INVESTIGATION OF THE MECHANISM OF ACTION OF IGF-II BINDING TO THE IGF-1R AND IR-A AND HOW SIGNALING PATHWAYS ARE PREFERENTIALLY ACTIVATED

By

Andrew Blyth
BBiotech, Sci(Hons)

Thesis
Submitted to Flinders University
for the degree of
Doctor of Philosophy

College of Medicine and Public Health
12 May 2023

ABSTRACT

The insulin like growth factor (IGF) system is essential for regulating metabolic and mitogenic signaling in mammalian cells. This system is particularly important during fetal development and postnatal growth. The IGF system is regulated by three ligands IGF-II, IGF-I and insulin all of which share a high degree of sequence and structural similarity. We are particularly interested in IGF-II which is capable of binding to the type 1 insulin like growth factor (IGF-1R) and insulin receptor isoform A (IR-A) with high affinity. IGF-II expression is commonly upregulated in cancer, as is the IR-A, resulting in cancer cell survival.

While insulin has been heavily studied for its role in diabetes, in comparison IGF-II has been poorly investigated. Sequence and structural similarities between insulin and IGF-II has indirectly advanced our understanding of IGF-II. However, there is a lack of information specific to IGF-II such as unique structural elements that contribute to differences in stability, receptor specificity and signaling outcome.

In mammalian insulins and IGFs there is a high degree of conservation of residues involved in receptor binding. The so-called ‘aromatic FFY/FYF’ motif found in insulin and the IGFs respectively, is an example of this. Removal of this motif in human insulin leads to essentially no receptor binding affinity. In recent years an influx of research into invertebrate insulins has led to the discovery of several insulins in the venom of cone snails which totally lack this aromatic FYF motif yet are still capable of binding and activating the human insulin receptor with high affinity. Structural and biochemical analysis of these cone snail venom insulins has revealed novel mechanisms of binding to the IR. These mechanisms have since been introduced into human insulin analogues lacking the FYF motif and the resultant analogues have been shown to have unique signaling kinetics compared to human insulin.

The overall aim of my thesis was to better understand residues of IGF-II that are responsible for structural integrity and stability. Also, I aimed to further interrogate the mechanism of IGF-II action through understanding how signaling pathways are preferentially activated. To do this IGF-II mutants lacking elements of the FYF motif were produced to investigate the role of the FYF motif in receptor binding by IGF-II. This is the first time an IGF-II analogue lacking all elements of the FYF motif has been produced. Also introduced into IGF-II were the equivalent mutations found in cone snail venom insulins that are responsible for the alternate mechanism of binding. From these

investigations several residues of IGF-II were identified that are important for protein folding, structure and stability. Furthermore, an IGF-II analogue with increased specificity for the IR-A but no change in signaling potency was characterised. Finally, by combining changes in the FYF motif and the equivalent mutations found in cone snail venom insulins that are responsible for the alternate mechanism of binding I sought to recapitulate the alternate binding mechanism utilised by cone snail insulins for insulin receptor binding. Characterisation of these IGF-II analogues provides valuable information toward developing highly selective small molecule antagonists with low signaling capabilities for the treatment of IGF dependent cancers.

STATEMENT OF ORIGINALITY

I certify that this thesis does not incorporate without acknowledgment any material previously submitted for a degree or diploma in any university; and the research within will not be submitted for any other future degree or diploma without the permission of Flinders University. To the best of my knowledge and belief it does not contain any material previously published or written by another person except where due reference is made in the text.

Signed...  .....

Date.....15 FEB 2023.....

ACKNOWLEDGEMENTS

To Professor Briony Forbes, I am extremely grateful for your supervision and guidance during my PhD. It has been a privilege to work with you over the past several years. Your perspective, advice and mentoring has been invaluable.

I would like to thank Professor Janni Petersen, Associate Professor Amy Reynolds, Dr Richard Van-Wegen for their guidance, encouragement and mentoring throughout my studies.

I am especially grateful for all of the amazing people who work and have studied in the Forbes laboratory. Carlie, you are a fantastic member of the laboratory, your teaching is exceptional, and you are a wealth of knowledge. I am so grateful for all of your help and guidance throughout my time in the lab. Maddie, it has been a delight teaching you and having you as a member of the lab. You make the workplace an especially fun and enjoyable place to work. Allanah, thank you for all of your assistance over the past few years, it has been so great to work with you. Sara, it has been wonderful having you in the lab over the past few years, thank you for your contributions to my project. And to everyone in the lab, thank you for indulging my coffee addiction!

I am extremely grateful to Dr Anya Arthurs for your guidance, mentoring (over many coffees) and support over the past few years.

Angus, thanks for the encouragement to actually start the PhD and consistent heckling throughout! The constant supply of coffee and sourdough bread was especially appreciated.

I would like to thank my partner, Nicole, for your support and understanding during the last year of my studies. You have certainly helped me with the final push to finish!

Thank you to all of my family and friends for their support and encouragement throughout the various stages of the PhD!

I acknowledge the contribution of the Australian Government Research Training Program Scholarship for its contribution to my candidature.

PUBLICATIONS

Chapter 1:

BLYTH, A., KIRK, N. S. & FORBES, B. E. 2020. Understanding IGF-II Action through Insights into Receptor Binding and Activation. Cells, 9, 2276.

Chapter 5:

BLYTH, A., ORTIZ, M., MERRIMAN, A., DELAINE, C. & FORBES, B. 2022. Determinants of IGF-II influencing stability, receptor binding and activation. Sci Rep, 12, 4695.

Other:

FORBES, B. E., **BLYTH, A.** & WIT, J. M. 2020. Disorders of IGFs and IGF-1R signaling pathways. Mol Cell Endocrinol, 518, 111035.

MITCHELL, M. L., HOSSAIN, M. A., LIN, F., PINHEIRO-JUNIOR, E. L., PEIGNEUR, S., WAI, D. C. C., DELAINE, C., **BLYTH, A. J**., FORBES, B. E., TYTGAT, J., WADE, J. D. & NORTON, R. S. 2021. Identification, Synthesis, Conformation and Activity of an Insulin-like Peptide from a Sea Anemone. Biomolecules, 11.

CONFERENCES & MEETINGS

A. Blyth, C. Delaine, A. Merriman, Michael Ortiz, Briony Forbes (2022). Determinants of IGF-II influencing stability and receptor binding and activation. *Australian peptides conference* – Queensland, Australia.

A. Blyth, C. Delaine, A. Merriman and B. E. Forbes (2019). Recombinant expression of Insulin-like growth factor-II analogues for investigation of receptor binding mechanisms - *Lorne proteins conference*. Lorne, Australia.

A. Blyth, C. Delaine and B. E. Forbes (2019). Investigation of alternate receptor binding mechanisms of the Insulin-like growth factor II*. Flinders University Cancer research day* - Adelaide, South Australia.

A. Blyth, C. Delaine and B. E. Forbes (2019). Investigation of alternate receptor binding mechanisms of the Insulin-like growth factor II. *Flinders University College of Medicine and Public Health emerging leaders showcase* – Adelaide, South Australia.

AWARDS & FUNDING

2019-2022

Australian Research Training Program Scholarship.

2019

Best poster award at the 2019 Flinders University Cancer research day - Adelaide, South Australia.

2022

Best poster award at the 2022 Australian peptides conference – Queensland, Australia.

College of Medicine and Public Health student conference travel grant.

ABBREVIATIONS

αCT	α-chain C-terminal region
Amp	Ampicillin
ATP	*Adenosine triphosphate*
BSA	Bovine serum albumin
Con-Ins G1	*Conus geographus* venom insulin 1
Con-Ins K1	*Conus kinoshitai* venom insulin 1
Con-Ins K2	*Conus kinoshitai* venom insulin 2
CR	Cysteine-rich domain
CryoEM	Cryogenic electron microscopy
CT	C-terminal domain
DMEM	Dulbecco's modified Eagle's medium high-glucose
DNA	Deoxyribose nucleic acid
DOI	Des octapeptide (B23–B30)-insulin
DTT	Dithiothreitol
E. coli	*Escherichia coli*
EDTA	Ethelene diamine tetra-acetate
Eu	Europium
FnIII-1	First fibronectin type-III domain
FnIII-2	Second fibronectin type-III domain
FnIII-3	Third fibronectin type-III domain
FPLC	Fast performance liquid chromatography

hIGF-II	human IGF-II
HPLC	High performance liquid chromatography
IB	Inclusion bodies
ID	Insert domain
IGF	Insulin-like growth factor
IGF-I	Insulin-like growth factor I
IGF-II	Insulin-like growth factor II
IGF-1R	Insulin-like growth factor 1 receptor
IGFBP	Insulin-like growth factor binding protein
IGF-1RZip	leucine-zippered receptor
Ins	Insert domain
IPTG	Isopropyl-β-D-thiogalactopyranoside
IR	insulin receptor
IR-A	Insulin receptor isoform A
IR-B	Insulin receptor isoform B
IRS	Insulin receptor substrate
JM	Juxta-membrane domain
LB	Luria broth
LCDV-1	Lymphocystis disease virus-1
L1	First leucine-rich repeat domain
L2	Second leucine-rich repeat domain
MALDI	Matrix assisted laser desorption ionisation
MAPK	Ras-mitogen-activated protein kinase

MOPS	4-Morpholinepropanesulfonic acid
NIM	Non-inducing medium
ODx	Optical density at x nm
PCR	Polymerase chain reaction
pGH	Porcine growth hormone
PI3K	Phosphoinositide 3-kinase
PKB	Protein kinase B
rpHPLC	Reverse phase high performance liquid chromatography
scLCDV1-VILP	Single chain Lymphocystis disease virus-1 viral insulin like peptide
SD	Standard deviation
sDHB	2,5-dihydroxybenzoic acid
SDS	Sodium dodecyl sulfate
SDS-PAGE	Sodium dodecyl sulfate poly acrylamide gel electrophoresis
S.E.M	Standard error of the mean
SFM	Serum free medium
TEMED	N,N,N,N,N-tetramethylethylene-diamine
TFA	Trifluoro acetic acid
TK	Tyrosine kinase
TM	Transmembrane domain
VILPs	Viral insulin-like peptides
95% CI	95% Confidence interval

TABLE OF CONTENTS

1 INTRODUCTION

The following chapter was published in *Cells:*

BLYTH, A., KIRK, N. S. & FORBES, B. E. 2020. Understanding IGF-II Action through Insights into Receptor Binding and Activation. Cells, 9, 2276.

Author contributions:

This review paper was produced from my PhD literature review. As fist author I was responsible for the overall planning of this review paper and wrote 85% of the published work. The structural biology section was written in collaboration with Dr Nicholas Kirk. Review of this work was performed by myself, Dr Kirk and Professor Forbes.

1.1 Introduction

The insulin-like growth factor (IGF) system controls metabolic and mitogenic responses in mammalian cells and importantly regulates embryonic growth and development as well as adult growth (Denley et al., 2005). The IGF system is regulated by three structurally similar ligands, IGF-I, IGF-II and insulin (Figure 1.1). These ligands act via one or more of the three related receptor tyrosine kinases: the two splice variants of the insulin receptor (IR-A and IR-B) and the type 1 Insulin-like growth factor receptor (IGF-1R). IR-B signaling is responsible for the classic IR metabolic activities. IGF-II is unique in that it can activate both IGF-1R and IR-A to promote cell growth and survival (Belfiore et al., 2017). There is evidence to suggest that IGF-II, IGF-I and insulin can promote both shared and unique signaling outcomes through IGF-1R and IR (Versteyhe et al., 2013, Morcavallo et al., 2011). However, IGF-II specific actions are generally attributed to tissue specific expression (Stylianopoulou et al., 1988, Jiao et al., 2013, Duan et al., 2010). IGF-II and IR-A overexpression has been found in several forms of cancer and has been shown to promote cell growth and survival (Holly et al., 2019, Belfiore et al., 2017). Despite its implication in cancer progression, the molecular mechanisms underlying IGF-II action are the least understood. For these reasons, IGF-II is the focus of this investigation.

IGF-II plays important roles in fetal growth and development, when it is most abundant (White et al., 2018, Randhawa and Cohen, 2005). Notably, IGF-II fetal plasma concentrations are several fold higher than that of IGF-I (Holland et al., 1997, Lee et al., 1993, Gluckman and Butler, 1983). Knockout of *Igf2* leads to a 60% reduction in weight at birth (DeChiara et al., 1990). IGF-II serum concentrations in many mammalian species decline rapidly after birth (Stylianopoulou et al., 1988, Gluckman and Butler, 1983, Soares et al., 1986). Interestingly, in adult mice, IGF-II serum levels are barely detectable, whereas in humans it is the more abundant IGF ligand (Blackburn et al., 1997, Fielder et al., 1990). In humans, the *IGF2* is maternally imprinted and only expressed from the paternal allele. Gain of methylation at the regulatory H19 locus on the paternal allele causes under-expression of IGF-II and results in undergrowth syndromes (Russell-Silver syndrome), which can include a variety of phenotypes including prenatal growth deficiency, facial dysmorphic features and developmental delay (Smith et al., 2007, Gicquel and Le Bouc, 2006). Alternatively, overexpression of IGF-II can produce an overgrowth syndrome (Beckwith–Wiedemann syndrome), which can include macroglossia, macrosomia, and abdominal wall defects (Smith et al., 2007, Gicquel and Le Bouc, 2006).

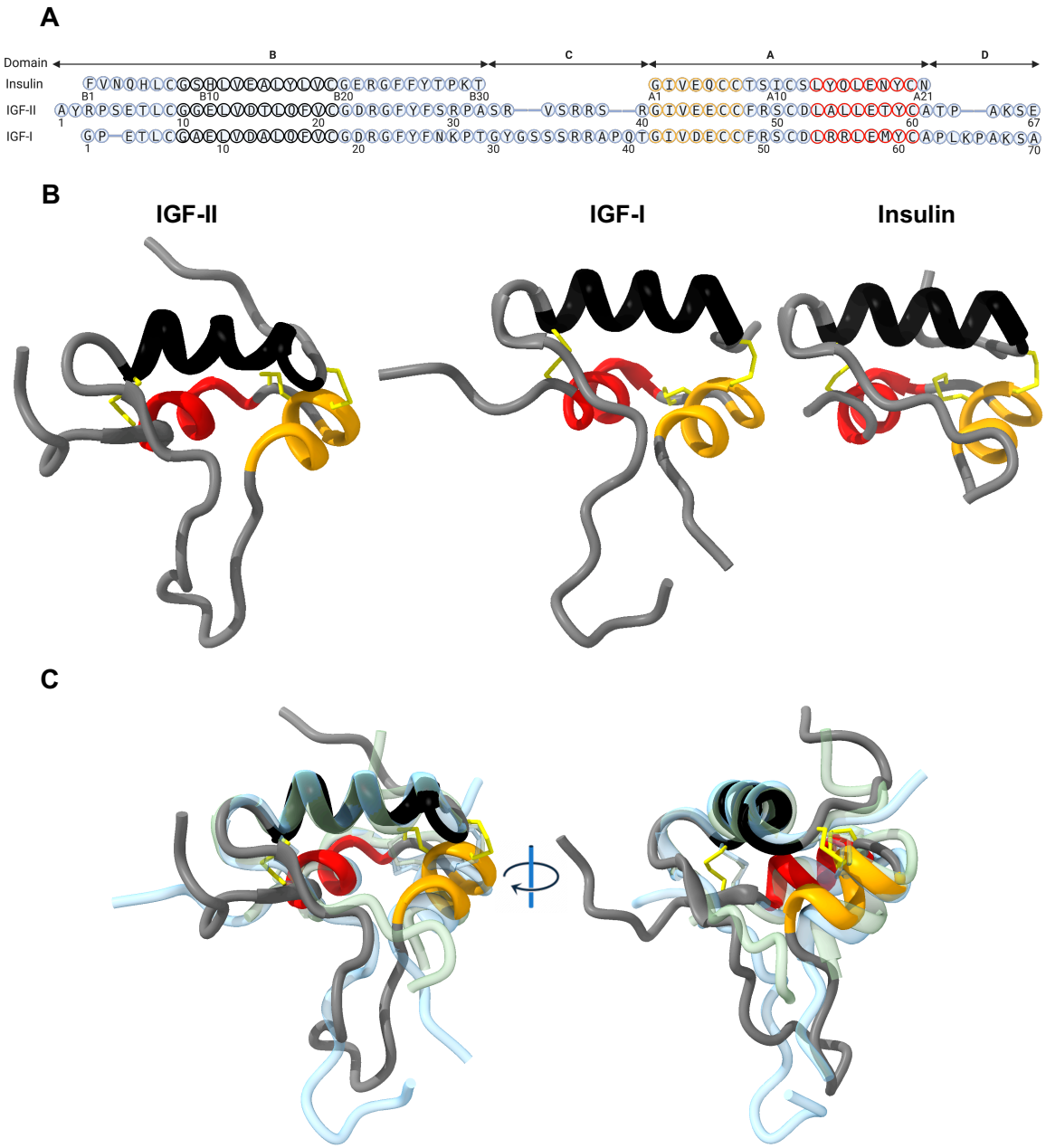


Figure 1.1: Sequence alignment and structural comparison of the IGFs and insulin.

(A) Sequence alignment of insulin, IGF-II and IGF-I. Each peptide has three helices; B-domain helix 1 (black circles), A-domain helix 2 (orange circles) and A-domain helix 3 (red circles). Residue numbers are indicated below each sequence. (B) IGF-II, IGF-I and insulin structures (PDB: 1IGL, 1GZR and 3I3Z respectively). Helix colouring as in A and disulfide bonds shown in yellow. (C) Structural alignment of IGF-II, IGF-I and insulin. IGF-II colouring is the same as in B, IGF-I is coloured in transparent light blue and insulin in transparent light green. The IGFs and insulin are highly similar in structure in their B- and A-domains. The C- and D-domains are unique to the IGFs. There is a lack of side chain density for IGF-I residues 36-38 in the C-domain.

At the tissue level, IGF-II promotes cell growth and survival. It regulates bone growth by promoting proper timing of chondrocyte maturation and perichondrial cell differentiation and survival (Yakar et al., 2018). Overexpression of *Igf2* in smooth muscle and pancreatic beta cells results in the development of cardiovascular defects and type 2 diabetes (Devedjian et al., 2000, Zaina and Nilsson, 2003). Conversely, knockout of placental *Igf2* leads to reduced placental growth and fetal growth restriction (Constancia et al., 2002). IGF-II is most abundant in the fetal and adult human brain, primarily produced by the choroid plexus but also the leptomeninges and endothelial cells (Ferron et al., 2015, Charalambous et al., 2004, Feil et al., 1994, DeChiara et al., 1991, Stylianopoulou et al., 1988). IGF-II has been identified in cerebral spinal fluid and has been found to promote neurogenesis in the subventricular and subgranular zone of the adult brain (Ziegler et al., 2012, Bracko et al., 2012, Lehtinen et al., 2011). Several investigations have also identified that IGF-II promotes stem cell self-renewal through activation of IR-A. For example, IGF-II:IR-A signaling supports neural stem cell maintenance and the expansion of neural progenitor cells (Ziegler et al., 2015). This role in stem cell self-renewal extends to other tissues, as identified using stem cell specific knockout of *Igf2* in young adults in which growth of intestinal stem cells is also inhibited (Ziegler et al., 2019).

IGF-II action is highly regulated by its interaction with soluble IGF binding proteins (IGFBPs), including IGF-II specific IGFBP-6. IGFBPs retain IGF-II in circulation and deliver it to target tissues (Forbes et al., 2012). In addition, the type 2 IGF receptor (IGF-2R, also called cation-independent mannose-6-phosphate receptor) is responsible for the control of circulating IGF-II levels, by binding to IGF-II with high affinity and targeting it for lysosomal degradation (Brown et al., 2008, Brown et al., 2009).

1.1.1 IGF-II and Cancer

It is well established that abnormal function of the IGF system promotes growth and metastasis of the 3 most commonly diagnosed cancers: breast, prostate and colorectal (Zha and Lackner, 2010, Belfiore and Malaguarnera, 2011, Livingstone, 2013). It also promotes growth and survival of brain, thyroid and ovarian cancers among others (Belfiore et al., 2017, Holly et al., 2019). Specifically, dysregulation of IGF-II expression has been associated with cancer progression (Belfiore et al., 2017). IGF-II expression is often upregulated in these cancers (Belfiore and Malaguarnera, 2011, Pollak, 2008, Livingstone, 2013) and often results in both autocrine and paracrine effects (Samani et al., 2007). For example, in the MDA-MB-157 breast cancer cell line, autocrine production of IGF-II stimulates cell growth though IR-A activation while expression in stromal and epithelial tissue of breast cancer specimens acts in both autocrine and paracrine

manners (Sciacca et al., 1999). Loss of imprinted IGF-II expression has been documented in many forms of cancer, leading to increased levels of intratumoural IGF-II, thereby promoting cell growth and tumorigenesis (Chao and D'Amore, 2008, Harris and Westwood, 2012, Livingstone, 2013). Interestingly, the mechanism by which loss of imprinting occurs has recently been investigated and found to involve overexpression of an intronic miRNA (miR-483-5p) found within the *IGF2* gene (Lui and Baron, 2013). miR-483-5p increases IGF-II transcription at the fetal promoter (Lui and Baron, 2013).

In cancer, IGF-II can act via IGF-1R and/or IR-A and these autocrine/paracrine signaling loops are regularly observed (Arcaro, 2013). IGF-1R, which promotes cell growth and survival, is also commonly upregulated in cancers such as breast, colorectal and prostate cancer (Arcaro, 2013, Pollak, 2008, Gallagher and LeRoith, 2010). In contrast to IR-B that signals through metabolic pathways, IR-A has mitogenic signaling capabilities that are important during development when IR-A is most abundantly expressed (Belfiore and Malaguarnera, 2011). IR-A is only expressed at very low levels in most adult cells (Denley et al., 2003). However, in malignant cells, including breast, thyroid, colon and prostate cancer, IR is over expressed with IR-A being the predominant isoform (Giorgino et al., 1991, Belfiore and Malaguarnera, 2011). IGF-II:IR-A signaling also supports maintenance of tumour stem and progenitor cells (Vella et al., 2019, Tominaga et al., 2017). Concomitant upregulation of both IGF-II and IR-A signaling thus provides cancer cells and tumour stem cells with an additional growth and survival mechanism (Belfiore et al., 2017).

1.1.2 IGF-II Signaling

The biological processes that IGF-II promotes result from activation of signaling pathways through its binding to the extracellular region of IR-A or IGF-1R. The overall mechanisms of binding of IGF-II, IGF-I and insulin to IGF-1R and IR are conserved. Receptor binding results in structural rearrangement of the receptor (further discussed below) causing autophosphorylation of the tyrosine kinase (TK) domains on the intracellular region of the receptor (Hubbard, 2004, Li et al., 2003, Favelyukis et al., 2001). Extensive studies conducted by Cabail et al (2015) have determined that in the unbound state, each monomer is autoinhibited by self-interaction of the activation loop within its TK active site, thereby precluding the binding of adenosine triphosphate (ATP). Upon ligand binding, structural rearrangement occurs allowing the juxtamembrane (JM) domain of one monomer to interact with the TK domain of the opposite monomer. This releases the autoinhibitory state and allows for the binding of ATP and subsequent substrate phosphorylation.

The first signaling step upon IGF-II, IGF-I, and insulin binding to their cognate receptors involves phosphorylation of three tyrosine residues within the activation loop of the TK domain (IGF-1R: Y1131, Y1135, and Y1136 and IR: Y1158, Y1162 and Y1163) (Laviola et al., 2007, Hubbard et al., 1994). Subsequently, residue Y950 (IGF-1R) or its equivalent Y960 (IR) is phosphorylated (De Meyts and Whittaker, 2002). This creates a docking site for IR substrates (IRS) and Shc (Figure 1.2), which are then phosphorylated (Hanke and Mann, 2009). Subsequent to receptor activation, two main signaling pathways are activated, the phosphoinositide 3-kinase (PI3K)-Akt/protein kinase B (PKB) pathway, responsible for metabolic responses and the Ras-mitogen-activated protein kinase (MAPK) pathway, resulting in mitogenic responses (cell growth, differentiation, and gene expression) (Taniguchi et al., 2006, Siddle, 2011, Hakuno and Takahashi, 2018).

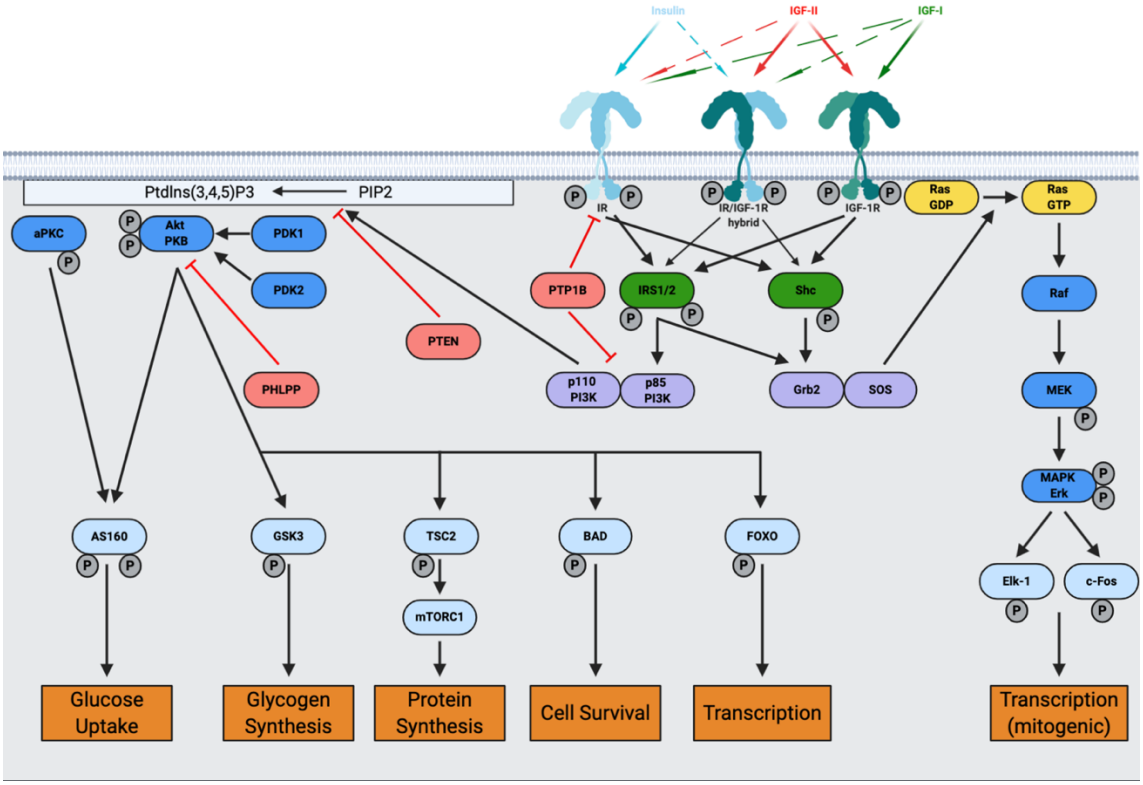


Figure 1.2: The Insulin and IGF system.

Insulin, IGF-I and IGF-II bind with different affinities to the IR-B, IR-A and IGF-1R (indicated by thickness of arrows). IGF-II binds with high affinity to both the IGF-1R and the IR-A and with low affinity to the IR-B. Upon receptor binding a structural change leads to activation of the intracellular tyrosine kinase domain and autophosphorylation (indicated by P). IRS1/2 and Shc adapter proteins are recruited, and two main signaling pathways are activated: the Akt/PKB and the Ras/MAPK pathways. Metabolic and mitogenic activities are promoted respectively (Adapted from: Siddle 2011).

1.1.3 How does IGF-II Bind and Activate IGF-1R and IR-A?

In order to understand how IGF-II promotes normal cell growth and survival and to develop ways to inhibit its action in cancer, a detailed knowledge of the molecular mechanisms underlying IGF-II receptor binding and activation is required. Our understanding so far has largely been derived through site-directed mutagenesis and comparative structural studies, with a recent cryogenic electron microscopy (cryoEM) study revealing the structure of IGF-II bound to IGF-1R (Xu et al., 2020). The details of our current understanding will now follow.

1.1.3.1 IGF-II Structure

IGF-II is a 67 amino acid single chain polypeptide with sequence and structural similarity to IGF-I (70 amino acids) and insulin (51 amino acid two-chain peptide) (Figure 1.1). Sequence alignments of the IGFs and insulin (Figure 1.1A) reveal 50% sequence homology between the B- and A-domains of the IGFs and the equivalent domains of insulin (Denley et al., 2005). The three-dimensional structure of insulin and the IGFs is also similar, as shown using structural overlays in Figure 1.1C. Three disulfide bonds hold together the specific three-dimensional structure, which comprises three α-helices (Figure 1.1B). IGF-I and IGF-II each comprise four domains: B, C, A and D (Torres et al., 1995). Insulin, in contrast, is a two-chained mature protein composed of A and B domains linked by two inter-chain disulfide bonds and one intra-chain disulfide bond within the A-chain (Figure 1.1B) (Smith et al., 2003). Alanine scanning mutagenesis first described two binding surfaces on each ligand, which were defined by a high affinity interaction (site 1, residues that have a >10-fold decrease in affinity when mutated to alanine) and a low affinity interaction (site 2, residues that have a 2-10-fold decrease in affinity when mutated to alanine) (Alvino et al., 2009, Gauguin et al., 2008a, Chen et al., 2000, Kristensen et al., 1997, De Meyts, 2015).

The oxidative folding pathways of insulin and IGF-I have been well characterised but IGF-II folding has been defined to a lesser extent. However, due to the sequence and structural similarities with IGF-II, IGF-I and insulin, it is expected that a similar folding pathway exists for IGF-II . The beginning of insulin and IGF-I folding likely involves the association of highly conserved hydrophobic residues (Weiss, 2009) and nascent interactions of a microdomain surrounding residues CA20-CB19 (C18-C61 for IGF-I). This microdomain orients the thiol groups to facilitate disulfide bond formation (Hua et al., 2006b, Narhi et al., 1993, Yan et al., 2003). Residue FB24 (F23 for IGF-I) of the insulin aromatic FFY motif is involved in this early microdomain formation. Subsequently

the interchain CA7-CB7 (C48-C6 in IGF-I) and the intra A-chain CA6-CA11 (C47-C52 in IGF-I) disulfides are formed (Narhi et al., 1993). The order in which these two bonds form, in both insulin and IGF-I, has been shown to be interchangeable. However, both are required for formation of the biologically active conformation (Hua et al., 1996, Hua et al., 2001). For insulin, several investigations have demonstrated that under oxidative conditions the native disulfide bonds are the predominant form, with essentially no detectable alternate disulfide bonded isoforms (Tang and Tsou, 1990, Katsoyannis and Tometsko, 1966, Qiao et al., 2001, Qiao et al., 2003). Loss of the A7-B7 or A6-A11 disulfide bonds results in significantly decreased stability and binding affinity (Hua et al., 2001, Hua et al., 1996). Disulfide bond isomers of insulin, A7-A11, A6-B7 insulin and A7-A7, A11-B7 insulin were shown to have essentially the same affinity as native insulin (Sieber et al., 1978).

Interestingly, under *in vitro* oxidative folding conditions, an alternate disulfide pairing of IGF-I (C18-C61, C6-C47, C48-C52) has been shown to occur in equilibrium with the native form of IGF-I (Hober et al., 1997, Hober et al., 1992, Hejnaes et al., 1992, Miller et al., 1993). This isoform has been extensively characterised and found to have a lower helical content compared to native IGF-I (Hober et al., 1992). In addition, the alternate disulfide pairing of IGF-I leads to a 200-fold reduction in affinity for the IGF-1R (Milner et al., 1995). Optimised refolding conditions have since been developed to favour the predominant formation of the native IGF-I structure using recombinant expression and subsequent chemical refolding techniques (King et al., 1992). As mentioned, there is considerably less known about IGF-II folding but there are reports of alternate disulfide bonded IGF-II isomers equivalent to the alternate disulfide pairing of IGF-I described above (Francis et al., 1993). The alternatively disulfide bonded IGF-II isomer has a 160-fold lower affinity compared to native IGF-II (Smith et al., 1989). Despite little being reported for native IGF-II folding, a similar recombinant folding method has been employed for IGF-II as used for IGF-I (Lien et al., 2001, Delaine et al., 2007).

1.1.3.2 Receptor Structure

IR-A, IR-B and IGF-1R are similar in amino acid sequence and structure and therefore comprise the same domains (Figure 1.3A): The first and second leucine-rich repeat domains (L1 and L2), cysteine-rich domain (CR), first, second and third fibronectin type-III domains (FnIII-1, 2, and 3), insert domain (ID), α-chain C-terminal region (αCT), transmembrane domain (TM), juxta-membrane domain, (JM), insert domain (Ins), tyrosine kinase (TK); C-terminal domain (CT) (schematic diagram shown in Figure 1.3A) (Adams et al., 2000). The two IR isoforms differ by the expression of exon 11, which encodes 12 amino acids that are absent in the IR-A splice variant. The receptors are

disulfide-linked (αβ)2 homodimers and the extracellular domains of each αβ monomer assemble in an anti-parallel, Λ-shaped conformation (Figure 1.3B and C), generating two equivalent ligand binding regions on each half of the receptor dimer. For both receptors, the α-chain C-terminal (αCT) helix of one monomer lies on the L1 surface of the opposing monomer to form site 1a (McKern et al., 2006, Xu et al., 2018). A prime symbol (′) will be used to designate A-domain from the opposite monomer. In the unbound (apo) state, the sites of membrane entry, including the FnIII stalks, are situated far apart thereby holding the intracellular tyrosine kinase domains in an inactive state, where the activation loop within TK domain is *cis*-autoinhibited (Figure 1.4A and B) (McKern et al., 2006, Xu et al., 2018, Hubbard et al., 1994).

1.1.3.3 Mechanism of receptor Binding and Activation

Upon ligand binding, the IGF-1R and IR undergo extensive structural change, whereby the αCT′ shifts to accommodate the ligand, which makes contact via its site 1 residues (Lou et al., 2006, Garrett et al., 1998). The FnIII stalks come close together, permitting dimerization of the intracellular tyrosine kinase domains to release the *cis*-autoinhibition of the TK domains (Figure 1.4C and D). Notably, such a conformation, as predicted by Kavran et al (2014), is essential for receptor activation. The dimerization of the TK domains of each monomer involves the JM region of one monomer interacting with the N-terminal lobe of the opposite monomer (Figure 1.4C and D), repositioning of the αC helix within this lobe. The repositioning of the αC helix allows for a salt bridge to form between residues K1030 and E1047 and facilitates ATP binding to residue K1030 and for substrate binding to residue D1132 (Cabail et al., 2015).

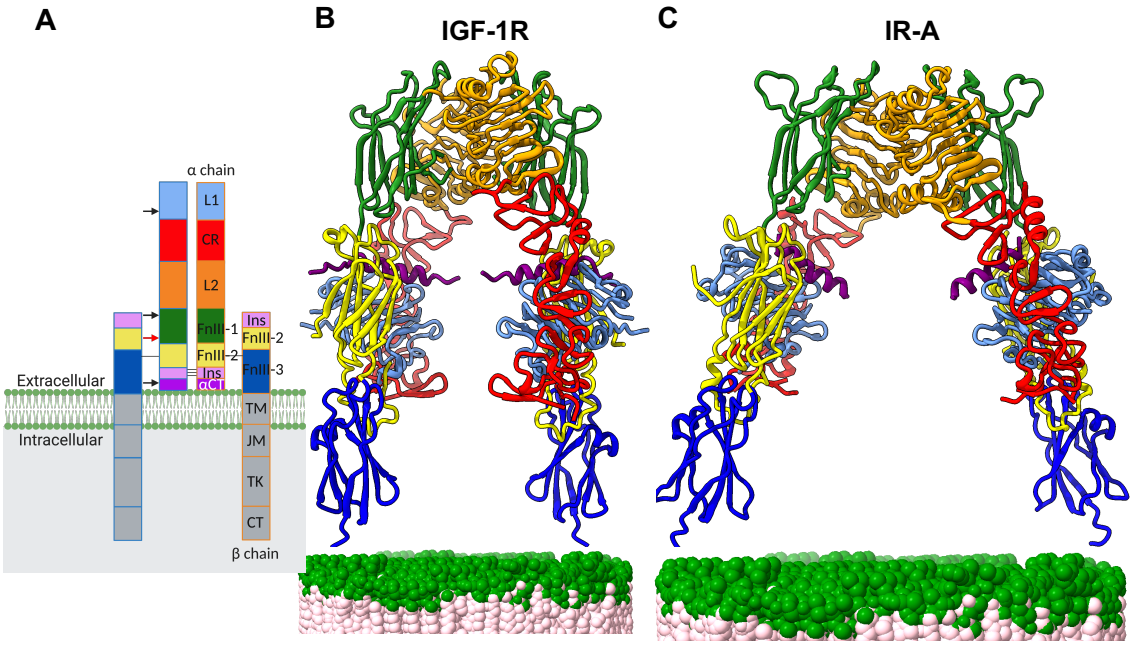


Figure 1.3: Solved structures of the IGF-1R and IR in their unbound state.

(A) Domain structure of the IGF-1R and IR tyrosine kinase receptors. Individual αβ monomers are indicated by blue or orange outline. The IGF-1R and IR have a high degree of sequence homology and therefore comprise the same domains: The first and second leucine-rich repeat domains (L1 and L2), cysteine-rich domain (CR), first, second and third fibronectin type-III domains (FnIII-1, 2, and 3), insert domain (Ins), α-chain C-terminal region (αCT), transmembrane domain (TM), juxta-membrane domain, (JM), Insert domain (Ins), tyrosine kinase (TK); C-terminal domain (CT) (Adams et al 2000). Arrows indicate regions involved in ligand binding (site 1, black; site 2, red). Crystal structure of the IGF-1R (B) and IR-A (C) ectodomain unbound (apo) (PDB: 5U8R and 4ZXB respectively). In the apo state each receptor adopts a similar structure with the FnIII-3 legs of the receptor positioned far apart and forming an open Λ-shape. The distance between the FnIII-3 legs is considerably shorter in the IGF-1R (~67 Å) compared to the IR-A (~120 Å).

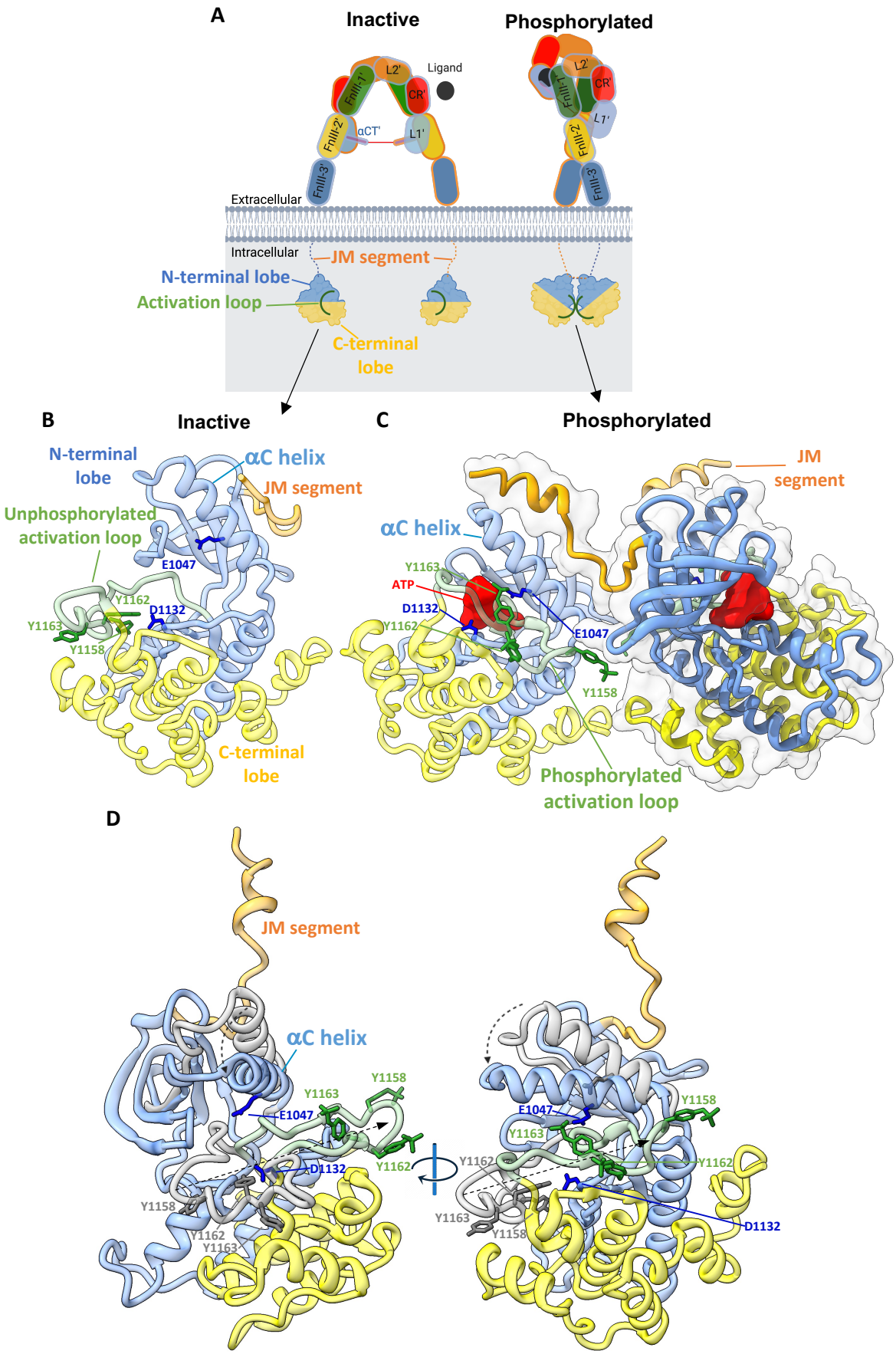

A
Inactive
Phosphorylated
Ligand
L2'
CR'
FnIII-1'
FnIII-2'
FnIII-3'
αCT'
L1'
L1'
Extracellular
Intracellular
JM segment
N-terminal lobe
Activation loop
C-terminal lobe
B
Inactive
N-terminal lobe
αC helix
JM segment
Unphosphorylated activation loop
E1047
Y1162
D1132
Y1163
Y1158
C-terminal lobe
C
Phosphorylated
JM segment
αC helix
Y1163
ATP
D1132
Y1162
E1047
Y1158
Phosphorylated activation loop
D
JM segment
αC helix
Y1158
Y1163
E1047
D1132
Y1162
Y1158
Y1162
Y1163
E1047
Y1158
Y1163
Y1162
Y1162
Y1163
Y1158
D1132

Figure 1.4: The tyrosine kinase domain of the insulin receptor in the inactive and phosphorylated state.

(A) Schematic diagram representing the structural rearrangement of the insulin receptor (IR) between apo state and receptor bound to a single ligand. In the unbound state the receptor FnIII legs are held apart. In the ligand bound state the receptor legs are brought together and cross over. (B) One kinase domain of the inactive insulin receptor (IR) dimer is shown (PDB:1IRK). N-terminal lobes of the TK domains, blue ribbon; C-terminal lobes of the TK domains, yellow ribbon; phosphorylated activation loops, green ribbon; upstream JM regions, orange ribbons. In the inactive state the insulin receptor (IR) activation loop within the TK domain is autoinhibited and residue Y1162 contacts the active site (D1132). (C) Phosphorylated IR TK domains, ribbon colouring as in (A) with one monomer surface filled in transparent grey (PDB:4XLV). The 'legs together' conformation in the ligand bound state promotes *trans* association of the juxtamembrane (JM) region with the αC helix of the opposing TK domain. The repositioning of the αC helix facilitates the formation of a salt bridge between K1030 and E1047 and therefore allows residue K1030 to bind to adenosine triphosphate (ATP, shown in red). (D) Overlay of one phosphorylated TK domain (coloured as in B and C) with the αC helix and activation loop of the inactive receptor shown in grey to demonstrate the reconfiguration of these regions upon ligand binding.

1.1.3.3.1 Site 1 and site 2 interaction

Site-directed mutagenesis studies originally defined two binding surfaces on the ligand and receptor, a high affinity interaction (site 1) and a low affinity interaction (site 2) (Alvino et al., 2009, Whittaker et al., 2008, Gauguin et al., 2008a, Whittaker et al., 2001, Chen et al., 2000, Kristensen et al., 1997, De Meyts, 2015). Solved receptor bound structures have provided additional detail of the binding interaction. The additional detail from structural studies has resulted in site 1 being further defined as site 1a (high affinity interaction) and site 1b interactions (low affinity interaction), with binding to both contributing to the overall high affinity binding at site 1. Site 1a on both the IR and IGF-1R involves ligand contact with the L1 domain and αCT′, while site 1b involves contact with the FnIII-1′ domain. For IGF-II bound to the IGF-1R the site 1a and site 1b contacts with the receptor are shown in Figure 1.5C and D. An as yet unexplained observation is the limited correlation of the site-directed mutagenesis data for IGF-II site 2 residues (Table 1.1) and their involvement in binding in the IGF-II:IGF-1R complex structure (Xu et al., 2020). A similar conundrum was revealed by the IGF-I:IGF-1R complex structure and the corresponding site-directed mutagenesis data (Li et al., 2019).

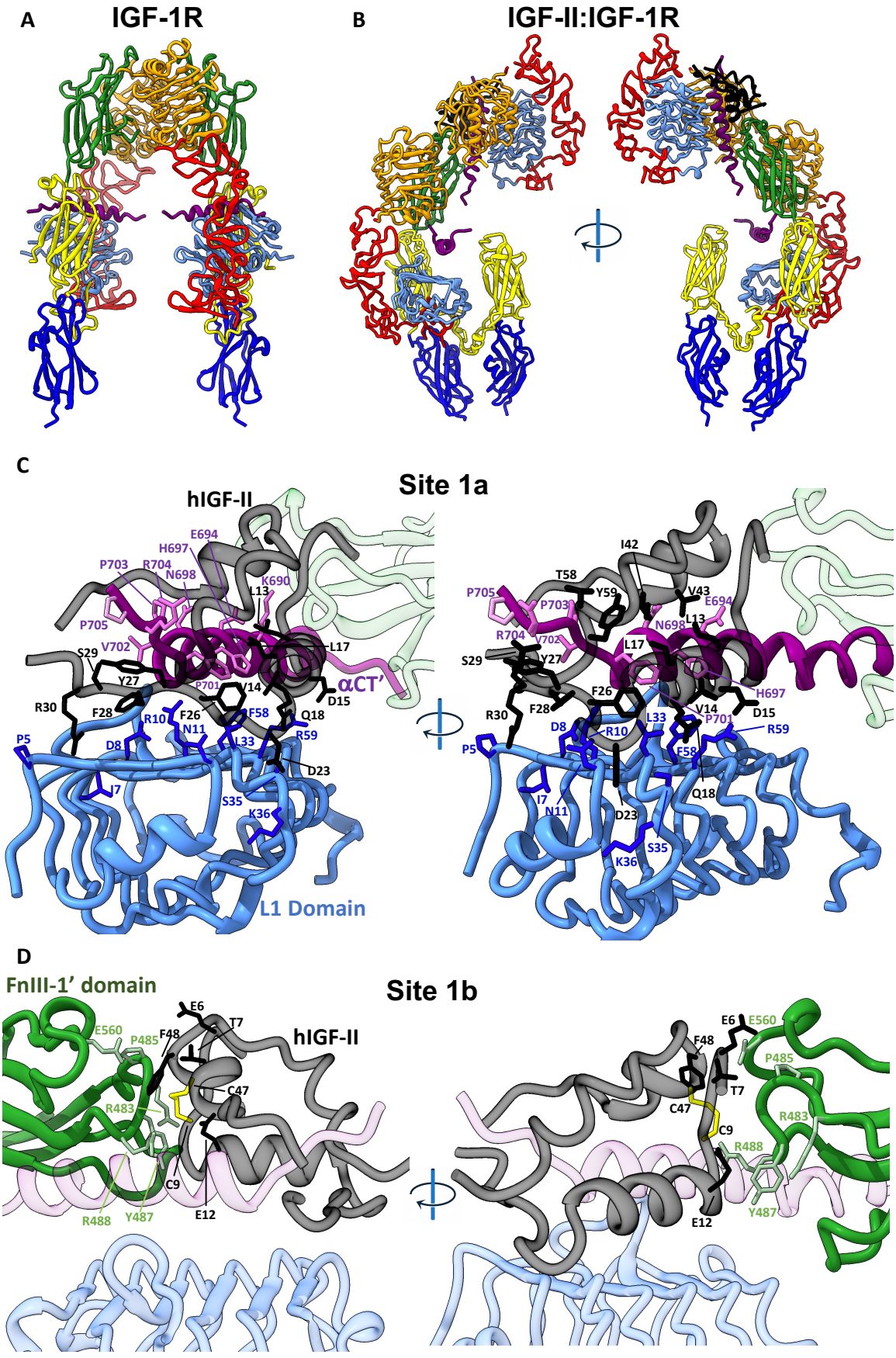

A
IGF-1R
B
IGF-II:IGF-1R
C
Site 1a
hIGF-II
E694
H697
P703
R704
N698
K690
L13
P705
V702
L17
S29
Y27
αCT'
P701
V14
D15
R30
F28
R10
F26
F58
Q18
N11
L33
R59
P5
D8
D23
I7
S35
K36
L1 Domain
I42
T58
Y59
V43
E694
P705
P703
N698
L13
R704
V702
L17
S29
Y27
F26
H697
R30
F28
V14
D15
D8
R10
L33
P701
R59
P5
F58
Q18
I7
N11
D23
S35
K36
D
FnIII-1' domain
Site 1b
E6
T7
E560
P485
F48
hIGF-II
C47
R483
C9
R488
Y487
E12
E6
E560
F48
P485
T7
C47
C9
R483
R488
Y487
E12

Figure 1.5: Structure of the IGF-1R unbound and with a single IGF-II ligand bound.

(A) Crystal structure of the IGF-1R ectodomain unbound (apo) (PDB: 5U8R). (B) Upon IGF-II (black) binding, a major structural rearrangement occurs resulting in a J-shape conformation of the receptor where the FnIII-3 legs are in close proximity (PDB: 6VWI and 6VWJ). The activated conformation (B) is stabilized by the ligand clipping the αCT′ and L1 domains together, interactions through site 1b on FnIII-1′, and potential salt bridges in the head region facilitated by ligand binding (between E687 (αCT′) and R335 (domain L2), between residues E693 (αCT′) and R488 (domain FnIII-1′), and between residues K690 (αCT′) and D489 (domain FnIII-1′)) (Xu et al 2020). (C) Zoom in of the site 1a ligand binding region between IGF-II and IGF1R involving IGF-II B-domain residues; C9, L13, V14, D15, L17, Q18, D23, F26, Y27, F28, S29, and R30 and the side chains of receptor domain L1 residues P5, I7, D8, R10, N11, L33, S35, K36, F58, and R59, and the side chains of receptor αCT′ residues H697, F701, V702, and P705. The IGF-II A-domain contacts the receptor αCT′ segment (and not domain L1), with the interaction mediated by the side chains of IGF-II residues I42, V43, T58, Y59, and T62 and the side chains of receptor αCT' residues K690′, E694′, H697′, N698′, F701′, V702′, P703′, and R704′. (C) Zoom in of the site 1b ligand binding region between IGF-II and IGF1R involving IGF-II residues E6, T7, C9, E12, C47, F48 and FnIII-1′ domain residues R483, P485, Y487, R488 and E560. (A), (B) (C) and (D) colouring scheme is the same as in Figure 1.3.

Table 1.1: Residues of human insulin identified in the site 2′ interaction with the FnIII-1′ domain.

Residues of the insulin and the IGFs proposed to be involved in the site 2′ interaction with the FnIII-1′ domain. Site 2′ residues identified by mutagenesis coloured blue and residues identified by mutagenesis and confirmed by structural investigations are coloured black. Binding site 2′ has been defined for the IGFs and insulin by mutagenesis studies. This interaction has been captured in structural studies for insulin bound to the IR. There are currently no reported structures depicting this interaction for either of the IGFs bound to the IGF-1R. The mutagenesis data for the IGFs suggests that a similar site 2′ interaction does occur.

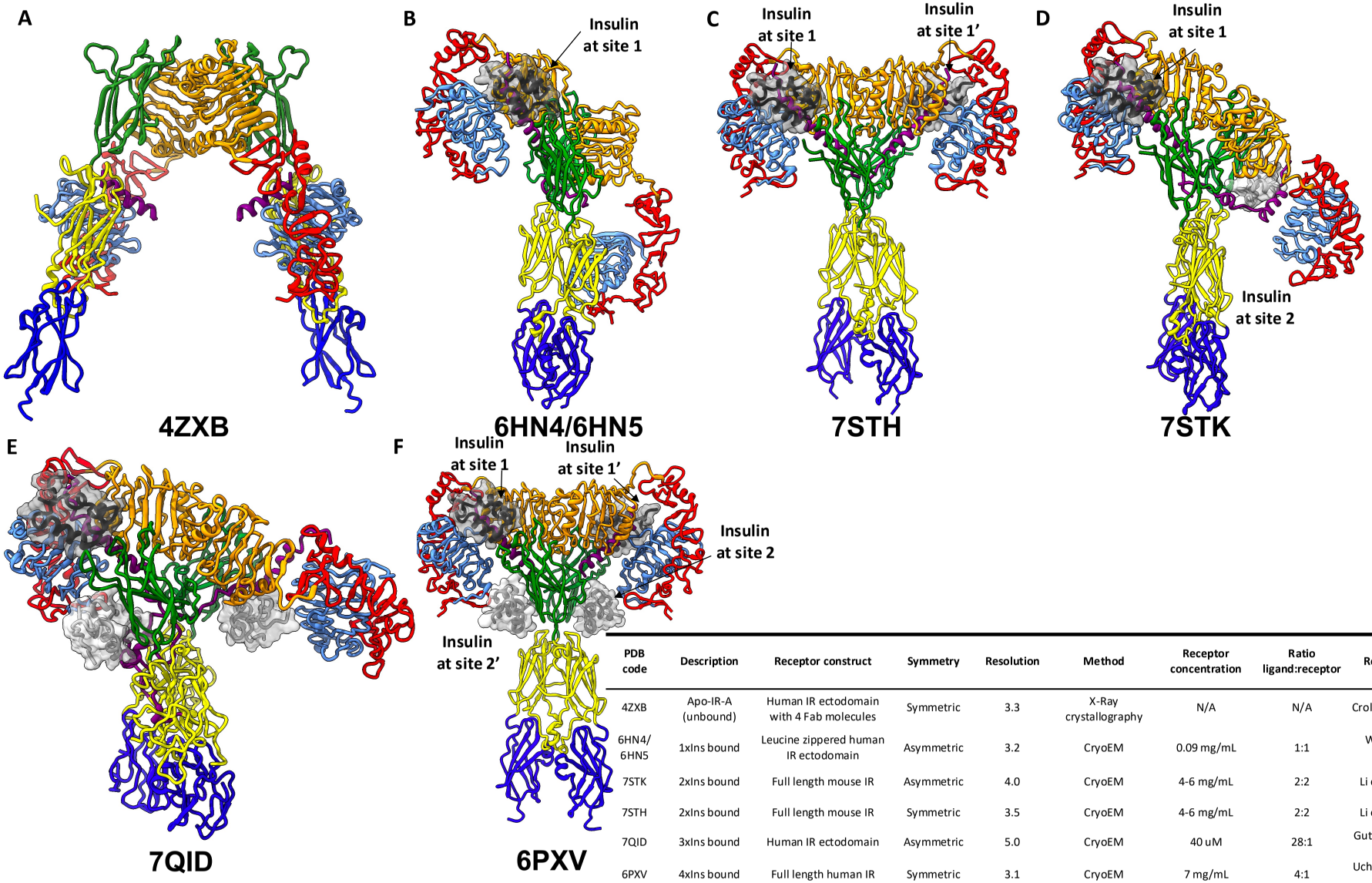
Insulin		IGF-II		IGF-I	
QB4	(Uchikawa et al., 2019)	-		-	
EB13	(Uchikawa et al., 2019, Kristensen et al., 1997)	D15	(Alvino et al., 2009)	D12	(Gauguin et al., 2008)
LB17	(Uchikawa et al., 2019, Kristensen et al., 1997)	F19	(Alvino et al., 2009)	F16	(Gauguin et al., 2008)
VB18	(Uchikawa et al., 2019)	-	-	-	-
LA13	(Uchikawa et al., 2019, Kristensen et al., 1997)	L53	(Alvino et al., 2009)	L54	(Gauguin et al., 2008)
YA14	(Uchikawa et al., 2019)	-	-	-	-
EA17	(Uchikawa et al., 2019, Kristensen et al., 1997)	E57	(Alvino et al., 2009)	E58	(Gauguin et al., 2008)
EB21	(Uchikawa et al., 2019)	-	-	-	-

For insulin:IR complexes there are several additional solved structures that have been reported with 2-4 insulins bound (structures shown in Figure 1.6). These structures were obtained by saturating the IR with excess insulin (conditions summarised in the corresponding table of Figure 1.6) and reveals for the first time the site 2 interaction on the FnIII-1′ domain (shown in Figure 1.7B), spanning residues 477-488 and 552-554 (Uchikawa et al., 2019). This interaction involves a distinctly different surface on the FnIII-1′ domain compared to site 1b. The range of receptor conformations, including symmetric and asymmetric (summarised in Figure 1.6), and varied number of insulins bound has raised debate over the series of binding events that occurs, particularly under physiological conditions. Investigations by Uchikawa et al (2019) and Li et al (2022) propose that the site 1 interaction occurs first, followed by site 2 engagement, which facilitates binding at site 1′ (Uchikawa et al., 2019). Contrary to this, several investigations have demonstrated that in the unbound state site 1 is inaccessible to ligand in this conformation as this region is observed to engage with elements of the FnIII-2′ domain (McKern et al., 2006, Whittaker et al., 2008, Croll et al., 2016). The site 2 interaction has therefore been proposed by several other groups to represent the initial point of IR contact (Gutmann et al., 2020, Lawrence, 2021). In the unbound state the site 2 interaction is solvent exposed and accessible for ligand binding (McKern et al., 2006, Whittaker et al., 2008, Croll et al., 2016). This binding event may facilitate the dissociation, of the site 1a L1 domain and αCT′ away from the FnIII-2′ domain (Gutmann et al., 2020, Lawrence, 2021). Further investigations are required to better understand exact order of binding events that occurs under physiological conditions, including the number of ligands involved, as this information will better inform us on how signaling activation is achieved. At physiologically relevant concentrations a single insulin is observed bound to the IR at site 1 (Weis et al., 2018) and results in receptor activation (Lee and Pilch, 1994, Shoelson et al., 1993).

While the site 2 interaction has yet to be visualised on the IGF-1R, mutagenesis data for IGF-I and IGF-II suggests that this interaction does occur (Table 1.1) (Alvino et al., 2011, Gauguin et al., 2008a). A ‘harmonic oscillator’ model of insulin receptor binding by insulin was described using a combination of biochemical and structural information (Kiselyov et al., 2009). This model defined the affinities of sites 1 and 2 as 6.41 nM and 399 nM, respectively, for insulin binding to the IR and 9.09 nM (site 1) and 490 nM (site 2), for IGF-I binding to the IGF-1R (Kiselyov et al., 2009).

1.1.3.3.2 First solved structure of IGF-II bound to the IGF-1R

Recently, a structure of the IGF-II:IGF-1R complex was determined using cryoEM to an average maximum resolution of 3.2 Å (Figure 1.5B) (Xu et al., 2020). This structure forms an asymmetric J-shaped conformation with the FnIII stalks together. The site 1a ligand binding interaction is similar to the previous insulin:IR and IGF-I:IGF-1R structures (Gutmann et al., 2020, Li et al., 2019, Uchikawa et al., 2019). The IGF-II molecule contacts the L1, L2, αCT′, and FnIII-1′ domains within the head region of the receptor (Figure 1.5C and D) (Xu et al., 2020). The L1-CR + (αCT′) module folds to the top of the receptor, permitting sparse interactions between IGF-II and the membrane-distal loops of FnIII-1′, facilitated by an outward rotation of domain L2 from its location in the apo ectodomain. The αCT′ helix on the L1 domain surface threads through the IGF-II C-domain loop (residues 33–40). The C-terminal segment of the IGF-II B-domain is displaced from the core of the ligand (normally packed against the core of the ligand in the unbound state) and engages with the receptor to make the site 1a interaction. The B-domain of IGF-II is stabilized by an interaction between IGF-II residue R30 and the hydroxyl group of IGF-1R residue Y28 and possibly a salt bridge between IGF-II residue R38 and IGF-1R residue E305. The ligand 'clips' onto the extended αCT′ helix in the active conformation, stabilizing a tight interaction between L1-CR-L2 and αCT′ with only sparse interactions between the ligand site and FnIII-1′ at site 1b (Figure 1.5C and D) (Xu et al., 2020). The site 1a interaction involves IGF-II side chains of residues of the B-domain (C9, L13, V14, L17, Q18, D23, F26, Y27, F28, S29, and R30) contacting the L1 and αCT′ segment and side chains of A-domain residues (I42, V43, E44, T58, Y59, and T62) contacting the αCT′ segment (but not the L1). These side chain interactions are similar in the IRZip:insulin and IGF-1R:IGF-I structures as the ligand sequences are highly conserved in these regions (Table 1.1). The site 1b interaction involves IGF-II B-domain residues E6, T7, C9 E12 and A-domain residues C47 and F48 which contact the FnIII-1′, thereby completing the definition of site 1b (Figure 1.5D).



PDB code	Description	Receptor construct	Symmetry	Resolution	Method	Receptor concentration	Ratio ligand:receptor	Reference
4ZXB	Apo-IR-A (unbound)	Human IR ectodomain with 4 Fab molecules	Symmetric	3.3	X-Ray crystallography	N/A	N/A	Croll et al 2016
6HN4/ 6HN5	1xIns bound	Leucine zippered human IR ectodomain	Asymmetric	3.2	CryoEM	0.09 mg/mL	1:1	Weis et al 2018
7STK	2xIns bound	Full length mouse IR	Asymmetric	4.0	CryoEM	4-6 mg/mL	2:2	Li et al 2022
7STH	2xIns bound	Full length mouse IR	Symmetric	3.5	CryoEM	4-6 mg/mL	2:2	Li et al 2022
7QID	3xIns bound	Human IR ectodomain	Asymmetric	5.0	CryoEM	40 uM	28:1	Gutmann et al 2020
6PXV	4xIns bound	Full length human IR	Symmetric	3.1	CryoEM	7 mg/mL	4:1	Uchikawa et al 2019

Figure 1.6: Structure of the IR unbound and with 1,2,3 and 4 insulins bound.

(A) Crystal structure of the apo insulin receptor (IR) ectodomain (PDB: 4ZXB). In the apo state the FnIII-3 legs of the receptor are positioned far apart forming an open Λ-Shape. (B) With a single insulin bound (black), an asymmetric J-shape conformation is formed where the FnIII-3 legs are in close proximity (PDB: 6HN4/6HN5). With two insulins bound, a symmetric shape can be formed (C), with each insulin bound to the site 1 and site 1' in the head region of the receptor (PDB: 7STH). Two insulins can also bind forming an asymmetric shape where one ligand is bound to site 1 and the other is bound to site 2' (D) (PDB: 7STK). Three insulins bound to the IR results in an asymmetric shape (E), where a single insulin is bound to each site 1 and site 1′ and the third insulin is bound to site 2′ (PDB: 7QID). Finally, a structure with four insulins bound to the IR has been solved, forming a symmetric shape with a single insulin bound to sites 1, 1′, 2 and 2′ (PDB: 6PXV). Shown in the table are the resolutions, final concentration of receptor and ligand: receptor stoichiometric ratio used for each cryoEM structure described in A-E. Notably, there are many different shaped IR complexes with up to four insulins bound, however, all structures with more than one insulin bound were prepared using stoichiometric ratios in excess of 1:1 ligand:receptor. Colouring scheme is the same as in Figure 1.3. Other structures exist with similar conformations as show in this figure however these structures shown here represent the highest resolution.

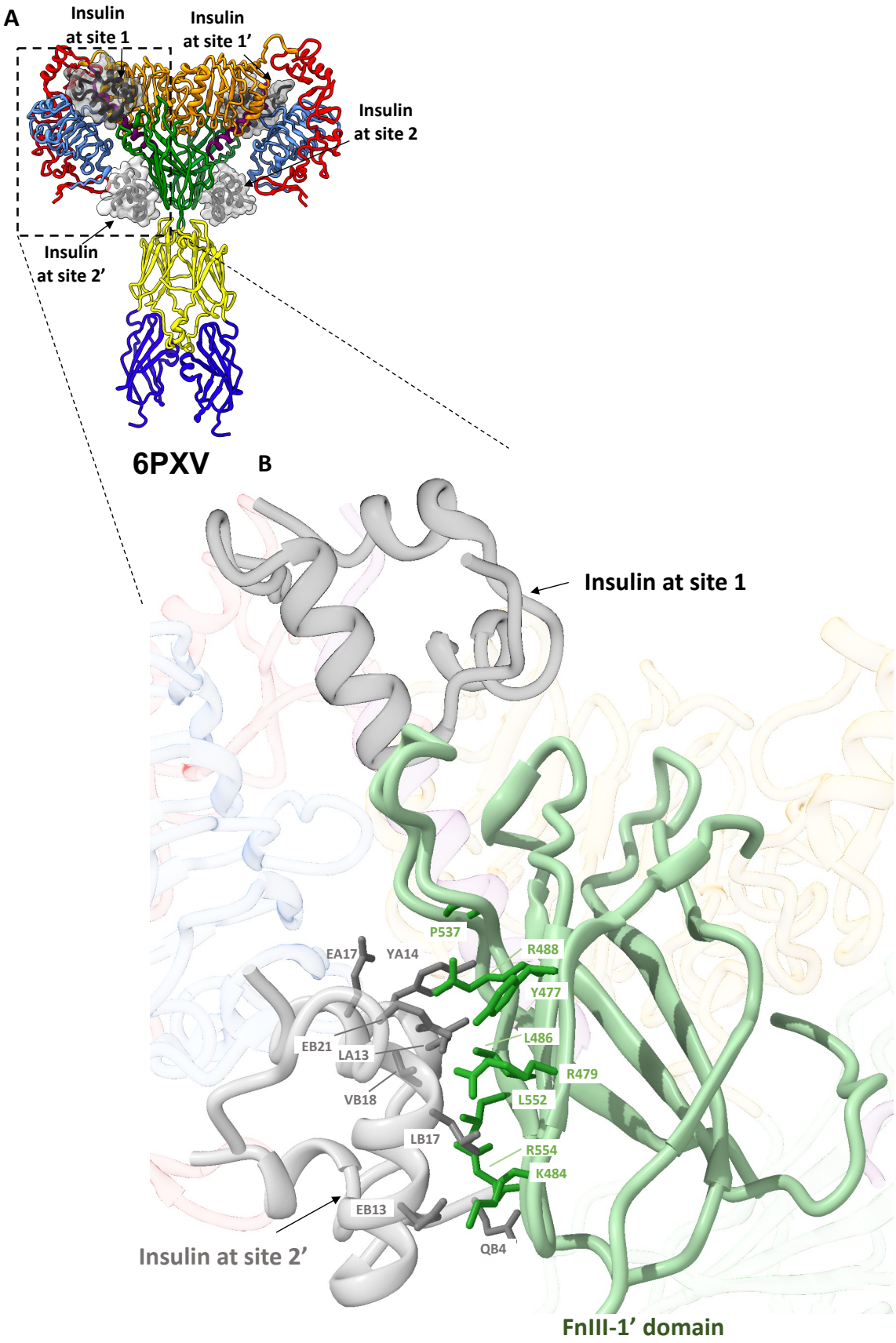

A
Insulin at site 1
Insulin at site 1'
Insulin at site 2
Insulin at site 2'
6PXV
B
Insulin at site 1
P537
EA17
YA14
R488
Y477
EB21
LA13
L486
R479
VB18
L552
LB17
R554
K484
EB13
QB4
Insulin at site 2'
FnIII-1' domain

Figure 1.7: Structure of the 4xInsulin bound to the IR and zoom in of site 2′ interaction.

(A) Structure of four insulins bound to the IR forming a symmetric shape with a single insulin bound to sites 1, 1′, 2 and 2′ (PDB: 6PXV). Domain colouring scheme as in Figure 1.3. (B) Zoom into the site 2′ interaction, as indicated by the dotted black box. Insulin residues QB4, GEB13, LB17, VB18, EB21, LA13, YA14 and EA17 contact receptor FnIII-1′ domain residues Y477′, R479′, K484′, L486′, R488′, P537′, L552′ and R554′ (Uchikawa et al 2019). This site 2′ interaction involves a different surface of the FnIII-1′ domain compared to the site 1b interaction.

This overall J-shaped conformation that brings the FnIII stalks together was seen previously in the Weis et al. insulin:IR (Figure 1.6B) and Li et al. IGF-I:IGF-1R studies (Weis et al., 2018, Li et al., 2019). The IGF-II:IGF-1R complex structure (Xu et al., 2020) was determined using a similar leucine-zippered receptor (IGF-1RZip) similar to that of the insulin receptor used in the Weis et al. study (Weis et al., 2018). The general topology of the single ligand bound asymmetric IGF-1RZip:IGF-II and IRZip:insulin structures also reflects that of the recently reported IGF-1R:IGF-I structure (Li et al., 2019), providing further evidence that this is the common activated IGF-1R conformation. The asymmetry observed in the activated structure is necessary for negative co-operativity, a hallmark of both IGF-1R and IR ligand binding summarized in a 'harmonic oscillator model' by Kiselyov et al (2009), whereby binding of a second ligand (to the unoccupied site 1 receptor binding pocket) accelerates the dissociation of the first bound ligand.

The major difference between the structures of IGF-II:IGF-1R and IGF-I:IGF-1R occurs in the respective growth factor C-domains and their interaction with the IGF-1R. In the IGF-1R complex, the IGF-II C-domain residues 33–36 are disordered, as are the adjacent receptor CR domain residues 258–265, suggesting that the C-domain is too short to form stable interactions with the receptor in this region (Figure 1.8) (Xu et al., 2020) (Xu et al., 2020). By contrast, the C-domain of IGF-I in IGF-1R:IGF-I is relatively well ordered, with IGF-I residue Y31 in its distal loop engaging receptor residues P5 and P256 (Figure 1.8) (Li et al., 2019). Although the resolution of the structure is low at IGF-I residues R36 and R37, they appear to contact IGF-1R L2 domain. With no equivalent to Y31, the IGF-II C-domain instead appears to be stabilized by self-interactions (a salt bridge with IGF-II residue E45 near the N-terminus of the first helix of the IGF-II A-domain, and a polar interaction with the IGF-II residue S39).

In summary, whilst IGF-II binds and activates IGF-1R through a similar mechanism to IGF-I, there are significant differences that likely explain their different binding affinities. Notably, the C-domain interactions are quite different, with IGF-II barely making receptor contact, whereas IGF-I C-domain contributes to binding affinity through several contacts. In addition, the lack of structural information for a site 2 interaction on the IGF-1R for both IGF-I and IGF-II may represent a major difference in receptor binding between the IGFs and insulin (for which a site 2 interaction has been captured). How this influences ligand specific signaling outcomes is still not understood. Importantly, no structure of IGF-II bound to IR-A has been reported.

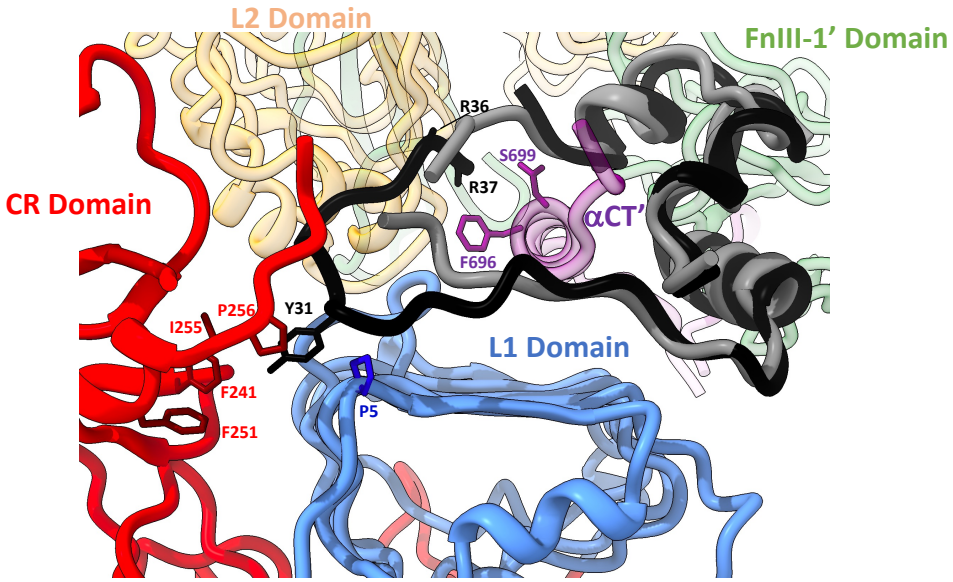


Figure 1.8: Comparison of IGF-II and IGF-I C-domain contacts with IGF-1R.

Interaction of IGF-II (grey) and IGF-I (black) C-domain with the IGF-1R (IGF-1R from 6PYH shown) (PDB: 6VWI and 6PYH and respectively). Domains of the IGF-1R coloured as in Figure 1.3. Residues of the CR and L1 domain engage with residue Tyr31 of IGF-I, for which there is no equivalent C-domain residue in IGF-II. Contact is also made between C-domain residues Arg36 and Arg37 of IGF-I and the L2 domain. Contact is also made between residue Q40 of IGF-I and F695 and S699 of the αCT′, however, there is no ordered structure for residue Q40 in this structure.

1.2 Implications of Structural Information for Developing Treatments for Disease

A detailed understanding of how IGF-II engages with its receptors and confers downstream signaling activation is essential for the development of drug therapies that target IGF action in cancer. Currently, most approaches target IGF action by directly blocking ligand binding to the IGF-1R using IGF-1R antibodies that inhibit ligand binding and stimulate receptor internalisation (Livingstone, 2013). Such inhibitors have been shown to reduce growth of IGF-II dependent cancers. However, increases in IGF-II:IR-A signaling can give rise to resistance to treatment (Belfiore et al., 2009, Gualberto and Pollak, 2009) (Belfiore et al., 2009, Gualberto and Pollak, 2009). This highlights the need for inhibitors of IGF-II acting via both IGF-1R and IR-A. Critical to this approach is being able to inhibit IGF-II without perturbing insulin signaling, and this highlights the need for structural data of IGF-II bound to IR-A. Such studies will further inform on how IGF-II is uniquely capable of binding and activating both IR-A and IGF-1R with high affinity and will suggest strategies to design inhibitors or allosteric regulators for the treatment of IGF-1R/IR-A regulated disease. A promising IGF-I analogue (R36E R37E IGF-I) has been shown to have antagonistic properties toward IGF-I signaling and suppression of tumorigenesis *in vivo* (Fujita et al., 2013, Saegusa et al., 2009). More recently, a viral insulin like peptide isolated from the lymphocystis disease virus-1 (LCDV-1) from the Iridoviridae family of DNA viruses has been shown to have antagonistic properties against the IGF-1R, which will be discussed further below (Altindis et al., 2018, Zhang et al., 2021). Such analogues demonstrate the promise and feasibility of targeting the ligand in attempts to slow IGF dependent cancer cell progression.

1.2.1 Alternative Mechanisms of IR Binding and Activation

The mechanism of insulin and IGF binding to their cognate receptors is highly conserved, involving site 1a contact with the L1 domain and αCT′, as well as contact with site 1b on the FnIII-1′ domain. Site 2 interaction, for insulin, involves contact on the FnIII-1′ domain. Mutagenesis data suggests this interaction likely occurs for the IGFs, but an analogous IGF-1R site 2 has yet to be visualised by structural studies. Residues of insulin and the IGFs that are involved in receptor engagement are highly conserved among vertebrate species (Irwin, 2021, Baral and Rotwein, 2019, Rotwein, 2017). There are however, subtle differences between vertebrate insulins and the IGFs that result in altered specificity for the IGF system receptors and signaling outcome. For example, in insulin, residue HB10 when mutated to a glutamate (as found at the equivalent residue in IGF-II, E12) increases affinity for the IGF-1R and IR-A to 1151% and 382%, respectively, compared to that of native insulin (Glendorf et al., 2012, Kaarsholm et al., 1993,

Kurtzhals et al., 2000). Interestingly, the mitogenic potency of this analogue also increases 257%.

