plots of the Objective Function (OF) on the x-axis vs the Procheck % score on the y-axis for each Ins peptide based on the templates.

A. 1MHJ

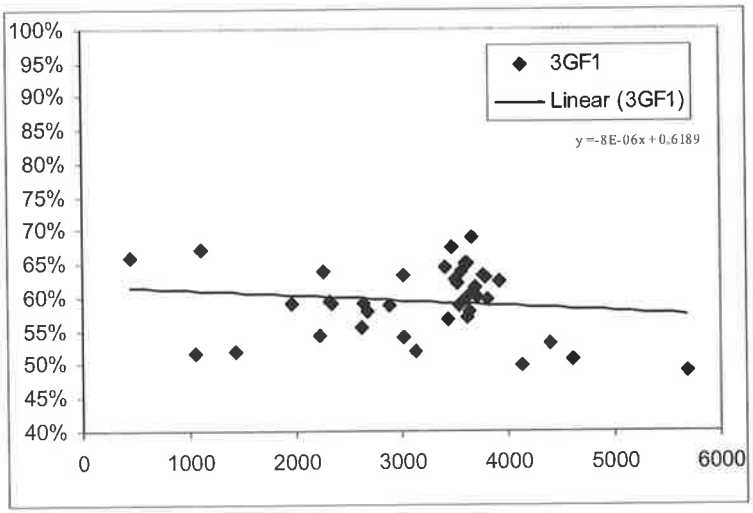
Human IGF-1 1IMX

Average Model OF value = 942
Number of ins peptide minimums = 20
Number with OF value <500 = 18
Average Model Procheck % (core) = 83



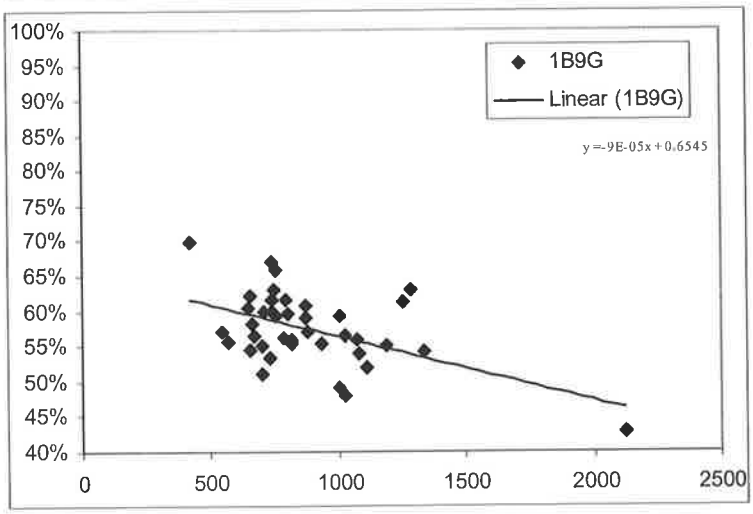
Human IGF-1 3GF1

Average Model OF value = 3162
Number of ins peptide minimums = 0
Number with OF value <500 = 1
Average Model Procheck % (core) = 59



Human IGF-1 1B9G

Average Model OF value = 870
Number of ins peptide minimums = 0
Number with OF value <500 = 3
Average Model Procheck % (core) = 58



Comparison of templates used in the comparative modelling.

Scatter plots of the Objective Function (OF) on the x-axis vs the Procheck % score on the y-axis for each Ins peptide based on the templates.

