Structure and Interactions of the C-terminal Domain of Insulin-like Growth Factor Binding Protein-2 (IGFBP-2)

By

Zhihe Kuang, B.Sc, M.Sc

A Thesis submitted to The University of Adelaide, South Australia
in fulfilment of the requirements for the Degree of
Doctor of Philosophy

School of Molecular & Biomedical Science
The University of Adelaide
Adelaide, South Australia
October, 2006

Table of Contents

Acknowledgements

I sincerely wish to thank my supervisors, Professor John Wallace from The University of Adelaide, and Professor Ray Norton from The Walter and Eliza Hall Institute of Medical Research, for the opportunities they generously made available to me to undertake a PhD project in their laboratories, and their enormous guidance, support and encouragement throughout my candidature. I am also extremely grateful for the generosity and patience of Dr Briony Forbes and Dr Shenggen Yao in teaching me the wonders of protein science and NMR spectroscopy, and a double thanks to both for their great contributions to the project.

I would like to thank Ms Kerrie McNeil, Professor Leon Bach, Dr David Keizer, Ms Christina Wang, Dr Peter Güntert, Mr Julian Thompson, and Dr Francine Carrick, for their various contributions, and very kind instructions and assistance throughout the project. Many thanks are due to Dr Mark Hinds and Dr Grant Booker for instructions and assistance.

A large debt of thanks is owing to all the members in the Wallace laboratory, in particular Dr Steven Polyak, Ms Carlie Delaine, Dr Mehrnaz Keyhanfar, Dr Teerakul Arpornsuwan, Dr Adam Denley and Ms Lisa Bailey, for teaching and helping me all the time. Thanks also to other staff and students in the Biochemistry division, especially to Mr Chris Cursaro, Dr Anita Merkel, and Dr Kasper Kowalski. Thank you to all the members in the Norton laboratory, in particular Dr Zhi-Ping Feng, Dr Andrew Low, Dr Jeff Babon, Dr Indu Chandrashekaran and Ms Jennifer Sabo, and to Mr Jakub Szarlat and Mr John McFarlane from the WEHI IT services, for their very generous advice and assistance.

I especially wish to thank my parents, whose love and support I have always relied on.

I also acknowledge receipt of an International Postgraduate Research Scholarship (IPRS) from The University of Adelaide.

Statement of Originality

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

I give consent for a copy of this thesis, when deposited in the University Library, to be available for loan and photocopying.

Zhihe Kuang

October 11th, 2006

Abstract

Insulin-like growth factor binding protein-2 (IGFBP-2) is the largest member of a family of six proteins (IGFBP-1 to 6) that bind insulin-like growth factors-I and -II (IGF-I/II) with high affinities. IGFBP molecules contain three domains of approximately equal length: the conserved cysteine-rich amino- and carboxyl-terminal domains, which are joined by a variable linker domain. The C-terminal domains of IGFBPs not only contribute to high-affinity IGF binding, but also confer binding specificity and have overlapping but variable interactions with many other molecules. At the time this project commenced, there was limited information on the structure-function relationships of the C-domain of IGFBPs.

In this thesis, the following N- and C-domain fragments were prepared in sufficient quantities for NMR studies: unlabelled, ^{15}N-labelled and ^{15}N/^{13}C-labelled $^{183\text{-}289}$IGFBP-2 (C-BP-2), unlabelled $^{141\text{-}289}$IGFBP-2 (Large-C-BP-2), and unlabelled and ^{15}N-labelled $^{1\text{-}138}$IGFBP-2 (N-BP-2). The IGF binding abilities of these fragments were assessed using BIAcore and cross-linking methods. Overall, the results indicated that C-BP-2 binds IGFs to only a limited extent, although the differences in IGF binding affinities among C-BP-2, N-BP-2 and full-length IGFBP-2 appeared to be larger in cross-linking studies than in BIAcore experiments.

The solution structure of C-BP-2 was determined using NMR spectroscopy. C-BP-2 has a thyroglobulin type 1 fold comprising an α-helix, a three-stranded anti-parallel β-sheet and three flexible loops. Compared to the structures of C-BP-6 (Headey et al., 2004a), C-BP-1 (Sala et al., 2005) and C-BP-4 (Sitar et al., 2006), which were reported during the course of the current study, the following structural differences that may affect IGF binding and have implications for other functional differences were found: C-BP-2 has (i) a longer disordered loop I, and (ii) an extended C-terminal tail, which is unstructured and very mobile. (iii) The length of the helix is identical to that of C-BP-6 but shorter than that of C-BP-1. (iv) An RGD motif is located in a solvent-exposed turn. The backbone dynamics of C-BP-2 were analysed using the reduced spectral density mapping approach based on the measured backbone amide ^{15}N relaxation parameters (R_1, R_2 and steady-state ^{15}N-{^{1}H} NOE), and the results were compared to the dynamic properties of C-BP-6 (Yao et al., 2004). C-BP-2 possesses significant fast time-scale

motions in the loops and termini, and may also have slow time-scale conformational or chemical exchange in the structured domain core and the loop II.

A more complete set of assignments for IGF-I ^{1}H-^{15}N HSQC spectra was obtained. The IGF and N-BP-2 binding sites on C-BP-2 were identified by NMR. Binding of C-BP-2 to the IGF·N-BP-2 binary complexes was significantly stronger than the binding of C-BP-2 to IGFs alone, switching from intermediate exchange to slow exchange, indicating there is cooperativity between N-BP-2 and C-BP-2 in IGF binding. Two possible structural mechanisms for this cooperativity were found: a possible conformational change of the Phe49-Leu54 region and the sidechain aromatic ring of Phe49 in IGF-I, and direct interaction between the N- and C-domains.

A pH-dependent heparin binding site on C-BP-2 was also identified by NMR. The heparin binding site is a patch containing the β-turn connecting the first and second strands, part of the third strand, and the beginning of the C-terminal tail. Lys227, His228, Asn232, Leu233, Lys234 and His271 are proposed to be the primary heparin binding residues. Protonation of His271 and His228 seems to be important for the binding, which occurs at slightly acidic pH (6.0) and is more significant at pH 5.5, but is largely suppressed at pH 7.4. Possible preferential binding of IGFBP-2 and its C-domain fragments to glycosaminoglycans in the acidic extracellular matrix of tumours may be related to their roles in cancer.

Abbreviations

^{125}I-IGF-II	radio-iodinated IGF-II tracer
ALS	acid labile subunit
bIGFBP-2	bovine IGFBP-2
β–ME	2-mercaptoethanol
BMRB	BioMagResBank
BSA	bovine serum albumin
C-BP-2	C-terminal domain of IGFBP-2
cDNA	complementary deoxyribonucleaic acid
CIP	calf intestinal alkaline phosphatase
dNTPs	deoxyribonucleotide triphosphates
DTT	dithiothreitol
E.coli	*Escherichia coli*
ECM	Extracellular matrix
EDC	N-ethyl-N'-[9-dimethylamino)propyl]carbodiimide
EDTA	ethylenediaminetetraacetic acid
FID	free induction decay
GAG	glycosaminoglycans
GH	growth hormone
HEPES	4-(2-hydroxyethyl)-1-piperazine-ethanesulphonic acid
His_6-tag	hexahistidine tag
HSQC	Heteronuclear Single Quantum Coherence
IGFBP	Insulin-Like Growth Factor Binding Protein
IGF-I	insulin-like growth factor I
IGF-II	insulin-like growth factor II
IGF-1R	insulin-like growth factor I receptor
IPTG	isopropyl-β–D-thiogalactopyranoside
IR	insulin receptor
LB	Luria broth
MHC	major histocompatibility complex
MMPs	matrix metalloproteinase

N-BP-2	N-terminal domain of IGFBP-2
NHS	N-hydroxysuccinimide
NMR	nuclear magnetic resonance
NOE	Nuclear Overhauser Effect
NOESY	Nuclear Overhauser Effect Spectroscopy
OD_{600}	optical density at 600 nm
PAGE	polyacrylamide gel electrophoresis
PAPP-A	pregnancy-associated plasma protein-A
PBS	phosphate-buffered saline
PCR	polymerase chain reaction
PDB	Protein Data Bank
PEG	polyethyleneglycol
ppm	parts per million
HPLC	reverse phase high performance liquid chromatography
RU	resonance units
SDS	sodium dodecyl sulphate
SOS	sucrose octasulfate
TEV	tobacco etch virus
TFA	trifluoroacetic acid
TOCSY	total correlated spectroscopy
Tris	2-amino-2-hydroxymethylpropane-1,3-diol
TROSY	Transverse Relaxation Optimized Spectroscopy
UV	ultra-violet

List of Figures

List of Tables

List of Publications

This thesis is based on the following publications:

Kuang, Z., Yao, S., Keizer, D., Wang, C., Bach, L., Forbes, B., Wallace, J., and Norton, R. Structure, dynamics and heparin binding of the C-terminal domain of insulin-like growth factor-binding protein-2 (IGFBP-2). ***J Mol Biol.*** *In press.* (doi:10.1016/j.jmb.2006.09.006)

Kuang, Z., Yao, S., McNeil, K., Bach, L., Forbes, B., Wallace, J., and Norton, R. Cooperativity between IGFBP-2 N- and C-terminal domains in IGF binding. *Manuscript in preparation.*

And conference proceedings:

Kuang, Z., Wang, C., Yao, S., Forbes, B., Bach, L., Wallace, J., and Norton, R. (2005) Solution structure of the C-terminal domain of insulin-like growth factor binding protein-2 (IGFBP-2). Proceedings of the 30th Annual Lorne Protein Conference. (Cowes, Phillip Island, Australia). Poster.

Kuang, Z., Yao, S., Keizer, D., Wang, C., Forbes, B., Bach, L., Wallace, J., and Norton, R. (2005) Structure of the C-terminal domain of insulin-like growth factor binding protein-2 (IGFBP-2) and its binding sites on IGFs. Proceedings of the ComBio Conference. (Adelaide, SA, Australia). Poster.

Kuang, Z., Yao, S., Keizer, D., Wang, C., McNeil, K., Forbes, B., Bach, L., Wallace, J., and Norton, R. (2005) Structure of the C-terminal domain of insulin-like growth factor binding protein-2 (IGFBP-2) and interactions with IGFs: an NMR study. Proceedings of the 1st Asia-Pacific NMR Symposium. (Yokohama, Japan). Poster. (Selected for a Young Scientist Travel Award).

Chapter 1

Introduction and Literature Review

1.1 An Introduction to the IGF System

1.1.1 Components of the IGF system: overview

The insulin-like growth factor (IGF) system is an important growth regulatory system in vertebrates. The IGFs exert their actions via cell surface receptors, and their actions are tightly regulated by IGF binding proteins and other molecules. Proper functioning of the IGF system is essential for normal growth and development, and abnormalities of the system are implicated in many diseases.

As illustrated in Figure 1.1, principal components of the IGF system include two peptide hormones (IGF-I and IGF-II), their receptors, and six high affinity IGF binding proteins (IGFBP-1 through -6) (Cohick and Clemmons, 1993; Jones and Clemmons, 1995; Stewart and Rotwein, 1996; Collett-Solberg and Cohen, 2000; Firth and Baxter, 2002; Denley et al., 2005a). Downstream signalling pathways underlying the functions of the receptors are important parts of the IGF system, but are not the scope of this thesis and thus are not shown in detail in the diagram.

IGF-I and IGF-II are single-chain polypeptide hormones consisting of 70 and 67 residues, respectively. IGF-I, IGF-II, and insulin share high similarity in sequences and three-dimensional structures. Despite these significant similarities and certain cross reactivities between IGFs and insulin for IGF and insulin receptors, each ligand can lead to specific signalling outcomes and is not able to fully compensate the loss of another ligand.