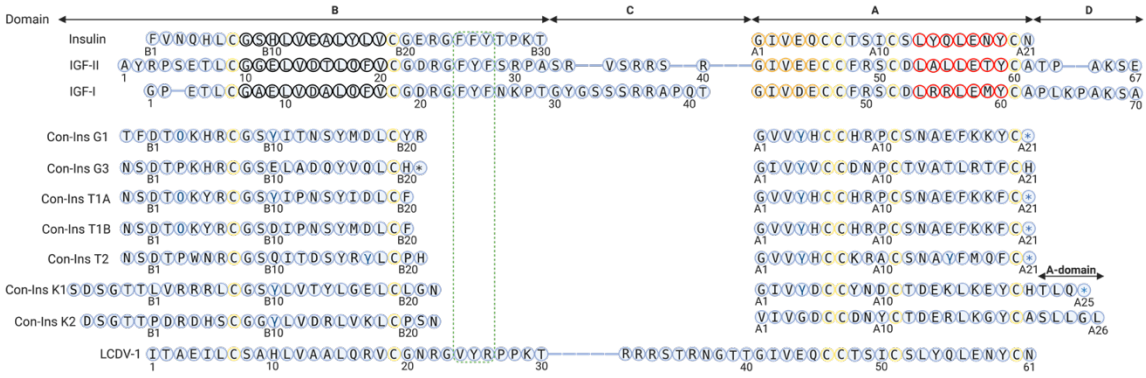


Figure 1.9: Sequence alignment of insulin, IGF-II, IGF-I, cone snail venom insulins and a viral insulin like peptide.

Sequence alignment of insulin, IGF-II, IGF-I, cone snail venom insulins identified from three species of cone snail *Conus. geographus*, *C. tulipa* and *C. kinoshitai* (Ahorukomeye et al 2019) and the viral insulin-like peptide lymphocystis disease virus-1 (LCDV-1) (Altindis et al., 2018). For human insulin and IGF-II peptides highlighted are the three helices; B-domain helix 1 (black circles), A-domain helix 2 (orange circles) and A-domain helix 3 (red circles). Residue numbers are indicated below each sequence and cysteine residues are coloured yellow. The green dotted box shows the FFY/FYF motif in insulin and the IGFs respectively and highlights the lack of this motif for the cone snail venom insulins. The viral insulin like peptide LCDV-1 has one of the three aromatic residues, a tyrosine at position Y25, equivalent to the Y27 and Y24 in IGF-II and IGF-I respectively. The cone snail venom insulins have several post translational modifications; γ, γ-carboxyglutamic acid; O, hydroxyproline; asterisk, C-terminal amidation.

More recently, it has been discovered that invertebrate insulins have less conservation of residues found to be important for receptor binding in human insulin yet they are still capable of binding to the IR with high affinity (Gronke et al., 2010, Li et al., 2020, Ahorukomeye et al., 2019). Amazingly, a number of cone snail venom insulins, isolated from *Conus geographus* and *C. kinoshitai,* have been recently found to lack the C-terminal segment of the B-chain (Figure 1.9). In human insulin, and indeed many other vertebrate insulins and IGFs, the C-terminal segment of the B-domain is essential for site 1a interaction. The importance of this segment was highlighted by attempts to remove the C-terminal 8 amino acids in human insulin (desoctainsulin or DOI), which led to a binding affinity of 0.09% for the IR and 0.002% for the IGF-1R, compared to human insulin (Weiss and Lawrence, 2018, Bromer and Chance, 1967, Cara et al., 1990). Astoundingly, *C. geographus* venom insulin 1 (Con-Ins G1) is capable of binding to the IR with high affinity through several residues which compensate for the loss of the FFY motif. Con-Ins G1 engages with receptor site 1a instead through residues YB15 and YB20 (Menting et al., 2016, Xiong et al., 2020). Site 1b is also strengthened in Con-Ins G1 through HA8, RA9 and a post translationally modified γ-carboxyl glutamate at position B10 (glutamic acid is the non-post translationally modified version (Safavi-Hemami et al., 2015)). This ligand was reported to have a similar affinity for the human IR as human insulin (Menting et al., 2016). These compensatory residues described above were incorporated into the human DOI analogue, termed Mini-Ins, which had a similar affinity to human insulin (Xiong et al., 2020). However, this high affinity binding was achieved with a weaker site 1a and stronger site 1b interaction compared to human insulin. Despite the altered site 1a and site 1b affinities, Mini-Ins retained equal potency of metabolic (Akt pathway) activation compared to human insulin. At lower concentrations Mini-Ins was less effective in promoting mitogenic signaling than human insulin (Xiong et al., 2020). These results suggest that alteration of the way in which insulins bind to the site 1a and 1b may influence signaling outcome.

C. kinoshitai venom insulin 1 (Con-Ins K1) is a cone snail venom insulin recently discovered with unique properties including an A-chain extension (T22, L23, Q24 and γ-carboxyl glutamate at position 25), a post translationally modified γ-carboxyl glutamate at position B10 and a leucine at position B20 (Xiong et al., 2021). Based on the Con-Ins K1 venom insulin, various combinations of A-chain extensions were investigated in the context of human DOI and found to re-establish high affinity binding for the IR, equal to that of human insulin (Xiong et al., 2021). This modified insulin analogue termed Vh-Ins has a four-residue helical A-chain extension; HA21, SA22, LA23 and QA24. This

extension forms a helix, extending the C-terminal helix of human insulin. Residue SA23 of the A-chain extension engages with the same pocket otherwise occupied by FB24 in human insulin (Xiong et al., 2021). LA23 also contributes to high affinity binding whilst residues HA21 and QA24 are less important. In contrast to Con-Ins G1, Vh-Ins has a high affinity for site 1a, and lack of side chain density for site 1b residues implies that the site 1b affinity is weakened (Xiong et al., 2021). While human insulin does not have an equivalent A-chain extension as found in Con-Ins K1, the IGFs do. The IGFs have a six amino acid extension at the C-terminus of the A-domain, known as the D-domain. The D-domain of IGF-II is equivalent to the A-chain extension found in Con-Ins K1, however, with a distinctly different amino acid sequence (Figure 1.9). The D-domain of the IGF-II contributes to a small extent to its high affinity for the IR-A (Denley et al., 2004).

Another source of unique insulin-like sequences has recently been described from lymphocystis disease virus-1 (LCDV-1) from the Iridoviridae family of DNA viruses (Altindis et al., 2018). These viral insulin-like peptides (VILPs), like the cone snail venom insulins, have a poorly conserved C-terminal segment of the B-chain, with only some elements of the aromatic FFY motif still present. These VILPs have more recently been shown to not only have moderate affinity for the human IGF-1R but also have potent antagonistic properties (Zhang et al., 2021). Recent structural analysis of a viral insulin like peptide (single chain Lymphocystis disease virus-1 viral insulin like peptide (scLCDV1-VILP)) bound to the IGF-1R found that the peptide engages with the L1 domain and αCT′ in a similar fashion to the IGFs. Interestingly the peptide also makes contact with FnIII-1′ domain residue H539′. This residue is equivalent to residue R544′ in the IR which forms part of the site 2 interaction (Gutmann et al., 2020). Interestingly, binding of this peptide resulted in further separation of the receptor legs compared to the unbound form (Moreau et al., 2022). The increased separation upon ligand binding was speculated to prevent receptor activation, consistent with the findings that scLCDV1-VILP has antagonistic properties (Moreau et al., 2022).

1.3 Aims and Significance

Overall, the IGF system has been heavily investigated, with the majority of research in this area involving insulin and IGF-I and to a much lesser extent, IGF-II. We have developed a detailed understanding of the site 1 interaction for each ligand bound to their cognate receptors. However, we still require a structure of IGF-II bound to the IR-A. The site 2 interaction has recently been observed for insulin bound to the IR at the FnIII-1′ domain and it is proposed that this interaction may represent the first site of contact for the ligand or an intermediate site to facilitate conformational change of the ligand and receptor. We are yet to observe this interaction for IGF-II or IGF-I bound to

the IGF-1R. The solved structures of unbound and receptor bound IGF-1R and IR reveal major conformational change occurs that leads to signaling activation. What we now require is a detailed understanding of the intermediate conformations that lead to signaling outcomes and additionally how metabolic and mitogenic signaling pathways are activated. Therefore, the overall aim of my PhD was to elucidate the molecular mechanisms underlying IGF-II receptor binding and activation leading to its unique signaling capabilities.

It has been shown that alterations to the human insulin sequence can lead to an alternate mechanism of receptor engagement leading to altered signaling outcomes. This has been demonstrated by incorporating Con-Ins G1 and Con-Ins K1 receptor binding residues into a site 1a minimised human insulin analogue, des octapeptide (B23–B30)-insulin (DOI). The resultant analogues, termed Mini-Ins and Vh-Ins, have shown that alternative mechanisms of receptor binding can be incorporated into human insulin analogues, and compensate for the missing FFY motif.

The overall goal of this project was to better understand the mechanism of IGF-II action and how signaling pathways are preferentially activated. To achieve this goal my aims were to develop site 1a minimised analogues of IGF-II and investigate if residues equivalent to the Con-Ins G1 and Con-Ins K1 residues that compensate for the loss of the aromatic FFY, can have the same compensatory effect when incorporated into site 1a minimised IGF-II analogues.

To achieve this overall goal, my PhD was based on three hypotheses:

Hypothesis 1: Mutation of the conserved FYF motif of IGF-II site 1a will not only affect IGF-1R/IR binding affinity but will also impact IGF-II structure and stability. A site 1a minimised IGF-II analogue could be achieved by introducing conservative substitutions at the IGF-II FYF motif.

Aim 1: To develop site 1a minimised IGF-II analogues by removing elements of the FYF motif.

Hypothesis 2: Cone snail venom insulin residues important for high affinity binding can be tolerated in IGF-II.

Aim 2: To individually investigate the effect of introducing Con-Ins G1, Con-Ins K1 or Con-Ins T1A alternate binding residues into human IGF-II

Hypothesis 3a: By combining the site 1a minimised IGF-II analogues with changes equivalent to the alternate binding mechanisms found in cone snail venom insulins, it

will be possible to compensate for the loss of the FYF motif in IGF-II.

Hypothesis 3b: Alternative modes of receptor contact will alter signaling bias, as is seen for Con-Ins G1.

Aim 3: To incorporate the alternate binding residues from Con-Ins G1 and Con-Ins K1 into the site 1a minimised IGF-II analogues and to test their ability to bind and activate the receptor.

By investigating alternate modes of receptor binding and determining the effect on signaling potency and bias, we can develop a detailed molecular understanding of how binding translates to preferential signaling activation. Ultimately, we aim to use this knowledge to develop small molecules/peptides with antagonistic properties to prevent IGF-II signaling through the IR-A to slow cancer growth and survival.

2 METHODS

2.1 Materials

2.1.1 Molecular Biology

The following reagents were purchased from New England Biolabs (MA, USA): 2-Log deoxyribose nucleic acid (DNA) Ladder, Antarctic Phosphatase, XbaI and EcoRI restriction enzymes, T4 DNA ligase, and dNTPs. GelRed® Nucleic Acid Gel Stain (10,000x) was purchased from Biotium. Ampicillin sodium salt (Amp), agarose, boric acid and ethylene diamine tetra-acetic acid (EDTA), sodium hydroxide, Bromophenol Blue, propan-2-ol, Tris-Ultrapure, 99.5% glycerol, β–mercaptoethanol, $RbCl_2$, 3-(N-Morpholino)propanesulfonic acid, 4-Morpholinepropanesulfonic acid (MOPS) and Xylene Cyanol, were obtained from Sigma-Aldrich (Merck). Big Dye (Version 3) Sequencing Reagent was purchased from Applied Biosystems. Bacto™ Tryptone and Bacto™ Yeast Extract were obtained from Becton Dickinson and Company. Bacto™ Agar was purchased from EMD Millipore. 100% ethanol, 100 % methanol, NaCl and acetic acid were obtained from ChemSupply. QIAprep Spin Miniprep Kit and QIAGEN® Gel Extraction Kit were purchased from QIAGEN®. Sodium dodecyl sulfate (SDS) was obtained from Boehringer Mannheim.

2.1.1.1 Solutions for Molecular Biology

2.1.1.1.1 20X TBE

1.79 M Boric acid, 1.78 M Tris, 63.6 mM EDTA in MilliQ H_2O pH 8.3

2.1.1.1.2 20X SB

760 mM Boric acid, 400 mM NaOH in MilliQ H_2O

2.1.1.1.3 6X DNA Loading Dye

1% (w/v) Bromophenol Blue, 50% Glycerol (v/v), 1% (w/v) Xylene Cyanol, 0.1% TBE (v/v) in MilliQ H_2O.

2.1.1.1.4 1X GelRed solution

1:100 dilution of 10,000x stock GelRed and 0.1 M NaCl in MQ H_2O

2.1.1.2 Kits for Molecular Biology

The QIAquick Gel Extraction kit for purification and extraction of DNA from agarose and the QIAprep spin Miniprep Kit for plasmid extraction and purification from bacterial cultures were purchased from QIAGEN® Pty. Ltd. (Victoria, Australia).

2.1.1.3 Bacterial Strains

TABLE 2.1: BACTERAL STRAINS USED FOR CLONING AND PROTEIN EXPRESSION.

Name	Genotype
E. coli DH5α	supE44 lacU169 hsdR17 recA1 endA1 gyrA96 λ^- relA1 [ϕ80lacZ m15]
E. coli BL21	fhuA2 [lon] ompT gal (λ DE3) [dcm] ΔhsdS λ DE3 = λ sBamHIo ΔEcoRI-B int::(lacI::PlacUV5::T7 gene1) i21 Δnin5

2.1.1.4 Bacterial Culture Media and Buffers

2.1.1.4.1 Luria Broth (LB)

5g/L bacto yeast extract, 10g/L bacto tryptone, 5g/L NaCl

2.1.1.4.2 LB ampicillin agar plates

LB agar ampicillin plates were produced by addition of 1.5% Bacto-agar to LB (Section 2.1.1.4.1) and ampicillin to a final concentration of 100 μg/mL.

2.1.1.4.3 Transformation Buffer 1

100 mM $RbCl_2$, 50 mM $MnCl_2$, 30 mM Potassium Acetate, 10 mM $CaCl_2$, 15% Glycerol, filter sterilized (0.2 μm).

2.1.1.4.4 Transformation Buffer 2

75 mM $CaCl_2$, 10 mM MOPS, 10 mM $RbCl_2$, 15% Glycerol, filter sterilized (0.2 μm).

2.1.1.4.5 BACTERIAL LYSIS BUFFER

10% β-Mercaptoethanol, 2% SDS in MilliQ H_2O.

2.1.1.5 Cloning Vector

The pGEM®-T Easy cloning vector system was used to amplify the insert DNA (Promega (Madison, Wisconsin USA)).

2.1.1.6 Expression Vectors

The pET32a expression vector (purchased from Novagen, Madison Wisconsin, USA) was used for the expression of IGF-II analogues.

2.1.1.7 Oligonucleotides

Primers used for the sequencing of IGF-II analogues are shown in Table 2.2.

TABLE 2.2:PRIMERS USED FOR THE SEQUENCING OF VECTOR CONSTRUCTS CONTAINING IGF-II ANALOGUE GENE SEQUENCE.

Primer	Sequence
T7 promoter (Sigma-Aldrich 69348)	5' TAA TAC GAC TCA CTA TAG GG 3'
T7 reverse (Sigma-Aldrich 69337)	5' GCT AGT TAT TGC TCA GCG G 3'

2.1.2 Expression and purification of IGF-II and IGF-II analogues

2.1.2.1 Protein electrophoresis

40% acrylamide, 2% Bis–acrylamide, Precision Plus Protein™ Dual Xtra standards and the protein gel apparatus Mini-PROTEAN® Electrophoresis Cell were purchased from BioRad Pty. Ltd. Coomassie Brilliant Blue R-250, N,N,N,N,N-tetramethylethylene-diamine (TEMED), β–mercaptoethanol, tricine, Bromophenol Blue, EDTA, Tris-HCl-(hydroxymethyl)-aminomethane (Tris-HCl) and urea were purchased from Sigma-Aldrich (Merck). Glycine was purchased from Bio-Strategy Lab Products.

2.1.2.2 Protein Electrophoresis Buffers

2.1.2.2.1 Glycine Running Buffer

1.918 M Gycine, 248 mM Tris, 1% SDS

2.1.2.2.2 Tricine Running Buffer

100 mM Tris, 10 mM Tricine, 0.1% SDS

2.1.2.2.3 2 x Protein Load Dye

125 mM Tris-HCl, 4% (w/v) SDS, 10% Glycerol, 0.1% Bromophenol Blue, pH 6.8

2.1.2.2.4 Gel Fixer Solution

50% (v/v) Ethanol, 10% (v/v) Acetic Acid

2.1.2.2.5 Gel Destain Solution

10% (v/v) Acetic Acid

2.1.2.2.6 Coomassie Blue Stain

50% (v/v) Ethanol, 10% (v/v) Acetic acid, 0.5% (w/v) Coomassie Brilliant Blue R-250

2.1.2.2.7 Gel Hydrator Solution

30% (v/v) Ethanol, 5% (v/v) Glycerol

2.1.2.3 Protein expression

Isopropyl-β-D-thiogalactopyranoside (IPTG) was purchased from Diagnostic Chemicals Limited. 2,5-dihydroxybenzoic acid (sDHB), HCl, Na_2SeO_3, Boric acid, $CoCl_2$, $CuCl_2$, was obtained from Sigma-Aldrich (Merck). $MnCl_2$ and KH_2PO_4 was obtained from BDH Chemicals. $ZnSO_4$ was purchased from Univar solutions. Trifluoracetic acid was purchased from Scharlau.

2.1.2.3.1 Protein expression buffers

2.1.2.3.1.1 Non-Inducing medium (NIM)

5g/L bacto yeast extract, 10g/L bacto tryptone, 50 mL/L 20X M, 2 mM $MgSO_4$, 0.2 mL/L1000X trace metals, 20 mL/L 50X 505

2.1.2.3.1.1.1 1000X TRACE METALS

60 mM HCl, 50 mM $FeCl_3$, 20 mM $CaCl_2$, 10 mM each of $MnCl_2$ and $ZnSO_4$, and 2 mM each of $CoCl_2$, $CuCl_2$, $NiCl_2$, Na_2MoO_4, Na_2SeO_3, and H_3BO_3

2.1.2.3.1.1.2 20X M

0.5 M KH_2PO_4, 0.5 M Na_2HPO_4, 1 M NH_4Cl and 100 mM Na_2SO_4

2.1.2.3.1.1.3 50X 505

2.5% (w/v) glucose, 25% (v/v) glycerol

2.1.2.4 Cell homogenisation and dissolution

2.1.2.4.1 CELL HOMOGENISATION AND DISSOLUTION BUFFERS

2.1.2.4.1.1 INCLUSION BODY WASH BUFFER

10 mM KH_2PO_4, 30 mM NaCl, pH 7.8

2.1.2.4.1.2 INCLUSION BODY DISSOLUTION BUFFER

8 M Urea, 0.1 M Tris, 40 mM Glycine, pH 2 (0.2 μm filtered)

2.1.2.5 Chromatography

2.1.2.5.1 HIGH PERFORMANCE LIQUID CHROMATOGRAPHY (HPLC) BUFFERS

2.1.2.5.1.1 HPLC BUFFER A

0.1 % Trifluoracetic acid (0.22 μm filtered)

2.1.2.5.1.2 HPLC BUFFER B

80 % Acetonitrile, 0.08 % Trifluoracetic acid (0.22 μm filtered)

2.1.2.6 Protein refolding and cleavage

2.1.2.6.1 10X REFOLD BUFFER

1 M Tris, 2 mM EDTA, pH 9.1 (0.22 μm filtered)

2.1.2.7 Mass spectrometry

2.1.2.7.1 sDHB MATRIX SOLUTION

50 g/L 2,5-dihydroxybenzoic acid (sDHB) in 0.1 % trifluoro acetic acid (TFA), 50 % acetonitrile.

2.1.3 Characterisation of IGF-II and IGF-II Analogue Biological Function

Hybridoma cells expressing antibodies specific for the IGF-1R α subunit (24-31) and IR α subunit (83-7) were a kind gift from Prof. K Siddle (Soos and Siddle, 1989, Ganderton et al., 1992). hIR-A over expressing L6 myoblasts were provided by Dr B.F. Hansen (Novo Nordisk A/S, Denmark). P6 cells (BALB/c3T3 cells overexpressing the human IGF-1R) were from Prof. R. Baserga (Pietrzkowski et al., 1992). [3H]-Thymidine was purchased from Perkin Elmer Life Sciences. R^- fibroblast cells (isolated from IGF-1R knockout mouse embryonic fibroblasts) were a gift from Prof. R. Baserga (Philadelphia, USA) (Sell et al., 1994) and R^- cells overexpressing hIR-A and hIR-B and were produced by Denley et al (Denley et al., 2004). Phospho Akt (T308) (New England Biolabs #9275S), phospho p44/42 MAPK (ERK1/2) (T202/Y204) (New England Biolabs #9101S) and mouse anti-β-tubulin (Invitrogen #32-2600), IRDye 800CW donkey anti rabbit IgG (Millenium Science 926-32213), IRDye 680RD donkey anti mouse (Millenium Science 926-68072) were used.

2.1.3.1 Cell lines and culture conditions

All cells were maintained at 37 °C in 5 % CO_2. R^-IR-A, R^-IR-B, L6 IR-A, and P6 cells were maintained in Dulbecco's modified Eagle's medium high-glucose (4.5 g/mL, DMEM) supplemented with 10% fetal calf serum, 100 units/mL penicillin, 100 μg/mL streptomycin and 0.5 % (250 μg/mL) G418. All cell culture media and supplements were purchased from Thermo Fisher Scientific Australia. Fetal calf serum was from Bovogen.

2.1.3.1.1 FUNCTIONAL CHARACTERISATION BUFFERS

2.1.3.1.1.1 1X TBST BUFFER

150 mM NaCl, 20 mM Tris pH 7.4 and 0.1% (v/v) Tween 20

2.1.3.1.1.2 BLOCKING BUFFER

3% bovine serum albumin (BSA) in 1x TBST (Section 2.1.3.1.1.1)

2.1.3.1.1.3 1X PBS BUFFER

1.46 mM KH_2PO_4, 136.9 mM NaCl, 2.68 mM KCl, 8.1 mM $NaHPO_4$ pH 7.4

2.2 Methods for Molecular Biology

2.2.1 Poly A Tailing of DNA Fragments

Poly A tailing of DNA fragments was performed to leave an adenosine overhang at the 3' end of each strand. This allows the fragment to be ligated into a vector that has been linearised with a single 3' thymidine overhang at each end. 100 ng of G-block DNA was poly A tailed in a 60 µL polymerase chain reaction (PCR) with 1x ThermoTaq buffer, 1.25 units of Taq polymerase and 10mM dATPs. The reaction was incubated at 72°C for 20 minutes using a PTC-200 Thermal Cycler (MJ Research).

2.2.2 TA Cloning of G-Block fragments

Linear pGEM®T Easy vector and poly A tailed DNA fragments (Section 2.2.1) were ligated as per the manufacturer's protocol. 25 ng of pGEM®T Easy vector, 0.5 µL of G-block DNA, 1.5 Weiss units of T4 DNA ligase and 1X Rapid Ligation Buffer were added to a PCR tube and incubated for 2 h at room temperature (Clark, 1988). 5 µL of the ligation reaction was transformed into 100 µL DH5α competent cells (Section 2.2.4).

2.2.3 Preparation of Competent *E. coli* Bacterial Cells

A single colony of *E. coli* cells from an LB agar streak plate was inoculated into 2 mL of LB (Section 2.1.1.4.1) and incubated at 37°C with shaking overnight. 330 µL of the overnight culture was then subcultured into 10 mL of fresh LB followed by incubation at 37°C with shaking for ~1.5 – 2 h until the OD at a wavelength of 600 nm reached 0.6. 5 mL of the culture was further subcultured into 100 mL of fresh, pre-warmed (at 37°C) LB in a 1 L flask followed by 1.5 h of incubation at 37°C with shaking. Next, the culture was evenly divided into four 50 mL sterile centrifuge tubes and incubated on ice for 5 minutes prior to centrifugation at 1,500 g for 5 minutes at 4°C. After centrifugation, the supernatant was removed from each tube and the cells were suspended in 10 mL of transformation buffer 1 (Section 2.1.1.4.3) per tube, then stored on ice for 5 minutes. The resuspended cells were then centrifuged at 1,500 g for 5 minutes at 4°C. Subsequently, the supernatant was removed and the cells were resuspended in 1 mL/tube of transformation buffer 2 (Section 2.1.1.4.4). Each tube was stored on ice for an additional 5 minutes prior to being aliquoted into 1.5 mL eppendorf tubes and stored at -80°C until required.

2.2.4 Bacterial Transformation

Approximately 100 pg - 10 ng of plasmid DNA was added to 100 μL of thawed DH5α cells or 50-100 ng was added to 100 μL of thawed BL21 competent *E. coli* cells and incubated on ice for 30 minutes. The cells were heat shocked at 42°C for 2 minutes, followed by immediate incubation on ice for 5 minutes. 1mL of LB (Section 2.1.1.4.1) was then added to the cells, aseptically, followed by incubation at 37°C for 30 minutes. The transformation reaction was then centrifuged at 1,500 g for 2 minutes then 1 mL of the supernatant was removed leaving 100 μL of supernatant and the pellet of cells. The cells were gently resuspended in the 100 μL of remaining supernatant and aseptically spread onto pre-warmed LB agar plates containing 100 μg/mL of ampicillin (Section 2.1.1.4.2). The plates were incubated for 16 h at 37°C. Single colonies were then selected and inoculated in to 5 mL of LB containing 100 μg/mL of ampicillin, cultured for 16 h at 37°C with agitation. For transformations using DH5α cells, DNA plasmid purification was then performed to extract the vector (DNA plasmid purification Section 2.2.6). Glycerol stocks were also prepared (Long term storage of plasmid DNA in DH5α and BL21 cells Section 2.2.5).

2.2.5 Long-Term Storage of Plasmid in DH5α and BL21 Cells

Cells containing the plasmid of interest were inoculated into 5 mL of LB (Section 2.1.1.4.1) containing 100 μg/mL of ampicillin and cultured for 16 h at 37°C with agitation. Equal volumes (200 μL) of 80% glycerol and the overnight culture were added to a 1.5 mL eppendorf tube and gently mixed by pipetting. The tubes were immediately stored at -80°C until required.

2.2.6 Plasmid Purification

Using DH5α cells from a glycerol stock, a streak plate was prepared (on an LB plate containing 100 μg/mL of ampicillin Section 2.1.1.4.2) and incubated at 37°C for 16 h. 5 mL of LB containing 100 μg/mL of ampicillin was inoculated with a single colony from the streak plate and then cultured for 16 h at 37°C with agitation. The culture was then transferred to sterile 15 mL tubes and centrifuged at 1,500 g for 5 minutes. The supernatant was carefully removed and the tube was centrifuged for an additional 2 minutes at 1,500 g. Any remaining supernatant was removed by pipetting. The plasmid DNA was extracted from the remaining cell pellet using the AccuPrep® Nano-Plus Plasmid DNA Extraction kit, according to the manufacturer's instructions. The plasmid

DNA was eluted in 50 μL of elution buffer.

2.2.7 Polymerase Chain Reaction (PCR) Amplification

PCR reactions were prepared (final concentration of each component Table 2.3), in a 50 μL total reaction volume using MilliQ H_2O. The PCR amplification of each IGF-II analogue was completed using primers listed in Table 2.1 and PCR cycling conditions shown in Table 2.4.

TABLE 2.3: COMPONENTS AND CONCENTRATIONS OF PCR REACTIONS FOR A 50 μL TOTAL REACTION.

Component	Concentration
Phusion HF Buffer	1X
dNTP's	200 μM
Forward primer	200 μM
Reverse primer	200 μM
Template DNA	1 pg – 10 ng
Phusion HF DNA polymerase	0.02 U/μΛ

TABLE 2.4: PCR CYCLING CONDITIONS FOR THE AMPLIFICATION OF IGF-II ANALOGUES IN THE PET32A VECTOR.

Cycle step	Temperature (°C)	Time (seconds)	Cycles
Initial denaturation	98	30	1
Denaturation	98	10	30
Annealing	62	30	
Extension	72	10	
Final extension	72	240	1

2.2.8 Mutagenesis reactions

Mutagenesis PCR reactions were prepared (final concentration of each component), in a 50 μL total reaction volume using MilliQ H_2O. The mutagenesis of each IGF-II analogue was completed using mutagenesis primers listed in Table 2.5 and PCR cycling conditions shown in Table 2.6.

TABLE 2.5: MUTAGENESIS PRIMER SEQUENCES FOR MUTATION OF IGF-II CONSTRUCTS. CODONS IN BOLD REPRESENT THE MUTATION TO BE INCORPORATED INTO THE IGF-II SEQUENCE. NON-BOLDED CODONS REPRESENT COMPLIMENTARY SEQUENCE TO THE IGF-II.

Mutation	Forward primer	Reverse primer
G22A	5' CAGTTCGTTTGC**GCG**GATCGTGGC 3'	5' GCC ACG ATC **CGC** GCA AAC GAA CTG 3'
G22L	5' CAGTTCGTTTGC**CTG**GATCGTGGC 3'	5' GCC ACG ATC **CAG** GCA AAC GAA CTG 3'
G22Y	5' CAGTTCGTTTGC**TAT**GATCGTGGC 3'	5' GCC ACG ATC **ATA** GCA AAC GAA CTG 3'
L17A	5' CTGGTGGATACC**GCG**CAGTTCGTTTGC3'	5' GCAAACGAACTG**CGC**GGTATCCACCAG 3'
L17Y	5' CTGGTGGATACC**TAT**CAGTTCGTTTGC 3'	5' GCAAACGAACTG**ATA**GGTATCCACCAG 3'

TABLE 2.6: PCR CYCLING CONDITIONS FOR MUTAGENESIS REACTIONS.

Cycle step	Temperature (°C)	Time (seconds)	Cycles
Initial denaturation	98	30	1
Denaturation	98	10	30
Annealing	62	30	
Extension	72	10	
Final extension	72	240	1

2.2.9 Agarose Gel Electrophoresis of DNA

DNA samples were diluted in 6x DNA loading buffer (Section 2.1.1.1.3) and MilliQ H_2O. The samples were electrophoresed through 1-2% agarose gels in a 1x TBE running buffer (Section 2.1.1.1.1) at 100 V in a Bio-Rad Sub-cell GT WIDE MINI running chamber. The agarose gels were stained using Biotium GelRed for 10 minutes and then washed with MilliQ H_2O twice followed by imaging using a BioRad Gel Doc EZ imager.

Small DNA fragments <300 bp were run on 2% agarose gels with 1x SB running buffer (Section 2.1.1.1.2).

2.2.10 Restriction Enzyme Digestion of DNA

Purified DNA was digested with 5-20 units of the required restriction enzymes, in NEB CutSmart Buffer (containing 50 mM potassium acetate, 20 mM tris-acetate, 10 mM magnesium acetate, 100 μg/mL BSA, pH 7.9) to a final volume of 50 μL with MilliQ H_2O. Digestion reactions were incubated at 37°C for 16 h. For vector digestions Antarctic phosphatase treatment (Section 2.2.11) was performed immediately after digestion. For digestion of IGF-II analogue fragments, termination of digestion was achieved by gel

electrophoresis (Section 2.2.9).

2.2.11 Antarctic Phosphatase Removal of 5' Phosphate from Digested Vector

Antarctic phosphatase removal of the 5' phosphate from the digested vector was achieved using New England Biolabs® Antarctic phosphatase, according to the manufacturers instructions. Briefly, 6 μL of 10X Antarctic phosphatase buffer and 3 μL of sterile MilliQ H_2O was added to the 50 μL digestion reaction. 1 μL of Antarctic phosphatase enzyme was added to the reaction followed by pipette mixing and incubation at 37°C for 30 minutes. The Antarctic phosphatase enzyme was subsequently heat inactivated at 80°C for 2 minutes.

2.2.12 Plasmid Ligation

DNA ligation was completed using insert fragments and vector constructs that had been digested with the same restriction enzymes, producing complimentary ends. Digested vector and insert at a 1:3 molar ratio, was added to 1x T4 DNA ligase buffer (containing 1 mM ATP, 50 mM tris HCl, 10 mM $MgCl_2$, 10 mM DTT, pH 7.5), 200 units of T4 DNA ligase, prepared to a total reaction volume of 10 μL with MilliQ H_2O. The ligation reaction was incubated for 16 h at 4°C, followed by transformation of 5 μL of the ligation reaction into competent DH5α cells (Section 2.2.4). For each ligation reaction a background control ligation was also prepared consisting of digested vector only (no insert DNA), which represents the amount of incompletely digested vector present in each ligation reaction. Single colonies were selected from the positive ligation (containing both vector and insert) reaction plate, inoculated into 5 mL of LB containing ampicillin and cultured overnight at 37°C with agitation. DNA plasmid purification of each culture was then prepared (Section 2.2.6) and diagnostic PCR reactions (Section 2.2.7) were performed on each analogue and electrophoresed on agarose gels (Section 2.2.9) to confirm that each insert was successfully ligated into the vector. Upon confirmation of expected sized PCR fragment, DNA sequencing (Sequencing 2.2.13) was then performed to confirm the correct sequence of the vector and insert.

2.2.13 Confirmation of Plasmid Sequence by DNA Sequencing

To a PCR tube 6 μL of Big Dye buffer, 2 μL of Big Dye terminator, 0.5 μL of 100 ng/μL primer, 250 ng of DNA and sterile MilliQ H_2O to 20 μL was added. The PCR tube was

flick mixed, pulse spun and subsequently placed into a Gene Works PTC-200 Peltier Thermal Cycler machine. PCR cycling conditions used are shown in Table 2.7.

TABLE 2.7: PCR CYCLING CONDITIONS FOR DNA SEQUENCING.

Cycle step	Temperature (°C)	Time (seconds)	Cycles
Initial denaturation	98	30	1
Denaturation	98	10	30
Annealing	51	30	
Extension	72	60	
Final extension	72	420	1

After the extension PCR reaction was complete the entire 20 μL reaction was pipetted into a 1.5 mL eppendorf tube with 80 μL of 75% isopropanol to precipitate the reaction. The tube was flick mixed then incubated at room temperature for 15 minutes. The DNA was then pelleted by centrifugation at 15,000 g for 30 minutes. The supernatant was carefully removed and 200 μL of fresh 75% isopropanol was added to the tube and vortexed. The DNA was then pelleted again by centrifugation at 15,000 g for 15 minutes followed by the supernatant being carefully removed. The tubes were kept open to dry before being sent to the Australian Genome Research Facility (AGRF) for sequencing.

2.2.14 DNA Purification of Gel Extract

Once DNA fragments had been separated on a preparative DNA gel (Section 2.2.9) and GelRed stained (Section 2.1.1.1.4) the gel was visualized using an ultraviolet transilluminator on a low wavelength UV setting. The bands corresponding to the expected size of DNA sequence were excised from the gel using a sterile scalpel. The excised gel was placed into an eppendorf tube and purification of the DNA was completed using an Invitrogen PureLink™ gel extraction kit following the manufacturer's

instructions. The DNA was eluted in 30 μL of elution buffer.

2.3 Methods for Expression and Purification of IGF-II and IGF-II analogues

2.3.1 Defined Minimal Medium Preparation

The defined minimal medium was prepared as follows: 10 g/L of tryptone and 5 g/L of yeast extract were added to a sterile 1 L bottle containing ~800 mL of MilliQ H_2O and dissolved. 50 mL of 20X M, and 20 mL of 50X 505 were added to the solution followed by 200 μL of 1000x trace metals and 2 mL of 1 M magnesium sulfate. MilliQ H_2O was added to a total of 1 L then the solution was filter sterilised.

2.3.2 Trial Expression of IGF-II Analogues in BL21 Expression Cells

The pET32a IGF-II analogues vector constructs were separately transformed into *E. coli* BL21 cells (Section 2.2.4) and plated onto LB agar selection plates containing 100 μg/mL of ampicillin (Section 2.1.1.4.2). The selection plates were incubated at 37°C for 16 h. A single colony from each selection plate was inoculated into 5 mL of non-inducing medium (NIM) (Section 2.1.2.3.1.1) containing 100 μg/mL of ampicillin followed by incubation at 37°C with shaking for 16 h. The OD_{600} of each culture was recorded using a OD_{600} from an Optizen POP UV/VIS spectrophotometer then sub-cultured into 5 mL of fresh NIM, with 100 μg/mL of ampicillin, to an OD_{600} of 0.2 A.U. The cultures were then incubated at 37°C with shaking for ~45 minutes until the OD_{600} reached 1.5 A.U. A 1 mL pre-induction sample of the culture was taken and stored in an eppendorf tube. Glycerol stocks were also prepared at this stage (Section 2.2.5).

Protein expression was then induced by addition of 0.1 mM IPTG followed by incubation at 37°C with shaking overnight. The following day a 1 mL induction sample was taken and stored in an eppendorf tube.

2.3.3 Preparation of Pre- and Post-Induction Samples

The pre- and post-induction samples were centrifuged at 15,000 g for 2 minutes The supernatant was removed, and the cell pellets were resuspended in 40 μL of bacterial lysis buffer per OD_{600} unit of the culture to standardize the concentration of bacterial proteins. The pre-and post-induction samples were run on 15% tricine SDS-polyacrylamide gels (SDS-PAGE) (Section 2.3.4). Successful induction was determined

by comparing the pre-and post-induction lanes.

2.3.4 Sodium Dodecyl Sulfate Polyacrylamide Gel Electrophoresis (SDS-PAGE)

SDS-PAGE analysis was completed using 5 μL of each protein sample with 5 μL of 2x protein loading dye (Section 2.1.2.2.3). The sample mixture was mixed by vortexing then boiled at 95°C for 5 minutes followed by centrifugation at 15,000 g for 2 minutes. Each protein sample (total of 10 μL each) was then loaded into separate lanes of a 15% Tricine SDS-PAGE gel, prepared using a Mini-PROTEAN® tetra handcast system, as per the manufacturer's instructions. The 15% Tricine gels were resolved at 100 V in Tricine running buffer (Section 2.1.2.2.2) for 1.5 h until the dye front had reached the bottom of the gel.

After electrophoresis the gels were placed into gel fixing solution (Section 2.1.2.2.4) for 15 minutes followed by staining with Coomassie Blue R-250 (Section 2.1.2.2.6) for 1 h, with slow agitation at 37°C. Next the non-specifically bound Coomassie Blue R-250 was removed from the gel using a destain solution (Section 2.1.2.2.5) for >16 h at 37°C. The gels were then incubated in a gel hydrating solution (Section 2.1.2.2.7) for 1 h before imaging using a Bio-Rad Gel Doc EZ imager.

2.3.5 Analytical reverse phase High Performance Liquid Chromatography (rpHPLC)

Analytical reverse phase-high performance liquid chromatography (rpHPLC) was used to determine the protein concentration and to monitor folding at each stage of protein refold and purification. Analysis was performed using a HICHROM C4 analytical column (214TP;5 μm, 2.1 x 100 mm, 300 Å) on an Agilent 1260 Infinity Quaternary liquid chromatography system. Buffer A: 0.1% TFA. Buffer B: 80% CH_3CN in 0.08% TFA. A 5 minute linear gradient of 20 – 25 % CH_3CN (25 – 31.3% buffer B, section 2.1.2.5.1.2), followed by a 25 – 45 % CH_3CN (31.3 – 56.3% buffer B, section 2.1.2.5.1.2) gradient over 20 minutes at flow rate of 0.5 mL/minute was used. Peptide elution was monitored by UV detection at 215 nm, and peak areas were used to determine peptide concentration.

2.3.6 Large Scale Expression of Peptides in *E. coli*

After confirmation of successful small-scale expression of IGF-II or IGF-II analogues

(Section 2.3.2) prepared glycerol stocks (Section 2.2.5) were used to inoculate 5 mL of NIM containing 100 μg/mL of ampicillin (Section 2.1.2.3.1.1) followed by incubation at 37°C with shaking for 5 h. 10 μL of culture was then inoculated into 50 mL of NIM containing 100 μg/mL of ampicillin and incubated for 16 h at 37°C with shaking. The following day the OD_{600} was recorded using an Optizen POP UV/VIS spectrophotometer. 500 mL of fresh NIM with 100 μg/mL of ampicillin was inoculated with culture to an OD_{600} of 0.2 A.U. and cultures were incubated at 37°C with shaking until the OD_{600} reached 1.5 A.U. Protein expression was then induced by addition of 0.1 mM IPTG followed by incubation at 37°C with shaking overnight. 1 mL pre- and post-induction samples were taken and prepared as per Section 2.3.3.

2.3.7 Lysis of *E.coli* by Homogenisation

Pelleted *E. coli* cells were resuspended in 30 mL of inclusion body (IB) wash buffer on ice (Section 2.1.2.4.1.1) prior to homogenisation at 30 kpsi using an E-1061 cell disrupter (Constant Systems Ltd). Samples were passed through the cell disrupter three times at 4°C to release the inclusion bodies. The lysed samples were centrifuged at 4,000 g for 10 minutes at 4°C. The supernatant was removed, and the pellet was washed using 5 mL of fresh IB wash buffer. The resuspended pellet was centrifuged at 10,000 g for 10 minutes. The supernatant was removed, and the pellet was again resuspended in 5 mL of fresh IB wash buffer. 1mL aliquots were pipetted into Eppendorf tubes and centrifuged at 10,000 g for 5 minutes at 4°C. Supernatant was removed and final pellets were stored at -20°C.

2.3.8 Gel Filtration

Gel filtration was initially performed to separate IGF-II or IGF-II analogues from host cell debris, including proteolytic enzymes. A GE HealthCare column (300 mm x 10 mm internal diameter column) packed with Superdex 75 (Sigma-Aldrich (Merck)) resin was equilibrated with MilliQ H_2O, then the column was sanitised by injection of 1 mL of 0.5 M NaOH and a column volume of MilliQ H_2O was then run through the column. Finally, the column was equilibrated in dissolution buffer (Section 2.1.2.4.1.2) containing 1.6 mM DTT. Each step was performed at a flowrate of 0.5 mL/minute using a GE Healthcare AKTA pure fast performance liquid chromatography unit (FPLC).

Inclusion bodies were dissolved in 700 μL dissolution buffer (Section 2.1.2.4.1.2) containing 20 mM dithiothreitol (DTT) pH 2 for 10 minutes at room temperature. The

dissolution solution was centrifuged at 15,000 g for 2 minutes then the supernatant was injected onto the Superdex 75 column equilibrated in dissolution buffer containing 1.6 mM DTT. The elution profile was monitored by U.V detection at 280 nm. 0.5 mL fractions were collected using a GE Healthcare fraction collector. Fractions were screened by analytical HPLC (Section 2.3.5).

2.3.9 Refolding of Fusion proteins

Refolding of IGF-II or IGF-II analogues was performed at a final peptide concentration of 0.5 μg/μL in 8 M Urea, 0.1 M Tris, 10 mM glycine, 5 mM EDTA, 1.6 mM DTT at pH 2. Refolding was initiated by addition of 1.25 mM 2-Hydroxyethyldisulfide. Refold was performed at room temperature for 30 minutes and was monitored by analytical HPLC at various timepoints.

2.3.10 α Lytic protease cleavage

At the end of refolding, IGF-II and IGF-II analogues were cleaved from their fusion partner (the first 11 amino acids of the porcine growth hormone (pGH 1-11)) by a mutant α-lytic protease called Prag A9 (Lien et al., 2001). This enzyme cleaves between the methionine of the PAPM linker and alanine which is the first residue of human IGF-II. The cleavage reaction was initiated by addition of α-lytic protease to the refold reaction to a final concentration of 1:600 enzyme:fusion protein. The cleavage reaction was monitored by analytical HPLC and cleavage was terminated by acidification to pH 2 using concentrated HCl.

2.3.11 Preparative reverse phase HPLC

Removal of the pGH leader and other impurities was performed using preparative HPLC. The acidified cleavage reaction was filtered using a 0.22 μm filter and loaded onto an Avantor® preparative column connected to an Agilent 1260 Infinity Quaternary liquid chromatography system. A 25-55% acetonitrile gradient (with 0.1% TFA) over 60 minutes at a flowrate of 2 mL/minute was used to separate the IGF-II or IGF-II analogues from the cleaved pGH leader. 0.5 mL fractions were collected and screened using analytical HPLC for pure IGF-II or IGF-II analogues. Pure IGF-II or IGF-II analogue fractions were pooled for freeze drying (Section 2.3.12). Concentration of the pooled peptides was determined using analytical HPLC.

2.3.12 Freeze drying

Pooled fractions of IGF-II or IGF-II analogues from preparative HPLC (Section 2.3.11) were aliquoted into 100, 50 or 20 μg in Eppendorf tubes. The samples were covered with parafilm and holes were made in the parafilm. Tubes were frozen on dry ice and then freeze dried using a Christ freeze drier for 16 h at <0.005 mbar at -80°C

2.3.13 Matrix assisted laser desorption ionisation mass spectrometry

The molecular weights of intact IGF-II analogues were determined using Matrix Assisted Laser Desorption Ionisation (MALDI) at Flinders Analytical and the University of South Australia, with the assistance of Dr Mark Condina and Dr Matthew Briggs. A MALDI Autoflextreme (Bruker Daltonics, Leipzig, Germany) equipped with a Nd: YAG laser (wavelength: 355 nm) was used. Purified and fully folded IGF-II analogues were subjected to mass spectral analysis using combinations of positive or negative ion with linear or reflectron mode. 1 μL of each sample was added to 1 μL superDHB (sDHB) matrix solution (Section 2.1.2.7.1), mixed, then 1 μL was spotted onto the MTP 384 ground steel target plate (Bruker Daltonics), and air dried. Calibrations were performed using Bruker Daltonics' Protein Calibration Standard I according to manufacturer's instructions.

2.4 Methods for functional characterisation

2.4.1 Receptor competition binding assays

IR and IGF-1R binding was measured essentially as described by Denley et al (Denley et al., 2004). Human IR isoform A (IR-A), isoform B (IR-B) and IGF-1R were solubilized from R^-IR-A, R^-IR-B and P6 cells, respectively. Cells were serum-starved for 4 h, in serum free medium (SFM) containing 1% BSA, before lysis in ice–cold lysis buffer (20 mM HEPES, 150 mM NaCl, 1.5 mM $MgCl_2$, 10% (v/v) glycerol, 1% (v/v) Triton X-100, 1 mM EGTA, and cOmplete protease inhibitor cocktail (Roche), pH 7.5)) for 1 h at 4 °C. Lysates were centrifuged for 10 minutes at 2,200 g, then 100 μL lysate was added per well to a white Greiner Lumitrac 600 96-well plate previously coated with anti-IGF-1R antibody 24-31 or anti-IR antibody 83-7 (250 ng/well in bicarbonate buffer pH 9.2). For IR-B binding assays approximately 3,000,000 fluorescent counts of europium-labelled insulin (Eu-insulin, prepared in-house) was added to each well along with increasing concentrations of unlabelled competitor in a final volume of 100 μL. For IR-A and IGF-1R binding assays approximately 3,000,000 fluorescent counts of europium-labelled

IGF-II (Eu-IGF-II, prepared in-house) was added. Reactions were incubated for 16 h at 4°C. Wells were washed three times with 1x TBST (20 mM Tris pH 7.5, 150 mM NaCl, and 0.1% (v / v) Tween 20), then 100 μL of DELFIA enhancement solution (PerkinElmer Life Sciences) was added to each well. After 10 minutes time-resolved fluorescence was measured using 340 nm excitation and 612 nm emission filters with a Victor X4, 2030 Multilabel Reader (Perkin Elmer). Assays were performed in triplicate in at least three independent experiments unless otherwise stated, and the data is an average of these.

2.4.2 Kinase receptor activation assays

Following a starvation period of 4 h in serum free DMEM containing 1% BSA, L6 IR-A cells were stimulated with 10 nM IGF-II or IGF-II analogue for various times (up to 20 minutes in DMEM containing 1% BSA). Lysates of cells were precipitated using 20% trichloroacetic acid and separated on a reducing 10% SDS-PAGE gel, transferred to nitrocellulose membranes and blocked using blocking buffer (Section 2.1.3.1.1.2) for 2 h before immunoblotting with primary antibodies (1:1000 phospho Akt (T308), 1:1000 phospho p44/42 MAPK (ERK1/2) (T202/Y204), 1:1000 mouse anti-β-tubulin) for 16 h at 4°C. Nitrocellulose membranes were washed with TBST (Section 2.1.3.1.1.1) 3x and then incubated with secondary antibodies (1:50,000 IRDye 800CW donkey anti rabbit IgG, 1:50,000 IRDye 680RD donkey anti mouse) for 1 h at room temperature. Following 3x TBST washes, the blots were scanned using an Odyssey CLx Imaging System (LI-COR Biosciences). Quantitation was performed with Image Studio Lite software. β-tubulin was used as a loading control and data was normalised to this. Activation is expressed as a fraction of the response to IGF-II at 10 minutes. A total of three independent experiments were performed, unless otherwise stated, and the data is an average of these.

2.4.3 DNA synthesis cell proliferation assays

DNA synthesis was carried out essentially as described by (Salic and Mitchison, 2008). Briefly, L6 IR-A cells were plated in a 96 well flat bottom plate ($32x10^4$ cells/well) and grown overnight at 37°C, 5% CO_2 as previously described (Gauguin et al., 2008b). Cells were starved in SFM for 2 h prior to treatment with increasing concentrations of insulin (Insulin, IGF-II or IGF-II analogues (0.01 - 300 nM) for 18 h in DMEM, 1 % BSA at 37°C and 5% CO_2. The cells were incubated with 10 μM of 5-Ethynyl-2'-deoxyuridine (EdU) for 4 h, washed with filtered PBS, 1% BSA and fixed in the dark for 15 minutes with 4% paraformaldehyde (PFA). Fixed cells were washed with PBS, 1% BSA and

permeabilised for 20 minutes with 0.5% Triton X-100. A click chemistry labelling cocktail (2 µM FAM-Azide 488, 100 mM Tris, pH 7.5, 4 mM $CuSO_4$, 100 mM sodium ascorbate) was added to the cells for 30 minutes at room temperature in the dark. Finally, cells were washed three times with PBS (Section 2.1.3.1.1.3) + 1% BSA and fluorescence was measured using 485 nm excitation and 535 nm emission filters with a PerkinElmer VICTOR X4 2030 Multilabel Reader. Assays were performed in triplicate in at least three independent experiments.

2.4.3.1 Statistical analyses

Statistical analyses were performed using the software package Prism v9.0.0 (GraphPad Software). Receptor binding assays (Section 2.4.1) were analysed using a nonlinear regression (one-site) model and Mean IC_{50} values were calculated. Receptor activation (Section 2.4.2) and DNA synthesis assays (Section 2.4.3) were performed using a two-way ANOVA with Dunnett's multiple comparison. Significance was accepted at $p < 0.033$.

2.4.4 Circular Dichroism analysis

Circular dichroism (CD) was performed as previously described (van Lierop et al., 2017). Briefly, CD spectra were recorded on a Jasco J-1500 CD spectrometer.

2.4.4.1 CD spectra

Spectra were measured from 300 to 180 nm with a 1.0 nm step size using a 1.0 s response time and 1.0 nm bandwidth in a quartz cuvette with a 0.2 cm path length. IGF-II and IGF-II analogues were resuspended in 10 mM acetic acid (pH 3) to a concentration of 0.05 µg/µL. To correct for background, the spectrum of buffer alone was subtracted from each sample spectra. The machine units collected, θ in millidegrees, was converted to the mean residue ellipticity (MRE), [θ] in degrees.cm^{2}dmol^{-1}residue^{-1}, as follows:

$$[\theta] = \theta \times \frac{(0.1 \times MRW)}{(P \times Conc.)}$$

The MRW is the protein mean weight ((atomic mass units/daltons) / number of residues)), P is pathlength (cm) and Conc. is protein concentration in mg/mL. $[\theta]_{222}$ is the molar ellipticity per residue at wavelength 222 nm. Helical content was calculated using the CDSSTR algorithm (Sreerama and Woody, 2000) for deconvolution against the reference protein database set SMP180. The program is available on the DICROWEB website (http://dichroweb.cryst.bbk.ac.uk/html/home.shtml).

2.4.4.1.1 THERMAL DENATURATION

Spectra were measured, for temperature denaturation studies, at a wavelength of 222 nm with a 1.0 nm step size using a 1.0 s response time and 1.0 nm bandwidth in a quartz cuvette with a 0.2 cm path length. Automated thermal control was used to increase the temperature by 2°C /minute at 1°C intervals at a peptide concentration of 0.05 μg/μL in 10 mM acetic acid (pH 3). MRE was calculated for the thermal denaturation profiles as done in Section 2.4.4.1.

3 IGF-II SITE 1A MINIMISED ANALOGUES

3.1 Introduction

Insulin, insulin like growth factor I (IGF-I) and IGF-II all share a high degree of sequence (Figure 3.1A) and structural similarity (Blyth et al., 2020). Site 1a contact is essential for high affinity binding of insulin and the IGFs to their cognate receptors (Lawrence, 2021). A key element of this site 1a interaction is the displacement of the C-terminal segment of the B-domain from the core of the ligand (Figure 3.1B and C) and the engagement of the aromatic FFY motif in insulin (FB24, FB25, YB26; IGF-I F24, Y25, F26; IGF-II F26, Y27, F28) with the L1 domain of the receptor (Figure 3.1E and F). FB24 of insulin 'anchors' insulin to the hydrophobic pocket of the L1 domain (comprising IR residues N15, L37, F39) and αCT residue F714, as shown in Figure 3.1E (Weis et al., 2018, Li et al., 2019, Xu et al., 2020, Menting et al., 2014, Xu et al., 2009). The importance of FB24 has also been demonstrated through extensive mutagenesis studies, having been mutated to each of the naturally occurring amino acids (with the exception of a cysteine) (Pandyarajan et al., 2014, Zakova et al., 2013, Mirmira and Tager, 1989). In general, it was concluded that aliphatic residues are best tolerated, with methionine, glycine and valine having essentially equal affinity for the insulin receptor (IR) as native IGF-II. Insulin residues FB25 and YB26 (Y27 and F28 for IGF-II and Y24 and F25 for IGF-I) have also been found to be important for receptor binding through mutagenesis studies. However, mutation of these residues in insulin is better tolerated compared to FB24 substitution (Zhang et al., 2003, Nakagawa and Tager, 1986, Mirmira et al., 1991, Pandyarajan et al., 2016). Structural investigations of the insulin:IR complex and the IGF:IGF-1R complexes depict these residues interacting with shallow depressions on the L1 domain and contacting the αCT (Weis et al., 2018, Li et al., 2019, Xu et al., 2020). Mutation of these residues to alanine results in 10% and 64% decrease in affinity for the IR respectively compared to human insulin (Kristensen et al., 1997, Kobayashi et al., 1984, Inouye et al., 1981, Chen et al., 2000).

Considerably fewer mutational studies have been performed on the FYF motifs of IGF-I and IGF-II compared to human insulin. Mutation of F26 in IGF-II (equivalent to FB24 in insulin) to a serine resulted in a 3.6-fold and 83.3-fold decrease in affinity for the IGF-1R and IR-A, respectively (Sakano et al., 1991). Y27L IGF-II mutants were reported to have a 25-95 fold decrease for the IGF-1R and 293-fold decrease for the IR-A (Sakano et al., 1991, Roth et al., 1991, Burgisser et al., 1991, Forbes et al., 2002). A Y27E IGF-II mutant was also reported with a 54-fold decreased affinity for the IGF-1R (Burgisser et al., 1991). Mutation of residue F28 to a leucine leads to an affinity of 2.8% and 6.2% for the IGF-

1R and IR-A, respectively, relative to native IGF-II (Alvino et al., 2009). For IGF-I a F23G mutant (equivalent to FB24 in insulin) had a 5-fold decrease in affinity for the IGF-1R (Hodgson et al., 1996). The Y24L IGF-I had a reduced binding affinity of 24- and 10-fold for the IGF-1R and IR-A respectively (Cascieri et al., 1988). Similarly, Y24S IGF-I had a reduced affinity of 16-fold and 2-fold for the IGF-1R and IR-A respectively (Cascieri et al., 1988).

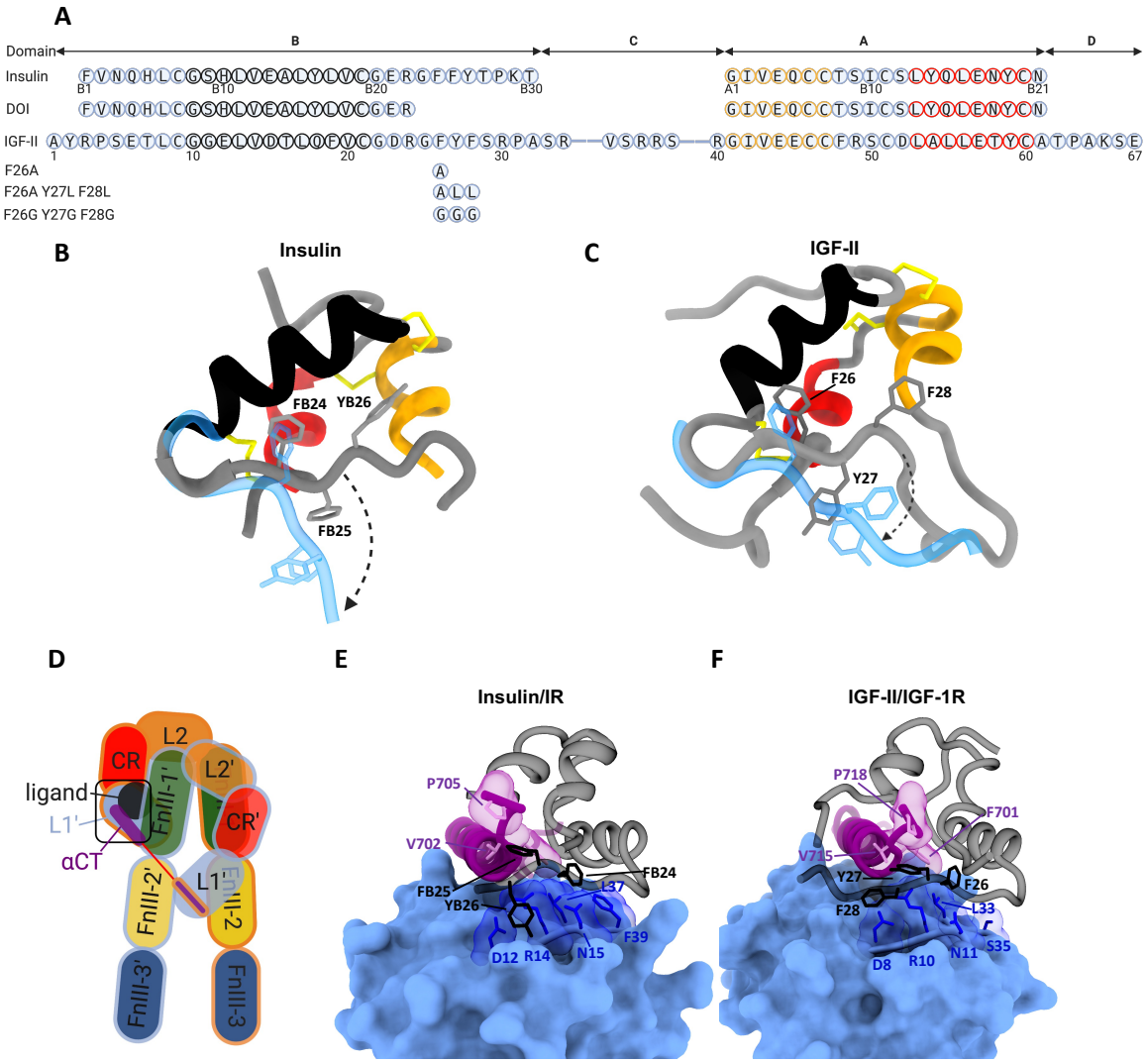

A
Domain
B
C
A
D
Insulin
B1
B10
B20
B30
A1
A10
A21
DOI
IGF-II
1
10
20
30
40
50
60
67
F26A
F26A Y27L F28L
F26G Y27G F28G
B
Insulin
FB24
YB26
FB25
C
IGF-II
F26
F28
Y27
D
L2
CR
L2'
ligand
FnIII-1'
CR'
L1'
αCT
FnIII-2'
L1'
FnIII-2
FnIII-3'
FnIII-3
E
Insulin/IR
P705
V702
FB24
L37
FB25
YB26
F39
D12
R14
N15
F
IGF-II/IGF-1R
P718
F701
V715
Y27
F26
F28
L33
S35
D8
R10
N11

Figure 3.1: Sequence and structural comparison of IGF-II and insulin.

(A) Sequence alignment of IGF-II, insulin, des-octapeptide (B23–B30)–insulin and three analogues; F26A IGF-II, F26A Y27L F28L IGF-II and F26G Y27G F28G IGF-II. Each peptide has three helices; B-domain helix 1 (black circles), A-domain helix 2 (orange circles) and A-domain helix 3 (red circles). Residue numbers are indicated below each sequence. (B) Insulin and (C) IGF-II unbound structures (PDB:1MSO and 1IGL respectively). Helix colouring as in A and disulfides shown in yellow. The C-terminal 9 amino acids of the B-domain from receptor bound insulin and IGF-II structures (PDB: 6HN4 and 6VWI) respectively is shown in transparent blue. Side chains of the FFY motif for insulin and FYF motif for IGF-II are shown and depicted is the displacement of this motif upon ligand binding to engage with the receptor. (D) Schematic diagram representing the IR and IGF-1R ligand bound extracellular domain structures. Individual αβ monomer outlines are coloured either blue or orange. Extracellular domains include the first and second leucine-rich repeat domains (L1 and L2), cysteine rich domain (CR), first, second and third fibronectin type-III domains (FnIII-1, -2, and -3), insert domain (ID), α-chain C-terminal region (αCT). Transmembrane and intracellular domains, including the tyrosine kinase domain, are not shown. (E) Insulin and (F) IGF-II bound to site 1 of their cognate receptors, as highlighted with the black box in (D) (from PDB: 6HN5 and 6VWI, respectively). L1 domains are surface filled in cornflower blue and αCT are coloured purple. The side chains of residues involved in contact with the FFY/FYF motif are shown and annotated.

The site 1a aromatic motif is also known for its important structural role. In the IGFs and insulin this motif forms part of the hydrophobic core of the protein (Torres et al., 1995, Blundell et al., 1972, Brzozowski et al., 2002). It has been proposed that the beginning of insulin and IGF folding involves the association of highly conserved hydrophobic residues (Weiss, 2009). In particular, nascent interactions of a microdomain surrounding residues CA20-CB19 are required to facilitate the correct orientation of the thiol groups for disulfide bond formation (Hua et al., 2006b, Narhi et al., 1993, Yan et al., 2003). This microdomain involves residue FB24 of the insulin aromatic FFY motif (F23 for IGF-I and F26 for IGF-II). Subsequently the interchain CA7-CB7 and the intra A-chain CA6-CA11 disulfides are formed. The order in which these two bonds form has been shown to be interchangeable and both are required for formation of the biologically active insulin conformation (Hua et al., 1996, Hua et al., 2001). A similar folding pathway has also been described for IGF-I. The first stable disulfide to form is C18-C61 followed by C48-C6 (Narhi et al., 1993). A high abundance of this two-disulfide folding intermediate (C18-C61, C48-C6) has been characterised (Miller et al., 1993, Hober et al., 1992, Rosenfeld et al., 1997). The final disulfide to form is C47-C52.

In IGF-I an alternative disulfide bonded structure has also been well documented (C18-C61, C6-C47, C48-C52 and shown in Figure 3.2A). This structure results in lower helical content compared to native IGF-I and a shifted spectral minimum is observed in circular dichroism (CD) studies compared to native IGF-I (supplementary Figure 7)(Hober et al., 1992). In addition, the alternate disulfide pairing of IGF-I leads to a 200-fold reduction in affinity for the IGF-1R (Milner et al., 1995). This alternate disulfide pairing of IGF-I has been shown to occur in equilibrium with the native form of IGF-I *in vitro* (Hober et al., 1997, Hober et al., 1992, Hejnaes et al., 1992). However, optimised refolding conditions during recombinant production, including the use of a leader peptide, have been developed to favour the predominant formation of the native IGF-I structure (King et al., 1992). Although there is considerably less known about IGF-II folding it is expected that a similar folding pathway occurs to form the C21-C60, C47-C9 and C46-C51 native disulfide bonds. There is evidence that an alternate disulfide bonded IGF-II isomer (C21-C60 followed by C46-C9 and C47-C51; Figure 3.2B) can form and it has a 160-fold decrease in affinity compared to native IGF-II (Smith et al., 1989). IGF-II folding has been optimised with a similar method to IGF-I, where a leader sequence is utilised to facilitate the predominant formation of the native IGF-II fold (Lien et al., 2001, Alvino et al., 2009).

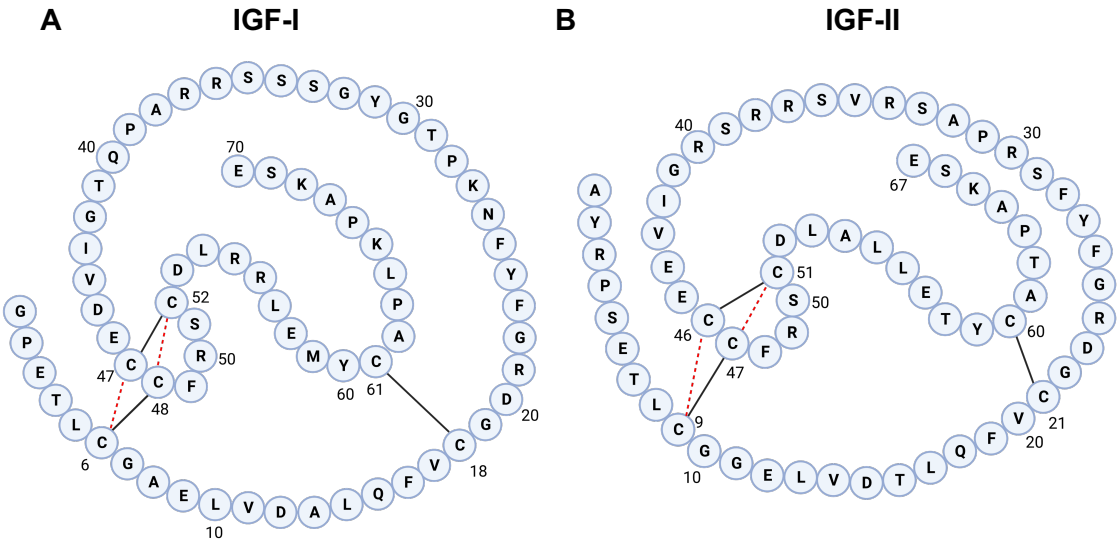


Figure 3.2: Disulfide pairing analogues of IGF-I and IGF-II.

The sequence of (A) IGF-I and (B) IGF-II are shown with native disulfide pairings represented with a solid black line and alternate disulfide bonds represented with red dashed lines.

Whilst the FFY aromatic motif is fundamental in folding, structure, receptor contact and is highly conserved among vertebrate insulin and IGFs (Conlon, 2001, Mirmira et al., 1991, Conlon, 2000, Irwin, 2021, Baral and Rotwein, 2019, Rotwein, 2017), this motif is not conserved in many invertebrate species (Gronke et al., 2010, Li et al., 2020, Ahorukomeye et al., 2019) and in some cases is totally absent. An example of this is seen in cone snail venom insulins which lack the C-terminal segment of the B-domain, including the key FFY motif (Ahorukomeye et al., 2019). These venom insulins have been shown to engage with the receptor through different contacts and still activate the human insulin receptor with relatively high affinity (Menting et al., 2016). In recent years there has been considerable research directed towards understanding alternate modes of IR and IGF-1R activation and several studies have correlated altered binding properties with altered signaling kinetics (Zhang et al., 2021, Xiong et al., 2021, Xiong et al., 2020). Cone snail venom insulin G1 (Con-Ins G1), which lacks the C-terminal 8 amino acids of the B-domain (including the FFY motif, shown in Figure 3.1A), binds to the IR with much higher affinity than expected through several key contacts not observed in human insulin (Xiong et al., 2020, Menting et al., 2016). A comparable mutant human insulin lacking the last 8 amino acids of the B-domain (des octapeptide (B23–B30)-insulin (DOI) binds very poorly to the IR (0.09% compared to native insulin) (Weiss and Lawrence, 2018, Bromer and Chance, 1967, Cara et al., 1990). A recent study introduced amino acid changes into DOI equivalent to those from Con-Ins G1 that make key receptor contacts (termed Mini-Ins) and restored similar binding affinity to human insulin (Xiong et al., 2020). In addition, Mini-Ins retained equal potency of metabolic (Akt pathway) activation and lower rate of mitogenic (ERK pathway) activation compared to human insulin.

Our laboratory aims to investigate the effect of introducing the same amino acid changes seen in Mini-Ins into IGF-II in the context of an IGF-II analogue that binds poorly to site 1a to determine if similar impact on receptor binding and signaling occurs. Understanding residue specific determinants of altered signaling bias is required to develop antagonists which can selectively block mitogenic action of IGF-II. The first step was to generate mutants with poor site 1a binding. Because IGF-II is a single chain ligand, total removal of the C-terminal segment of the B-domain, as in DOI, is not possible. Instead, mutation of the site 1a must be used. In insulin, the FYF motif has been shown to be important for receptor binding, protein folding and formation of the core of the protein. However, Con-

Ins G1 and Mini Ins, both lacking the C-terminal 8 amino acids, demonstrate that stable and highly active insulins can be produced despite the lack of the FFY motif (Ahorukomeye et al., 2019, Menting et al., 2016). In this chapter we designed three variations of site 1a minimised IGF-II analogues to determine which peptide maintained native IGF-II like structure but significantly reduced binding affinity.

In the IGFs and insulin the first F residue (F26) of the aromatic FYF motif acts as an anchor to the L1 domain of the receptor. Alanine mutants of the equivalent residue in insulin (FB24) results in 1-5% affinity for the IR compared to human insulin (Mirmira et al., 1991, Kristensen et al., 1997, Mirmira and Tager, 1989). As mentioned above, mutation of residue FB24 in insulin results in the largest decrease in receptor binding affinity compared to mutation of residue FB25 and YB26 (reviewed in (Weiss and Lawrence, 2018)). Mutation of residues FB25 and YB26 to alanine results in 10% and 64% decrease in affinity, respectively, for the IR compared to human insulin (Kristensen et al., 1997, Kobayashi et al., 1984). In the current investigation the first site 1a mutant was a single amino acid change of F26 of IGF-II to alanine. This mutant was used to test the effect on receptor binding with the expectation that this would have minimal impact on the IGF-II core structure. There is evidence to suggest that alanine mutation of insulin residue B24 can be compensated by a single residue register shift where the neighbouring phenylalanine FB25 can engage in the hydrophobic pocket of the L1 domain that FB24 would otherwise occupy (Rege et al., 2020a). This register shift results in essentially equal binding affinity for the IR as human insulin. Although F26 in IGF-II does not neighbour another phenylalanine, but rather a tyrosine (Y27), it is possible that this highly similar hydrophobic residue may compensate for the loss of F26, as observed for insulin. In case a register shift also occurs in the F26A IGF-II mutant, a second combination F26A Y27L F28L IGF-II mutant was also produced. The leucine residues at positions 27 and 28 were selected based on previous individual mutations made in IGF-II previously which resulted in a 94.3-fold and 6.4-fold decrease in IGF-1R binding respectively (Sakano et al., 1991, Delaine et al., 2007). It was anticipated that the hydrophobic leucine residues would stabilise the core of the protein but result in a low affinity interaction with the receptor. To investigate total removal of the FYF motif a F26G Y27G F28G IGF-II analogue was also produced, this mutant would inform on the contribution of the FYF motif to the disulfide pairing and stabilisation of the hydrophobic core of IGF-II. The expectation was that the IGF-II FYF motif would play a similar role to the insulin FFY motif and that the F26G Y27G F28G IGF-II analogue may not fold due to the total loss of hydrophobic residues to participate in initial core formation.

In summary, the aim of this chapter was to generate site 1a mutants of IGF-II that would then act as the basis for exploration in subsequent chapters of whether alternate binding mechanisms analogous to those used by cone snail insulins could be introduced into IGF-II. The design, expression, purification and subsequent receptor binding and signaling properties of three IGF-II analogues is discussed. The mutants investigated are F26A IGF-II, F26A Y27L F28L IGF-II and F26G Y27G F28G IGF-II.

3.2 Results

The results are presented in two parts. Native IGF-II production is first described to provide an understanding of the optimised conditions used for generating folded and active IGF-II. The biophysical and functional assays conducted to monitor folding and biological activity are also introduced. This provides a comparison for subsequent refolding and characterisation of the site 1a knockout analogues.

3.2.1 Production of IGF-II

3.2.1.1 Expression and Dissolution of IGF-II

IGF-II expression has previously been optimised in our laboratory using an expression plasmid encoding the 11 N-terminal amino acids of the porcine growth hormone (pGH(1-11)) as a fusion partner and with a protease cleavage recognition site (Lien et al., 2001). The pGH(1-11) fusion partner facilitates high level expression and inclusion body formation, increased solubility and increased efficiency of refold during refold steps (Francis et al., 1992). IGF-II expression was induced by addition of IPTG (Section 2.3.6). Pre- and post-induction samples were separated on a 15% tricine gel under reducing conditions, followed by staining with Coomassie blue (Section 2.4.3). Successful expression of pGH(1-11) IGF-II was confirmed by the presence of a 9.3 kDa band (expected size of pGH(1-11) IGF-II) in the post induction sample but not in the pre induction sample. Inclusion bodies were released from *E.coli* by high pressure homogenisation (Section 2.3.7), solubilised (Section 2.3.8 and Figure 3.3A), then separated using size exclusion chromatography (Section 2.3.8) to separate from any remaining high molecular weight host cell material including proteolytic enzymes (Figure 3.3B).

3.2.1.2 Refolding, cleavage and purification of IGF-II

Refold of pGH(1-11) IGF-II was performed in a final concentration of 2.5 M urea and 0.5 mM DTT (Section 2.3.9) (Figure 3.3C). Refold was initiated by the addition of the oxidising agent 2-hydroxyethyldisulfide, which created an oxidising environment that

facilitated the formation of the three disulfide bonds present in pGH(1-11) IGF-II and formation of the biologically active structure. As shown in Figure 3.3C, the major species at T=3 hours (18.0 minute retention time) has a decreased retention time compared to T=0, characteristic of a transition from a reduced state to a folded state. An intermediate folding state of pGH(1-11) IGF-II was observed at 19.5 minutes, which has been previously detected when folding pGH(1-11) IGF-I (Milner et al., 1996). The pGH(1-11) leader sequence was cleaved from IGF-II using α-LP (Section 2.3.10) (Figure 3.3D). As shown in Figure 3.3D, upon cleavage of folded pGH(1-11) IGF-II two peaks at 12.0 minutes and 15.0 minutes were generated. Consistent with previous data from our laboratory, mature IGF-II has a retention time of approximately 12.0 minutes under these chromatographic conditions (Figure 3.3E), while the 15.0 minute peak represents the cleaved pGH(1-11) leader sequence (Denley, 2004). A single peak representing mature IGF-II was detected by analytical rpHPLC following a preparative chromatography step (Section 2.3.11) and freeze drying.

3.2.2 Characterisation of IGF-II

3.2.2.1 Mass spectrometry and circular dichroism

Mass spectrometry was used to confirm the correct mass of IGF-II (Supplementary Table 2). Circular dichroism was then used to analyse the secondary structure of IGF-II. IGF-II circular dichroism (CD) spectrum was measured in 10 mM acetic acid at a concentration of 0.05 μg/μL (Section 2.4.4.1) (Figure 3.4A). The CD spectrum of IGF-II is consistent with previous studies (Delaine et al., 2007, Blyth et al., 2022) and calculated helical content (Table in Figure 3.4A) presented here is also consistent with published solution structures of IGF-II (Torres et al., 1995).

3.2.2.2 Thermal denaturation studies

Thermal denaturation studies of IGF-II were also performed at a concentration of 0.05 μg/μL in 10 mM acetic acid (Section 2.4.4.1.1) and α-helical content was monitored by circular dichroism at the helix sensitive wavelength 222 nm from 20°C to 90°C. As shown in Figure 3.4B, at the beginning of thermal denaturation (20±0.1°C) IGF-II had a $[\theta]_{222}$ value of -10.5×10^3, consistent with the CD spectrum shown in Figure 3.4A. At completion of thermal denaturation (90±0.1°C) IGF-II had a $[\theta]_{222}$ value of -7.26×10^3. A simple linear regression model was fit to the data (shown in the accompanying table in Figure 3.4B). The rate of change (slope) in $[\theta]_{222}$ for IGF-II was 0.06 $[\theta]_{222}$/°C. Expressed as a percentage change from 20-90±0.1°C, IGF-II has a 30.9% change in ellipticity

demonstrating unfolding of the protein. Shown in Supplementary Figure 3 are the $[\theta]_{222}$ values from 90°C to 20°C.