A. 1IMX
B. 3GF1
C. 1B9G.

D. Tables of Z and Q scores of models before and after simulated annealing

BEFORE SIMULATED ANNEALING																	
z-scores		1APH	1B2F	1B9G	1BON	1EFE	1HLS	1IGL	1IMX	1MHI	1MHJ	3GF1	3LRI	6RLX	9INS		MINIMUM	AVG
	Res.	bInsulin	pInsulin	hIGF-1	Bombyxin-II	hminiproins	hInsulin	hIGF-2	hIGF-1	hInsulin	hInsulin	hIGF-1	hLR3IGF-1	hRelaxin	pInsulin			
Daf-28	54	-4.96	-4.02	-5.67	-4.65	-3.45	-3.92	-5.34	-4.67	-5.05	-4.15	-5.06	-5.28	-4.85	-5.35		-5.67	-4.74
Ins-01	80	-4.57	-3.76	-4.67	-3.37	-3	-4.1	-3.38	-4.14	-4.34	-4.24	-3.77	-3.23	-3.84	-4.06		-4.67	-3.89
Ins-02	53	-5.89	-6.1	-6.17	-5.09	-3.74	-5.47	-5.82	-5.35	-6.35	-4.43	-5.83	-5.15	-4.97	-6.06		-6.35	-5.46
Ins-03	51	-5.2	-5.65	-5.55	-5.8	-3.86	-5.65	-4.98	-6.42	-7.35	-5.1	-5.38	-5.74	-6.57	-5.78		-7.35	-5.65
Ins-04	52	-5.46	-5.42	-5.76	-5.37	-3.37	-5.1	-6.35	-4.92	-6.51	-5.82	-6.22	-5.91	-5.46	-5.15		-6.51	-5.49
Ins-05	52	-5.39	-5.07	-6.37	-5.25	-3.64	-4.1	-6.13	-4.79	-5.89	-5.68	-5.57	-6.28	-5.71	-5.58		-6.37	-5.39
Ins-06	50	-6.22	-6.16	-7.46	-6.88	-4.76	-5.93	-6.38	-7.58	-7.17	-6.32	-6.94	-7.5	-6.87	-6.57		-7.58	-6.62
Ins-07	50	-5.58	-5.9	-6.74	-6.62	-3.92	-5.67	-5.94	-5.87	-6.78	-6.55	-6.05	-5.75	-5.84	-5.75		-6.78	-5.93
Ins-08	50	-5.68	-5.61	-6.81	-7.17	-4.51	-6.03	-6.34	-5.72	-6.83	-6.04	-6.62	-5.66	-6.93	-5.6		-7.17	-6.11
Ins-09	57	-4.93	-4.56	-5.51	-5.02	-2.48	-4.28	-4.44	-4.26	-5.48	-4.12	-3.97	-4	-4.79	-4.78		-5.51	-4.47
Ins-10	55	-2.43	-2.52	-3.88	-2.41	-1.36	-3.67	-3.23	-3.07	-3.91	-2.61	-2.23	-2	-5.91	-2.78		-5.91	-3.00
Ins-11	53	-3.66	-4.18	-5.04	-3.81	-2.94	-3.57	-3.27	-3.81	-4.79	-4.23	-4.5	-3.65	-5.04	-4.23		-5.04	-4.05
Ins-12	53	-4.54	-4.55	-5.71	-5.74	-3.03	-4.14	-4.17	-3.9	-5.5	-4.9	-4.59	-3.9	-4.88	-4.49		-5.74	-4.57
Ins-13	48	-5.19	-5.15	-6.11	-4.4	-3.86	-5.08	-4.88	-4.22	-6.05	-4.7	-5.33	-4.75	-4.61	-5.53		-6.11	-4.99
Ins-14	53	-4.68	-4.4	-5.93	-5.07	-2.74	-4.22	-4.65	-4.17	-4.87	-4.51	-5.29	-4.54	-5.62	-4.86		-5.93	-4.68
Ins-15	51	-4.92	-5.35	-5.58	-4.94	-3.19	-4.8	-4.11	-4.68	-5.48	-4.96	-5.67	-4.27	-5.41	-5.42		-5.67	-4.91
Ins-16	50	-4.28	-4.04	-4.54	-4.69	-3.13	-2.9	-3.64	-3.12	-4.46	-3.48	-4.19	-3.15	-4.5	-4.67		-4.69	-3.91
Ins-17	67	-5.17	-5.02	-4.75	-5.75	-3.99	-4.93	-5.69	-5.32	-5.28	-4.98	-4.39	-4.08	-5.37	-4.79		-5.75	-4.97
Ins-18	68	-4.14	-4.77	-3.8	-4.35	-3	-3.86	-4.21	-3.31	-4.16	-3.72	-4.17	-3.33	-4.28	-3.68		-4.77	-3.91