Cell surface receptors for IGF-I and IGF-II include IGF-I receptor (IGF-1R), IGF-II receptor (IGF-2R), insulin receptor exon 11- (IR-A), and insulin receptor exon 11+ (IR-B). IGF-1R, IR-A and IR-B are four-subunit transmembrane tyrosine kinase receptors. IGF-I, IGF-II and insulin can bind to the extracellular regions of IGF-1R and activate its intracellular tyrosine kinase activity, resulting in cellular responses such as proliferation, differentiation, migration and anti-apoptosis (Cohick and Clemmons,

1993; Jones and Clemmons, 1995; Collett-Solberg and Cohen, 2000). The metabolic actions of insulin are generally mediated by the exon 11+ isoform of IR (IR-B), which has low affinity for IGFs. On the other hand, IGF-II can bind to the exon 11- isoform of IR (IR-A) with high affinity, and IR-A activation lead to mitogenic effects that are similar to those of IGF-1R (Denley et al., 2003). Furthermore, there are functional hybrid receptors formed between IGF-1R and either IR-A or IR-B, but their roles are still unclear (Denley et al., 2005a).

The IGF-2R is a 300 kDa protein consisting of 15 extracellular domains, a 23-residue transmembrane domain, and a 163-residue intracellular domain (Kornfeld, 1992; Braulke, 1999). IGF-2R is also known as the cation-independent mannose-6-phophate receptor, because it binds some proteins containing mannose-6-phosphate. Unlike IGF-1R and IR, IGF-2R has no signalling transduction capacity. IGF-2R sequesters IGF-II, resulting in internalisation and degradation of IGF-II (Kornfeld, 1992; Braulke, 1999).

IGFBP-1 to -6 are ~30 kDa proteins that bind IGFs with high affinity, and this family of proteins have been extensively reviewed (Jones and Clemmons, 1995; Rajaram et al., 1997; Hwa et al., 1999; Baxter, 2000; Duan, 2002; Firth and Baxter, 2002; Mohan and Baylink, 2002; Rosenzweig, 2004; Bach et al., 2005). IGFBPs transport IGFs in the circulation, and regulate the bioavailability and biological actions of IGFs. In addition, IGFBP have a broad spectrum of IGF-independent functions. IGFBP molecules contain conserved and cysteine-rich N- and C-terminal domains, connected by variable linker domains; both the N- and C-domains contribute to high affinity IGF binding. Proteolytic cleavage of IGFBPs, which mostly occur in the linker domains, and binding of IGFBPs to the extracellular matrix can decrease IGF binding affinity and thus release the IGFs. Posttranslational modifications of IGFBPs, including glycosylation and phosphorylation, can also alter their IGF binding affinity.

The IGF system also includes other molecules that interact with the abovementioned principal components and affect their activities, such as the acid labile subunit (ALS) and the proteases that cleave IGFBPs. Up to 70% of the IGFs in circulation is found in a ternary complex form with IGFBP-3 and ALS (Boisclair et al., 2001). During the past few years, an increasing number of other molecules have been found to interact with IGFBPs, especially via the C-domain of IGFBPs. These interactions either modulate IGF binding or mediate IGF-independent actions, and will be described in section 1.5.

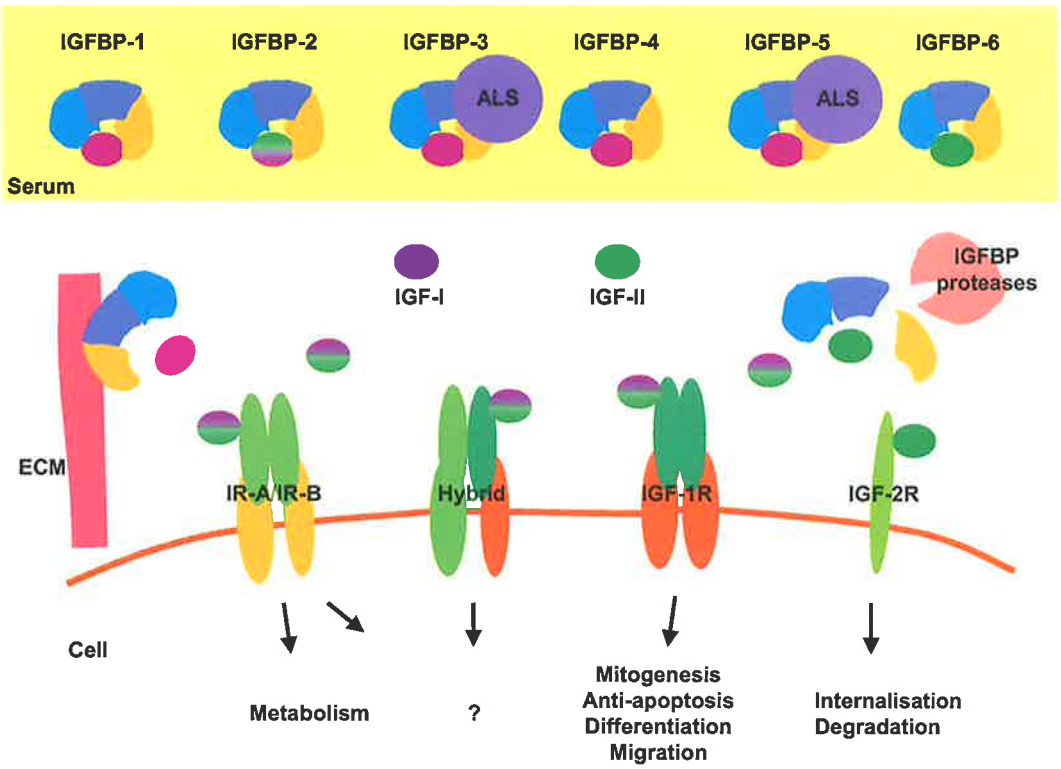


Figure 1.1 Schematic representation of the IGF system

Schematic representation of the IGF system. The IGF system consists of the receptors (IGF-1R, two IR isoforms, IGF-1R:IR hybrids and IGF-2R), the peptides (IGF-I, IGF-II, and insulin) and six high affinity IGF binding proteins (IGFBPs). IGFs circulate mainly in IGF:IGFBP-3:ALS complex. Release of IGFs from IGFBPs occurs upon IGFBP proteolysis or extracellular matrix (ECM) binding. This figure was adapted from (Denley et al., 2005a).

1.1.2 Molecular interactions of the IGF system

The IGFs, their receptors and the IGF binding proteins exert their biological functions through the protein-protein interactions between these molecule and the downstream signalling pathways of the receptors.

The binding affinities of IGF-I, IGF-II and insulin for different receptors and IGFBPs have been measured extensively, and were reviewed recently by Denley and co-workers (Denley et al., 2005a). It is worth noting that the binding affinity values for a particular interaction, measured by different groups using different assays, can vary significantly. Thus, the values of different protein interactions seem to be more comparable if they are measured in the same study.

In the IGF system, the proteins sharing sequence and structure similarities, such as IGF-I and IGF-II, or six IGFBPs, have overlapping interacting targets. On the other hand, differences in the amino acid sequences may result in differences of the binding sites and thus binding abilities. Therefore, different ligands do have binding preferences to different receptors or binding proteins, and vice versa.

Overall, IGF-I and IGF-II have comparable or higher binding affinities (K_D ~0.1-10 nM) for IGFBPs than for their receptors, whereas insulin does not bind IGFBPs. IGFBP-6 has markedly higher (~20 to 100-fold) binding affinity for IGF-II than for IGF-I, while IGFBP-2 has slightly higher (~2-fold) binding affinity for IGF-II than for IGF-I.

IGF-I has ~1000-fold higher binding affinity for IGF-1R (ED_{50} ~0.2 nM) than for both IR-A and IR-B, but IGF-II has comparable IGF-1R and IR-A binding affinities (ED_{50} ~0.6-0.9 nM), which are ~10-fold higher than its IR-B binding affinity. IGF-I has ~3-fold higher IGF-1R binding affinity compared with IGF-II, whereas IGF-II has ~100-fold or ~30-fold higher IR-A or IR-B binding affinities compared with IGF-I. IGF-II has >1000-fold higher binding affinity for IGF-2R, compared with IGF-I. On the other hand, insulin has >500-fold higher binding affinity for IRs than for IGF-1R, and insulin does not bind IGF-2R.

Understanding the structural determinants of these protein interactions is of great interest, and manipulation of the interactions based on such knowledge would have important implications in areas such as cancer therapy. Numerous studies have therefore been conducted to identify the important residues in IGFs, their receptors, and

IGFBPs for these binding interactions. Some of these findings will be reviewed in sections 1.3 and 1.4.

Furthermore, molecular interactions among these proteins can be either enhanced or reduced by certain mechanisms. For example, the interactions may involve more than two molecules, either forming more stable multiple protein complexes, or resulting in inhibition by competing a common binding surface. The interactions can also result in a conformational change of the protein structure and thus either facilitate or impair other protein interactions. Such dynamic modulations of protein interactions appear to play important roles in the interactions between IGFs and IGFBPs, which are the subjects of this study, but were previously less understood.

1.1.3 Expression and regulation of IGFs and IGFBPs

IGFs are growth-promoting peptides functioning in endocrine, paracrine and autocrine manners (Cohick and Clemmons, 1993; Jones and Clemmons, 1995; Collett-Solberg and Cohen, 2000). The *IGF1* gene is located in chromosome 12 in humans. IGF-I plays an important role in both prenatal and postnatal development, and its expression is higher during puberty and adulthood before decreasing in the elderly. Systemically circulating IGF-I is mainly generated by the liver under the control of growth hormone (GH) released by the pituitary gland. Many other adult organs, including kidney, lung, and bone, also produce IGF-I. The *IGF2* gene is located in chromosome 11 in humans. Interestingly, there is significant difference in the IGF-II expression pattern between mouse and human. In the mouse, IGF-II is expressed during embryogenesis but its expression declines rapidly after birth and only remains minimally in some tissues of the brain. However, in humans, IGF-II is expressed at a high level in the foetus and continues to be expressed and secreted by the liver after birth, and in fact its concentration in serum is higher than that of IGF-I. IGF-II expression is not regulated by GH, and its regulation is less understood.

In humans, both *IGFBP1* and *IGFBP3* genes are located in chromosome 7, whereas both *IGFBP2* and *IGFBP5* genes are located in chromosome 2, and *IGFBP4* and *IGFBP6* genes are located in 17 and 12, respectively (Bach and Parker, 2003). IGFBPs are produced by a variety of biological tissues and can be found in various biological fluids (Rajaram et al., 1997). *IGFBP2* and *IGFBP6* genes lack the TATA

box. Different functional *cis* elements have been found in the promoter regions of IGFBP genes, suggesting that variable expression regulations exist for expression of IGFBP proteins (Bach and Parker, 2003). Basal *IGFBP2* gene expression is regulated by transcription factor Sp1 (Boisclair et al., 1993). *IGFBP2* gene contains putative nuclear factor κB binding sites and hypoxia response elements (HREs), and its expression is increased by NF-κB (Cazals et al., 1999). Interestingly, serum IGFBP-2 levels are increased in transgenic mice overexpressing IGF-II, suggesting a regulation of IGFBP-2 expression by IGF-II (Wolf et al., 1994; Blackburn et al., 1997).

1.1.4 Biological actions of IGFs

Despite the high similarities in both ligands and receptors between IGFs and insulin, insulin mainly exerts metabolic actions via the insulin receptor IR-B, whereas cellular actions of IGF-I and -II via IGF-1R are mainly mitogenic.