3.2.2.3 Competition binding assays

IGF-1R and IR-A competition binding assays were performed using europium labelled IGF-II (Eu-IGF-II) and increasing concentrations of competing IGF-II or insulin (Figure 3.5A and B). As expected, IGF-II had an IC_{50} of 0.626 nM for the IGF-1R, consistent with previous reports (Blyth et al., 2022, Surinya et al., 2008). Insulin was found to have a an IC_{50} of 15.9 nM for the IGF-1R, 25.4-fold lower than IGF-II. Conversely, insulin has an IC_{50} of 0.226 nM for the IR-A, 7.43-fold greater than that of IGF-II (IC_{50} of 1.68 nM for the IR-A), which is similar to previous reports using Europium labelled IGF-II (Blyth et al., 2022, Alvino et al., 2009).

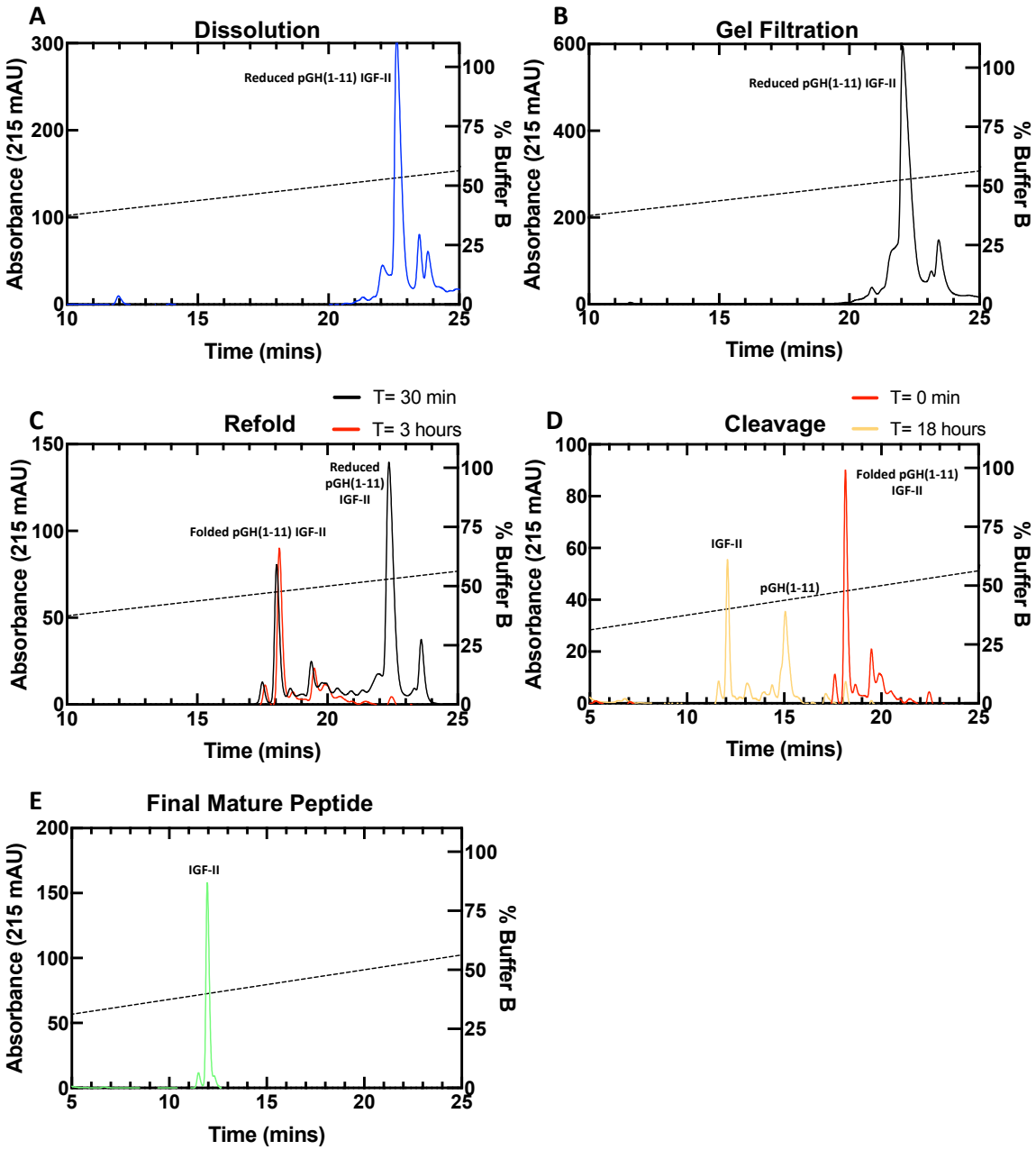
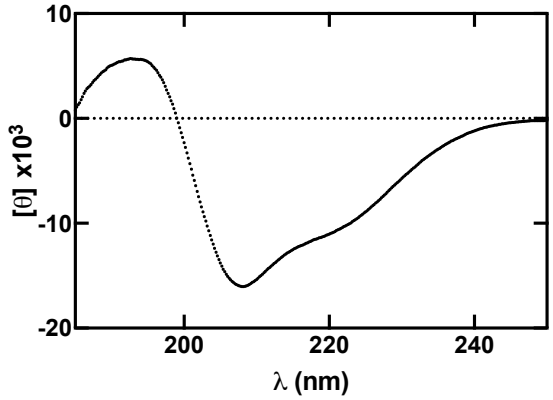


Figure 3.3: Purification of human IGF-II.

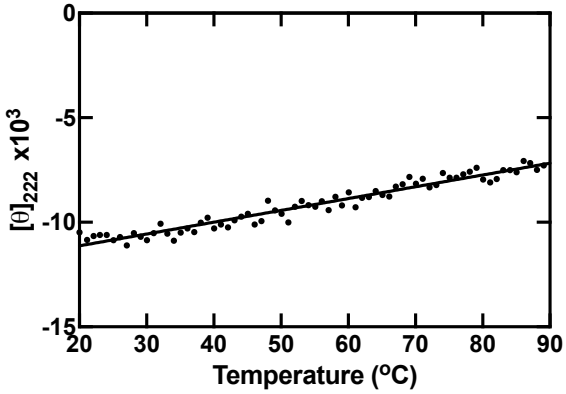
Analytical reverse phase HPLC (using a C4 column and a 25 – 45 % CH_3CN (31.3 – 56.3% buffer B) over 20 minutes (Section 2.3.5)) was used to monitor each stage of protein purification. Shown are the absorbance chromatograms at 215nm (left Y axis). Dotted black line represents the % of buffer B (right Y axis). Samples analysed were (A) dissolution, (B) post gel filtration, (C) the refold T=30 minutes (solid black line) and T=3 hours (red), (D) α-LP cleavage T=0 minutes (red) and T=18 hours (orange) with mature IGF-II at 10.5 min and (E) final mature IGF-II peptide post preparative HPLC purification (green).

A



IGF-II concentration (μg/μL)	Helix	Strand	Turns	Unordered
0.2	8	38	12	42
0.1	23	23	14	40
0.05	41	13	16	30

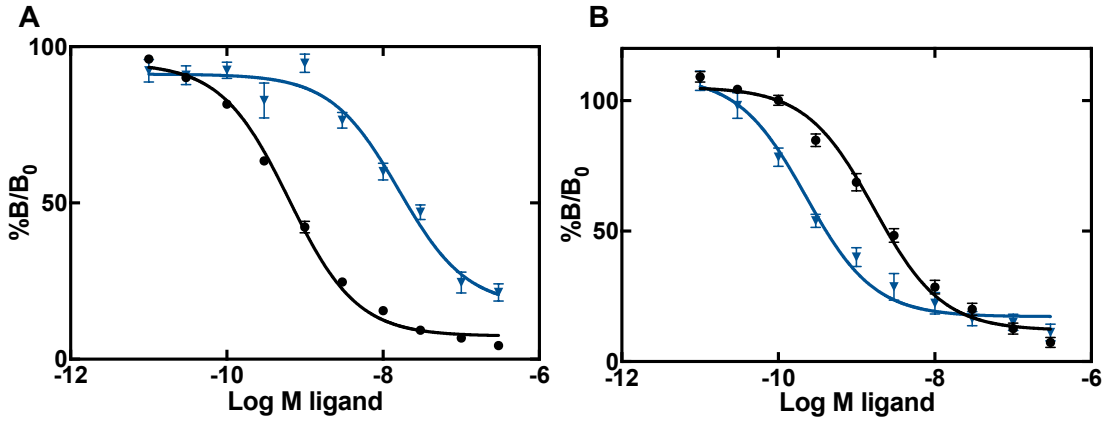
B



Analogue	R^2	Slope	$[\theta]_{222}x10^3$		% change from 20°C to 90°C
			Start (20°C)	End (90°C)	
IGF-II	0.9462	0.06	-10.5	-7.26	30.9

Figure 3.4: Analysis of IGF-II secondary structure and thermal stability using circular dichroism.

(A) Secondary structure analysis of IGF-II was performed using circular dichroism (CD) at 0.05 μg/μL in 10 mM acetic acid, mean residue ellipticity [θ] was measured from 260 - 180 nm (Section 2.4.4.1). Percent helical content shown in the table was performed using Dichroweb and was calculated using the CDSSTR algorithm for deconvolution against the reference protein database set SMP180 (Sreerama and Woody 2000). (B) Thermal stability was monitored by measuring mean residue ellipticity [θ] at 222 nm from 20°C to 90°C (at 1±0.1°C increments) (Section 2.4.4.1.1). Temperature denaturation data were analysed using a linear regression line of best fit. R^2 values, slope and percentage change in helical content from 20°C to 90°C are shown in the table.



Analogue	IC_{50} (nM)	
	IGF-1R	IR-A
IGF-II	0.626	1.68
Insulin	15.9	0.226

Figure 3.5: Competition of IGF-II or insulin with europium-labelled IGF-II for binding the IGF-1R and IR-A.

Immunocaptured IGF-1R (A) or IR-A (B) was incubated with europium-labelled IGF-II (Eu-IGF-II) and increasing concentrations of competing IGF-II or insulin. Time-resolved fluorescence was measured as described in Section 2.4.1. Results are expressed as a percentage of Eu-IGF-II bound in the absence of competing ligand (B/B_0). Data shown are the mean ±S.E.M. n = ≥ 3 independent experiments each with triplicate technical replicates. Error bars are shown when greater than the size of the symbols. Calculated IC_{50} values are shown in the table.

3.2.3 Expression of site 1a IGF-II analogues

Three IGF-II analogues were produced and were characterised to determine which most effectively reduced receptor binding affinity whilst maintaining structural integrity:

1. F26A IGF-II
2. F26A Y27L F28L IGF-II
3. F26G Y27G F28G IGF-II

Each IGF-II analogue was expressed with a pGH(1-11) fusion partner as described above for IGF-II (Section 1.2.1.1 and Supplementary Figures 1 and 2), and inclusion bodies were released from *E. coli* by homogenisation (section 2.3.7).

3.2.4 Production of F26A IGF-II

3.2.4.1 Expression Dissolution and Refolding of F26A IGF-II

pGH(1-11) F26A IGF-II partially pure inclusion bodies were solubilised (Figure 3.6A) and separated by size exclusion chromatography (Figure 3.6B) (Section 2.3.8). Refolding was performed under the same conditions used for IGF-II (Section 3.2.1.2 and Section 2.3.9). As shown in Figure 3.6C at T=0 minutes reduced pGH(1-11) F26A IGF-II had a retention time of 21.0 minutes. After refolding for 3 hours, three folding states were observed (17.0, 18.0 and 18.5 minutes). Based on the folding of native pGH(1-11) IGF-II (Section 1.2.1.2) it was predicted that the peak at 17.0 minutes represented fully folded pGH(1-11) F26A IGF-II and the peaks at 18.0 and 18.5 minutes represented intermediate folding states. As the equivalent residue FB24 in insulin is well documented to have a role in the formation of the hydrophobic core the presence of multiple intermediate peaks in the F26A IGF-II refold was not unexpected.

3.2.4.2 Cleavage and purification of F26A IGF-II

At 18 hours of cleavage (Figure 3.6D) the pGH(1-11) leader sequence was observed at 14.0 minutes. It was expected that the 10.9 minute species was mature F26A IGF-II based on experience in the production of IGF-II analogues in our laboratory. However, several other species were also observed with retention times between 11.6 and 14 minutes. Based on refolding of pGH(1-11) IGF-II (Lien et al., 2001) these species are likely to be intermediate folding states of F26A IGF-II. The 10.8 minute peak and two potential intermediate folding states (11.7 and 13.3 minutes were captured during preparative HPLC and freeze dried (as described for IGF-II in Section 3.2.1.2) for characterisation (purified fractions shown in Figure 3.6E).

3.2.5 Characterisation of F26A IGF-II Purified Fractions

3.2.5.1 Mass spectrometry and circular dichroism

MALDI mass spectrometry (Section 2.3.13) confirmed that the masses of each isolated fraction matched the expected mass of F26A IGF-II (Supplementary Table 1). Secondary structure analysis was then performed on each isolated fraction using CD analysis (Section 2.4.4.1). Shown in Figure 3.7A the CD profile of F26A IGF-II 10.8 minute fraction most closely resembled that of native IGF-II and the calculated helical content was also similar (as shown in the corresponding table to Figure 3.7A). Notably, the CD profiles of the 11.7 and 13.3 minute fractions had a slight shift in their profiles to the left in comparison to IGF-II (Figure 3.7A), which was most noticeable between 180 -210 nm. In addition, the calculated helical content of these isolated fractions was markedly lower than that of both IGF-II and the F26A IGF-II 10.8 minute fraction (corresponding table of Figure 3.7A). The observation that F26A IGF-II 11.7 and 13.3 minute fractions had markedly lower helical content is consistent with the expectation that these fractions represent folding intermediates of F26A IGF-II.

3.2.5.2 Thermal denaturation studies

Thermal denaturation studies were performed for each of the purified fractions (section 2.3.13). The F26A IGF-II 10.8 minute fraction started at 20°C with a $[\theta]_{222}$ of -9.29×10^3, consistent with the corresponding CD spectrum shown in Figure 3.7A. The thermal denaturation of F26A IGF-II 10.8 minute fraction followed a similar pattern to IGF-II (shown in Figure 3.7B) and at completion (90±0.1°C) both IGF-II and F26A IGF-II 10.8 minute peak had similar $[\theta]_{222}$ values (-7.26×10^3 and -6.94×10^3 respectively), demonstrating unfolding of both proteins. Both F26A IGF-II fractions 11.7 and 13.3 exhibited no change in ellipticity due to their low starting helical content, consistent with these species being folding intermediates. Shown in Supplementary Figure 4 are the $[\theta]_{222}$ values from 90°C to 20°C.

3.2.5.3 Competition binding assays

The ability of each of the F26A IGF-II isolated fractions to bind the IGF-1R was next tested using competition binding assays (Section 2.4.1). As indicated in Figure 3.8, F26A IGF-II fractions 11.7 and 13.3 were performed only as n=1 as both were poor at competing for binding to the IGF-1R, further confirming that these fractions were folding intermediates (<0.1% relative binding affinity compared to IGF-II; Figure 3.8).

Conversely, F26A IGF-II 10.8 minute fraction retained 11.5% binding affinity for the IGF-1R compared to that of IGF-II (an 8.66 fold decrease in affinity).

Competition assays for the IR-A (Figure 3.16B) were also performed using Eu-IGF-II. Based on the results described above for each isolated fraction, only the correctly folded F26A IGF-II (fraction 10.8) was tested on the IR-A. It was found that F26A IGF-II 10.8 minute fraction retained an affinity of <0.1% for the IR-A compared to human IGF-II.

In summary, these data suggest that the F26A IGF-II 10.8 minute fraction represents the correctly folded state with correct mass, highly similar secondary structure to native IGF-II and relatively high affinity for the IGF-1R compared to IGF-II. F26A IGF-II fractions 11.7 and 13.3, whilst having the correct mass, had marked reduction of helical content and bound with very low affinity to the IGF-1R, confirming that they were most likely intermediate folding states of F26A IGF-II.

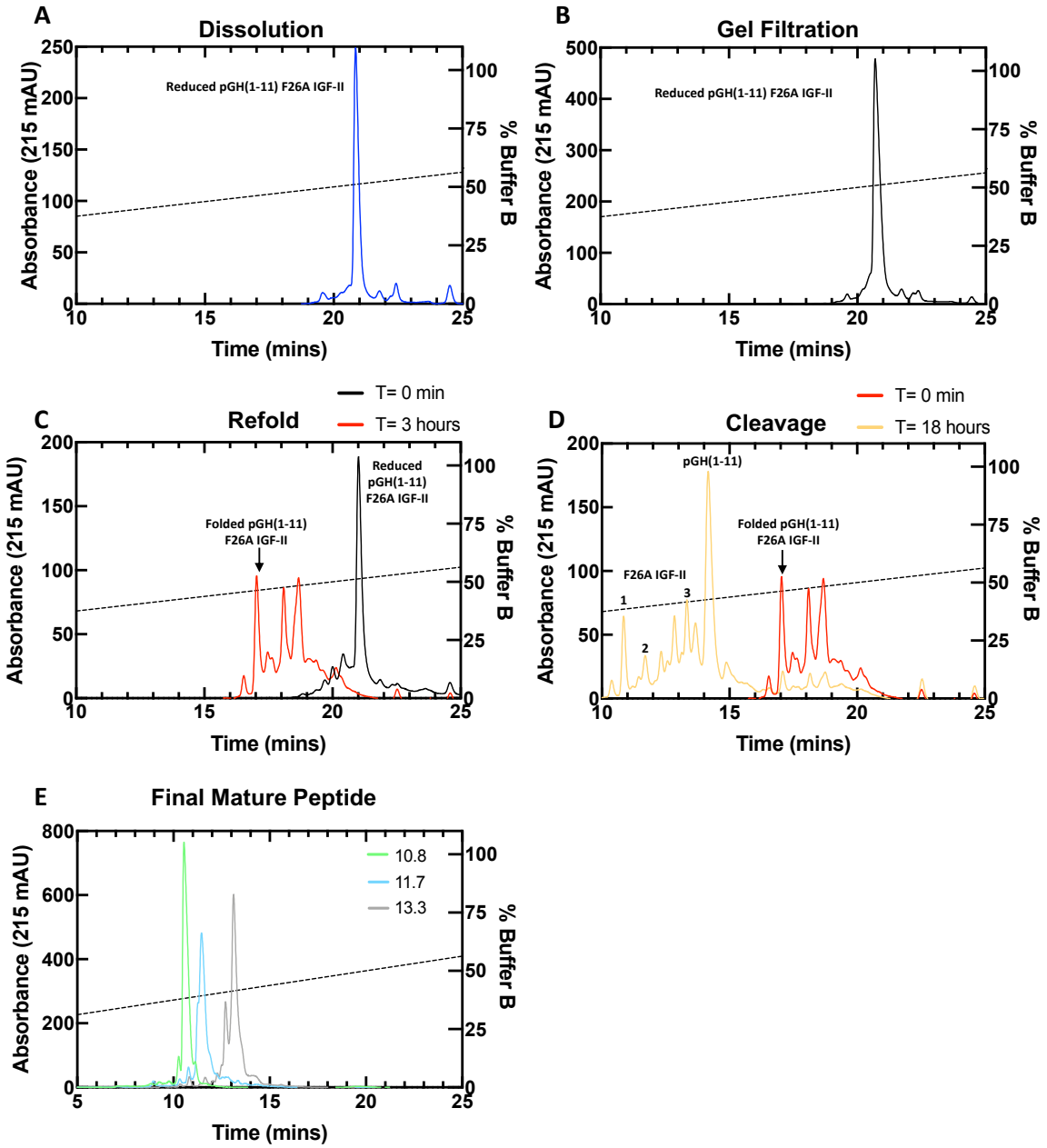
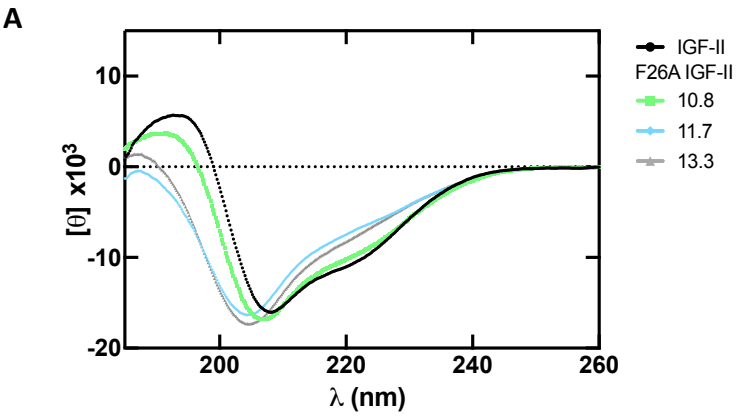
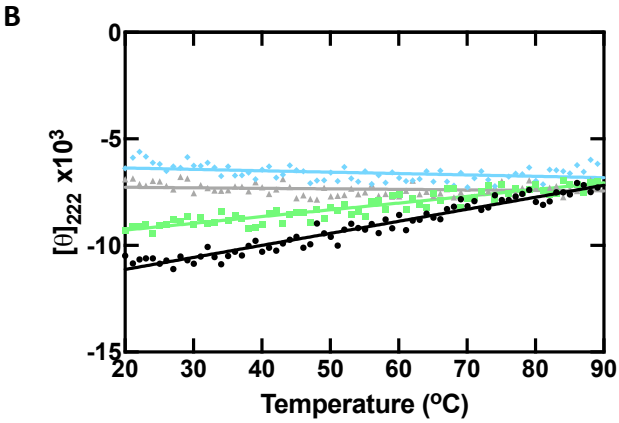


Figure 3.6: Purification of F26A IGF-II.

Analytical reverse phase HPLC (using a C4 column and a 25 – 45 % CH_3CN (31.3 – 56.3% buffer B) over 20 minutes (Section 2.3.5)) was used to monitor each stage of protein purification. Shown are the absorbance chromatograms at 215nm (left Y axis). Dotted black line represents the % of buffer B (right Y axis). Samples analysed were (A) dissolution, (B) post gel filtration, (C) the refold T=0 minutes (solid black line) and T=3 hours (red), (D) α-LP cleavage T=0 minutes (red) and T=18 hours (orange). (E) Three purified fractions of F26A IGF-II (numbered 1-3 in D) with retention times of 10.8 minutes (green, peak 1 in D), 11.7 minutes (blue, peak 2 in D) and 13.3 minutes (grey, peak 3 in D).



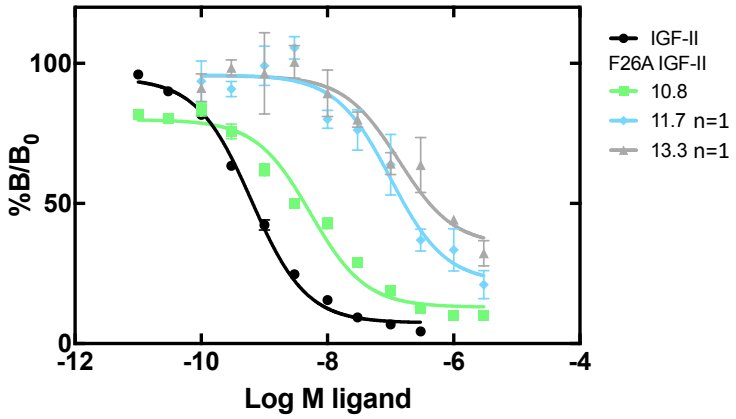
Analogue	Helix	Strand	Turns	Unordered
IGF-II	41	13	16	30
F26A IGF-II 10.8 min	37	13	16	35
F26A IGF-II 11.7 min	22	22	14	41
F26A IGF-II 13.3 min	22	20	16	43



Analogue	R^2	Slope	$[\theta]_{222}x10^3$		% change from 20°C to 90°C
			20°C	90°C	
IGF-II	0.9462	0.06	-10.5	-7.26	30.9
F26A IGF-II 10.8 min	0.8621	0.03	-9.29	-6.94	25.3
F26A IGF-II 11.7 min	0.1292	-0.001	-6.38	-6.09	4.55
F26A IGF-II 13.3 min	0.03127	-0.002	-6.89	-6.56	4.79

Figure 3.7: Analysis of F26A IGF-II secondary structure and thermal stability using circular dichroism.

(A) Secondary structure analysis of purified F26A IGF-II samples was performed using circular dichroism (CD) at 0.05 μg/μL in 10 mM acetic acid, mean residue ellipticity [θ] was measured from 260 - 180 nm (Section 2.4.4.1). Percent helical content shown in the table was derived using Dichroweb and was calculated using the CDSSTR algorithm for deconvolution against the reference protein database set SMP180 (Sreerama and Woody 2000). (B) Thermal stability was monitored by measuring mean residue ellipticity [θ] at 222 nm from 20°C to 90°C (at 1±0.1°C increments)(Section 2.4.4.1.1). Temperature denaturation data were analysed using a linear regression line of best fit. R^2 values, slope and percentage change in helical content from 20°C to 90°C are shown in the table.



Analogue	IC_{50} (nM)	Relative Binding affinity (%)	Fold Change
IGF-II	0.626	100	-
F26A IGF-II 10.8 min	5.42	11.5	8.66
F26A IGF-II 11.7 min	>100	ND	ND
F26A IGF-II 13.3 min	>100	ND	ND

Figure 3.8: Competition of F26A IGF-II and europium-labelled IGF-II for binding the IGF-1R.

Immunocaptured IGF-1R was incubated with europium-labelled IGF-II (Eu-IGF-II) and increasing concentrations of competing F26A IGF-II. Time-resolved fluorescence was measured as described in section 2.4.1. Results are expressed as a percentage of Eu-IGF-II bound in the absence of competing ligand (B/B_0). Data shown are the mean ±S.E.M. n = ≥ 3 independent experiments each with triplicate technical replicates (unless otherwise indicated on the graph). Error bars are shown when greater than the size of the symbols. Calculated IC_{50} values are shown in the table.

3.2.6 Production of F26A Y27L F28L IGF-II

3.2.6.1 Expression, Dissolution and Refolding of F26A Y27L F28L IGF-II

F26A Y27L F28L IGF-II inclusion bodies were solubilised (Figure 3.9A), separated by size exclusion chromatography (Figure 3.9B) and refolded under the same conditions as described for IGF-II in Section 3.2.1. As shown in Figure 3.9C at T=0 minutes reduced pGH(1-11) F26A Y27L F28L IGF-II had a retention time of 21.2 minutes. After refolding for 30 minutes there were three major peaks. As described for F26A it was expected that the 17.4 minute peak represents correctly folded pGH(1-11) F26A Y27L F28L IGF-II and the peaks to the right, including 18.2 and 18.8 minute peaks, represented folding intermediates as described for F26A IGF-II (section 3.2.4.2)

3.2.6.2 cleavage and purification of F26A Y27L F28L IGF-II

At 10 hours of cleavage (Figure 3.9D) the 11.1 minute peak was expected to be correctly folded mature F26A Y27L F28L IGF-II. It is interesting to note that despite more changes in the F26A Y27L F28L IGF-II analogue compared to F26A IGF-II, there were fewer folding intermediates of lower abundance in the final cleavage sample of F26A Y27L F28L IGF-II (Figure 3.9D) compared to F26A IGF-II (Figure 3.6D). The 11.1 minute peak was captured during preparative HPLC (Section 2.3.11) and freeze dried (Section 2.3.12) for characterisation (purified fractions shown in Figure 3.9E).

3.2.7 Characterisation of F26A Y27L F28L IGF-II

3.2.7.1 Mass spectrometry and circular dichroism

MALDI mass spectrometry confirmed that the mass of the purified 11.1 minute peak matched the expected mass of F26A Y27L F28L IGF-II (Supplementary Table 1). The CD profile of F27A Y27L F28L IGF-II (Figure 3.10A) resembled that of native IGF-II. Although there was a slight shift of F27A Y27L F28L IGF-II to the left between 180 – 210 nm. The calculated helical content of F26A Y27L F28L IGF-II and IGF-II were also highly similar (40 and 41% respectively, as shown in the corresponding table to Figure 3.10A). This suggests that the 11.1 minute peak most likely represents correctly folded peptide and that the slight shift in spectrum may be attributed to the mutations introduced into this analogue.

3.2.7.2 Thermal denaturation studies

Thermal denaturation studies were performed for F26A Y27L F28L IGF-II as stated above for IGF-II (Section 3.2.2.2). As shown in Figure 3.10B and corresponding table, F26A Y27L F28L IGF-II and IGF-II began with $[\theta]_{222}$ values of -9.43 and -10.5,

respectively, and followed similar denaturation patterns, ending at 90±0.1°C with identical $[\theta]_{222}$ values (-7.26), demonstrating similar denaturation properties. Shown in Supplementary Figure 5 are the $[\theta]_{222}$ values from 90°C to 20°C.

3.2.7.3 Competition binding assays

The ability of F27A Y27L F28L IGF-II to bind the IGF-1R receptor was tested using competition binding assays (Section 2.4.1). In comparison to native IGF-II, F27A Y27L F28L IGF-II was found to retain <0.1% affinity for the IGF-1R compared to that of native IGF-II (Figure 3.11 and corresponding Table). Competition assays for the IR-A (Figure 3.16B) using Eu-IGF-II found that F26A Y27L F28L IGF-II had an affinity <0.01% compared to human IGF-II.

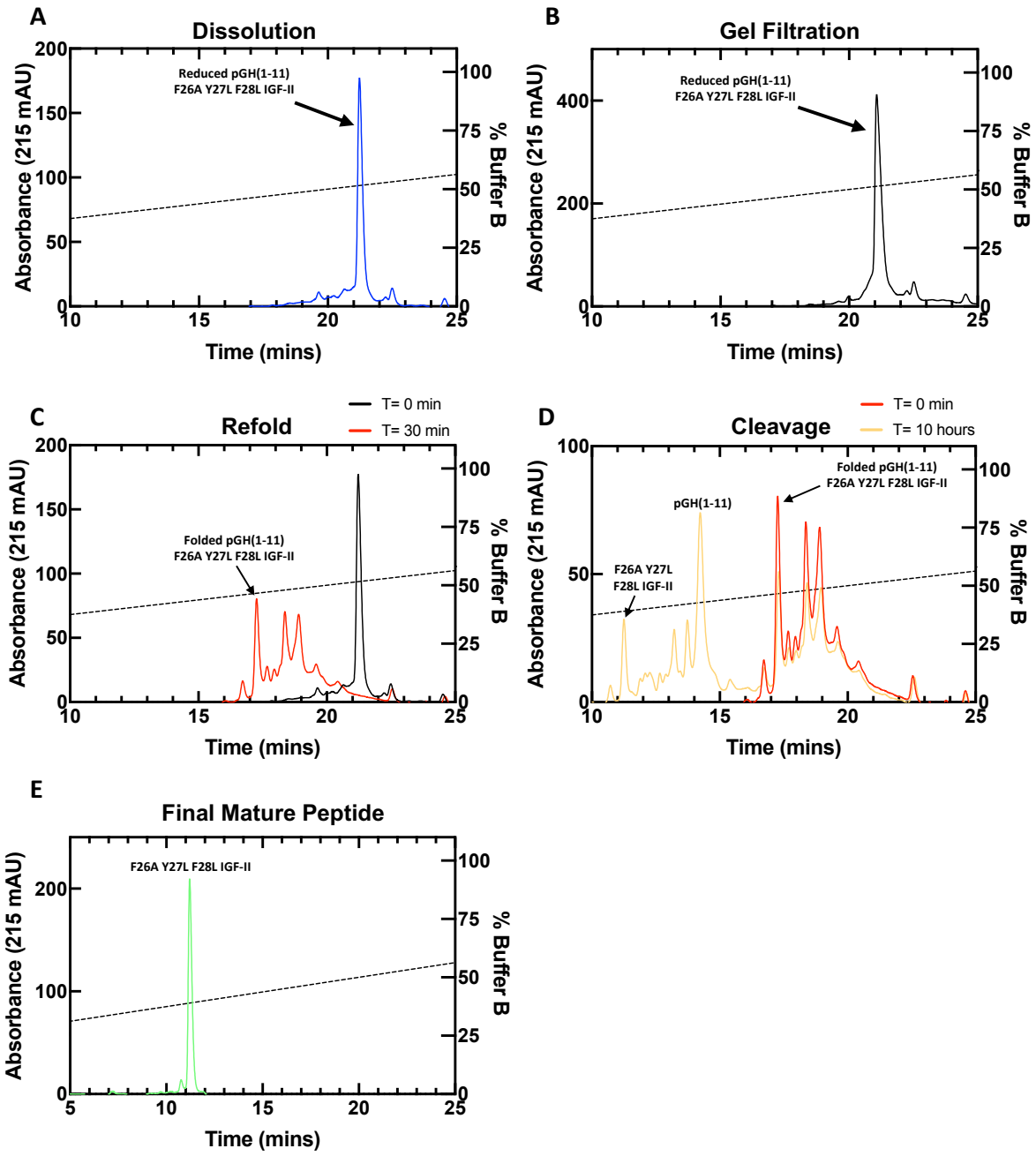
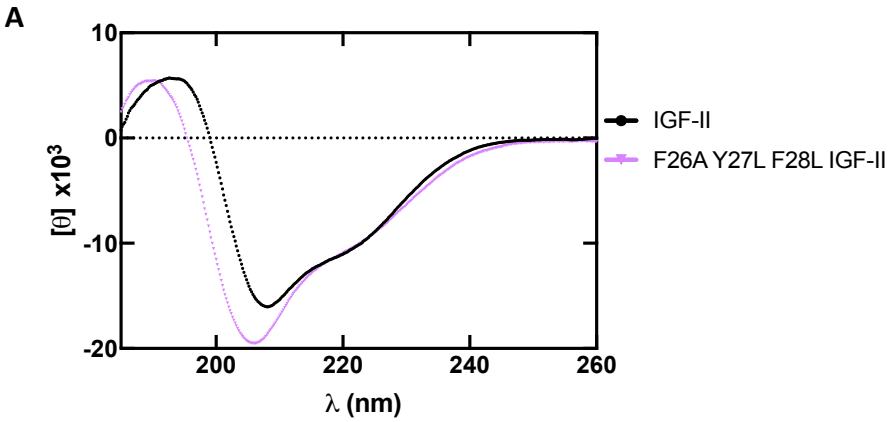
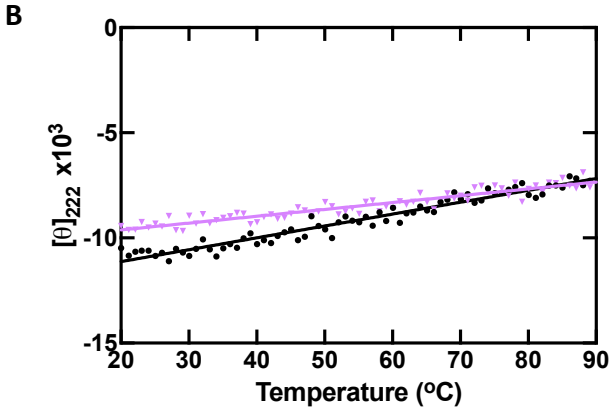


Figure 3.9: Purification of F26A Y27L F28L IGF-II.

Analytical reverse phase HPLC (using a C4 column and a 25 – 45 % CH_3CN (31.3 – 56.3% buffer B) over 20 minutes (Section 2.3.5)) was used to monitor each stage of protein purification. Shown are the absorbance chromatograms at 215 nm (left Y axis). Dotted black line represents the % of buffer B (right Y axis). Samples analysed were (A) dissolution, (B) post gel filtration, (C) the refold T=0 minutes (solid black line) and T=30 minutes (red), (D) α-LP cleavage T=0 minutes (red) and T=10 hours (orange) with mature F26A Y27L F28L IGF-II at 11.1 minutes and (E) final mature F26A Y27L F28L IGF-II peptide post preparative HPLC purification (green).



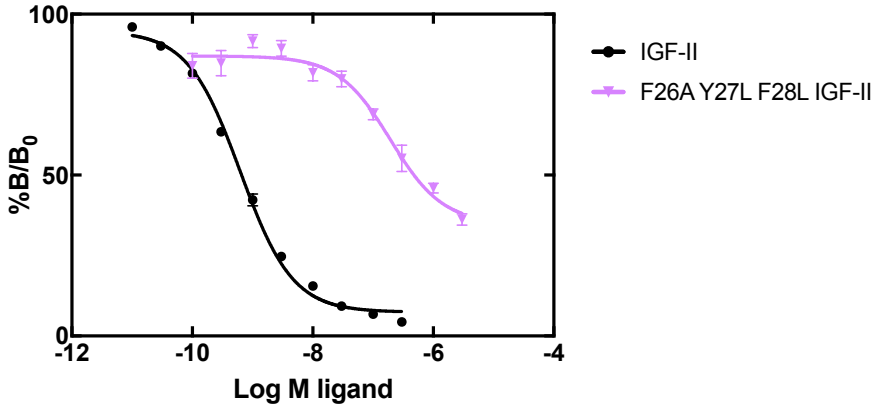
Analogue	Helix	Strand	Turns	Unordered
IGF-II	41	13	16	30
F26A Y27L F28L IGF-II	40	12	15	33



Analogue	R^2	Slope	$[\theta]_{222}x10^3$		% change from 20°C to 90°C
			20°C	90°C	
IGF-II	0.9462	0.06	-10.5	-7.26	30.9
F26A Y27L F28L IGF-II	0.8842	0.03	-9.43	-7.26	23.0

Figure 3.10: Analysis of F26A Y27L F28L IGF-II secondary structure and thermal stability using circular dichroism.

(A) Secondary structure analysis of F26A Y27L F28L IGF-II was performed using circular dichroism at 0.05 μg/μL in 10 mM acetic acid, mean residue ellipticity [θ] was measured from 260 - 180 nm (Section 2.4.4.1). Percent helical content shown in the table was derived using Dichroweb and was calculated using the CDSSTR algorithm for deconvolution against the reference protein database set SMP180 (Sreerama and Woody 2000). (B) Thermal stability was monitored by measuring mean residue ellipticity [θ] at 222 nm from 20°C to 90°C (at 1±0.1°C increments) (Section 2.4.4.1.1). Temperature denaturation data were analysed using a linear regression line of best fit. R^2 values, slope and percentage change in helical content from 20°C to 90°C are shown in the table.



Analogue	IC_{50} (nM)	Relative Binding affinity (%)	Fold Change
IGF-II	0.626	100	-
F26A Y27L F28L IGF-II	>100	ND	ND

Figure 3.11: Competition of F26A Y27L F28L IGF-II and europium-labelled IGF-II for binding the IGF-1R.

Immunocaptured IGF-1R was incubated with europium-labelled IGF-II (Eu-IGF-II) and increasing concentrations of competing F26A Y27L F28L IGF-II. Time-resolved fluorescence was measured as described in Section 2.4.1. Results are expressed as a percentage of Eu-IGF-II bound in the absence of competing ligand (B/B_0). Data shown are the mean ±S.E.M. n = ≥ 3 independent experiments each with triplicate technical replicates. Error bars are shown when greater than the size of the symbols. Calculated IC_{50} values are shown in the table.

3.2.8 Production of F26G Y27G F28G IGF-II

Reduced pGH(1-11) F26G Y27G F28G IGF-II, post dissolution (Figure 3.12A) and post gel filtration (Section 2.3.8)(Figure 3.12B) had a retention time of 19.9 minutes. As shown in Figure 3.12C there was a shift in reduced pGH(1-11) F26G Y27G F28G IGF-II from 19.9 minutes at T=0 to several peaks at T=20 hours. The first major peak at 16.1 minutes was expected to be fully folded pGH(1-11) F26G Y27G F28G IGF-II while peaks to the right were expected to be folding intermediates. After 18 hours of cleavage, there were several cleaved folding intermediates of F26G Y27G F28G IGF-II. Based on the retention times of other site 1a IGF-II analogues described above it was expected that the 9.6 minute peak (labelled as '3' in Figure 3.12D) was the fully folded form of mature F26G Y27G F28G IGF-II. Interestingly, two major peaks were also observed at 7.1 and 8.2 minutes. While these peaks have a much faster retention time than expected for a mature fully folded IGF-II analogue, they were purified during preparative HPLC (Section 2.3.11) along with the 9.6 minute peak for further characterisation (purified fractions are shown in Figure 3.12E).

3.2.9 Characterisation of F26G Y27G F28G IGF-II purified fractions

3.2.9.1 Mass spectrometry and circular dichroism

The correct mass of F26G Y27G F28G IGF-II was confirmed using MALDI mass spectrometry (Supplementary Table 1). CD studies showed significantly reduced α-helical content for all isolated peaks (Figure 3.13A). Secondary structure predictions indicated that F26G Y27G F28G IGF-II peaks 7.1 and 8.2 helical content had been reduced to 5 and 8% respectively, compared to IGF-II. This was consistent with a shift in the minimum at 208 nm (for native IGF-II) to a lower wavelength for these isolated peaks. However, F26G Y27G F28G IGF-II peak 9.6 retained helical content of approximately 20%. There is a still a notable difference between the minimum for IGF-II compared to F26G Y27G F28G IGF-II peak 9.6, albeit smaller than that of the other isolated peaks.

3.2.9.2 Thermal denaturation studies

Due to the near total loss of helical content in F26G Y27G F28G IGF-II peaks 7.1 and 8.2 thermal denaturation studies were not performed on these analogues. Thermal denaturation studies (Section 2.4.4.1.1) were performed on F26G Y27G F28G IGF-II peak 9.6 (Figure 3.13B). Consistent with CD studies, this analogue began with $[\theta]_{222}$ value of -5.64 at 20$\pm$0.1°C, compared to -10.5 for that of IGF-II. F26G Y27G F28G IGF-

II peak 9.6 essentially exhibited no change in ellipticity due to its low starting helical content. Shown in Supplementary Figure 6 are the $[\theta]_{222}$ values from 90°C to 20°C.

3.2.9.3 Competition binding assays

Consistent with the above finding that F26G Y27G F28G IGF-II peaks 7.1 and 8.2 had significantly lower helical content, competition binding assays (Section 2.4.1) demonstrate that these isolated fractions did not bind the IGF-1R, even at very high concentrations of up to 3 μM (Figure 3.14). However, it was found that F26G Y27G F28G IGF-II peak 9.6 did bind weakly to the IGF-1R at high concentrations, retaining <0.06% affinity compared to that of IGF-II.

These data reveal that isolated peaks 7.1 and 8.2 are likely folding intermediates of F26G Y27G F28G IGF-II with confirmed low helical content using CD studies and essentially no affinity for the IGF-1R up to a concentration of 3 μM. Based on the fact that it bound IGF-1R to a certain extent, peak 9.6 may represent the correctly folded structure of F26G Y27G F28G IGF-II, although further analyses would be required to definitively prove this. Due to the major changes made to the motif essential for maintaining structure and binding it is not surprising that α-helical content remains low and binding affinity for the IGF-1R is weak.

Competition assays for the IR-A (Figure 3.16B) found that F26G Y27G F28G IGF-II had no detectable affinity to a concentration of 3 μM.

3.2.10 Comparison of site 1a minimised IGF-II analogues

Three site 1a minimised IGF-II analogues were purified and characterised to determine the effect of the introduced mutations on both structure and function. Of the three analogues, F26A IGF-II retained the highest affinity for the IGF-1R, indicating preservation of the core structure and that Y27 and F28 residues of the FYF motif still contributed to binding. Interestingly, this analogue has a marked reduction in affinity for the IR-A (Figure 3.16B and accompanying table).

F26A Y27L F28L IGF-II appeared to retain the most similar secondary structure to human IGF-II, (Figure 3.15A and accompanying Table), while having <0.1% reduction in affinity for the IGF-1R and even lower affinity for the IR-A, compared to human IGF-II. This suggests that this combined mutation of the IGF-II FYF motif effectively preserves the structure of the analogue and knocks out the binding of this motif to the receptors. Finally, while F26G Y27G F28G IGF-II had an even lower binding affinity for the IGF-1R,

the CD studies suggest that these introduced mutations drastically altered the structure of the analogue.

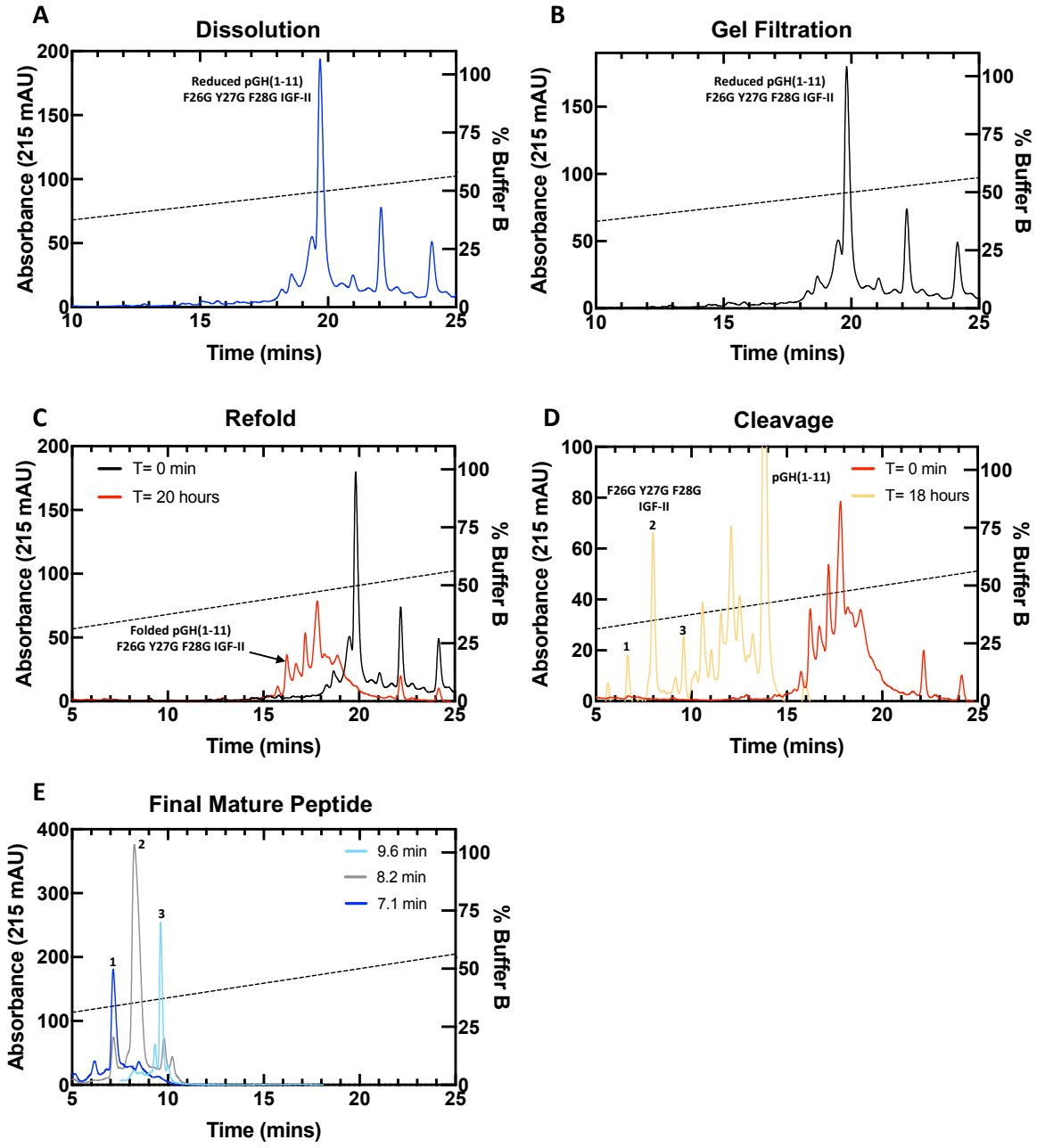
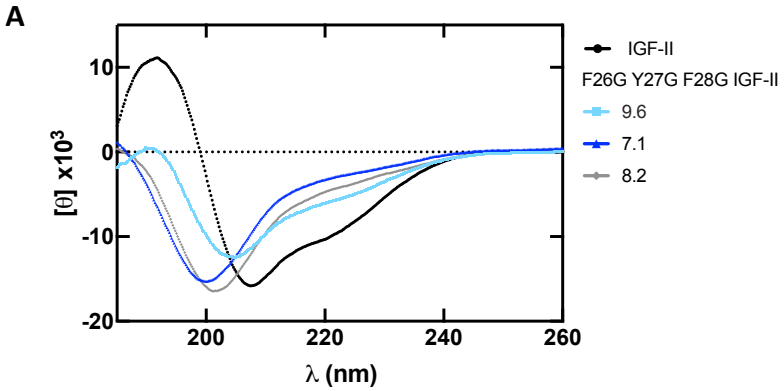
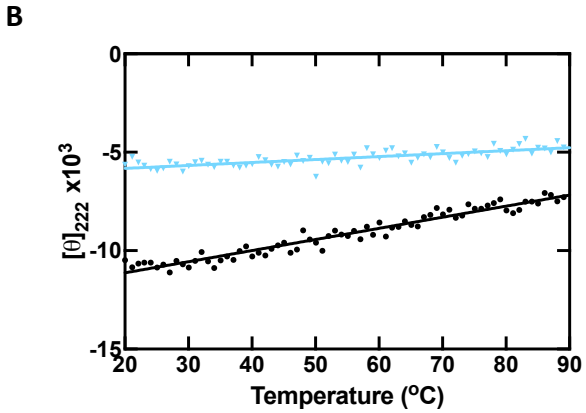


Figure 3.12: Purification of F26G Y27G F28G IGF-II.

Analytical reverse phase HPLC (using a C4 column and a 25 – 45 % CH_3CN (31.3 – 56.3% buffer B) over 20 minutes (Section 2.3.5)) was used to monitor each stage of protein purification. Shown are the absorbance chromatograms at 215 nm (left Y axis). Dotted black line represents the % of buffer B (right Y axis). (A) Dissolution, (B) post gel filtration, (C) the refold T=0 minutes (solid black line) and T=20 hours (red), (D) α-LP cleavage T=0 minutes (red) and T=18 hours (orange). (E) Three purified fractions of F26G Y27G F28G IGF-II (numbered 1-3 in D) with retention times of 7.1 minutes (green, peak 1 in D), 8.2 minutes (blue, peak 2 in D) and 9.6 minutes (grey, peak 3 in D).



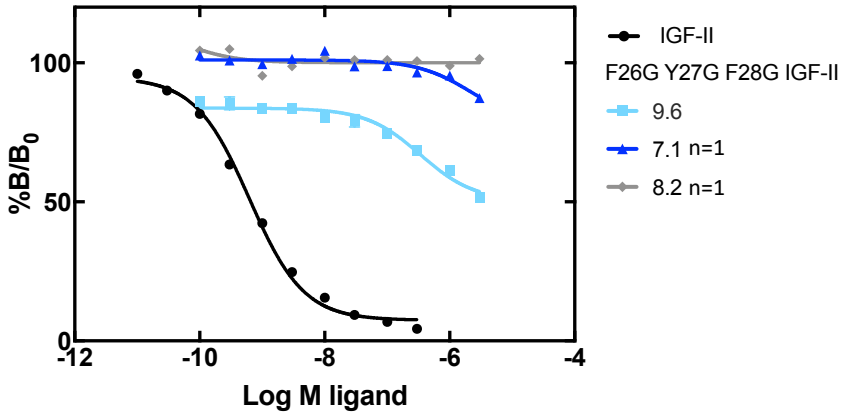
Analogue	Helix	Strand	Turns	Unordered
IGF-II	41	13	16	30
F26G Y27G F28G IGF-II 7.1 min	5	33	12	49
F26G Y27G F28G IGF-II 8.2 min	8	30	13	49
F26G Y27G F28G IGF-II 9.6 min	20	22	15	44



Analogue	R^2	Slope	$[\theta]_{222}x10^3$		% change from 20°C to 90°C
			20°C	90°C	
IGF-II	0.9462	0.06	-10.5	-7.26	30.9
F26G Y27G F28G IGF-II 9.6 min	0.6034	0.02	-5.64	-4.54	19.5

Figure 3.13: Analysis of F26G Y27G F28G IGF-II secondary structure and thermal stability using circular dichroism.

(A) Secondary structure analysis of F26G Y27G F28G IGF-II was performed using circular dichroism at 0.05 μg/μL in 10 mM acetic acid, mean residue ellipticity [θ] was measured from 260 - 180 nm (Section 2.4.4.1). Percent helical content shown in the table was derived using Dichroweb and was calculated using the CDSSTR algorithm for deconvolution against the reference protein database set SMP180 (Sreerama and Woody 2000). (B) Thermal stability was monitored by measuring mean residue ellipticity [θ] at 222 nm from 20°C to 90°C (at 1±0.1°C increments) (Section 2.4.4.1.1). Temperature denaturation data were analysed using a linear regression line of best fit. R^2 values, slope and percentage change in helical content from 20°C to 90°C are shown in the table.

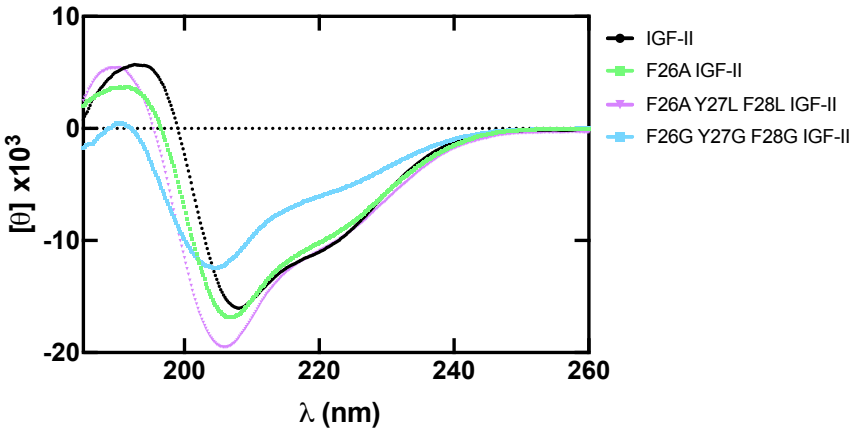


Analogue	IC_{50} (nM)	Relative Binding affinity (%)	Fold Change
IGF-II	0.626	100	-
F26G Y27G F28G IGF-II 7.1 min	ND	ND	ND
F26G Y27G F28G IGF-II 8.2 min	ND	ND	ND
F26G Y27G F28G IGF-II 9.6 min	>1000	ND	ND

Figure 3.14: Competition of F26G Y27G F28G IGF-II and europium-labelled IGF-II for binding the IGF-1R.

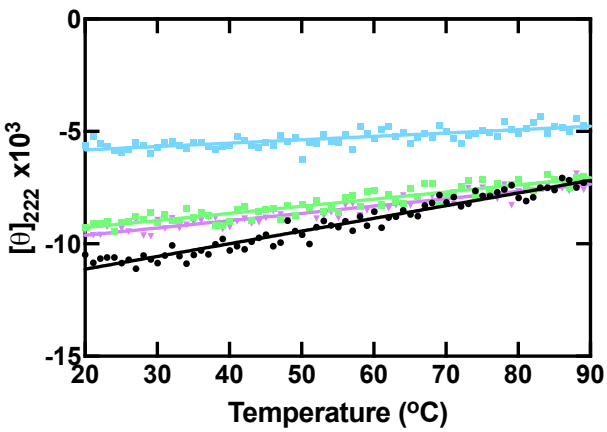
Immunocaptured IGF-1R was incubated with europium-labelled IGF-II (Eu-IGF-II) and increasing concentrations of competing F26G Y27G F28G IGF-II. Time-resolved fluorescence was measured as described in Section 2.4.1. Results are expressed as a percentage of Eu-IGF-II bound in the absence of competing ligand (B/B_0). Data shown are the mean ±S.E.M. n = ≥ 3 independent experiments each with triplicate technical replicates (unless otherwise indicated on the graph). Error bars are shown when greater than the size of the symbols. Calculated IC_{50} values are shown in the table.

A



Analogue	Helix	Strand	Turns	Unordered
IGF-II	41	13	16	30
F26A IGF-II	37	13	16	35
F26A Y27L F28L IGF-II	40	12	15	33
F26G Y27G F28G IGF-II	20	22	15	44

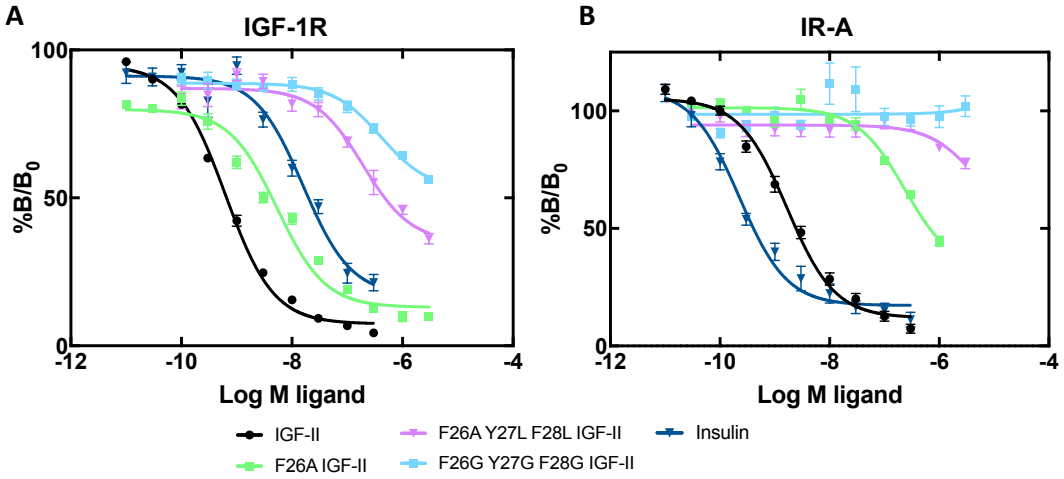
B



Analogue	R^2	Slope	$[\theta]_{222}x10^3$		% change from 20°C to 90°C
			20°C	90°C	
IGF-II	0.9462	0.06	-10.5	-7.26	30.9
F26A IGF-II	0.8621	0.03	-9.29	-6.94	25.3
F26A Y27L F28L IGF-II	0.8842	0.03	-9.43	-7.26	23.0
F26G Y27G F28G IGF-II	0.6034	0.02	-5.64	-4.54	19.5

Figure 3.15: Secondary structure analysis and thermal stability of IGF-II site 1a minimised IGF-II analogues.

(A) Secondary structure analysis of IGF-II site 1a minimised IGF-II analogues was performed using circular dichroism at 0.05 μg/μL in 10 mM acetic acid, mean residue ellipticity [θ] was measured from 260 - 180 nm (Section 2.4.4.1). Percent helical content shown in the table was derived using Dichroweb and was calculated using the CDSSTR algorithm for deconvolution against the reference protein database set SMP180 (Sreerama and Woody 2000). (B) Thermal stability was monitored by measuring mean residue ellipticity [θ] at 222 nm from 20°C to 90°C (at 1±0.1°C increments) (Section 2.4.4.1.1). Temperature denaturation data were analysed using a linear regression line of best fit. R^2 values, slope and percentage change in helical content from 20°C to 90°C are shown in the table.



Analogue	IGF-1R			IR-A		
	IC_{50} (nM)	95% CI (nM)	Relative binding affinity (% of IGF-II)	IC_{50} (nM)	95% CI (nM)	Relative binding affinity (% of IGF-II)
IGF-II	0.626	0.568-0.689	100	1.68	1.36-2.09	100
Insulin	15.9	10.7-23.7	3.94	0.226	0.156-0.329	743
F26A IGF-II	5.42	4.11-7.15	11.5	>100	ND	ND
F26A Y27L F28L IGF-II	>100	ND	ND	>1000	ND	ND
F26G Y27G F28G IGF-II	>1000	ND	ND	ND	ND	ND

Figure 3.16: Competition of site 1a minimised IGF-II analogues and europium-labelled IGF-II for the IGF-1R and IR-A.

Immunocaptured (A) IGF-1R or (B) IR-A was incubated with europium-labelled IGF-II (Eu-IGF-II) and increasing concentrations of competing IGF-II, insulin or IGF-II analogue. Time-resolved fluorescence was measured as described in section (Section 2.4.1). Results are expressed as a percentage of Eu-IGF-II bound in the absence of competing ligand (B/B_0). Data shown are the mean ±S.E.M. n = ≥ 3 independent experiments each with triplicate technical replicates. Error bars are shown when greater than the size of the symbols. Calculated IC_{50} values, 95% confidence interval (95% CI) relative binding affinity relative to human IGF-II are shown in the table.

3.3 Discussion

Contact of insulin, IGF-I and IGF-II at site 1a is essential for high affinity binding to their cognate receptors. A key element of this site 1a interaction is the displacement of the C-terminal segment of the B-domain from the core of the ligand and the 'anchoring' to the receptor through contact of insulin FB24 of the FFY motif (Menting et al., 2014) (F26 for IGF-II and F23 for IGF-I) to the L1 domain and αCT′. Insulin residues FB25 and YB26 of this motif also contact the L1 domain and the αCT′ (Kristensen et al., 1997, Kobayashi et al., 1984, Inouye et al., 1981, Chen et al., 2000, Pandyarajan et al., 2016, Nakagawa and Tager, 1986). This aromatic motif is highly conserved among vertebrate insulin and IGFs (Conlon, 2001, Mirmira et al., 1991, Conlon, 2000) but it is not conserved in many invertebrate species (Gronke et al., 2010, Li et al., 2020, Ahorukomeye et al., 2019) and in some cases it is totally absent. An example of this is seen in cone snail venom insulins which lack the C-terminal segment of the B-domain, including the key FFY motif (Ahorukomeye et al., 2019). These venom insulins have been shown to engage with the receptor through different contacts and still activate the human insulin receptor with high affinity (Menting et al., 2016).

Here we sought to investigate if the high affinity site 1a interaction of IGF-II could be removed through mutation of the FYF motif to allow for alternative modes of receptor engagement to be investigated. We designed three site 1a minimised IGF-II analogues; F26A IGF-II, F26A Y27L F28L IGF-II and F26G Y27G F28G IGF-II to determine which mutations maintained nativelike IGF-II structure and effectively reduced binding affinity. These analogues were expressed as fusion peptides in *E. coli* in inclusion bodies. Native IGF-II refolding in optimised conditions promotes the formation of the correctly folded structure of IGF-II and post cleavage of the fusion partner results in a single predominant peak representing fully folded mature IGF-II (Figure 3.3D) (Lien et al., 2001). Conversely, for the site 1a IGF-II mutants, several folding intermediates were formed during the refold Figure 3.6D, 3.9D and 3.12D. The folding pathway of single chain proinsulin has been well described (Hua et al., 2006a, Weiss, 2009). Under oxidative conditions conserved hydrophobic residues surrounding CA20 and CB19 form a microdomain which orient the CA20 and CB19 thiols to allow disulfide pairing (this is the first disulfide to form) (Weiss, 2009, Hua et al., 2006b). Mutations of insulin residues known to be involved in the microdomain surrounding the CA19-CB20 disulfide bond have reported inefficient folding of the resultant protein (Huang et al., 2004, Weiss et al., 2002a). The FFY motif of insulin

is essential for formation of the core of the protein and residue FB24 likely contributes to the microdomain surrounding CA19-CB20. Based on sequence and structural similarities between single chain proinsulin and IGF-II it is expected that the FYF motif in IGF-II plays a similar role in formation of the protein core early in the oxidative refolding pathway but this has not been investigated for IGF-II. It is likely that perturbation of this aromatic motif hinders efficient formation of the hydrophobic core and formation of the C21-C60 disulfide bond, resulting in folding intermediates being formed.

In the case of F26A IGF-II, it was observed that isolated peaks 11.7 and 13.3 had significant shifts in their spectral minimum from 208 nm (as seen for native IGF-II) to a lower wavelength (Figure 3.7A) in addition to attenuated helical content. These shifts in the CD spectral properties for these isolated fractions is consistent with incorrect disulfide pairings, which has been documented in IGF-I (Hober et al., 1992). In IGF-I folding, both correctly paired native IGF-I (C18-C61, C6-C48 and C47-C52) and disulfide-swap IGF-I (C18-C61, C6-C47 and C48-C52) can form, as confirmed by peptide mapping and sequencing of isolated folding products (Hober et al., 1992, Miller et al., 1993). Disulfide-swap IGF-I has attenuated helical content and a shifted spectral minimum in CD studies (Hober et al., 1992), and binds the IGF-1R with a 10-fold reduction in affinity (Miller et al., 1993). Unfortunately, time did not permit the confirmation of the disulfide bonding pattern of the F26A IGF-II 11.7 and 13.3 peak peptides, which would ideally involve a pepsin peptide mapping/mass spectrometry approach (Milner et al., 1999) or NMR studies (Miller et al., 1993) as used to identify disulfide-swap IGF-I. In the future an alternative method to chemical refolding could be used to generate correctly folded IGF-II and analogues. For example, an *E. coli* periplasmic expression system has been developed which facilitates native disulfide bond formation (Klint et al., 2013). This expression and folding method would involve the expression of IGF-II analogues as a fusion partner with a signal peptide to direct the analogue to the periplasm of *E. coli* for folding in the presence of disulfide isomerase proteins (Dsb), which have been shown to assist in correct disulfide pairing in protein structures (Manta et al., 2019).

The F26G Y27G F28G IGF-II isolated peaks 7.1 and 8.2 are also likely to represent mismatched folding intermediate structures, whereas F26G Y27G F28G IGF-II fraction 9.6 is possibly the correctly folded peptide. As expected, total removal of the aromatic FYF and replacement with glycine residues does significantly alter the secondary structure and binding affinity of the IGF-II analogue (as observed in Figure 3.15A and

Figure 3.16). This makes it difficult to definitively prove correct folding through the evidence gathered as described above. Interestingly, total removal of the equivalent motif in an insulin analogue (and other elements of the site 1a), des-octapeptide (B23–B30) -insulin (DOI), results in near total loss of affinity for the IGF-1R and low affinity for the IR-A compared to human insulin (<0.0006% for the IGF-1R and 0.09% for the IR) (Cara et al., 1990, Bromer and Chance, 1967). Structural and biochemical data suggest that the FYF motif in IGF-II are the main contacts with the L1 domain (and also make contact with the αCT′). However, in insulin, there are additional high affinity receptor contacts. Residue YB16 in insulin makes a high affinity π-π interaction with F39 of the L1 domain and is unique to insulin and the IR (equivalent residue in IGF-1R is S39) (Weis et al., 2018, Gauguin et al., 2008b).

In human insulin, for the C-terminal segment of the B-chain, including the FFY motif, to fold back against the B-chain helix a β turn involving residues B20-B23 is essential (Smith et al., 2003). In mammalian insulins glycine residues are highly conserved at residues B20 and B23 to facilitate this β turn (Nakagawa et al., 2006). Only a small side chain such as alanine has been shown to be accommodated at this position in human insulin, albeit with low yield using a yeast expression system (Kristensen et al., 1997). Interestingly, when the C-terminal segment of the B-chain in human insulin is removed, a β turn is no longer required, as depicted in structural investigations by Bao et al (1997) and indeed more recent investigations have demonstrated that the loss of this C-terminal segment allows for bulky side chins such as that of tyrosine and leucine to be tolerated at position B20 (Xiong et al., 2020). In the case of IGF-II, the single chain structure of this ligand prevented total removal of the equivalent C-terminal segment of the B-domain and therefore the type 1 β turn is still required in this structure to allow the FYF motif to fold back against the b domain helix. It is therefore not surprising that our attempts to introduce a large bulky aromatic tyrosine residue led to a poorly structured analogue.

Binding data reveals that F26A IGF-II retains 11.3% affinity for the IGF-1R and ≤0.1% for the IR-A compared to native IGF-II (Figure 3.16). Together these data suggest that the FFY motif of IGF-II is more critical for IR binding than for the IGF-1R. Residue B24 in insulin has also been mutated to each of the naturally occurring amino acids (with the exception of a cysteine) (Pandyarajan et al., 2014, Zakova et al., 2013, Mirmira and Tager, 1989). While FB24A insulin was found to have 1% affinity for the IR, FB24G insulin was shown to have equal affinity for the IR as human insulin (Pandyarajan et al., 2014, Zakova et al., 2013, Mirmira and Tager, 1989). Further investigations to determine

how this analogue is capable of binding to the IR despite the loss of a the key anchoring residue (FB24) concluded that a single residue register shift may occur in the FB24G insulin analogue, which allows the neighbouring F25 to occupy the same hydrophobic pocket as FB24 (Rege et al., 2020a). The nM affinity of F26A IGF-II for the IGF-1R may be attributed to a register shift, however further investigation would be required to determine this. As done for register shift investigations in insulin, molecular modelling could be used to determine if a register shift occurs in F26A IGF-II whereby the neighbouring YB27 occupies the hydrophobic pocket normally occupied by F26.

In summary, three site 1a minimised IGF-II analogues were produced and characterised in order to investigate alternative modes of receptor binding and engagement. Of the three mutants, two maintained native like IGF-II secondary structure, F26A IGF-II and F26A Y27L F28L IGF-II. F26A IGF-II retained the highest binding affinity for the IGF-1R of all the mutants, followed by F26A Y27L F28L IGF-II. The third analogue, F26G Y27G F28G IGF-II, with the most perturbation of the aromatic FYF motif resulted in markedly altered secondary structure compared to IGF-II and the lowest binding affinity. While F26G Y27G F28G IGF-II is ideal for investigating alternate modes of receptor binding, as it has total perturbation of the FYF motif, the poor yield of folded material is likely to be problematic when introducing further mutations. However, F26A IGF-II was much higher in yield and retained high affinity for the IGF-1R. This is likely due to retention of some elements of the FYF motif. For these reasons, subsequent chapters will explore alternate binding mechanisms in both F26A IGF-II and F26G Y27G F28G IGF-II mutants.

4 SINGLE MUTANTS OF IGF-II BASED ON CON-INS G1 ALTERNATE BINDING CONTACTS

4.1 Introduction

Understanding residue specific determinants of altered signaling bias is required to develop antagonists which can selectively block mitogenic action of IGF-II. There are a few examples in IGF-II of single residue mutations that alter signaling properties. However, in recent years many invertebrate insulins have been described with entirely novel binding mechanisms and unique signaling properties (Menting et al., 2016, Xiong et al., 2021). Notably, cone snail insulins Con-Ins G1 and Con-Ins K1 lack the FFY motif that is key for high affinity site 1a binding and yet they bind with relatively high affinity due to alternate compensating residues. Our laboratory now aims to investigate IGF-II altered signaling properties by removing the classic binding site in human IGF-II and introducing these newly discovered modes of receptor contact. Working toward this goal, in the previous chapter (Chapter 3) the production of three site 1a minimised IGF-II analogues was described, producing for the first time, IGF-II analogues lacking the entire FYF motif. These analogues provide the framework for investigating alternate modes of receptor engagement. In this chapter individual residues in native IGF-II were substituted with residues identified in cone snail venom insulins as being able to compensate for the missing FFY (in insulin) motif. The effect of these mutations on folding, receptor binding and activation was investigated. Ultimately these mutations will be incorporated into the site 1a minimised IGF-II analogues produced in Chapter 3 to test whether they can compensate for the lack of the FFY motif.

As described in detail in Chapter 3, in human insulin and insulin like growth factors I and II (IGF-I and IGF-II) the FFY/FYF motif is crucial for high affinity interaction with the L1 domain and αCT′ (site 1a). A lower affinity interaction is also made with the FnIII-1′ domain (site 1b) (Blyth et al., 2022, Lawrence, 2021). The FFY motif has been heavily studied in insulin, in particular showing that residue FB24 is critical for anchoring the ligand to a hydrophobic pocket on the surface of the L1 domain (Pandyarajan et al., 2014, Weis et al., 2018, Menting et al., 2013, Mirmira and Tager, 1989). Residues FB25 and YB26 engage with shallow depressions on the surface of the L1 domain and are also observed to interact with the αCT′ (the same is observed for IGF-II and IGF-I binding to the IGF-1R). (Weis et al., 2018, Li et al., 2019, Xu et al., 2020). Sequence comparisons of vertebrate insulins and IGFs demonstrate the importance of this FFY/FYF motif as it is highly conserved among the vast majority of vertebrates (Irwin, 2021, Baral and Rotwein, 2019, Rotwein, 2017).

More recent studies investigating invertebrate insulins and IGFs have revealed that the

aromatic FFY/FYF motif is much less conserved (Gronke et al., 2010, Li et al., 2020, Ahorukomeye et al., 2019). Structural investigations of these invertebrate insulins have revealed novel modes of binding to the IR despite lack of conservation of the FFY/FYF motif. Cone snail venom insulins Con-Ins G1 and Con-Ins K1 completely lack the C-terminal segment of the B-domain including the FFY motif (Figure 4.1) (Xiong et al., 2021, Menting et al., 2016). Despite this Con-Ins G1 has a similar affinity for the human IR to human insulin (Menting et al., 2016). It engages with receptor site 1a through residues YB15 and YB20 instead of the FFY motif (Figure 4.2B) (Menting et al., 2016). However, the affinity for site 1a is significantly lower compared to human insulin. Con Ins G1 residues HA8, RA9 and a post translationally modified γ-carboxyl glutamate at position B10 contribute to an increase in affinity for site 1b (Safavi-Hemami et al., 2015).(Figure 4.2C). The influence of the three post translational modifications on Con-Ins G1s affinity was found to only result in a minor increase in affinity compared to their non-modified counterparts (Menting et al., 2016).

Removal of the 8 amino acids from the C-terminus of the B-domain from human insulin, termed des octapeptide (B23–B30)-insulin (DOI), results in an analogue with an affinity of 0.09% for the IR and 0.002% for the IGF-1R, compared to human insulin (Cara et al., 1990). Recently, the key binding residues identified in Con-Ins G1 were incorporated into the DOI analogue (Figure 4.3A and B). This analogue, now termed Mini-Ins, had a similar affinity to human insulin (Xiong et al., 2020); however, with a weaker site 1a and stronger site 1b interaction compared to human insulin. Despite the altered site 1a and site 1b affinities, Mini-Ins retained an equal potency in stimulation of metabolic (Akt pathway) activation compared to human insulin. The kinetics of mitogenic activation (ERK pathway) by Mini-Ins appeared to be delayed compared to human insulin, but an equal maximal response was stimulated (Xiong et al., 2020).

Another cone snail venom insulin of interest is Con-Ins K1. This peptide has also been found to have residues that compensate for the missing FFY motif (Figure 4.2 C and D). Con-Ins K1 has a unique A-chain extension; T22, L23, Q24 and γ-carboxyl glutamate at position 25 (Xiong et al., 2021). This analogue also has a post translationally modified γ-carboxyl glutamate at position B10 and a leucine at position B20. Various combinations of A-chain extensions were also investigated in the context of DOI (Xiong et al., 2021). One variation of the A-chain extension, termed Vh-Ins, was found to increase affinity 400-fold, to equal that of human insulin. (Xiong et al., 2021). Vh-Ins has a four-residue helical A-chain extension (HA21, SA22, LA23 and QA24) (Figure 4.3D). Residue LA23

of this A-chain extension was found to engage with the same pocket otherwise occupied by FB24 in human insulin (Xiong et al., 2021) (Figure 4.4). Through generation of a series of analogues with changes at each residue of the HSLQ motif LA23 was also confirmed to be important for high affinity binding whilst residues HA21 and QA24 were less important. In contrast to Con-Ins G1, Vh-Ins has a high affinity for the site 1a, and structural data implies that site 1b was weakened, with only EB10 observed to contact the with FnIII-1′ domain (residue R539) (Xiong et al., 2021). The IGFs possess a similar domain (D-domain) to the A-chain extension found in Con-Ins K1; however, with no sequence similarity (Figure 4.3D). The D-domain of the IGF-II contributes, to a small extent, to its high affinity for the IR-A (Denley et al., 2004). Based on the Vh-Ins findings, replacement of the IGF-II D-domain with an A61, S62 L63, Q64 sequence was investigated to see if this would alter overall affinity and/or change signaling outcome. An alanine was maintained at position 61 in an attempt to preserve the local structure neighbouring the C60-C21 disulfide bond.

Our laboratory now aims to introduce into IGF-II substitutions equivalent to the residues that compensate for the loss of the FFY motif in Con-Ins G1 and Con-Ins K1. Structures of human IGF-II and Con-Ins G1 were used to assess the potential impact of introducing these residues into IGF-II. As shown in Figure 4.5, native IGF-II has an arginine at position 49 (equivalent position to insulin SA9). An arginine is also found in Con-Ins G1 (residue RA9). At position 12 (HB10 in insulin) IGF-II has a glutamate, which is the same as the non-post translationally modified version of the γ-carboxyl glutamate at position B10 in Con-Ins G1.

Residue G22 in IGF-II, equivalent to YB20 in Con-Ins G1, is required to form a type I turn, allowing the C-terminus of the B-domain to fold antiparallel against the B-chain helix (Torres et al., 1995, Blundell et al., 1972) (Figure 4.5). This residue is highly conserved, suggesting it is important for maintaining structure (Baral and Rotwein, 2019). Attempts to introduce another small side chain, alanine, into the equivalent position in insulin was tolerated, albeit with very poor yield using a yeast expression system (Kristensen et al., 1997). However, we speculated that due to the conservation of a glycine at this position in IGF-II, attempts to mutate it to a significantly larger side chain, such as a tyrosine as found in Con-Ins G1, would impact the ability of the B-domain helix to fold back against the B-domain helix and hence hinder the formation of a stable structure.