Ins-19	57	-5.24	-5.37	-5.27	-5.2	-4.45	-4.81	-4.81	-4.48	-6.38	-5.01	-5.31	-4.14	-4.29	-5.26		-6.38	-5.00
Ins-20	50	-6.47	-6.36	-7.24	-6.58	-4.67	-5.75	-5.25	-5.86	-6.42	-5.22	-5.96	-5.5	-6.42	-5.69		-7.24	-5.96
Ins-21	63	-5.09	-5.46	-6	-4.27	-3.61	-5.1	-3.77	-3.9	-6.07	-5.81	-3.23	-3.92	-4.24	-5.41		-6.07	-4.71
Ins-22	57	-6.12	-5.33	-6.56	-5.77	-3.95	-4.6	-5.24	-5.14	-5.43	-5.44	-5.52	-3.83	-5.3	-5.51		-6.56	-5.27
Ins-23	54	-5.17	-4.67	-5.81	-5.94	-3.24	-4.92	-5.06	-5.49	-5.6	-5.26	-4.78	-4.17	-4.63	-5.58		-5.94	-5.02
Ins-24	53	-5.23	-4.04	-5.48	-5.69	-4.25	-4.6	-6.04	-4.88	-5.79	-4.36	-5.42	-5.69	-5.48	-4.79		-6.04	-5.12
Ins-25	52	-5.72	-5.87	-6.25	-6.17	-4.38	-5.08	-5.71	-4.87	-5.82	-4.7	-5.92	-6.08	-6.27	-5.57		-6.27	-5.60
Ins-26	53	-4.42	-4.82	-4.91	-4.89	-3.96	-4.6	-4.67	-5.96	-5.33	-4.5	-5.25	-5.38	-5.74	-4.95		-5.96	-4.96
Ins-27	53	-4.4	-4.56	-4.69	-3.97	-3.72	-4.42	-4.3	-4.34	-5.25	-3.44	-4.06	-5.27	-5.15	-3.89		-5.27	-4.39
Ins-28	49	-6.48	-7	-6.29	-5.99	-4.69	-6.2	-4.58	-5.6	-7	-5.75	-6.5	-4.93	-5.39	-7.12		-7.12	-5.97
Ins-29	54	-6.19	-6.73	-6.63	-7.01	-4.43	-5.27	-6.69	-5.58	-7.09	-6	-6.6	-6.14	-6.9	-5.91		-7.09	-6.23
Ins-30	53	-4.32	-4.24	-4.89	-5.45	-3.66	-4.93	-4.7	-5.27	-5.22	-4.02	-5.16	-5.04	-5.77	-4.16		-5.77	-4.77
Ins-31a	56	-5.66	-5.38	-5.82	-3.91	-4.49	-5.19	-5.23	-4.56	-5.81	-5.05	-4.83	-4.1	-3.76	-5.73		-5.82	-4.97
Ins-31b	57	-5.41	-5.86	-6.62	-5.42	-5.37	-5.7	-4.48	-4.65	-7.02	-3.72	-5.2	-4.51	-4.5	-5.9		-7.02	-5.31
Ins-31c	57	-5.73	-5.74	-6.09	-6.41	-5.19	-5.82	-5.04	-5.62	-6.26	-4.33	-5.73	-4.51	-5.99	-5.34		-6.41	-5.56
Ins-32	52	-6.52	-5.91	-5.82	-5.61	-5.85	-5.93	-6.32	-6.54	-6.77	-5.84	-5.83	-4.85	-6.21	-5.96		-6.77	-6.00
Ins-33	53	-5.08	-5.15	-6.19	-5.35	-3.89	-5.42	-5.19	-4.95	-6.2	-5.4	-5.17	-5.35	-5.33	-4.75		-6.20	-5.24
Ins-34	60	-4.3	-4.57	-3.97	-4.01	-3.55	-3.75	-4.82	-4.08	-5.23	-3.5	-4.91	-3.84	-4.76	-4.23		-5.23	-4.25
Ins-35	54	-3.38	-4.14	-4.93	-4.32	-4.23	-4.08	-4.98	-4.4	-5	-3.02	-4.4	-4.3	-4.77	-3.78		-5.00	-4.27
Ins-36	70	-4.05	-3.26	-3.42	-2.16	-2.23	-3.94	-2.1	-2.45	-4.55	-3.94	-2.19	-2.51	-3.39	-4.06		-4.55	-3.16
Ins-37	55	-5.12	-4.86	-6.43	-4.82	-4.56	-3.81	-5.15	-4.97	-5	-4.44	-4.91	-4.62	-5.95	-4.81		-6.43	-4.96
	MIN	-6.52	-7	-7.46	-7.17	-5.85	-6.2	-6.69	-7.58	-7.35	-6.55	-6.94	-7.5	-6.93	-7.12			
	AVG	-5.07475	-5.03875	-5.634	-5.133	-3.8085	-4.7835	-4.927	-4.82275	-5.73725	-4.73225	-5.06625	-4.67125	-5.29225	-5.08825			