IGFs stimulate cell proliferation by increasing DNA, protein, and carbohydrate synthesis and progression of cell cycles (Cohick and Clemmons, 1993; Jones and Clemmons, 1995; Collett-Solberg and Cohen, 2000). IGFs also have potent anti-apoptotic effects. Forbes and co-workers showed that, the binding affinities of IGF-I and -II analogues for IGF-1R correlate with their protective effects against serum withdrawal-induced apoptosis of rat pheochromocytoma cells (Forbes et al., 2002). In addition, IGFs can stimulate terminal differentiation in some cell types, including myoblasts, chondrocytes, and neural cells, as well as regulate other hormone secretion, and induce chemotactic migration of a variety of cells (Cohick and Clemmons, 1993; Jones and Clemmons, 1995; Collett-Solberg and Cohen, 2000).

Transgenic mice overexpressing human IGF-I had a ~30% increase in body weight, which was not apparent until four weeks after birth. GH expression was suppressed, reflecting a negative feedback. Circulating IGF-I levels were elevated but were significantly lower than that in the GH transgenic mice (Mathews et al., 1988).

IGF-I knockout mice and IGF-II knockout mice had similar prenatal phenotypes, and both had 60% of normal birth weight. However, survival was normal in the mice lacking IGF-II, but IGF-I knockout mice had a marked increase in neonatal death rate. IGF-I and -II double knockout mice had only 30% of normal birth weight, and all the mice died within minutes of birth (Baker et al., 1993).

IGF-1R knockout mice had 45% of normal birth weight, and also died immediately after birth (Liu et al., 1993). Interestingly, IGF-I and IGF-1R double knockout mice were identical to those lacking receptor only, whereas IGF-II and IGF-1R double knockout mice were more severely affected and were identical to IGF-I and IGF-II double knockout mice (Liu et al., 1993).

1.1.5 Biological functions of IGFBPs

Recently it has been shown that IGFBPs are multifunctional proteins. Their biological functions, which have been reviewed extensively (Rajaram et al., 1997; Baxter, 2000; Duan, 2002; Firth and Baxter, 2002; Mohan and Baylink, 2002), can be divided into two major groups, i.e., IGF-dependent functions and IGF-independent functions.

IGFBPs were identified as IGF carrier proteins in the circulation. More than 99% of the IGFs in the circulation or the extra-cellular space is associated with IGFBPs. Binding to IGFBPs prolongs the half-life of IGFs, so IGFBPs effectively sequester IGF for storage and potential endocrine function (Gulve and Dice, 1989). Furthermore, IGFBPs modulate IGF actions by influencing the amount and localization of free IGF molecules that are able to interact with the receptors. Binding to IGFBPs inhibits IGF actions, because IGF-IGFBP complexes cannot activate the IGF-I receptor (Firth and Baxter, 2002). In certain situations, however, IGFBPs can also enhance IGF actions by some incompletely understood mechanisms, such as co-localising the IGF molecules to the receptor-abundant cell surface and promoting the subsequent release of IGFs (Firth and Baxter, 2002).

In addition to both inhibitory and stimulatory IGF actions, IGFBPs exert functions that are independent of the binding of IGFs and the activation of IGF receptors by IGFs. The first report of such functions was the stimulatory effect of IGFBP-1 on CHO cell migration, mediated by the RGD motif in its C-domain (Jones et al., 1993b). An homologous RGD motif is present in IGFBP-2 but not IGFBP-3 to -6, and similar effects of IGFBP-2 via integrins have been confirmed recently (Schutt et al., 2004; Wang et al., 2006). Many of the IGF-independent functions reported so far are of IGFBP-3 and -5. It was reported that these IGFBPs inhibit breast cancer cell growth and potentiate apoptosis by modulating the expression of Bax and Bcl-2, but

involvement of IGFs and their receptors was not excluded in those studies (Butt et al., 2000). IGFBP-3 and -5 were found to translocate into the nucleus and induce apoptosis in some cell types (Lee and Cohen, 2002). Some molecular interactions are involved in this function, which will be reviewed in section 1.5.6. The N-domain of IGFBP-5 has transactivation activity mediated by several conserved acidic residues. IGFBP-2 and -3, but not -1 or -6 also have strong transactivation activity (Zhao et al., 2006). IGFBP-5 and IGFBP-3 were shown to have growth-inhibitory effects in IGF-1R knockout mice fibroblast cells (Valentinis et al., 1995), whereas recombinant human IGFBP-5 stimulated bone formation in IGF-I knockout mice (Miyakoshi et al., 2001). Recently, it is found that IGFBP-5 directly activates tissue plasminogen activator (tPA) to activate plasminogen, resulting in cell death induced by cleavage of focal adhesions, and this ability is independent on IGF-I or heparin (Sorrell et al., 2006). IGF-independent functions of IGFBP-4 and -6 are less well established.

IGFBP knockout mice have few phenotypic manifestations. For example, mice lacking IGFBP-2 did not show large phenotypic alterations, with a ~30% reduction of spleen weight of adult males being the major morphological difference (Pintar et al., 1996). Combinatory knockouts of up to three of the binding proteins did not have significant effect, which is probably due to the functional compensation by other IGFBPs (Pintar et al., 1999). In contrast, transgenic mice overexpressing IGFBPs show generalized or localized growth retardation, which would be expected from inhibition of IGF actions (Schneider et al., 2000; Silha and Murphy, 2002). In addition, impaired glucose homeostasis and fertility have been observed in some IGFBP overexpressing mice but less consistently. Interestingly, these IGFBP overexpression mice have not suggested an important physiological role for the IGF-independent functions. However, such functions, if any, may have been masked or overlooked in the data (Schneider et al., 2000; Silha and Murphy, 2002).

1.1.6 Implications of the IGF system in health and diseases

The IGF system is a major regulator of both prenatal and postnatal somatic growth and cellular proliferation (Monzavi and Cohen, 2002). In addition, recent studies have demonstrated an important role of the insulin/IGF signalling pathways in the control of

longevity; disruption of these pathways significantly extended lifespan in several animal models (Janssen and Lamberts, 2004; Bartke, 2005; Katic and Kahn, 2005).

In humans, dysregulation of the GH/IGF axis can lead to growth disorders. Growth hormone insensitivity syndrome (GHIS, or Laron syndrome) is caused by GH resistance and IGF-I deficiency, whereas over-production of GH and consequently of IGF-I can lead to acromegaly (Monzavi and Cohen, 2002). Furthermore, IGFs have hypoglycaemic and insulin-sensitising actions, and dysregulation of these actions is involved in diabetes and its complications (Monzavi and Cohen, 2002). Administration of IGF-I has been the subject of trials, for the treatment of GHIS and diabetes patients (Monzavi and Cohen, 2002; Savage et al., 2004; Rosenfeld, 2005), and displacement of IGF from IGFBPs using IGF analogues (Loddick et al., 1998) was proposed to be a neuroprotective treatments for stroke.

Even more importantly, the role of the IGF system in tumourigenesis has become a topic of great interest in cancer research (Furstenberger and Senn, 2002; Denley et al., 2003; LeRoith and Roberts, 2003; Pollak et al., 2004; Foulstone et al., 2005; Yakar et al., 2005). High serum concentrations of IGF-I and low serum concentrations of IGFBP-3 are associated with an increased risk of breast, prostate, colon, and lung cancers. IGF-1R is commonly, although not always, overexpressed in many tumours. Most primary tumours and transformed cell lines overexpress IGF-II, which may act in autocrine and paracrine manners. Overexpression of IGF-II often results from loss of imprinting (LOI) of the *IGF2* gene (Christofori et al., 1995; Okamoto et al., 1997), which has been found to be a potential marker of colorectal cancer risk (Cui et al., 2003). Mutation in the IGF-2R gene occurs in some cancers (De Souza et al., 1995). Several cancer cell types overexpressing IGF-II preferentially express IR-A, which has a high affinity for IGF-II, and activation of IR-A by IGF-II results in mitogenic effects and cancer progression (Denley et al., 2003). Interestingly, IGFBP-2 appears to play paradoxical roles in tumourigenesis, which will be reviewed in section 1.2.2. Targeting the IGF system as an approach to cancer therapy is currently underway (Perks and Holly, 2003; Pollak et al., 2004; Yakar et al., 2005; Yee, 2006). In particular, tyrosine kinase inhibitors and IGF-1R antibodies, as well as the IGF antibodies, recombinant IGFBPs and soluble IGF-2R proteins are being developed to block IGF binding to IGF-1R or IGF-1R signalling.

1.2 Biological Actions of IGFBP-2

1.2.1 IGF-dependent and -independent actions

In terms of IGF-dependent actions, IGFBP-2 generally inhibits IGF actions, especially those of IGF-II, possibly reflecting its slightly higher IGF-II binding affinities (Jones and Clemmons, 1995; Firth and Baxter, 2002). Overexpression of IGFBP-2 in human embryonic kidney fibroblasts resulted in inhibition of cell proliferation, which can be reversed by addition of exogenous IGFs (Hoflich et al., 1998). IGFBP-2 also inhibits IGF-I actions *in vivo*, since crossing with IGFBP-2 transgenic mice significantly reduced the growth of the giant GH transgenic mice, which had 2- to 3-fold higher serum IGF-I levels (Hoeflich et al., 2001a). Purified human IGFBP-2 inhibited IGF-II-stimulated DNA synthesis (Reeve et al., 1993).

On the other hand, stimulatory effects of IGFBP-2 have also been reported. Recombinant porcine IGFBP-2 increased both basal and IGF-induced proliferation of porcine endometrial glandular epithelial cells, and both IGF-dependent and independent stimulations were possible in that model (Badinga et al., 1999). Overexpression of IGFBP-2 in mouse adrenocortical tumour Y-1 cells increased cell proliferation, possibly in an IGF-independent manner (Hoeflich et al., 2000).

Recent reports showed that integrin interaction by IGFBP-2, mediated by the RGD motif in its C-domain, is an important IGF-independent action of IGFBP-2 that can affect adhesion, migration, and proliferation in some cell types (Jones et al., 1993b; Schutt et al., 2004; Wang et al., 2006). This will be reviewed in detail in section 1.5.5.

Furthermore, there has been evidence suggesting some novel biological functions of IGFBP-2 under certain circumstances, but so far they are unclear and future investigations are required. IGFBP-2 and a C-domain fragment were found in peri/nuclear fractions of various tissues isolated from IGFBP-2 transgenic and non-transgenic mice (Hoeflich et al., 2004). Increased intracellular colocalization and interaction of IGFBP-2 with the cyclin-dependent kinase inhibitor $p21^{CIP1/WAF1}$ was found during serum deprivation growth inhibition of mouse lung epithelial cells (Terrien et al., 2005).

1.2.2 Roles of IGFBP-2 in Cancer

Increased serum and tissue IGFBP-2 levels were found in patients with many malignancies, such as breast, prostate, colon, lung, CNS, adrenal glands, and ovary cancers (Hoeflich et al., 2001b). A positive correlation between the tumour grade and level of IGFBP-2 expression has also been described in some of these cancers, in contrast to some other IGFBPs such as IGFBP-3 (Hoeflich et al., 2001b). IGFBP-2 is expressed at high level only in glioblastoma and not in mid-grade or low-grade gliomas, and strong expression is associated with poor patient survival in diffuse gliomas (Zhang et al., 2002). Highly elevated serum IGFBP-2 and low IGFBP-3 at diagnosis of childhood acute lymphoblastic leukaemia are correlated with a higher risk of lack of remission or a relapse (Vorwerk et al., 2005).

Nevertheless, whether IGFBP-2 promotes or protects against tumour progression is unclear, because both inhibitory and stimulatory effects of IGFBP-2 on cancer cells have been reported (Hoeflich et al., 2001b). This probably reflects the fact that, as mentioned above, in addition to both inhibition and enhancement of IGF actions, IGFBP-2 exerts IGF-independent functions, and both IGF-dependent and IGF-independent activities vary according to the system under examination (Hoeflich et al., 2001b).