Residue L17 in IGF-II, equivalent to YB15 in Con-Ins G1, is directed toward the core of the protein (Figure 4.5). This residue is also highly conserved, suggesting it is likely

required for maintaining structure of the core of the protein (Baral and Rotwein, 2019). Mutation of this residue to an alanine in insulin was not tolerated (Kristensen et al., 1997), suggesting that attempts to mutate this residue in IGF-II may also not be tolerated. Based on these assessments, it was predicted that mutation of residues G22 and L17 in IGF-II would affect folding/stability.

To investigate the effect of Con-Ins G1 alternate binding residues on the biophysical properties and biological activity of IGF-II, several analogues were investigated; namely L17A IGF-II, L17Y IGF-II, G22A IGF-II, G22L IGF-II and G22Y IGF-II. Each was analysed for impact on structure and their ability to activate the IR-A and IGF-1R. To investigate the effect of A-domain modifications as seen in Con-Ins K1, replacement of the IGF-II D-domain with an A61, S62, L63, Q64 sequence was subsequently investigated. The only difference of this modified D-domain to the one tested in Vh-Ins is an alanine at the beginning (A60) instead of a histidine (HA21), to preserve the native structure neighbouring the C60 disulfide bond.

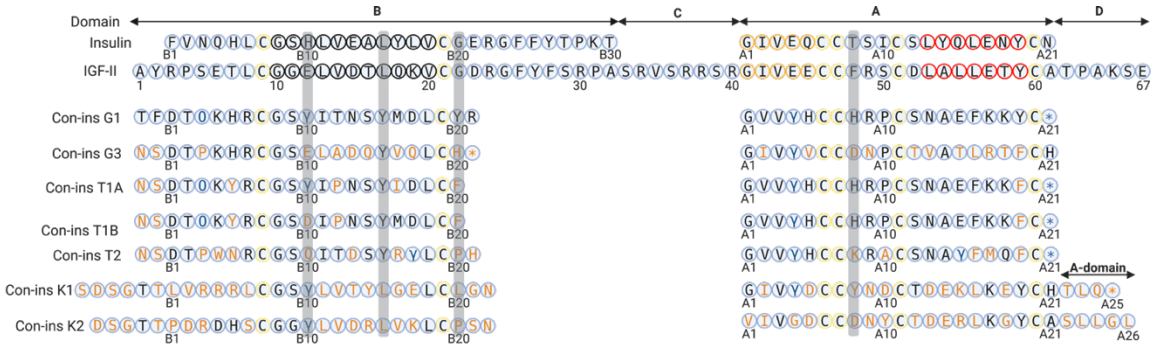


Figure 4.1: Sequence alignment of insulin, IGF-II and cone snail venom insulins.

Sequence alignment of insulin, IGF-II and cone snail venom insulins identified from three species of cone snail *Conus. geographus*, *C. tulipa* and *C. kinoshitai* (Ahorukomeye et al 2019). For human insulin and IGF-II peptides, the three helices; B-domain helix 1 (black circles), A-domain helix 2 (orange circles) and A-domain helix 3 (red circles) are shown. Residue numbers are indicated below each sequence and cysteine residues are coloured yellow. For the cone snail venom insulins, residues that differ from the Con-Ins G1 sequence are highlighted in orange letters. Grey boxes represent residues in Con-Ins G1 that compensate for the lack of the FFY motif. The A-chain extensions of Con-Ins K1 and K2, compensate for the loss of the FFY motif. The cone snail venom insulins have several post translational modifications; γ, γ-carboxyglutamic acid; O, hydroxyproline; asterisk, C-terminal amidation.

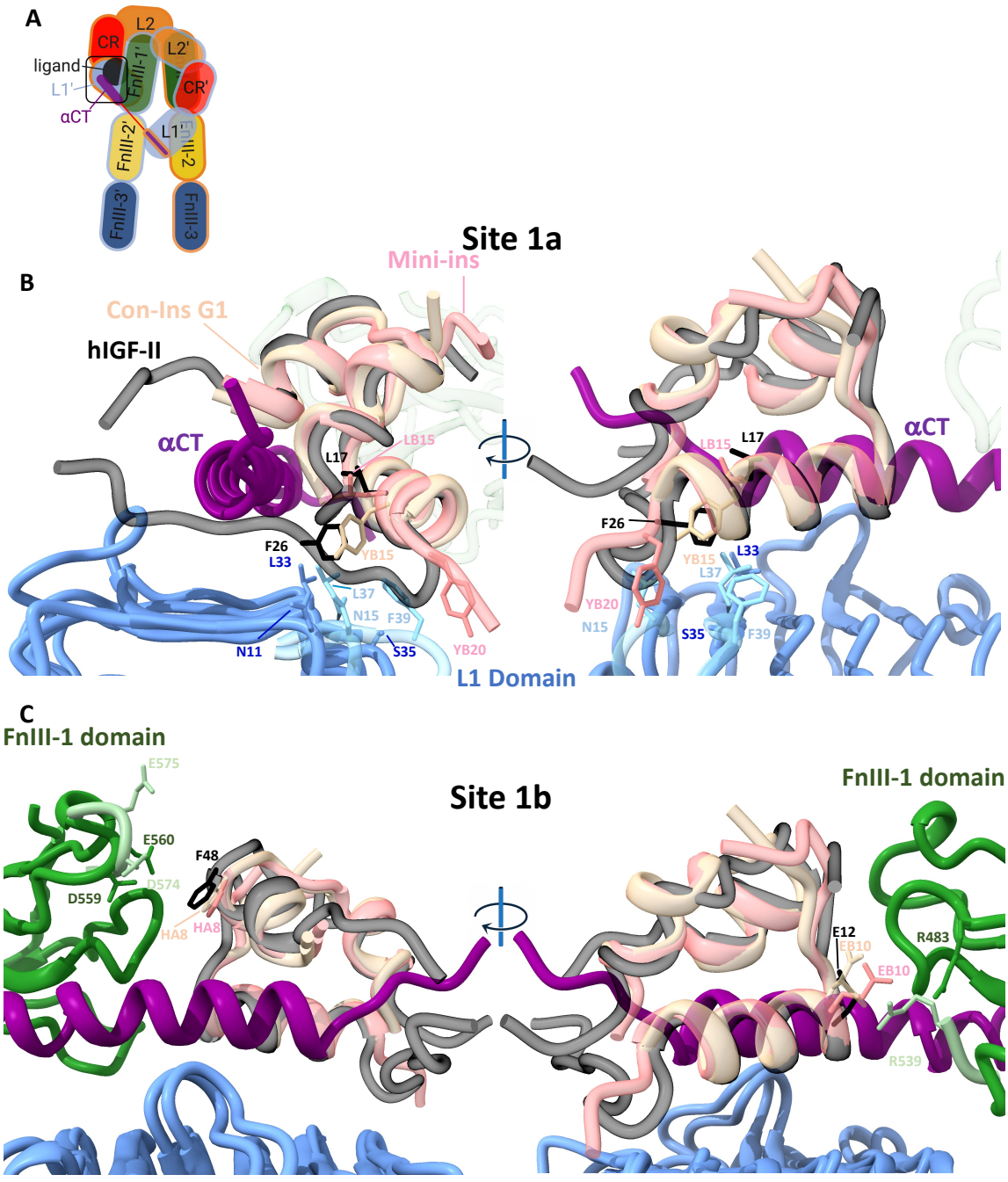

A
L2
CR
L2'
ligand
FnIII-1'
CR'
L1'
αCT
FnIII-2'
L1'
FnIII-2
FnIII-3'
FnIII-3
Site 1a
B
Mini-ins
Con-Ins G1
hIGF-II
αCT
LB15
L17
F26
L33
YB15
L37
N15
F39
N11
S35
YB20
L1 Domain
C
FnIII-1 domain
E575
E560
F48
D574
D559
HA8
Site 1b
FnIII-1 domain
E12
EB10
R483
R539

Figure 4.2: Structural comparison of IGF-II, Con-Ins G1 and Mini-Ins receptor engagement.

(A) Schematic diagram representing the IR and IGF-1R ligand bound extracellular domain structures. Individual αβ monomer outlines are coloured either orange or blue and domains from the second monomer are denoted with a prime (e.g. L1'). Extracellular domains include the first (light blue) and second (orange) leucine-rich repeat domains (L1 and L2), cysteine-rich domain (CR, red), first (green), second (yellow) and third (dark green) fibronectin type-III domains (FnIII-1, -2, and -3), insert domain (ID, red line) and α-chain C-terminal region (αCT, purple). The ligand is a black circle. Transmembrane and intracellular domains, including the tyrosine kinase domain, are not shown. (B) IGF-II (grey) bound to site 1a of the IGF-1R (from PDB: 6VWI) overlaid with Con-Ins G1 (PDB: 6VEQ, tan) and Mini-Ins (PDB: 6VET, salmon). The colouring of the receptor domains is as in (A). The site 1a L1 domain residues contacting F26 of IGF-II are shown (equivalent residues of the IR L1 domain are shown in light blue, from PDB: 6VEQ). Side chains are shown of residues LB15/YB15 of Con-Ins G1 and Mini-Ins, and YB20 of Mini-Ins (YB20 of Con-Ins G1 is not resolved in this structure). (C) IGF-II bound to site 1b of the IGF-1R (from PDB: 6VWI) with Con-Ins G1 and Mini-Ins overlayed with the same colouring scheme as described for B. Side chains of residues involved in site 1b contact are shown. Equivalent residues of the IR FnIII-1′ domain are indicated in light green (from PDB: 6VEQ).

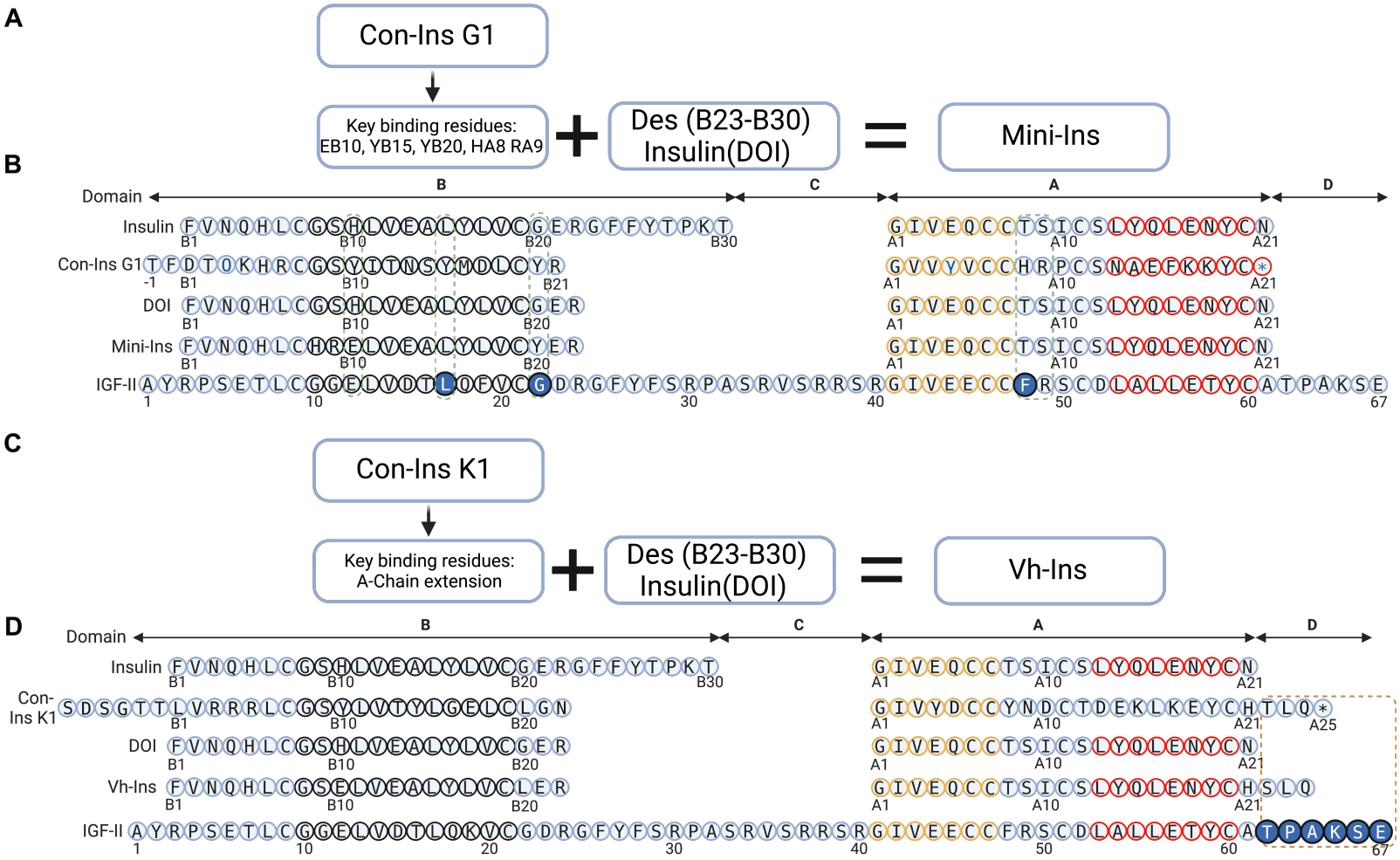

A
Con-Ins G1
Key binding residues:
EB10, YB15, YB20, HA8 RA9
Des (B23-B30)
Insulin(DOI)
Mini-Ins
B
Domain
B
C
A
D
Insulin
Con-Ins G1
DOI
Mini-Ins
IGF-II
C
Con-Ins K1
Key binding residues:
A-Chain extension
Des (B23-B30)
Insulin(DOI)
Vh-Ins
D
Domain
Insulin
Con-
Ins K1
DOI
Vh-Ins
IGF-II

Figure 4.3: Schematic representations of the construction of Mini-Ins and Vh-Ins based on identified compensatory residues for the loss of the FYF motif.

(A) Schematic diagram representing the process used to produce Mini-Ins. Con-Ins G1 compensates for the lack of the FFY motif through several residues including YB15, YB20, HA8, RA9 and a post translationally modified γ-carboxyl glutamate at position B10 (shown in the green dotted boxes), forming an alternate binding mechanism compared to human insulin. These residues were explored in a human insulin analogue lacking the C-terminal 8 amino acids (including the FFY motif) of the B-chain (des-octapeptide (B23–B30)–insulin (DOI)). Mini-Ins therefore represents a human insulin analogue with an alternate mode of receptor binding. (B) Sequence alignment of each peptide described in A. Each peptide has three helices; B-domain helix 1 (black circles), A-domain helix 2 (orange circles) and A-domain helix 3 (red circles). Residue numbers are indicated below each sequence. The cone snail venom insulins have several post translational modifications; γ, γ-carboxyglutamic acid; O, hydroxyproline; asterisk, C-terminal amidation. The three green dotted boxes highlight alternate binding residues in Con-Ins G1. The equivalent three IGF-II residues shown in solid blue were investigated in this chapter. (C) Schematic diagram representing the process used to produce Vh-Ins. Con-Ins K1 compensates for the lack of the FFY motif through an A-chain extension, highlighted in the orange dotted box, including T22, L23, Q24 and γ-carboxyl glutamate at position 25. The A-chain extension residues were incorporated into a DOI analogue, Vh-Ins. Vh-Ins represents a human insulin analogue with an alternate mode of receptor binding. (D) Sequence alignment of each peptide described above with helices coloured as in B. The equivalent residues in IGF-II form the D-domain and are shown in solid blue. This chapter explores replacement of the IGF-II D-domain with the A-domain residues of Vh-Ins.

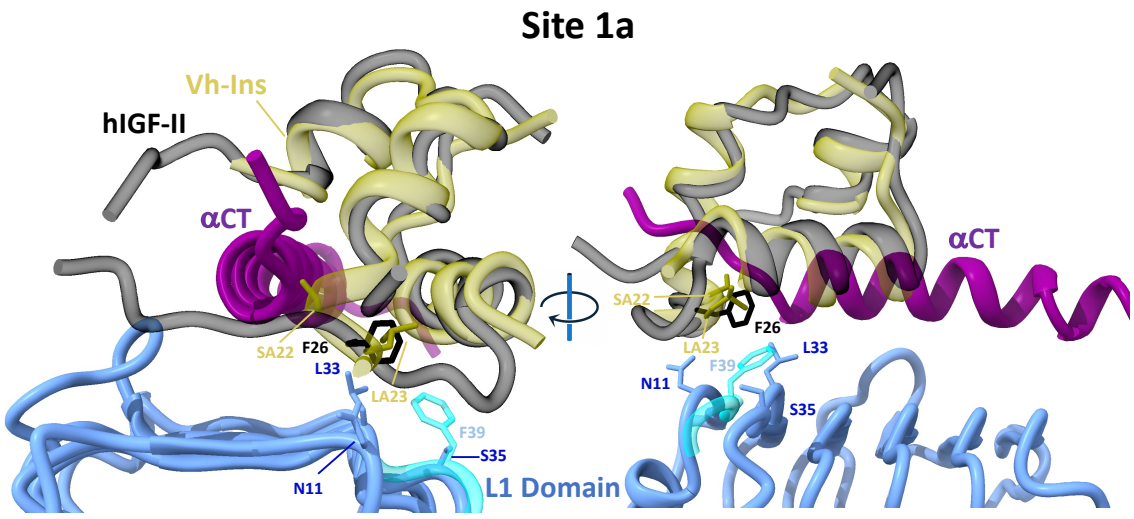


Figure 4.4: Structural comparison of IGF-II and Vh-Ins receptor engagement.

IGF-II (grey) bound to site 1a of the IGF-1R (from PDB: 6VWI) overlayed with Vh-Ins (olive green) (PDB:7MQO). Residues of the L1 domain that contact residue F26 of IGF-II are highlighted (equivalent residues of the IR L1 domain are shown in light blue, from PDB: 7MQO). Vh-Ins side chains compensating for the lack of the FFY motif when bound to the IR are highlighted and include SA22 and LA23. LA23 engages with the equivalent pocket on the IR L1 domain as occupied by IGF-II F26 on the IGF-1R.

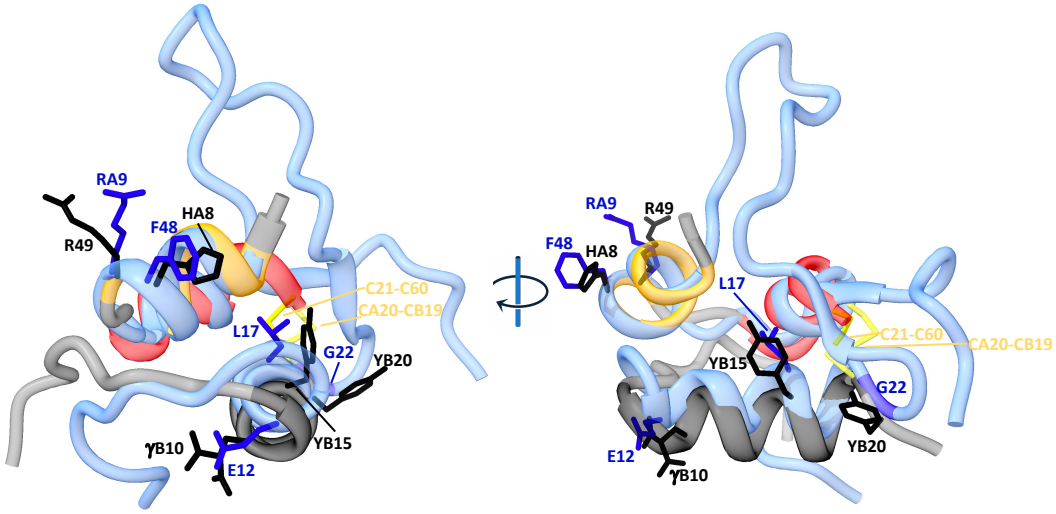


Figure 4.5: Structural comparison of IGF-II and Con-Ins G1.

(B) Overlay of Con-Ins G1 (PDB: 5JYQ (grey) and IGF-II and (PDB: 1IGL (blue)). Helix colouring as in A for Con-Ins G1 and IGF-II shown in light blue. Side chains of key receptor contacts of Con-Ins G1 are shown in black and equivalent residues in IGF-II are shown in dark blue. Residue γB10 in Con-Ins G1 is a post translationally modified glutamic acid residue (γ-carboxyglutamic acid). Therefore, E12 in IGF-II is the non-post translationally modified version of γ-carboxyglutamic acid. Residue LB15 of Con-Ins G1 and equivalent residue L17 in IGF-II are positioned in the core of the protein. YB20 in Con-Ins G1 is adjacent to the CA20-CB19 disulfide bond (shown in yellow). In IGF-II the equivalent residue is G22 adjacent to C21-C60. Con-Ins G1 residue HA8 and equivalent residue of IGF-II, F48, are positioned on the outer surface of the ligand. IGF-II already has an arginine at position 49, as observed in Con-Ins G1 (RA9).

4.2 Results

4.2.1 Production of G22 IGF-II and L17 IGF-II analogues

4.2.1.1 Expression and Dissolution of IGF-II analogues

Each analogue was expressed as a fusion protein including 11 N-terminal amino acids of the porcine growth hormone (pGH(1-11)) as a fusion partner and a cleavage recognition site (Lien et al., 2001). The pGH(1-11) expression was induced by addition of IPTG (Section 2.3.6). Pre and post induction samples were taken (Section 2.3.6) and separated on a 15% tricine gel under reducing conditions, followed by staining with Coomassie blue (Section 2.3.4). Successful expression of each analogue was confirmed by the presence of a 9.3 kDa band (expected size of pGH(1-11) IGF-II analogue) in the post induction sample but not in the pre induction sample (Supplementary Figures 8 and 9). The presence of the pGH(1-11) fusion partner facilitated formation of inclusion bodies during expression which were released from *E.coli* by high pressure homogenisation (Section 2.3.7). Insoluble inclusion bodies were separated from soluble host cell material by centrifugation (Section 2.3.7). pGH(1-11) IGF-II analogue inclusion bodies were subsequently solubilised (Section 2.3.8) (Figure 4.6A - 4.11A). As shown in the accompanying table of Figure 4.12 the L17 IGF-II analogues had similar protein concentration at the dissolution stage compared to native IGF-II, whilst the G22 IGF-II analogues all had lower concentrations. The dissolution reactions were acidified to pH 2 prior to loading onto a size exclusion chromatography column (Section 2.3.8) to separate IGF-II analogues from high molecular weight host cell material including proteolytic enzymes (Figure 4.6B - 4.11B). The L17 IGF-II analogues and native IGF-II all had relatively minor decreases in yield between dissolution and gel filtration, whereas significant losses occurred at this stage for the G22 IGF-II analogues (Figure 4.12 and accompanying table). In particular, only 13.4% of G22Y IGF-II peptide remained after gel filtration compared to the dissolution stage.

4.2.1.2 Refolding, cleavage and purification of IGF-II analogues

Refold of each pGH(1-11) IGF-II analogue was performed at a final concentration of 2.5 M urea and 0.5 mM DTT (Section 2.3.9) (Figure 4.6C - 4.11C). Refold was initiated by the addition of the oxidising agent 2-hydroxyethyldisulfide, which facilitated the formation of the three disulfide bonds and subsequently formation of the biologically active structure. As shown in Figure 4.6C – 4.11C, at the completion of refold there is a decrease in retention time of each analogue compared to T=0, characteristic of a transition from a reduced state to a more hydrophilic structure. Intermediate folding

states of each IGF-II analogue were observed during the refold. As described in Chapter 3, intermediate folding states were characterised. Based on these investigations we have been able to gain an understanding of the folding landscape of IGF-II analogues. In this chapter we have applied the same understanding of folding intermediates to select the mature, fully folded IGF-II analogue in each case. Preparative reverse phase HPLC (Section 2.3.11) was used to separate the cleaved leader sequence from each mature IGF-II analogue followed by freeze drying (Section 2.3.12) of the mature IGF-II. Interestingly, L17A IGF-II and L17Y IGF-II analogues had major losses of peptide during the refold (Figure 4.12 and accompanying table) with an 83.7% and 82.0% decrease in peptide yield, respectively (between gel filtration and completion of refold). The pGH(1-11) leader sequence was cleaved from each analogue using α-LP (Section 2.3.10) (Figures 4.6D - 4.11D). Upon cleavage of folded pGH(1-11) IGF-II analogues, the pGH(1-11) leader sequence was observed (Figures 4.6D - 4.11D). The final mature IGF-II analogue is indicated on each Figure (Figures 4.6E - 5.11E). Post cleavage, L17A IGF-II and L17Y IGF-II appeared to have several intermediate folding states suggesting disulfide shuffling during cleavage, whilst the G22 IGF-II analogues all appeared to have one major peak. Mass spectrometry and circular dichroism were performed on all analogues.

4.2.1.3 Mass spectrometry and circular dichroism

The mass of each IGF-II analogue was equivalent to the mass expected, indicating the formation of the 3 disulfide bonds (Supplementary Table 2). Circular dichroism (CD) spectra were recorded for each IGF-II analogue between 180 and 260 nm (Section 2.4.4.1). L17A IGF-II and L17Y IGF-II exhibited attenuated helical content in comparison to native IGF-II (both with 23% helical content compared to 41% for IGF-II, as shown in Figure 4.13A and corresponding Table). The spectral minima of the L17 IGF-II analogues were also shifted to a lower wavelength. Similarly, G22L IGF-II and G22Y IGF-II had attenuated helical content (23 and 19% respectively), compared to IGF-II (41%), with slight shifts in the spectral minima and maxima (Figure 4.14A and accompanying Table). Conversely, G22A IGF-II (Figure 4.14A) and ASLQ IGF-II (Figure 4.15A) were observed to have an increase in helical content (48% and 49%, respectively) compared to IGF-II and similar spectral minima. The marked reduction in helical content for the majority of these analogues indicates that the introduced mutations markedly affected secondary structure**.** The analogues produced in this chapter follow a similar pattern of folding states as the analogues characterised in Chapter 3, therefore, the peaks isolated here were believed to represent the most folded state for these analogues. The decrease in helical

content could be attributed to the mutations introduced affecting secondary structure. However, there is a possibility that these isolated fractions represent intermediate folding states and additional investigations would be required to definitively demonstrate correct disulfide pairing.

4.2.1.4 Thermal denaturation studies

Thermal denaturation studies were preformed from 20°C to 90°C and helical content was monitored at the helix sensitive wavelength 222 nm (Section 2.4.4.1.1). At the beginning of thermal denaturation (20±0.1°C) all IGF-II analogues had $[\theta]_{222}$ values consistent with CD spectra (Figure 4.13B - 4.15B and corresponding Tables). The rate of change (slope) in $[\theta]_{222}$ value with increasing temperature is lower for both L17A IGF-II and L17Y IGF-II, compared to native IGF-II which is attributed to their lower starting helical content (Figure 4.13B and corresponding Table). At the completion of temperature denaturation (90±0.1°C) L17A and L17Y IGF-II had similar $[\theta]_{222}$ values as native IGF-II. G22L IGF-II and G22Y IGF-II (Figure 4.14B and corresponding Table) had similar starting helical content to the L17 IGF-II analogues. However, at completion of temperature denaturation (90±0.1°C) G22A IGF-II and G22Y IGF-II had lower $[\theta]_{222}$ values, indicating further loss of helical content. G22A IGF-II and ASLQ IGF-II (Figure 4.15B), exhibited a similar change (slope) in $[\theta]_{222}$ as native IGF-II, with both analogues maintaining similar helical content to IGF-II at the completion of temperature denaturation (90±0.1°C). Refer to Supplementary Figures 10 and 11 for the $[\theta]_{222}$ values from 90°C to 20°C.

4.2.1.5 IGF-1R and IR-A competition binding assays

IGF-1R and IR-A competition binding assays were performed using europium labelled IGF-II (Eu-IGF-II) and increasing concentrations of competing ligand (Section 2.4.1). When binding to the IGF-1R, L17A IGF-II had a lower affinity (12.3%) than native IGF-II. L17Y IGF-II, however, had even lower affinity (<1%) for the IGF-1R compared to native IGF-II (Figure 4.16A and accompanying Table). Both L17A IGF-II and L17Y IGF-II also exhibited similar decreases in affinities for the IR-A, with 10.2% and <0.01% binding affinities compared to IGF-II (Figure 4.16B and accompanying Table).

G22L IGF-II and G22Y IGF-II also had lower affinities for the IGF-1R compared to native IGF-II (12.9 and 13.5%, respectively) (Figure 4.17A and accompanying Table). The effects of the G22L IGF-II and G22Y IGF-II mutations on IR-A binding affinities were somewhat greater than those on the IGF-1R. G22L IGF-II and G22Y IGF-II bound to the IR-A with 5.03% and 3.27% affinities compared to IGF-II. Strikingly, G22A IGF-II bound

to the IGF-1R with the same affinity as native IGF-II (111%), but with a 2-fold lower affinity for the IR-A (58.5% compared to native IGF-II). Finally, the addition of the ASLQ motif at the end of the D-domain resulted in only a 2-3 fold lower affinity compared to IGF-II for both the IGF-1R and IR-A respectively (Figure 4.18 and accompanying Table).

4.2.1.6 Activation of Akt and ERK1/2 signaling

Potency of activation of metabolic and mitogenic signaling was measured through immunoblotting for phosphorylated Akt (T308) and phosphorylated ERK1/2, respectively (Section 2.4.2). For the L17 IGF-II analogues a dose response assay was used (0, 5, 10, 50, 100 and 500 nM). L6 rat skeletal myoblasts overexpressing the human IR-A (L6 IR-A) were treated with varying concentrations of IGF-II or L17 IGF-II analogue for 10 minutes. At each concentration, IGF-II potently activates phosphorylation of Akt (Figure 4.19B). L17A IGF-II weakly activates phosphorylation of Akt at 5 and 10 nM but at 50,100 and 500 nM it is equipotent to IGF-II at stimulating phosphorylation of Akt. As expected, L17Y IGF-II poorly activated phosphorylation of Akt at all concentrations, consistent with its low binding affinity for the IR-A. Phosphorylation of ERK1/2 followed a similar pattern for all analogues (Figure 4.19C), where IGF-II was the most potent at all concentrations. L17A IGF-II was less potent compared to IGF-II. Finally, L17Y IGF-II was the least potent at activating ERK1/2 phosphorylation at all concentrations, however, this was only tested once (n=1) and further testing would be required to determine statistical significance.

For the G22 IGF-II analogues potency of activation of metabolic and mitogenic signaling was also measured through immunoblotting (Section 2.4.2) for phosphorylated Akt (T308) and phosphorylated ERK1/2, respectively, however, in this case a time course assay was used (Figure 4.20). L6 rat skeletal myoblasts overexpressing the human IR-A (L6 IR-A) were treated with 10 nM of each G22 IGF-II analogue or IGF-II for 0, 0.5, 1, 3, 5, and 10 min. G22A IGF-II and IGF-II displayed a similar trend of Akt (T308) and ERK1/2 phosphorylation (Figure 4.20B and Figure 4.20C). G22L IGF-II and G22Y IGF-II poorly activated Akt (T308) and ERK1/2 phosphorylation at all time points compared to native IGF-II. This was only performed once and would need to be repeated (n=3) to determine statistical significance.

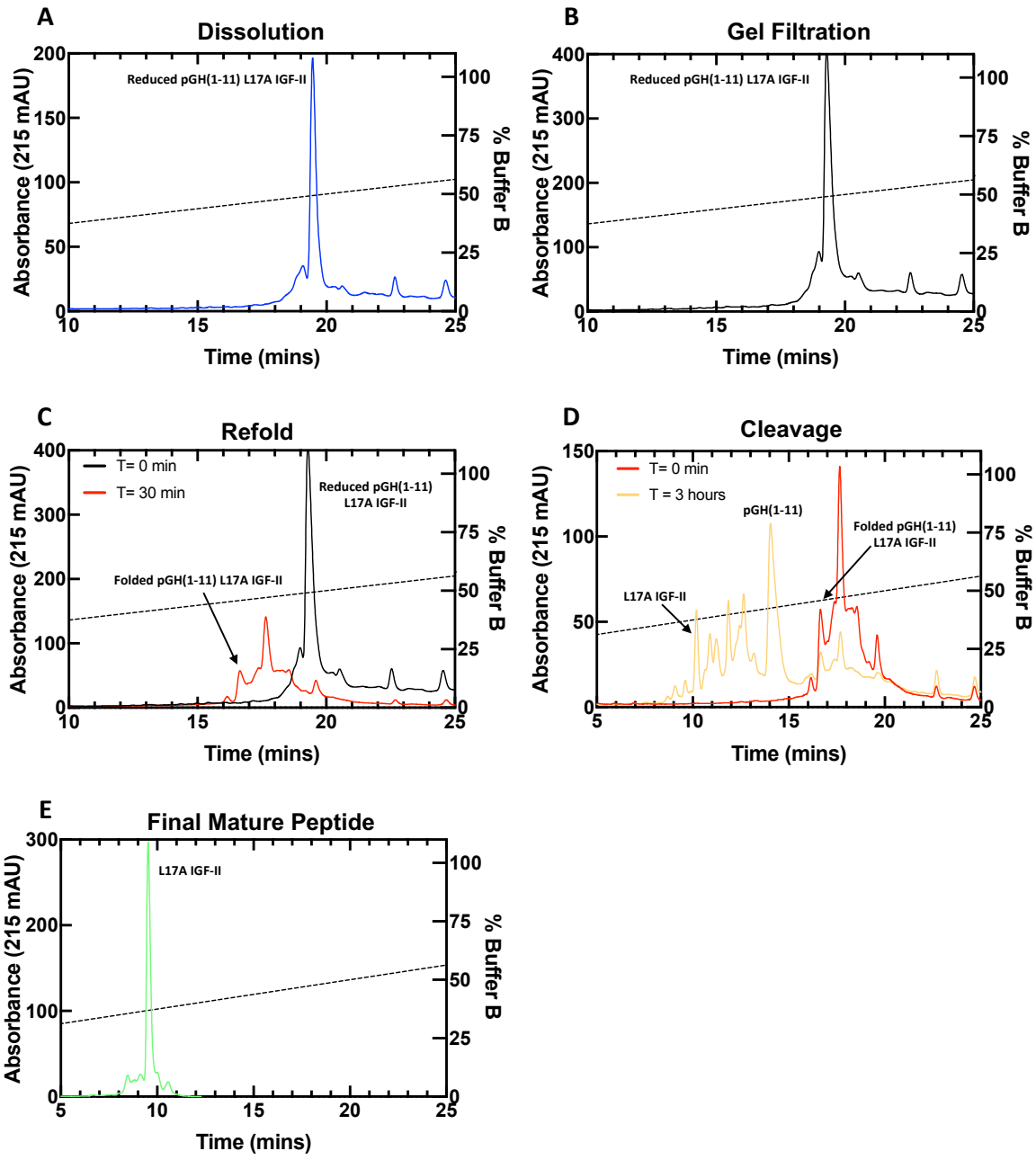


Figure 4.6: Purification of L17A IGF-II.

Analytical reverse phase HPLC (using a C4 column and a 25 – 45 % CH_3CN (31.3 – 56.3% buffer B) over 20 minutes (Section 2.3.5)) was used to monitor each stage of protein purification. Shown are the absorbance chromatograms at 215 nm of reduced pGH(1-11) L17A IGF-II (left Y axis). Dotted black line represents the % of buffer B (right Y axis). (A) Dissolution, (B) post gel filtration, (C) the refold T=0 minutes (solid black line) and T=30 minutes (red), (D) α-LP cleavage T=0 minutes (red) and T=3 hours (orange) with mature L17A IGF-II at 9.8 min and (E) final mature L17A IGF-II peptide post preparative HPLC purification (green).

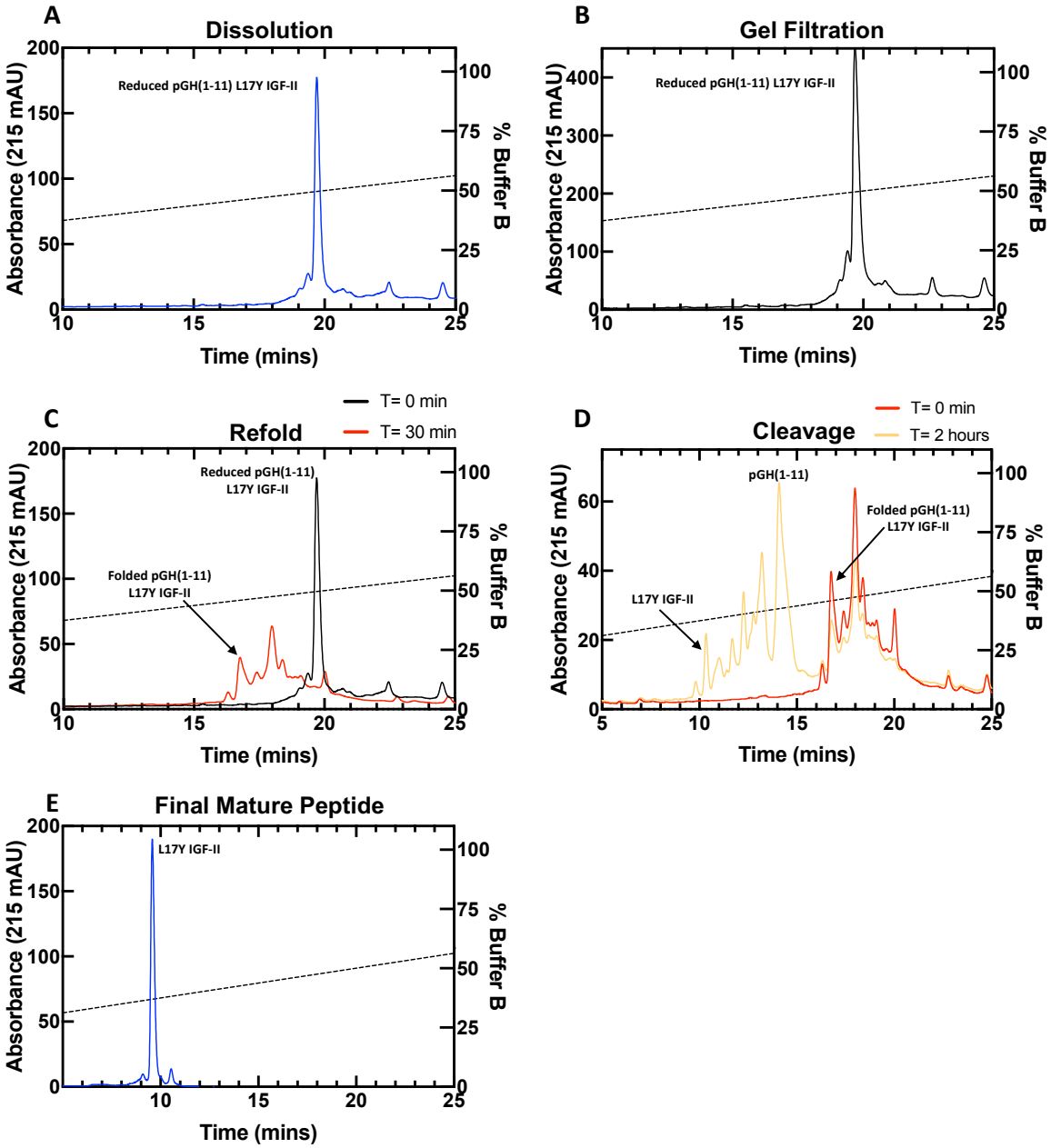


Figure 4.7: Purification of L17Y IGF-II.

Analytical reverse phase HPLC (using a C4 column and a 25 – 45 % CH_3CN (31.3 – 56.3% buffer B) over 20 minutes (Section 2.3.5)) was used to monitor each stage of protein purification. Shown are the absorbance chromatograms at 215 nm of reduced pGH(1-11) L17Y IGF-II (left Y axis). Dotted black line represents the % of buffer B (right Y axis). (A) Dissolution, (B) post gel filtration, (C) the refold T=0 minutes (solid black line) and T=30 minutes (red), (D) α-LP cleavage T=0 minutes (red) and T=2 hours (orange) with mature L17Y IGF-II at 9.8 min and (E) final mature L17Y IGF-II peptide post preparative HPLC purification (green).

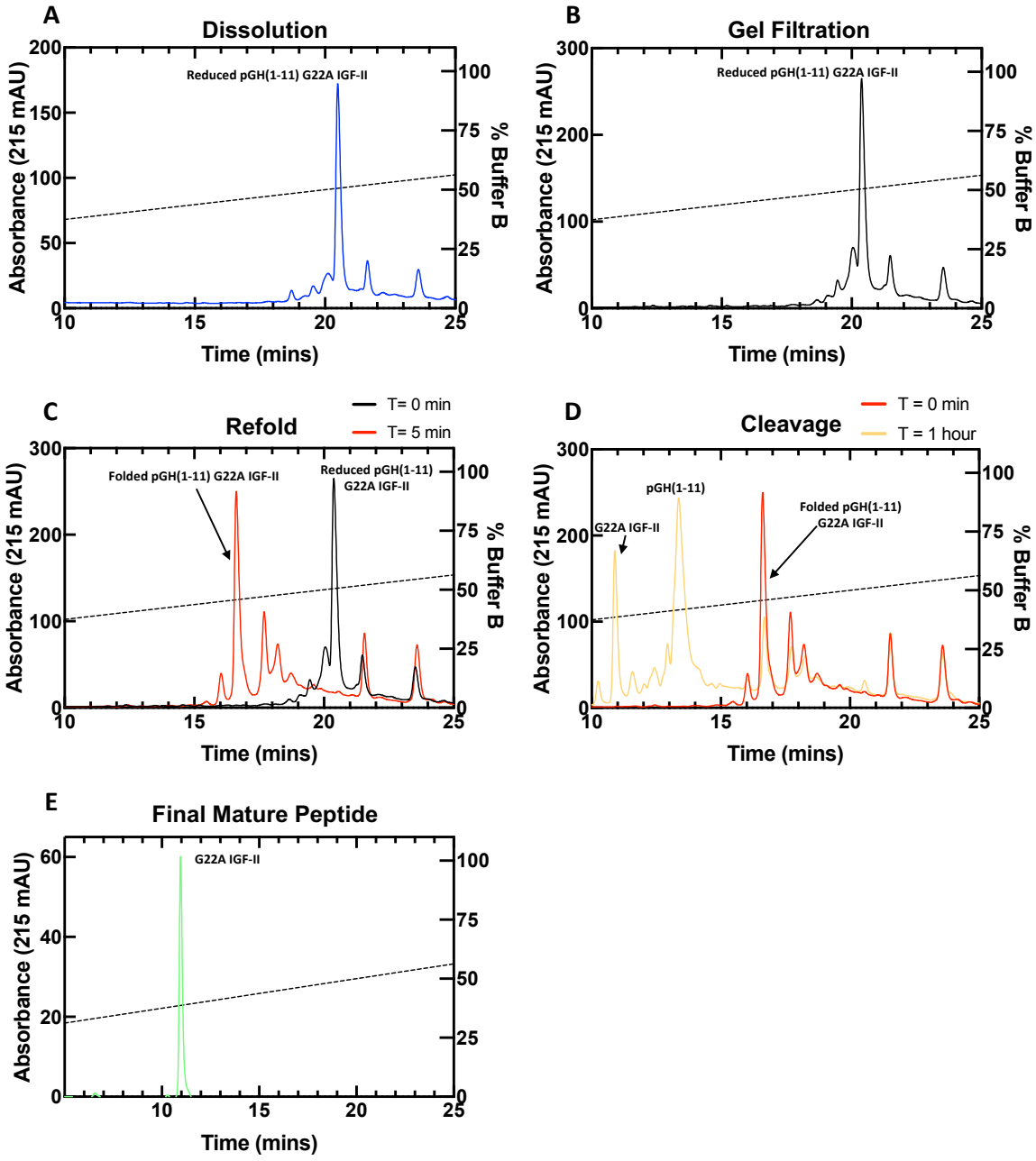


Figure 4.8: Purification of G22A IGF-II.

Analytical reverse phase HPLC (using a C4 column and a 25 – 45 % CH_3CN (31.3 – 56.3% buffer B) over 20 minutes (Section 2.3.5)) was used to monitor each stage of protein purification. Shown are the absorbance chromatograms at 215 nm of reduced pGH(1-11) G22A IGF-II (left Y axis). Dotted black line represents the % of buffer B (right Y axis). (A) Dissolution, (B) post gel filtration, (C) the refold T=0 minutes (solid black line) and T=5 minutes (red), (D) α-LP cleavage T=0 minutes (red) and T=1 hour (orange) with mature G22A IGF-II at 11.0 min and (E) final mature G22A IGF-II peptide post preparative HPLC purification (green).

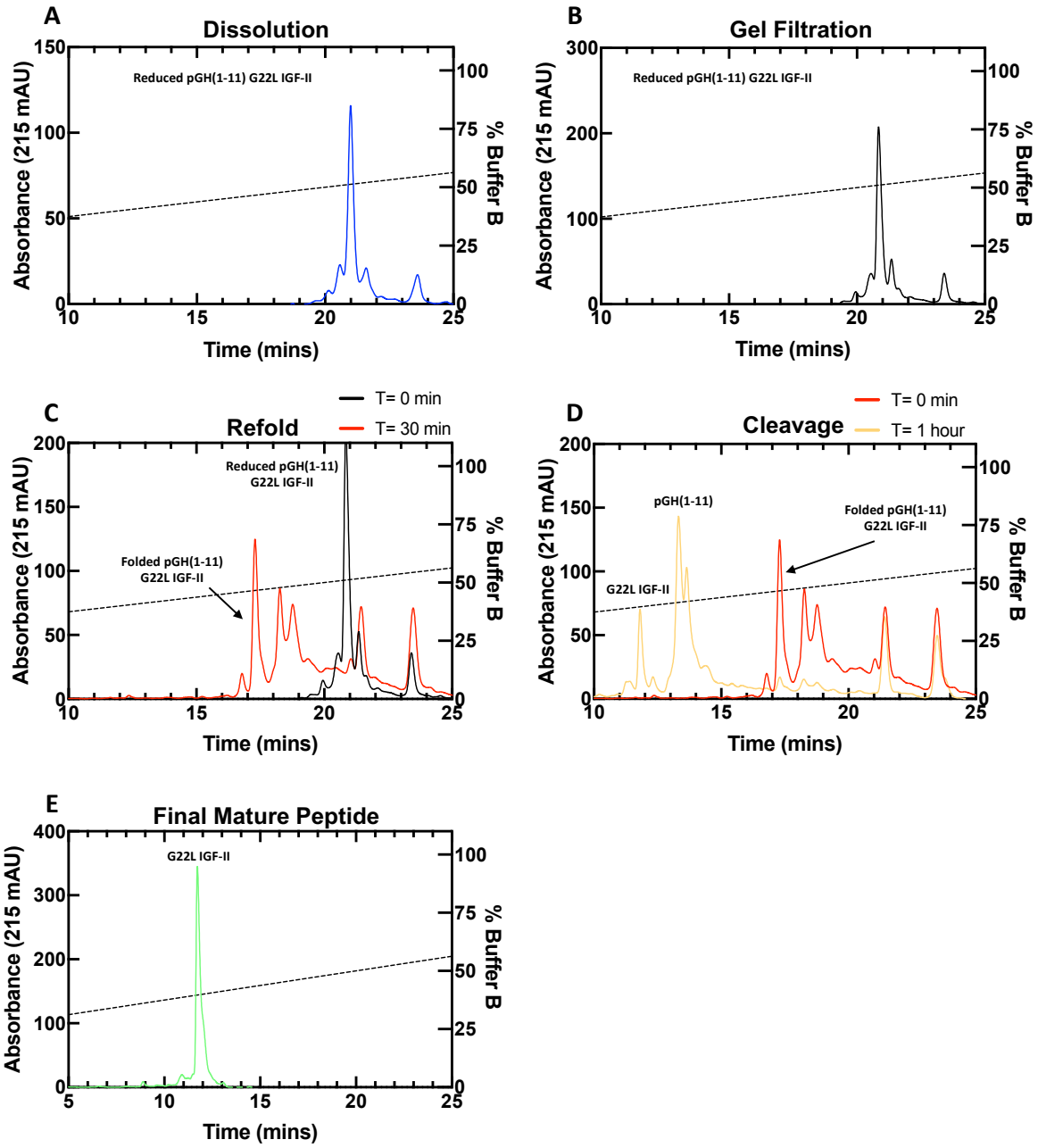


Figure 4.9: Purification of G22L IGF-II.

Analytical reverse phase HPLC (using a C4 column and a 25 – 45 % CH_3CN (31.3 – 56.3% buffer B) over 20 minutes (Section 2.3.5)) was used to monitor each stage of protein purification. Shown are the absorbance chromatograms at 215 nm of reduced pGH(1-11) G22L IGF-II (left Y axis). Dotted black line represents the % of buffer B (right Y axis). (A) Dissolution, (B) post gel filtration, (C) the refold T=0 minutes (solid black line) and T=30 minutes (red), (D) α-LP cleavage T=0 minutes (red) and T=1 hour (orange) with mature G22L IGF-II at 11.9 min and (E) final mature G22L IGF-II peptide post preparative HPLC purification (green).

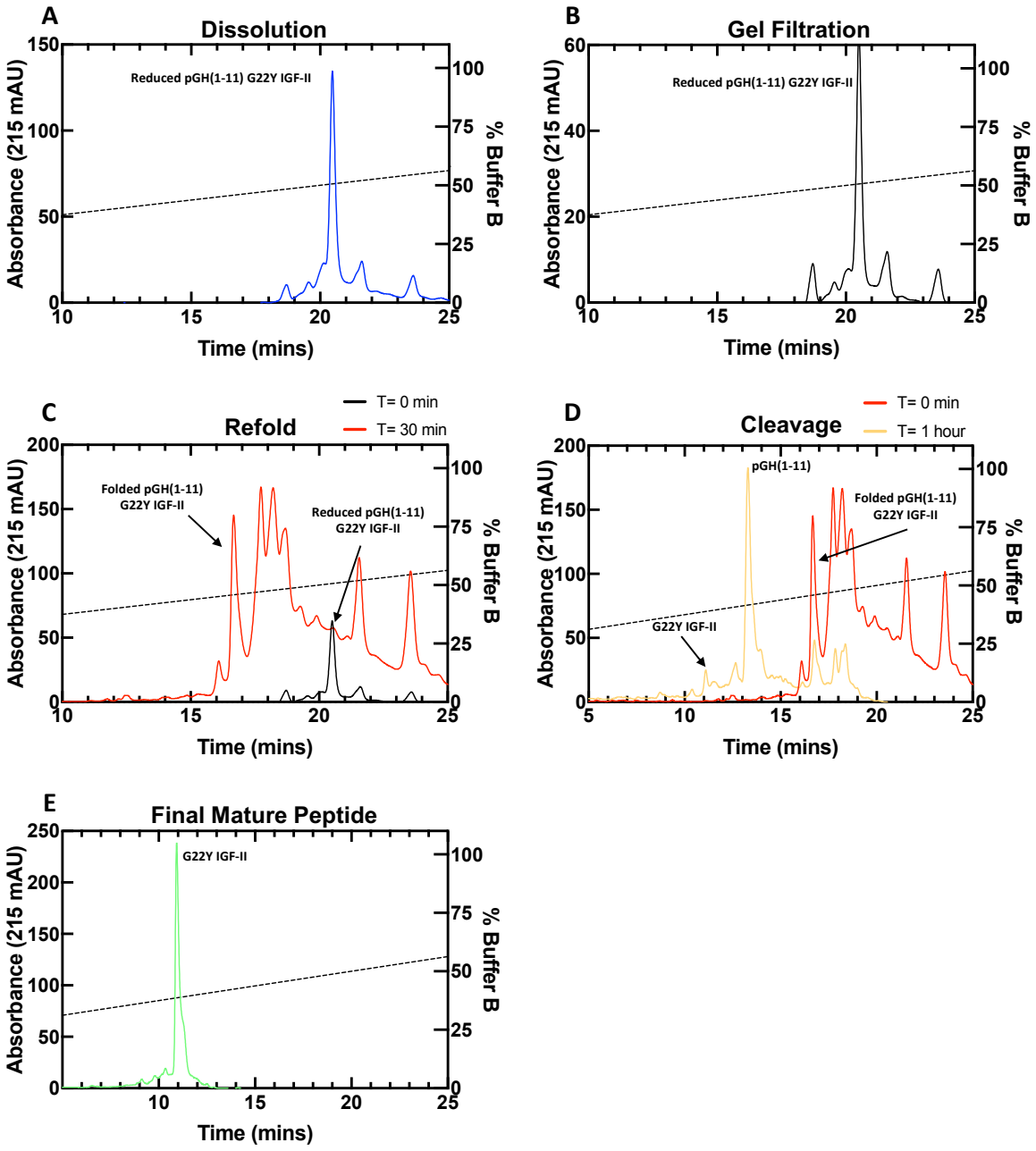


Figure 4.10: Purification of G22Y IGF-II.

Analytical reverse phase HPLC (using a C4 column and a 25 – 45 % CH_3CN (31.3 – 56.3% buffer B) over 20 minutes (Section 2.3.5)) was used to monitor each stage of protein purification. Shown are the absorbance chromatograms at 215 nm of reduced pGH(1-11) G22Y IGF-II (left Y axis). Dotted black line represents the % of buffer B (right Y axis). (A) Dissolution, (B) post gel filtration, (C) the refold T=0 minutes (solid black line) and T=30 minutes (red), (D) α-LP cleavage T=0 minutes (red) and T=1 hour (orange) with mature G22Y IGF-II at 11.0 min and (E) final mature G22Y IGF-II peptide post preparative HPLC purification (green).

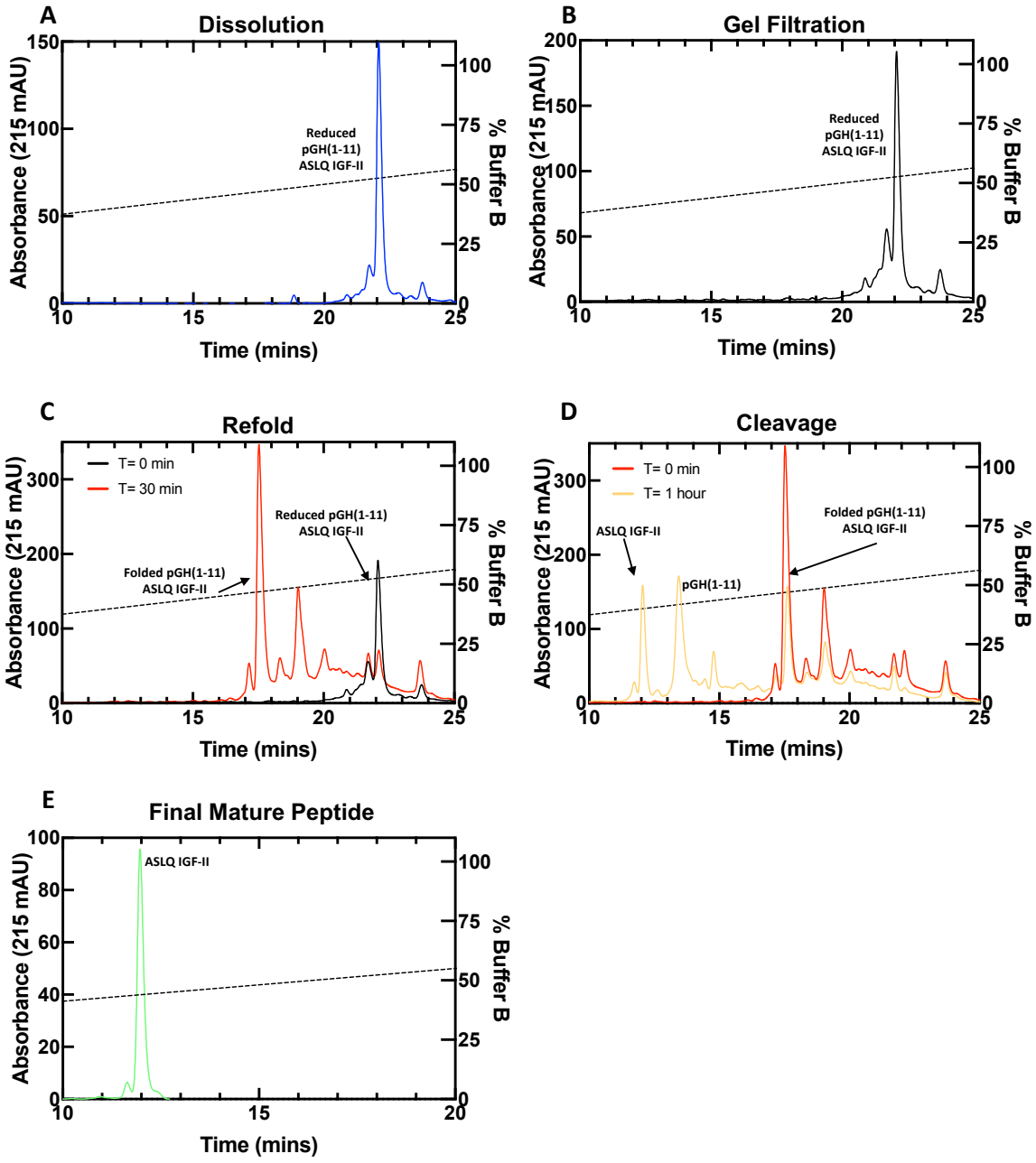
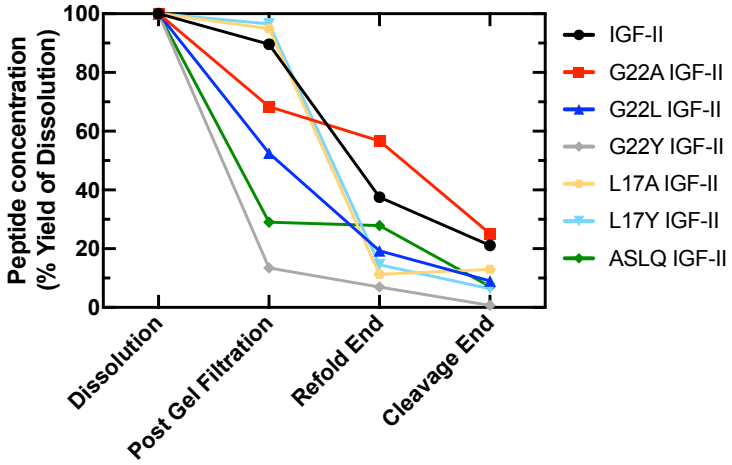


Figure 4.11: Purification of ASLQ IGF-II.

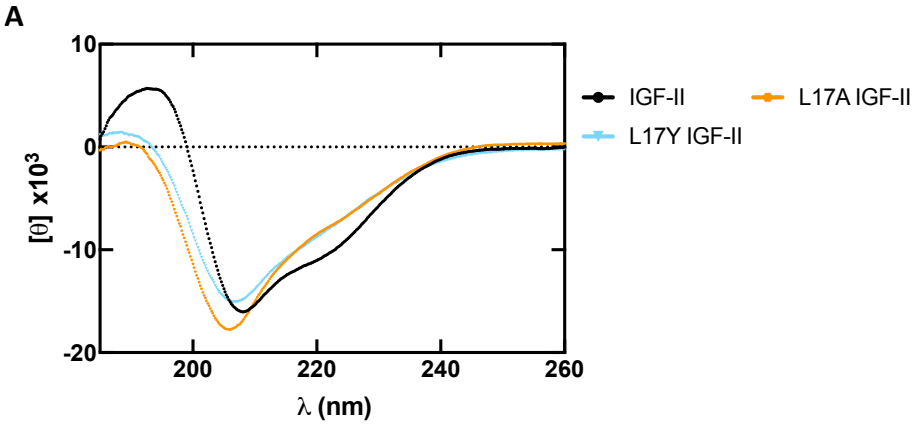
Analytical reverse phase HPLC (using a C4 column and a 25 – 45 % CH_3CN (31.3 – 56.3% buffer B) over 20 minutes (Section 2.3.5)) was used to monitor each stage of protein purification. Shown are the absorbance chromatograms at 215 nm of reduced pGH(1-11) ASLQ IGF-II (left Y axis). Dotted black line represents the % of buffer B (right Y axis). (A) Dissolution, (B) post gel filtration, (C) the refold T=0 minutes (solid black line) and T=30 minutes (red), (D) α-LP cleavage T=0 minutes (red) and T=1 hour (orange) with mature ASLQ IGF-II at 12.0 min and (E) final mature ASLQ IGF-II peptide post preparative HPLC purification (green).



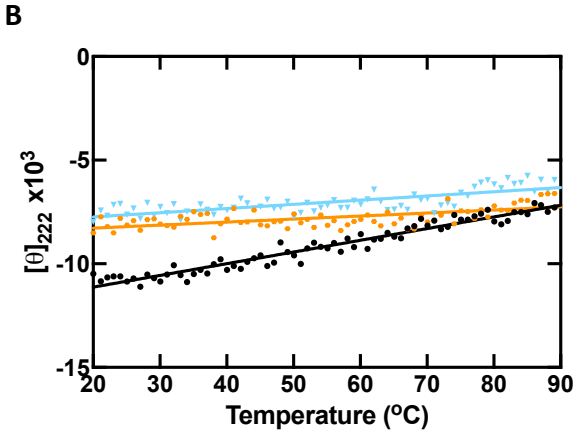
Analogue	Peptide conc. (mg/L of starting ferment) at dissolution	Peptide concentration (% of Dissolution)		
		Gel Filtration	Refold End	Cleavage End
IGF-II	49.4	89.6	37.5	21.1
G22A IGF-II	31.1	68.4	56.7	25.1
G22L IGF-II	32.5	52.5	19.3	9.03
G22Y IGF-II	39.7	13.4	6.95	0.709
L17A IGF-II	54.0	95.0	11.3	12.9
L17Y IGF-II	56.3	96.5	14.5	6.34
ASLQ IGF-II	37.4	29.0	27.8	7.16

Figure 4.12: Summary of protein purification yields for G22 IGF-II, L17 IGF-II and ASLQ IGF-II analogues.

Yield of IGF-II and each IGF-II analogue at dissolution, post gel filtration, refold end and cleavage end. Yield is expressed as a percentage of the peptide mass (μg) after dissolution (100%). Table of the protein concentration (mg/mL) at dissolution and then yield at subsequent steps, expressed as a percentage of the starting concentration of each peptide at dissolution.



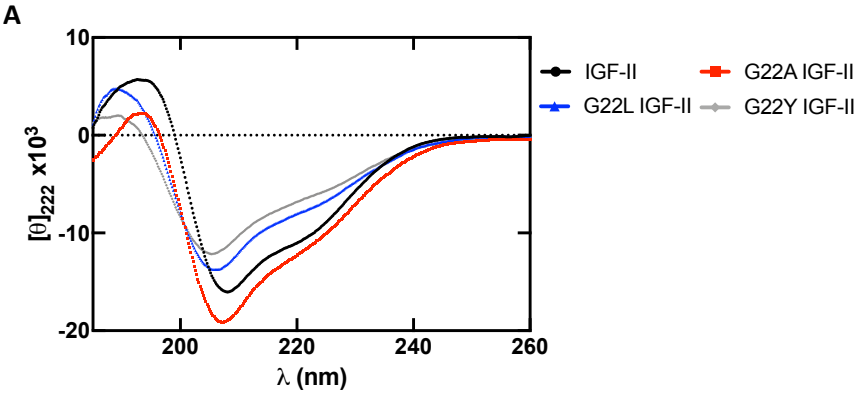
Analogue	Helix	Strand	Turns	Unordered
IGF-II	41	13	16	30
L17A IGF-II	23	23	15	40
L17Y IGF-II	23	22	15	41



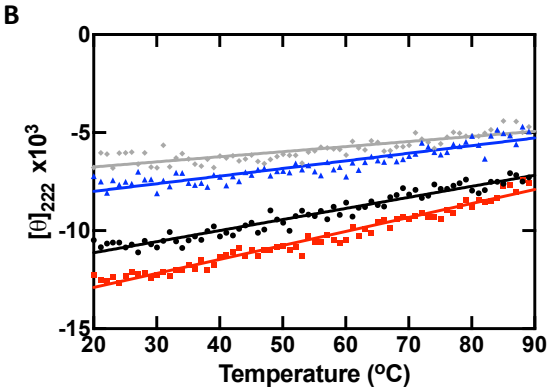
Analogue	R^2	Slope	$[\theta]_{222}x10^3$		% change from 20°C to 90°C
			20°C	90°C	
IGF-II	0.9462	0.06	-10.5	-7.26	30.9
L17A IGF-II	0.3879	0.01	-8.51	-6.48	23.9
L17Y IGF-II	0.5946	0.02	-7.98	-6.22	22.1

Figure 4.13: Analysis of L17A IGF-II and L17Y IGF-II secondary structure and thermal stability using circular dichroism.

(A) Secondary structure analysis of the L17 IGF-II analogues was performed using circular dichroism at 0.05 μg/μL in 10 mM acetic acid and mean residue ellipticity [θ] was measured from 260 - 180 nm (Section 2.4.4.1). Percent helical content shown in the table was derived using Dichroweb and was calculated using the CDSSTR algorithm for deconvolution against the reference protein database set SMP180 (Sreerama and Woody 2000). (B) Thermal stability was monitored by measuring mean residue ellipticity [θ] at 222 nm from 20°C to 90°C (at 1±0.1°C increments) (Section 2.4.4.1.1). Temperature denaturation data were analysed using a linear regression line of best fit. R^2 values, slope and percentage change in helical content from 20°C to 90°C are shown in the table.



Analogue	Helix	Strand	Turns	Unordered
IGF-II	41	13	16	30
G22A IGF-II	48	11	16	25
G22L IGF-II	23	24	13	40
G22Y IGF-II	19	24	14	44

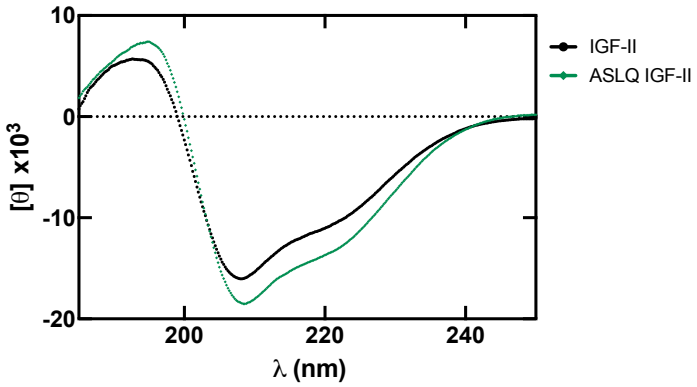


Analogue	R^2	Slope	$[\theta]_{222}x10^3$		% change from 20°C to 90°C
			20°C	90°C	
IGF-II	0.9462	0.06	-10.5	-7.26	30.9
G22A IGF-II	0.9655	0.07	-12.3	-7.31	40.6
G22L IGF-II	0.8388	0.04	-7.21	-4.69	35.0
G22Y IGF-II	0.7146	0.03	-6.78	-4.20	38.1

Figure 4.14: Analysis of secondary structure and thermal stability of G22 IGF-II analogues using circular dichroism.

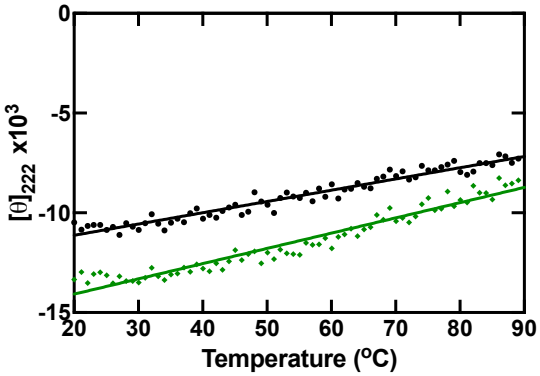
(A) Secondary structure analysis of the G22 IGF-II analogues was performed using circular dichroism at 0.05 μg/μL in 10 mM acetic acid and mean residue ellipticity [θ] was measured from 260 - 180 nm (Section 2.4.4.1). Percent helical content shown in the table was derived using Dichroweb and was calculated using the CDSSTR algorithm for deconvolution against the reference protein database set SMP180 (Sreerama and Woody 2000). (B) Thermal stability was monitored by measuring mean residue ellipticity [θ] at 222 nm from 20°C to 90°C (at 1±0.1°C increments)(Section 2.4.4.1.1). Temperature denaturation data were analysed using a linear regression line of best fit. R^2 values, slope and percentage change in helical content from 20°C to 90°C are shown in the table.

A



Analogue	Helix	Strand	Turns	Unordered
IGF-II	41	13	16	30
ASLQ IGF-II	49	12	16	23

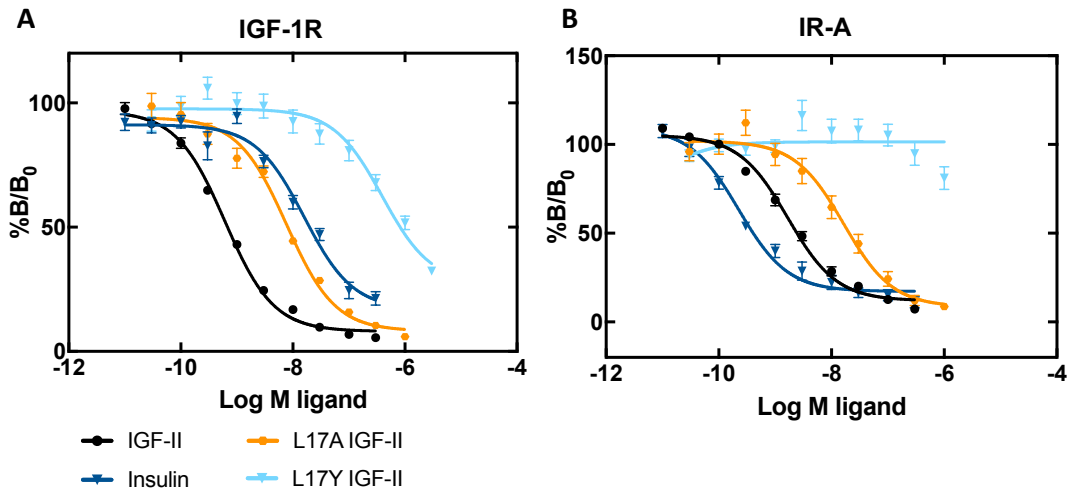
B



Analogue	R^2	Slope	Ellipticity value (222 nm) 10^3		% change from 20°C to 90°C
			20°C	90°C	
IGF-II	0.9462	0.06	-10.5	-7.26	30.9
ASLQ IGF-II	0.9382	0.08	-13.3	-8.55	35.7

Figure 4.15: Analysis of secondary structure and thermal stability of ASLQ IGF-II analogues using circular dichroism.

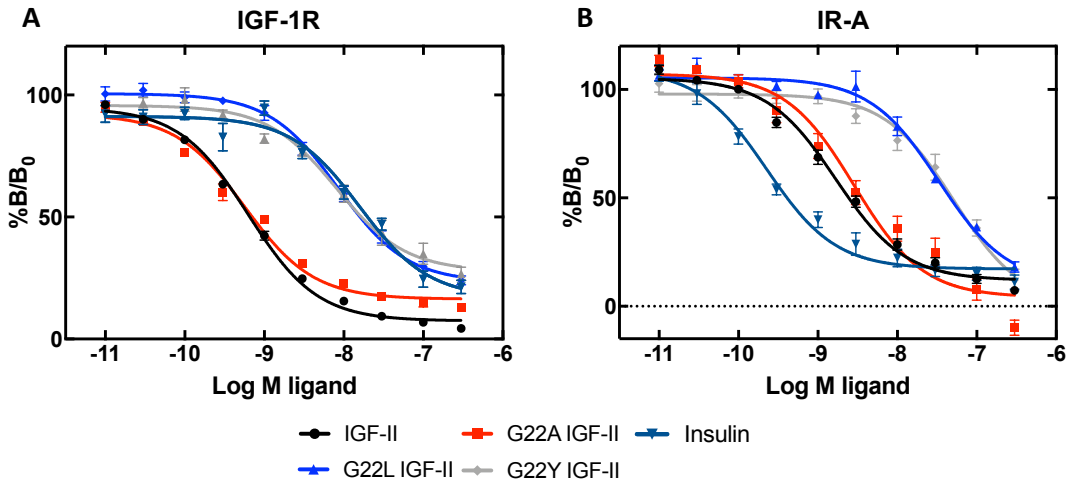
(A) Secondary structure analysis of the ASLQ IGF-II analogues was performed using circular dichroism at 0.05 μg/μL in 10 mM acetic acid and mean residue ellipticity [θ] was measured from 260 - 180 nm (Section 2.4.4.1). Percent helical content shown in the table was derived using Dichroweb and was calculated using the CDSSTR algorithm for deconvolution against the reference protein database set SMP180 (Sreerama and Woody 2000). (B) Thermal stability was monitored by measuring mean residue ellipticity [θ] at 222 nm from 20°C to 90°C (at 1±0.1°C increments)(Section 2.4.4.1.1). Temperature denaturation data were analysed using a linear regression line of best fit. R^2 values, slope and percentage change in helical content from 20°C to 90°C are shown in the table.



Analogue	IGF-1R			IR-A		
	IC_{50} (nM)	95% CI (nM)	Relative binding affinity (% of IGF-II)	IC_{50} (nM)	95% CI (nM)	Relative binding affinity (% of IGF-II)
IGF-II	0.626	0.568-0.689	100	1.68	1.36-2.09	100
Insulin	15.9	10.7-23.7	3.94	0.226	0.156-0.329	743
L17A IGF-II	7.68	5.73-10.3	0.0815	16.5	10.6-26.4	10.2
L17Y IGF-II	>100	ND	ND	>1000	ND	ND

Figure 4.16: Competition of L17A and L17Y IGF-II with europium-labelled IGF-II for binding the IGF-1R and IR-A

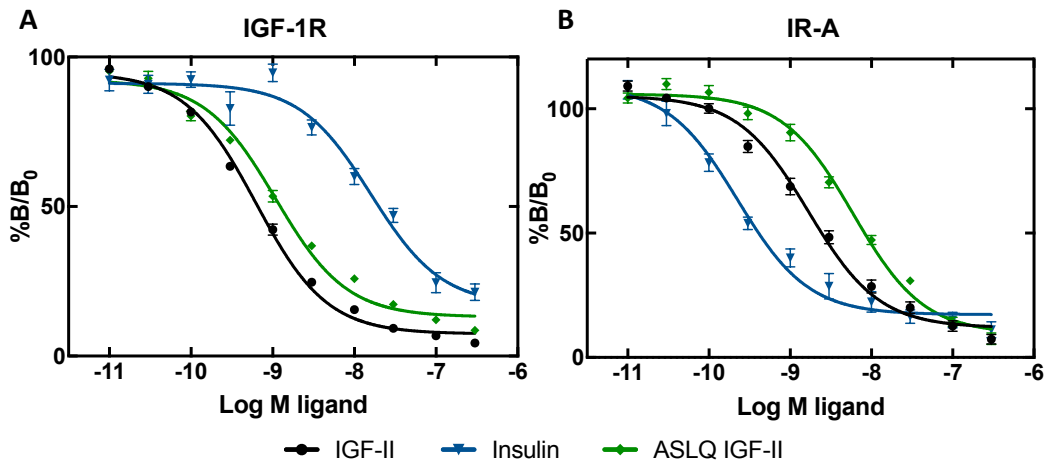
Competition binding of each analogue with the IGF-1R (A) or IR-A (B). Immunocaptured IGF-1R or IR-A was incubated with europium-labelled IGF-II (Eu-IGF-II) and increasing concentrations of competing L17A IGF-II or L17Y IGF-II. Time-resolved fluorescence was measured as described in Section 2.4.1. Results are expressed as a percentage of Eu-IGF-II bound in the absence of competing ligand (B/B_0). Data shown are the mean ±S.E.M. n = ≥ 3 independent experiments each with triplicate technical replicates. Error bars are shown when greater than the size of the symbols. Calculated IC_{50} values, 95% confidence interval (95% CI) relative binding affinity relative to human IGF-II are shown in the table.



Analogue	IGF-1R			IR-A		
	IC_{50} (nM)	95% CI (nM)	Relative binding affinity (% of IGF-II)	IC_{50} (nM)	95% CI (nM)	Relative binding affinity (% of IGF-II)
IGF-II	0.626	0.568-0.689	100	1.68	1.36-2.09	100
Insulin	15.9	10.7-23.7	3.94	0.226	0.156-0.329	743
G22A IGF-II	0.565	0.439-0.726	111	2.87	1.79-4.64	58.5
G22L IGF-II	8.08	5.37-12.1	7.75	33.4	23.9-47.0	5.03
G22Y IGF-II	8.45	6.50-11.0	7.41	51.3	31.4-85.4	3.27

Figure 4.17: Competition of G22 IGF-II analogues with europium-labelled IGF-II for binding the IGF-1R.

Competition binding of each analogue with the IGF-1R (A) or IR-A (B). Immunocaptured IGF-1R or IR-A was incubated with europium-labelled IGF-II (Eu-IGF-II) and increasing concentrations of competing G22A IGF-II, G22L IGF-II or G22Y IGF-II. Time-resolved fluorescence was measured as described in Section 2.4.1. Results are expressed as a percentage of Eu-IGF-II bound in the absence of competing ligand (B/B_0). Data shown are the mean ±S.E.M. n = ≥ 3 independent experiments each with triplicate technical replicates. Error bars are shown when greater than the size of the symbols. Calculated IC_{50} values, 95% confidence intervals (95% CI), relative binding affinity (binding affinities relative to human IGF-II as a percent) are shown in the table.