BEFORE SIMULATED ANNEALING																		
q-scores		1APH	1B2F	1B9G	1BON	1EFE	1HLS	1IGL	1IMX	1MHI	1MHJ	3GF1	3LRI	6RLX	9INS		MINIMUM	AVG
	Res.	bInsulin	pInsulin	hIGF-1	Bombyxin-II	hminiproins	hInsulin	hIGF-2	hIGF-1	hInsulin	hInsulin	hIGF-1	hLR3IGF-1	hRelaxin	pInsulin			
Daf-28	54.00	-1.24	-1.01	-1.42	-1.17	-0.86	-0.98	-1.34	-1.17	-1.27	-1.04	-1.27	-1.32	-1.22	-1.34		-1.42	-1.19
Ins-01	80.00	-1.04	-0.86	-1.07	-0.77	-0.68	-0.94	-0.77	-0.94	-0.99	-0.97	-0.86	-0.74	-0.88	-0.93		-1.07	-0.89
Ins-02	53.00	-1.48	-1.54	-1.55	-1.28	-0.94	-1.38	-1.47	-1.35	-1.60	-1.12	-1.47	-1.30	-1.25	-1.53		-1.60	-1.37
Ins-03	51.00	-1.32	-1.44	-1.41	-1.48	-0.98	-1.44	-1.27	-1.63	-1.87	-1.30	-1.37	-1.46	-1.67	-1.47		-1.87	-1.44
Ins-04	52.00	-1.38	-1.37	-1.46	-1.36	-0.85	-1.29	-1.61	-1.25	-1.65	-1.47	-1.57	-1.50	-1.38	-1.30		-1.65	-1.39
Ins-05	52.00	-1.36	-1.28	-1.61	-1.33	-0.92	-1.04	-1.55	-1.21	-1.49	-1.44	-1.41	-1.59	-1.45	-1.41		-1.61	-1.36
Ins-06	50.00	-1.59	-1.57	-1.91	-1.76	-1.22	-1.52	-1.63	-1.94	-1.83	-1.62	-1.77	-1.92	-1.76	-1.68		-1.94	-1.69
Ins-07	50.00	-1.43	-1.51	-1.72	-1.69	-1.00	-1.45	-1.52	-1.50	-1.73	-1.67	-1.55	-1.47	-1.49	-1.47		-1.73	-1.51
Ins-08	50.00	-1.45	-1.43	-1.74	-1.83	-1.15	-1.54	-1.62	-1.46	-1.75	-1.54	-1.69	-1.45	-1.77	-1.43		-1.83	-1.56
Ins-09	57.00	-1.22	-1.13	-1.36	-1.24	-0.61	-1.06	-1.10	-1.05	-1.36	-1.02	-0.98	-0.99	-1.18	-1.18		-1.36	-1.11
Ins-10	55.00	-0.61	-0.63	-0.97	-0.60	-0.34	-0.92	-0.81	-0.77	-0.98	-0.65	-0.56	-0.50	-1.47	-0.69		-1.47	-0.75
Ins-11	53.00	-0.92	-1.05	-1.27	-0.96	-0.74	-0.90	-0.82	-0.96	-1.21	-1.07	-1.13	-0.92	-1.27	-1.07		-1.27	-1.02
Ins-12	53.00	-1.14	-1.15	-1.44	-1.45	-0.76	-1.04	-1.05	-0.98	-1.39	-1.23	-1.16	-0.98	-1.23	-1.13		-1.45	-1.15
Ins-13	48.00	-1.34	-1.33	-1.58	-1.14	-1.00	-1.31	-1.26	-1.09	-1.56	-1.21	-1.38	-1.23	-1.19	-1.43		-1.58	-1.29
Ins-14	53.00	-1.18	-1.11	-1.49	-1.28	-0.69	-1.06	-1.17	-1.05	-1.23	-1.14	-1.33	-1.14	-1.42	-1.22		-1.49	-1.18
Ins-15	51.00	-1.25	-1.36	-1.42	-1.26	-0.81	-1.22	-1.05	-1.19	-1.39	-1.26	-1.44	-1.09	-1.38	-1.38		-1.44	-1.25
Ins-16	50.00	-1.09	-1.03	-1.16	-1.20	-0.80	-0.74	-0.93	-0.80	-1.14	-0.89	-1.07	-0.81	-1.15	-1.19		-1.20	-1.00
Ins-17	67.00	-1.23	-1.19	-1.13	-1.37	-0.95	-1.17	-1.35	-1.27	-1.26	-1.18	-1.04	-0.97	-1.28	-1.14		-1.37	-1.18
Ins-18	68.00	-0.98	-1.13	-0.90	-1.03	-0.71	-0.91	-1.00	-0.78	-0.99	-0.88	-0.99	-0.79	-1.01	-0.87		-1.13	-0.93
Ins-19	57.00	-1.30	-1.33	-1.30	-1.29	-1.10	-1.19	-1.19	-1.11	-1.58	-1.24	-1.31	-1.02	-1.06	-1.30		-1.58	-1.24