1.3 Structure of the IGFs, and receptor and IGFBP binding determinants

1.3.1 Structure of IGFs

As shown in Figure 1.2, IGF-I and IGF-II consist of 70 and 67 amino acid residues, respectively. They can be divided into four domains, designated B (IGF-I: Gly1-Thr29; IGF-II: Ala1-Phe28), C (IGF-I: Gly30-Thr41; IGF-II: Ser29-Arg40), A (IGF-I: Gly42-Cys62; IGF-II: Gly41-Ala61), and D (IGF-I: Ala63-Ala70; IGF-II: Thr62-Glu67), from the N-terminus (Rotwein, 1991). The sequence conservation between IGF-I and IGF-II is high, with 70% of the B and A domain residues of human IGF-I and IGF-II being identical to each other and 50% of the B and A domain residues are identical to the

corresponding residues of insulin (Rotwein, 1991). The C domain of IGFs is analogous to the C domain of mammalian proinsulin, which is cleaved during the processing and is not present in the mature insulin, while the D domain at the C-terminus of IGFs has no insulin equivalents. The prepropeptides of IGFs also contain an N-terminal signal sequence and two alternative forms of a C-terminal E domain that result from alternative mRNA splicing; both the signal peptide and the additional E domains are cleaved during post-translational processing (de Pagter-Holthuizen et al., 1986).

Initial structural models of the IGFs were proposed by Blundell and co-workers, based on the crystal structures of insulin (Blundell et al., 1978; Blundell et al., 1983). Since then, solution structures of IGF-I (Cooke et al., 1991; Sato et al., 1993) and IGF-II (Terasawa et al., 1994; Torres et al., 1995), IGF-I in complex with a peptide derived from phage display (Schaffer et al., 2003), as well as crystal structures of IGF-I in the presence of detergent deoxy big CHAPS (Vajdos et al., 2001) or SB12 (Brzozowski et al., 2002) have been determined. More recently, crystal structures of IGF-I in complex with an N-terminal fragment of human IGFBP-5 (mini-N-BP-5, Ala40 to Ile92) (Zeslawski et al., 2001), the whole N-domain of IGFBP-4 (N-BP-4) (Siwanowicz et al., 2005), and both N- and C-domains of IGFBP-4 (Sitar et al., 2006) have been solved. Both IGF-I and IGF-II have similar structures to that of insulin in well-defined equivalent regions (Figure 1.2). The global structure of IGF-I is maintained both free and in complex with detergent or N-domain of IGFBPs, although certain structural differences are evident in some regions.

Both IGF-I and -II contain three helices, stabilised by three disulfide bonds that are essential for structural integrity. The B domain consists of a short flexible stretch followed by the first α-helix (B1) (IGF-I: Ala8-Cys18; IGF-II: Gly11-Cys21), β-turn (IGF-I: Gly19-Gly22; IGF-II: Gly22-Gly25) and a loop (IGF-I: Phe23-Asn26; IGF-II: Phe26-Ala32). The C domain (IGF-I: Gly30-Thr41; IGF-II: Ser33-Arg40) connecting the B and A domains is a flexible region in the free form. The A domain contains two antiparallel α-helices, the A1-helix (IGF-I: Gly42-Cys48; IGF-II: Ile42-Arg49) and the A2-helix (IGF-I: Leu54-Cys61; IGF-II: Leu53-Tyr59), stabilized by disulfide bonds (IGF-I: Cys47-Cys52; IGF-II: Cys46-Cys51) and separated by structured loop regions. The A1 and A2 helices are stabilized with the B helix by one disulfide bond linking the beginning of the B-helix to the C-terminal end of the A1-helix (IGF-I: Cys6-Cys48 and IGF-II: Cys9-Cys47) and another connecting the C-terminal end of the B-helix to the

C-terminal end of the A2-helix (IGF-I: Cys18-Cys61 and IGF-II: Cys21-Cys60). The D domain (IGF-I: Pro63-Ala70; IGF-2: Thr62-Glu67) following the A2 helix and comprising the C-terminus of the molecules is flexible.

The solution structures of IGF-II, which are the only available structures of IGF-II so far, are relatively divergent among the ensemble of structures, probably resulting from the flexibility of the molecule together with low NMR spectral quality due to aggregation (Terasawa et al., 1994; Torres et al., 1995). Major differences between IGF-I and IGF-II occur at the start of the third helix and in flexible regions including the C domain and the N- and C-termini. Structures of IGF analogues have also been studied. For example, long-[Arg3]-IGF-I, which is a potent IGF-I analogue with lower binding affinity for IGFBPs, has an essentially identical tertiary fold compared with the native protein (Laajoki et al., 2000).

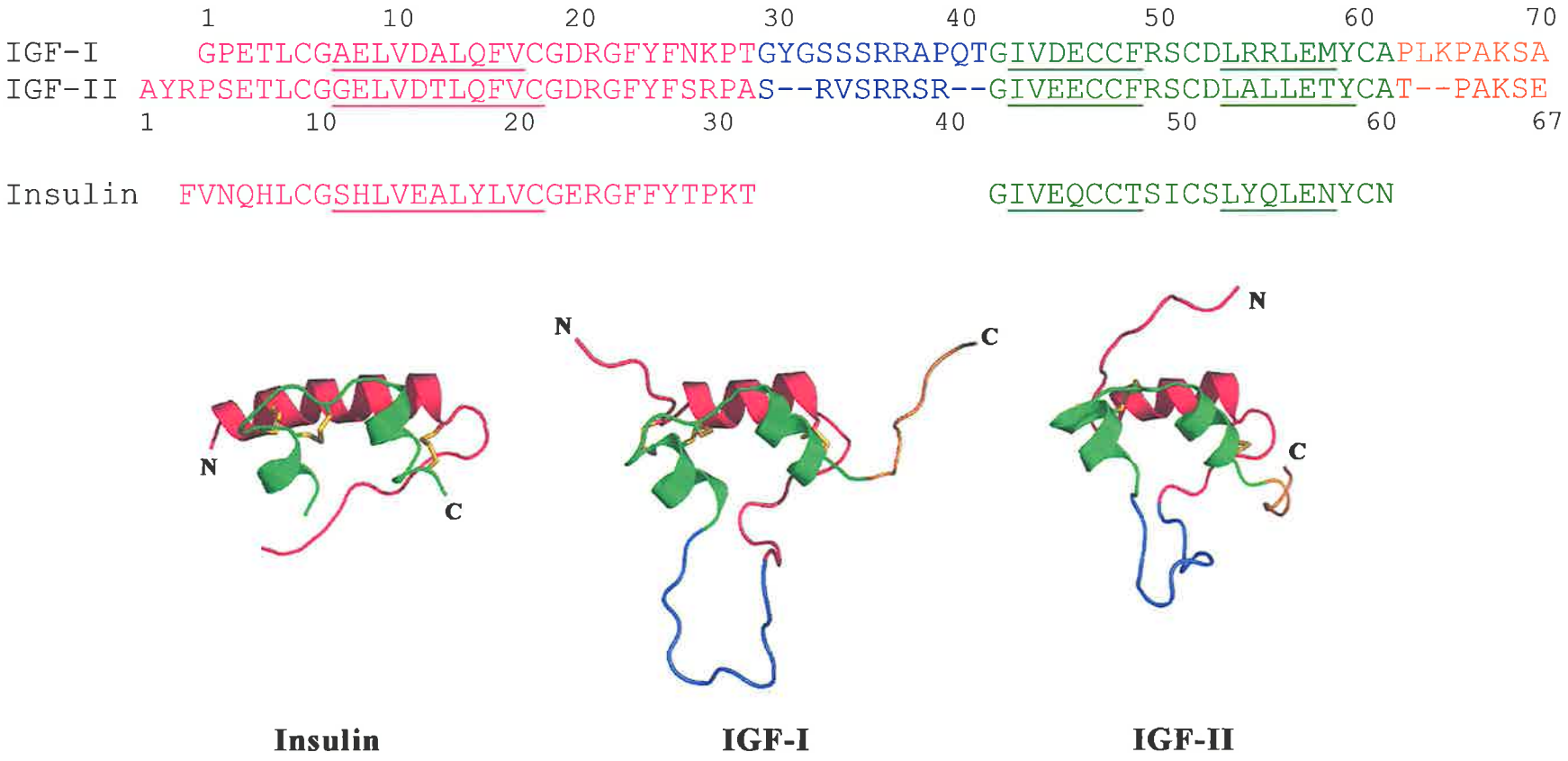


Figure 1.2 Sequences and structures of IGF-I, IGF-II and insulin

Sequential alignments of human IGF-I, IGF-II, and insulin, and ribbon views of IGF-I (PDB 1PMX (Schaffer et al., 2003)), IGF-II (PDB 1IGL (Torres et al., 1995)), and insulin (PDB 1ZEH (Ciszak et al., 1995)) structures. Residue number for IGF-I and IGF-II are given. B, C, A, and D domains are colored pink, blue, green, and orange, respectively. The underlined residues form helical regions.

1.3.2 The IGF-1R, IGF-2R, IR-A and IR-B binding residues of IGFs

IGF residues that are important for receptor binding have been reviewed in detail recently (Carrick et al., 2002; Denley et al., 2005a). In this section, these residues are summarized according to the domains of IGFs. Emphasis is placed the IGF-1R binding residues, which are also mapped onto the IGF-I surface in Figure 1.3, while residues for IGF-2R, IR-A, and IR-B binding are mentioned briefly.

Some B domain residues are important for IGF-1R binding. While deletions of the first three residues of IGF-I or first six residues of IGF-II did not affect IGF-1R binding, des(1-4) and des(1-5) IGF-I (Bagley et al., 1989; Francis et al., 1993) and des(1-7) and des(1-8) IGF-II (Roth et al., 1991; Hashimoto et al., 1995) had a large reduction in IGF-IR affinity. Ala8, Glu9, Val11, Arg21, Phe23, and Tyr24 are important for IGF-IR binding (Cascieri et al., 1988; Bayne et al., 1990; Hodgson et al., 1995; Shooter et al., 1996), while Ala substitutions of Val11, Asp12, Gln15, and Phe16 may have altered the B helix resulting in reduced IGF-1R binding (Jansson et al., 1997). Corresponding residues of IGF-II, including Phe26 and Tyr27, seem to form a similar surface for IGF-1R binding (Roth et al., 1991; Sakano et al., 1991). Phe26 and Tyr27 of IGF-II are also important for IGF-2R binding (Roth et al., 1991; Sakano et al., 1991). Residues Thr4, Ala8, Glu9, Val11, Phe23 and Tyr24 in IGF-I, and corresponding residues in IGF-II, are important for IR binding (Sakano et al., 1991; Shooter et al., 1996).

The C domain of IGF-I and -II is important in IGF-1R binding. Replacement of the entire C domain by a $(Gly)_4$ linker reduced the IGF-1R binding up to 30-fold (Bayne et al., 1989). Arg36, Arg37, and Tyr31 are important for the binding (Bayne et al., 1990; Zhang et al., 1991; Jansson et al., 1998). Recently, Denley and co-workers showed that C domain of IGF-I contributes to the IGF-1R binding preference of IGF-I over IGF-II (Denley et al., 2004). On the contrary, C domain residues of IGF-I do not seem to contribute to IGF-2R binding (Bayne et al., 1989; Bayne et al., 1990). Interestingly, removal of IGF-I C domain, and mutations of Tyr31, or Arg36 and Arg37, increased IR binding affinity of IGF-I (Bayne et al., 1989; Bayne et al., 1990; Zhang et al., 1991). The C domain of IGF-I and -II also contributes to the IR isoform binding difference between IGF-I and -II (Denley et al., 2004).