Analogue	IGF-1R			IR-A		
	IC_{50} (nM)	95% CI (nM)	Relative binding affinity (% of IGF-II)	IC_{50} (nM)	95% CI (nM)	Relative binding affinity (% of IGF-II)
IGF-II	0.626	0.568-0.689	100	1.68	1.36-2.09	100
Insulin	15.9	10.7-23.7	3.94	0.226	0.156-0.329	743
ASLQ IGF-II	1.08	0.894-13.1	58.0	5.95	4.80-7.38	28.2

Figure 4.18: Competition of ASLQ IGF-II analogues with europium-labelled IGF-II for binding the IGF-1R.

Competition binding of each analogue with the IGF-1R (A) or IR-A (B). Immunocaptured IGF-1R or IR-A was incubated with europium-labelled IGF-II (Eu-IGF-II) and increasing concentrations of competing ASLQ IGF-II or IGF-II. Time-resolved fluorescence was measured as described in Section 2.4.1. Results are expressed as a percentage of Eu-IGF-II bound in the absence of competing ligand (B/B_0). Data shown are the mean ±S.E.M. n = ≥ 3 independent experiments each with triplicate technical replicates. Error bars are shown when greater than the size of the symbols. Calculated IC_{50} values, 95% confidence interval (95% CI) relative binding affinity relative to human IGF-II are shown in the table.

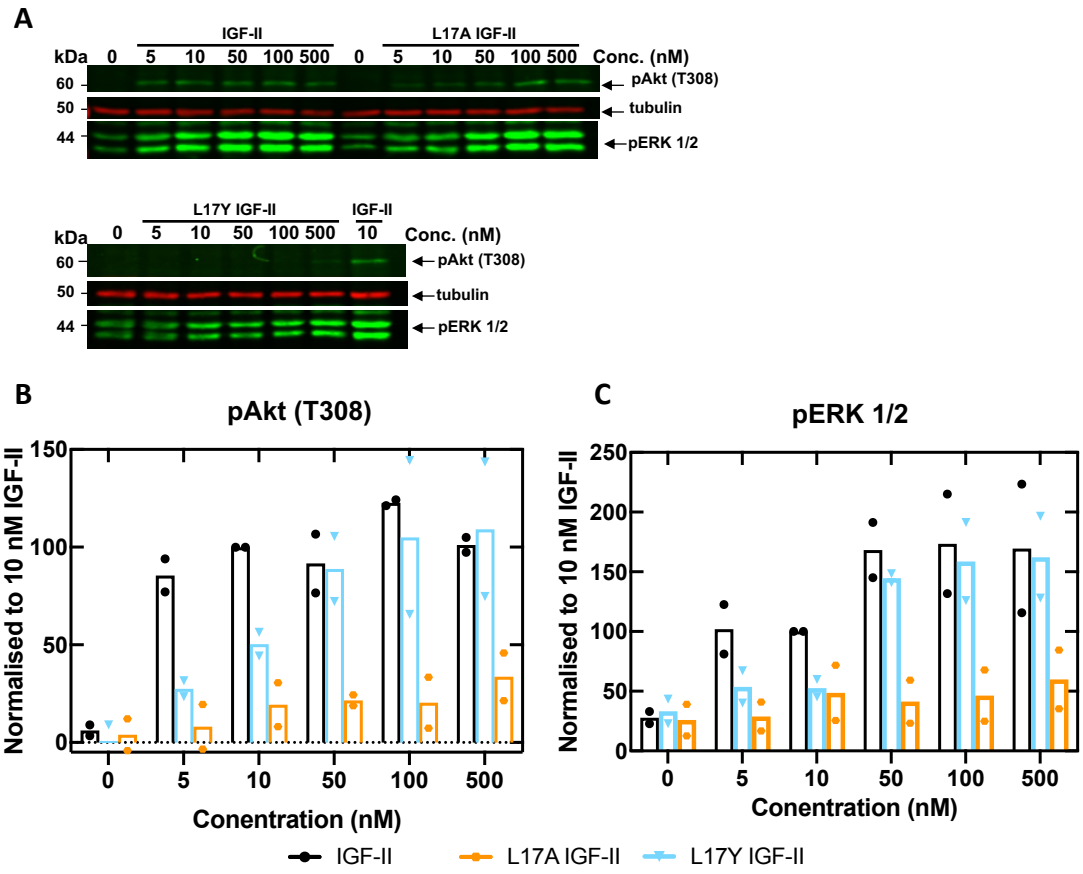


Figure 4.19: Induction of Akt and ERK phosphorylation upon IR-A activation by IGF-II and L17 IGF-II analogues.

L6 IR-A cells were treated with IGF-II, L17A IGF-II or L17Y IGF-II at 10, 50, 100 and 500 nM for 10 min (Section 2.4.2). Whole cell lysates were prepared and subjected to SDS-PAGE followed by immunoblotting for phosphorylated Akt (pAkt (T308)) and phosphorylated ERK 1/2 (pERK 1/2) (A). Each blot included lanes from cells untreated (at 10 min) and treated with 10 nM IGF-II for 10 min. Densitometric quantitation of n = 2 independent experiments are shown as a bar graph with individual values from each experiment. Relative pAkt (B) and pERK 1/2 (C) are expressed as a fraction of the level detected when cells were stimulated with 10 nM IGF-II for 10 min. There is no change in total Akt and ERK 1/2 at these time points. In each case, pAkt (T308) and pERK 1/2 was first normalized against the loading control (β-tubulin). Refer to Supplementary Figure 13 for uncropped Western Blot images.

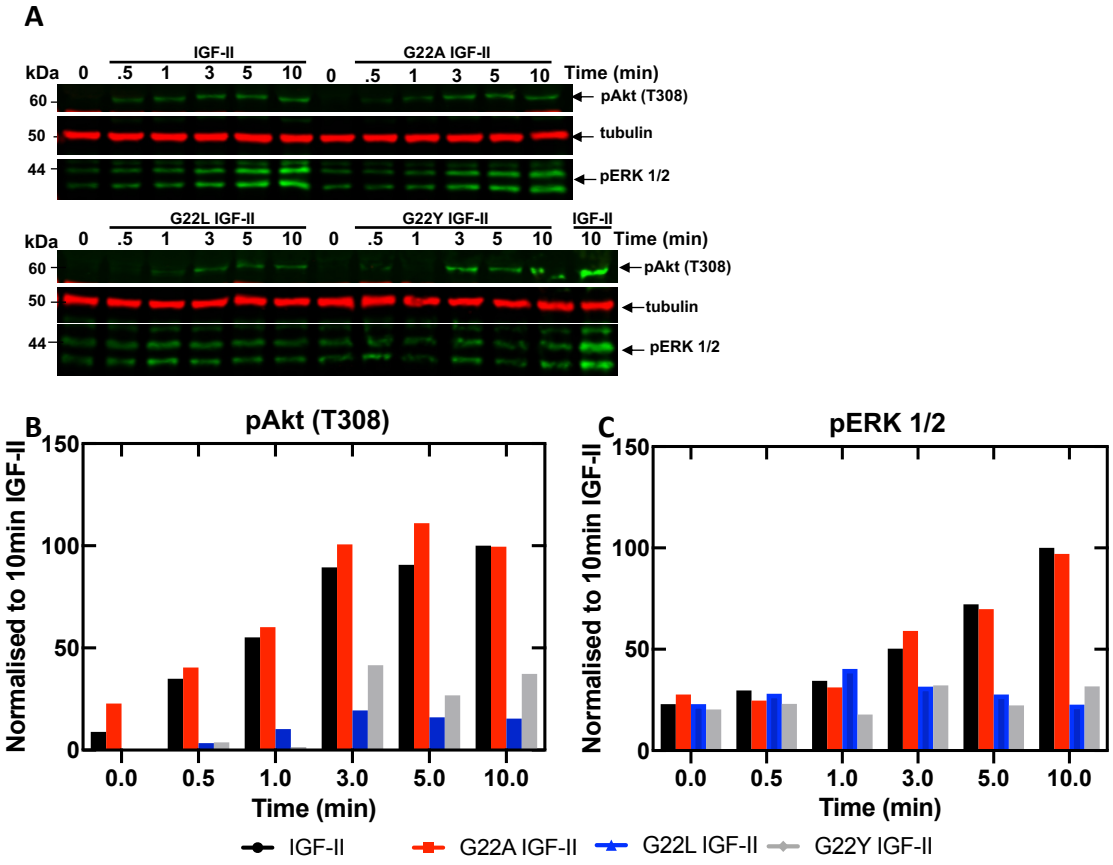


Figure 4.20: Induction of Akt and ERK phosphorylation upon IR-A activation by IGF-II and G22 IGF-II analogues.

L6 IR-A cells were treated with IGF-II, G22A IGF-II, G22L IGF-II, or G22Y IGF-II at 10 nM for 0, 0.5, 1, 3, 5 and 10 min (Section 2.4.2). Whole cell lysates were prepared and subjected to SDS-PAGE followed by immunoblotting for phosphorylated Akt (pAkt (T308)) (B), phosphorylated ERK 1/2 (pERK 1/2) (C). Blot included lanes from cells untreated (basal = 0 min) and treated with 10 nM IGF-II for 10 min. This experiment was completed as n=1. Relative pAkt (T308) (C) and pERK (D) are expressed as a fraction of the level detected when cells were stimulated with 10 nM IGF-II for 10 min. In each case, pAkt (T308) and pERK 1/2 was first normalized against the loading control (β-tubulin). Refer to Supplementary Figure 12 for uncropped Western Blot images.

4.3 Discussion

To compensate for the lack of the FFY motif in the C-terminus of the B-domain, Con-Ins G1 has 5 notable amino acid changes that allow for reasonably high affinity IR binding, namely YB15, YB20 (which contact site 1a and bind to the region normally occupied by FB24) and HA8, RA9 and EB10 (which contact site 1b) (Xiong et al., 2020, Menting et al., 2016). In contrast, Con-Ins K1 insulin does not have a tyrosine at position 15 or 20 but has an A-chain extension that interacts with site 1a to compensate for the lack of the FFY motif (Xiong et al., 2021). Incorporating these substitutions into human DOI (Con-Ins G1-like insulin (Mini-Ins) and Con-Ins K1-like insulin (Vh-Ins)), in both cases leads to the ability to bind the IR with the same affinity as human insulin (Xiong et al., 2021, Xiong et al., 2020). However, on closer analysis Mini-Ins is seen to have a reduced affinity (128-fold reduction) for site 1a compared to native insulin (Xiong et al., 2020). Instead, Mini-Ins is proposed to have a stronger affinity for site 1b. This mode of binding leads to slower kinetics of mitogenic (ERK1/2 pathway) activation (Xiong et al., 2020). Conversely, Vh-Ins, has strong site 1a interactions (9.91-fold reduction in affinity), weak site 1b interactions and the same mitogenic potency as human insulin (Xiong et al., 2021). Based on the results of these engineered insulin analogues we now speculate that strength of interaction at site 1a and site 1b may be a determinant of signaling outcome.

Here we sought to ascertain what effect the key contacts from Con-Ins G1 and Con-Ins K1 have on binding and signaling potency in the context of IGF-II. We first investigated several IGF-II mutants at residue G22, equivalent to residues YB20 in Con-Ins G1 and LB20 in Con-Ins K1 which contribute to site 1a contact. A glycine is highly conserved at the equivalent position in insulin-like peptides across vertebrate species. Its conservation reflects the critical role this residue plays in facilitating the type 1 β turn of the C-terminus of the B-domain, which allows the end of the B-domain to fold back against the B-domain helix (Torres et al., 1995) (Figure 4.5). The same role has been reported for the equivalent residue in insulin (GB20) (Blundell et al., 1972, Timofeev et al., 2010, Nakagawa et al., 2006). Chiral mutagenesis of insulin demonstrated the importance of D-glycine amino acids for the two β turns in the B-chain; L-glycine amino acids are unfavourable (Nakagawa et al., 2006, Nakagawa et al., 2005, Menting et al., 2014). Interestingly, it was also found that substitution of residue GB20 to L-alanine impeded disulfide pairing (Nakagawa et al., 2006).

It has been established for insulin that a conserved microdomain surrounding the CA20-

CB19 disulfide bond exists, which forms early during oxidative refolding and aligns the CA20 and CB19 thiols to allow for disulfide pairing (Weiss, 2009, Hua et al., 2006b, Huang et al., 2004, Weiss et al., 2002b). In the case of IGF-II, while this phenomenon has not been confirmed experimentally, structural similarities would suggests that a similar microdomain likely forms around the equivalent C18-C61 disulfide bond. Given this, it was not surprising that instead of glycine at IGF-II position 22, which neighbours the C18-C61 disulfide bond, the introduction of the more bulky and hydrophobic leucine or tyrosine caused significant alteration to the efficiency of folding (Figure 4.12) and the final secondary structure (Figure 4.14A). In Con-Ins G1 and indeed many of the cone snail venom insulins, a range of residues are tolerated at position B20 as the β turn is no longer required due to the lack of the 8 C-terminal amino acids of the B-chain (Ahorukomeye et al., 2019, Menting et al., 2016). This β turn is essential in human insulin, to allow the C-terminal segment of the B-chain to fold back against the B-chain helix. However, when the C-terminal 8 amino acids are removed in human insulin, the type 1 β turn is no longer essential and residues such as tyrosine can be accommodated at position B20 (Xiong et al., 2020). In the context of IGF-II, here we demonstrate that the type 1 β turn is essential for IGF-II structure and it is likely that introduction of bulky side chains such as that of tyrosine and leucine perturb this turn from forming. The effect of these mutations could be further investigated using structural determination by nuclear magnetic resonance (NMR).

The introduction of alanine at position 22 resulted in a higher efficiency of refolding compared to that of native IGF-II (Figure 4.12), suggesting the integrity of this microdomain is maintained and its role in folding is not perturbed by introduction of alanine. The equivalent mutation in insulin (GB20A insulin) was reported to have low yield (~10%) using a yeast expression system (Kristensen et al., 1997). Although different methods of refolding were employed in the current study compared to the Kristensen paper, it appears that an alanine at position 22 of IGF-II may enhance the efficiency of refolding. It must be noted that the expression and purification of all analogues in this investigation was performed once, and repetition of this would be required to confirm these findings. Overall, the IGF-II G22 mutants confirm the importance of maintaining the local structure surrounding the C18-C61 disulfide bond for efficient folding.

Another interesting finding was that G22A IGF-II resulted in essentially equal affinity (111%) for the IGF-1R compared to native IGF-II. A GB20A insulin mutant resulted in an

affinity 270% for the IR compared to that of human insulin (Kristensen et al., 1997), whereas G22A IGF-II had a 2-fold lower affinity for the IR-A compared to native IGF-II. Evidently introduction of an alanine into B20 of insulin has a greater effect on IR binding than it does for IGF-II binding to the IGF-1R. Additionally, our CD data implies that an alanine at position 22 of IGF-II facilitates the extension of the B-domain helix. Indeed, the same conclusion was made for GB20 insulin (Kristensen et al., 1997), but the mechanisms by which this helix extension leads to increased receptor binding affinity of GB20A insulin or why the equivalent mutation in IGF-II leads to a lowering of affinity are not fully understood.

The ability of each of these G22 IGF-II analogues to stimulate phosphorylation of Akt and ERK1/2 was tested and found that G22A IGF-II had essentially equal potency of Akt and ERK1/2 activation as native IGF-II. G22L IGF-II and G22Y IGF-II had lower potency of activation of Akt and ERK1/2 activation compared to native IGF-II however further testing is required to determine the statistical significance. There are no reports of the ability of the equivalent insulin analogue (B20A insulin (Kristensen et al., 1997)) to stimulate downstream signaling.

In native IGF-II and insulin, residue L17 and LB15, respectively, are located in the middle of the B-domain helix with each side chain positioned toward the core of the protein (Torres et al., 1995, Blundell et al., 1972). Mutation of LB15 to alanine in insulin was not tolerated using a yeast expression system (Kristensen et al., 1997). Combined these structural and mutational data for insulin implicate this residue as being essential for maintenance of the core structure. In this investigation, an alanine and a tyrosine were introduced into position 17 of IGF-II. During the refolding stage there were major losses of L17A IGF-II and L17Y IGF-II, respectively (Figure 4.12). In addition, the final purified peptides had markedly lower helical content compared to native IGF-II (Figure 4.13), suggesting that the mutations introduced affect secondary structure. Interestingly under the conditions used for refolding there were very few folding intermediates observed with native IGF-II but in folding the L17A IGF-II and L17Y IGF-II multiple intermediates were evident. It was certainly anticipated that some of these introduced mutations may not be tolerated due to the high degree of conservation of these residues.

Overlays of IGF-II and Con-Ins G1 (Figure 4.21) reveal that a potential steric clash might arise following introduction of a bulkier tyrosine at IGF-II residue 17 with F26 of the IGF-II FYF motif. Residue F26 of IGF-II appears to occupy the same space as the side chain of YB15 Con-Ins G1. Such a steric clash is avoided in Con-Ins G1 due to the lack of the

FFY motif within the missing last 8 residues of the B-chain. Thus, tyrosine at residue 17 of IGF-II may be better tolerated if the equivalent site 1a residues, including F26 and other elements of the FYF motif, are mutated to smaller side chains such as alanine. This hypothesis will be explored in Chapter 6. In future experiments we aim to further explore the structure of these analogues. While we have confirmed the presence of the three disulfide bonds through mass spectrometry, next we aim to confirm the correct disulfide bond connectivity through additional NMR studies, as used to identify disulfide-swap IGF-I (Miller et al., 1993), or through digestion of folded IGF-II by pepsin and subsequent mass spectrometry (Milner et al., 1999).

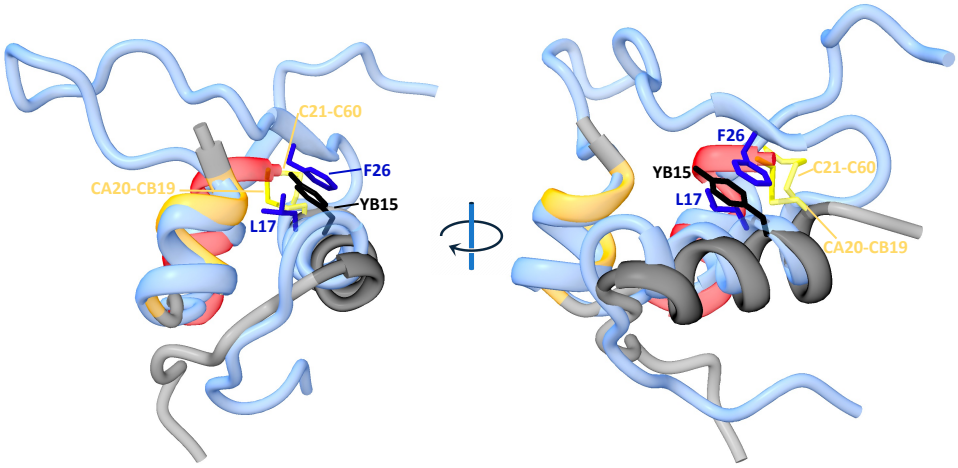


Figure 4.21: Structural overlay of IGF-II and Con-Ins G1 to compare the spatial position of F26 of IGF-II and YB15 of Con-Ins G1.

(B) Con-Ins G1 (grey) and IGF-II (blue) overlay (PDB: 5JYQ and 1IGL respectively). Con-Ins G1 helices coloured: B-domain helix 1 (black), A-domain helix 2 (orange) and A-domain helix 3 (red). Side chain of residue YB15 in Con-Ins G1 is positioned into the core of the protein and occupies a similar space as residue F26 in IGF-II. These residues are also positioned close to the CA20-CB19 (insulin)/ C21-C60 (IGF-II) disulfide bond (shown in yellow).

IGF-I has an affinity for the IR-A 15% of that of IGF-II. However, replacement of IGF-I D-domain with IGF-II D-domain increases affinity to 22%, compared to IGF-II (Denley et al., 2004). Replacement of IGF-I C-domain with that of IGF-II results in an affinity of 28% for the IR-A, compared to IGF-II. Interestingly, with both C- and D-domains of IGF-II introduced into IGF-I a near equal affinity for the IR-A (93%), compared to IGF-II is achieved (Denley et al., 2004). Total removal of the D-domain in IGF-II was reported to reduce binding affinity for the IGF-1R to 18% compared to native IGF-II (Roth et al., 1991). An interesting finding from the cone snail venom insulins characterised in recent years is the presence of an A-chain extension in Con-Ins K1, equivalent to the D-domain in IGF-II. While solution structures of IGF-II depict the D-domain having no ordered structure (Torres et al., 1995), structures of a modified human insulin analogue with a variation of an A-chain extension based on Con-Ins K1 was found to result in an extension of the A-chain helix (Xiong et al., 2021). Here we explored the replacement of the native IGF-II D-domain with variation of the Vh-Ins A-chain extension, A61, S62, L63, Q64 IGF-II. CD studies of ASLQ IGF-II infer that this modified D-domain also extends the IGF-II C-terminal A-chain helix.

ASLQ IGF-II bound to the IGF-1R and IR-A with a lower affinity relative to native IGF-II (2-3 fold). It is interesting that this modified D-domain retains slightly higher affinity for the IGF-1R than the IR-A. Previous reports suggest that the IGF-II D-domain has a minor influence on its specificity for the IR-A (Denley et al., 2004). Structural overlays of IGF-II and Vh-Ins bound to their cognate receptors (Figure 4.4) indicates that residue LA23 of the Vh-Ins A-chain extension occupies the same space as F26 in IGF-II (Xiong et al., 2021). Replacement of the native IGF-II D-domain with the modified D-domain (ASLQ) was tolerated in IGF-II from a structural perspective, as confirmed by CD studies. In Chapter 6 the aim was to then test if the modified D-domain is capable of engaging with the hydrophobic pocket of the L1 domain in place of residue F26, as was seen for the D-domain of Vh-Ins, which compensates for the missing FFY motif by engaging with the same hydrophobic pocket as FB24 (Xiong et al., 2021). This was achieved by introducing the modified D-domain into the site 1a minimised IGF-II analogues described in Chapter 3.

In summary, the aim of this chapter was to investigate several residues identified in Con-Ins G1 and Con-Ins K1, that compensate for the loss of the aromatic FFY motif, in the context of IGF-II. Of the mutants generated G22Y and L17Y were chosen to explore in

the context of site 1 minimised analogues produced in Chapter 3. The other mutations need to be further investigated for their folding status before being incorporated into a site 1 minimised IGF-II analogues.

5 INVESTIGATION OF THE ROLE OF RESIDUE 18 OF IGF-II IN PEPTIDE STABILITY, RECEPTOR BINDING AND ACTIVATION

The following chapter was published in *Scientific Reports:*

BLYTH, A., ORTIZ, M., MERRIMAN, A., DELAINE, C. & FORBES, B. 2022. Determinants of IGF-II influencing stability, receptor binding and activation. Sci Rep, 12, 4695.

Author contributions:

Conceptualization, AB and BF. Analogue production, AB and MO. Receptor binding assays, CD AB and MO. Receptor activation assays, CD. DNA synthesis assays, AM. Analysis, AB. Writing, AB and BF. Review and editing, AB, BF, CD, AM.

5.1 Introduction

As described in chapter 4, several residues of *Conus geographus* venom insulin 1 (Con-Ins G1) and *C. kinoshitai* venom insulin 1 (Con-Ins K1), that have been shown to compensate for the loss of the site 1a aromatic FFY motif, were individually investigated in IGF-II. Sequence alignments of human insulin, IGF-II and cone snail venom insulins reveal major variation in the residue located at position B16 (position 18 in IGF-II) (Figure 4.1A). In human insulin, a tyrosine has been found to be important for IR specificity (Gauguin et al., 2008b). At the equivalent residue in IGF-II, Con-Ins G1 and *C. tulipa* venom insulin 1A (Con-Ins T1A) a glutamine, methionine or isoleucine, respectively, is found. The apparent lack of conservation at this position sparked our interest. We aimed to mutate residue Q18 of IGF-II to the equivalent residues found in human insulin, Con-Ins G1 and Con-Ins T1A to determine their effect on receptor binding and biological activity.

A series of studies have implicated the human insulin residue YB16, positioned within the insulin B-chain α-helix (Figure 5.1C), as a key residue in ligand structural stability and dimerization. It plays a role in maintenance of the B-chain helix structure, with mutation to proline perturbing disulfide pairing and impairing secretion in mammalian cells (Sun et al., 2020). Furthermore, deletion of this residue has been linked to neonatal diabetes, presumably due to a similar perturbation of the B-chain helical structure (Colombo et al., 2008). Not only is the residue involved in maintenance of structural integrity but it also plays a role in insulin dimerization. Mutation to histidine or alanine results in a monomeric insulin (Chen et al., 2000, Ludvigsen et al., 1994). Interestingly, in the cone snail venom insulins Con-Ins G1 and Con-Ins T1A a methionine and an isoleucine, respectively are located at position B16. Due to the importance of a tyrosine residue at this position in human insulin we were interested in investigating the effect of mutating the equivalent residue in human IGF-II, a non-aromatic glutamine at position Q18 (Figure 5.1C), to a methionine and a leucine as found in Con-Ins G1 and Con-Ins T1A and additionally to a tyrosine as found in human insulin. In this study we explore the impact of these mutations on stability and aggregation.

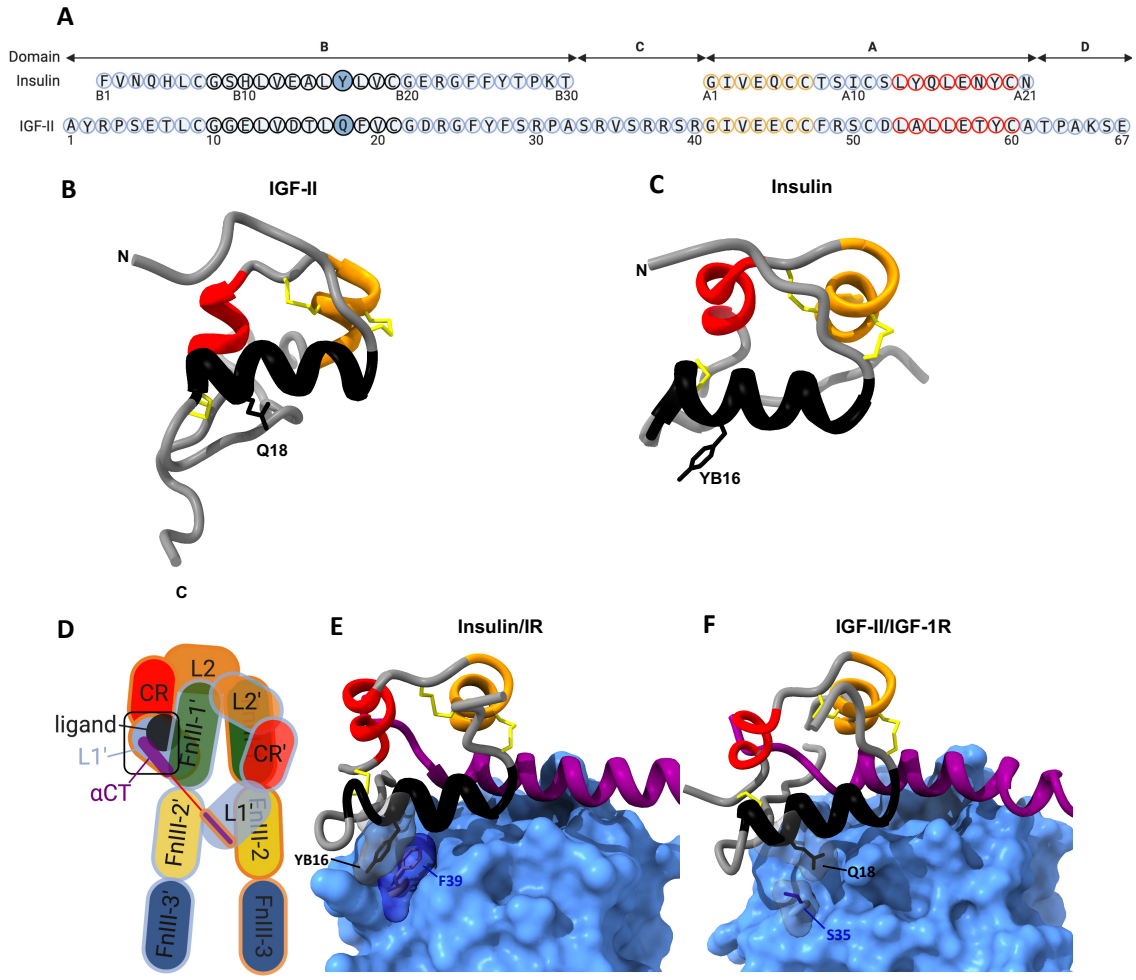

A
Domain
B
C
A
D
Insulin
B1
B10
B20
B30
A1
A10
A21
IGF-II
1
10
20
30
40
50
60
67
B
IGF-II
N
Q18
C
C
Insulin
N
YB16
D
CR
L2
L2'
ligand
L1'
αCT
FnIII-1'
CR'
FnIII-2'
L1'
FnIII-2
FnIII-3'
FnIII-3
E
Insulin/IR
YB16
F39
F
IGF-II/IGF-1R
Q18
S35

Figure 5.1: Sequence and structural comparison of IGF-II and insulin.

(A) Sequence alignment of IGF-II and insulin. Domains are indicated above. Each peptide has three alpha helices; B-chain helix 1 (black circles), A-chain helix 2 (orange circles) and A-chain helix 3 (red circles). Residue numbers are indicated below each sequence. Insulin YB16 and IGF-II Q18 residues are in solid filled circles. (B) IGF-II and (C) insulin ribbon structures (PDB: 1IGL and 1MSO respectively) show the three disulfide bonds and helices coloured as in (A). Side chains of YB16 in insulin and Q18 in IGF-II are shown. (D) Schematic diagram representing the IR and IGF-1R ligand bound extracellular domain structures. Individual $\alpha\beta$ monomer outlines are coloured either blue or orange. Extracellular domains include the first and second leucine-rich repeat domains (L1 and L2), cysteine-rich domain (CR), first, second and third fibronectin type-III domains (FnIII-1, -2, and -3), insert domain (ID), α-chain C-terminal region (αCT). Transmembrane and intracellular domains, including the tyrosine kinase domain, are not shown. (E) Insulin and (F) IGF-II bound to site 1 of their cognate receptors, as highlighted with the black box in (D) (from PDB: 6HN5 and 6VWI, respectively). L1 domains are surface filled in cornflower blue, α-chain C-terminal regions (αCT) are coloured purple. The side chains of insulin residue YB16 and F39 of the IR L1 domain and the equivalent residues Q18 in IGF-II and S35 of the IGF-1R L1 domain are shown with transparent surface fill.

Our previous studies have identified sequence differences between IGF-II and insulin that account for their receptor binding specificities. Determinants in the IGF-II B-, A- and C-domains contribute to high affinity binding to the IGF-1R and also account for the 8-10-fold lower affinity for the IR-A compared to insulin (Hexnerova et al., 2016, Gauguin et al., 2008b). Site-directed mutagenesis studies of the IGF-II and insulin ligands have defined two receptor binding sites located on opposite surfaces of the ligand. The so-called “site 1” has been characterized as the primary binding site and involves a high affinity interaction with the receptor, whereas site 2 residues have been characterized as contributing to a low affinity interaction (Alvino et al., 2009, De Meyts et al., 1978, De Meyts, 2015). The site 1 surfaces play similar roles in binding to both the IGF-1R and the IR (Weis et al., 2018, Xu et al., 2020).

The IGF-1R and IR have identical domain structures (Figure 5.1D) comprising 2α and 2β subunits. Ligand binding mediates major structural rearrangement of the receptors from an open ‘Λ-shape’ to the activated ‘J-shape’ (Xu et al., 2020, Weis et al., 2018, Lawrence, 2021, Blyth et al., 2020). How specific interactions between particular ligand residues and the receptor result in activation of the downstream signaling pathways is only partially understood. Structural studies of the ligand-bound receptors have clearly defined the ‘primary’ site 1 interaction with the L1 domain from one monomer and the αCT′ domain from the opposite monomer (Figure 5.1E and 51F) (Blyth et al., 2020). There is also a low affinity interaction with receptor FnIII-1′ domain from the opposite monomer involving one of the previously defined site 2 residues (E12 IGF-II or HB10 insulin) (Blyth et al., 2020). The remaining ligand site 2 residues do not contact the receptor in the ‘J-shaped’ fully activated structure and have been proposed to be involved in a transient interaction prior to the final active signaling conformation (Xu et al., 2020, Weis et al., 2018, Lawrence, 2021, Gutmann et al., 2020, Scapin et al., 2018, Uchikawa et al., 2019, Li et al., 2019).

Insulin residue YB16 is involved in site 1 contact with the IR (Figure 5.1E and 5.1F). Mutation to either alanine or histidine results in a 2-3-fold decrease in IR binding affinity (Chen et al., 2000, Ludvigsen et al., 1994). A change to glutamine (as found in IGF-II) leads to a significant (10-fold) decrease in binding affinity for the IR-A (Schwartz et al., 1985). This highlights a point of difference between the interaction of insulin and IGF-II with the IR. The YB16 of insulin contacts residue F39 of the L1 domain of the IR-A through a high affinity π-π interaction, as revealed by recent cryo-electron microscopy (cryoEM) structures (Figure 5.1E and 5.1F) (Weis et al., 2018). The change in insulin to

a non-aromatic glutamine is unable to substitute for the tyrosine to generate such a high affinity interaction. Interestingly the glutamine substitution also results in a significantly reduced metabolic potency, suggesting this interaction influences IR activation (Schwartz et al., 1985), although mutation to alanine has only minor impact on IR metabolic activity (Chen et al., 2000).

Residue YB16 plays such an important role in the biophysical and functional properties of human insulin and is highly conserved among most vertebrate insulins. Due to the lack of conservation of a tyrosine at this residue in many cone snail venom insulins including Con-Ins G1 and Con-Ins T1A we were interested in understanding the effect of mutating the equivalent residue in IGF-II. Unexpectedly, mutation of Q18 to tyrosine had a major effect on analogue production due to aggregation problems and also the mutation affected IGF-IIs thermal stability. We uncover a unique role of IGF-II residue 18 as a determinant of receptor binding specificity, but surprisingly changes at Q18 do not influence signaling outcomes.

5.2 Results

5.2.1 Production and characterization of IGF-II analogues

To investigate the roles of IGF-II residue Q18 in the biophysical properties, receptor binding and receptor signaling three IGF-II analogues were produced; Q18I IGF-II, Q18M IGF-II and Q18Y IGF-II. Each IGF-II analogue was expressed in *E. coli* as a fusion protein, including 11 N-terminal amino acids of porcine growth hormone and a cleavage recognition site (Supplementary Figure 14), refolded (Section 2.3.9) and the mature peptide purified (Section 2.3.11) at sufficient quantities for biophysical and functional assays. Mass spectrometry (Section 2.3.13) confirmed correct masses for each analogue (Supplementary Table 3). Expression efficiencies of fusion peptide analogues in *E. coli* were similar (Supplementary Figure 14) and equivalent to native IGF-II. Figures 5.2- 5.4 show the analytical HPLC profiles during purification of each analogue including, dissolution, gel filtration, refold, cleavage and final purified peptide. Figure 5.5 summarises the protein yield at several major steps of the protein purification. Notably, the concentration of all IGF-II fusion peptide analogues at the dissolution stage was markedly lower than for the IGF-II fusion peptide (Figure 5.5). Thereafter the yields post gel filtration were 38.1% (IGF-II) and above for all peptides. At the completion of refold, yields were similar between IGF-II (20.4%) Q18I (20.0%) and Q18M IGF-II (28.9%). However, Q18Y IGF-II fusion peptide yield after refold compared to at dissolution was

considerably lower (3.7%) than for IGF-II (20.4%) (Figure 5.5). This loss of Q18Y IGF-II fusion protein is likely due to precipitation, which was visually evident during refold. This observation suggests that the Q18Y mutation is poorly tolerated, and only a small proportion of protein is capable of folding into the energetically favourable native-like conformation.

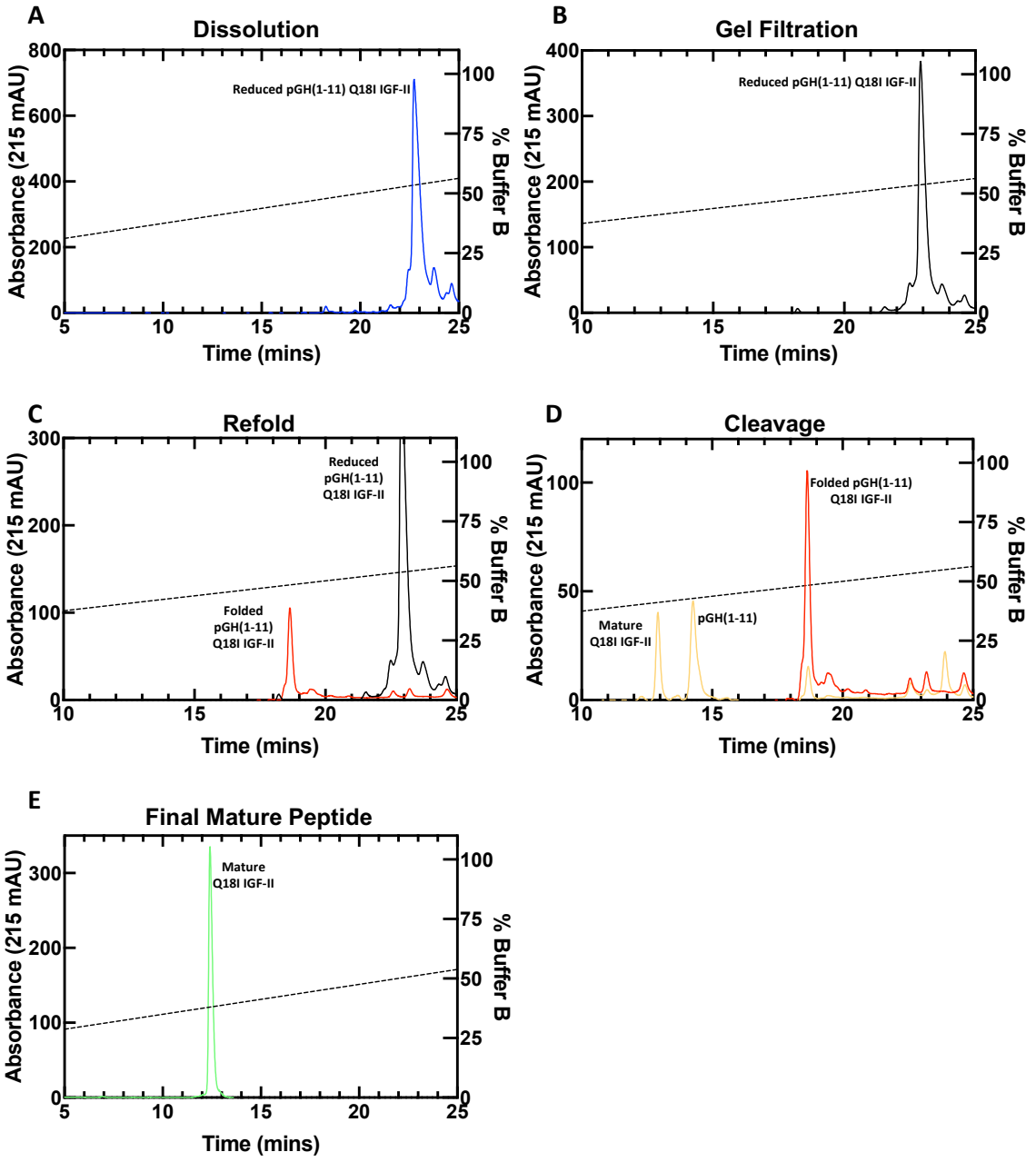


Figure 5.2: Purification of Q18I IGF-II.

Analytical reverse phase HPLC (using a C4 column and a 25 – 45 % CH_3CN (31.3 – 56.3% buffer B) over 20 minutes (Section 2.3.5)) was used to monitor each stage of protein purification. Shown are the absorbance chromatograms at 215 nm of reduced pGH(1-11) Q18I IGF-II (left Y axis). Dotted black line represents the % of buffer B (right Y axis). (A) Dissolution, (B) post gel filtration, (C) the refold T=0 minutes (solid black line) and T=30 minutes (red), (D) α-LP cleavage T=0 minutes (red) and T=10 hours (orange) with mature Q18I IGF-II at 13.5 minutes and (E) final mature Q18I IGF-II peptide post preparative HPLC purification (green).

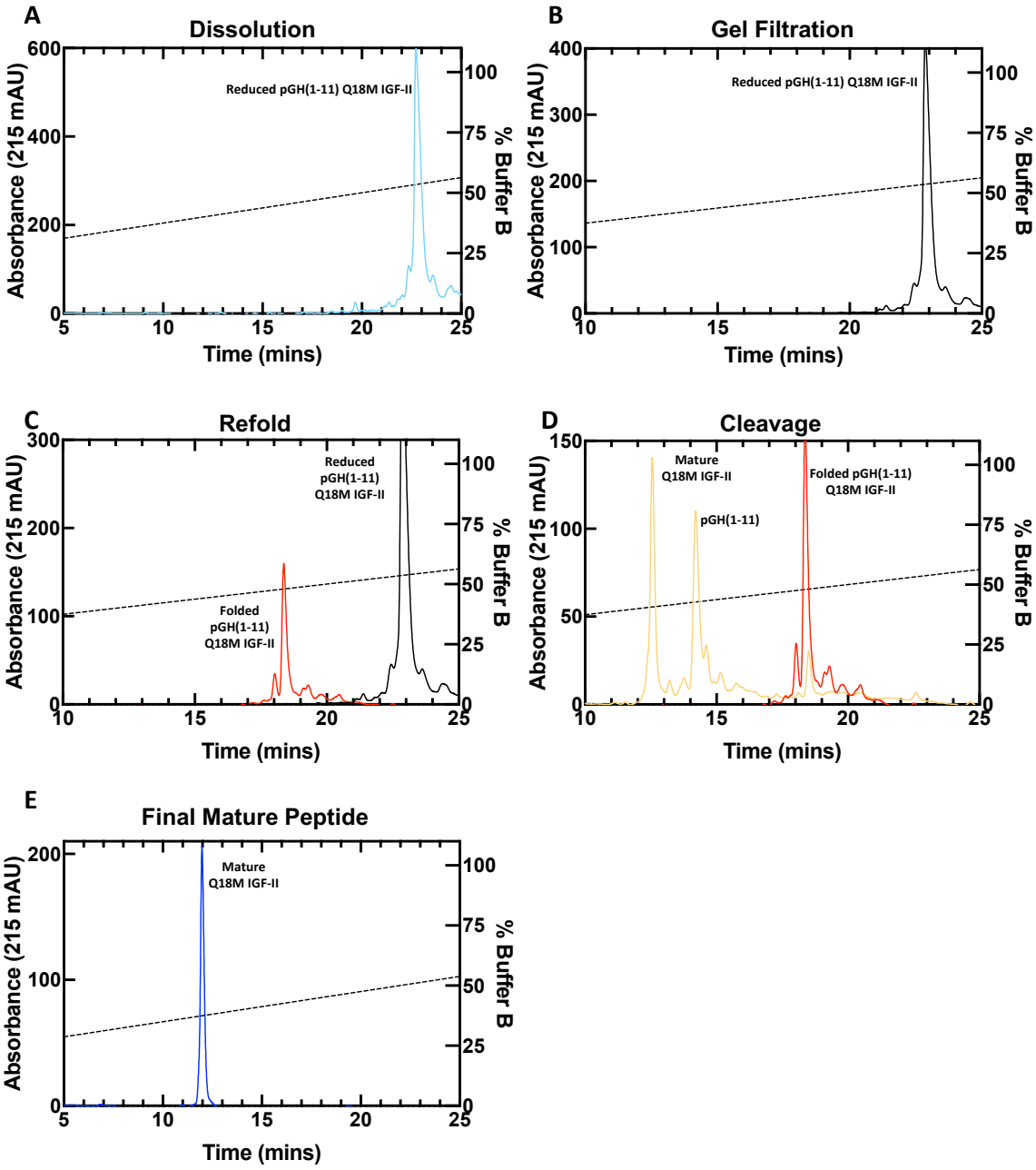


Figure 5.3: Purification of Q18M IGF-II.

Analytical reverse phase HPLC (using a C4 column and a 25 – 45 % CH_3CN (31.3 – 56.3% buffer B) over 20 minutes (Section 2.3.5)) was used to monitor each stage of protein purification. Shown are the absorbance chromatograms at 215 nm of reduced pGH(1-11) Q18M IGF-II (left Y axis). Dotted black line represents the % of buffer B (right Y axis). (A) Dissolution, (B) post gel filtration, (C) the refold T=0 minutes (solid black line) and T=30 minutes (red), (D) α-LP cleavage T=0 minutes (red) and T=10 hours (orange) with mature Q18M IGF-II at 12 minutes and (E) final mature Q18M IGF-II peptide post preparative HPLC purification (blue).

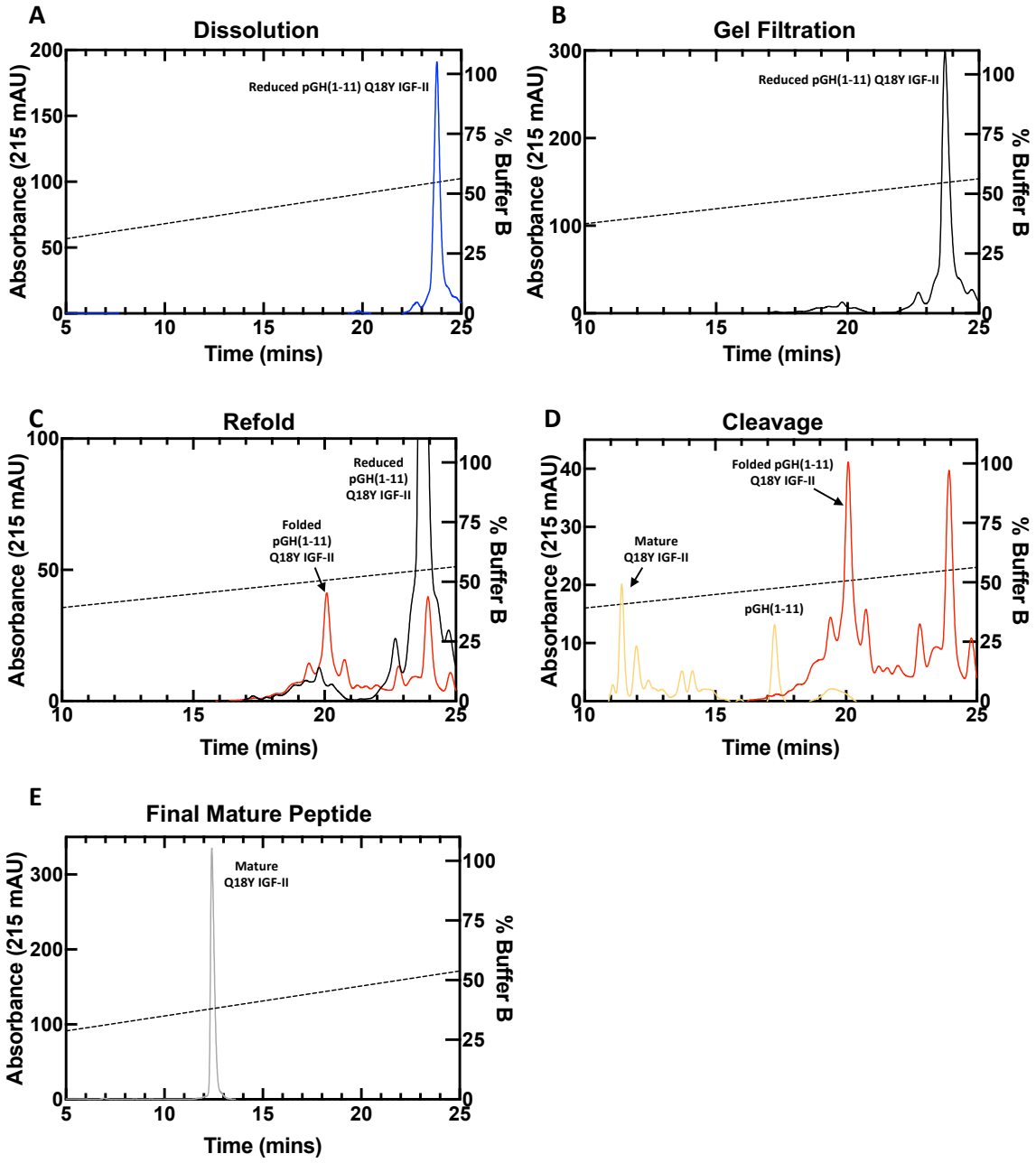
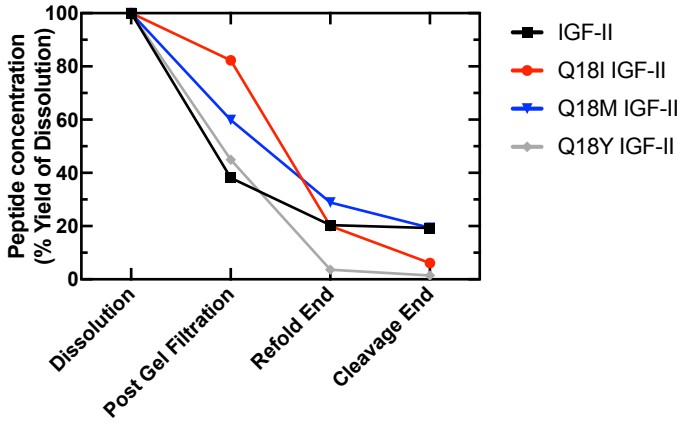


Figure 5.4: Purification of Q18Y IGF-II.

Analytical reverse phase HPLC (using a C4 column and a 25 – 45 % CH_3CN (31.3 – 56.3% buffer B) over 20 minutes (Section 2.3.5)) was used to monitor each stage of protein purification. Shown are the absorbance chromatograms at 215 nm of reduced pGH(1-11) Q18Y IGF-II (left Y axis). Dotted black line represents the % of buffer B (right Y axis). (A) Dissolution, (B) post gel filtration, (C) the refold T=0 minutes (solid black line) and T=30 minutes (red), (D) α-LP cleavage T=0 minutes (red) and T=10 hours (orange) with mature Q18Y IGF-II at 11 minutes and (E) final mature Q18Y IGF-II peptide post preparative HPLC purification (green).



Analogue	Peptide conc. (mg/L of starting ferment) at dissolution	Peptide concentration (% of Dissolution)		
		Gel Filtration	Refold End	Cleavage End
IGF-II	120	38.1	20.4	2.20
Q18I IGF-II	44.5	82.3	20.0	2.51
Q18M IGF-II	69.0	59.9	28.9	9.46
Q18Y IGF-II	65.6	45.0	3.7	1.16

Figure 5.5: IGF-II analogue production yields.

Concentration at gel filtration and refold steps in the purification of Q18 IGF-II analogue fusion proteins. Yield is expressed as a percentage of the concentration after dissolution (100%). Table of protein concentration at dissolution (mg/L) and then yield at subsequent steps expressed as a percentage of each of the analogue concentrations at dissolution.

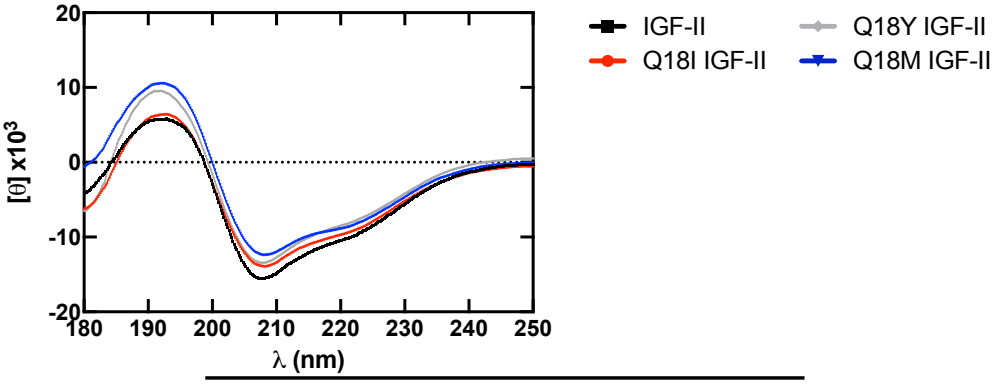
5.2.1.1 Circular dichroism studies

Circular dichroism (CD) spectra were recorded between 180 nm and 250 nm (Figure 5.6A) (Section 2.4.4.1). Q18I, Q18M and Q18Y IGF-II spectra all exhibited attenuated ellipticity at 222 nm, compared to IGF-II, indicating partial loss of α-helical content 30. The α-helical content calculated based on the CD spectra obtained for IGF-II was 38%, consistent with previous reports (Torres et al., 1995). The -helical content of Q18I IGF-II was calculated to be 35%, whereas Q18Y and Q18M IGF-II were found to have equal α-helical content of 32%.

5.2.1.2 Temperature denaturation

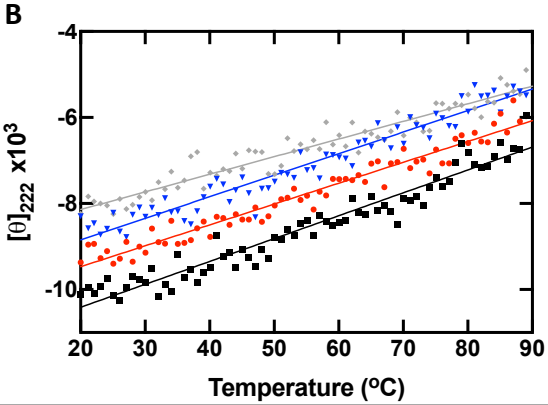
Temperature denaturation assays (Section 2.4.4.1.1) were performed from 20°C -90°C, for each analogue (Figure 5.6B) and α-helical content was monitored by CD at the helix sensitive wavelength 222 nm. Consistent with the 222 nm values obtained for the CD spectra (Figure 5.6A) IGF-II had the lowest ellipticity, followed by Q18I, Q18M and Q18Y IGF-II. As shown in Figure 5.6B, the rate of change (slope) in ellipticity with increasing temperature is similar between Q18I IGF-II, Q18M IGF-II and IGF-II (also expressed as % change from 20°C to 90°C in the accompanying table), whilst the rate of change in ellipticity for Q18Y IGF-II was the lowest at this temperature indicating this analogue is the least thermodynamically stable.

A



Analogue	Helix	Strand	Turns	Unordered
IGF-II	38	13	15	34
Q18I IGF-II	35	13	16	37
Q18M IGF-II	32	17	15	36
Q18Y IGF-II	32	17	15	36

B



Analogue	R^2	Slope	$[\theta]_{222}x10^3$		% change from 20°C to 90°C
			20°C	90°C	
IGF-II	0.9394	0.05	-10.5	-6.8	35.2
Q18I IGF-II	0.9544	0.05	-9.5	-6.1	35.9
Q18M IGF-II	0.9404	0.05	-8.8	-5.3	39.7
Q18Y IGF-II	0.9262	0.04	-7.9	-4.3	45.2

Figure 5.6: Analysis of Q18 IGF-II analogues and IGF-II secondary structure and thermal stability using circular dichroism.

(A) Secondary structure analysis of Q18 IGF-II analogues and IGF-II was performed using circular dichroism at 0.05 μg/μL in 10 mM acetic acid, mean residue ellipticity [θ] was measured from 260 - 180 nm (Section 2.4.4.1). Percent helical content shown in the table was performed using Dichroweb and was calculated using the CDSSTR algorithm for deconvolution against the reference protein database set SMP180 (Sreerama and Woody 2000). (B) Thermal stability was monitored by measuring mean residue ellipticity [θ] at 222 nm from 20°C to 90°C (at 1±0.1°C increments) (Section 2.4.4.1.1). Temperature denaturation data were analysed using a linear regression line of best fit. R^2 values, slope and percentage change in helical content from 20°C to 90°C are shown in the table.

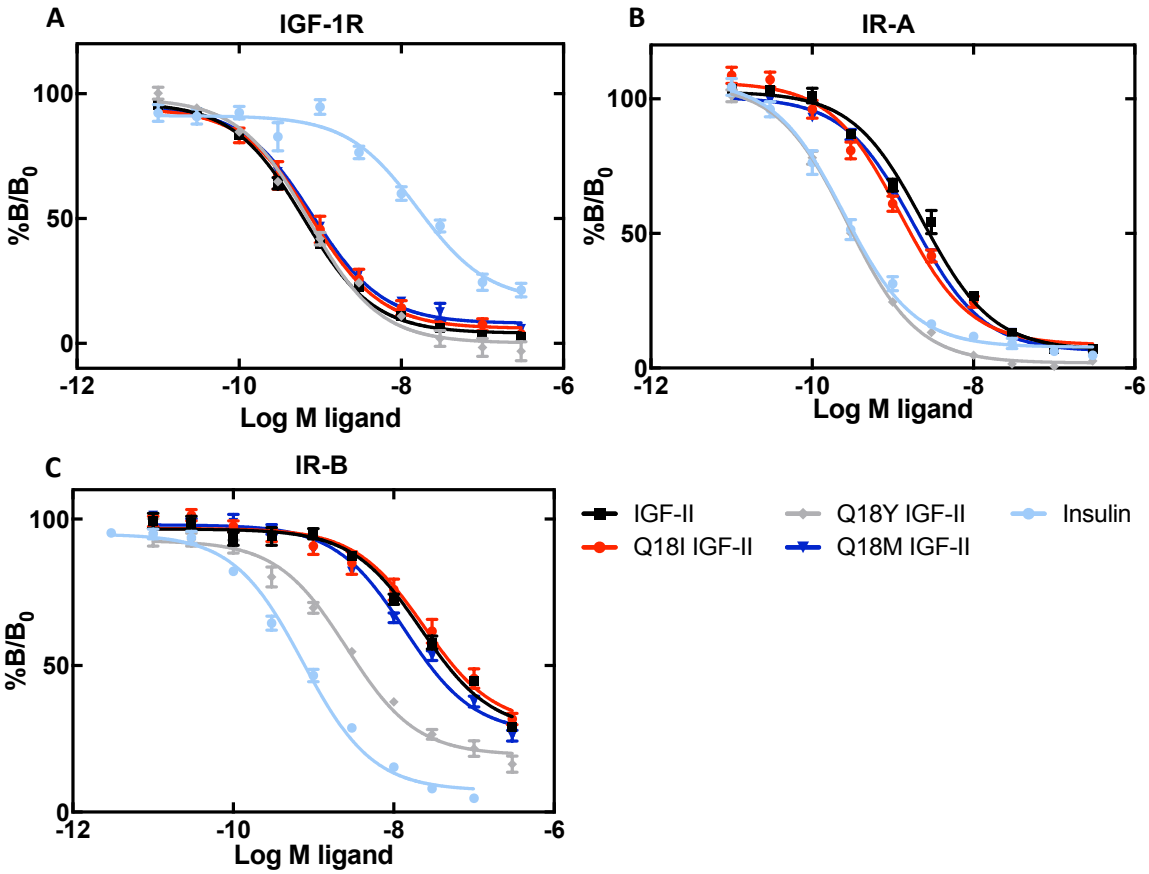
5.2.1.3 Immunocaptured receptor binding assays

IGF-1R binding assays were performed (Section 2.4.1) using europium (Eu) labelled IGF-II and increasing concentrations of competing ligands (Figure 5.7A). As expected, insulin had a low affinity for the IGF-1R with an IC_{50} value of 15.5 nM, which is similar to previous reports using Eu labelled IGF-I (Denley et al., 2004). All three Q18 analogues bound the IGF-1R with similar affinities to IGF-II (IC50 0.53nM). Q18I IGF-II had the highest binding affinity for the IGF-1R, with an IC_{50} value of 0.45 nM, and Q18M and Q18Y IGF-II had similar IC_{50} values of 0.84 and 0.66 nM, respectively.

IR-A binding assays were also performed using Eu labelled IGF-II (Figure 5.7B). As expected, insulin bound with a 9.40-fold higher affinity to the IR-A than IGF-II, consistent with current literature 31,32. Interestingly, there was little effect on binding with the Q18I and Q18M substitutions (Figure 5.7B), whereas substitution with tyrosine resulted in an 8.48-fold increase in affinity compared to IGF-II, to an IC_{50} similar to insulin. A similar observation was made in IR-B binding assays performed with Eu labelled insulin and increasing concentrations of competing ligands (Figure 5.7C). Insulin bound with a 30.0-fold higher affinity to the IR-B than IGF-II, consistent with current literature (Denley et al., 2004). Again, there was minimal effect of the isoleucine and methionine substitutions on IR-B binding, whereas Q18Y IGF-II had an 8.54-fold increased binding affinity compared to IGF-II. However, Q18Y IGF-II still did not bind IR-B as well as insulin (a 3.51-fold difference in affinities). Overall, these binding assays highlight the importance of the tyrosine residue in high affinity binding to the insulin receptor.

5.2.1.4 Activation of Akt and ERK signaling

Potency of activation of metabolic and mitogenic signaling was also measured through immunoblotting phosphorylated Akt (T308) and ERK respectively (Figure 5.8) (Section 2.4.2). L6 rat skeletal myoblasts overexpressing the human IR-A (L6 IR-A) were treated with 10 nM of either Q18 IGF-II analogues or IGF-II and a time course was performed (0,0.5, 1, 3, 5, 10 and 20 min). Each IGF-II analogue was equipotent to IGF-II at stimulating Akt phosphorylation (Figure 5.8A, 5.8B and 5.8D), except for Q18I IGF-II at t = 5 when there was a small but significant difference to IGF-II ($p = 0.01$ to 0.05). Each IGF-II analogue was similarly found to be equipotent to IGF-II in stimulating ERK phosphorylation (Figure 5.8A, 5.8B and 5.8C), except for T=5 min Q18M IGF-II ($p = 0.01$ to 0.05) and Q18Y IGF-II ($p \leq 0.01$) when there was a small but significant difference to IGF-II. It was surprising that Q18Y IGF-II was essentially equipotent to IGF-II despite its 8.48-fold greater IR-A binding affinity.



Analogue	IGF-1R			IR-A			IR-B		
	IC_{50} (nM)	95% CI (nM)	Relative binding affinity (% of IGF-II)	IC_{50} (nM)	95% CI (nM)	Relative binding affinity (% of IGF-II)	IC_{50} (nM)	95% CI (nM)	Relative binding affinity (% of IGF-II)
IGF-II	0.622	0.568-0.689	100	2.34	1.94-2.82	100	22.2	16.6-22.9	100
Insulin	15.9	10.7-23.7	3.91	0.249	0.199-0.314	940	0.741	0.628-0.874	2996
Q18I IGF-II	0.781	0.588-1.03	79.6	1.23	0.980-1.54	190	24.1	15.5-37.4	92.1
Q18M IGF-II	0.800	0.685-0.935	77.8	1.82	1.63-2.03	129	13.6	10.5-17.5	163.2
Q18Y IGF-II	0.730	0.583-0.914	85.2	0.276	0.247-0.309	848	2.60	1.92-3.51	854

Figure 5.7: Competition of Q18 IGF-II analogues and europium-labelled IGF-II for the IGF-1R, IR-A and IR-B.

Competition binding of insulin, IGF-II and IGF-II analogues with europium-labelled IGF-II for the IGF-1R (A) and IR-A (B), and europium-labelled insulin for the IR-B (C) (Section 2.4.1). Immunocaptured IGF-IR, IR-A or IR-B were incubated with europium-labelled IGF-II or insulin and increasing concentrations of the competitive ligands insulin, IGF-II, Q18I IGF-II, Q18M IGF-II or Q18Y IGF-II. Time-resolved fluorescence was measured as described in Section 2.4.1. Results are expressed as a percentage of Eu-IGF-II bound in the absence of competing ligand (B/B_0). Data shown are the mean ±S.E.M. n = ≥ 3 independent experiments each with triplicate technical replicates (unless otherwise indicated on the graph). Error bars are shown when greater than the size of the symbols. Calculated IC_{50} values, 95% confidence interval (95% CI) relative binding affinity relative to human IGF-II are shown in the table.

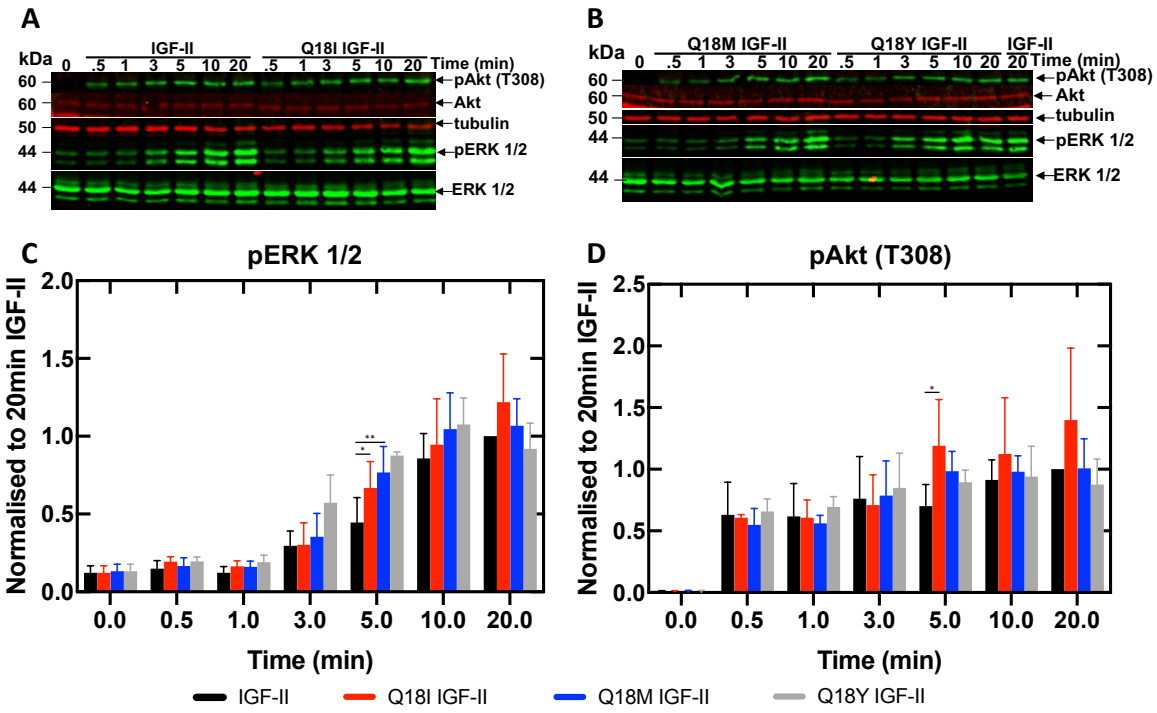
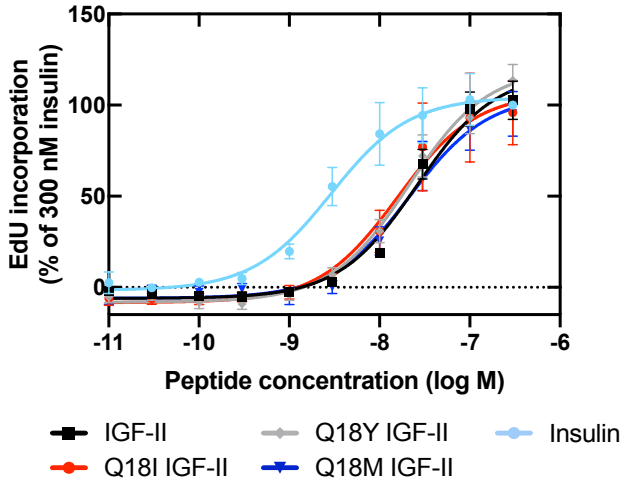


Figure 5.8: Induction of Akt and ERK phosphorylation upon IR-A activation by IGF-II and Q18 IGF-II analogues.

L6 IR-A cells were treated with IGF-II, Q18I IGF-II, Q18M IGF-II, or Q18Y IGF-II at 10 nM for 0, 0.5, 1, 3, 5, 10, 20 minutes (Section 2.4.2). Whole cell lysates were prepared and subjected to SDS-PAGE followed by immunoblotting for phosphorylated Akt (pAkt T308), total Akt T308, phosphorylated ERK (pERK 1/2) and total ERK 1/2 (A and B). Each blot included lanes from cells untreated (basal = 0 minutes) and treated with 10 nM IGF-II for 20 minutes. Densitometric quantitation of n = ≥ 3 independent experiments are shown as a bar graph of the mean ± S.E. Relative pAkt T308 (C) and pERK 1/2 (D) are expressed as a fraction of the level detected when cells were stimulated with 10 nM IGF-II for 20 minutes. There is no change in total Akt T308 and ERK 1/2 at these time points. In each case, pAkt T308 and pERK 1/2 was first normalized against the loading control (β - tubulin). Refer to Supplementary Figures 15 and 16 for uncropped Western Blot images. A two-way ANOVA with Dunnett's multiple comparison was performed. The only significantly different pAkt responses in (C) were IGF-II versus Q18I IGF-II at t = 5 min (*p value 0.01 to 0.05). The only significantly different pERK 1/2 responses in (D) were at t = 5 min for IGF-II versus Q18M IGF-II (*p value 0.01 to 0.05 and IGF-II versus Q18Y IGF-II **p ≤ 0.01).

5.2.1.5 DNA synthesis assays

Mitogenic potency of each analogue was measured (Section 2.4.3) using a DNA synthesis assay performed using L6 IR-A cells. Insulin potently activated DNA synthesis, 9-fold more potently than IGF-II (Figure 5.9). All IGF-II analogues were equipotent to IGF-II in activating DNA synthesis. This result is consistent with the equal ability of the analogues to activate ERK 1/2 (and suggests the small differences seen at 5 minutes (Figure 8B and D) are not biologically significant), and also demonstrates the unexpected low potency of Q18Y IGF-II despite its IR-A binding affinity being equal to insulin.



Analogue	EC_{50} (nM)	Relative DNA synthesis (% of IGF-II)	Fold change
IGF-II	26.1	100	-
Insulin	2.90	900	9
Q18I IGF-II	15.7	166	1.67
Q18M IGF-II	21.5	121	1.21
Q18Y IGF-II	21.2	123	1.23

Figure 5.9: DNA synthesis in response to increasing concentrations of insulin, IGF-II and IGF-II analogues.

Serum starved L6 rat skeletal myoblasts overexpressing the IR-A were treated with 0.01 - 300 nM insulin, IGF-II, Q18I IGF-II, Q18M IGF-II, or Q18Y IGF-II for 18 hours. Cells were incubated for 4 hours with 10 µM 5-Ethynyl-2'-deoxyuridine (EdU). 2 µM FAM-Azide 488, 100 mM Tris, pH 7.5, 4 mM CuSO4, 100 mM sodium ascorbate was added to the cells for 30 min. Fluorescence was measured using 485 nm excitation and 535 nm emission filters using a Multilabel Reader (Section 2.4.3). Data is shown as a percentage incorporation of 300 nM Insulin. Data shown are the mean ± S.E. n= ≥ 3 independent experiments. A two-way ANOVA with Dunnett's multiple comparison was performed. Insulin versus IGF-II (****, $p < 0.0001$).

5.3 Discussion

There is extensive understanding of the biophysical and functional properties of insulin and the roles played by individual residues in receptor binding and activation, whereas there is relatively little known about the determinants of IGF-II aggregation, stability and function through the IGF-1R and IR-A. Here we sought to understand the effect of mutating residue Q18 in insulin to that of methionine, isoleucine and tyrosine found in the equivalent position in Con-Ins G1 and Con-Ins T1A and human insulin receptively. We sought to ascertain whether the residue Q18 plays similar roles as the equivalent YB16 residue of human insulin, which is involved in maintaining structural stability and insulin dimerization as well as being key in receptor binding. Additionally, mutation of this residue to methionine and isoleucine as found in Con-Ins G1 and Con-Ins T1A was done to determine what effect these residues have on receptor binding and activation. Three analogues of IGF-II, Q18I IGF-II, Q18M IGF-II and Q18Y IGF-II were expressed as fusion peptides in E. coli in inclusion bodies. Interestingly the fusion peptides of all analogues were more difficult to solubilize than IGF-II fusion peptide and more peptide remained as insoluble aggregate at this step. Also, the refold step for the Q18Y IGF-II fusion peptide was significantly less efficient than for all other peptides due to marked precipitation observed under the refold conditions. Seemingly, the presence of a hydrophobic residue, and in particular the aromatic ring structure of tyrosine, at this position in the B-chain promotes self-association. This is consistent with the role of the equivalent residue B16 of insulin in dimerization, as demonstrated by substitution with alanine or histidine, rendering insulin monomeric (Chen et al., 2000). One would predict that introduction of glutamine into position B16 of insulin would reduce self-association, although this has not been reported.

A previously unrecognized role for Q18 in maintaining IGF-II B-domain helical stability was revealed in this study, with a lower helical content observed for all purified analogues. The lower helical content of Q18I IGF-II is consistent with the known destabilization of the helical conformation by the branched side chain of isoleucine (O'Neil and DeGrado, 1990). Methionine and tyrosine are evidently not as favourable as glutamine at this position in IGF-II, perhaps due to their larger bulk compared to glutamine. Q18Y IGF-II was also noticeably less thermodynamically stable. Whilst tyrosine is present in the equivalent position in insulin there are differences in the i, i+3 interactions, with a glutamate and asparagine in the i+3 position of insulin and IGF-II, respectively. These differences may account for the tolerance of tyrosine at this position

in insulin with respect to thermodynamic stability.

Molecular details of the specificity of IR and IGF-1R binding have been elucidated through the analysis of the three IGF-II Q18 analogues. Earlier mutagenesis of the IR residue F39 to serine (found at the equivalent IGF-1R residue S35) (Kjeldsen et al., 1994) or alanine (Williams et al., 1995) pointed towards a role of F39 in insulin binding. However, mutagenesis of the IGF-1R S35 to alanine did not disrupt IGF-I binding (Whittaker et al., 2001). It was not until recent structures of the insulin:IR 21(Weis et al., 2018) and IGF-II:IGF-1R (Xu et al., 2020) complexes were solved that it became clear this may be a site of difference in the binding mechanisms at each receptor and a determinant of ligand receptor binding specificity. Site-directed mutagenesis of insulin pointed towards this (Chen et al., 2000, Ludvigsen et al., 1994, Schwartz et al., 1985) but until now this has not conclusively been demonstrated using IGF-II.

Here we show that Q18 does not play a major role in IGF-1R binding as there was very little impact of the IGF-II Q18 substitutions on IGF-1R binding despite the biophysical changes observed. This is interesting given that structural analyses show Q18 is adjacent to side chains of L33, S35 and R59 IGF-1R (Xu et al., 2020), but supports the S35A IGF-1R mutagenesis data showing no effect on IGF-I binding (Whittaker et al., 2001). Consistent with this, the equivalent residue Q15 of IGF-I is in close proximity to the IGF-1R residues L33 and R59 in the IGF-I:IGF-1R structure (Xu et al., 2018) and IGF-I mutation of Q15 to serine (Shooter et al., 1996), alanine (Jansson et al., 1997) or glutamate (Jansson et al., 1997) had little effect on IGF-1R binding (1.25-, 2- and 1.4-fold decrease respectively). Thus, we have now conclusively demonstrated that these interactions are not contributing greatly to the overall IGF-1R site 1a binding affinity.

Structures of the insulin:IR-A reveal a π-π interaction between insulin YB16 and IR F39 (Figure 5.10B and 5.10D) (Weis et al., 2018). The importance of this π-π interaction for high affinity ligand binding to the IR was previously highlighted by mutation of insulin YB16 to glutamine or alanine which significantly reduced binding affinity (Gauguin et al., 2008b, Schwartz et al., 1985) whereas mutation to phenylalanine or tryptophan only led to a modest change in IR affinity, presumably as the π-π interaction was maintained 39 (Hu et al., 1993). Here, through introduction of a tyrosine at residue 18 of IGF-II we show that by emulating the π-π interaction between insulin YB16 and IR F39 and we greatly increased the affinity of IGF-II for both IR-A and IR-B (Figure 5.7B and C respectively). Indeed, the single residue change of Q18 to tyrosine was sufficient to confer to IGF-II

the same IR-A binding affinity as insulin, whereas substitution with isoleucine, methionine (Figure 5.7B) or alanine (Alvino et al., 2009) had little effect. Gauguin et al (2008b) reported that a sextuple IGF-II analogue (T7H, T16A, Q18Y, F48T, S50I, T58N IGF-II) had a similar affinity to insulin for the IR-A, and our current study would suggest that the Q18Y substitution is largely responsible for its increased binding affinity. The affinity of IGF-I for the IR was also increased by introduction of a tyrosine at the equivalent position in the Q15Y, F16L IGF-I analogue (Bayne et al., 1988), again highlighting the importance of a π-π interaction at this position. Overall, the evidence provided here conclusively demonstrates the key role of the interaction between insulin YB16 and IR F39 in high affinity binding and receptor binding specificity.