Ins-20	50.00	-1.65	-1.63	-1.85	-1.68	-1.19	-1.47	-1.34	-1.50	-1.64	-1.33	-1.52	-1.41	-1.64	-1.45		-1.85	-1.52
Ins-21	63.00	-1.23	-1.32	-1.45	-1.03	-0.87	-1.23	-0.91	-0.94	-1.47	-1.40	-0.78	-0.95	-1.02	-1.31		-1.47	-1.14
Ins-22	57.00	-1.51	-1.32	-1.62	-1.43	-0.98	-1.14	-1.30	-1.27	-1.34	-1.35	-1.37	-0.95	-1.31	-1.36		-1.62	-1.30
Ins-23	54.00	-1.30	-1.17	-1.46	-1.49	-0.81	-1.23	-1.27	-1.38	-1.40	-1.32	-1.20	-1.05	-1.16	-1.40		-1.49	-1.26
Ins-24	53.00	-1.32	-1.02	-1.38	-1.43	-1.07	-1.16	-1.52	-1.23	-1.46	-1.10	-1.37	-1.43	-1.38	-1.21		-1.52	-1.29
Ins-25	52.00	-1.45	-1.49	-1.58	-1.56	-1.11	-1.29	-1.45	-1.23	-1.47	-1.19	-1.50	-1.54	-1.59	-1.41		-1.59	-1.42
Ins-26	53.00	-1.11	-1.21	-1.24	-1.23	-1.00	-1.16	-1.18	-1.50	-1.34	-1.13	-1.32	-1.36	-1.45	-1.25		-1.50	-1.25
Ins-27	53.00	-1.11	-1.15	-1.18	-1.00	-0.94	-1.11	-1.08	-1.09	-1.32	-0.87	-1.02	-1.33	-1.30	-0.98		-1.33	-1.11
Ins-28	49.00	-1.67	-1.80	-1.62	-1.54	-1.21	-1.59	-1.18	-1.44	-1.80	-1.48	-1.67	-1.27	-1.38	-1.83		-1.83	-1.53
Ins-29	54.00	-1.55	-1.69	-1.66	-1.76	-1.11	-1.32	-1.68	-1.40	-1.78	-1.50	-1.65	-1.54	-1.73	-1.48		-1.78	-1.56
Ins-30	53.00	-1.09	-1.07	-1.23	-1.37	-0.92	-1.24	-1.18	-1.33	-1.31	-1.01	-1.30	-1.27	-1.45	-1.05		-1.45	-1.20
Ins-31a	56.00	-1.41	-1.34	-1.45	-0.97	-1.12	-1.29	-1.30	-1.13	-1.44	-1.25	-1.20	-1.02	-0.93	-1.42		-1.45	-1.23
Ins-31b	57.00	-1.34	-1.45	-1.64	-1.34	-1.33	-1.41	-1.11	-1.15	-1.74	-0.92	-1.29	-1.12	-1.11	-1.46		-1.74	-1.31
Ins-31c	57.00	-1.42	-1.42	-1.51	-1.59	-1.28	-1.44	-1.25	-1.39	-1.55	-1.07	-1.42	-1.12	-1.48	-1.32		-1.59	-1.37
Ins-32	52.00	-1.65	-1.50	-1.47	-1.42	-1.48	-1.50	-1.60	-1.66	-1.71	-1.48	-1.48	-1.23	-1.57	-1.51		-1.71	-1.52
Ins-33	53.00	-1.28	-1.30	-1.56	-1.35	-0.98	-1.37	-1.31	-1.25	-1.56	-1.36	-1.30	-1.35	-1.34	-1.20		-1.56	-1.32
Ins-34	60.00	-1.05	-1.12	-0.97	-0.98	-0.87	-0.92	-1.18	-1.00	-1.28	-0.85	-1.20	-0.94	-1.16	-1.03		-1.28	-1.04
Ins-35	54.00	-0.85	-1.04	-1.24	-1.08	-1.06	-1.02	-1.25	-1.10	-1.25	-0.76	-1.10	-1.08	-1.20	-0.95		-1.25	-1.07
Ins-36	70.00	-0.95	-0.77	-0.80	-0.51	-0.52	-0.93	-0.49	-0.58	-1.07	-0.93	-0.52	-0.59	-0.80	-0.96		-1.07	-0.74
Ins-37	55.00	-1.28	-1.21	-1.60	-1.20	-1.14	-0.95	-1.29	-1.24	-1.25	-1.11	-1.23	-1.15	-1.48	-1.20		-1.60	-1.24
	MIN	-1.67	-1.80	-1.91	-1.83	-1.48	-1.59	-1.68	-1.94	-1.87	-1.67	-1.77	-1.92	-1.77	-1.83			
	AVG	-1.27	-1.26	-1.41	-1.29	-0.95	-1.20	-1.23	-1.21	-1.44	-1.18	-1.27	-1.17	-1.33	-1.27			