IGF-I A domain residues Val44, Met59, Tyr60 and Leu62 are important for IGF-1R binding (Bayne et al., 1990; Shooter et al., 1996; Denley et al., 2005b). In particular,

The Val44Met mutation was found to cause severe mental and growth retardation in a patient (Walenkamp et al., 2005). Although this mutant maintains a native fold, it has impaired IGF-1R and IR binding ability (Denley et al., 2005b). A domain residues that have been shown to be unimportant for IGF-1R binding include Phe49, Arg50, Ser51, Arg55, and Arg56 of IGF-I (Jansson et al., 1998) and Ala54 and Leu55 of IGF-II (Forbes et al., 2001). Interestingly, IGF-II A domain residues, including Phe48, Arg49, and Ser50 form a major IGF-2R binding site (Sakano et al., 1991). Substitutions of IGF-1 A domain residues by corresponding residues in insulin significantly reduced IGF-2R binding (Cascieri et al., 1989), whereas substitutions of IGF-II Ala54 and Leu55 by corresponding residues in IGF-I, Arg54 and Arg55, decreased IGF-2R binding (Forbes et al., 2001). Substitutions of A domain residues of IGF-I by the corresponding residues of insulin resulted in increased IR binding affinity (Cascieri et al., 1989). IGF-I residues Val44 (Denley et al., 2005b) and Tyr60 (Bayne et al., 1990) are very important for IR binding.

The contribution of the D domain of IGF-I and -II to IGF-1R binding is less clear. Deletion of the D domain in IGF-I did not affect IGF-1R binding in one study (Bayne et al., 1989), but in another study Ala substitutions of Lys65 and Lys68 decreased the IGF-1R affinity up to 10-fold (Zhang et al., 1991). Deletion of the D domain in IGF-II reduced IGF-1R binding 5-fold (Roth et al., 1991; Sakano et al., 1991). Swapping the D domain of IGF-I to IGF-II slightly increased IGF-1R binding affinity for IGF-IDI and slightly reduced that for IGF-IIDI (Denley et al., 2004). Deletion of the D domain in IGF-II reduced its IGF-2R binding by two-fold (Roth et al., 1991; Sakano et al., 1991), whereas deletion of the D domain in IGF-I did not affect its IGF-2R binding and slightly increased its IR binding affinity (Bayne et al., 1989). The D domain also plays a minor role in IR binding difference between IGF-I and -II (Denley et al., 2004).

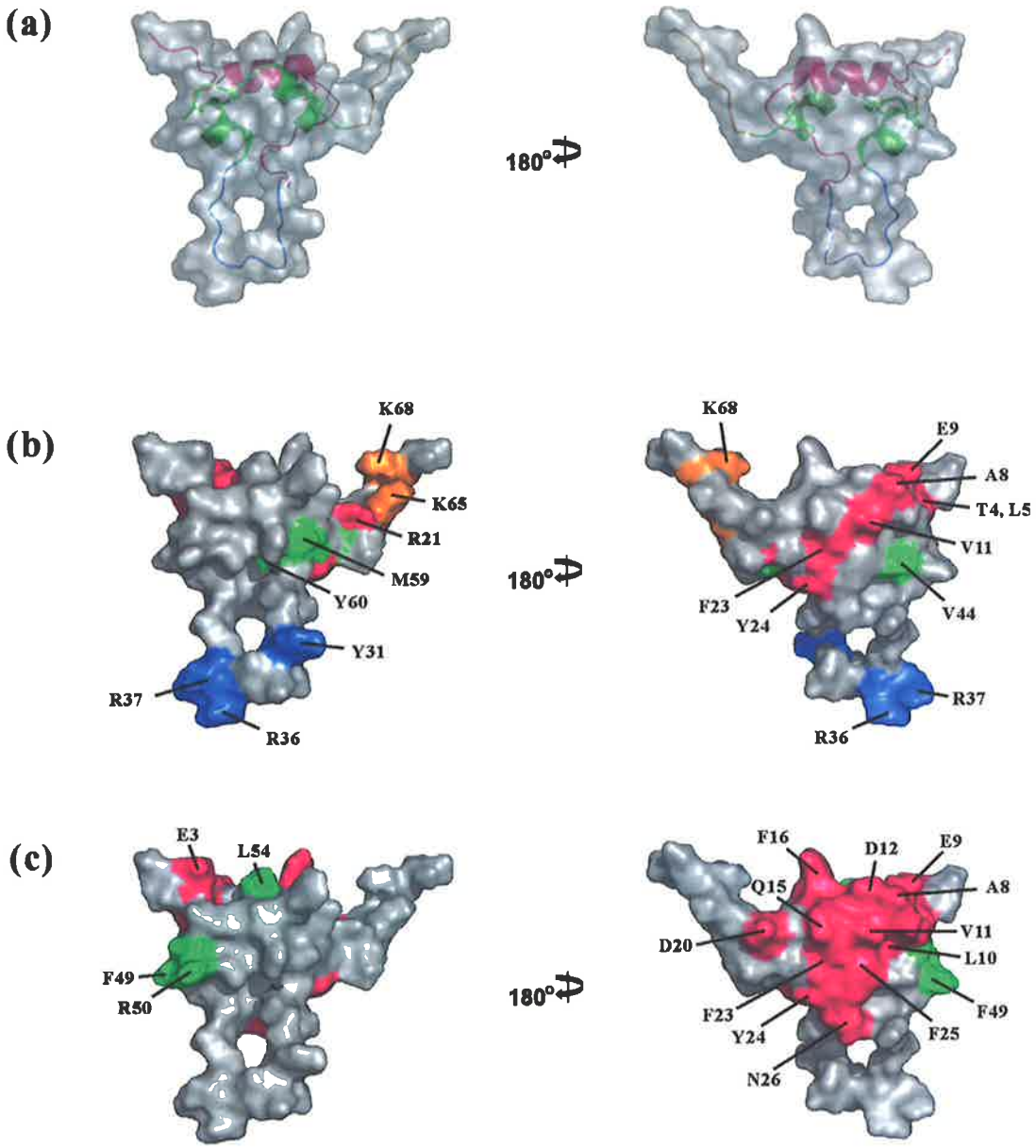


Figure 1.3 IGF-1R and IGFBP binding residues on IGF-I

(a) Transparent surface model of IGF-I with ribbon view inside. (b) IGF-I residues important for IGF-1R binding are colored according to their domain location (Figure 1.2). (c) IGF-I residues important for IGFBP binding are colored according to their domain location.

1.3.3 The IGFBP binding residues of IGFs

The interactions between IGFs and IGFBPs have been investigated extensively. Previous studies, using mutagenesis as well as other approaches, have identified specific residues of IGFs that are involved in IGFBP binding.

In the light of the fact that insulin has high similarities in primary sequence and three-dimensional structure compared to IGFs but is incapable of binding IGFBPs, Swapping of homologous domains, and substitution of residues between insulin and IGFs have been used to identify IGF residues that contribute to IGFBP binding (De Vroede et al., 1985; Cascieri et al., 1988; Clemmons et al., 1992; Oh et al., 1993). Interestingly, although the mature insulin molecule does not have counterparts of the IGF C and D domains, the major IGFBP binding site on IGFs does not include these domains. On the contrary, switching the B domain between insulin and IGF-I conferred IGFBP binding ability on insulin and reduced IGFBP binding of IGF-I (De Vroede et al., 1985; Cascieri et al., 1988). Substitution of residue Phe49, Arg50, and Ser51 in IGF-I (Clemmons et al., 1992; Oh et al., 1993) and Phe48, Arg49 and Ser50 in IGF-II (Bach et al., 1993) with corresponding residues in insulin dramatically decreased IGFBP binding affinity.

A variant form of IGF-I lacking the first three residues, des(1-3)IGF-I, isolated from bovine colostrum, has reduced IGFBP binding affinity (Francis et al., 1988), as well as an increased stimulatory effect on protein synthesis in L6 rat myoblast cells (Ballard et al., 1987). The importance of Glu3 of IGF-I in binding was confirmed in sequential deletions from the N-terminus (Bagley et al., 1989). The corresponding residue in IGF-II, Glu6, is also required for IGFBP binding (Francis et al., 1993).

Another two B domain residues, Gln15 and Phe16, are also involved in IGFBP binding identified by alanine screening and point mutation analysis. When mutated in combination with Glu3 and Thr4, binding to serum IGFBP is decreased by up to 600 fold (Ballard et al., 1987). Although Gln15 and Phe16 were shown to be important for structural integrity (Jansson et al., 1997), mutants that did not have grossly altered structures still exhibited reduced affinity for IGFBPs (Magee et al., 1999). Alanine scan of IGF-I also revealed that Glu3 and Phe49 are very important for IGFBP-1 binding (Dubaquie and Lowman, 1999); this study identified additional residues involved in

IGFBP-1 and -3 binding, including Gly7, Leu10 and Phe25 (Dubaquie and Lowman, 1999).

Using a chemical iodination approach, Moss and co-workers showed the accessibilities of Tyr60 of IGF-I and Tyr59 of IGF-II were decreased 2- and 6-fold, respectively, when IGFs were bound to IGFBP-2, indicating their involvement (Moss et al., 1991). By contrast, Tyr24 and Tyr31 of IGF-I and Tyr2 and Tyr27 in IGF-II in IGF-II were not protected in the IGF-IGFBP-2 complexes (Moss et al., 1991). A similar method, hydrogen/deuterium exchange, has been used to probe the IGF-I residues that are protected from the solvent when it is in complexed with IGFBP-1 and suggested that parts of the regions Gly1-Glu9 and Phe49-Leu54 are involved in IGFBP-1 binding, consistent with results obtained with other approaches (Ehring, 1999). Using biotinylation of IGF-I, Robinson and Rosenzweig reported that the biotin moieties NHS-linker was inaccessible to Gly1 and Lys27, but not to Lys65 and Lys68 when IGF-I is bound to IGFBP-2 or IGFBP-3 (Robinson and Rosenzweig, 2004).

NMR spectroscopy has also been used to identify IGF residues that are significantly perturbed upon IGFBP binding (Jansson et al., 1998; Carrick et al., 2005). Jansson and co-workers reported that IGF-I residues Pro2, Glu3, Gly7, Gly19, Pro28-Gly30, Gly32, Arg36, Arg37, Gln40-Gly42, Pro63, Lys65, and Lys68-Ala70 were perturbed upon binding to IGFBP-1, and subsequent alanine substitution of arginine residues Arg36, Arg37 and Arg50 affected IGFBP-1 binding (Jansson et al., 1998). More recently, Carrick and co-workers studied the chemical shift perturbations of IGF-I and IGF-II upon binding to full-length IGFBP-2 and a N-domain fragment of IGFBP-2 (Carrick et al., 2005). IGF-I residues Glu9, Leu14, Gln15, Phe16, Gly19 Phe25, Leu54 and Leu57 and IGF-II residues Gly11, Phe26, Phe28, Glu45, Ser50, Leu55, Glu57, and Thr58 exhibited large shifts or disappeared (Carrick et al., 2005).

During the course of this thesis, crystal structures of IGF-I·N-BP-4·C-BP-4 ternary complexes were reported (Siwanowicz et al., 2005; Sitar et al., 2006), which will be discussed in the following Chapters. Collectively, the IGFBP binding site on IGF-I are formed by the B and A domain residues Glu3, Thr4, Cys6, Gly7, Ala8, Glu9, Leu10, Val11, Asp12, Gln15, Phe16, Asp20, Phe23, Tyr24, Phe25, Asn26, Val44, Cys48, Phe49, Arg50, Ser51, Leu54, and Tyr60, which are mapped on the IGF-I surface in Figure 1.3. Equivalent residues on IGF-II also form its IGFBP binding site. It can be seen that the IGFBP binding site partially overlaps with the IGF-1R binding site, such that binding to IGFBPs blocks the binding of IGFs to the receptors.