Following the initial binding event insulin and IGF-II activate the PI3K/Akt and ERK/MAP kinase signaling pathways with potencies in line with their IR-A binding affinities (Rajapaksha and Forbes, 2015). To our surprise, despite Q18Y IGF-II binding IR-A with equal affinity to insulin and an 8.48-fold higher affinity than IGF-II it was only able to activate Akt and ERK1/2 signaling to the same extent as IGF-II. This was also reflected in the DNA synthesis assay where Q18Y IGF-II was equipotent with IGF-II, Q18I IGF-II and Q18M IGF-II, with a 7.31-fold lower potency compared to insulin (Figure 5.9). We conclude that the formation of a π-π interaction with IR F39 essentially made no impact on the signaling outcome of IGF-II.

We have sought to explain why Q18Y IGF-II does not activate the IR proportionately to its binding affinity through the knowledge provided by recent cryoEM structures of insulin bound to the IR, as there is currently no structure available of the IGF-II:IR complex. In what is believed to be the fully activated IR structure, insulin binds at site 1 to the IR L1 domain from one monomer and the αCT′ and FnIII-1′ domains from the opposite monomer (Weis et al., 2018). Of the contacts made between insulin residues and IR site 1, most involve side chains of residues in either the αCT′ alone, both the αCT′ and L1 or the FnIII-1′ (Figure 5.10A and 5.10B). It is uncommon for contact to be made only with the L1 domain, as is the case for the tyrosine π-π interaction with IR F39. Mutation of insulin residues primarily binding the αCT′ negatively affect αCT′ interaction and significantly reduce IR activation (Shoelson et al., 1983). This is seen, for example, upon mutation to alanine of A-chain insulin residues VA3 and YA19 that contact IR residues F710 and F714, and residue FB25 that contacts IR V715. Similarly, simultaneous binding of insulin to both IR monomers through αCT′ and L1 interaction is critical for high affinity binding and IR activation (Hua et al., 1993). Insulin FB24 within the critical FB24, FB25,

YB26 motif is regarded as a critical anchor residue that contacts αCT′ F714 and the L1 domain (Pandyarajan et al., 2014).

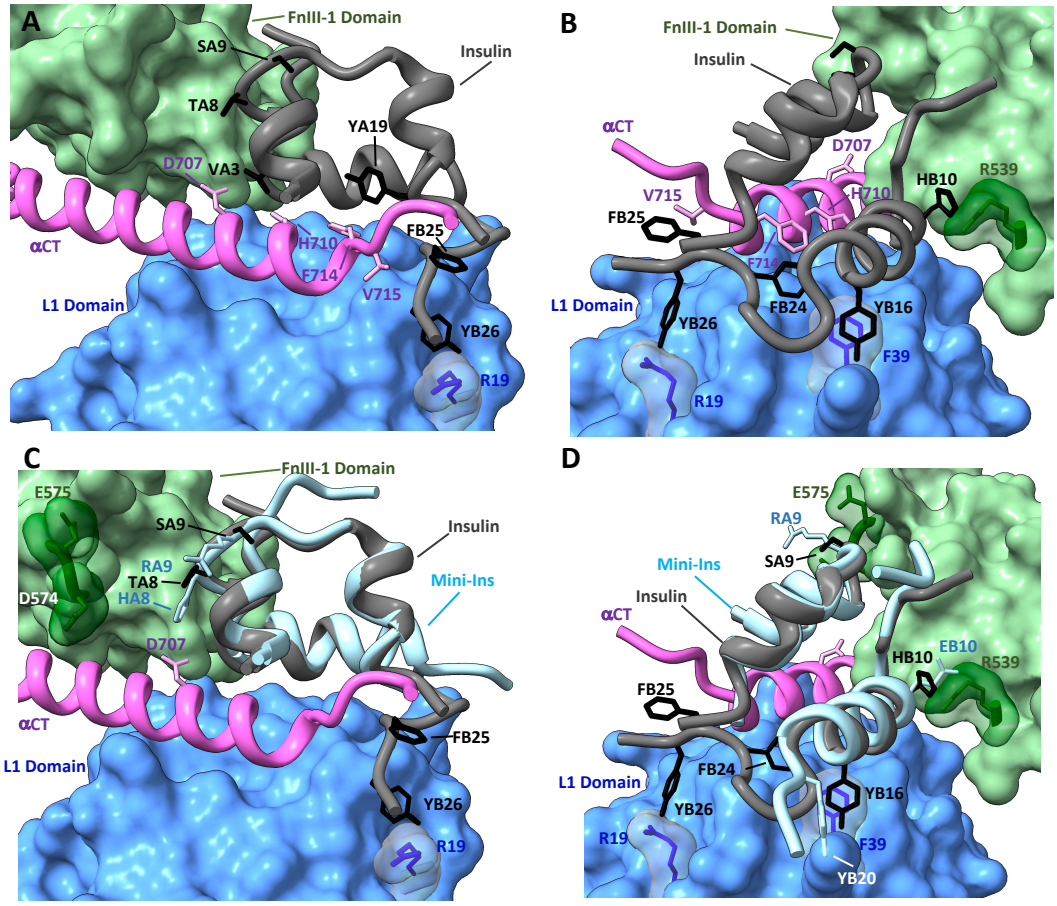


Figure 5.10: Comparison of Insulin and IGF-II binding to their cognate receptors at site 1.

(A and B) CryoEM structure of insulin bound to the IR (PDB: 6HN5) highlighting the site 1 interaction. Insulin binds to the IR primary binding site 1a (αCT′ (purple) and L1 domain (surface filled cornflower blue) of opposite monomers), with key contacts involving insulin (black) residues VA3, YA19 and FB25 with αCT′ resides H710, F714 and V715. Additionally, insulin residues FB24, FB25 and YB26 (FFY motif) contact residues R19 and YB16 of the L1 domain. An additional interaction is made between insulin residue HB10 and R539 of the FnIII-1′ domain (shown in B). This has been termed site 1b. (C and D) CryoEM structure of insulin bound to the IR with mini-ins overlayed (PDB: 6HN5 and 6VET respectively). In site 1a binding, lack of the FFY motif in mini-Ins is compensated through GB20Y mutation which likely interacts with residue F39 of the L1 domain. The site 1b interaction with FnIII-1′ domain R539 is strengthened through a HB10E mutation and additional contacts are made through residue E575 and mini-ins residues SA9R and TA8H.

Further interactions between insulin residues and the FnIII-1′ contribute to the overall binding affinity and can influence activation and downstream signaling. For example, HB10 approaches IR FnIII-1′ residue R539 in the fully activated structure, and mutation of HB10 to asparagine 45(Kurtzhals et al., 2000) or glutamate (Glendorf et al., 2008, Kaarsholm et al., 1993, Schwartz et al., 1987) results in 2-4-fold increased IR affinity leading to greater metabolic potency. Similarly, the recently developed Mini-Ins analogue (HA8, RA9, EB10, YB20 des-octapeptide) designed based on structural features of a cone snail insulin-like peptide (Xiong et al., 2020, Menting et al., 2016) has increased FnIII-1′ affinity that results in altered signaling. Not only does Mini-Ins residue EB10 form the salt bridge with R539 (Figure 5.10D) (Xiong et al., 2020) but residues HA8 and RA9 make additional contacts with FnIII-1′ domain residues D574 and E575, respectively, that are not observed in the insulin:IR complex (Figure 5.10C). The strengthened affinity for the FnIII-1′ domain results in equal IR binding affinity to insulin by compensating for the lack of the site 1a B24 anchor residue. Mini-Ins has a potency in Akt (T308) activation similar to insulin but a reduced ability to activate ERK 1/2.

Unlike FnIII-1′ interactions, evidently in the case of Q18Y IGF-II increasing the affinity through the L1 domain interaction alone does not influence IR signaling. However, by introducing the Q18Y change and concomitantly strengthening the interaction at the FnIII-1′, as is seen with the sextuple IGF-II analogue, signaling potency can be increased to equal that of insulin (Gauguin et al., 2008b). Other residues that contact the L1 alone, including SB9, EB21 and YB26 (YB26 shown in Figure 5.10), also don't contribute greatly to IR binding affinity or biological activity (Glendorf et al., 2008, Nakagawa et al., 2006, Zakova et al., 2008) supporting the proposal that an L1 only interaction has a lesser impact on activation than the interactions with the αCT′ and FnIII-1′ domains of the opposite monomer. In conclusion, our findings provide some insight into the mechanism of IR activation enabled through specific ligand:receptor interactions. However, a full understanding of the mechanisms driving signaling bias will require further investigation.

In summary, the overall aim of this PhD is to investigate IGF-II altered signaling properties by removing the classic binding site in human IGF-II and introduce these newly discovered modes of receptor contact. As described in chapter 3 three site 1a minimised IGF-II analogues were produced, lacking some or all of the elements of the FYF motif. These analogues provide the framework for investigating alternate modes of receptor engagement. In Chapter 4 and in this chapter we individually explored residues found in cone snail venom insulins that compensate for the loss of the FFY motif in insulin

and in this chapter we further explored a residue that is poorly conserved among many cone snail venom insulins. As described in Chapter 4 and in this chapter, single mutations based on cone snail venom insulin peptides were explored in native IGF-II to first gain an understanding of their effect on structure, stability, receptor binding and where appropriate, activation.

With the aim of recapitulating the alternate binding modes found in the cone snail insulins in IGF-II the following chapter (Chapter 6) will explore some of the mutations described in this in combination with site 1a 'minimised' IGF-II analogues produced in Chapter 3. This approach investigates whether these combined mutations could compensate for the lack of site classic 1a binding (through the FYF motif) in an analogous way to the compensating function they adopt in cone snail insulins. It is expected that these residues would convey high affinity binding to mutants lacking the ability to bind through the site 1a FYF motif.

6 SITE 1A MINIMISED IGF-II ANALOGUES AND CONE SNAIL VENOM INSULIN ALTERNATE BINDING MUTATIONS

6.1 Introduction

In light of novel insulins being identified with unique modes of insulin receptor (IR) engagement and unique signaling kinetics (including cone snail venom insulins) the Forbes laboratory has been interested in incorporating alternate binding modes into the insulin-like growth factor II (IGF-II). Testing alternative modes of receptor engagement in IGF-II may also provide us with unique insights into how signaling bias is achieved. By better understanding these mechanisms underlying signaling outcome we can use this knowledge to develop antagonists which can selectively block mitogenic action of IGF-II in cancer.

We are interested in two distinctly different modes of compensating for the loss of the FFY motif in cone snail venom insulins. The first, identified in Con-Ins G1, has been shown to engage with receptor site 1a through residues YB15 and YB20 (Figure 4.2B) (Menting et al., 2016). YB15 engages with the same hydrophobic pocket of the L1 domain as residue FB24 of the aromatic FYF motif in human insulin. Con-Ins G1 binds to site 1 b with a higher affinity than human insulin due to HA8, RA9 and a post translationally modified γ-carboxyl glutamate at position B10 (glutamic acid is the non-post translationally modified version (Safavi-Hemami et al., 2015)) (Figure 4.2C). As described in Chapter 4, these residues from Con-Ins G1 have been incorporated into a human insulin analogue also lacking the FFY motif, termed des octapeptide (B23–B30)-insulin (DOI). The resultant analogue now termed Mini-Ins, was found to have a weaker site 1a interaction but stronger site 1b interaction compared to human insulin (Xiong et al., 2020). Despite the differences in strength of interaction at site 1a and 1b, compared to human insulin, this analogue had an affinity similar to human insulin for the IR. The mitogenic potency of Mini-Ins appeared to be delayed for Mini-Ins compared to human insulin, however, resulted in equal maximal response (Xiong et al., 2020).

The second mode of compensating for the loss of the FFY motif in cone snail venom insulins comes from the Con-Ins K1 insulin, which compensates for the missing FFY motif through an A-chain extension; T22 L23, Q24 and γ-carboxyl glutamate 25 (Xiong et al., 2021). A variation of this A-chain extension (HA21 SA22, LA23 and QA24) was investigated in the context of human DOI and residue LA23 of this A-chain extension was also found to engage with the same pocket otherwise occupied by FB24 in human insulin (Xiong et al., 2021). LA23 was also found to be important for high affinity binding. Vh-Ins was shown to have a high affinity for the site 1a and weak affinity for site 1b (Xiong et al., 2021).

Briefly summarising the previous chapters, Chapter 3 described the production of several site 1a minimised IGF-II analogues which now provide a template into which can be incorporated residue changes that might compensate for the lack of the FYF motif. Of these analogues produced in Chapter 3, F26A IGF-II and F26G Y27G F28G IGF-II mutants were determined to be the most appropriate site 1a minimised analogues for testing of alternate binding mechanisms. In Chapter 4 and the unique binding contacts from Con-Ins G1 and Con-Ins K1, which compensate for the loss of the FFY motif in insulin, were individually tested in human IGF-II to gain a baseline understanding of their effect on structure, stability, binding and where necessary, signaling. Of the residues investigated from Con-Ins G1 residues L17Y and G22Y will be further explored in this chapter in the context of the site 1a minimised IGF-II analogues (Chapter 3). As discussed in Chapter 4 F48H IGF-II mutation has already been described in the literature as having an increased receptor binding affinity (Machackova et al., 2018) and therefore individual mutation of this residue was not required. However, this residue will be explored in combination with residues L17Y and G22Y in the context of the site 1a minimised IGF-II analogues in an attempt to compensate for the loss of the FYF motif in IGF-II. In addition, the A-chain extension explored in Chapter 4, based on Con-Ins K1 will also be explored in this Chapter in the context of site 1a minimised IGF-II analogues (Chapter 3).

As described in Chapter 3, F26A IGF-II and F26G Y27G F28G IGF-II were selected as the most appropriate site 1a minimised IGF-II analogues for testing alternate modes of receptor binding. F26A IGF-II has preservation of some elements of the aromatic motif (Y27 and F28) and is therefore likely to maintain some native like structure, whereas F26G Y27G F28G IGF-II, with no preservation of the aromatic FYF motif, may allow for the Con-Ins G1 residues to fully compensate for the loss of the FYF motif. The combination analogues investigated in this chapter were as follows:

1. Con-Ins G1 compensatory residues + F26A IGF-II site 1a minimised analogue:

- F26A G22Y F48H L17Y IGF-II
- F26A F48H L17Y IGF-II

2. Con-Ins G1 compensatory residues + F26G Y27G F28G IGF-II site 1a minimised analogue

- F26G Y27G F28G G22Y F48H IGF-II

3. Con-Ins K1 compensatory residues + F26G Y27G F28G IGF-II site 1a minimised analogue

- F26G Y27G F28G IGF-II site 1a minimised analogue
- F26G Y27G F28G ASLQ IGF-II

6.2 Results

6.2.1 Production of combination IGF-II analogues

6.2.1.1 Expression and Dissolution of combination IGF-II analogues

Each analogue was expressed as a fusion protein that included 11 N-terminal amino acids of the porcine growth hormone (pGH(1-11)) and a cleavage recognition site (Lien et al., 2001). The pGH(1-11) expression was induced by addition of IPTG (section 2.3.6). Pre- and post-induction samples were taken and separated on a 15% tricine gel under reducing conditions, followed by staining with Coomassie blue (Section 2.3.4). Successful expression of each analogue was confirmed by the presence of a 9.3 kDa band (expected size of pGH(1-11) IGF-II analogue) in the post induction sample but not in the pre induction sample (Supplementary Figures 17 - 19). Insoluble inclusion bodies were separated from soluble host cell material by centrifugation (Section 2.3.7). Inclusion bodies were subsequently solubilised (Section 2.3.8) (Figure 6.1-6.5A). The dissolution reactions were acidified to pH 2 prior to loading onto a size exclusion chromatography column (Figure 6.1 - 6.5B).

6.2.1.2 Refolding, cleavage and purification of IGF-II analogues

Refold of each pGH(1-11) IGF-II analogue was performed at a final concentration of 2.5 M urea and 0.5 mM DTT (Section 2.3.9) (Figure 6.1 - 6.5C). Refold was initiated by the addition of the oxidising agent 2-hydroxyethyldisulfide, which facilitated the formation of the three disulfide bonds and subsequently formation of the biologically active structure. As shown in Figures 6.1 - 6.5C, at the completion of refold there is a decrease in retention time of each analogue compared to T=0, characteristic of a transition from a reduced state to a more hydrophilic structure. Intermediate folding states of each IGF-II analogue were observed during the refold, as observed for IGF-II analogues described in Chapters 3 and 4. The pGH(1-11) leader sequence was cleaved from each analogue using α-LP (Section 2.3.10) (Figures 6.1-6.5D). Upon cleavage of the folded pGH(1-11) IGF-II analogues a peak representing cleaved pGH(1-11) can be observed (as indicated in Figure 6.1-6.5D). Not surprisingly, several peaks were formed, representing cleaved IGF-II analogues and their intermediate folding forms. Where necessary, as in Chapter 3, several peaks were isolated during the preparative HPLC stage (Section 2.3.11) and screened for secondary structure, activity and correct mass (peaks screened are indicated on each Figure 6.1-6.5D). Final isolated peaks were freeze dried (Section 2.3.12) for characterisation. Preparative reverse phase HPLC (Section 2.3.11) was used to separate the cleaved leader sequence from each mature IGF-II analogue followed by

freeze drying (Section 2.3.12) of the mature IGF-II analogue (Figure 6.1E – 6.5E).

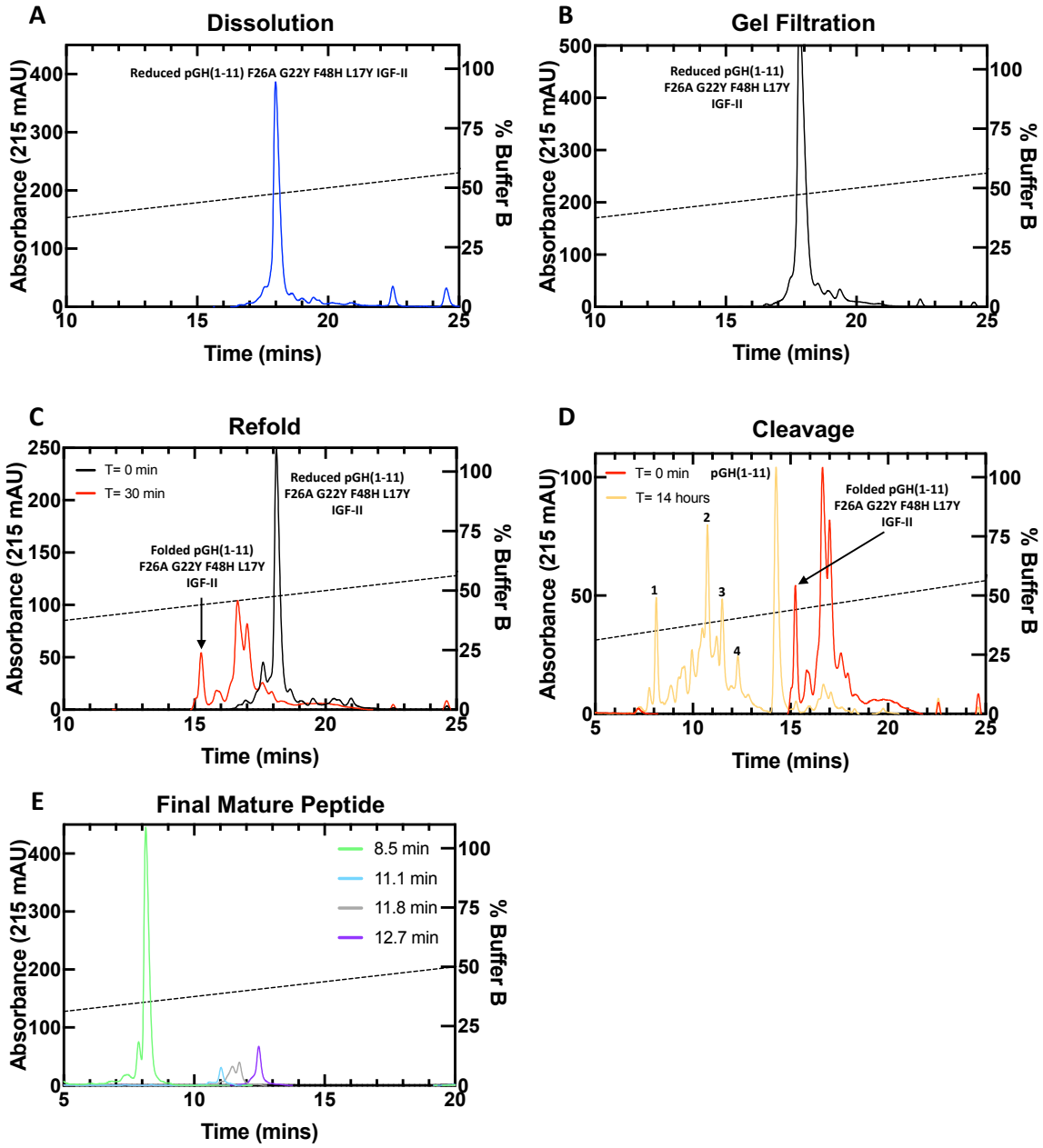


Figure 6.1: Purification of F26A G22Y F48H L17Y IGF-II.

Analytical reverse phase HPLC (using a C4 column and a 25 – 45 % CH_3CN (31.3 – 56.3% buffer B) over 20 minutes (Section 2.3.5)) was used to monitor each stage of protein purification. Shown are the absorbance chromatograms at 215 nm (left Y axis). Dotted black line represents the % of buffer B (right Y axis). Samples analysed were (A) dissolution, (B) post gel filtration, (C) the refold T=30 minutes (solid black line) and T=3 hours (red), (D) α-LP cleavage T=0 minutes (red) and T=18 hours (orange). (E) Three purified fractions of F26A G22Y F48H L17Y IGF-II (numbered 1-4 in D) with retention times of 8.5 minutes (green, peak 1 in D), 11.1 minutes (blue, peak 2 in D), 11.8 minutes (grey, peak 3 in D) and 12.7 minutes (purple, peak 4 in D).

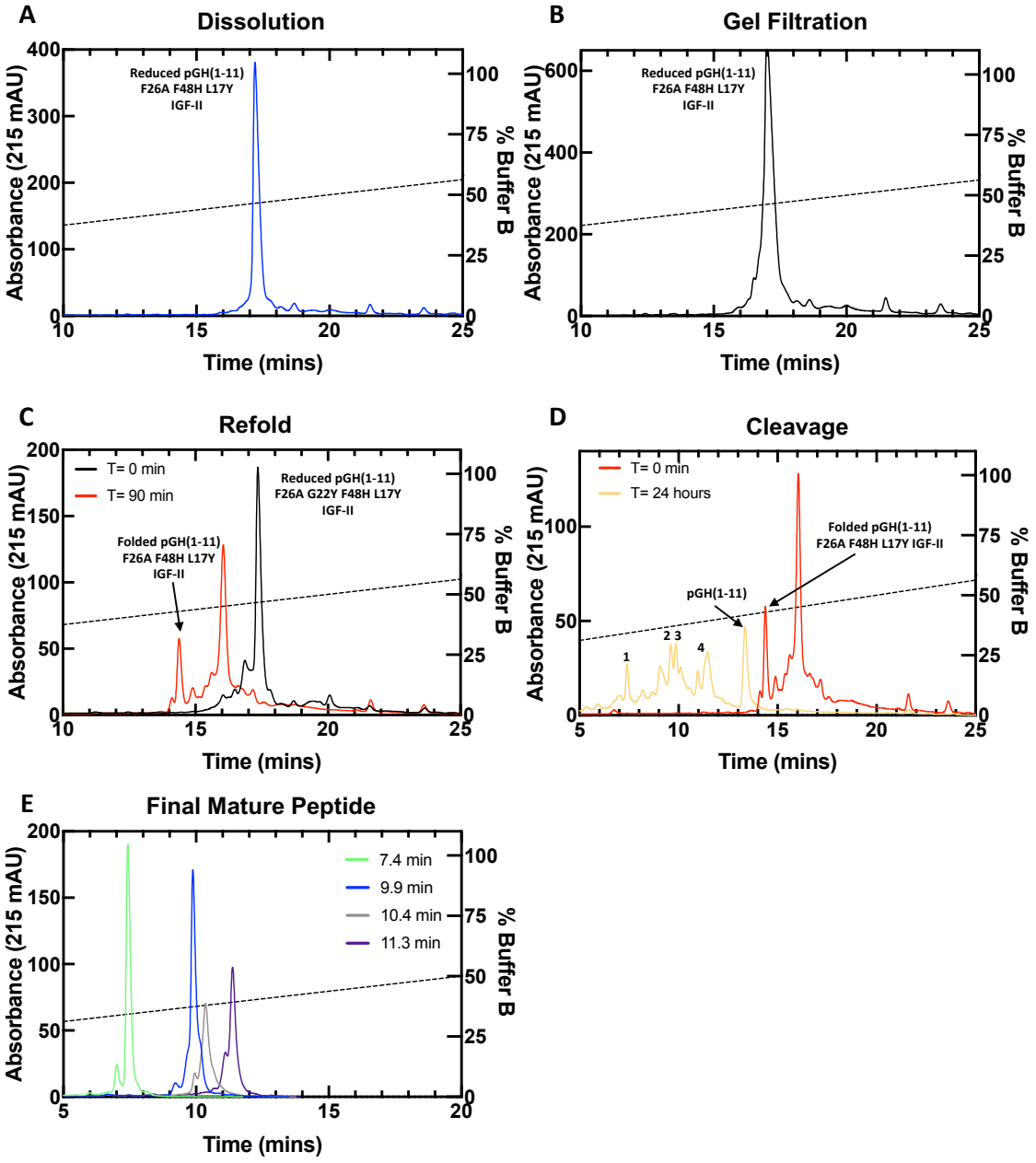


Figure 6.2: Purification of F26A F48H L17Y IGF-II.

Analytical reverse phase HPLC (using a C4 column and a 25 – 45 % CH_3CN (31.3 – 56.3% buffer B) over 20 minutes (Section 2.3.5)) was used to monitor each stage of protein purification. Shown are the absorbance chromatograms at 215nm (left Y axis). Dotted black line represents the % of buffer B (right Y axis). Samples analysed were (A) dissolution, (B) post gel filtration, (C) the refold T=90 minutes (solid black line) and T=3 hours (red), (D) α-LP cleavage T=0 minutes (red) and T=18 hours (orange). (E) Three purified fractions of F26A F48H L17Y IGF-II (numbered 1-4 in D) with retention times of 7.4 minutes (green, peak 1 in D), 9.9 minutes (blue, peak 2 in D), 10.4 minutes (grey, peak 3 in D) and 11.3 minutes (purple, peak 4 in D).

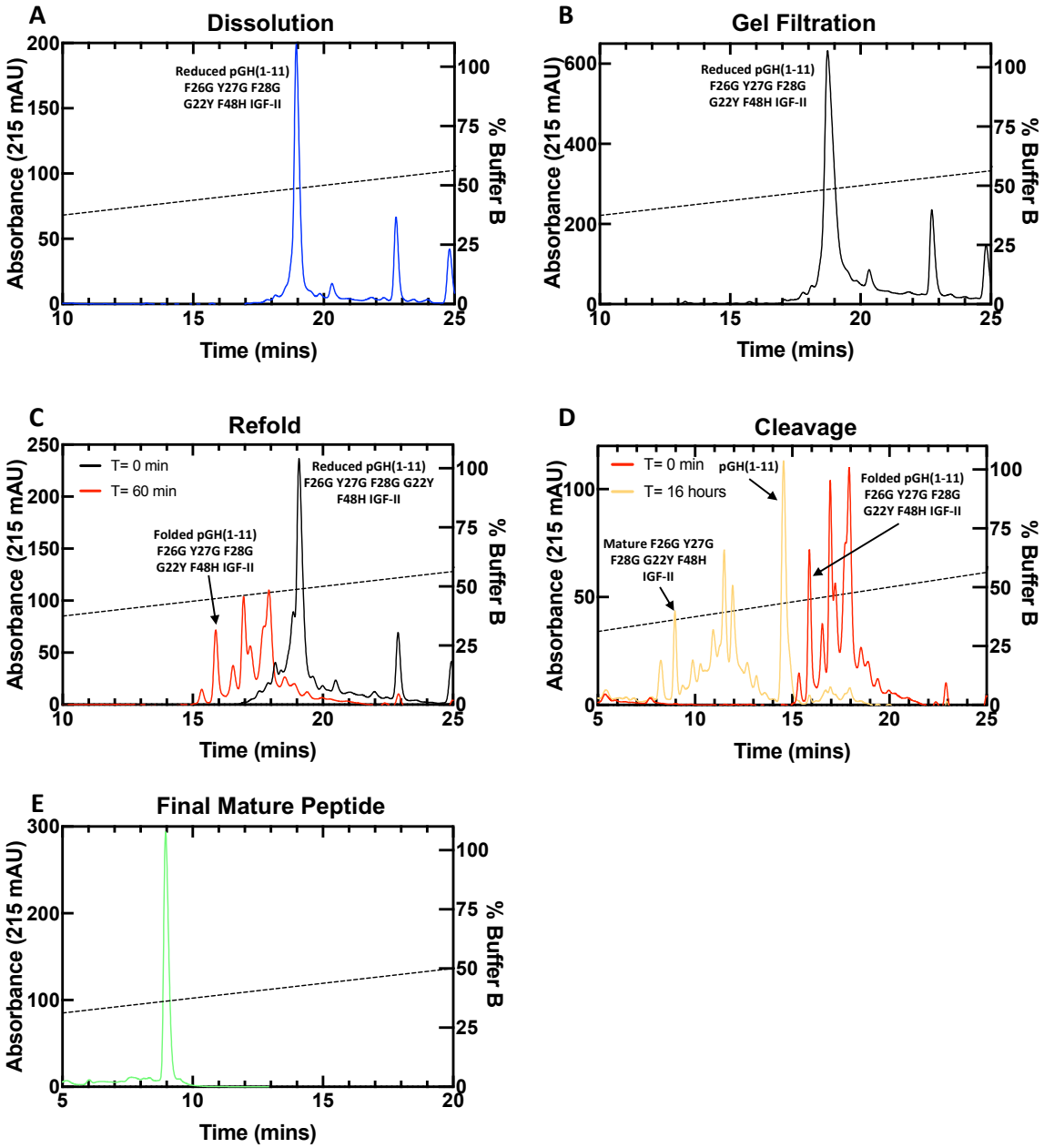


Figure 6.3: Purification of F26G Y27G F28G G22Y F48H IGF-II.

Analytical reverse phase HPLC (using a C4 column and a 25 – 45 % CH_3CN (31.3 – 56.3% buffer B) over 20 minutes (Section 2.3.5)) was used to monitor each stage of protein purification. Shown are the absorbance chromatograms at 215nm (left Y axis). Dotted black line represents the % of buffer B (right Y axis). Samples analysed were (A) dissolution, (B) post gel filtration, (C) the refold T=60 minutes (solid black line) and T=3 hours (red), (D) α-LP cleavage T=0 minutes (red) and T=16 hours (orange) with mature F26G Y27G F28G G22Y F48H IGF-II at 10.5 min and (E) final mature F26G Y27G F28G G22Y F48H IGF-II peptide post preparative HPLC purification (green).

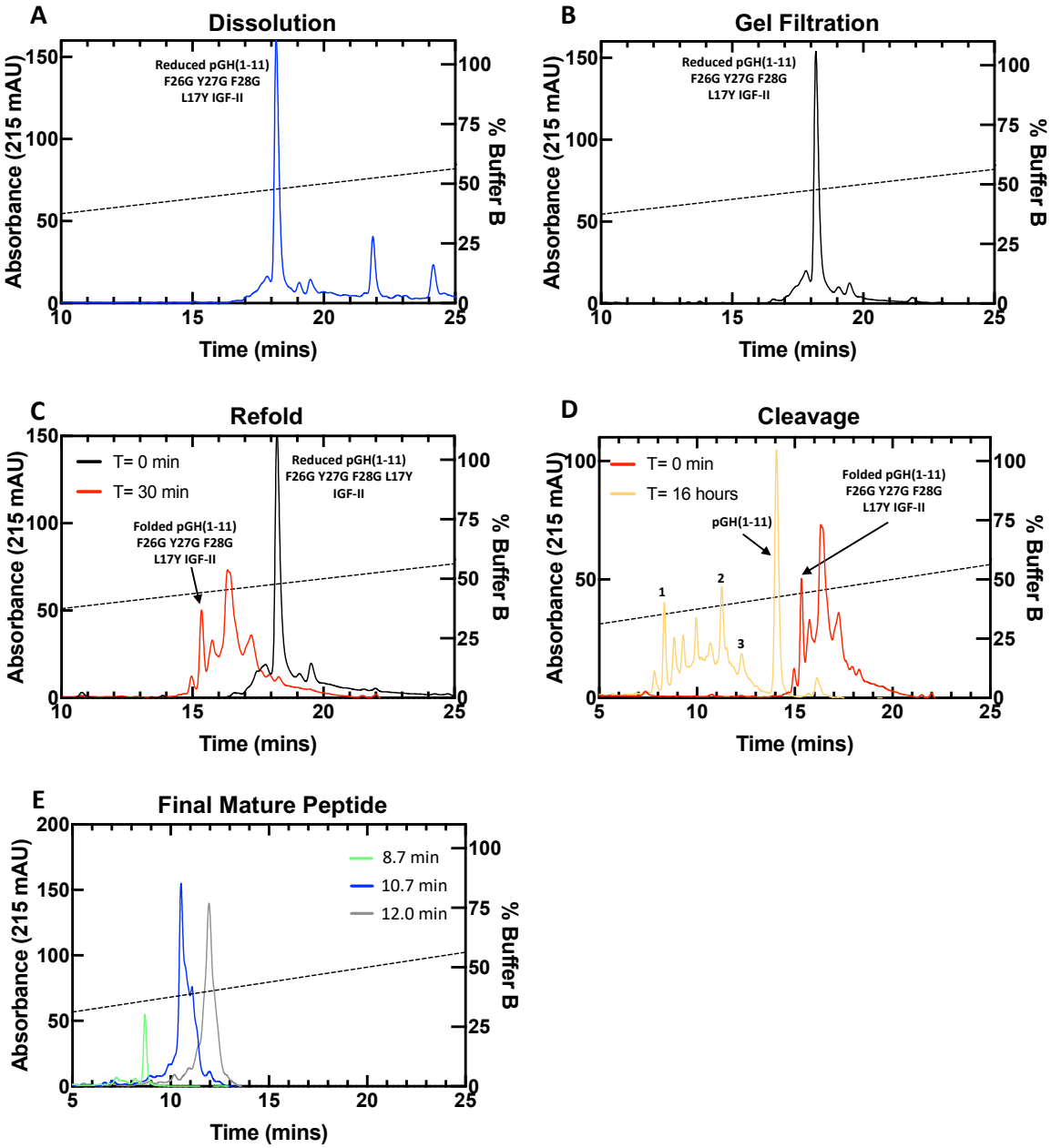


Figure 6.4: Purification of F26G Y27G F28G L17Y IGF-II.

Analytical reverse phase HPLC (using a C4 column and a 25 – 45 % CH_3CN (31.3 – 56.3% buffer B) over 20 minutes (Section 2.3.5)) was used to monitor each stage of protein purification. Shown are the absorbance chromatograms at 215nm (left Y axis). Dotted black line represents the % of buffer B (right Y axis). Samples analysed were (A) Dissolution, (B) post gel filtration, (C) the refold T=90 minutes (solid black line) and T=3 hours (red), (D) α-LP cleavage T=0 minutes (red) and T=16 hours (orange). (E) Three purified fractions of F26G Y27G F28G L17Y IGF-II (numbered 1-3 in D) with retention times of 8.7 minutes (green, peak 1 in D), 10.7 minutes (blue, peak 2 in D) and 12.0 minutes (grey, peak 3 in D).

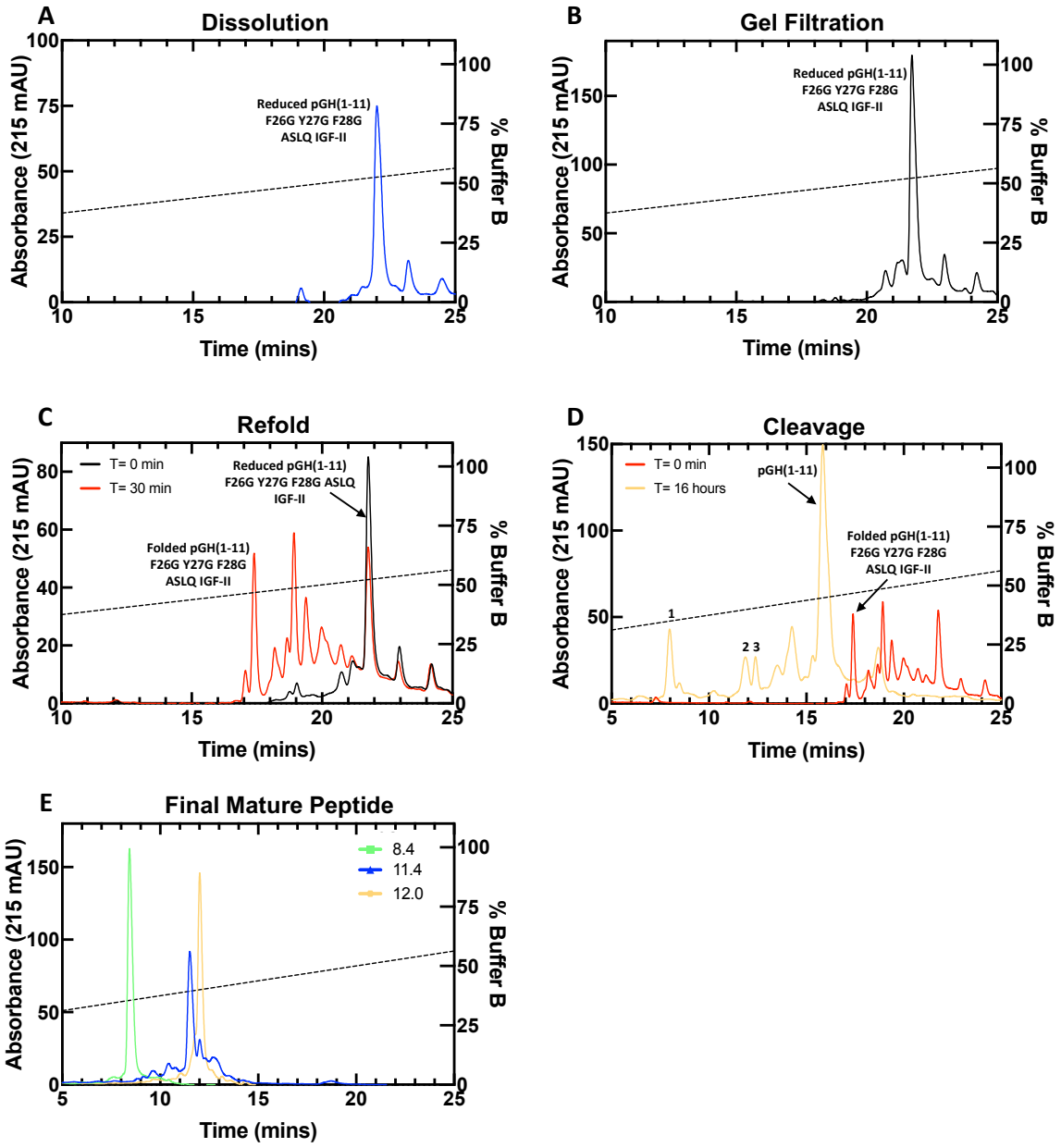


Figure 6.5: Purification of F26G Y27G F28G ASLQ IGF-II.

Analytical reverse phase HPLC (using a C4 column and a 25 – 45 % CH_3CN (31.3 – 56.3% buffer B) over 20 minutes (Section 2.3.5)) was used to monitor each stage of protein purification. Shown are the absorbance chromatograms at 215nm (left Y axis). Dotted black line represents the % of buffer B (right Y axis). Samples analysed were (A) Dissolution, (B) post gel filtration, (C) the refold T=30 minutes (solid black line) and T=3 hours (red), (D) α-LP cleavage T=0 minutes (red) and T=16 hours (orange). (E) Three purified fractions of F26G Y27G F28G ASLQ IGF-II (numbered 1-3 in D) with retention times of 8.4 minutes (green, peak 1 in D), 11.4 minutes (blue, peak 2 in D) and 12.0 minutes (orange, peak 3 in D).

6.2.2 Characterisation of IGF-II analogues

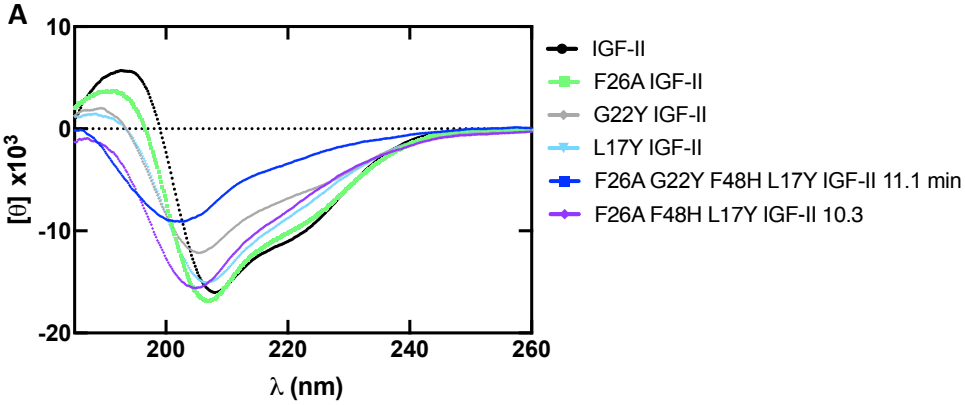
As mentioned above, where necessary, several peaks were isolated for secondary structure determination, IGF-1R binding assay screens and mass spectrometry to confirm mass of the analogues. CD profiles and IGF-1R binding assays of the screened fractions are shown in Supplementary Figures 20 - 26. Those peaks determined to have the correct mass, highest binding affinity and helical content most similar to native IGF-II are listed below and were taken on for further analysis in this chapter:

- F26A G22Y F48H L17Y IGF-II **11.1 min peak**
- F26A F48H L17Y IGF-II **10.3 min peak**
- F26G Y27G F28G G22Y F48H IGF-II **9.0 min peak**
- F26G Y27G F28G L17Y IGF-II **8.7 min peak**
- F26G Y27G F28G ASLQ **12.0 min peak**

6.2.2.1 Mass spectrometry and circular dichroism

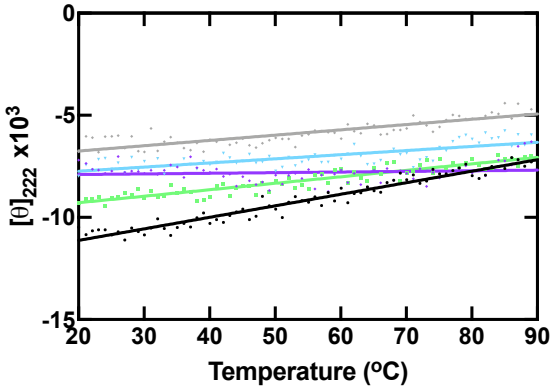
The mass of each IGF-II analogue was confirmed (Supplementary Table 4). Circular dichroism (CD) spectra between 180 and 260 nm were recorded for each IGF-II analogue (Section 2.4.4.1). (Figure 6.6A and corresponding Table). F26A IGF-II (characterised in Chapter 3) and its CD profile is shown in Figure 6.6A for comparison. F26A IGF-II was found to have a helical content of 37%, similar to that of native IGF-II. Individual mutation of G22Y and L17Y result in 19 and 23% helical content respectively (described in Chapter 4). Introduction of all three Con-Ins G1 alternate binding residues (F26A G22Y F48H L17Y IGF-II 11.1 min peak) results in near total loss of helical content (6%) and shift in the spectral minimum to a lower wavelength compared to native IGF-II. However, without the G22Y mutation (F26A F48H L17Y IGF-II 10.3 min peak) the resultant analogue has 17% helical content and a only small shift in the spectral minimum to a lower wavelength compared to native IGF-II.

A



Analogue	Helix	Strand	Turns	Unordered
IGF-II	41	13	16	30
L17Y IGF-II	23	22	15	41
G22Y IGF-II	19	24	14	44
F26A IGF-II 10.8 min	37	13	16	35
F26A G22Y F48H L17Y IGF-II 11.1 min	6	35	11	47
F26A F48H L17Y IGF-II 10.3 min	17	23	15	46

B



Analogue	R^2	Slope	$[\theta]_{222}x10^3$		% change from 20°C to 90°C
			20°C	90°C	
IGF-II	0.9462	0.0564	-10.5	-7.26	30.9
L17Y IGF-II	0.5946	0.0202	-7.98	-6.22	22.1
G22Y IGF-II	0.7146	0.0259	-6.78	-4.20	38.1
F26A IGF-II 10.8 min	0.8621	0.0317	-9.29	-6.94	25.3
F26A F48H L17Y IGF-II 10.3 min	0.01769	0.00305	-7.20	-7.59	5.42

Figure 6.6: Analysis of IGF-II analogue secondary structure and thermal stability.

(A) Secondary structure analysis of each IGF-II analogue was performed using circular dichroism at 0.05 μg/μL in 10 mM acetic acid, mean residue ellipticity [θ] was measured from 260 - 180 nm (Section 2.4.4.1). Percent helical content shown in the table was performed using Dichroweb and was calculated using the CDSSTR algorithm for deconvolution against the reference protein database set SMP180 (Sreerama and Woody 2000). (B) Thermal stability was monitored by measuring mean residue ellipticity [θ] at 222 nm from 20°C to 90°C (at 1±0.1°C increments) (Section 2.4.4.1.1). CD profiles and thermal stability of single mutants (Chapter 4) and site 1a minimised IGF-II mutants (Chapter 3) are also shown for comparison. Temperature denaturation was not performed for F26A G22Y F48H L17Y IGF-II.

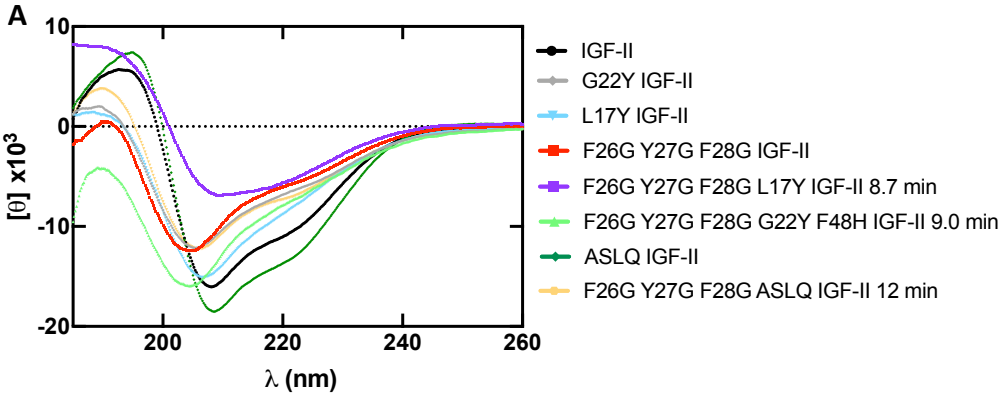
F26G Y27G F28G IGF-II site 1a minimised analogue (Chapter 3) resulted in a marked reduction of helical content (20%) compared to native IGF-II (41%) (Figure 6.7A). Not surprisingly, addition of L17Y to this site 1a minimised analogue further reduced helical content to 15%. Individual mutation of L17Y in IGF-II (Chapter 4) led to a helical content of 23%. Interestingly, introduction of both G22Y and F48H mutations (F26G Y27G F28G G22Y F48H IGF-II) increased helical content compared to the parent site 1a IGF-II analogue (F26G Y27G F28G IGF-II) (Figure 6.7A). Finally, as described in Chapter 4, ASLQ IGF-II increased helical content to 49% compared to 41% for native IGF-II. This modified D-domain in the context of F26G Y27G F28G IGF-II resulted in a similar helical content (21%) compared to the parent site 1a minimised analogue alone (F26G Y27G F28G IGF-II with a helical content of 20%).

6.2.2.2 Thermal denaturation studies

Thermal denaturation studies were performed from 20°C to 90°C and helical content was monitored at the helical sensitive wavelength 222 nm (Section 2.4.4.1.1). At the beginning of thermal denaturation (20±0.1°C) all IGF-II analogues had $[\theta]_{222}$ values consistent with CD spectra (Figure 6.6B – 6.7B and corresponding Tables). F26A F48H L17Y IGF-II analogue resulted in a change in helical content from 20°C to 90°C of 5.42% but this is likely due to its low starting helical content.

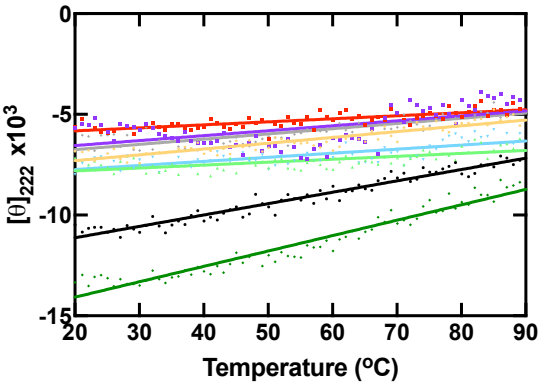
The F26G Y27G F28G IGF-II site 1a minimised analogues are shown in Figure 6.7B. All variations of the F26G Y27G F28G IGF-II site 1a minimised analogues have lower $[\theta]_{222}$ values at the completion of temperature denaturation (90±0.1°C) compared to the beginning (20±0.1°C) indicating loss of helical content. At completion of temperature denaturation, the F26G Y27G F28G ASLQ analogue had a similar $[\theta]_{222}$ value whereas all other F26G Y27G F28G IGF-II analogues had lower values compared to native IGF-II, indicating further loss of helical content.

A



Analogue	Helix	Strand	Turns	Unordered
IGF-II	41	13	16	30
L17Y IGF-II	23	22	15	41
G22Y IGF-II	19	24	14	44
F26G Y27G F28G IGF-II 9.6 min	20	22	15	44
F26G Y27G F28G G22Y F48H IGF-II	23	25	12	40
F26G Y27G F28G L17Y IGF-II 8.7 min	15	35	12	37
ASLQ IGF-II	49	12	16	23
F26G Y27G F28G ASLQ IGF-II 12.0 min	21	24	14	42

B



Analogue	R^2	Slope	$[\theta]_{222}x10^3$		% change from 20°C to 90°C
			20°C	90°C	
IGF-II	0.9462	0.0564	-10.5	-7.26	30.9
L17Y IGF-II	0.5946	0.0202	-7.98	-6.22	22.1
G22Y IGF-II	0.7146	0.0259	-6.78	-4.20	38.1
F26G Y27G F28G IGF-II 9.6 min	0.6034	0.0150	-5.64	-4.54	19.5
F26G Y27G F28G G22Y F48H IGF-II 9.0 min	0.4491	0.0145	-7.16	-5.95	16.9
F26G Y27G F28G L17Y IGF-II 8.7 min	0.3093	0.0246	-5.66	-4.06	28.3
ASLQ IGF-II	0.9382	0.0764	-13.3	-8.55	35.7
F26G Y27G F28G ASLQ IGF-II 12.0 min	0.8247	0.0288	-6.83	-5.14	24.7

Figure 6.7: Analysis of IGF-II analogue secondary structure and thermal stability.

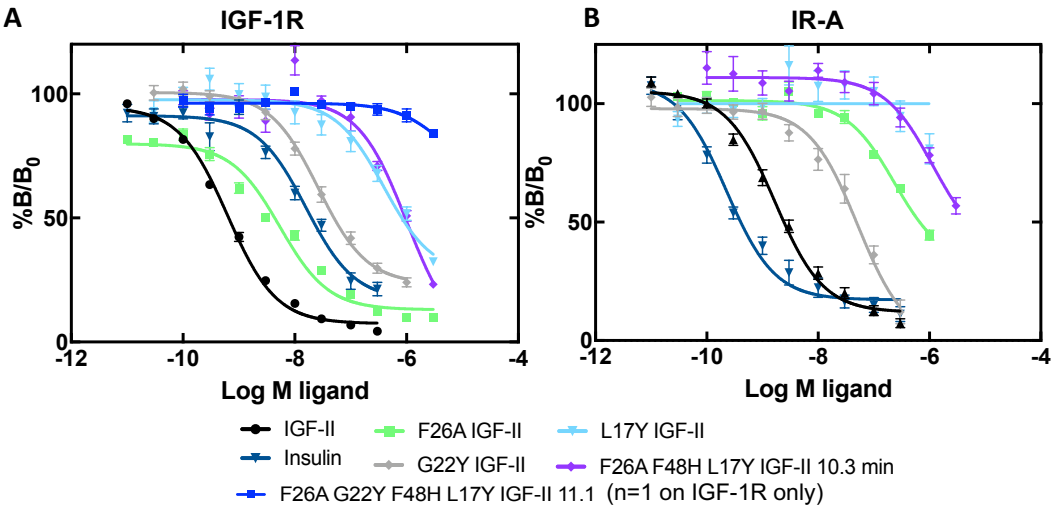
(A) Secondary structure analysis of each IGF-II analogue was performed using circular dichroism (CD) at 0.05 μg/μL in 10 mM acetic acid, mean residue ellipticity [θ] was measured from 260 - 180 nm (Section 2.4.4.1). Percent helical content shown in the table was performed using Dichroweb and was calculated using the CDSSTR algorithm for deconvolution against the reference protein database set SMP180 (Sreerama and Woody 2000). (B) Thermal stability was monitored by measuring mean residue ellipticity [θ] at 222 nm from 20°C to 90°C (at 1±0.1°C increments) (Section 2.4.4.1.1). CD profiles and thermal stability of single mutants (Chapter 4) and site 1a minimised IGF-II mutants (Chapter 3) are also shown for comparison.

6.2.2.3 IGF-1R competition binding assays

IGF-1R and IR-A competition binding assays were performed using europium labelled IGF-II (Eu-IGF-II) and increasing concentrations of competing ligand (Section 2.4.1). F26A IGF-II site 1a minimised alone had a decreased affinity for the IGF-1R and IR-A 11.5% and <1% respectively, compared to native IGF-II (discussed in Chapter 3) (Figure 6.8). Individual G22Y and L17Y IGF-II mutants also had reduced affinities for both receptors (Figure 6.8). F26A G22Y F48H L17Y IGF-II had near total loss of affinity for the IGF-1R and IR-A. Interestingly, removal of the G22Y (F26A F48H L17Y IGF-II) re-establishes binding somewhat, although the affinities were still only <1% for the IGF-1R and <0.01% for the IR-A compared to native IGF-II. Notably, all mutants shown in Figure 6.8 resulted in a more drastic loss of affinity for the IR-A compared to their affinities for the IGF-1R.

F26G Y27G F28G IGF-II site 1a minimised analogue had <0.01% affinity for the IGF-1R, compared to native IGF-II, and total loss of binding affinity for the IR-A (Figure 6.9) Addition of L17Y to this analogue (F26G Y27G F28G L17Y IGF-II) further reduced affinity for the IGF-1R compared to native IGF-II and also resulted in total loss of affinity for the IR-A. Interestingly, F26G Y27G F28G G22Y F48H IGF-II was able to compete with Eu-IGF-II over the concentration range of 10-300 nM albeit poorly, with 300nM competition resulting in 60% binding of Eu-IGF-II. However, at 1000 and 3000 nM competing F26G Y27G F28G G22Y F48H IGF-II there was an apparent increase in Eu-IGFII binding. This infers that at the highest concentrations more Eu-IGF-II can bind. One possibility to explain this is that the competing F26G Y27G F28G G22Y F48H IGF-II maybe aggregating at those concentrations. However, further investigation would be required to confirm this.

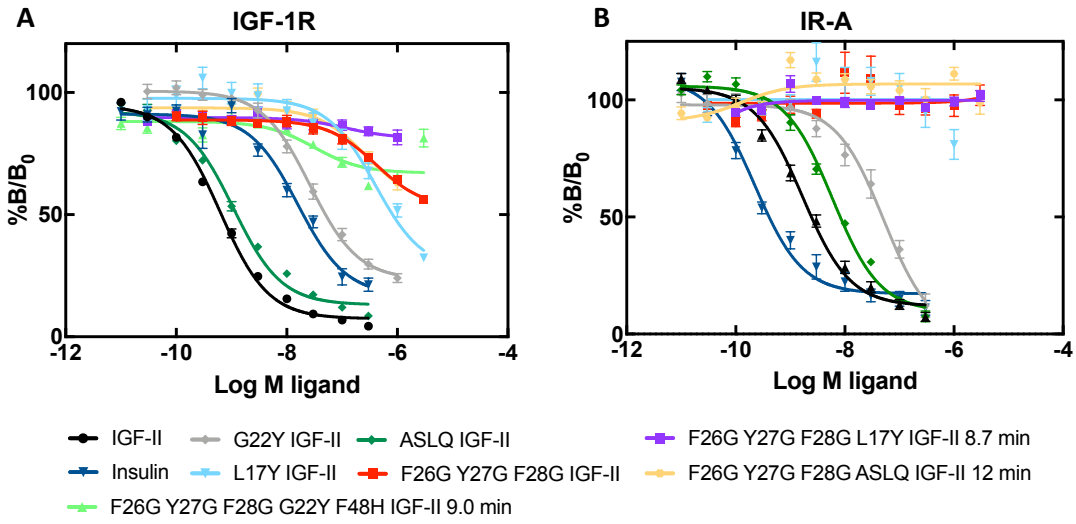
ASLQ IGF-II bound the IGF-1R with relatively high affinity (58.0% compared to native IGF-II) but there was more of an impact on affinity for the IR-A (28.2% compared to IGF-II) (discussed in Chapter 4)(Figure 6.9). F26G Y27G F28G ASLQ IGF-II had an affinity for the IGF-1R of <0.1% compared to native IGF-II, equal to that of the parent site 1a minimised IGF-II analogue (F26G Y27G F28G IGF-II).



Analogue	IGF-1R			IR-A		
	IC_{50} (nM)	95% CI (nM)	Relative binding affinity (% of IGF-II)	IC_{50} (nM)	95% CI (nM)	Relative binding affinity (% of IGF-II)
IGF-II	0.626	0.568-0.689	100	1.68	1.36-2.09	100
Insulin	15.9	10.7-23.7	3.94	0.226	0.156-0.329	743
F26A IGF-II	5.42	4.11-7.15	11.5	>100	ND	ND
L17Y IGF-II	>100	ND	ND	>1000	ND	ND
G22Y IGF-II	8.45	6.50-11.0	7.41	51.3	31.4-85.4	3.27
F26A G22Y F48H L17Y IGF-II	ND	ND	ND	ND	ND	ND
F26A F48H L17Y IGF-II	>100	ND	ND	>1000	ND	ND

Figure 6.8: Competition of IGF-II analogues with europium-labelled IGF-II for binding the IGF-1R and IR-A.

Immunocaptured IGF-1R or IR-A was incubated with europium-labelled IGF-II (Eu-IGF-II) and increasing concentrations of competing IGF-II, insulin or IGF-II analogue. Time-resolved fluorescence was measured as described in Section 2.4.1. Results are expressed as a percentage of Eu-IGF-II bound in the absence of competing ligand (B/B_0). Data shown are the mean ±S.E.M. n = ≥ 3 independent experiments each with triplicate technical replicates (unless otherwise indicated on the graph). Error bars are shown when greater than the size of the symbols. Calculated IC_{50} values are shown in the table. Binding curves and calculated IC_{50} values of single mutants (Chapter 4) and site 1a minimised IGF-II mutants (Chapter 3) are also shown for comparison.



Analogue	IGF-1R			IR-A		
	IC_{50} (nM)	95% CI (nM)	Relative binding affinity (% of IGF-II)	IC_{50} (nM)	95% CI (nM)	Relative binding affinity (% of IGF-II)
IGF-II	0.626	0.568-0.689	100	1.68	1.36-2.09	100
Insulin	15.9	10.7-23.7	3.94	0.226	0.156-0.329	743
F26G Y27G F28G IGF-II	>1000	ND	ND	ND	ND	ND
L17Y IGF-II	>100	ND	ND	>1000	ND	ND
G22Y IGF-II	8.45	6.50-11.0	7.41	51.3	31.4-85.4	3.27
ASLQ IGF-II	1.08	0.894-13.1	58.0	5.95	4.80-7.38	28.2
F26G Y27G F28G ASLQ IGF-II	>1000	ND	ND	ND	ND	ND
F26G Y27G F28G L17Y IGF-II	ND	ND	ND	ND	ND	ND
F26G Y27G F28G G22Y F48H IGF-II	ND	ND	ND	ND	ND	ND

Figure 6.9: Competition of IGF-II analogues with europium-labelled IGF-II for binding the IGF-1R and IR-A.

Immunocaptured IGF-1R (A) or IR-A (B) was incubated with europium-labelled IGF-II (Eu-IGF-II) and increasing concentrations of competing IGF-II, insulin or IGF-II analogue. Time-resolved fluorescence was measured as described in Section 2.4.1. Results are expressed as a percentage of Eu-IGF-II bound in the absence of competing ligand (B/B_0). Data shown are the mean ±S.E.M. n = ≥ 3 independent experiments each with triplicate technical replicates (unless otherwise indicated on the graph). Error bars are shown when greater than the size of the symbols. Calculated IC_{50} values are shown in the table. Binding curves and calculated IC_{50} values of single mutants (Chapter 4) and site 1a minimised IGF-II mutants (Chapter 3) are also shown for comparison.

6.3 Discussion

IGF-II is a single-chained ligand, with a C-domain connecting the B- and A-domains. This is the major difference between IGF-II and insulin. While insulin also possesses a C-domain, upon expression this domain is proteolytically cleaved to give a mature two-chained molecule (Kemmler et al., 1971). To preserve the single chain structure of IGF-II when producing site 1a minimised IGF-II analogues in Chapter 3 a mutational approach was used to remove elements of the FYF motif.

In this chapter we first explored the alternate binding residues of Con-Ins G1 in the context of the F26A IGF-II site 1a minimised analogue. F26A IGF-II alone was shown through CD studies to retain native-like IGF-II secondary structure (discussed in detail in Chapter 3). Three additional mutations were made to the F26A IGF-II analogue (L17Y, G22Y and F48H) to mimic the Con-Ins G1 alternate binding mechanism (F26A G22Y F48H L17Y IGF-II). This peptide was found to have essentially no helical content or affinity for the IGF-1R or IR-A. However, a second analogue was produced, without the tyrosine mutation at position 22 (F26A F48H L17Y IGF-II) and this analogue was shown to have 17% helical content and <0.1% affinity for the IGF-1R and 0.01% affinity for the IR-A. It was found in Chapter 4 that mutation of residue G22 to a tyrosine in IGF-II was poorly tolerated and this was most likely attributed to the type 1 β turn being required for IGF-II structure. Here in the context of the site 1a minimised IGF-II analogues it is likely that the β turn at residues 22-25 have been perturbed by introduction of a tyrosine at position 22. In the context of Mini-Ins and Con-Ins G1 a tyrosine residue can be accommodated at position B20, most likely due to the lack of the C-terminal 8 amino acids of the B-chain.

In Chapter 4 we speculated, using structural overlays of native IGF-II and Con-Ins G1, that residues F26 and L17Y, together in IGF-II, may result in a steric clash for the same spatial position in the protein and during receptor engagement. However, it was anticipated that in the context of the site 1a minimised IGF-II analogues, where the F26 has been mutated, this may allow L17Y to engage with the hydrophobic pocket of the L1 domain. However, in both F26A G22Y F48H L17Y IGF-II and F26A F48H L17Y IGF-II there appeared to be no increase in binding affinity compared to L17Y IGF-II analogue tested in Chapter 4.

In this chapter the alternative binding mechanism of Con-Ins G1 was also combined with the F26G Y27G F28G IGF-II site 1a minimised analogue. F26G Y27G F28G L17Y IGF-

II was first tested and found that L17Y further reduced helical content (15%), compared to the site 1a minimised analogue alone (F26G Y27G F28G IGF-II, 20% helical content) and had essentially no affinity for the IGF-1R or IR-A. Mutations to the equivalent residue in insulin (LB15) were unable to be efficiently folded using a yeast expression system (Kristensen et al., 1997) and structures of human insulin and native IGF-II depict the side chain of this residue buried in the core of the protein (Torres et al., 1995, Blundell et al., 1972). Here we were able to produce a small proportion of folded material. However, this analogue was found to have significantly attenuated helical content in comparison to native IGF-II and essentially no biological activity. YB15 in Con-Ins G1 is observed to engage with the hydrophobic pocket of the L1 domain and when mutated to alanine results in a ~10-fold reduction in Akt signaling potency (Xiong et al., 2020). In contrast, YB15 in the context of DOI does not increase Akt signaling potency compared to DOI (Xiong et al., 2020). These results are consistent with our observation that in the context of site the 1a minimised IGF-II analogue (F26G Y27G F28G IGF-II equivalent to DOI), L17Y does not alter biological activity. A major difference in the binding of native IGF-II and human insulin is the requirement of the IGF-II C-domain to thread through the αCT′ upon binding to site 1a (Xu et al., 2020). Mature insulin does not contain a C-domain. In the context of the site 1a minimised IGF-II analogues with mutations from Con-Ins G1, discussed above, we speculate that a constraint resulting from the threading of the C-domain through the αCT′ precludes the L17Y residue from engaging with the hydrophobic pocket of the L1 domain.

Shown in Figure 6.10 are the 20 conformations of native IGF-II obtained using nuclear magnetic resonance (NMR) (Torres et al., 1995). In the native IGF-II structure there is a high degree of variability in the structure of the C-domain in each of the structures suggesting that this region is highly flexible (Torres et al., 1995). In native IGF-II the FYF motif is located at the N-terminus of the C-domain and has a well-defined position in the protein, with residues F26 and F28 interacting with the core of the protein (Figure 6.10) (Torres et al., 1995). Based on these solution structures of native IGF-II it is anticipated that the FYF motif in IGF-II is particularly important in this ligand to engage with the protein core to stabilise the C-domain and C-terminal segment of the B-domain, which are highly flexible. It is therefore not surprising that our attempts to mutate the FYF motif in IGF-II to glycine residues (F26G Y27G F28G) resulted in markedly altered secondary structure compared to native IGF-II. The resultant analogue likely having additional flexibility and disorder in this region due to the loss of the hydrophobic side chains of the FYF motif. Indeed, insulin residue FB24 of insulin is also known to be required for protein

folding, including formation of the B19-A20 disulfide bond (Rege et al., 2020b).

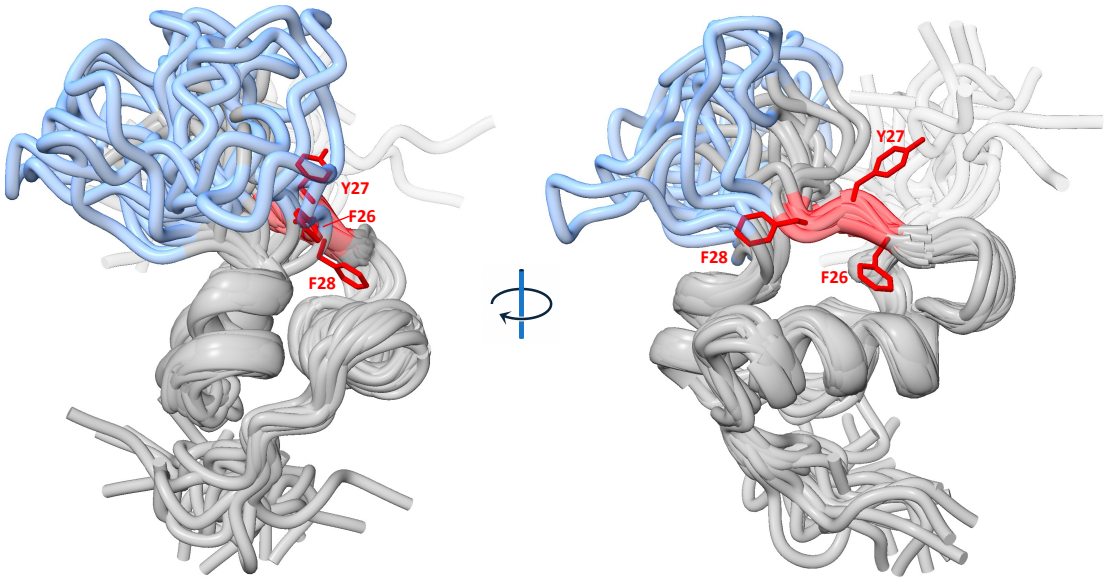


Figure 6.10: IGF-II solution structures.

Solution structures of IGF-II determined using nuclear magnetic resonance (NMR) (PDB: 1IGL). Each of the twenty conformations of IGF-II are shown with the C-domain coloured in cornflower blue. The aromatic FYF motif is coloured in red with side chains shown. The C-domain of IGF-II has a high degree of conformational flexibility. In twenty conformations the aromatic FYF motif is well defined with residues F28 and F26 interacting with the core of the protein.

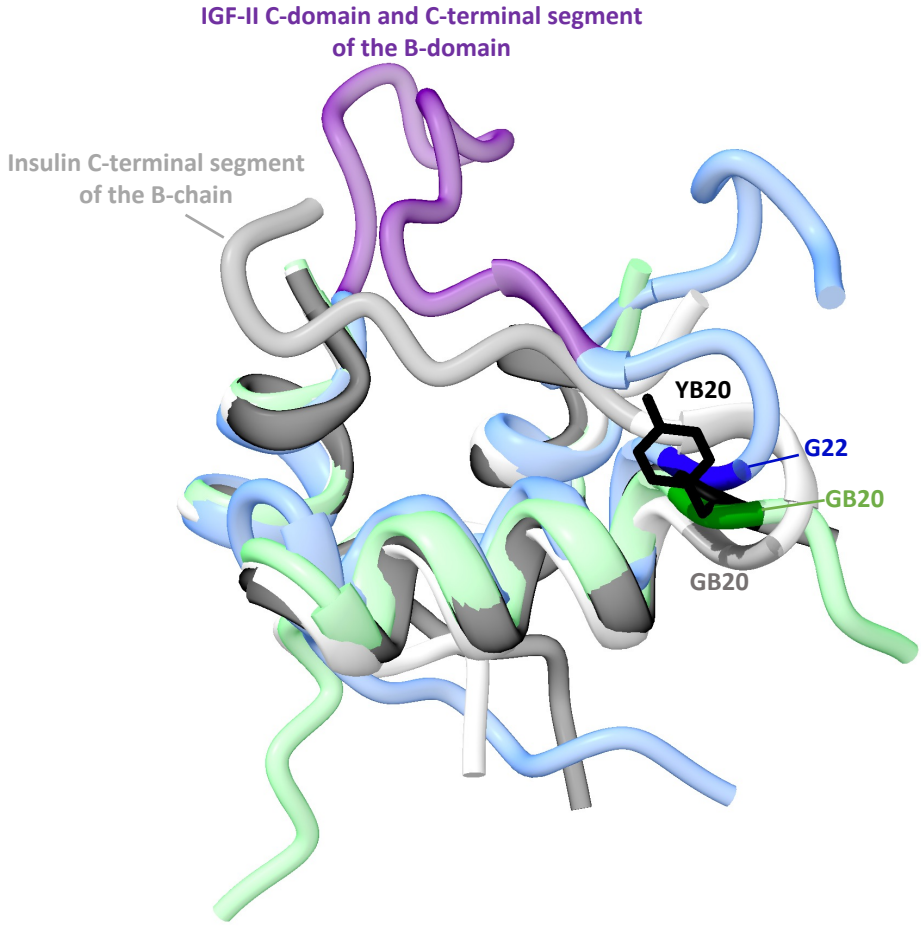


Figure 6.11: Structural alignment of des heptapeptide insulin, human insulin and native IGF-II.

Structural alignment of des heptapeptide insulin (green PDB: 1DEI), human insulin (white PDB: 3I3Z) and native IGF-II (cornflower blue PDB: 1IGL). Human insulin is a two chained ligand. Residues GB20, EB21, RB22 and GB23 form a type 1 β turn to facilitate the C-terminal segment of the B-chain (grey) to fold back against the B-chain helix. When the C-terminal segment of the B-chain in human insulin is removed as done in desheptapeptide (B24-B30) insulin, residues GB20, EB21, RB22 and GB23 no longer form a β turn. Native IGF-II is a single chained ligand, with a flexible C-domain (coloured in purple) connecting to the C-terminal segment of the B-domain. Due to the presence of the C-domain the type 1 β turn including residue G22 will likely be required to allow for the C-terminal segment of the B-domain to fold back against the B-domain helix. Therefore it is not surprising that attempts to mutate residue G22 to leucine and tyrosine were poorly tolerated as they likely perturbed the formation of the β turn.

As described above residue G22 in native IGF-II is involved in a type 1 β turn (Torres et al., 1995). In human insulin an equivalent β turn is also essential for protein structure, especially the folding of the C-terminal segment of the B-domain back against the B-chain helix. Solved structures of a human insulin analogue lacking the C-terminal 7 amino acids of the B-chain (desheptapeptide (B24-B30) insulin) revealed that residue GB20 (and residues EB21, RB22 and GB23) no longer forms a β turn (Figure 6.11) (Bao et al., 1997). Therefore, it is not surprising that mutations to residue GB20 are tolerated in insulin analogues lacking the C-terminal segment of the B-chain. Comparing these structures (Figure 6.11) to that of native IGF-II provides clues as to why these mutations are poorly tolerated in the single chains IGF-II structure. The presence of the flexible C-domain and the C-terminal segment of the B-domain means there is still a requirement for the type 1 β turn. Consistent with the finding presented in this chapter, introduction of large side chained amino acids such as tyrosine and leucine to position 22 in IGF-II perturbs the native IGF-II structure, as the β turn likely does not form correctly. Future work may be performed to entirely remove the C-domain of these IGF-II analogues, (including the C-terminal 8 amino acids of the B-domain) and explore the ability of a two chained IGF-II analogue to bind and activate the IGF-1R and IR-A.

A major difference between Vh-Ins and Mini-Ins is the affinity for site 1a. Mini-Ins has a weaker affinity for site 1a than Vh-Ins and insulin (57.1-128-fold lower than insulin), whereas Vh-Ins has a higher affinity for site 1a than Mini-Ins (only 9.91- fold lower than insulin) (Xiong et al., 2021, Xiong et al., 2020). Conversely, Mini-Ins is shown through mutagenesis to have a stronger interaction with site 1b compared to human insulin. In addition, molecular modelling of Mini-Ins bound to the human IR depicts interactions between the side chains of RA9 and EB10 with the FnIII-1′ domain of the receptor (Xiong et al., 2020). For Vh-Ins there is a lack of side chain density at site 1b, suggesting that is has a weaker interaction at site 1b (Xiong et al., 2021).

As described in detail in Chapter 4 we produced an IGF-II analogue with the native D-domain replaced with an A61, S62 L63, Q64 IGF-II sequence, with the key elements of the Vh-Ins modified A-chain. This analogue also had increase in helical content (48%), likely by extension of the C-terminal A-domain helix. In addition, this analogue maintained high affinity for the IGF-1R. Structural overlays of native IGF-II and Vh-Ins (Figure 4.4) suggest that the C-terminal A-chain extension likely occupies a similar space as the FYF motif in IGF-II. We therefore postulated that in the context of an otherwise native IGF-II structure (ASLQ IGF-II), the FYF motif is likely to be interacting with the L1

domain of the receptor. In this chapter we investigated the modified D-domain in the context of the site 1a minimised analogue (F26G Y27G F28G ASLQ IGF-II) to determine what effect this modified D-domain had when the FYF motif was absent. This analogue had essentially the same helical content as the site 1a minimised analogue (F26G Y27G F28G IGF-II) alone and bound to the IGF-1R with the same affinity. Both the site 1a minimised analogue (F26G Y27G F28G IGF-II) and F26G Y27G F28G ASLQ IGF-II had essentially no affinity for the IR-A. The similarity in binding affinities of these analogues suggests that site 1a contact has not been re-established by introduction of the modified D-domain. However, further investigation would be required to confirm this. Isothermal titration calorimetry (ITC) could be used to determine the affinity of these analogues for site 1a.

In human insulin, several investigations have highlighted the importance of FB24 in insulin for early folding events. Mutation of this residue to any of the 19 other naturally occurring amino acids was shown to result in impaired biosynthesis and secretion in HEK293T cells (Rege et al., 2020b). However, a study published by Hua et al (1996) demonstrated that peptide synthesis could be used to produce B24Y insulin, which was also shown to be biologically active. There is a possibility that in this investigation the mutation of residue F26 in IGF-II (equivalent to residue FB24 in insulin) is also affecting the nascent folding events during the chemical refold. Peptide synthesis may be used in future investigations to determine if a more native like IGF-II secondary structure can be formed.

In summary, in this chapter we have explored two mechanisms of compensation for the removal of elements of site 1a, as found in Con-Ins G1 and Con-Ins K1. Using site 1a minimised IGF-II analogues produced in Chapter 4 we introduced equivalent residues in Con-Ins G1 and Con-Ins K1 that were found to compensate for the missing FFY motif in insulin. We show here that in IGF-II a high affinity binding interaction could not be re-established.