AFTER SIMULATED ANNEALING																		
z-scores		1APH	1B2F	1B9G	1BON	1EFE	1HLS	1IGL	1IMX	1MHI	1MHJ	3GF1	3LRI	6RLX	9INS	MINIMUM	AVG
	Res.	bInsulin	pInsulin	hIGF-1	Bombyxin-II	hminiproins	hInsulin	hIGF-2	hIGF-1	hInsulin	hInsulin	hIGF-1	hLR3IGF-1	hRelaxin	pInsulin		
Daf-28	54	-5.45	-4.76	-5.95	-4.72	-6.48	-5.53	-5.40	-5.00	-4.71	-5.71	-5.89	-5.28	-5.03	-5.87	-6.48	-5.41
Ins-01	80	-4.27	-4.18	-4.02	-3.57	-3.90	-4.51	-4.13	-3.62	-3.88	-4.07	-3.78	-3.08	-4.67	-4.58	-4.67	-4.02
Ins-02	53	-6.06	-6.07	-5.90	-4.83	-5.20	-5.86	-6.18	-5.41	-5.82	-4.78	-5.50	-4.49	-5.14	-5.90	-6.18	-5.51
Ins-03	51	-6.35	-6.02	-5.55	-6.16	-5.61	-6.44	-5.44	-6.40	-6.77	-4.78	-5.25	-5.95	-5.69	-5.26	-6.77	-5.83
Ins-04	52	-5.46	-6.32	-5.55	-5.48	-5.71	-6.38	-5.97	-5.02	-6.26	-5.15	-6.50	-5.49	-6.23	-6.05	-6.50	-5.83
Ins-05	52	-6.51	-5.99	-5.84	-6.67	-5.98	-5.18	-7.26	-5.68	-4.70	-6.14	-6.15	-7.28	-6.55	-5.57	-7.28	-6.11
Ins-06	50	-7.37	-6.96	-7.69	-6.65	-7.20	-6.54	-7.46	-7.01	-6.64	-5.98	-7.97	-6.47	-7.25	-6.57	-7.97	-6.98
Ins-07	50	-5.20	-5.58	-5.10	-7.08	-5.51	-6.99	-6.25	-5.69	-5.37	-6.50	-6.38	-4.57	-6.71	-6.07	-7.08	-5.93
Ins-08	50	-6.60	-6.36	-6.50	-6.56	-6.06	-6.49	-5.91	-5.94	-5.74	-6.36	-6.54	-4.20	-6.94	-6.81	-6.94	-6.22
Ins-09	57	-5.30	-5.29	-6.12	-4.44	-4.74	-5.67	-4.74	-5.32	-5.12	-5.07	-5.31	-4.13	-5.86	-5.22	-6.12	-5.17
Ins-10	55	-6.06	-4.25	-4.72	-4.05	-4.69	-4.53	-4.56	-5.28	-3.93	-4.03	-4.60	-3.87	-5.25	-4.01	-6.06	-4.56
Ins-11	53	-4.02	-5.43	-4.56	-5.58	-4.17	-4.40	-3.31	-3.14	-4.50	-4.49	-4.70	-3.32	-5.13	-4.24	-5.58	-4.36
Ins-12	53	-5.28	-5.29	-5.88	-5.36	-4.95	-4.62	-4.77	-4.01	-5.22	-5.36	-4.97	-3.73	-4.81	-4.50	-5.88	-4.91
Ins-13	48	-5.85	-5.61	-6.15	-4.25	-5.34	-5.06	-4.92	-4.99	-4.90	-5.60	-5.40	-5.13	-6.08	-4.73	-6.15	-5.29
Ins-14	53	-4.64	-4.53	-5.47	-4.56	-4.51	-4.46	-3.91	-4.93	-4.97	-4.59	-5.38	-4.72	-5.92	-4.29	-5.92	-4.78
Ins-15	51	-5.04	-5.41	-4.98	-4.88	-5.37	-5.02	-3.46	-4.90	-5.57	-5.64	-4.41	-5.23	-6.00	-5.24	-6.00	-5.08
Ins-16	50	-4.90	-4.89	-5.56	-5.61	-4.70	-4.03	-4.69	-3.95	-4.42	-4.77	-4.65	-3.89	-5.31	-4.18	-5.61	-4.68
Ins-17	67	-5.23	-5.87	-4.36	-4.71	-4.99	-6.46	-6.25	-5.42	-5.25	-4.19	-5.16	-5.29	-6.36	-6.13	-6.46	-5.41
Ins-18	68	-4.38	-4.20	-4.66	-3.39	-4.41	-3.65	-4.42	-3.38	-3.90	-4.15	-4.75	-3.61	-5.13	-3.93	-5.13	-4.14
Ins-19	57	-5.30	-4.91	-4.79	-5.05	-5.07	-5.49	-4.16	-5.17	-6.07	-5.34	-5.43	-5.10	-4.20	-4.80	-6.07	-5.06
Ins-20	50	-6.39	-6.47	-5.35	-5.25	-7.65	-6.05	-5.99	-5.31	-5.45	-6.31	-6.76	-4.97	-6.05	-6.31	-7.65	-6.02
Ins-21	63	-5.17	-6.11	-5.59	-4.55	-4.87	-5.35	-4.89	-3.89	-4.68	-5.80	-4.80	-5.05	-4.62	-6.09	-6.11	-5.10
Ins-22	57	-6.54	-5.54	-5.79	-6.48	-6.99	-5.96	-5.50	-6.09	-4.30	-4.50	-5.79	-5.02	-6.76	-5.79	-6.99	-5.79
Ins-23	54	-6.25	-5.73	-6.13	-6.79	-5.93	-5.70	-5.41	-5.46	-5.81	-5.26	-4.75	-5.01	-4.79	-6.07	-6.79	-5.65
Ins-24	53	-6.45	-5.75	-6.50	-5.76	-4.85	-5.79	-4.23	-6.16	-5.53	-4.82	-5.69	-6.54	-3.88	-5.08	-6.54	-5.50
Ins-25	52	-6.17	-6.26	-4.45	-6.66	-5.80	-6.04	-7.10	-5.68	-6.05	-7.08	-7.19	-5.85	-4.97	-6.03	-7.19	-6.10
Ins-26	53	-5.29	-5.49	-5.47	-4.64	-6.04	-5.36	-4.93	-5.64	-5.08	-3.96	-5.07	-5.58	-4.56	-5.11	-6.04	-5.16
Ins-27	53	-4.83	-4.20	-5.67	-4.84	-5.63	-4.34	-5.16	-4.74	-5.53	-3.82	-5.65	-4.87	-4.96	-5.34	-5.67	-4.97
Ins-28	49	-6.34	-6.55	-6.19	-4.82	-4.63	-6.04	-4.21	-5.49	-6.16	-7.98	-6.40	-6.20	-6.29	-7.29	-7.98	-6.04
Ins-29	54	-5.98	-7.24	-5.98	-7.59	-6.71	-6.44	-7.05	-5.01	-6.23	-6.03	-7.92	-6.32	-6.70	-6.23	-7.92	-6.53
Ins-30	53	-4.78	-5.55	-5.74	-5.92	-5.62	-5.45	-4.43	-5.28	-6.18	-4.58	-4.67	-5.52	-6.30	-5.12	-6.30	-5.37
Ins-31a	56	-5.79	-5.75	-4.88	-3.69	-5.43	-4.52	-5.37	-4.81	-5.63	-4.50	-5.06	-4.56	-4.31	-5.61	-5.79	-4.99
Ins-31b	57	-6.06	-7.01	-6.89	-6.35	-5.30	-5.94	-4.84	-5.46	-6.15	-5.24	-5.34	-5.39	-5.05	-6.60	-7.01	-5.83
Ins-31c	57	-6.57	-5.67	-6.75	-7.20	-5.25	-5.66	-4.44	-5.40	-5.97	-5.51	-6.58	-5.85	-5.59	-6.28	-7.20	-5.91
Ins-32	52	-6.94	-6.57	-6.15	-6.40	-6.55	-6.74	-6.53	-7.00	-6.45	-6.05	-6.16	-4.81	-5.12	-6.46	-7.00	-6.28
Ins-33	53	-5.70	-6.34	-6.26	-5.58	-5.10	-5.69	-4.69	-5.94	-6.57	-5.00	-5.29	-5.59	-4.70	-6.25	-6.57	-5.62
Ins-34	60	-4.13	-5.36	-4.73	-5.30	-5.51	-4.77	-4.54	-4.73	-5.06	-3.59	-5.42	-5.09	-4.99	-5.03	-5.51	-4.88
Ins-35	54	-4.63	-4.08	-4.35	-5.11	-4.74	-4.23	-4.16	-4.45	-4.17	-3.10	-4.89	-3.03	-4.24	-4.75	-5.11	-4.28
Ins-36	70	-4.22	-3.18	-4.85	-2.08	-3.80	-4.41	-3.19	-4.14	-3.53	-3.07	-3.11	-3.11	-3.53	-4.08	-4.85	-3.59
Ins-37	55	-4.49	-5.04	-6.01	-6.11	-6.02	-3.47	-5.21	-5.14	-5.35	-5.28	-6.23	-5.00	-5.87	-5.23	-6.23	-5.32
	MIN	-7.37	-7.24	-7.69	-7.59	-7.65	-6.99	-7.46	-7.01	-6.77	-7.98	-7.97	-7.28	-7.25	-7.29	-7.98	-7.40
	AVG	-5.55	-5.55	-5.58	-5.37	-5.43	-5.38	-5.13	-5.15	-5.34	-5.10	-5.54	-4.95	-5.44	-5.47	-5.58	-5.35