It is worth noting that some IGF mutations affect binding to different IGFBPs to varying degree. For example, substitution of IGF-II residues Phe48-Ser50 with the corresponding residues from insulin reduced binding to IGFBP-1, -5, and -6 more than 50-fold and to IGFBP-4 15- to 50-fold, but to IGFBP-2 and IGFBP-3 only 6- to 12-fold (Bach et al., 1993). The Glu9Lys mutation in IGF-I decreased its affinities for IGFBP-2 and -6 to a much larger degree than to other IGFBPs (Magee et al., 1999). In addition, alanine substitutions of IGF-I residues Glu3 and F49 reduced binding to IGFBP-1 by 34- and 100-fold, respectively, but did not affect or reduced only 4-fold the binding affinity for IGFBP-3 (Dubaquie and Lowman, 1999). It therefore appears that the binding sites on IGF surfaces for different IGFBPs vary subtly, and/or, differences in the IGFBP structures lead to different tolerance for particular IGF mutations.

1.4 Structure and IGF Binding Determinants of the IGFBPs

1.4.1 Domain organization and disulfide bonds of IGFBPs

Sequence alignments of six IGF binding proteins are shown in Figure 1.4, and sequences of IGFBP-2 of different species are shown in Figure 1.5. As illustrated in Figure 1.6, according to the distinct regions of amino acid sequence homology, IGFBP molecules are generally divided into three domains of approximately equal length, the highly conserved cysteine-rich amino- and carboxyl-terminal domains, joined by a non-conserved linker domain, which differs between each IGFBP species (Hwa et al., 1999). There are intra-domain disulfide bonds within the N- and C-domains, but there are no inter-domain disulfide bonds. Disulfide bond patterns are identical in the C-terminal domains of IGFBP-2 (Forbes et al., 1998), IGFBP-4 (Chelius et al., 2001), and IGFBP-6 (Neumann et al., 1998; Neumann and Bach, 1999), and in the latter parts of the N-terminal domains of all six binding proteins (Hashimoto et al., 1997; Kalus et al., 1998; Neumann et al., 1998; Neumann and Bach, 1999; Standker et al., 2000; Chelius et al., 2001) but different for the first eight or six cysteines between IGFBP-4 (Siwanowicz et al., 2005) and IGFBP-6 (Neumann et al., 1998; Neumann and Bach, 1999).

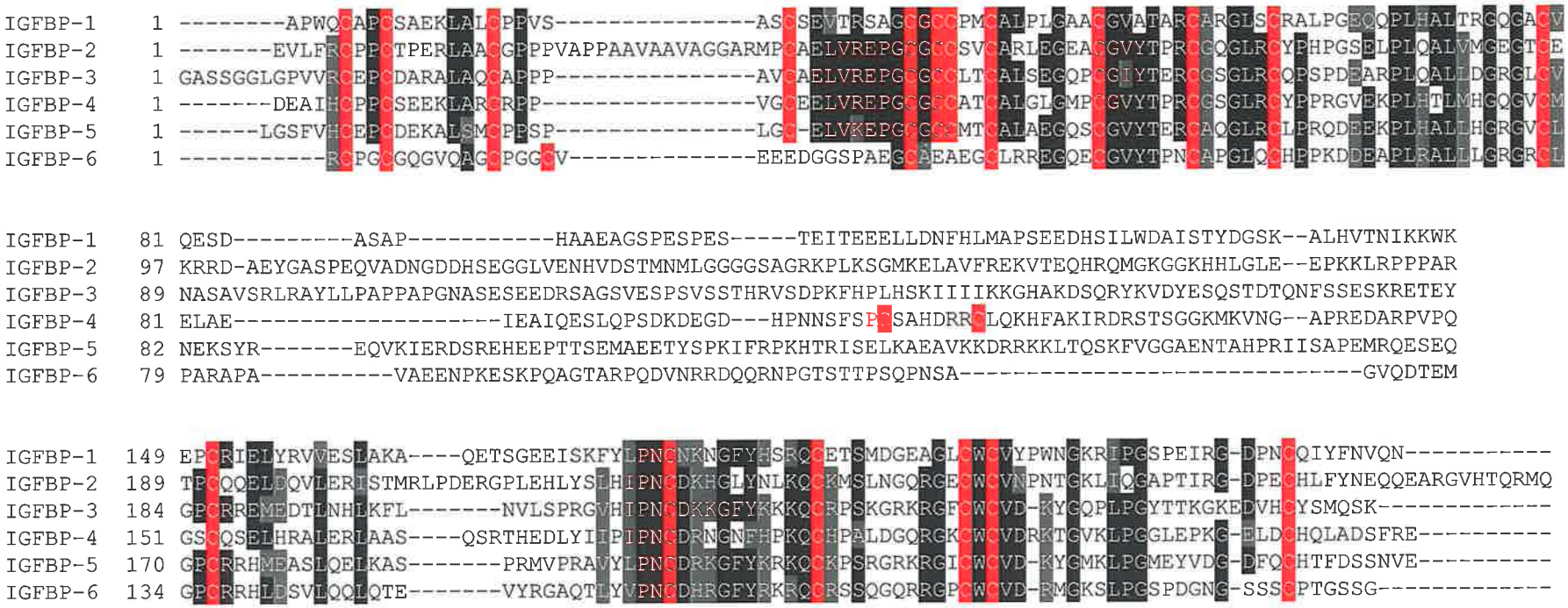


Figure 1.4 Sequence alignments of six IGF binding proteins

Sequences are of human proteins. IGFBP-1 (Swiss-Prot: P08833); IGFBP-2 (P18065); IGFBP-3 (P17936); IGFBP-4 (P22692); IGFBP-5 (P24593); IGFBP-6 (P24592).

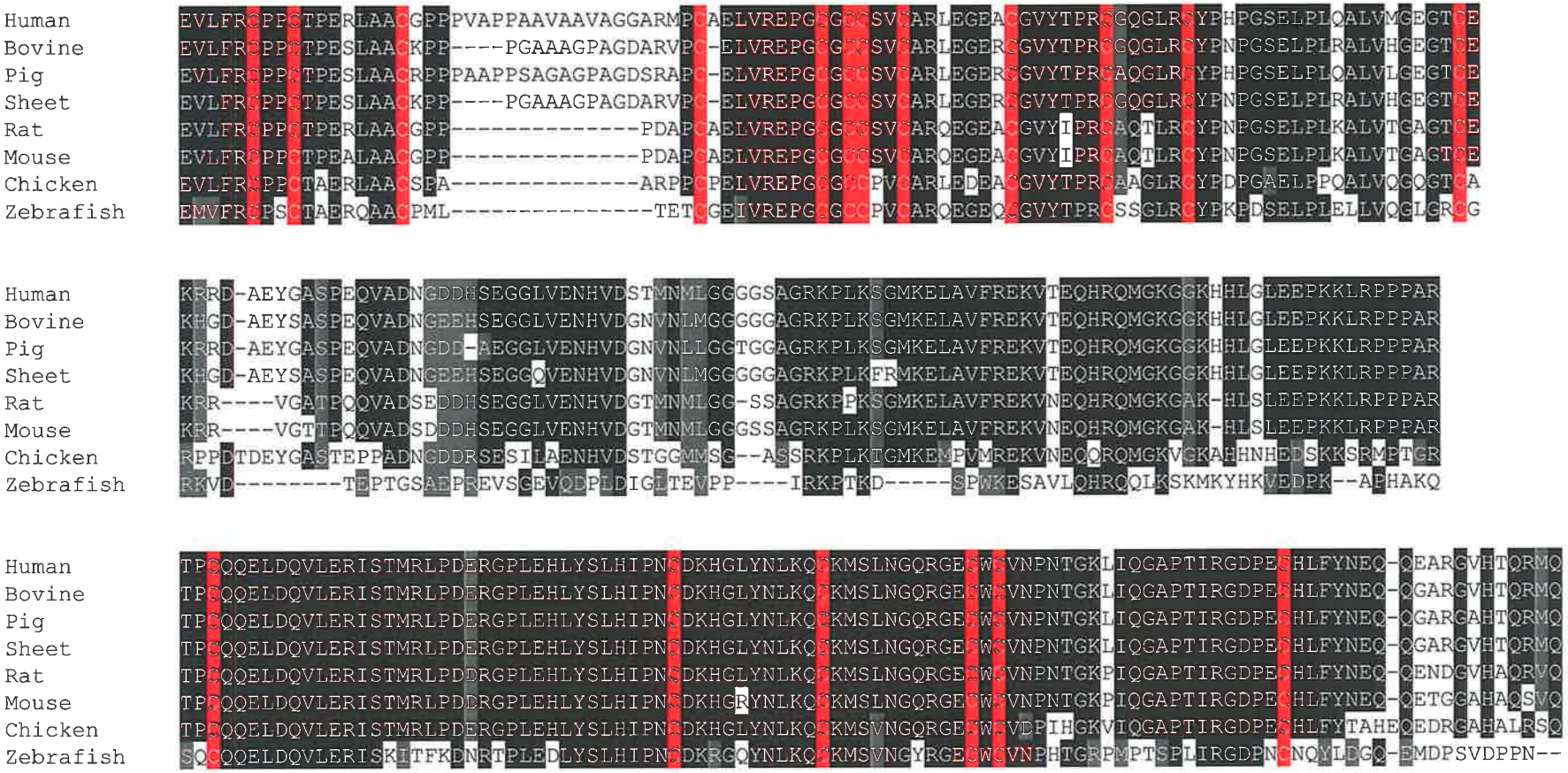


Figure 1.5 Sequence alignments of IGFBP-2 of several different species

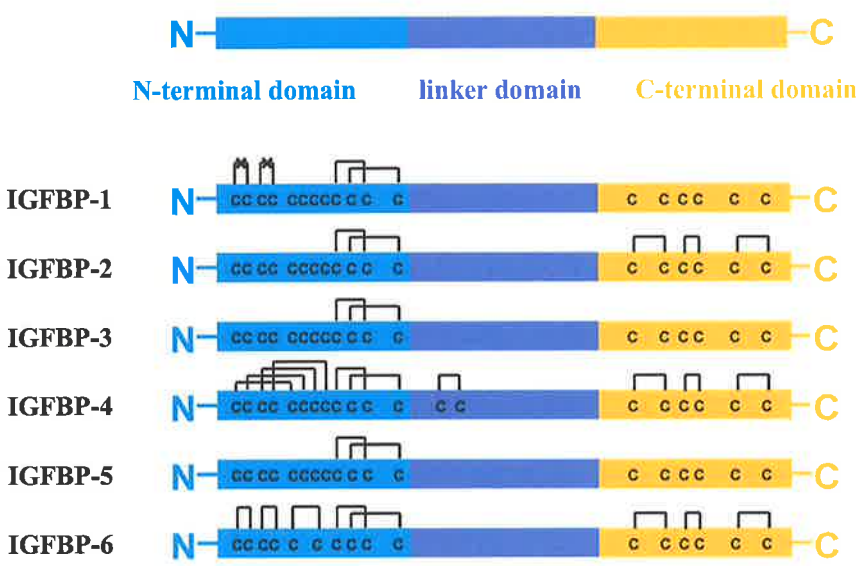


Figure 1.6 Domain organization and disulfide bonds of IGFBPs

IGFBP molecules can be divided into three domains. Conserved cysteine residues are represented by "C", and known disulfide bonds shown by connecting lines. References are in section 1.4.1. "X" in IGFBP-1 indicates the cysteines are not bonded.