7 FINAL DISCUSSION

7.1 Introduction

A considerable amount of research at the molecular level has been directed toward understanding insulin and its interaction with the IR with the aim to develop drug therapies for diabetes that better emulate normal insulin release in the human body. The high degree of sequence and structural similarities between insulin and IGF-II has indirectly advanced our understanding of IGF-II through the plethora of insulin studies. However, IGF-II specific research has made it clear that despite the similarities between the ligands there are unique structural elements that contribute to differences in receptor specificity and signaling outcome. Indeed, this has been highlighted by studies emerging during the course of my PhD including the first structure of IGF-II bound to the IGF-1R solved by cryogenic electron microscopy (cryoEM) (Xu et al., 2020). Now, armed with solved structures of each of the three ligands in complex with their cognate receptors, we can further interrogate the subtle differences that lead to the altered signaling properties of each ligand. The overall goal of my PhD project was to better understand the mechanism of IGF-II action and how signaling pathways are preferentially activated. The ultimate goal of our laboratory is to use this knowledge to develop inhibitors of IGF-II action in cancer. To achieve the overall goal of my PhD we explored the effect of introducing alternate binding mechanisms into IGF-II, based on knowledge from cone snail venom insulins, to determine the effect on receptor binding and activation. These investigations also led to a greater understanding of native IGF-II structure and function.

While the key receptor binding contacts on insulin and the IGFs have been well mapped via mutagenesis and structural investigations, several residues have been found to contribute to altered signaling potency. The most studied example of this is residue histidine B10 in insulin. Mutation of this residue to acidic side chains, including glutamate (HB10E insulin) and aspartate (HB10D insulin) leads to an increased binding affinity for both the IGF-1R and IR, and also leads to a disproportionate increase in mitogenic potency compared to its metabolic potency (Kaarsholm et al., 1993, Xiong et al., 2020, Schwartz et al., 1987, Kurtzhals et al., 2000, Brange et al., 1988, Glendorf et al., 2012). Importantly, the HB10D analogue was found to have an increased IR residence time (Glendorf et al., 2012, Drejer et al., 1991) resulting in sustained receptor phosphorylation and receptor internalisation (Hamel et al., 1999, Drejer, 1992).

In recent years, the identification of cone snail venom insulins, led to the discovery of the smallest known insulin in nature, which despite the loss of the C-terminal segment of the B-chain, including the aromatic (FB24 FB25 YB26) motif and other key elements of the

site 1a, still bind the IR (Xiong et al., 2021, Xiong et al., 2020). These discoveries shifted research interest in this field toward insulins that bind to the receptor with distinctly different mechanisms, compared to human insulin, in an effort to uncover novel analogues with unique signaling properties. Two separate investigations have since introduced the alternate binding mechanisms of cone snail venom insulins into the DOI analogue and restored human insulin-like affinity for the IR (Xiong et al., 2021, Xiong et al., 2020). These alternative mechanisms of IR and IGF-1R binding and activation sparked my interest. Based on several cone snail insulins including Con-Ins G1 and Con-Ins K1, I aimed to incorporate the key alternative binding contacts into IGF-II and investigate the effect on signaling.

7.2 Discussion of key findings and future work

To investigate the effect of introducing alternate binding mechanisms to compensate for the minimised 1a (FYF motif) in IGF-II a stepwise approach was used. Firstly, several site 1a minimised analogues of IGF-II were produced, each with mutations to the aromatic FYF motif, which is critical for high affinity binding to site 1a of the receptor (L1 domain and αCT′). Biophysical and receptor binding analyses of the site 1a minimised analogues provided new insights into the importance of the FYF motif in IGF-II structural stability and folding as well as the role of this motif in IGF-IIs receptor binding specificity. In addition, these site 1a minimised analogues of IGF-II provided us with low affinity templates into which potential alternate binding mechanisms could be incorporated. Prior to their incorporation into the site 1a minimised analogues, each Con-Ins G1 and Con-Ins K1 residue that was found to compensate for the lack of the FFY motif in insulin, was individually mutated in human IGF-II to test their effect on structure, stability and biological activity. These investigations led to a more comprehensive understanding of residues important for IGF-II folding, structure and stability. Finally, residues of Con-Ins G1 and Con-Ins K1, that compensate for the loss of the FFY motif, were incorporated into the site 1a minimised IGF-II analogues. From these investigations we have for the first time attempted to remove key elements of the classical site 1a binding motif in IGF-II (FYF motif) and introduce alternative binding mechanisms that compensate for the loss of this motif, based off Con-Ins G1 and Con-Ins K1.

7.2.1 Site 1a minimised analogues

Removal of the site 1a FYF motif of IGF-II was particularly challenging for several reasons. The first challenge is that IGF-II is a single chain mature peptide. Insulin,

however, is a two-chained mature peptide, with the aromatic FFY motif positioned at the C-terminus of the B-chain. Removal of the aromatic motif was therefore possible by truncating the B-domain, as done in (des octapeptide (B23–B30)-insulin (DOI) (Weiss and Lawrence, 2018, Bromer and Chance, 1967, Cara et al., 1990). For IGF-II, to maintain the single chain structure a mutational approach was used to perturb the FYF motif. While there is a detailed understanding of the importance of this aromatic motif in insulin, there was little information specific to IGF-II and indeed there have only been a few attempts at mutating residues of this region in IGF-II (Sakano et al., 1991, Roth et al., 1991, Burgisser et al., 1991, Forbes et al., 2002). These investigations showed that mutation of residues F26 or Y27 led to a greater reduction in affinity for the IR-A than the IGF-1R, whilst mutation of residue F28 resulted in a similar reduction in affinity for both the IGF-1R and IR-A (Sakano et al., 1991, Roth et al., 1991, Burgisser et al., 1991, Forbes et al., 2002). Since alanine scanning mutagenesis has not been performed on this motif in IGF-II it is difficult to make comparisons on the relative importance of each residue in receptor binding. In this investigation I have provided additional information toward the importance of the FYF motif in IGF-II. In particular, I found that all mutations of the IGF-II FYF motif resulted in a more marked loss of affinity for the IR-A than the IGF-1R, even when taking the lower affinity of IGF-II for IR-A into account. We therefore concluded that the aromatic FYF motif of IGF-II plays a more important role in IR-A binding than in IGF-1R binding. Indeed, these findings are consistent with insulin studies that have demonstrated the importance of the FFY motif in IR binding (Zakova et al., 2013).

Interestingly, mutation of residue FB24 of insulin (FB24A, equivalent residue to F26 of IGF-II) results in a 278-fold reduction in affinity for the IR (Pandyarajan et al., 2014, Kobayashi et al., 1984), whereas the F26A IGF-II analogue only has a ~10-fold lower affinity. The difference between these equivalent insulin and IGF-II mutants is surprising, particularly when the side chains occupy the same position (Figure 7.1). A particularly interesting observation from the Mini-Ins study suggests that the affinity of the site 1a interaction (L1 domain and αCT′, refer to Figure 1.5C) can influence the affinity of the site 1b interaction (FnIII-1′ refer to Figure 1.5D). Mutation of HA8 in Mini-Ins to that of human insulin (threonine) decreases affinity >20 fold (Xiong et al., 2020) and yet the effect of the mutating the same residue in the context of human insulin is much less: mutation of residue TA8 in human insulin to histidine only resulted in a 2-3-fold increase in affinity for the IR (Kaarsholm et al., 1993, Schaffer, 1994). We hypothesise that the lower affinity of Mini-Ins for the site 1a (128-fold decrease compared to human insulin)

enables a stronger site 1b interaction with the FnIII-1′ domain (Xiong et al., 2020, Kiselyov et al., 2009). Site 1a and 1b are located on opposite surfaces of the ligand, the lowering of the site 1a affinity in Mini-Ins may allow for the ligand to engage more strongly with site 1b including the histidine at position A8 (Xiong et al., 2020). Indeed, the overall 10-fold reduction in affinity of F26A IGF-II for the IR-A does not inform us on the specific site 1a and site 1b affinity and it would be informative to individually determine this. We speculate that F26A IGF-II has a lower affinity at site 1a and a higher affinity at site 1b, compared to human insulin, afforded by the phenylalanine at position 48 (equivalent to residue TA8 in insulin).

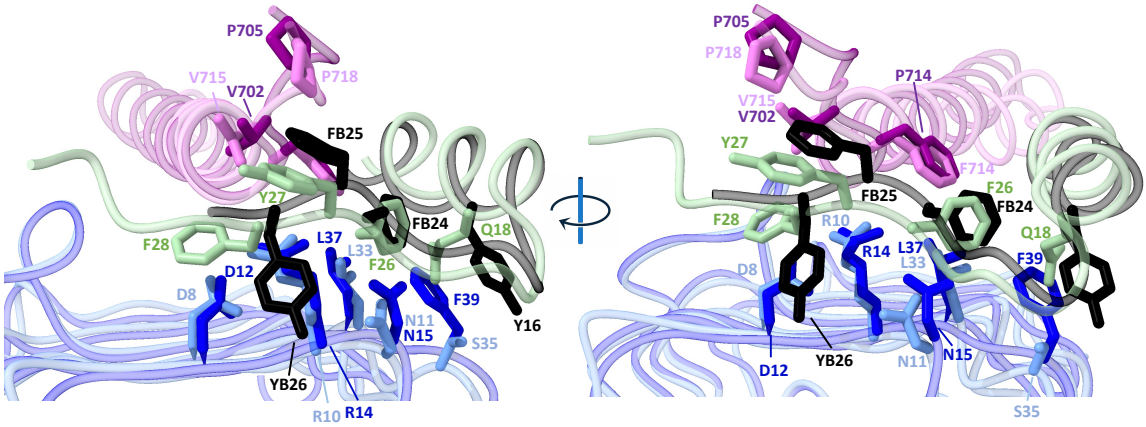


Figure 7.1: Overlay of Insulin and IGF-II B-domains interacting with site 1a.

Insulin (black) and IGF-II (green) B-domain bound α helix and C-terminus interacting with site 1a of their cognate receptors (from PDB: 6HN5 and 6VWI, respectively). L1 domains light blue (IGF-1R) and dark blue (IR), α-chain C-terminal regions (αCT′) are coloured light purple (IGF-1R) and dark purple (IR). The side chains of residues involved in contact with the FFY/FYF motif and residue YB16/Q18 are shown and annotated.

We are currently in the process of investigating the effect of these site 1a minimised analogues on downstream signaling. For analogues with the most significant loss of binding we expect there to be essentially no signaling response. Here I showed that F26A IGF-II retains high affinity for the IGF-1R but has low affinity for the IR-A. I propose that study of the signaling in response to F26A IGF-II may provide some interesting insight into the role of this residue in signaling responses. Interestingly mutation of the F28 in the FYF motif to leucine resulted in an equal reduction in affinity for the IGF-1R and IR-A but a >10 fold decrease in potency of IGF-1R activation and very potency of IR-A activation (Alvino et al., 2009). These results implicate the FYF motif in influencing potency of IR-A activation. Therefore, it would be interesting to determine if the F26A IGF-II analogue produced in this thesis also conferred altered signaling responses.

In future investigations we aim to further investigate the F26A IGF-II analogue, which has an overall 10-fold reduction in affinity for the IR-A, to determine if the interaction at site 1b affinity is stronger compared to native IGF-II. Furthermore, we aim to determine the effect of site 1a perturbation on signaling outcome.

7.2.2 Q18 IGF-II analogues

The interaction between YB16 and IR F39 is important for high affinity IR-A binding. This study demonstrated that a similar high affinity interaction, recapitulated by introduction of tyrosine at residue 18 in IGF-II, increased the affinity of IGF-II to equal that of insulin but surprisingly this does not impact on downstream signaling. Based on the results presented in this thesis and other investigations several additional questions have been raised that we aim to further investigate in future experiments:

1. What involvement does the L1 domain have on signaling outcome?
2. Is mitogenic signaling potency determined by the site 1b strength of interaction?
3. Is dissociation rate/ internalization affected by increased affinity to the L1 domain?

In the high affinity interaction of IGF-II at site 1a residue F26 anchors the ligand to the L1 domain in an equivalent manner to FB24 of insulin. Further strengthening the affinity to the IR L1 domain as demonstrated by Q18Y IGF-II appears to have little effect on signaling outcome (magnitude or bias) despite increasing IR binding affinity. Given this, we propose that signaling outcome is determined elsewhere on the receptor. Site 1b has already been implicated in determining signaling outcome. In IGF-II, mutation of site 1b

residue F48 to a histidine increases binding affinity for the IGF-1R and IR-A 261% and 537% respectively compared to native IGF-II. Signaling potency was found to proportionally increase in conjunction with their reported binding affinities (Machackova et al., 2018). Another example is seen in insulin where mutation of residue HB10 to a glutamine results in 382% increase in affinity and 257% increase in mitogenic potency compared to human insulin (Glendorf et al., 2012).

A study published by Gauguin et al (2008b) explored several mutations of IGF-II including a Q18Y mutation in the context of multiple changes (T7H, T16A, Q18Y, F48T, S50I, T58N IGF-II). This analogue, termed sextuple IGF-II, was shown to only have a minor increase in affinity for the IR-A compared to native IGF-II. However, this analogue displayed equal potency of IR-A activation as human insulin, higher than that of native IGF-II. We have shown in this thesis that an individual mutation of Q18 to a tyrosine results in no change in IR-A activation compared to native IGF-II. This led us to speculate that the increase in signaling potency of sextuple IGF-II is achieved through the other mutations made in this analogue. Mutation of insulin residue AB14 (equivalent to T16 of IGF-II) to threonine increased IR-A binding affinity ~2 fold (Gauguin et al., 2008b). As an alanine substitution at IGF-II residue 16 has not been made it is difficult to predict if this substitution contributes to the slight increase in IR-A binding affinity in context of sextuple IGF-II. As shown in Figure 7.2 residues F48, T7 and possibly S50 form part of the site 1b interaction. Mutation of IGF-II residues F48, R49, S50 to those found in insulin (T, S, I, respectively) had only a minor effect on IGF-1R and IR binding (Sakano et al., 1991), suggesting the F48T and S50I changes in sextuple IGF-II may contribute to the slight increase in IR-A binding. Interestingly, mutation of IGF-II F48 to histidine led to a 2.6-fold increase in affinity for the IGF-1R and a 5.4-fold increase in binding affinity for the IR-A compared to native IGF-II (Machackova et al., 2018). Mutation of the equivalent insulin residue TA8 to a histidine leads to a ~3-fold increase in affinity for the IR (Kaarsholm et al., 1993). There are no known mutations of IGF-II residue T7 in literature. However, the equivalent residue in insulin HB5 when mutated to threonine (as found in IGF-II) results in a 4.24-fold decrease in affinity for the IR (Gauguin et al., 2008b). In IGF-I residue T4 (equivalent to residue T7 in IGF-II) when mutated to a histidine (as found in insulin) had little effect on IGF-1R binding but ~7-fold higher IR-A binding affinity (Shooter et al., 1996, Denley et al., 2005). While there are no reported single analogues of IGF-II residue S50, mutation of the equivalent insulin residue IA10 to serine results in a ~5-fold reduction in IR binding affinity (Gauguin et al., 2008b). Taken together this information would suggest the mutations introduced into sextuple IGF-II at site 1 b may contribute to

the slightly greater binding affinity for the IR-A. Residue T58 contributes to site 1a binding as it contacts the αCT′ domain (Xu et al., 2020) (Figure 1.5C). Mutation of NA18 to a threonine (as found in the equivalent position in IGF-II) leads to a 2.3-fold increase in affinity for the IR (Palsgaard, 2003). It would be informative to individually mutate these residues in IGF-II (T7 to histidine, S50 to isoleucine and T58 to asparagine) to determine if they are responsible for sextuple IGF-IIs increased potency of IR-A activation.

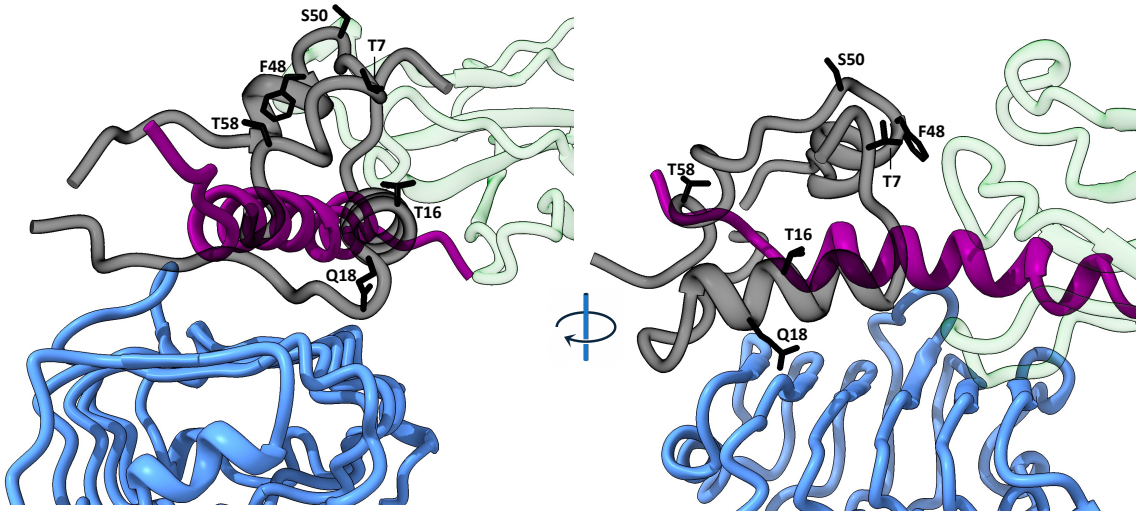


Figure 7.2: IGF-II interacting with site 1a and site 1b.

IGF-II (black) bound to the IGF-1R (from PDB: 6VWI). Only the L1 domain (light blue), αCT′ (purple) and FnIII-1′ domains are shown (green). The side chains of T7, T16, Q18, I50, T58 and F48 are shown in black. In Sextuple IGF-II these residues were mutated (T7H, T16A, Q18Y, S50I, T58N, F48T). Residue Q18 of IGF-II interacts with the L1 domain, forming a part of the site 1a interaction. Residue T58 interacts with the αCT′, also forming part of the site 1a interaction. Residue T7 and F48 of IGF-II interacts with the FnIII-1′ domain, forming part of the site 1b interaction. Residues T16 and S50 have not been implicated in receptor binding.

Our next step would be to investigate the effect of strengthening and weakening the site 1b interaction in combination with the Q18Y mutation and determine the effect on signalling outcome, dissociation rate and receptor internalisation. The following analogues presented in Table 7.1 would help answer the above research questions. Site 1b affinity could be measured using isothermal titration calorimetry assays (ITC) as described by Li et al (2022) using only the FnIII-1 domain. The affinities for site 1a would also be best measured by ITC as previously described, using the L1-CR-L2/αCT (Li et al., 2022, Xiong et al., 2020, Menting et al., 2009). This technique was utilised by our collaborators, at the Walter and Eliza Hall Institute of Medical Research (WEIHI), to measure the affinity of insulin analogues for IR site 1a using the first three domains of the receptor (L1-CR L2) and the αCT peptide (Menting et al., 2009).

In future investigations we aim to further interrogate the role of the L1 domain of the IGF-1R and IR by producing L1 domain mutants to determine the effect on binding and signalling. Previous studies reported domain swap mutants of the IR and IGF-1R (Schumacher et al., 1993). However, these were limited by the lack of structural information and the role of L1 domain in ligand binding specificity remained unclear. Since this publication structural data has provided further detail on the similarities and differences between the IGF-1R and IR L1 domains. Using this structural information to identify major points of difference between the IR and IGF-1R L1 domains, future investigations could interrogate the contribution of specific residues in the L1 domains that are unique to the IR or IGF-1R. As described above ITC assays using the first three domains of the receptor (L1-CR L2) and the αCT peptide have been routinely used to measure the affinity of ligands for IR site 1a (Menting et al., 2009).

Our characterisation of each analogue in this thesis involved circular dichroism analysis to determine secondary structure. While this technique is a very useful tool, our next step for these analogues is to determine how the introduced side chains are positioned within the structure of the protein. We will approach these structural investigations in a stepwise manner. We have recently begun molecular dynamics simulation (in collaboration with researchers from Latrobe University), of these analogues. Following this we will then perform modelling of these analogues in complex with site 1a and 1b of the IGF-1R. With the first known structure of a single IGF-II ligand in complex with the IGF-1R using cryoEM we can utilise this solved structure to model the analogues, produced in this thesis, at site 1a and 1b. If unique binding properties are observed at site 1a or site 1b using molecular dynamic simulations then we would then progress to structural analyses

using cryoEM to confirm these findings.

Table 7.1: Proposed IGF-II analogues for future investigation based on the findings presented in this thesis.

Proposed IGF-II analogues for future investigation. Based on the results presented in this thesis and the work of others we hypothesise the following changes to receptor binding affinity, ligand dissociation rate, receptor internalization and signalling potency, compared to native IGF-II.

Analogue	Site 1a affinity	Site 1b affinity	Dissociation rate	Rate of IR internalization	Signalling potency
F26A IGF-II	Weakened	No change	Increased	Decreased	Proportional decrease in Akt and ERK due to drop in affinity
F26A F48H IGF-II	Weakened	Strengthened	Increased	Decreased	Decreased
Q18Y F48H IGF-II	Strengthened	Strengthened	Decreased	Increased	Increase in mitogenic
Q18Y E12A IGF-II	Strengthened	Weakened	Increased	Decreased	Decreased

Currently, there are no solved structures of IGF-II bound to the IR-A, despite its high affinity binding to this receptor and the implication of both IGF-II and the IGF-1R in cancer progression. We therefore aim to obtain structures of IGF-II, and IGF-II analogues produced in this thesis, bound to the IR-A as well as the IGF-1R. Structural detail of IGF-II bound to the IR-A would be a valuable tool for comparing to the recently solved structure of IGF-II bound to the IGF-1R (Xu et al., 2020). These structural comparisons may be able to tease out how preferential signalling outcomes are achieved.

7.2.3 Effect of single mutations on IGF-II folding and stability

Until now there has been little detail regarding IGF-II structural stability. My investigation into single mutations of IGF-II, based on cone snail venom insulin binding mechanisms, led to the understanding of several key residues that are important for IGF-II structure and stability. In Chapter 4, two residues which have been implicated in Con-Ins G1s alternate binding mechanism were explored. We demonstrated for the first time in IGF-II, that the residues L17 and G22 have important structural roles in human IGF-II. Mutation of residue L17 to alanine drastically affects the structure of the protein and reduces binding affinity to both the IGF-1R and IR. The same was the case for L17Y mutation. While G22A IGF-II had essentially no change in affinity for the IGF-1R and IR-A more radical mutations at this position, such as introduction of an aromatic tyrosine or a leucine, disrupted the native like structure of the protein and consequently reduced binding affinity for both the IGF-1R and IR-A.

In Chapter 4, D-domain modifications in IGF-II were explored by introduction of a sequence equivalent to A-chain extension observed in Con-Ins K1 and modified A-chain extension in Vh-Ins. The solved structure of Vh-Ins revealed that the modified A-chain extension (HA21 SA22, LA23 QA24) extended the C-terminal A-chain helix (Xiong et al., 2021). Here the IGF-II D-domain was replaced with an A, S, L, Q motif (A61, S62 L63, Q64 IGF-II), with the equivalent HA21 residue being kept as an alanine to maintain the local structure of the neighbouring C60 residue. Circular dichroism analysis indicated that this modified D-domain extends the C-terminal A-domain helix, although structural analysis will need to be performed to confirm this. This modified D-domain had a minor effect on IGF-1R binding but more drastically reduced affinity for the IR-A (Chapter 4). Applying similar modifications into a DOI analogue (lacking the C-terminal 8 amino acids of the B-chain) resulted in essentially the same affinity as human insulin for the IR and

IGF-1R (Xiong et al., 2021).

A major aspect of this PhD thesis has involved recombinant protein expression and chemical refolding IGF-II refolding. The effect of mutations on protein structure was explored. In many cases the IGF-II analogues had low helical content, which also correlated to poor refolding efficiency. Indeed, we also found two analogues that had increased in helical content and increased folding efficiency. These included G22A IGF-II and ASLQ IGF-II. While ASLQ IGF-II had a major loss of peptide at the steps between dissolution and gel filtration there was essentially no loss of peptide between gel filtration and the completion of the chemical refold (Figure 4.11). Even in the case of native IGF-II, in which this chemical refolding method has previously been optimised, we see major reduction in peptide between gel filtration and the completion of chemical refolding (Figure 5.11). For both analogues (G22A IGF-II and ASLQ IGF-II) it appears that the increase in helical content results in increased folding efficiency. However, for both analogues there was negligible effect on binding affinity, although for ASLQ IGF-II there was a reduction in affinity for the IR-A (discussed below). The equivalent mutation to G22A IGF-II in insulin (GB20A) also led to an increase in helical content and it was confirmed that the B-chain helix was extended by a turn (Kristensen et al., 1997). Interestingly, it was found that this insulin analogue actually had a significant decrease in yield using a yeast expression system (Kristensen et al., 1997). Our findings, in this thesis, that increases in the helical content of IGF-II analogues leads to increased efficiency of refolding may be useful in further exploring, especially in the context of analogues that are known to have poor folding efficiency such as the combination analogues summarised below.

7.2.4 Site 1a minimised IGF-II analogues and cone snail venom insulin alternate binding mechanisms

The final stage of my project involved applying the Con-Ins G1 alternate binding residues and Con-Ins K1 alternate binding residues to the site 1a minimised IGF-II analogues (Chapter 3). A significant challenge of this stage was the folding of these combined analogues. Attempts to apply the Con-Ins G1 residues into the F26G Y27G F28G IGF-II analogue all resulted in markedly reduced secondary structure compared to native IGF-II. We concluded that the FYF motif in native IGF-II is essential for stabilising the flexible C-domain by engaging with the hydrophobic core of the protein (As shown in Figure 6.10). The loss of these hydrophobic elements in IGF-II analogues, particularly in F26G Y27G F28G mutations, likely resulted in further instability in this region of the ligand.

Indeed, an investigation by Rege et al (2020b) revealed that residue FB24 of the aromatic FFY motif in human insulin is essential for nascent folding events. It was also found that each of the 19 possible mutations at position B24 in human insulin led to impaired biosynthesis, using HEK293T cells (Rege et al., 2020b). However, structural analysis of FB24Y insulin, produced using peptide synthesis, was found to have a highly similar structure as human insulin and was biologically active (Hua et al., 1996). In this thesis we used recombinant protein expression systems and chemical refolding techniques to produce each IGF-II analogue. We have shown that loss of the aromatic elements of this motif in IGF-II also impede folding.

We have demonstrated that the single chain structure of native IGF-II and the presence of a highly flexible C-domain likely makes production of correctly folded mutants with a perturbed FYF motif challenging. Using a recombinant expression system and chemical refolding techniques we were unable to produce a native like folded IGF-II analogue. The cone snail insulins and DOI derivatives incorporating cone snail changes can be successfully generated using chemical synthesis and by combining the two separate chains, with the use of protective groups and regioselective deprotection to form disulfide bonds in a controlled manner (Xiong et al., 2021, Xiong et al., 2020). It is interesting to consider that *in vivo* the cone snail venom insulins are likely produced as a single chained propeptide (Safavi-Hemami et al., 2015) that is protected from aggregation and is folded and processed into the mature two-chained insulin with the assistance of molecular chaperons. Interestingly, several chaperone proteins in cone snail venom glands including protein-disulfide isomerase (PDI), peptidyl-prolyl *cis-trans* isomerase and immunoglobulin-binding-protein (BiP) have been shown to accelerate the oxidative folding, conformation and function of proteins in cone snail venom (Safavi-Hemami et al., 2012). We would predict that if expressed recombinantly cone snail insulin propeptides would be unstable and difficult to fold, as we saw for F26G Y27G F28G IGF-II and the other mutants with changes at the FFY motif. Conversely, it would therefore be interesting to test the ability of the IGF-II analogues explored in this thesis to fold in the presence of the chaperone proteins described above.

We have demonstrated in this thesis that the FYF motif in native IGF-II likely clips down the flexible C-domain and C-terminal segment of the B-domain. Removal of elements of the FYF motif leads to further flexibility of the C-domain and the C-terminal segment of the B-domain and impacts the ability of these analogues to fold into a native like IGF-II structure. In mammalian cells mature human insulin is derived from a single chain

precursor termed proinsulin that includes a C-domain (Dodson and Steiner, 1998). The C-domain connects the B- and A-chains together and is 35 amino acids in length (IGF-II C-domain is 8 amino acids in length). This C-domain in proinsulin is required for early folding events in pancreatic β-cells and has proinsulin reduced fibrillation compared to mature human insulin (Weiss, 2009). Proinsulin has approximately 5-fold lower affinity for the IR-A compared to mature human insulin (Malaguarnera et al., 2012). Interestingly, several studies have explored shortened C-domains in insulin and have reported several analogues with equal affinity as mature human insulin (Glidden et al., 2018a, Glidden et al., 2018b, Hua et al., 2008, Smith et al., 2022b). Structural analysis of single chain insulins has found that the C-domain also threads through the αCT (Smith et al., 2022a) much like the IGFs. Molecular dynamics simulations of varying C-domain lengths in single chain insulins revealed that short C-domains will likely impair the ability to thread through the αCT (Smith et al., 2022a). These simulations are consistent with the findings that a 2 amino acid linker between the B- and A-domains in human insulin results in an affinity of 0.5% for the IR compared to mature human insulin (Kobayashi et al., 1989). It would be interesting to explore variations of the C-domain length in IGF-II in the context of the combination analogues described above, to determine if shortened C-domain length in these analogues stabilises the structure.

While during my PhD I was unable to restore native IGF-II like affinity via an alternative mechanism than the classic site 1a binding mechanism, I believe that I have made a significant contribution to the understanding of IGF-II structure and function through the investigations presented here. In particular, I have made significant progress in the understanding of IGF-II structure, stability, receptor binding specificity.

7.3 Future directions

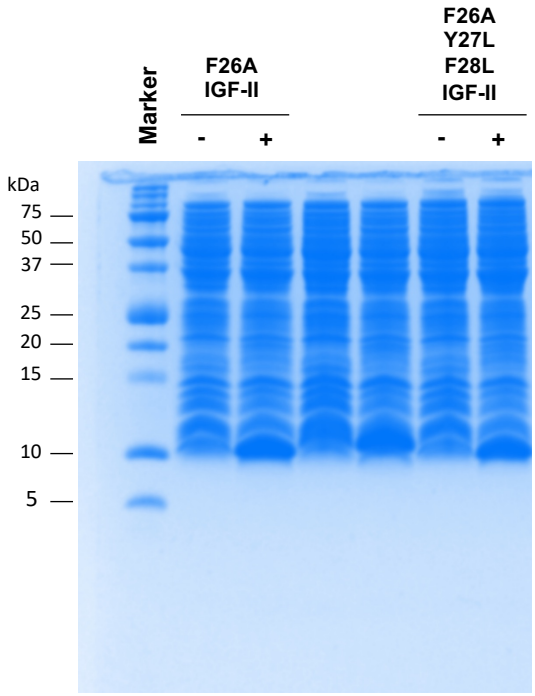
Understanding of the molecular detail of insulin like growth factor system has progressed enormously in recent years. There are a multitude of structures of insulin bound to the insulin receptor that each reveal additional detail of the mechanism of receptor binding (Lawrence, 2021). These structures have raised debate and speculation around an additional site of interaction between insulin and the insulin receptor. While we have had mutagenesis data to suggest that this interaction occurs, only recent insulin:IR structures obtained using cryogenic electron microscopy (cryoEM), have captured this site of interaction. We require a more detailed understanding of this site of interaction including the order in which the binding events occur. While this second site interaction has been

identified by mutagenesis studies as occurring in the IGFs there are currently no solved structures of either IGF ligand at this site on the IGF-1R. Further investigation is required to determine if this site 2 interaction does in fact occur for the IGFs binding to the IGF-1R or if this represents a major difference in IR and IGF-1R binding events.

Future research efforts in the field of IGF research will also involve better understanding how receptor binding leads to signalling activation. In the literate and presented in this thesis we have many examples of insulin and insulin like growth factor analogues with unique signalling properties. However, we still do not fully understand how altered binding leads to altered signalling outcome. In recent solved structures of ligand bound to the IR and IGF-1R there has been speculation that the distance between the receptor legs in the activated state may correlate to signalling bias.

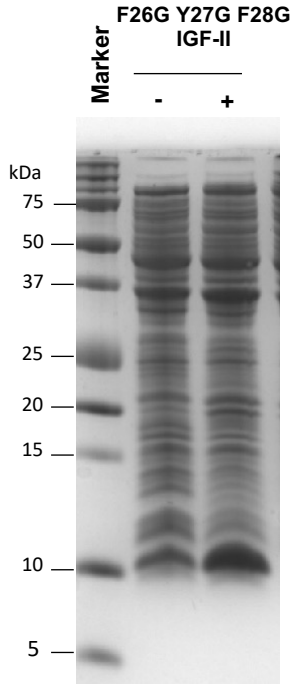
Finally, a major future area for the advancement of our understanding of IGF-II function is to obtain solved structures of IGF-II bound to the IR-A. IGF-II is the only ligand that is capable of binding to the IR-A and IGF-1R with high affinity. Having molecular detail of IGF-II bound to each receptor will provide a detailed understanding of how high affinity binding for each receptor is achieved. This knowledge will be valuable for the development of drug therapies targeted toward a single receptor.

8 SUPPLEMENTARY MATERIAL



Supplementary Figure 1: Sodium dodecyl-sulfate polyacrylamide gel electrophoresis (SDS-PAGE) of induced and uninduced whole cell lysates during *E. coli fermentation.*

Coomassie-stained reducing 15% tricine SDS-PAGE of F26A IGF-II and F26A Y27L F28L IGF-II protein expression in *E. coli* BL21 cells. Non induced (-) and Isopropyl β-d-1-thiogalactopyranoside (IPTG) induced (+) whole cell lysates.



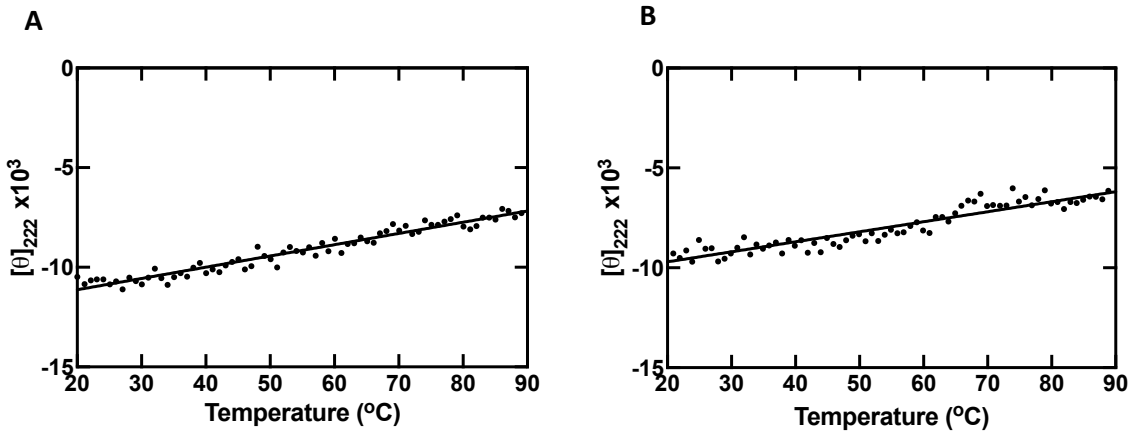
Supplementary Figure 2: Sodium dodecyl-sulfate polyacrylamide gel electrophoresis (SDS-PAGE) of induced and uninduced whole cell lysates during *E. coli fermentation.*

Coomassie-stained reducing 15% tricine SDS-PAGE of F26G Y27G F28G IGF-II protein expression in *E. coli* BL21 cells. Non induced (-) and Isopropyl β-d-1-thiogalactopyranoside (IPTG) induced (+) whole cell lysates.

Supplementary Table 1: Matrix assisted laser desorption ionisation (MALDI) mass spectrometry of intact IGF-II analogues.

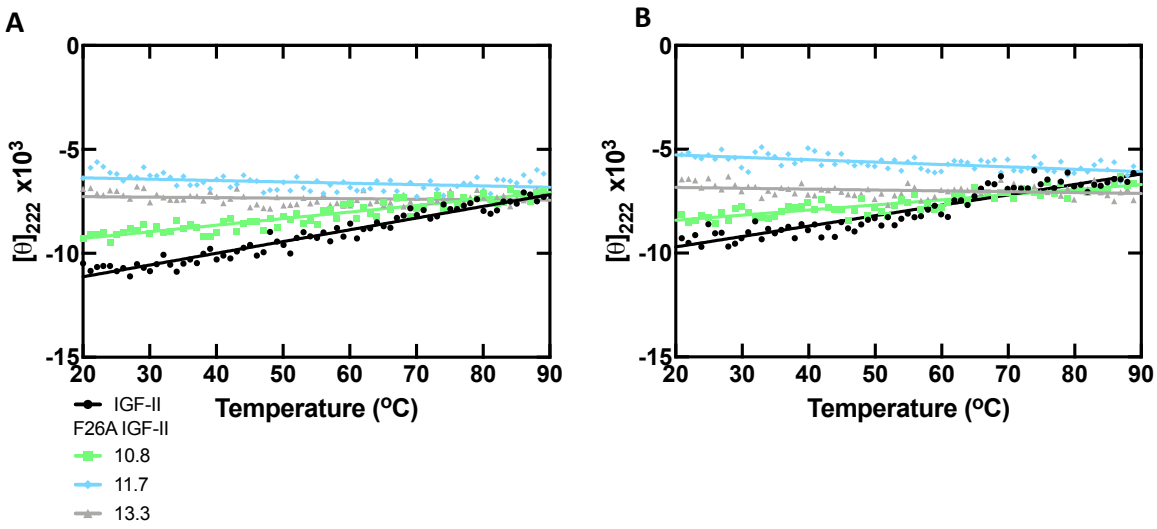
Confirmation of final folded mass of intact IGF-II analogues were determined using MALDI mass spectrometry (Section 2.3.13).

Analogue	HPLC retention time (min)	Average mass (Daltons)	
		Theoretical	Observed
F26A IGF-II	10.8	7393.344	7393.577
	11.7		7393.683
	13.3		7393.531
F26A Y27L F28L IGF-II	11.4	7309.310	7309.434
F26G Y27G F28G IGF-II	7.1	7183.069	6700.386
	8.2		7712.729
	9.6		7183.263



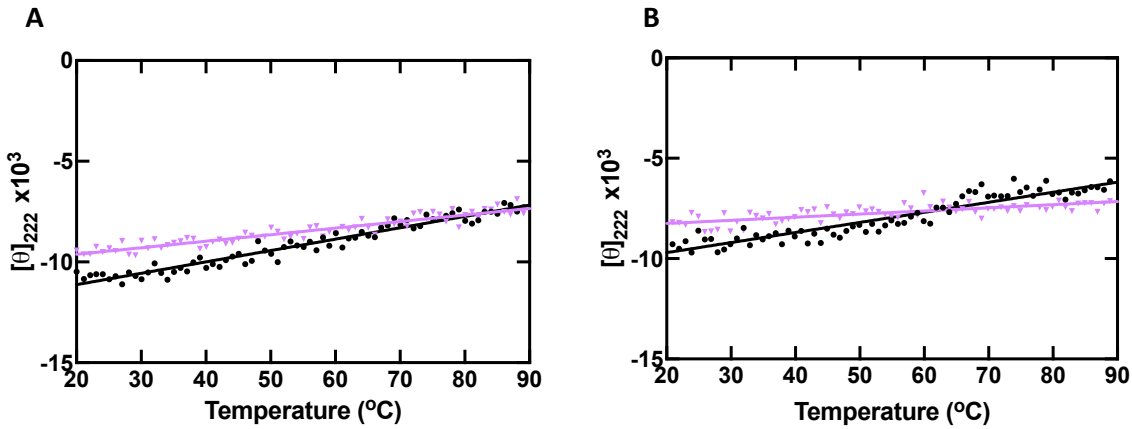
Supplementary Figure 3: Temperature denaturation of human IGF-II.

Temperature denaturation studies of human IGF-II were performed in 10 mM acetic acid at a concentration of 0.05 μg/μL. Mean residue ellipticity [θ] at 222 nm was measured from 20°C to 90°C (at 1±0.1°C increments) (A), followed by measurement from 90°C to 20°C (at 1±0.1°C increments) (B).



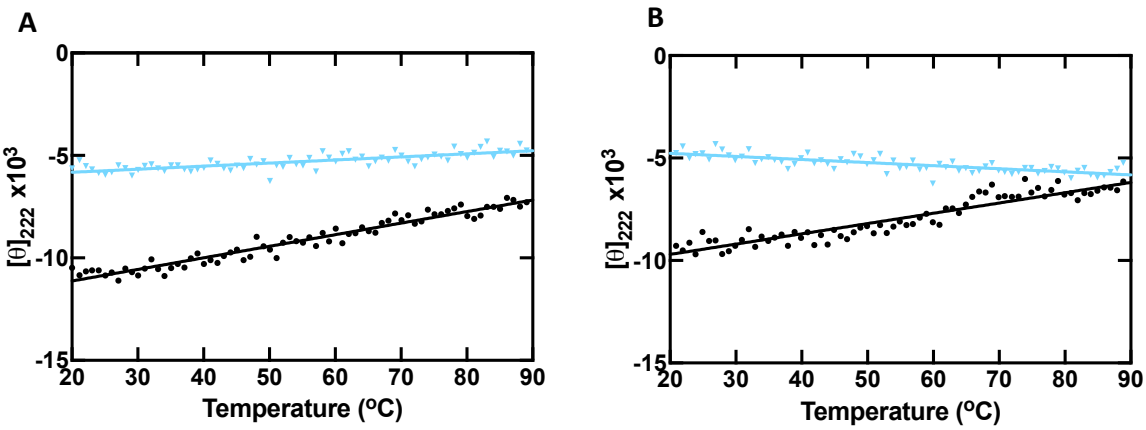
Supplementary Figure 4: Temperature denaturation of F26A IGF-II and human IGF-II.

Temperature denaturation studies of F26A IGF-II purified fractions and human IGF-II were performed in 10 mM acetic acid at a concentration of 0.05 μg/μL. Mean residue ellipticity [θ] at 222 nm was measured from 20°C to 90°C (at 1±0.1°C increments) (A). Mean residue ellipticity [θ] at 222 nm was also measured from 90°C to 20°C (at 1±0.1°C increments) (B).



Supplementary Figure 5: Temperature denaturation of F26A Y27L F28L IGF-II and human IGF-II.

Temperature denaturation studies of F26A Y27L F28L IGF-II (purple) and human IGF-II (black) were performed in 10 mM acetic acid at a concentration of 0.05 μg/μL. Mean residue ellipticity [θ] at 222 nm was measured from 20°C to 90°C (at 1±0.1°C increments) (A). Mean residue ellipticity [θ] at 222 nm was also measured from 90°C to 20°C (at 1±0.1°C increments) (B).



Supplementary Figure 6: Temperature denaturation of F26G Y27G F28G IGF-II and human IGF-II.

Temperature denaturation studies of F26G Y27G F28G IGF-II (blue) and human IGF-II (black) were performed in 10 mM acetic acid at a concentration of 0.05 μg/μL. Mean residue ellipticity [θ] at 222 nm was measured from 20°C to 90°C (at 1±0.1°C increments) (A). Mean residue ellipticity [θ] at 222 nm was also measured from 90°C to 20°C (at 1±0.1°C increments) (B).

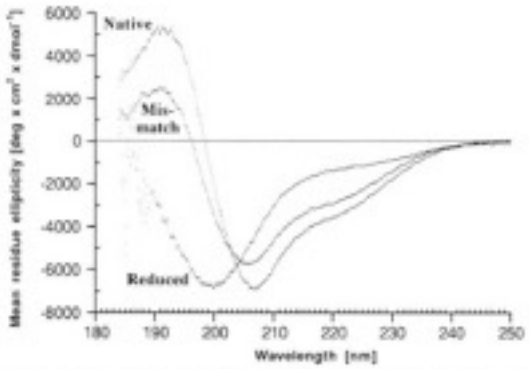
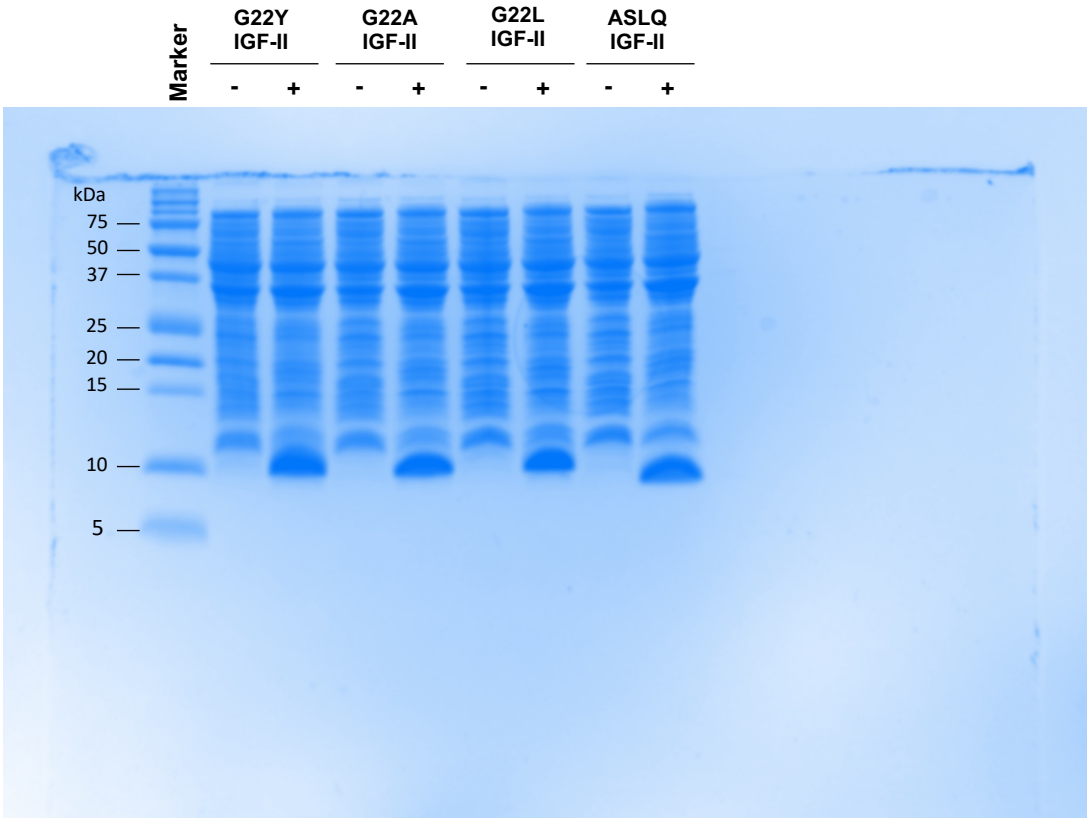


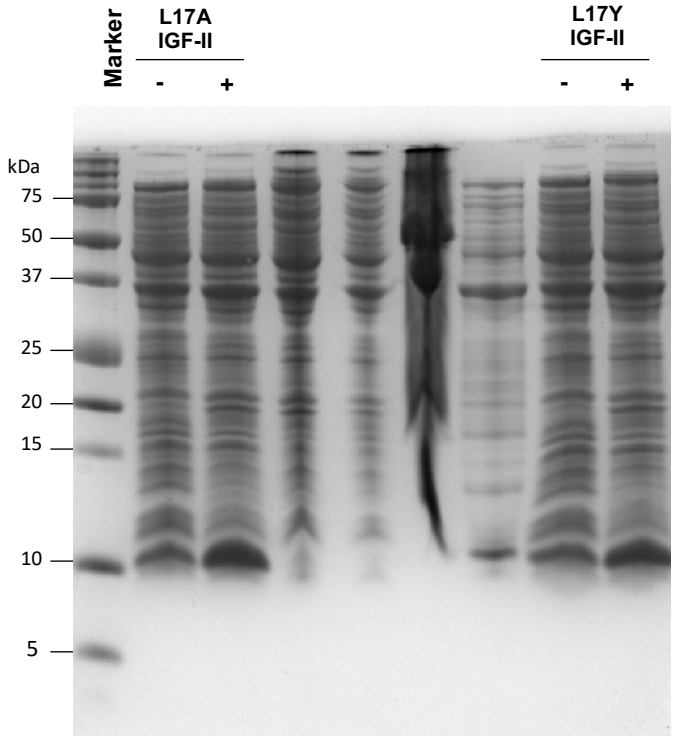
FIGURE 3: Far-UV circular dichroism spectra of reduced (Reduced), mismatched (Mis-match), and native (Native) IGF-I from 184 to 250 nm as described under Materials and Methods.

Supplementary Figure 7. Excerpt of Figure 3 from Hober et al (1991) of the CD spectra of reduced and mis-matched IGF-I compared to native IGF-I.



Supplementary Figure 8: Sodium dodecyl-sulfate polyacrylamide gel electrophoresis (SDS-PAGE) of induced and uninduced whole cell lysates during *E. coli fermentation.*

Coomassie-stained reducing 15% tricine SDS-PAGE of G22Y IGF-II, G22A IGF-II, G22L IGF-II and ASLQ IGF-II protein expression in *E. coli* BL21 cells. Non induced (-) and Isopropyl β-d-1-thiogalactopyranoside (IPTG) induced (+) whole cell lysates.



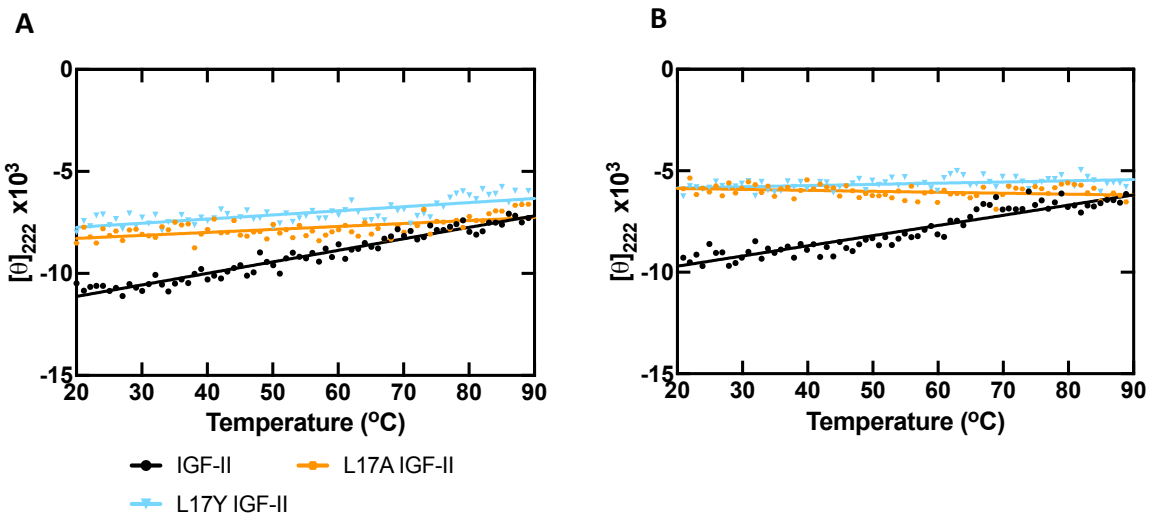
Supplementary Figure 9: Sodium dodecyl-sulfate polyacrylamide gel electrophoresis (SDS-PAGE) of induced and uninduced whole cell lysates during *E. coli fermentation.*

Coomassie-stained reducing 15% tricine SDS-PAGE of L17A IGF-II and L17Y IGF-II protein expression in *E. coli* BL21 cells. Non induced (-) and Isopropyl β-d-1-thiogalactopyranoside (IPTG) induced (+) whole cell lysates.

Supplementary Table 2: Matrix assisted laser desorption ionisation (MALDI) mass spectrometry of intact IGF-II analogues.

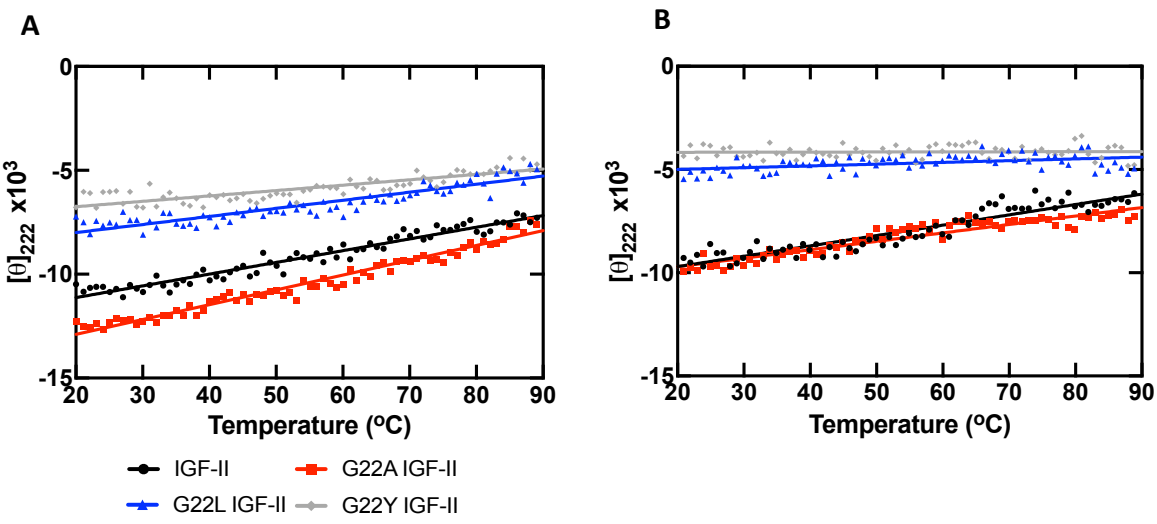
Confirmation of final folded mass of intact IGF-II analogues were determined using MALDI mass spectrometry (section 2.3.13).

Analogue	HPLC retention time (min)	Average mass (Daltons)	
		Theoretical	Observed
IGF-II	10.5	7469.442	7469.632
G22A IGF-II	11.0	7483.469	7484.396
G22L IGF-II	12.0	7525.549	7523.714
G22Y IGF-II	11.2	7575.566	7574.980
L17A IGF-II	9.8	7427.361	7427.485
L17Y IGF-II	9.8	7519.458	7517.106



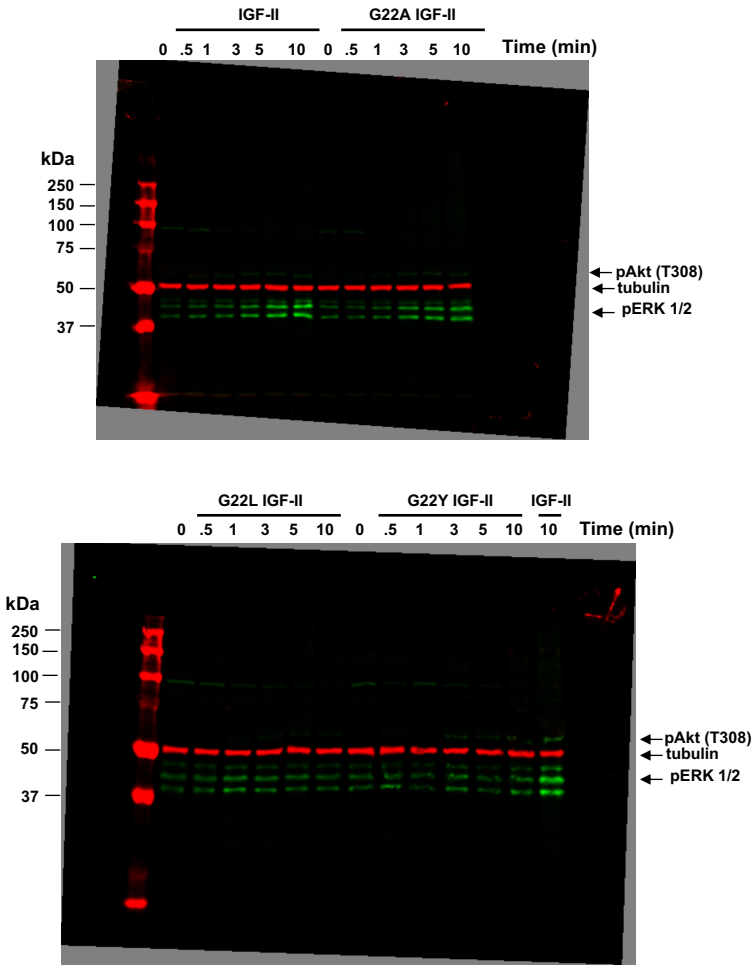
Supplementary Figure 10: Temperature denaturation of L17A IGF-II, L17Y IGF-II and human IGF-II.

Temperature denaturation studies of L17A IGF-II, L17Y IGF-II and human IGF-II were performed in 10 mM acetic acid at a concentration of 0.05 μg/μL. Mean residue ellipticity [θ] at 222 nm was measured from 20°C to 90°C (at 1±0.1°C increments) (A). Mean residue ellipticity [θ] at 222 nm was also measured from 90°C to 20°C (at 1±0.1°C increments) (B).

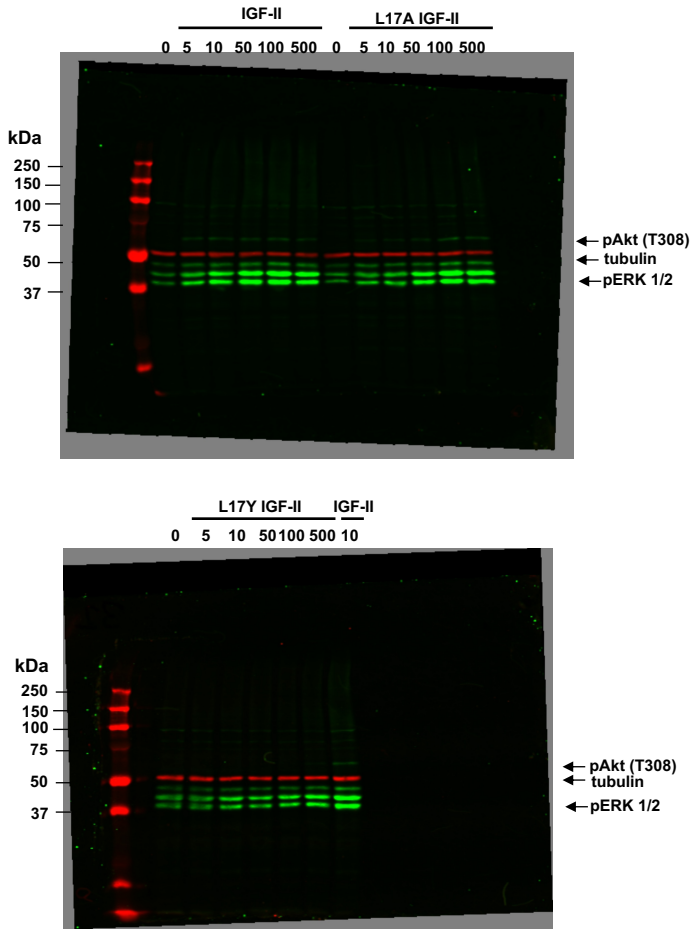


Supplementary Figure 11: Temperature denaturation of G22A IGF-II, G22L IGF-II, G22Y IGF-II and human IGF-II.

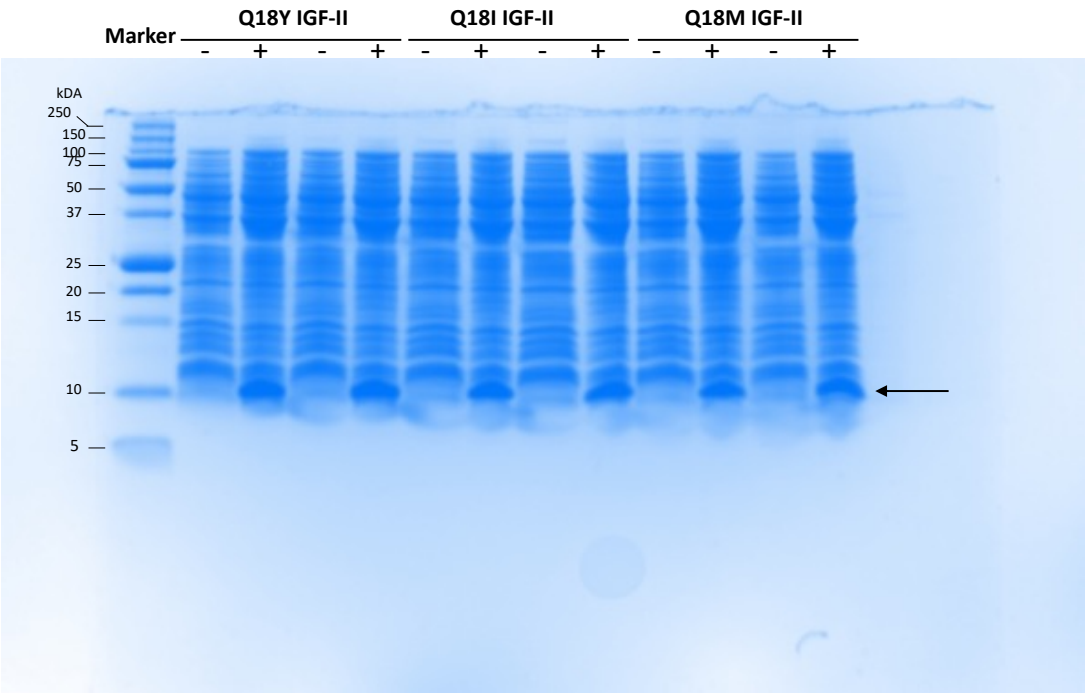
Temperature denaturation studies of G22A IGF-II, G22L IGF-II, G22Y IGF-II and human IGF-II were performed in 10 mM acetic acid at a concentration of 0.05 μg/μL. Mean residue ellipticity [θ] at 222 nm was measured from 20°C to 90°C (at 1±0.1°C increments) (A). Mean residue ellipticity [θ] at 222 nm was also measured from 90°C to 20°C (at 1±0.1°C increments) (B).



Supplementary Figure 12: Uncropped Western Blot of Akt/ ERK phosphorylation upon IR-A activation by IGF-II and G22 IGF-II analogues (shown in Figure 4.20).



Supplementary Figure 13: Uncropped Western Blot of Akt/ ERK phosphorylation upon IR-A activation by IGF-II and L17 IGF-II analogues (shown in Figure 4.19).



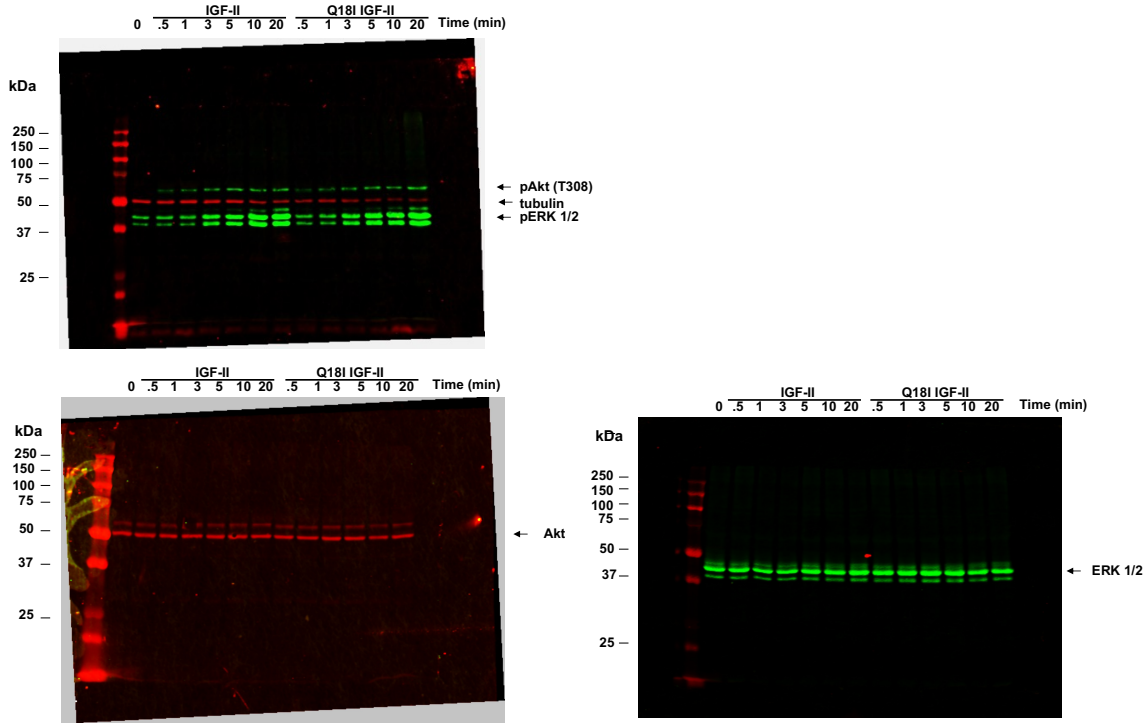
Supplementary Figure 14: Sodium dodecyl-sulfate polyacrylamide gel electrophoresis (SDS-PAGE) of induced and uninduced whole cell lysates during *E. coli fermentation.*

Coomassie-stained reducing 15% tricine SDS-PAGE of Q18I IGF-II, Q18M IGF-II and Q18Y IGF-II protein expression in *E. coli* BL21 cells. Non induced (-) and Isopropyl β-d-1-thiogalactopyranoside (IPTG) induced (+) whole cell lysates.

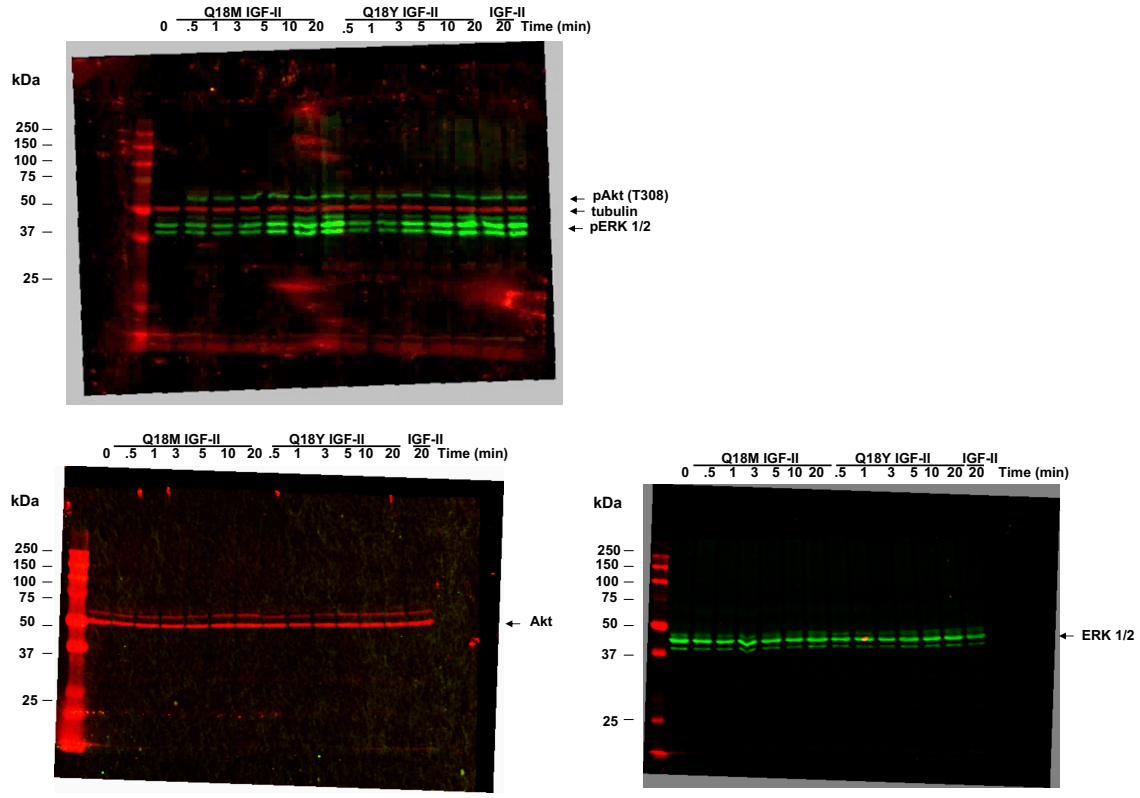
Supplementary Table 3: Matrix assisted laser desorption ionisation (MALDI) mass spectrometry of intact IGF-II analogues.

Confirmation of final folded mass of intact IGF-II analogues were determined using MALDI mass spectrometry (section 2.3.13).

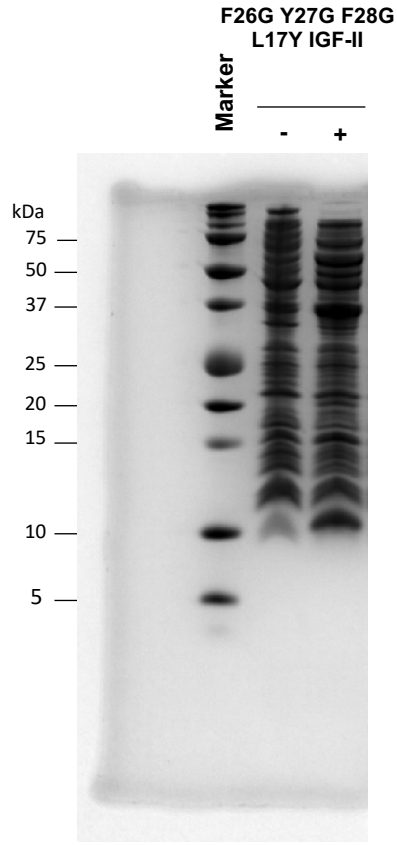
Analogue	Monoisotopic mass (Daltons)	
	Theoretical	Observed
IGF-II	7470.56	7470.53
Q18I IGF-II	7455.58	7456.05
Q18M IGF-II	7473.54	7473.71
Q18Y IGF-II	7505.56	7506.18



Supplementary Figure 15: Uncropped Akt/ ERK activation Western Blot shown in Figure 5.8.

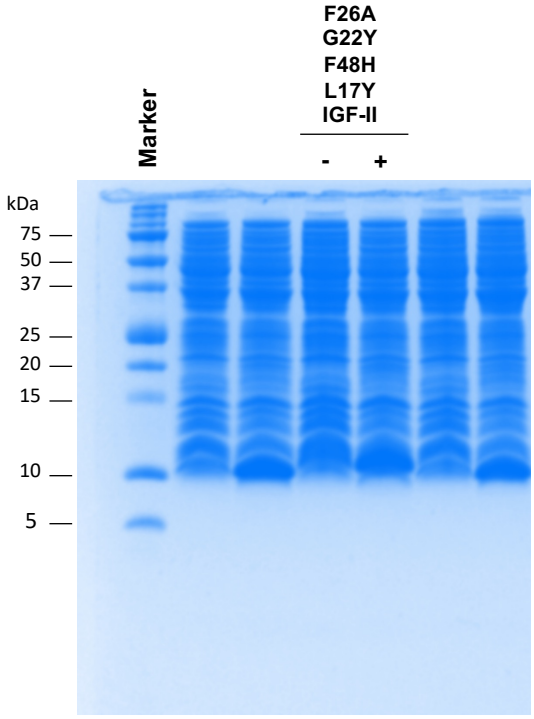


Supplementary Figure 16: Uncropped Akt/ ERK activation Western Blot shown in Figure 5.8.



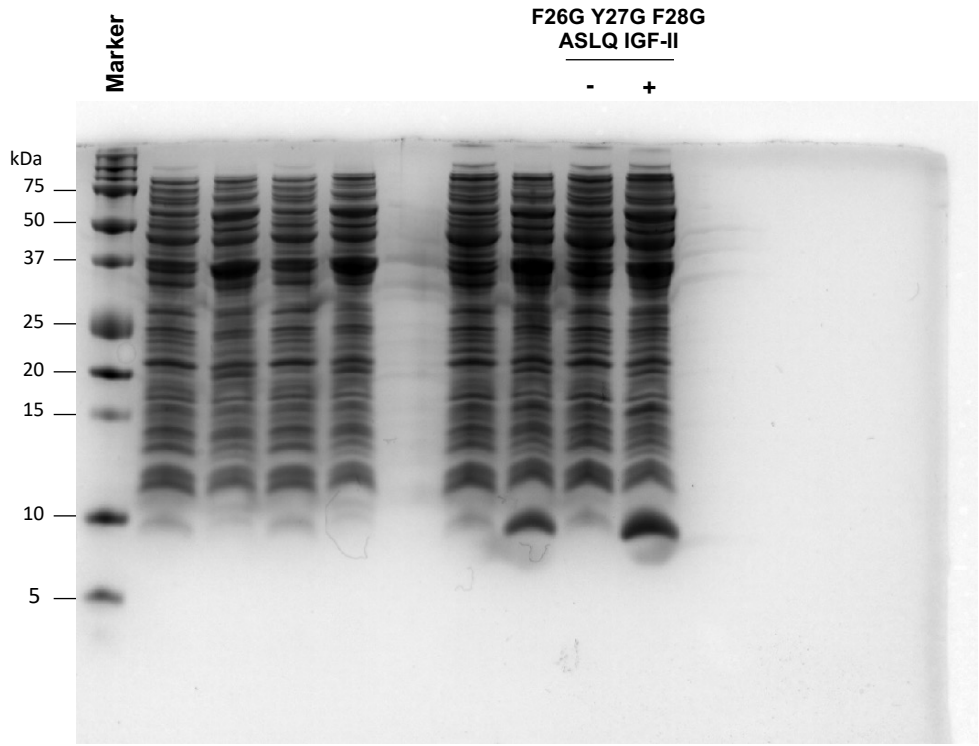
Supplementary Figure 17: Sodium dodecyl-sulfate polyacrylamide gel electrophoresis (SDS-PAGE) of induced and uninduced whole cell lysates during *E. coli fermentation.*

Coomassie-stained reducing 15% tricine SDS-PAGE of F26G Y27G F28G L17Y IGF-II protein expression in *E. coli* BL21 cells. Non induced (-) and Isopropyl β-d-1-thiogalactopyranoside (IPTG) induced (+) whole cell lysates.



Supplementary Figure 18: Sodium dodecyl-sulfate polyacrylamide gel electrophoresis (SDS-PAGE) of induced and uninduced whole cell lysates during *E. coli fermentation.*

Coomassie-stained reducing 15% tricine SDS-PAGE of F26A G22Y F48H L17Y IGF-II protein expression in *E. coli* BL21 cells. Non induced (-) and Isopropyl β-d-1-thiogalactopyranoside (IPTG) induced (+) whole cell lysates.



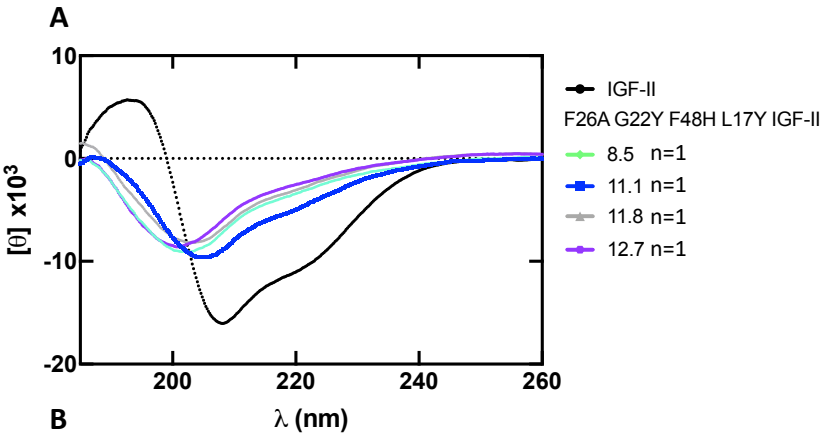
Supplementary Figure 19: Sodium dodecyl-sulfate polyacrylamide gel electrophoresis (SDS-PAGE) of induced and uninduced whole cell lysates during *E. coli fermentation.*

Coomassie-stained reducing 15% tricine SDS-PAGE of F26G Y27G F28G ASLQ IGF-II protein expression in *E. coli* BL21 cells. Non induced (-) and Isopropyl β-d-1-thiogalactopyranoside (IPTG) induced (+) whole cell lysates.

Supplementary Table 4: Matrix assisted laser desorption ionisation (MALDI) mass spectrometry of intact IGF-II analogues.

Confirmation of final folded mass of intact IGF-II analogues were determined using MALDI mass spectrometry (section 2.3.13).