AFTER SIMULATED ANNEALING																		
q-scores		1APH	1B2F	1B9G	1BON	1EFE	1HLS	1IGL	1IMX	1MHI	1MHJ	3GF1	3LRI	6RLX	9INS	MINIMUM	AVG
	Res.	bInsulin	pInsulin	hIGF-1	Bombyxin-II	hminiproins	hInsulin	hIGF-2	hIGF-1	hInsulin	hInsulin	hIGF-1	hLR3IGF-1	hRelaxin	pInsulin		
Daf-28	54	-1.37	-1.19	-1.49	-1.18	-1.62	-1.39	-1.35	-1.25	-1.18	-1.43	-1.48	-1.32	-1.26	-1.47	-1.62	-1.36
Ins-01	80	-0.97	-0.95	-0.92	-0.81	-0.89	-1.03	-0.94	-0.83	-0.89	-0.93	-0.86	-0.70	-1.07	-1.05	-1.07	-0.92
Ins-02	53	-1.53	-1.53	-1.49	-1.22	-1.31	-1.48	-1.56	-1.36	-1.47	-1.20	-1.39	-1.13	-1.29	-1.49	-1.56	-1.39
Ins-03	51	-1.62	-1.53	-1.41	-1.57	-1.43	-1.64	-1.38	-1.63	-1.72	-1.22	-1.34	-1.51	-1.45	-1.34	-1.72	-1.48
Ins-04	52	-1.38	-1.60	-1.40	-1.39	-1.45	-1.61	-1.51	-1.27	-1.58	-1.30	-1.65	-1.39	-1.58	-1.53	-1.65	-1.47
Ins-05	52	-1.65	-1.52	-1.48	-1.69	-1.51	-1.31	-1.84	-1.44	-1.19	-1.55	-1.56	-1.84	-1.66	-1.41	-1.84	-1.55
Ins-06	50	-1.88	-1.78	-1.97	-1.70	-1.84	-1.67	-1.91	-1.79	-1.70	-1.53	-2.04	-1.65	-1.85	-1.68	-2.04	-1.78
Ins-07	50	-1.33	-1.43	-1.30	-1.81	-1.41	-1.79	-1.60	-1.45	-1.37	-1.66	-1.63	-1.17	-1.72	-1.55	-1.81	-1.52
Ins-08	50	-1.69	-1.63	-1.66	-1.68	-1.55	-1.66	-1.51	-1.52	-1.47	-1.63	-1.67	-1.07	-1.77	-1.74	-1.77	-1.59
Ins-09	57	-1.31	-1.31	-1.51	-1.10	-1.17	-1.40	-1.17	-1.32	-1.27	-1.25	-1.31	-1.02	-1.45	-1.29	-1.51	-1.28
Ins-10	55	-1.51	-1.06	-1.18	-1.01	-1.17	-1.13	-1.14	-1.32	-0.98	-1.01	-1.15	-0.97	-1.31	-1.00	-1.51	-1.14
Ins-11	53	-1.01	-1.37	-1.15	-1.41	-1.05	-1.11	-0.83	-0.79	-1.13	-1.13	-1.18	-0.84	-1.29	-1.07	-1.41	-1.10
Ins-12	53	-1.33	-1.33	-1.48	-1.35	-1.25	-1.16	-1.20	-1.01	-1.31	-1.35	-1.25	-0.94	-1.21	-1.13	-1.48	-1.24
Ins-13	48	-1.51	-1.45	-1.59	-1.10	-1.38	-1.31	-1.27	-1.29	-1.27	-1.45	-1.39	-1.33	-1.57	-1.22	-1.59	-1.37
Ins-14	53	-1.17	-1.14	-1.38	-1.15	-1.14	-1.12	-0.98	-1.24	-1.25	-1.16	-1.36	-1.19	-1.49	-1.08	-1.49	-1.20
Ins-15	51	-1.28	-1.38	-1.27	-1.24	-1.37	-1.28	-0.88	-1.25	-1.42	-1.43	-1.12	-1.33	-1.53	-1.33	-1.53	-1.29
Ins-16	50	-1.25	-1.25	-1.42	-1.43	-1.20	-1.03	-1.20	-1.01	-1.13	-1.22	-1.19	-0.99	-1.36	-1.07	-1.43	-1.20
Ins-17	67	-1.24	-1.40	-1.04	-1.12	-1.19	-1.54	-1.49	-1.29	-1.25	-1.00	-1.23	-1.26	-1.51	-1.46	-1.54	-1.29
Ins-18	68	-1.04	-1.00	-1.10	-0.80	-1.05	-0.87	-1.05	-0.80	-0.92	-0.98	-1.13	-0.86	-1.22	-0.93	-1.22	-0.98
Ins-19	57	-1.31	-1.21	-1.18	-1.25	-1.25	-1.36	-1.03	-1.28	-1.50	-1.32	-1.34	-1.26	-1.04	-1.19	-1.50	-1.25
Ins-20	50	-1.63	-1.65	-1.37	-1.34	-1.96	-1.55	-1.53	-1.36	-1.39	-1.61	-1.73	-1.27	-1.55	-1.61	-1.96	-1.54
Ins-21	63	-1.25	-1.47	-1.35	-1.10	-1.18	-1.29	-1.18	-0.94	-1.13	-1.40	-1.16	-1.22	-1.12	-1.47	-1.47	-1.23
Ins-22	57	-1.62	-1.37	-1.43	-1.60	-1.73	-1.47	-1.36	-1.51	-1.06	-1.11	-1.43	-1.24	-1.67	-1.43	-1.73	-1.43
Ins-23	54	-1.57	-1.44	-1.54	-1.70	-1.49	-1.43	-1.36	-1.37	-1.46	-1.32	-1.19	-1.26	-1.20	-1.52	-1.70	-1.42
Ins-24	53	-1.62	-1.45	-1.64	-1.45	-1.22	-1.46	-1.07	-1.55	-1.39	-1.21	-1.43	-1.65	-0.98	-1.28	-1.65	-1.39
Ins-25	52	-1.56	-1.58	-1.13	-1.69	-1.47	-1.53	-1.80	-1.44	-1.53	-1.79	-1.82	-1.48	-1.26	-1.53	-1.82	-1.54
Ins-26	53	-1.33	-1.38	-1.38	-1.17	-1.52	-1.35	-1.24	-1.42	-1.28	-1.00	-1.28	-1.41	-1.15	-1.29	-1.52	-1.30
Ins-27	53	-1.22	-1.06	-1.43	-1.22	-1.42	-1.09	-1.30	-1.19	-1.39	-0.96	-1.42	-1.23	-1.25	-1.34	-1.43	-1.25
Ins-28	49	-1.63	-1.68	-1.59	-1.24	-1.19	-1.55	-1.08	-1.41	-1.58	-2.05	-1.64	-1.59	-1.62	-1.87	-2.05	-1.55
Ins-29	54	-1.50	-1.81	-1.50	-1.90	-1.68	-1.61	-1.77	-1.26	-1.56	-1.51	-1.99	-1.58	-1.68	-1.56	-1.99	-1.64
Ins-30	53	-1.20	-1.40	-1.45	-1.49	-1.42	-1.37	-1.12	-1.33	-1.56	-1.15	-1.18	-1.39	-1.59	-1.29	-1.59	-1.35
Ins-31a	56	-1.44	-1.43	-1.21	-0.92	-1.35	-1.12	-1.33	-1.19	-1.40	-1.12	-1.26	-1.13	-1.07	-1.39	-1.44	-1.24
Ins-31b	57	-1.50	-1.73	-1.70	-1.57	-1.31	-1.47	-1.20	-1.35	-1.52	-1.30	-1.32	-1.33	-1.25	-1.63	-1.73	-1.44
Ins-31c	57	-1.63	-1.40	-1.67	-1.78	-1.30	-1.40	-1.10	-1.34	-1.48	-1.36	-1.63	-1.45	-1.38	-1.55	-1.78	-1.46
Ins-32	52	-1.76	-1.66	-1.56	-1.62	-1.66	-1.71	-1.65	-1.77	-1.63	-1.53	-1.56	-1.22	-1.30	-1.63	-1.77	-1.59
Ins-33	53	-1.44	-1.60	-1.58	-1.41	-1.28	-1.43	-1.18	-1.50	-1.65	-1.26	-1.33	-1.41	-1.18	-1.57	-1.65	-1.42
Ins-34	60	-1.01	-1.31	-1.16	-1.29	-1.35	-1.17	-1.11	-1.16	-1.24	-0.88	-1.32	-1.24	-1.22	-1.23	-1.35	-1.19
Ins-35	54	-1.16	-1.02	-1.09	-1.28	-1.19	-1.06	-1.04	-1.12	-1.05	-0.78	-1.23	-0.76	-1.06	-1.19	-1.28	-1.07
Ins-36	70	-0.99	-0.75	-1.14	-0.49	-0.89	-1.04	-0.75	-0.97	-0.83	-0.72	-0.73	-0.73	-0.83	-0.96	-1.14	-0.85
Ins-37	55	-1.12	-1.26	-1.50	-1.52	-1.50	-0.87	-1.30	-1.28	-1.34	-1.32	-1.55	-1.25	-1.46	-1.31	-1.55	-1.33
	MIN	-1.88	-1.81	-1.97	-1.90	-1.96	-1.79	-1.91	-1.79	-1.72	-2.05	-2.04	-1.84	-1.85	-1.87	-1.60	
	AVG	-1.39	-1.39	-1.40	-1.34	-1.36	-1.35	-1.28	-1.29	-1.34	-1.28	-1.39	-1.24	-1.36	-1.37		