1.4.2 The IGF binding site on IGFBP N-domain

IGF binding activity has been reported for both naturally occurring N-domain fragments generated by proteolysis and recombinant N-domain fragments, including IGFBP-1 (Huhtala et al., 1986), IGFBP-2 (Carrick et al., 2001; Mark et al., 2005), IGFBP-3 (Hashimoto et al., 1997; Lalou et al., 1997; Devi et al., 2000; Galanis et al., 2001), IGFBP-4 (Chernausek et al., 1995; Qin et al., 1998; Standker et al., 2000; Vorwerk et al., 2002; Siwanowicz et al., 2005), IGFBP-5 (Huhtala et al., 1986; Andress et al., 1993; Chernausek et al., 1995; Kubler et al., 1998; Standker et al., 1998) and IGFBP-6 (Headey et al., 2004c). These fragments exhibited 5- to 100-fold lower affinities than the whole molecules, and also exhibited significantly reduced or no inhibition of IGF action. At high concentrations, a N-domain fragment of IGFBP-3 inhibited IGF-1R autophosphorylation by sequestration of IGF-I (Devi et al., 2000). Interestingly, a number of other proteins contain domains sharing sequence similarities with the N-terminal domain of IGFBPs, and were hence named IGFBP related proteins (IGFBPrPs) (Hwa et al., 1999). These proteins bind IGFs with at least 100-fold lower affinities than the IGFBPs.

Using a chemical iodination approach, Hobba and co-workers found that Tyr60 of IGFBP-2 was protected from iodination in the IGF-I-IGFBP-2 and IGF-II-IGFBP-2 complex, suggesting its involvement in binding (Hobba et al., 1996). Subsequently, these authors revealed that both aromatic side-chain and hydrogen bonding potential of Tyr60 were important structural determinants of IGF binding (Hobba et al., 1998). Alanine substitution of the adjacent valine, Val59Ala, also altered the binding characteristics of IGFBP-2, whilst alanine substitutions for Thr61, Pro62 and Arg63 did not reduce IGF binding affinities (Hobba et al., 1998). The importance of corresponding residues of IGFBP-3 (Imai et al., 2000), IGFBP-4 (Imai et al., 1997; Qin et al., 1998) and IGFBP-5 (Kalus et al., 1998) in IGF binding have been confirmed.

Analysis of deletion mutants of IGFBP-4 indicated that the IGF binding domain in IGFBP-4 involves a hydrophobic motif (Leu72-Met80) (Byun et al., 2001). In IGFBP-3, single substitutions of Ile56, Leu80 and Leu81 to glycines resulted in 2- to 5-fold decrease in IGF-I binding. Greater reduction was seen when both Leu80 and Leu81 were mutated, and when all of these three residues were substituted, binding affinity was lost completely (Buckway et al., 2001).

NMR analysis of the binding interaction between IGF-I and mini-N-BP-5 by Kalus and co-workers showed the largest IGF-I induced chemical shifts for residues Val49, Tyr50, Pro62 and Lys68 to Leu75 (Kalus et al., 1998). Several multiple site-directed mutagenesis experiments have been conducted to confirm this structural information. Imai and co-workers showed that the IGFBP-5 mutant (Lys68Asn/Pro69Gln/Leu70Gln/Leu73Gln/Leu74Gln) and the corresponding IGFBP-3 mutant had ~1000-fold and >1000-fold reductions in affinity, respectively, in solution binding assays (Imai et al., 2000). Interestingly, the same IGFBP-5 mutant was recreated and was shown to have 60-fold decreased affinity in BIAcore analysis but its affinity was too low to be derived in solution binding assay (Shand et al., 2003). This discrepancy may have been due to differences in the techniques employed by those two groups. Similarly, the IGF binding affinity of a human IGFBP-3 mutant containing six alanine substitutions (Ile56Ala/Tyr57Ala/Arg75Ala/Leu77Ala/Leu80Ala/Leu81Ala) was analysed by solution binding assay and shown to be reduced more than 80-fold in solution binding assays (Hong et al., 2002).

1.4.3 Structures of IGF·N-BP complexes

The NMR structure of mini-N-BP-5, a 53-residue fragment encompassing the last 2 disulfide bonds of the N-domain of IGFBP-5, reported by Kalus and co-workers adopted a rigid globular structure consisting of a centrally located three-stranded anti-parallel β-sheet, stabilized by two inside packed disulfide bridges (Kalus et al., 1998).

Subsequently, these authors solved the 2.1 Å crystal structure of IGF-I-mini-N-BP-5 complex (Zeslawski et al., 2001). The fold of the uncomplexed mini-N-BP-5 determined by NMR is preserved in this complex. Residues in a solvent-exposed loop, which undergo slow exchange on the millisecond time scale in the free form (Renner and Holak, 2001), were rigidified in the complex (Zeslawski et al., 2001). The interactions between IGF-I and mini-N-BP-5 are mainly hydrophobic and consist of interlaced protruding side chains of IGF-I and solvent-exposed hydrophobic side chains of mini-N-BP-5 (Zeslawski et al., 2001). The side chains of IGF-I residues, Phe16, Leu54 and Glu3 are inserted deep into a cleft on the mini-N-BP-5 formed by site chains of Arg53 and Arg59 on the solvent-exposed site of the molecule and by Leu70 and Leu74 on the opposite inner site, with a base formed by Cys60 and Leu61. Phe16

makes direct contacts with the backbone and side chain of Val49 and Cys60 of mini-N-BP-5. The hydrophobic cluster is closed on the solvent side by side chains of Glu3 and Glu9 of IGF-I and His71 and Tyr50 of mini-N-BP-5. In addition, Arg59 of mini-N-BP-5 makes hydrogen bonds with Glu58 of IGF-I (Zeslawski et al., 2001).

More recently, Siwanowicz and co-workers reported the crystal structure of IGF-I·N-BP-4 complex (Figure 1.7), which showed the structure of the N-terminal part of a IGFBP N-domain and the correct linkage of the first four disulfide bonds in IGFBP-4 that may also be true for IGFBP-1, -2, -3 and -5 (Siwanowicz et al., 2005). The second subdomain of N-BP-4 corresponding to mini-N-BP-5 had an almost identical structure and very similar hydrophobic interactions with IGF-I compared with mini-N-BP-5 (Siwanowicz et al., 2005). The N-terminal subdomain of N-BP-4 has a disulfide bond ladder-like structure consisting of a short two-stranded β-sheet (Siwanowicz et al., 2005). This subdomain has interactions with the second subdomain of N-BP-4 as well as IGF-I, resulting in a "L" shape of the whole N-BP-4 (Siwanowicz et al., 2005). This subdomain provided several additional contacts to IGF-I. Thus, Ala3 and Ile4 of N-BP-4 made major interaction by filling the hydrophobic patch formed by IGF-I Phe23 Tyr24, and Phe25 and the interaction is further strengthened by interaction between Arg28 of N-BP-4 and Asp20 of IGF-I as well as hydrogen bonds between Ile4 and Phe23 (Siwanowicz et al., 2005). The more recent IGF-I·N-BP-4·C-BP-4 complex structure (Sitar et al., 2006) will be discussed in the following Chapters.

1.4.4 The IGF-binding site on IGFBP C-domain

As with the N-terminal domain, both naturally occurring and recombinant IGFBP C-domain fragments of IGFBP-2 (Wang et al., 1988; Ho and Baxter, 1997), IGFBP-3 (Spencer and Chan, 1995; Devi et al., 2000; Galanis et al., 2001; Vorwerk et al., 2002), IGFBP-4 (Standker et al., 2000), and IGFBP-5 (Standker et al., 1998) have been shown to possess partial IGF binding activity. For example, a C-terminal fragment of IGFBP-2 (169-289) isolated from human milk had less than 10% binding affinity to IGF-I and 25% binding affinity to IGF-II compared with wide-type human IGFBP-2 in solution binding assay (Ho and Baxter, 1997). Deletions of C-domain residues 222-284 of bovine IGFBP-2 dramatically reduced its IGF binding affinity (Forbes et al., 1998). BIAcore analysis of recombinant bovine IGFBP-2 N- and C-domain fragments by

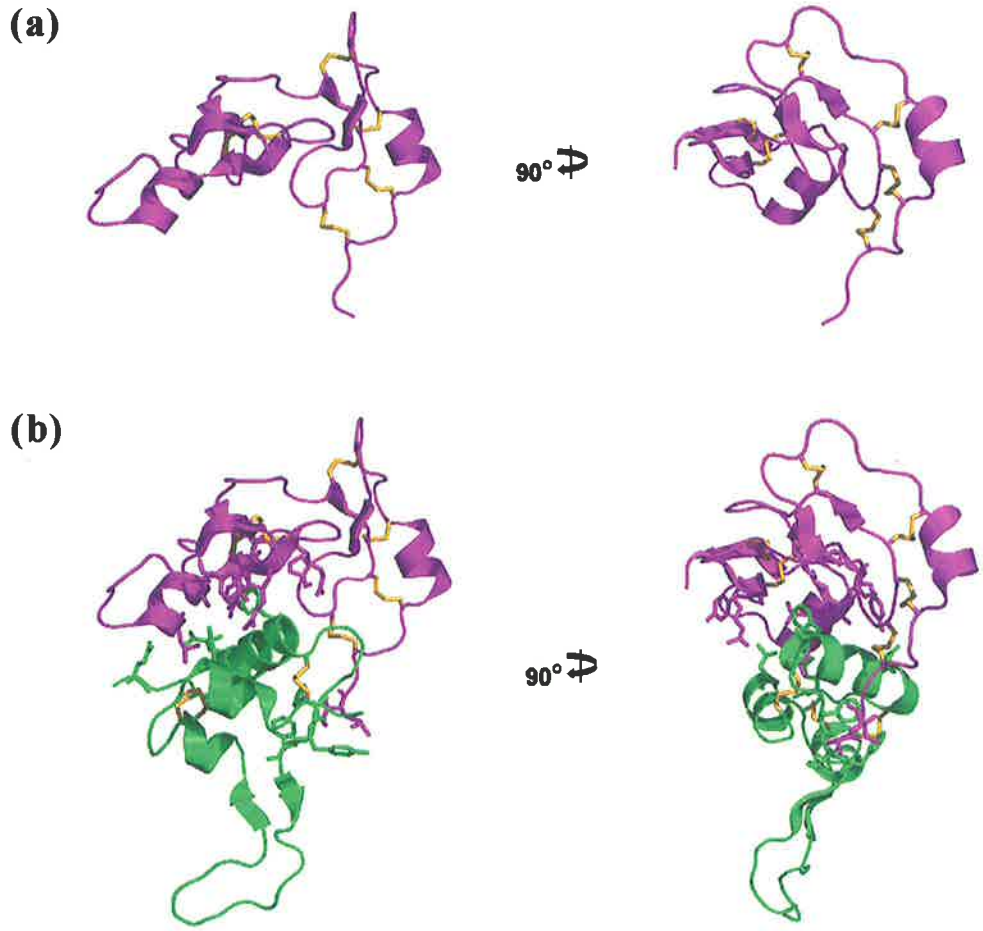


Figure 1.7 Structure of the IGF-I·N-BP-4 complex

PDB 1WQJ (Siwanowicz et al., 2005). N-BP-4 (pink) and IGF-I (green) are shown as ribbon models. Disulfide bonds are shown (gold). (a) Structure of N-BP-4; (b) Structure of the complex. Sidechains of major interacting residues are shown.

Carrick and co-workers showed that the C-terminal domain possesses a binding site, which is essential for the stability of the complex and must be combined with the fast association binding site in the N-terminal domain to form one high affinity binding site (Carrick et al., 2001). No inhibition of IGF-I induced IGF-1R phosphorylation was detectable for IGFBP-3 C-domain fragments, despite their ability to bind IGF (Devi et al., 2000).

Although the majority of previous studies showed the reduced IGF binding affinity of C-domain fragments, the reduction proportions compared to full-length IGFBPs vary greatly. In addition, there have been some reports in which C-domain fragments did not bind IGFs (Kalus et al., 1998; Qin et al., 1998; Fernandez-Tornero et al., 2005; Siwanowicz et al., 2005). Differences may have been due to the different structural integrity of these fragments and variations in the sensitivity of the binding assays performed.