Analogue	HPLC retention time (min)	Average mass (Daltons)	
		Theoretical	Observed
F26A G22Y F48H L17Y IGF-II	8.5	7539.449	7539.917
	11.1		7539.952
	11.8		7539.167
	12.7		7618.298
F26A F48H L17Y IGF-II	7.4	7433.325	7433.799
	9.9		7434.462
	10.4		7434.674
	11.3		7589.318
F26G Y27G F28G L17Y IGF-II	8.9	7233.085	7183.263
F26G Y27G F28G G22Y F48H IGF-II	9.0	7279.157	7280.261
ASLQ IGF-II	12.0	7184.142	7184.458
F26G Y27G F28G ASLQ IGF-II	8.4	6897.769	6926.638
	11.4		6926.638
	12.0		-

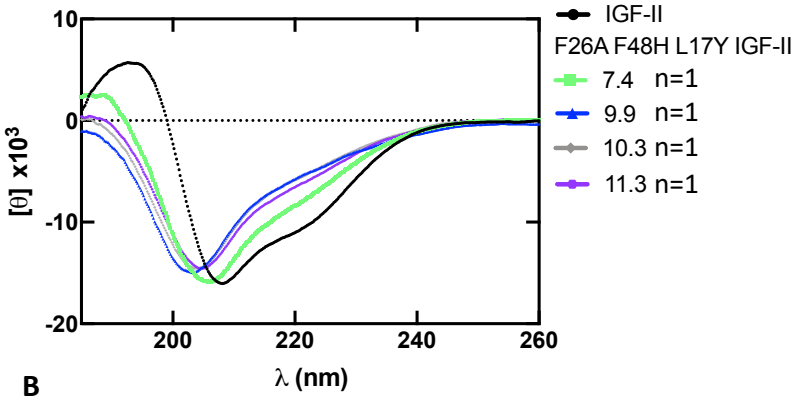


Analogue	Helix	Strand	Turns	Unordered
IGF-II	41	13	16	30
F26A G22Y F48H L17Y IGF-II 8.5 min	8	33	12	46
F26A G22Y F48H L17Y IGF-II 11.1 min	6	35	11	47
F26A G22Y F48H L17Y IGF-II 11.8 min	5	34	9	52
F26A G22Y F48H L17Y IGF-II 12.7 min	5	37	12	46

Supplementary Figure 20: Analysis of F26A G22Y F48H L17Y IGF-II secondary structure.

(A) Secondary structure analysis of F26A G22Y F48H L17Y IGF-II was performed using circular dichroism at 0.05 μg/μL in 10 mM acetic acid, mean residue ellipticity [θ] was measured from 260 - 180 nm. Percent helical content shown in the table was performed using Dichroweb and was calculated using the CDSSTR algorithm for deconvolution against the reference protein database set SMP180 (Sreerama and Woody 2000). (B) Thermal stability was monitored by measuring mean residue ellipticity [θ] at 222 nm from 20°C to 90°C (at 1±0.1°C increments).

A

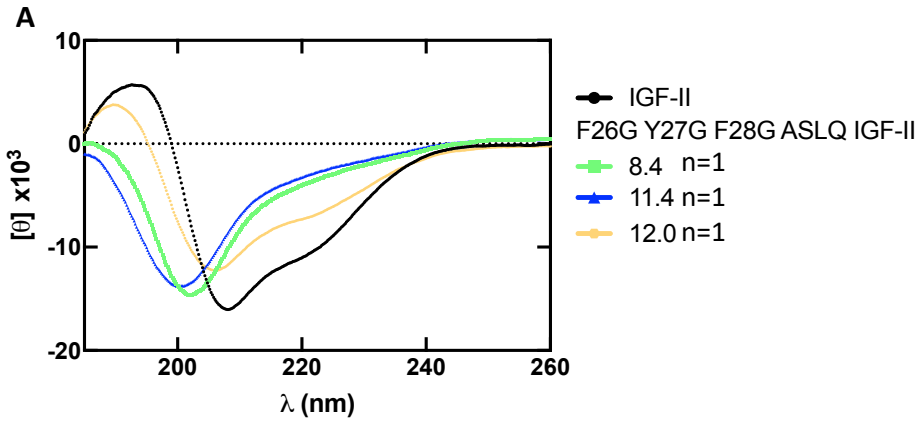


B

Analogue	Helix	Strand	Turns	Unordered
IGF-II	41	13	16	30
F26A F48H L17Y IGF-II 7.4 min	22	20	14	43
F26A F48H L17Y IGF-II 9.9 min	18	21	16	46
F26A F48H L17Y IGF-II 10.3 min	17	23	15	46
F26A F48H L17Y IGF-II 11.7 min	19	22	15	45

Supplementary Figure 21: Analysis of F26A F48H L17Y IGF-II secondary structure and thermal stability using circular dichroism.

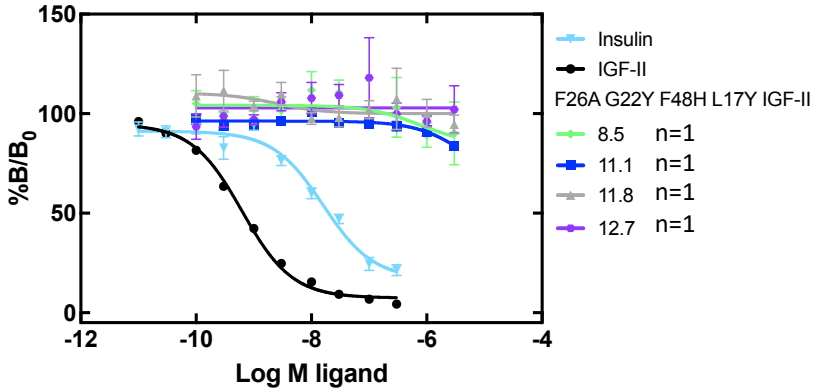
(A) Secondary structure analysis of the F26A F48H L17Y IGF-II was performed using circular dichroism at 0.05 μg/μL in 10 mM acetic acid, mean residue ellipticity [θ] was measured from 260 - 180 nm. Percent helical content shown in the table was performed using Dichroweb and was calculated using the CDSSTR algorithm for deconvolution against the reference protein database set SMP180 (Sreerama and Woody 2000). (B) Thermal stability was monitored by measuring mean residue ellipticity [θ] at 222 nm from 20°C to 90°C (at 1±0.1°C increments). Temperature denaturation data were analysed using a linear regression line of best fit. R^2 values, slope and percentage change in helical content from 20°C to 90°C are shown in the table.



Analogue	Helix	Strand	Turns	Unordered
IGF-II	41	13	16	30
F26G Y27G F28G ASLQ IGF-II 8.4	8	29	11	52
F26G Y27G F28G ASLQ IGF-II 11.4	5	34	13	48
F26G Y27G F28G ASLQ IGF-II 12.0	21	24	14	42

Supplementary Figure 22: Analysis of F26G Y27G F28G ASLQ IGF-II secondary structure and thermal stability using circular dichroism.

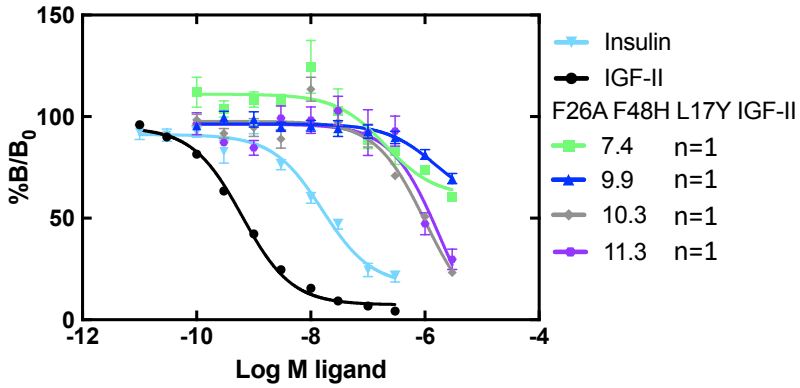
(A) Secondary structure analysis of the F26G Y27G F28G ASLQ IGF-II was performed using circular dichroism at 0.05 μg/μL in 10 mM acetic acid, mean residue ellipticity [θ] was measured from 260 - 180 nm. Percent helical content shown in the table was performed using Dichroweb and was calculated using the CDSSTR algorithm for deconvolution against the reference protein database set SMP180 (Sreerama and Woody 2000). (B) Thermal stability was monitored by measuring mean residue ellipticity [θ] at 222 nm from 20°C to 90°C (at 1±0.1°C increments). Temperature denaturation data were analysed using a linear regression line of best fit. R^2 values, slope and percentage change in helical content from 20°C to 90°C are shown in the table.



Analogue	IC_{50} (nM)	Relative Binding affinity (%)	Fold Change
IGF-II	0.626	100	-
Insulin	15.9	3.94	25.4
F26A G22Y F48H L17Y IGF-II 8.5 min	ND	ND	ND
F26A G22Y F48H L17Y IGF-II 11.1 min	ND	ND	ND
F26A G22Y F48H L17Y IGF-II 11.8 min	ND	ND	ND
F26A G22Y F48H L17Y IGF-II 12.7 min	ND	ND	ND

Supplementary Figure 23: Competition of F26A G22Y F48H L17Y IGF-II with europium-labelled IGF-II for binding the IGF-1R.

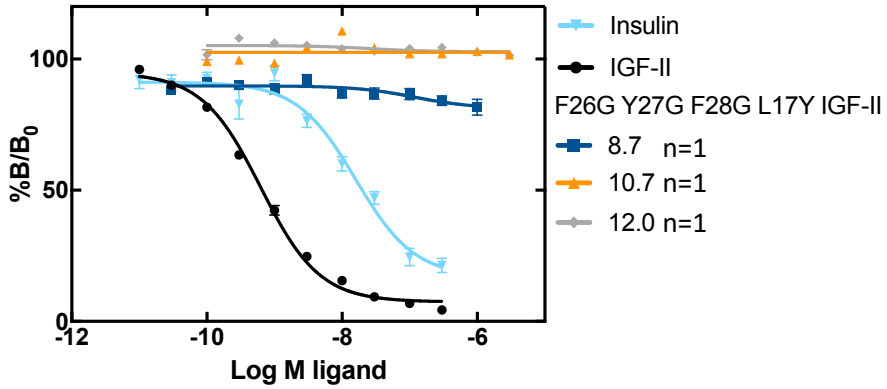
Competition binding of each analogue for the IGF-1R. Immunocaptured IGF-1R was incubated with europium-labelled IGF-II (Eu-IGF-II) and increasing concentrations of competing F26A G22Y F48H L17Y IGF-II. Time-resolved fluorescence was measured as described in section 2.4.1. Results are expressed as a percentage of Eu-IGF-II bound in the absence of competing ligand (B/B_0). Data shown are the mean ±S.E.M. n = ≥ 3 independent experiments each with triplicate technical replicates (unless otherwise stated on the Figure). Error bars are shown when greater than the size of the symbols. Calculated IC_{50} values, % binding affinity and fold change relative to human IGF-II are shown in the table.



Analogue	IC_{50} (nM)	Relative Binding affinity (%)	Fold Change
IGF-II	0.626	100	-
Insulin	15.9	3.94	25.4
F26A F48H L17Y IGF-II 7.4 min	ND	ND	ND
F26A F48H L17Y IGF-II 9.9 min	ND	ND	ND
F26A F48H L17Y IGF-II 10.3 min	>100	≤0.620	>161
F26A F48H L17Y IGF-II 11.3 min	>100	≤0.620	>161

Supplementary Figure 24: Competition of F26A F48H L17Y IGF-II with europium-labelled IGF-II for binding the IGF-1R.

Competition binding of each analogue with the IGF-1R. Immunocaptured IGF-1R was incubated with europium-labelled IGF-II (Eu-IGF-II) and increasing concentrations of competing F26A F48H L17Y IGF-II. Time-resolved fluorescence was measured as described in section 2.4.1. Results are expressed as a percentage of Eu-IGF-II bound in the absence of competing ligand (B/B_0). Data shown are the mean ±S.E.M. n = ≥ 3 independent experiments each with triplicate technical replicates (unless otherwise stated on the Figure). Error bars are shown when greater than the size of the symbols. Calculated IC_{50} values, % binding affinity and fold change relative to human IGF-II are shown in the table.



Analogue	IC_{50} (nM)	Relative Binding affinity (%)	Fold Change
IGF-II	0.626	100	-
Insulin	15.9	3.94	25.4
F26G Y27G F28G L17Y IGF-II 8.7 min	ND	ND	ND
F26G Y27G F28G L17Y IGF-II 10.7 min	ND	ND	ND
F26G Y27G F28G L17Y IGF-II 12 min	ND	ND	ND

Supplementary Figure 25: Competition of F26G Y27G F28G L17Y IGF-II with europium-labelled IGF-II for binding the IGF-1R.

Competition binding of each analogue for the IGF-1R. Immunocaptured IGF-1R was incubated with europium-labelled IGF-II (Eu-IGF-II) and increasing concentrations of competing F26G Y27G F28G L17Y IGF-II. Time-resolved fluorescence was measured as described in section 2.4.1. Results are expressed as a percentage of Eu-IGF-II bound in the absence of competing ligand (B/B_0). Data shown are the mean ±S.E.M. n = ≥ 3 independent experiments each with triplicate technical replicates (unless otherwise stated on the Figure). Error bars are shown when greater than the size of the symbols. Calculated IC_{50} values, % binding affinity and fold change relative to human IGF-II are shown in the table.

9 REFERENCE LIST

ADAMS, T. E., EPA, V. C., GARRETT, T. P. & WARD, C. W. 2000. Structure and function of the type 1 insulin-like growth factor receptor. *Cell Mol Life Sci,* 57**,** 1050-93.

AHORUKOMEYE, P., DISOTUAR, M. M., GAJEWIAK, J., KARANTH, S., WATKINS, M., ROBINSON, S. D., FLOREZ SALCEDO, P., SMITH, N. A., SMITH, B. J., SCHLEGEL, A., FORBES, B. E., OLIVERA, B., HUNG-CHIEH CHOU, D. & SAFAVI-HEMAMI, H. 2019. Fish-hunting cone snail venoms are a rich source of minimized ligands of the vertebrate insulin receptor. *Elife,* 8.

ALTINDIS, E., CAI, W., SAKAGUCHI, M., ZHANG, F., GUOXIAO, W., LIU, F., DE MEYTS, P., GELFANOV, V., PAN, H., DIMARCHI, R. & KAHN, C. R. 2018. Viral insulin-like peptides activate human insulin and IGF-1 receptor signaling: A paradigm shift for host-microbe interactions. *Proc Natl Acad Sci U S A,* 115**,** 2461-2466.

ALVINO, C. L., MCNEIL, K. A., ONG, S. C., DELAINE, C., BOOKER, G. W., WALLACE, J. C., WHITTAKER, J. & FORBES, B. E. 2009. A Novel Approach to Identify Two Distinct Receptor Binding Surfaces of Insulin-like Growth Factor II. *Journal of Biological Chemistry,* 284**,** 7656-7664.

ALVINO, C. L., ONG, S. C., MCNEIL, K. A., DELAINE, C., BOOKER, G. W., WALLACE, J. C. & FORBES, B. E. 2011. Understanding the Mechanism of Insulin and Insulin-Like Growth Factor (IGF) Receptor Activation by IGF-II. *Plos One,* 6.

ARCARO, A. 2013. Targeting the insulin-like growth factor-1 receptor in human cancer. *Front Pharmacol,* 4**,** 30.

BAO, S. J., XIE, D. L., ZHANG, J. P., CHANG, W. R. & LIANG, D. C. 1997. Crystal structure of desheptapeptide(B24-B30)insulin at 1.6 A resolution: implications for receptor binding. *Proc Natl Acad Sci U S A,* 94**,** 2975-80.

BARAL, K. & ROTWEIN, P. 2019. The insulin-like growth factor 2 gene in mammals: Organizational complexity within a conserved locus. *PLoS One,* 14**,** e0219155.

BAYNE, M. L., APPLEBAUM, J., CHICCHI, G. G., HAYES, N. S., GREEN, B. G. & 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., FRASCA, F., PANDINI, G., SCIACCA, L. & VIGNERI, R. 2009. Insulin receptor isoforms and insulin receptor/insulin-like growth factor receptor hybrids in physiology and disease. *Endocr Rev,* 30**,** 586-623.

BELFIORE, A. & MALAGUARNERA, R. 2011. Insulin receptor and cancer. *Endocr Relat Cancer,* 18**,** R125-47.

BELFIORE, A., MALAGUARNERA, R., VELLA, V., LAWRENCE, M. C., SCIACCA, L., FRASCA, F., MORRIONE, A. & VIGNERI, R. 2017. Insulin Receptor Isoforms in Physiology and Disease: An Updated View. *Endocr Rev,* 38**,** 379-431.

BLACKBURN, A., DRESSENDORFER, R. A., BLUM, W. F., ERHARD, M., BREM, G., STRASBURGER, C. J. & WOLF, E. 1997. Interactions of insulin-like growth factor (IGF)-II and growth hormone in vivo: circulating levels of IGF-I and IGF-binding proteins in transgenic mice. *Eur J Endocrinol,* 137**,** 701-8.

BLUNDELL, T., DODSON, G., HODGKIN, D. & MERCOLA, D. 1972. Insulin: The Structure in the Crystal and its Reflection in Chemistry and Biology by. *Advances in protein chemistry,* 26**,** 279-402.

BLYTH, A., ORTIZ, M., MERRIMAN, A., DELAINE, C. & FORBES, B. 2022. Determinants of IGF-II influencing stability, receptor binding and activation. *Sci Rep,* 12**,** 4695.

BLYTH, A. J., KIRK, N. S. & FORBES, B. E. 2020. Understanding IGF-II Action through Insights into Receptor Binding and Activation. *Cells,* 9**,** 2276.

BRACKO, O., SINGER, T., AIGNER, S., KNOBLOCH, M., WINNER, B., RAY, J., CLEMENSON, G. D., JR., SUH, H., COUILLARD-DESPRES, S., AIGNER, L., GAGE, F. H. & JESSBERGER, S. 2012. Gene expression profiling of neural stem cells and their neuronal progeny reveals IGF2 as a regulator of adult hippocampal neurogenesis. *J Neurosci,* 32**,** 3376-87.

BRANGE, J., RIBEL, U., HANSEN, J. F., DODSON, G., HANSEN, M. T., HAVELUND, S., MELBERG, S. G., NORRIS, F., NORRIS, K., SNEL, L. & ET AL. 1988. Monomeric insulins obtained by protein engineering and their medical implications. *Nature,* 333**,** 679-82.

BROMER, W. W. & CHANCE, R. E. 1967. Preparation and characterization of desoctapeptide-insulin. *Biochim Biophys Acta,* 133**,** 219-23.

BROWN, J., DELAINE, C., ZACCHEO, O. J., SIEBOLD, C., GILBERT, R. J., VAN BOXEL, G., DENLEY, A., WALLACE, J. C., HASSAN, A. B., FORBES, B. E. & JONES, E. Y. 2008. Structure and functional analysis of the IGF-II/IGF2R interaction. *Embo Journal,* 27**,** 265-276.

BROWN, J., JONES, E. Y. & FORBES, B. E. 2009. Interactions of IGF-II with the IGF2R/cation-independent mannose-6-phosphate receptor mechanism and biological outcomes. *Vitam Horm,* 80**,** 699-719.

BRZOZOWSKI, A. M., DODSON, E. J., DODSON, G. G., MURSHUDOV, G. N., VERMA, C., TURKENBURG, J. P., DE BREE, F. M. & DAUTER, Z. 2002. Structural origins of the functional divergence of human insulin-like growth factor-I and insulin. *Biochemistry,* 41**,** 9389-97.

BURGISSER, D. M., ROTH, B. V., GIGER, R., LUTHI, C., WEIGL, S., ZARN, J. & HUMBEL, R. E. 1991. Mutants of human insulin-like growth factor II with altered affinities for the type 1 and type 2 insulin-like growth factor receptor. *J Biol Chem,* 266**,** 1029-33.

CABAIL, M. Z., LI, S., LEMMON, E., BOWEN, M. E., HUBBARD, S. R. & MILLER, W. T. 2015. The insulin and IGF1 receptor kinase domains are functional dimers in the activated state. *Nat Commun,* 6**,** 6406.

CARA, J. F., MIRMIRA, R. G., NAKAGAWA, S. H. & 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.

CASCIERI, M. A., CHICCHI, G. G., APPLEBAUM, J., HAYES, N. S., GREEN, B. G. & 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.

CHAO, W. & D'AMORE, P. A. 2008. IGF2: epigenetic regulation and role in development and disease. *Cytokine Growth Factor Rev,* 19**,** 111-20.

CHARALAMBOUS, M., MENHENIOTT, T. R., BENNETT, W. R., KELLY, S. M., DELL, G., DANDOLO, L. & WARD, A. 2004. An enhancer element at the Igf2/H19 locus drives gene expression in both imprinted and non-imprinted tissues. *Dev Biol,* 271**,** 488-97.

CHEN, H., SHI, M., GUO, Z. Y., TANG, Y. H., QIAO, Z. S., LIANG, Z. H. & FENG, Y. M. 2000. Four new monomeric insulins obtained by alanine scanning the dimer-forming surface of the insulin molecule. *Protein Eng,* 13**,** 779-82.

CLARK, J. M. 1988. Novel non-templated nucleotide addition reactions catalyzed by procaryotic and eucaryotic DNA polymerases. *Nucleic Acids Res,* 16**,** 9677-86.

COLOMBO, C., PORZIO, O., LIU, M., MASSA, O., VASTA, M., SALARDI, S., BECCARIA, L., MONCIOTTI, C., TONI, S., PEDERSEN, O., HANSEN, T., FEDERICI, L., PESAVENTO, R., CADARIO, F., FEDERICI, G., GHIRRI, P., ARVAN, P., IAFUSCO, D., BARBETTI, F., EARLY ONSET DIABETES STUDY GROUP OF THE ITALIAN SOCIETY OF PEDIATRIC, E. & DIABETES 2008. Seven mutations in the human insulin gene linked to permanent neonatal/infancy-onset diabetes mellitus. *J Clin Invest,* 118**,** 2148-56.

CONLON, J. M. 2000. Molecular evolution of insulin in non-mammalian vertebrates. *American Zoologist,* 40**,** 200-212.

CONLON, J. M. 2001. Evolution of the insulin molecule: insights into structure-activity and phylogenetic relationships. *Peptides,* 22**,** 1183-93.

CONSTANCIA, M., HEMBERGER, M., HUGHES, J., DEAN, W., FERGUSON-SMITH, A., FUNDELE, R., STEWART, F., KELSEY, G., FOWDEN, A., SIBLEY, C. & REIK, W. 2002. Placental-specific IGF-II is a major modulator of placental and fetal growth. *Nature,* 417**,** 945-8.

CROLL, T. I., SMITH, B. J., MARGETTS, M. B., WHITTAKER, J., WEISS, M. A., WARD, C. W. & LAWRENCE, M. C. 2016. Higher-Resolution Structure of the Human Insulin Receptor Ectodomain: Multi-Modal Inclusion of the Insert Domain. *Structure,* 24**,** 469-76.

DE MEYTS, P. 2015. Insulin/receptor binding: the last piece of the puzzle? What recent progress on the structure of the insulin/receptor complex tells us (or not) about negative cooperativity and activation. *Bioessays,* 37**,** 389-97.

DE MEYTS, P., VAN OBBERGHEN, E. & ROTH, J. 1978. Mapping of the residues responsible for the negative cooperativity of the receptor-binding region of insulin. *Nature,* 273**,** 504-9.

DE MEYTS, P. & WHITTAKER, J. 2002. Structural biology of insulin and IGF1 receptors: implications for drug design. *Nat Rev Drug Discov,* 1**,** 769-83.

DECHIARA, T. M., EFSTRATIADIS, A. & ROBERTSON, E. J. 1990. A growth-deficiency phenotype in heterozygous mice carrying an insulin-like growth factor II gene disrupted by targeting. *Nature,* 345**,** 78-80.

DECHIARA, T. M., ROBERTSON, E. J. & EFSTRATIADIS, A. 1991. Parental imprinting of the mouse insulin-like growth factor II gene. *Cell,* 64**,** 849-59.

DELAINE, C., ALVINO, C. L., MCNEIL, K. A., MULHERN, T. D., GAUGUIN, L., DE MEYTS, P., JONES, E. Y., BROWN, J., WALLACE, J. C. & FORBES, B. E. 2007. A novel binding site for the human insulin-like growth factor-II (IGF-II)/mannose 6-phosphate receptor on IGF-II. *J Biol Chem,* 282**,** 18886-94.

DENLEY, A. 2004. *Investigation into the molecular specificity of the IGFs: a chimeric approach.* Doctor of Philosophy, The University of Adelaide.

DENLEY, A., BONYTHON, E. R., BOOKER, G. W., COSGROVE, L. J., FORBES, B. E., WARD, C. W. & WALLACE, J. C. 2004. Structural determinants for high-affinity binding of insulin-like growth factor II to insulin receptor (IR)-A, the exon 11 minus isoform of the IR. *Mol Endocrinol,* 18**,** 2502-12.

DENLEY, A., COSGROVE, L. J., BOOKER, G. W., WALLACE, J. C. & FORBES, B. E. 2005. Molecular interactions of the IGF system. *Cytokine Growth Factor Rev,* 16**,** 421-39.

DENLEY, A., WALLACE, J. C., COSGROVE, L. J. & FORBES, B. E. 2003. The insulin receptor isoform exon 11- (IR-A) in cancer and other diseases: a review. *Horm Metab Res,* 35**,** 778-85.

DEVEDJIAN, J. C., GEORGE, M., CASELLAS, A., PUJOL, A., VISA, J., PELEGRIN, M., GROS, L. & BOSCH, F. 2000. Transgenic mice overexpressing insulin-like growth factor-II in beta cells develop type 2 diabetes. *J Clin Invest,* 105**,** 731-40.

DODSON, G. & STEINER, D. 1998. The role of assembly in insulin's biosynthesis. *Curr Opin Struct Biol,* 8**,** 189-94.

DREJER, K. 1992. The bioactivity of insulin analogues from in vitro receptor binding to in vivo glucose uptake. *Diabetes Metab Rev,* 8**,** 259-85.

DREJER, K., KRUSE, V., LARSEN, U. D., HOUGAARD, P., BJORN, S. & GAMMELTOFT, S. 1991. Receptor binding and tyrosine kinase activation by insulin analogues with extreme affinities studied in human hepatoma HepG2 cells. *Diabetes,* 40**,** 1488-95.

DUAN, C., REN, H. & GAO, S. 2010. Insulin-like growth factors (IGFs), IGF receptors, and IGF-binding proteins: roles in skeletal muscle growth and differentiation. *Gen Comp Endocrinol,* 167**,** 344-51.

FAVELYUKIS, S., TILL, J. H., HUBBARD, S. R. & MILLER, W. T. 2001. Structure and autoregulation of the insulin-like growth factor 1 receptor kinase. *Nat Struct Biol,* 8**,** 1058-63.

FEIL, R., WALTER, J., ALLEN, N. D. & REIK, W. 1994. Developmental control of allelic methylation in the imprinted mouse Igf2 and H19 genes. *Development,* 120**,** 2933-43.

FERRON, S. R., RADFORD, E. J., DOMINGO-MUELAS, A., KLEINE, I., RAMME, A., GRAY, D., SANDOVICI, I., CONSTANCIA, M., WARD, A., MENHENIOTT, T. R. & FERGUSON-SMITH, A. C. 2015. Differential genomic imprinting regulates paracrine and autocrine roles of IGF2 in mouse adult neurogenesis. *Nat Commun,* 6**,** 8265.

FIELDER, P. J., THORDARSON, G., TALAMANTES, F. & ROSENFELD, R. G. 1990. Characterization of insulin-like growth factor binding proteins (IGFBPs) during gestation in mice: effects of hypophysectomy and an IGFBP-specific serum protease activity. *Endocrinology,* 127**,** 2270-80.

FORBES, B. E., HARTFIELD, P. J., MCNEILL, K. A., SURINYA, K. H., MILNER, S. J., COSGROVE, L. J. & WALLACE, J. C. 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 - Correlation of binding affinity with ability to prevent apoptosis. *European Journal of Biochemistry,* 269**,** 961-968.

FORBES, B. E., MCCARTHY, P. & NORTON, R. S. 2012. Insulin-like growth factor binding proteins: a structural perspective. *Front Endocrinol (Lausanne),* 3**,** 38.

FRANCIS, G. L., APLIN, S. E., MILNER, S. J., MCNEIL, K. A., BALLARD, F. J. &

WALLACE, J. C. 1993. Insulin-like growth factor (IGF)-II binding to IGF-binding proteins and IGF receptors is modified by deletion of the N-terminal hexapeptide or substitution of arginine for glutamate-6 in IGF-II. *Biochem J,* 293 (Pt 3)**,** 713-9.

FRANCIS, G. L., ROSS, M., BALLARD, F. J., MILNER, S. J., SENN, C., MCNEIL, K. A., WALLACE, J. C., KING, R. & WELLS, J. R. 1992. Novel recombinant fusion protein analogues of insulin-like growth factor (IGF)-I indicate the relative importance of IGF-binding protein and receptor binding for enhanced biological potency. *J Mol Endocrinol,* 8**,** 213-23.

FUJITA, M., IEGUCHI, K., CEDANO-PRIETO, D. M., FONG, A., WILKERSON, C., CHEN, J. Q., WU, M., LO, S. H., CHEUNG, A. T., WILSON, M. D., CARDIFF, R. D., BOROWSKY, A. D., TAKADA, Y. K. & TAKADA, Y. 2013. An integrin binding-defective mutant of insulin-like growth factor-1 (R36E/R37E IGF1) acts as a dominant-negative antagonist of the IGF1 receptor (IGF1R) and suppresses tumorigenesis but still binds to IGF1R. *J Biol Chem,* 288**,** 19593-603.

GALLAGHER, E. J. & LEROITH, D. 2010. The proliferating role of insulin and insulin-like growth factors in cancer. *Trends in Endocrinology and Metabolism,* 21**,** 610-618.

GANDERTON, R. H., STANLEY, K. K., FIELD, C. E., COGHLAN, M. P., SOOS, M. A. & SIDDLE, K. 1992. A monoclonal anti-peptide antibody reacting with the insulin receptor beta-subunit. Characterization of the antibody and its epitope and use in immunoaffinity purification of intact receptors. *Biochem J,* 288 (Pt 1)**,** 195-205.

GARRETT, T. P., MCKERN, N. M., LOU, M., FRENKEL, M. J., BENTLEY, J. D., LOVRECZ, G. O., ELLEMAN, T. C., COSGROVE, L. J. & WARD, C. W. 1998. Crystal structure of the first three domains of the type-1 insulin-like growth factor receptor. *Nature,* 394**,** 395-9.

GAUGUIN, L., DELAINE, C., ALVINO, C. L., MCNEIL, K. A., WALLACE, J. C., FORBES, B. E. & DE MEYTS, P. 2008a. Alanine scanning of a putative receptor binding surface of insulin-like growth factor-I. *Journal of Biological Chemistry,* 283**,** 20821-20829.

GAUGUIN, L., KLAPROTH, B., SAJID, W., ANDERSEN, A. S., MCNEIL, K. A., FORBES, B. E. & DE MEYTS, P. 2008b. Structural basis for the lower affinity of the insulin-like growth factors for the insulin receptor. *J Biol Chem,* 283**,** 2604-13.

GICQUEL, C. & LE BOUC, Y. 2006. Hormonal regulation of fetal growth. *Horm Res,* 65 Suppl 3**,** 28-33.

GIORGINO, F., BELFIORE, A., MILAZZO, G., COSTANTINO, A., MADDUX, B., WHITTAKER, J., GOLDFINE, I. D. & VIGNERI, R. 1991. Overexpression of Insulin-Receptors in Fibroblast and Ovary Cells Induces a Ligand-Mediated Transformed Phenotype. *Molecular Endocrinology,* 5**,** 452-459.

GLENDORF, T., KNUDSEN, L., STIDSEN, C. E., HANSEN, B. F., HEGELUND, A. C., SORENSEN, A. R., NISHIMURA, E. & KJELDSEN, T. 2012. Systematic evaluation of the metabolic to mitogenic potency ratio for B10-substituted insulin analogues. *PLoS One,* 7**,** e29198.

GLENDORF, T., SORENSEN, A. R., NISHIMURA, E., PETTERSSON, I. & KJELDSEN, T. 2008. Importance of the solvent-exposed residues of the insulin B chain alpha-helix for receptor binding. *Biochemistry,* 47**,** 4743-51.

GLIDDEN, M. D., ALDABBAGH, K., PHILLIPS, N. B., CARR, K., CHEN, Y. S., WHITTAKER, J., PHILLIPS, M., WICKRAMASINGHE, N. P., REGE, N., SWAIN, M., PENG, Y., YANG, Y., LAWRENCE, M. C., YEE, V. C., ISMAIL-BEIGI, F. & WEISS, M. A. 2018a. An ultra-stable single-chain insulin analog resists thermal inactivation and exhibits biological signaling duration equivalent to the native protein. *J Biol Chem,* 293**,** 47-68.

GLIDDEN, M. D., YANG, Y., SMITH, N. A., PHILLIPS, N. B., CARR, K., WICKRAMASINGHE, N. P., ISMAIL-BEIGI, F., LAWRENCE, M. C., SMITH, B. J. & WEISS, M. A. 2018b. Solution structure of an ultra-stable single-chain insulin analog connects protein dynamics to a novel mechanism of receptor binding. *J Biol Chem,* 293**,** 69-88.

GLUCKMAN, P. D. & BUTLER, J. H. 1983. Parturition-related changes in insulin-like growth factors-I and -II in the perinatal lamb. *J Endocrinol,* 99**,** 223-32.

GRONKE, S., CLARKE, D. F., BROUGHTON, S., ANDREWS, T. D. & PARTRIDGE, L. 2010. Molecular Evolution and Functional Characterization of Drosophila Insulin-Like Peptides. *Plos Genetics,* 6.

GUALBERTO, A. & POLLAK, M. 2009. Emerging role of insulin-like growth factor receptor inhibitors in oncology: early clinical trial results and future directions. *Oncogene,* 28**,** 3009-3021.

GUTMANN, T., SCHAFER, I. B., POOJARI, C., BRANKATSCHK, B., VATTULAINEN, I., STRAUSS, M. & COSKUN, U. 2020. Cryo-EM structure of the complete and ligand-saturated insulin receptor ectodomain. *J Cell Biol,* 219.

HAKUNO, F. & TAKAHASHI, S. I. 2018. IGF1 receptor signaling pathways. *J Mol Endocrinol,* 61**,** T69-T86.

HAMEL, F. G., SIFORD, G. L., FAWCETT, J., CHANCE, R. E., FRANK, B. H. & DUCKWORTH, W. C. 1999. Differences in the cellular processing of AspB10 human insulin compared with human insulin and LysB28ProB29 human insulin. *Metabolism,* 48**,** 611-7.

HANKE, S. & MANN, M. 2009. The phosphotyrosine interactome of the insulin receptor family and its substrates IRS-1 and IRS-2. *Mol Cell Proteomics,* 8**,** 519-34.

HARRIS, L. K. & WESTWOOD, M. 2012. Biology and significance of signalling pathways activated by IGF-II. *Growth Factors,* 30**,** 1-12.

HEJNAES, K. R., BAYNE, S., NORSKOV, L., SORENSEN, H. H., THOMSEN, J., SCHAFFER, L., WOLLMER, A. & SKRIVER, L. 1992. Development of an optimized refolding process for recombinant Ala-Glu-IGF-1. *Protein Eng,* 5**,** 797-806.

HEXNEROVA, R., KRIZKOVA, K., FABRY, M., SIEGLOVA, I., KEDROVA, K., COLLINSOVA, M., ULLRICHOVA, P., SRB, P., WILLIAMS, C., CRUMP, M. P., TOSNER, Z., JIRACEK, J., VEVERKA, V. & ZAKOVA, L. 2016. Probing Receptor Specificity by Sampling the Conformational Space of the Insulin-like Growth Factor II C-domain. *J Biol Chem,* 291**,** 21234-21245.

HOBER, S., FORSBERG, G., PALM, G., HARTMANIS, M. & NILSSON, B. 1992. Disulfide exchange folding of insulin-like growth factor I. *Biochemistry,* 31**,** 1749-56.

HOBER, S., UHLEN, M. & NILSSON, B. 1997. Disulfide exchange folding of disulfide mutants of insulin-like growth factor I in vitro. *Biochemistry,* 36**,** 4616-22.

HODGSON, D. R., MAY, F. E. & WESTLEY, B. R. 1996. Involvement of phenylalanine 23 in the binding of IGF-1 to the insulin and type I IGF receptor. *Regul Pept,* 66**,** 191-6.

HOLLAND, M. D., HOSSNER, K. L., WILLIAMS, S. E., WALLACE, C. R., NISWENDER, G. D. & ODDE, K. G. 1997. Serum concentrations of insulin-like growth factors and placental lactogen during gestation in cattle. I. Fetal profiles. *Domest Anim Endocrinol,* 14**,** 231-9.

HOLLY, J. M. P., BIERNACKA, K. & PERKS, C. M. 2019. The Neglected Insulin: IGF-II, a Metabolic Regulator with Implications for Diabetes, Obesity, and Cancer. *Cells,* 8.

HU, S. Q., BURKE, G. T. & KATSOYANNIS, P. G. 1993. Contribution of the B16 and B26 tyrosine residues to the biological activity of insulin. *J Protein Chem,* 12**,** 741-7.

HUA, Q. X., HU, S. Q., FRANK, B. H., JIA, W., CHU, Y. C., WANG, S. H., BURKE, G. T., KATSOYANNIS, P. G. & WEISS, M. A. 1996. Mapping the functional surface of insulin by design: structure and function of a novel A-chain analogue. *J Mol Biol,* 264**,** 390-403.

HUA, Q. X., LIU, M., HU, S. Q., JIA, W., ARVAN, P. & WEISS, M. A. 2006a. A conserved histidine in insulin is required for the foldability of human proinsulin: structure and function of an ALAB5 analog. *J Biol Chem,* 281**,** 24889-99.

HUA, Q. X., MAYER, J. P., JIA, W., ZHANG, J. & WEISS, M. A. 2006b. The folding

nucleus of the insulin superfamily: a flexible peptide model foreshadows the native state. *J Biol Chem,* 281**,** 28131-42.

HUA, Q. X., NAKAGAWA, S. H., JIA, W., HU, S. Q., CHU, Y. C., KATSOYANNIS, P. G. & WEISS, M. A. 2001. Hierarchical protein folding: asymmetric unfolding of an insulin analogue lacking the A7-B7 interchain disulfide bridge. *Biochemistry,* 40**,** 12299-311.

HUA, Q. X., NAKAGAWA, S. H., JIA, W., HUANG, K., PHILLIPS, N. B., HU, S. Q. & WEISS, M. A. 2008. Design of an active ultrastable single-chain insulin analog: synthesis, structure, and therapeutic implications. *J Biol Chem,* 283**,** 14703-16.

HUA, Q. X., SHOELSON, S. E., INOUYE, K. & WEISS, M. A. 1993. Paradoxical structure and function in a mutant human insulin associated with diabetes mellitus. *Proc Natl Acad Sci U S A,* 90**,** 582-6.

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. & 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.

HUBBARD, S. R. 2004. Juxtamembrane autoinhibition in receptor tyrosine kinases. *Nat Rev Mol Cell Biol,* 5**,** 464-71.

HUBBARD, S. R., WEI, L., ELIS, L. & HENDRICKSON, W. A. 1994. Crystal-Structure of the Tyrosine Kinase Domain of the Human Insulin-Receptor. *Nature,* 372**,** 746-754.

INOUYE, K., WATANABE, K., TOCHINO, Y., KOBAYASHI, M. & SHIGETA, Y. 1981. Semisynthesis and properties of some insulin analogs. *Biopolymers,* 20**,** 1845-58.

IRWIN, D. M. 2021. Evolution of the mammalian insulin (Ins) gene; Changes in proteolytic processing. *Peptides,* 135**,** 170435.

JANSSON, M., UHLEN, M. & NILSSON, B. 1997. Structural changes in insulin-like growth factor (IGF) I mutant proteins affecting binding kinetic rates to IGF binding protein 1 and IGF-I receptor. *Biochemistry,* 36**,** 4108-17.

JIAO, S., REN, H., LI, Y., ZHOU, J., DUAN, C. & LU, L. 2013. Differential regulation of IGF-I and IGF-II gene expression in skeletal muscle cells. *Mol Cell Biochem,* 373**,** 107-13.

KAARSHOLM, N. C., NORRIS, K., JORGENSEN, R. J., MIKKELSEN, J., LUDVIGSEN, S., OLSEN, O. H., SORENSEN, A. R. & HAVELUND, S. 1993. Engineering stability of the insulin monomer fold with application to structure-activity relationships. *Biochemistry,* 32**,** 10773-8.

KATSOYANNIS, P. G. & TOMETSKO, A. 1966. Insulin synthesis by recombination of A and B chains: a highly efficient method. *Proc Natl Acad Sci U S A,* 55**,** 1554-61.

KAVRAN, J. M., MCCABE, J. M., BYRNE, P. O., CONNACHER, M. K., WANG, Z., RAMEK, A., SARABIPOUR, S., SHAN, Y., SHAW, D. E., HRISTOVA, K., COLE, P. A. & LEAHY, D. J. 2014. How IGF-1 activates its receptor. *Elife,* 3.

KEMMLER, W., PETERSON, J. D. & STEINER, D. F. 1971. Studies on the conversion of proinsulin to insulin. I. Conversion in vitro with trypsin and carboxypeptidase B. *J Biol Chem,* 246**,** 6786-91.

KING, R., WELLS, J. R., KRIEG, P., SNOSWELL, M., BRAZIER, J., BAGLEY, C. J., WALLACE, J. C., BALLARD, F. J., ROSS, M. & FRANCIS, G. L. 1992. Production and characterization of recombinant insulin-like growth factor-I (IGF-I) and potent analogues of IGF-I, with Gly or Arg substituted for Glu3, following their expression in Escherichia coli as fusion proteins. *J Mol Endocrinol,* 8**,** 29-41.

KISELYOV, V. V., VERSTEYHE, S., GAUGUIN, L. & DE MEYTS, P. 2009. Harmonic oscillator model of the insulin and IGF1 receptors' allosteric binding and activation. *Mol Syst Biol,* 5**,** 243.

KJELDSEN, T., WIBERG, F. C. & 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.

KLINT, J. K., SENFF, S., SAEZ, N. J., SESHADRI, R., LAU, H. Y., BENDE, N. S., UNDHEIM, E. A., RASH, L. D., MOBLI, M. & KING, G. F. 2013. Production of recombinant disulfide-rich venom peptides for structural and functional analysis via expression in the periplasm of E. coli. *PLoS One,* 8**,** e63865.

KOBAYASHI, M., OHGAKUO, S., IWASAKI, M., MAEGAWA, H., WATANBAE, N., TAKADA, Y., SHIGETA, Y. & INOUE, K. 1984. Changes in receptor binding, biological activity and immunoreactivity of insulin caused by replacing the residues B23-B26 with alanine. *Biomedical research,* 5**,** 267-272.

KOBAYASHI, M., SASAOKA, T., SUGIBAYASHI, M., IWANISHI, M. & SHIGETA, Y. 1989. Receptor binding and biologic activity of biosynthetic human insulin and mini-proinsulin produced by recombinant gene technology. *Diabetes Res Clin Pract,* 7**,** 25-8.

KRISTENSEN, C., KJELDSEN, T., WIBERG, F. C., SCHAFFER, L., HACH, M., HAVELUND, S., BASS, J., STEINER, D. F. & ANDERSEN, A. S. 1997. Alanine scanning mutagenesis of insulin. *J Biol Chem,* 272**,** 12978-83.

KURTZHALS, P., SCHAFFER, L., SORENSEN, A., KRISTENSEN, C., JONASSEN, I., SCHMID, C. & TRUB, T. 2000. Correlations of receptor binding and metabolic and mitogenic potencies of insulin analogs designed for clinical use. *Diabetes,* 49**,** 999-1005.

LAVIOLA, L., NATALICCHIO, A. & GIORGINO, F. 2007. The IGF-I signaling pathway. *Curr Pharm Des,* 13**,** 663-9.

LAWRENCE, M. C. 2021. Understanding insulin and its receptor from their three-dimensional structures. *Mol Metab,* 52**,** 101255.

LEE, C. Y., CHUNG, C. S. & SIMMEN, F. A. 1993. Ontogeny of the porcine insulin-like growth factor system. *Mol Cell Endocrinol,* 93**,** 71-80.

LEE, J. & PILCH, P. F. 1994. The insulin receptor: structure, function, and signaling. *Am J Physiol,* 266**,** C319-34.

LEHTINEN, M. K., ZAPPATERRA, M. W., CHEN, X., YANG, Y. J., HILL, A. D., LUN, M., MAYNARD, T., GONZALEZ, D., KIM, S., YE, P., D'ERCOLE, A. J., WONG, E. T., LAMANTIA, A. S. & WALSH, C. A. 2011. The cerebrospinal fluid provides a proliferative niche for neural progenitor cells. *Neuron,* 69**,** 893-905.

LI, D., CHEN, X., ZHU, F. & CHEN, K. 2020. Insulin-like peptides in model insects. *Isj-Invertebrate Survival Journal,* 17**,** 186-195.

LI, J., CHOI, E., YU, H. & BAI, X. C. 2019. Structural basis of the activation of type 1 insulin-like growth factor receptor. *Nat Commun,* 10**,** 4567.

LI, J., PARK, J., MAYER, J. P., WEBB, K. J., UCHIKAWA, E., WU, J., LIU, S., ZHANG, X., STOWELL, M. H. B., CHOI, E. & BAI, X. C. 2022. Synergistic activation of the insulin receptor via two distinct sites. *Nat Struct Mol Biol,* 29**,** 357-368.

LI, S., COVINO, N. D., STEIN, E. G., TILL, J. H. & HUBBARD, S. R. 2003. Structural and biochemical evidence for an autoinhibitory role for tyrosine 984 in the juxtamembrane region of the insulin receptor. *J Biol Chem,* 278**,** 26007-14.

LIEN, S., MILNER, S. J., GRAHAM, L. D., WALLACE, J. C. & FRANCIS, G. L. 2001. Linkers for improved cleavage of fusion proteins with an engineered alpha-lytic protease. *Biotechnol Bioeng,* 74**,** 335-43.

LIVINGSTONE, C. 2013. IGF2 and cancer. *Endocr Relat Cancer,* 20**,** R321-39.

LOU, M., GARRETT, T. P., MCKERN, N. M., HOYNE, P. A., EPA, V. C., BENTLEY, J. D., LOVRECZ, G. O., COSGROVE, L. J., FRENKEL, M. J. & WARD, C. W. 2006. The first three domains of the insulin receptor differ structurally from the insulin-like growth factor 1 receptor in the regions governing ligand specificity. *Proc Natl Acad Sci U S A,* 103**,** 12429-34.

LUDVIGSEN, S., ROY, M., THOGERSEN, H. & KAARSHOLM, N. C. 1994. High-resolution structure of an engineered biologically potent insulin monomer, B16 Tyr-->His, as determined by nuclear magnetic resonance spectroscopy. *Biochemistry,* 33**,** 7998-8006.

LUI, J. C. & BARON, J. 2013. Evidence that Igf2 down-regulation in postnatal tissues and up-regulation in malignancies is driven by transcription factor E2f3. *Proc Natl Acad Sci U S A,* 110**,** 6181-6.

MACHACKOVA, K., CHRUDINOVA, M., RADOSAVLJEVIC, J., POTALITSYN, P., KRIZKOVA, K., FABRY, M., SELICHAROVA, I., COLLINSOVA, M., BRZOZOWSKI, A. M., ZAKOVA, L. & JIRACEK, J. 2018. Converting Insulin-like Growth Factors 1 and 2 into High-Affinity Ligands for Insulin Receptor Isoform A by the Introduction of an Evolutionarily Divergent Mutation. *Biochemistry,* 57**,** 2373-2382.

MALAGUARNERA, R., SACCO, A., VOCI, C., PANDINI, G., VIGNERI, R. & BELFIORE, A. 2012. Proinsulin binds with high affinity the insulin receptor isoform A and predominantly activates the mitogenic pathway. *Endocrinology,* 153**,** 2152-63.

MANTA, B., BOYD, D. & BERKMEN, M. 2019. Disulfide Bond Formation in the Periplasm of Escherichia coli. *EcoSal Plus,* 8.

MCKERN, N. M., LAWRENCE, M. C., STRELTSOV, V. A., LOU, M. Z., ADAMS, T. E., LOVRECZ, G. O., ELLEMAN, T. C., RICHARDS, K. M., BENTLEY, J. D., PILLING, P. A., HOYNE, P. A., CARTLEDGE, K. A., PHAM, T. M., LEWIS, J. L., SANKOVICH, S. E., STOICHEVSKA, V., DA SILVA, E., ROBINSON, C. P., FRENKEL, M. J., SPARROW, L. G., FERNLEY, R. T., EPA, V. C. & WARD, C. W. 2006. Structure of the insulin receptor ectodomain reveals a folded-over conformation. *Nature,* 443**,** 218-21.

MENTING, J. G., GAJEWIAK, J., MACRAILD, C. A., CHOU, D. H., DISOTUAR, M. M., SMITH, N. A., MILLER, C., ERCHEGYI, J., RIVIER, J. E., OLIVERA, B. M., FORBES, B. E., SMITH, B. J., NORTON, R. S., SAFAVI-HEMAMI, H. & LAWRENCE, M. C. 2016. A minimized human insulin-receptor-binding motif revealed in a Conus geographus venom insulin. *Nat Struct Mol Biol,* 23**,** 916-920.

MENTING, J. G., WARD, C. W., MARGETTS, M. B. & LAWRENCE, M. C. 2009. A thermodynamic study of ligand binding to the first three domains of the human insulin receptor: relationship between the receptor alpha-chain C-terminal peptide and the site 1 insulin mimetic peptides. *Biochemistry,* 48**,** 5492-500.

MENTING, J. G., WHITTAKER, J., MARGETTS, M. B., WHITTAKER, L. J., KONG, G. K., SMITH, B. J., WATSON, C. J., ZAKOVA, L., KLETVIKOVA, E., JIRACEK, J., CHAN, S. J., STEINER, D. F., DODSON, G. G., BRZOZOWSKI, A. M., WEISS, M. A., WARD, C. W. & LAWRENCE, M. C. 2013. How insulin engages its primary binding site on the insulin receptor. *Nature,* 493**,** 241-5.

MENTING, J. G., YANG, Y., CHAN, S. J., PHILLIPS, N. B., SMITH, B. J., WHITTAKER, J., WICKRAMASINGHE, N. P., WHITTAKER, L. J., PANDYARAJAN, V., WAN, Z. L., YADAV, S. P., CARROLL, J. M., STROKES, N., ROBERTS, C. T., JR., ISMAIL-BEIGI, F., MILEWSKI, W., STEINER, D. F., CHAUHAN, V. S., WARD, C. W., WEISS, M. A. & LAWRENCE, M. C. 2014. Protective hinge in insulin opens to enable its receptor engagement. *Proc Natl Acad Sci U S A,* 111**,** E3395-404.

MILLER, J. A., NARHI, L. O., HUA, Q. X., ROSENFELD, R., ARAKAWA, T., ROHDE, M., PRESTRELSKI, S., LAUREN, S., STONEY, K. S., TSAI, L. & ET AL. 1993. Oxidative refolding of insulin-like growth factor 1 yields two products of similar thermodynamic stability: a bifurcating protein-folding pathway. *Biochemistry,* 32**,** 5203-13.

MILNER, S. J., CARVER, J. A., BALLARD, F. J. & FRANCIS, G. L. 1999. Probing the disulfide folding pathway of insulin-like growth factor-I. *Biotechnol Bioeng,* 62**,** 693-703.

MILNER, S. J., FRANCIS, G. L., WALLACE, J. C., MAGEE, B. A. & BALLARD, F. J. 1995. Mutations in the B-domain of insulin-like growth factor-I influence the oxidative folding to yield products with modified biological properties. *Biochem J,* 308 (Pt 3)**,** 865-71.

MILNER, S. J., THOMAS, S. M., BALLARD, F. J. & FRANCIS, G. L. 1996. Optimization of the hydroxylamine cleavage of an expressed fusion protein to produce recombinant human insulin-like growth factor (IGF)-I. *Biotechnol Bioeng,* 50**,** 265-72.

MIRMIRA, R. G., NAKAGAWA, S. H. & 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.

MIRMIRA, R. G. & TAGER, H. S. 1989. Role of the phenylalanine B24 side chain in directing insulin interaction with its receptor. Importance of main chain conformation. *J Biol Chem,* 264**,** 6349-54.

MORCAVALLO, A., GASPARI, M., PANDINI, G., PALUMMO, A., CUDA, G., LARSEN, M. R., VIGNERI, R. & BELFIORE, A. 2011. Research resource: New and diverse substrates for the insulin receptor isoform A revealed by quantitative proteomics after stimulation with IGF-II or insulin. *Mol Endocrinol,* 25**,** 1456-68.

MOREAU, F., KIRK, N. S., ZHANG, F., GELFANOV, V., LIST, E. O., CHRUDINOVA, M., VENUGOPAL, H., LAWRENCE, M. C., JIMENEZ, V., BOSCH, F., KOPCHICK, J. J., DIMARCHI, R. D., ALTINDIS, E. & RONALD KAHN, C. 2022. Interaction of a viral insulin-like peptide with the IGF-1 receptor produces a natural antagonist. *Nat Commun,* 13**,** 6700.

NAKAGAWA, S. H., HUA, Q. X., HU, S. Q., JIA, W., WANG, S., KATSOYANNIS, P. G. & WEISS, M. A. 2006. Chiral mutagenesis of insulin. Contribution of the B20-B23 beta-turn to activity and stability. *J Biol Chem,* 281**,** 22386-22396.

NAKAGAWA, S. H. & TAGER, H. S. 1986. Role of the phenylalanine B25 side chain in directing insulin interaction with its receptor. Steric and conformational effects. *J Biol Chem,* 261**,** 7332-41.

NAKAGAWA, S. H., ZHAO, M., HUA, Q. X., HU, S. Q., WAN, Z. L., JIA, W. & WEISS, M. A. 2005. Chiral mutagenesis of insulin. Foldability and function are inversely

regulated by a stereospecific switch in the B chain. *Biochemistry,* 44**,** 4984-99.

NARHI, L. O., HUA, Q. X., ARAKAWA, T., FOX, G. M., TSAI, L., ROSENFELD, R., HOLST, P., MILLER, J. A. & WEISS, M. A. 1993. Role of native disulfide bonds in the structure and activity of insulin-like growth factor 1: genetic models of protein-folding intermediates. *Biochemistry,* 32**,** 5214-21.

O'NEIL, K. T. & DEGRADO, W. F. 1990. A thermodynamic scale for the helix-forming tendencies of the commonly occurring amino acids. *Science,* 250**,** 646-51.

PALSGAARD, J. 2003. Receptor Binding Properties of Twenty-five Different Hystricomorph-like Human Insulin Analogs. *Masters thesis University of Copenhagen, Denmark*.

PANDYARAJAN, V., PHILLIPS, N. B., REGE, N., LAWRENCE, M. C., WHITTAKER, J. & WEISS, M. A. 2016. Contribution of TyrB26 to the Function and Stability of Insulin: STRUCTURE-ACTIVITY RELATIONSHIPS AT A CONSERVED HORMONE-RECEPTOR INTERFACE. *J Biol Chem,* 291**,** 12978-90.

PANDYARAJAN, V., SMITH, B. J., PHILLIPS, N. B., WHITTAKER, L., COX, G. P., WICKRAMASINGHE, N., MENTING, J. G., WAN, Z. L., WHITTAKER, J., ISMAIL-BEIGI, F., LAWRENCE, M. C. & WEISS, M. A. 2014. Aromatic anchor at an invariant hormone-receptor interface: function of insulin residue B24 with application to protein design. *J Biol Chem,* 289**,** 34709-27.

PIETRZKOWSKI, Z., LAMMERS, R., CARPENTER, G., SODERQUIST, A. M., LIMARDO, M., PHILLIPS, P. D., ULLRICH, A. & BASERGA, R. 1992. Constitutive expression of insulin-like growth factor 1 and insulin-like growth factor 1 receptor abrogates all requirements for exogenous growth factors. *Cell Growth Differ,* 3**,** 199-205.

POLLAK, M. 2008. Insulin and insulin-like growth factor signalling in neoplasia. *Nature Reviews Cancer,* 8**,** 915-928.

QIAO, Z. S., GUO, Z. Y. & FENG, Y. M. 2001. Putative disulfide-forming pathway of porcine insulin precursor during its refolding in vitro. *Biochemistry,* 40**,** 2662-8.

QIAO, Z. S., MIN, C. Y., HUA, Q. X., WEISS, M. A. & FENG, Y. M. 2003. In vitro refolding of human proinsulin. Kinetic intermediates, putative disulfide-forming pathway folding initiation site, and potential role of C-peptide in folding process. *J Biol Chem,* 278**,** 17800-9.

RAJAPAKSHA, H. & FORBES, B. E. 2015. Ligand-binding affinity at the insulin receptor isoform-A and subsequent IR-A tyrosine phosphorylation kinetics are important determinants of mitogenic biological outcomes. *Frontiers in Endocrinology,* 6.

RANDHAWA, R. & COHEN, P. 2005. The role of the insulin-like growth factor system in prenatal growth. *Mol Genet Metab,* 86**,** 84-90.

REGE, N. K., LIU, M., DHAYALAN, B., CHEN, Y. S., SMITH, N. A., RAHIMI, L., SUN, J., GUO, H., YANG, Y., HAATAJA, L., PHILLIPS, N. F. B., WHITTAKER, J., SMITH, B. J., ARVAN, P., ISMAIL-BEIGI, F. & WEISS, M. A. 2020a. "Register-shift" insulin analogs uncover constraints of proteotoxicity in protein evolution. *J Biol Chem,* 295**,** 3080-3098.

REGE, N. K., LIU, M., YANG, Y., DHAYALAN, B., WICKRAMASINGHE, N. P., CHEN, Y. S., RAHIMI, L., GUO, H., HAATAJA, L., SUN, J., ISMAIL-BEIGI, F., PHILLIPS, N. B., ARVAN, P. & WEISS, M. A. 2020b. Evolution of insulin at the edge of foldability and its medical implications. *Proc Natl Acad Sci U S A,* 117**,** 29618-29628.

ROSENFELD, R. D., MILLER, J. A., NARHI, L. O., HAWKINS, N., KATTA, V., LAUREN, S., WEISS, M. A. & ARAKAWA, T. 1997. Putative folding pathway of insulin-like growth factor-I. *Arch Biochem Biophys,* 342**,** 298-305.

ROTH, B. V., BURGISSER, D. M., LUTHI, C. & HUMBEL, R. E. 1991. Mutants of human insulin-like growth factor II: expression and characterization of analogs with a substitution of TYR27 and/or a deletion of residues 62-67. *Biochem Biophys Res Commun,* 181**,** 907-14.

ROTWEIN, P. 2017. Diversification of the insulin-like growth factor 1 gene in mammals. *PLoS One,* 12**,** e0189642.

SAEGUSA, J., YAMAJI, S., IEGUCHI, K., WU, C. Y., LAM, K. S., LIU, F. T., TAKADA, Y. K. & TAKADA, Y. 2009. The direct binding of insulin-like growth factor-1 (IGF-1) to integrin alphavbeta3 is involved in IGF-1 signaling. *J Biol Chem,* 284**,** 24106-14.

SAFAVI-HEMAMI, H., GAJEWIAK, J., KARANTH, S., ROBINSON, S. D., UEBERHEIDE, B., DOUGLASS, A. D., SCHLEGEL, A., IMPERIAL, J. S., WATKINS, M., BANDYOPADHYAY, P. K., YANDELL, M., LI, Q., PURCELL, A. W., NORTON, R. S., ELLGAARD, L. & OLIVERA, B. M. 2015. Specialized insulin is used for chemical warfare by fish-hunting cone snails. *Proc Natl Acad Sci U S A,* 112**,** 1743-8.

SAFAVI-HEMAMI, H., GORASIA, D. G., STEINER, A. M., WILLIAMSON, N. A., KARAS, J. A., GAJEWIAK, J., OLIVERA, B. M., BULAJ, G. & PURCELL, A. W. 2012. Modulation of conotoxin structure and function is achieved through a multienzyme complex in the venom glands of cone snails. *J Biol Chem,* 287**,** 34288-303.

SAKANO, K., ENJOH, T., NUMATA, F., FUJIWARA, H., MARUMOTO, Y., HIGASHIHASHI, N., SATO, Y., PERDUE, J. F. & FUJITA-YAMAGUCHI, Y. 1991. The design, expression, and characterization of human insulin-like growth factor II (IGF-II) mutants specific for either the IGF-II/cation-independent mannose 6-phosphate receptor or IGF-I receptor. *J Biol Chem,* 266**,** 20626-35.

SALIC, A. & MITCHISON, T. J. 2008. A chemical method for fast and sensitive detection

of DNA synthesis in vivo. *Proc Natl Acad Sci U S A,* 105**,** 2415-20.

SAMANI, A. A., YAKAR, S., LEROITH, D. & BRODT, P. 2007. The role of the IGF system in cancer growth and metastasis: overview and recent insights. *Endocr Rev,* 28**,** 20-47.

SCAPIN, G., DANDEY, V. P., ZHANG, Z., PROSISE, W., HRUZA, A., KELLY, T., MAYHOOD, T., STRICKLAND, C., POTTER, C. S. & CARRAGHER, B. 2018. Structure of the insulin receptor-insulin complex by single-particle cryo-EM analysis. *Nature,* 556**,** 122-125.

SCHAFFER, L. 1994. A model for insulin binding to the insulin receptor. *Eur J Biochem,* 221**,** 1127-32.

SCHUMACHER, R., SOOS, M. A., SCHLESSINGER, J., BRANDENBURG, D., SIDDLE, K. & ULLRICH, A. 1993. Signaling-competent receptor chimeras allow mapping of major insulin receptor binding domain determinants. *J Biol Chem,* 268**,** 1087-94.

SCHWARTZ, G. P., BURKE, G. T. & KATSOYANNIS, P. G. 1987. A superactive insulin: [B10-aspartic acid]insulin(human). *Proc Natl Acad Sci U S A,* 84**,** 6408-11.

SCHWARTZ, G. P., WONG, D., BURKE, G. T., DE VROEDE, M. A., RECHLER, M. M. & KATSOYANNIS, P. G. 1985. Glutamine B16 insulin: reduced insulin-like metabolic activity with moderately preserved mitogenic activity. *Journal of protein chemistry,* 4**,** 185-197.

SCIACCA, L., COSTANTINO, A., PANDINI, G., MINEO, R., FRASCA, F., SCALIA, P., SBRACCIA, P., GOLDFINE, I. D., VIGNERI, R. & BELFIORE, A. 1999. Insulin receptor activation by IGF-II in breast cancers: evidence for a new autocrine/paracrine mechanism. *Oncogene,* 18**,** 2471-9.

SHOELSON, S., HANEDA, M., BLIX, P., NANJO, A., SANKE, T., INOUYE, K., STEINER, D., RUBENSTEIN, A. & TAGER, H. 1983. Three mutant insulins in man. *Nature,* 302**,** 540-3.

SHOELSON, S. E., LEE, J., LYNCH, C. S., BACKER, J. M. & PILCH, P. F. 1993. BpaB25 insulins. Photoactivatable analogues that quantitatively cross-link, radiolabel, and activate the insulin receptor. *J Biol Chem,* 268**,** 4085-91.

SHOOTER, G. K., MAGEE, B., SOOS, M. A., FRANCIS, G. L., SIDDLE, K. & WALLACE, J. C. 1996. Insulin-like growth factor (IGF)-I A- and B-domain analogues with altered type 1 IGF and insulin receptor binding specificities. *J Mol Endocrinol,* 17**,** 237-46.

SIDDLE, K. 2011. Signalling by insulin and IGF receptors: supporting acts and new players. *J Mol Endocrinol,* 47**,** R1-10.

SIEBER, P., EISLER, K., KAMBER, B., RINIKER, B., RITTEL, W., MARKI, F. & DE GASPARO, M. 1978. Synthesis and biological activity of two disulphide bond isomers of human insulin: [A7-A11,A6-B7-cystine]- and [A6-A7,A11-B7-cystine]insulin (human). *Hoppe Seylers Z Physiol Chem,* 359**,** 113-23.

SMITH, A. C., CHOUFANI, S., FERREIRA, J. C. & WEKSBERG, R. 2007. Growth regulation, imprinted genes, and chromosome 11p15.5. *Pediatr Res,* 61**,** 43R-47R.

SMITH, G. D., PANGBORN, W. A. & BLESSING, R. H. 2003. The structure of T6 human insulin at 1.0 A resolution. *Acta Crystallogr D Biol Crystallogr,* 59**,** 474-82.

SMITH, M. C., COOK, J. A., FURMAN, T. C. & OCCOLOWITZ, J. L. 1989. Structure and activity dependence of recombinant human insulin-like growth factor II on disulfide bond pairing. *J Biol Chem,* 264**,** 9314-21.

SMITH, N. A., MENTING, J. G., WEISS, M. A., LAWRENCE, M. C. & SMITH, B. J. 2022a. Single-chain insulin analogs threaded by the insulin receptor aCT domain. *Biophysical Journal,* 121**,** 4063-4077.

SMITH, N. A., MENTING, J. G., WEISS, M. A., LAWRENCE, M. C. & SMITH, B. J. 2022b. Single-chain insulin analogs threaded by the insulin receptor alphaCT domain. *Biophys J,* 121**,** 4063-4077.

SOARES, M. B., TURKEN, A., ISHII, D., MILLS, L., EPISKOPOU, V., COTTER, S., ZEITLIN, S. & EFSTRATIADIS, A. 1986. Rat insulin-like growth factor II gene. A single gene with two promoters expressing a multitranscript family. *J Mol Biol,* 192**,** 737-52.

SOOS, M. A. & SIDDLE, K. 1989. Immunological relationships between receptors for insulin and insulin-like growth factor I. Evidence for structural heterogeneity of insulin-like growth factor I receptors involving hybrids with insulin receptors. *Biochem J,* 263**,** 553-63.

SREERAMA, N. & WOODY, R. W. 2000. Estimation of protein secondary structure from circular dichroism spectra: comparison of CONTIN, SELCON, and CDSSTR methods with an expanded reference set. *Anal Biochem,* 287**,** 252-60.

STYLIANOPOULOU, F., HERBERT, J., SOARES, M. B. & EFSTRATIADIS, A. 1988. Expression of the insulin-like growth factor II gene in the choroid plexus and the leptomeninges of the adult rat central nervous system. *Proc Natl Acad Sci U S A,* 85**,** 141-5.

SUN, J., XIONG, Y., LI, X., HAATAJA, L., CHEN, W., MIR, S. A., LV, L., MADLEY, R., LARKIN, D., ANJUM, A., DHAYALAN, B., REGE, N., WICKRAMASINGHE, N. P., WEISS, M. A., ITKIN-ANSARI, P., KAUFMAN, R. J., OSTROV, D. A., ARVAN, P. & LIU, M. 2020. Role of Proinsulin Self-Association in Mutant INS Gene-Induced Diabetes of Youth. *Diabetes,* 69**,** 954-964.

SURINYA, K. H., FORBES, B. E., OCCHIODORO, F., BOOKER, G. W., FRANCIS, G. L., SIDDLE, K., WALLACE, J. C. & COSGROVE, L. J. 2008. An investigation of the ligand binding properties and negative cooperativity of soluble insulin-like growth factor receptors. *J Biol Chem,* 283**,** 5355-63.

TANG, J. G. & TSOU, C. L. 1990. The insulin A and B chains contain structural information for the formation of the native molecule. Studies with protein disulphide-isomerase. *Biochem J,* 268**,** 429-35.

TANIGUCHI, C. M., EMANUELLI, B. & KAHN, C. R. 2006. Critical nodes in signalling pathways: insights into insulin action. *Nat Rev Mol Cell Biol,* 7**,** 85-96.

TIMOFEEV, V. I., CHUPROV-NETOCHIN, R. N., SAMIGINA, V. R., BEZUGLOV, V. V., MIROSHNIKOV, K. A. & KURANOVA, I. P. 2010. X-ray investigation of gene-engineered human insulin crystallized from a solution containing polysialic acid. *Acta Crystallogr Sect F Struct Biol Cryst Commun,* 66**,** 259-63.

TOMINAGA, K., SHIMAMURA, T., KIMURA, N., MURAYAMA, T., MATSUBARA, D., KANAUCHI, H., NIIDA, A., SHIMIZU, S., NISHIOKA, K., TSUJI, E. I., YANO, M., SUGANO, S., SHIMONO, Y., ISHII, H., SAYA, H., MORI, M., AKASHI, K., TADA, K. I., OGAWA, T., TOJO, A., MIYANO, S. & GOTOH, N. 2017. Addiction to the IGF2-ID1-IGF2 circuit for maintenance of the breast cancer stem-like cells. *Oncogene,* 36**,** 1276-1286.

TORRES, A. M., FORBES, B. E., APLIN, S. E., WALLACE, J. C., FRANCIS, G. L. & 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.

UCHIKAWA, E., CHOI, E., SHANG, G., YU, H. & BAI, X. C. 2019. Activation mechanism of the insulin receptor revealed by cryo-EM structure of the fully liganded receptor-ligand complex. *Elife,* 8.

VAN LIEROP, B., ONG, S. C., BELGI, A., DELAINE, C., ANDRIKOPOULOS, S., HAWORTH, N. L., MENTING, J. G., LAWRENCE, M. C., ROBINSON, A. J. & FORBES, B. E. 2017. Insulin in motion: The A6-A11 disulfide bond allosterically modulates structural transitions required for insulin activity. *Sci Rep,* 7**,** 17239.

VELLA, V., NICOLOSI, M. L., GIULIANO, M., MORRIONE, A., MALAGUARNERA, R. & BELFIORE, A. 2019. Insulin Receptor Isoform A Modulates Metabolic Reprogramming of Breast Cancer Cells in Response to IGF2 and Insulin Stimulation. *Cells,* 8.

VERSTEYHE, S., KLAPROTH, B., BORUP, R., PALSGAARD, J., JENSEN, M., GRAY, S. G. & DE MEYTS, P. 2013. IGF-I, IGF-II, and Insulin Stimulate Different Gene Expression Responses through Binding to the IGF-I Receptor. *Front Endocrinol (Lausanne),* 4**,** 98.

WEIS, F., MENTING, J. G., MARGETTS, M. B., CHAN, S. J., XU, Y., TENNAGELS, N.,

WOHLFART, P., LANGER, T., MULLER, C. W., DREYER, M. K. & LAWRENCE, M. C. 2018. The signalling conformation of the insulin receptor ectodomain. *Nat Commun,* 9**,** 4420.

WEISS, M. A. 2009. The structure and function of insulin: decoding the TR transition. *Vitam Horm,* 80**,** 33-49.

WEISS, M. A. & LAWRENCE, M. C. 2018. A thing of beauty: Structure and function of insulin's "aromatic triplet". *Diabetes Obes Metab,* 20 Suppl 2**,** 51-63.

WEISS, M. A., NAKAGAWA, S. H., JIA, W., XU, B., HUA, Q. X., CHU, Y. C., WANG, R. Y. & KATSOYANNIS, P. G. 2002a. Protein structure and the spandrels of San Marco: insulin's receptor-binding surface is buttressed by an invariant leucine essential for its stability. *Biochemistry,* 41**,** 809-19.

WEISS, M. A., WAN, Z., ZHAO, M., CHU, Y. C., NAKAGAWA, S. H., BURKE, G. T., JIA, W., HELLMICH, R. & KATSOYANNIS, P. G. 2002b. Non-standard insulin design: structure-activity relationships at the periphery of the insulin receptor. *J Mol Biol,* 315**,** 103-11.

WHITE, A., LOUEY, S., CHANG, E. I., BOEHMER, B. H., GOLDSTROHM, D., JONKER, S. S. & ROZANCE, P. J. 2018. A 1 week IGF-1 infusion decreases arterial insulin concentrations but increases pancreatic insulin content and islet vascularity in fetal sheep. *Physiol Rep,* 6**,** e13840.

WHITTAKER, J., GROTH, A. V., MYNARCIK, D. C., PLUZEK, L., GADSBOLL, V. L. & 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, L., HAO, C. L., FU, W. & WHITTAKER, J. 2008. High-Affinity Insulin Binding: Insulin Interacts with Two Receptor Ligand Binding Sites. *Biochemistry,* 47**,** 12900-12909.

WILLIAMS, P. F., MYNARCIK, D. C., YU, G. Q. & 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.

XIONG, X., BLAKELY, A., KIM, J. H., MENTING, J., SCHAFER, I., SCHUBERT, H., AGRAWAL, R., GUTMANN, T., DELAINE, C. & ZHANG, Y. 2021. Visualization of insulin receptor activation by a novel insulin analog with elongated a chain and truncated b chain.

XIONG, X., MENTING, J. G., DISOTUAR, M. M., SMITH, N. A., DELAINE, C. A., GHABASH, G., AGRAWAL, R., WANG, X., HE, X., FISHER, S. J., MACRAILD, C. A., NORTON, R. S., GAJEWIAK, J., FORBES, B. E., SMITH, B. J., SAFAVI-HEMAMI, H., OLIVERA, B., LAWRENCE, M. C. & CHOU, D. H. 2020. Author Correction: A structurally minimized yet fully active insulin based on cone-snail venom insulin principles. *Nat Struct Mol Biol,* 27**,** 683.

XU, B., HUANG, K., CHU, Y. C., HU, S. Q., NAKAGAWA, S., WANG, S., WANG, R. Y., WHITTAKER, J., KATSOYANNIS, P. G. & WEISS, M. A. 2009. Decoding the cryptic active conformation of a protein by synthetic photoscanning: insulin inserts a detachable arm between receptor domains. *J Biol Chem,* 284**,** 14597-608.

XU, Y., KIRK, N. S., VENUGOPAL, H., MARGETTS, M. B., CROLL, T. I., SANDOW, J. J., WEBB, A. I., DELAINE, C. A., FORBES, B. E. & LAWRENCE, M. C. 2020. How IGF-II Binds to the Human Type 1 Insulin-like Growth Factor Receptor. *Structure,* 28**,** 786-798 e6.

XU, Y. B., KONG, G. K. W., MENTING, J. G., MARGETTS, M. B., DELAINE, C. A., JENKIN, L. M., KISELYOV, V. V., DE MEYTS, P., FORBES, B. E. & LAWRENCE, M. C. 2018. How ligand binds to the type 1 insulin-like growth factor receptor. *Nature Communications,* 9.

YAKAR, S., WERNER, H. & ROSEN, C. J. 2018. Insulin-like growth factors: actions on the skeleton. *J Mol Endocrinol,* 61**,** T115-T137.

YAN, H., GUO, Z. Y., GONG, X. W., XI, D. & FENG, Y. M. 2003. A peptide model of insulin folding intermediate with one disulfide. *Protein Sci,* 12**,** 768-75.

ZAINA, S. & NILSSON, J. 2003. Insulin-like growth factor II and its receptors in atherosclerosis and in conditions predisposing to atherosclerosis. *Curr Opin Lipidol,* 14**,** 483-9.

ZAKOVA, L., KAZDOVA, L., HANCLOVA, I., PROTIVINSKA, E., SANDA, M., BUDESINSKY, M. & JIRACEK, J. 2008. Insulin analogues with modifications at position B26. Divergence of binding affinity and biological activity. *Biochemistry,* 47**,** 5858-68.

ZAKOVA, L., KLETVIKOVA, E., VEVERKA, V., LEPSIK, M., WATSON, C. J., TURKENBURG, J. P., JIRACEK, J. & BRZOZOWSKI, A. M. 2013. Structural Integrity of the B24 Site in Human Insulin Is Important for Hormone Functionality. *Journal of Biological Chemistry,* 288**,** 10230-10240.

ZHA, J. P. & LACKNER, M. R. 2010. Targeting the Insulin-like Growth Factor Receptor-1R Pathway for Cancer Therapy. *Clinical Cancer Research,* 16**,** 2512-2517.

ZHANG, F., ALTINDIS, E., KAHN, C. R., DIMARCHI, R. D. & GELFANOV, V. 2021. A viral insulin-like peptide is a natural competitive antagonist of the human IGF-1 receptor. *Mol Metab,* 53**,** 101316.

ZHANG, Z., TANG, Y., YAO, S., ZHU, S. & FENG, Y. 2003. Protein engineering of insulin: Two novel fast-acting insulins [B16Ala]insulin and [B26Ala]insulin. *Sci China C Life Sci,* 46**,** 474-80.

ZIEGLER, A. N., FENG, Q., CHIDAMBARAM, S., TESTAI, J. M., KUMARI, E.,

ROTHBARD, D. E., CONSTANCIA, M., SANDOVICI, I., COMINSKI, T., PANG, K., GAO, N., WOOD, T. L. & LEVISON, S. W. 2019. Insulin-like Growth Factor II: An Essential Adult Stem Cell Niche Constituent in Brain and Intestine. *Stem Cell Reports,* 12**,** 816-830.

ZIEGLER, A. N., LEVISON, S. W. & WOOD, T. L. 2015. Insulin and IGF receptor signalling in neural-stem-cell homeostasis. *Nat Rev Endocrinol,* 11**,** 161-70.

ZIEGLER, A. N., SCHNEIDER, J. S., QIN, M., TYLER, W. A., PINTAR, J. E., FRAIDENRAICH, D., WOOD, T. L. & LEVISON, S. W. 2012. IGF-II promotes stemness of neural restricted precursors. *Stem Cells,* 30**,** 1265-76.