E. Free Energy calculated by NAMD using the CHARM22 forcefield after simulated annealing

Free Energy	1APH	1B2F	1B9G	1BON	1EFE	1HLS	1IGL	1IMX	1MHI	1MHJ	3GF1	3LRI	6RLX	9INS			
	bInsulin	pInsulin	hIGF-1	Bombyxin-II	hminiproins	hInsulin	hIGF-2	hIGF-1	hInsulin	hInsulin	hIGF-1	hLR3IGF-1	hRelaxin	pInsulin		Minimum	Average
Daf-28	-540.85	-561.19	-493.59	-549.63	-533.54	-668.21	-565.70	-667.80	-562.40	-605.44	-641.76	-535.70	-631.79	-633.72		-493.59	-668.21
Ins-01	-1202.99	-1337.83	-1361.04	-1408.60	-1320.55	-1278.81	-1357.35	-1371.71	-1281.41	-1230.40	-1488.93	-1175.36	-1413.07	-1417.58		-1175.36	-1488.93
Ins-02	-993.45	-860.07	-1039.95	-1083.33	-858.88	-993.21	-960.58	-971.43	-1010.43	-1107.93	-977.42	-1011.90	-1099.67	-854.25		-854.25	-1107.93
Ins-03	-315.18	-262.53	-169.02	-213.49	-223.18	-199.49	-395.76	-266.69	-172.81	-320.21	-278.11	-114.21	-122.00	-267.74		-114.21	-395.76
Ins-04	-281.54	-367.32	-336.37	-404.16	-400.87	-336.89	-450.06	-359.36	-420.35	-354.47	-498.97	-411.77	-410.64	-421.65		-281.54	-498.97
Ins-05	-420.26	-290.95	-311.02	-260.94	-141.89	-364.79	-255.08	-375.14	-195.30	-268.08	-388.14	-238.28	-340.92	-359.77		-141.89	-420.26
Ins-06	-708.57	-621.65	-555.01	-760.56	-614.95	-651.40	-643.91	-634.48	-727.62	-620.05	-659.51	-698.30	-795.97	-726.69		-555.01	-795.97
Ins-07	-955.18	-942.15	-969.84	-877.14	-1078.01	-935.51	-752.36	-1058.22	-898.45	-912.52	-847.47	-583.59	-977.06	-916.61		-583.59	-1078.01
Ins-08	-423.18	-612.10	-641.24	-482.29	-361.84	-401.16	-495.83	-524.75	-631.79	-633.45	-607.73	-335.35	-527.05	-579.04		-335.35	-641.24
Ins-09	-744.80	-775.07	-551.67	-611.52	-753.87	-622.33	-781.20	-704.18	-905.06	-715.74	-755.37	-850.40	-660.85	-823.23		-551.67	-905.06
Ins-10	-67.89	-202.25	-153.94	-221.69	-151.92	-75.30	-122.92	-64.71	-169.98	-234.63	-78.85	-161.10	-248.74	-206.96		-64.71	-248.74
Ins-11	-591.70	-535.02	-693.77	-451.52	-373.92	-536.95	-733.40	-658.21	-606.40	-532.23	-814.10	-711.71	-707.88	-633.09		-373.92	-814.10
Ins-12	-401.18	-379.77	-309.03	-406.66	-270.49	-244.19	-367.76	-204.96	-452.29	-257.06	-436.24	-381.71	-383.70	-480.24		-204.96	-480.24
Ins-13	-602.30	-429.58	-466.83	-419.57	-470.11	-552.38	-430.72	-478.63	-558.03	-578.37	-643.48	-552.85	-561.16	-511.07		-419.57	-643.48
Ins-14	-437.81	-374.00	-256.59	-232.87	-239.82	-268.16	-212.06	-197.05	-376.54	-269.48	-315.35	-240.64	-277.97	-242.96		-197.05	-437.81
Ins-15	-786.29	-694.18	-801.89	-747.55	-563.75	-716.73	-820.85	-660.07	-728.80	-706.17	-832.09	-690.18	-865.71	-814.03		-563.75	-865.71
Ins-16	-521.50	-551.44	-715.57	-430.98	-376.14	-603.62	-489.40	-362.75	-705.21	-424.19	-558.83	-624.30	-477.27	-429.40		-362.75	-715.57
Ins-17	-1064.56	-1024.20	-1114.82	-1035.78	-1050.64	-1176.94	-1297.40	-1124.48	-1110.26	-1257.92	-1065.31	-1032.79	-1241.15	-1088.08		-1024.20	-1297.40
Ins-18	-1235.67	-1306.86	-1142.52	-1195.82	-1185.06	-1253.69	-1376.43	-1306.97	-1165.04	-1046.77	-1539.11	-1153.25	-1211.31	-1220.52		-1046.77	-1539.11
Ins-19	-873.31	-688.18	-952.23	-959.76	-842.44	-756.64	-876.34	-955.02	-886.57	-666.94	-958.15	-1092.24	-923.00	-674.74		-666.94	-1092.24
Ins-20	-416.18	-591.99	-496.73	-534.29	-436.27	-611.19	-651.58	-376.96	-518.84	-443.28	-632.08	-478.99	-860.87	-514.88		-376.96	-860.87
Ins-21	151.75	12.64	129.17	147.34	109.46	48.42	106.23	99.59	-309.48	-47.78	56.21	-76.63	-259.69	48.96		151.75	-309.48
Ins-22	-192.10	-268.02	-180.57	-261.24	-183.42	-66.83	-304.37	-161.06	-363.32	-303.96	-435.21	-114.05	-210.59	-228.30		-66.83	-435.21
Ins-23	-266.78	-303.85	-189.98	-276.48	-238.56	-313.27	-209.72	-219.79	-371.47	-210.13	-168.30	-119.47	-292.77	-373.16		-119.47	-373.16
Ins-24	-95.70	-78.85	-26.91	-75.44	-204.25	-52.33	-73.12	-99.88	-114.59	-61.39	-24.98	92.67	-88.25	-82.04		92.67	-204.25
Ins-25	-858.78	-822.07	-757.01	-771.77	-577.35	-646.42	-762.12	-741.77	-786.11	-823.48	-772.92	-751.28	-608.11	-703.32		-577.35	-858.78
Ins-26	-739.10	-687.90	-727.28	-699.92	-707.57	-710.59	-546.12	-640.06	-682.49	-589.65	-748.23	-571.58	-652.50	-646.45		-546.12	-748.23
Ins-27	-435.58	-402.54	-490.64	-243.52	-394.46	-309.33	-380.77	-239.76	-301.79	-505.65	-289.65	-356.10	-321.58	-324.55		-239.76	-505.65
Ins-28	-345.92	-363.59	-293.08	-318.98	-220.50	-405.71	-524.02	-196.10	-325.15	-401.49	-270.45	-347.63	-380.34	-372.85		-196.10	-524.02
Ins-29	-429.76	-440.72	-314.01	-434.26	-317.69	-418.27	-409.42	-310.22	-377.98	-341.62	-473.19	-368.05	-455.20	-361.01		-310.22	-473.19
Ins-30	-108.53	-55.69	19.75	-76.93	-22.63	-49.64	-145.82	-86.04	-174.85	90.47	-120.20	-215.89	-145.47	-86.32		90.47	-215.89
Ins-31a	-1454.71	-1317.39	-1284.62	-1571.86	-1196.55	-1462.70	-1270.18	-1196.37	-1523.44	-1502.19	-1464.80	-1174.47	-1520.31	-1372.21		-1174.47	-1571.86
Ins-31b	-948.07	-910.20	-847.07	-1021.39	-829.97	-874.81	-946.80	-762.35	-1023.16	-770.77	-900.36	-741.31	-1001.62	-1128.26		-741.31	-1128.26
Ins-31c	-773.17	-707.81	-469.69	-594.24	-556.71	-598.50	-749.18	-617.46	-769.44	-643.49	-621.26	-497.40	-634.74	-535.46		-469.69	-773.17
Ins-32	-994.44	-1005.50	-891.58	-1052.54	-673.91	-934.18	-955.66	-867.11	-874.85	-662.14	-1137.72	-717.38	-1064.33	-910.67		-662.14	-1137.72
Ins-33	-399.53	-418.81	-381.23	-446.69	-333.97	-242.13	-572.57	-565.28	-278.78	-345.88	-570.22	-261.39	-381.76	-332.45		-242.13	-572.57
Ins-34	-1156.35	-1024.83	-941.82	-1008.53	-982.69	-990.63	-1160.25	-1024.10	-911.42	-951.78	-1218.89	-1039.71	-1126.15	-1180.94		-911.42	-1218.89
Ins-35	-456.14	-485.98	-495.40	-448.20	-453.85	-444.37	-418.81	-409.84	-420.89	-619.81	-545.74	-314.67	-302.18	-563.01		-302.18	-619.81
Ins-36	-1503.12	-1429.41	-1449.51	-1453.56	-1350.17	-1691.72	-1464.18	-1301.31	-1430.87	-1564.23	-1403.21	-1619.33	-1688.42	-1402.32		-1301.31	-1691.72
Ins-37	-557.03	-668.12	-518.32	-636.69	-804.45	-846.66	-699.28	-573.50	-601.60	-561.81	-727.03	-483.98	-426.86	-860.73		-426.86	-860.73
Minimum	-1503.12	-1429.41	-1449.51	-1571.86	-1350.17	-1691.72	-1464.18	-1371.71	-1523.44	-1564.23	-1539.11	-1619.33	-1688.42	-1417.58			
Average	-628.69	-619.67	-591.06	-613.33	-554.68	-606.43	-638.07	-580.90	-643.13	-600.05	-671.33	-568.81	-657.47	-630.78			