Mutations in IGFBP-5 of the IGFBP-2 Gly229 and Gln235 individually to Lys and Ala decreased IGF-binding to 7-flod and 5-fold, respectively (Bramani et al., 1999; Song et al., 2000); combined mutations of equivalent residues in IGFBP-3 and -5 resulted in a larger reduction in IGF binding affinity (Shand et al., 2003; Yan et al., 2004). Mutations of ^{228}KGRKR in IGFBP-3 (equivalent to IGFBP-2 ^{240}LNGQR) decreased binding affinity for IGF-I and IGF-II by 2- to 3-fold (Firth et al., 1998). Deletions of the residues immediately downstream the CWCV motif were found not to affect IGF binding in some reports (Forbes et al., 1998), but were shown to decrease IGF binding in another (Qin et al., 1998). Mutations of the ^{253}KED in IGFBP-3 (equivalent to IGFBP-2 ^{256}RGD) to RGD decreased IGF binding affinity 4- to 6-fold (Firth et al., 1998), but complete deletions of the equivalent residues and the subsequent residues towards the C-terminus in IGFBP-2 (Forbes et al., 1998) and IGFBP-4 (Qin et al., 1998) did not affect IGF binding. Using the photoaffinity labelling approach, Horney and co-workers located two regions in the C-domain of IGFBP-2 (Gly212-Lys227 and Gly266-Arg287) that were thought to be involved in IGF-I interaction (Horney et al., 2001).

During the course of this thesis, others have used NMR (Headey et al., 2004a; Yao et al., 2004) and crystallography (Siwanowicz et al., 2005; Sitar et al., 2006) to reveal the IGF binding by IGFBP C-domains. These recent reports will be discussed in following chapters.

1.4.5 The linker domain

Most evidence so far does not support an IGF binding site in the variable linker domain of IGFBPs. However, it is expected that, linking the N- and C-domain binding sites generates higher IGF binding affinity than a single binding site provides.

Although deletion of residues 89-184 of hIGFBP-3 reduced IGF binding affinity to 40-fold (Firth et al., 1998), the $^{95\text{-}154}$IGFBP-5 did not have detectable binding affinity for IGF by BIAcore analysis (Kalus et al., 1998), and deletion of residues 94-119 and 121-141 of IGFBP-4 did not alter IGF binding activity (Qin et al., 1998). Mutations at the proteolytic cleavage sites, Lys120Asn, His121Asn, Lys134Gly (Rees et al., 1998), Lys134Gln, Lys136Gln, Met135Leu and Met135Glu (Conover et al., 2001) of IGFBP-4 and Lys138Asn/Lys139Asn double mutation of IGFBP-5 (Imai et al., 1997) did not affect the IGF-I binding affinity. Similarly, mutations of phosphorylation sites of IGFBP-1, Ser98Ala and Ser101Ala, did not affect IGF binding affinity (Jones et al., 1991).

1.5 Modifications and Other Molecular Interactions of IGFBPs

Proteolysis and post-translational modifications of IGFBPs may affect their IGF binding, and thus regulate IGF actions. Furthermore, there is a rapidly growing body of evidence showing that IGFBPs and/or IGF-IGFBP complexes interact with other molecules, as illustrated in Figure 1.8, and these interactions can modulate IGF actions and/or generate IGF-independent functions (Firth and Baxter, 2002; Ricort, 2004). Different IGFBPs have overlapping but variable such interactions and majority of these interactions are mediated by their C-domain (Bach et al., 2005).

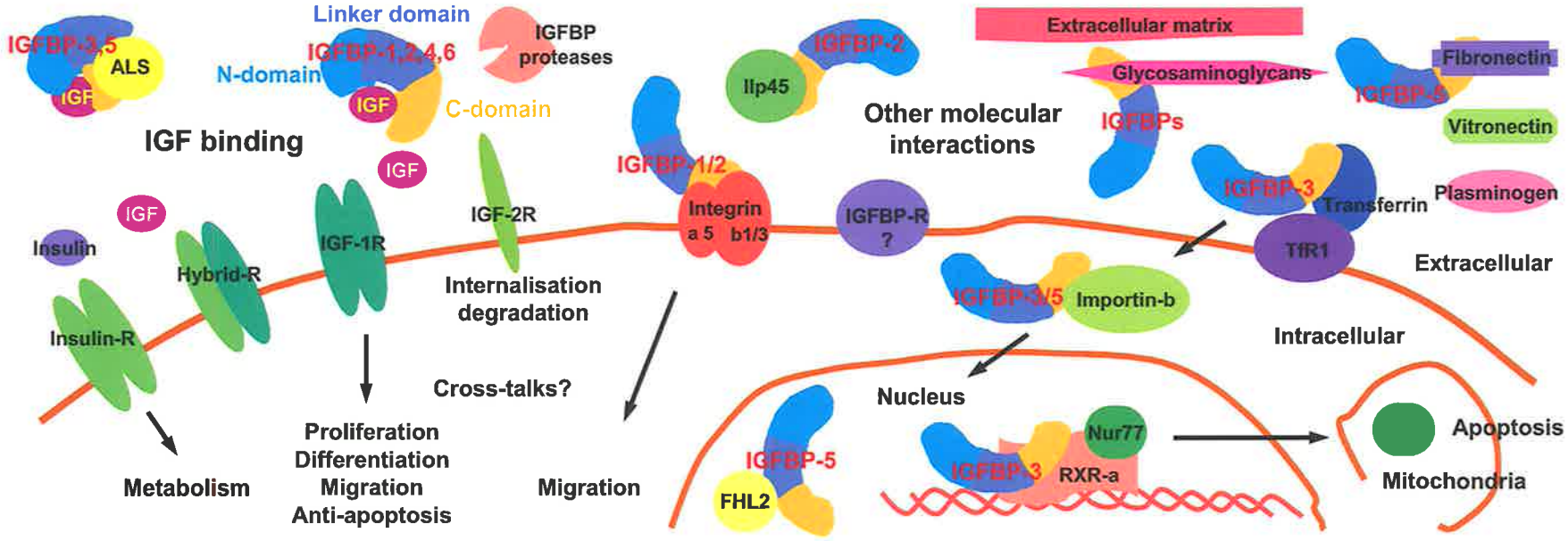


Figure 1.8 Molecular interactions of IGFBPs

A schematic summary of the known molecular interactions of IGFBPs. Reviews are available (Firth and Baxter, 2002; Bach et al., 2005). References are in section 1.5.

1.5.1 Proteolysis of IGFBPs

Proteolysis is a major mechanism for IGF releasing from IGFBPs, and plays an important role in controlling the bioavailability of IGFs in physiological and pathological states (Bunn and Fowlkes, 2003). Proteolytic cleavage has been demonstrated for all six IGFBPs, and the majority of the cleavage occurs within the linker domain. Overlapping but variable proteases for each IGFBP have been reported, including metalloproteinases such as the pregnancy-associated plasma protein A (PAPP-A), and the matrix metalloproteinase (MMPs), serine proteinases such as the complement component C1s, and cathepsins. Numerous proteolytic IGFBP fragments with reduced affinity for IGFs have been reported. These fragments may also have IGF-independent actions. An increase in IGFBP proteolysis occur in some physiological states such as pregnancy and wound healing, and pathological conditions including cancer, diabetes, and arthritis.

Proteolytic degradation of IGFBPs is regulated by multiple mechanisms (Bunn and Fowlkes, 2003). Changes in the expression and distribution of IGFBP cleaving proteases and their inhibitors would directly modulate the rate of IGFBP degradation. Covalent post-translational modifications of IGFBPs can influence their proteolysis. For example, phosphorylation of IGFBP-1 and -3, and glycosylation of IGFBP-6 inhibited their proteolytic degradation *in vitro* (Bunn and Fowlkes, 2003). Since both the phosphorylation or glycosylation sites and the protease cleavage sites are located in the linker domains, these post-translational modifications may generate steric hindrance or charge repulsion for the proteases.

Non-covalent interactions, such as ligand and ECM binding, can also affect IGFBP proteolysis (Bunn and Fowlkes, 2003). The presence of IGF-II enhances the PAPP-A cleavage of IGFBP-4 (Byun et al., 2000; Vorwerk et al., 2002), and IGFBP-2 (Monget et al., 2003). IGF-II facilitates the cleavage by binding to IGFBP-4 rather than directly interacting with the protease (Vorwerk et al., 2002). IGFs also enhance proteolytic degradation of IGFBP-2 by a serine protease in the conditioned media from porcine smooth muscle cells (Gockerman and Clemmons, 1995). In contrast, IGF-II slightly inhibited IGFBP-5 degradation by PAPP-A (Laursen et al., 2001) and PAPP-A2 (Overgaard et al., 2001). Conformational change of the IGFBP molecules upon IGF binding may either expose or obstruct the protease cleavage sites, but the exact

structural mechanisms are unclear. Similarly, changes of susceptibility to proteolytic degradation also occur upon IGFBPs binding to the ECM. Binding to fibroblast ECM was found to protect IGFBP-5 from proteolytic degradation (Jones et al., 1993a), but, in another study, binding of IGFBP-5 to fibronectin increased proteolytic degradation of IGFBP-5 (Xu et al., 2004).

1.5.2 Post-translational modifications of IGFBPs

IGFBPs have post-translational modifications such as phosphorylation and glycosylation, which may influence IGF binding and/or IGFBP proteolysis. Most of the modifications occur in the linker domain.

IGFBP-1, -3 and -5 can be phosphorylated at some serine residues in their linker domains (Coverley and Baxter, 1997). The effects on IGF binding, however, vary according to the IGFBP species. Although the structural mechanism is yet unknown, phosphorylation of human IGFBP-1 residues Ser101, Ser119 and Ser169 increased IGF-I binding affinity 6-fold (Jones et al., 1991), but this was not observed in a study of rat IGFBP-1 (Peterkofsky et al., 1998). Phosphorylation of IGFBP-3 did not alter its IGF binding affinity (Hoeck and Mukku, 1994), but it decreased the affinity of IGFBP-3 for ALS and cell surface (Coverley and Baxter, 1997). However, it was found in another study that phosphorylation of IGFBP-3 by deoxyribonucleic acid DNA-dependent protein kinase reduced IGF-I binding and enhanced nuclear accumulation (Schedlich et al., 2003).

Glycosylation has been reported for IGFBP-3 (Zapf et al., 1988), IGFBFP-4 (Ceda et al., 1991), IGFBP-5 (Conover and Kiefer, 1993), and IGFBP-6 (Bach et al., 1992). Glycosylation does not appear to alter IGF binding affinity of these IGFBPs, but decreases the cell surface association of IGFBP-3 (Firth and Baxter, 1999) and inhibits the glycosaminoglycan and cell membrane binding, as well as chymotrypsin and trypsin proteolysis of IGFBP-6 (Neumann et al., 1998; Marinaro et al., 2000).

Recently it was shown that nuclear localized IGFBP-3 is highly polyubiquitinated at multiple lysine residues in its C-domain, and is subject to ubiquitin/proteasome-dependent degradation (Santer et al., 2006).

1.5.3 Heparin, glycosaminoglycan and extracellular matrix binding

Binding of IGFBPs to heparin and other glycosaminoglycan, as well as other components of the extracellular matrix, can alter the IGF binding and proteolytic cleavage of IGFBPs. Glycosaminoglycan binding of IGFBPs was investigated initially with respect to an 18-residue region that contains 10 arginine or lysine residues in the C-domains of IGFBP-3 and -5. Arai and co-workers showed that heparin, heparan sulfate, and dermatan sulfate, but not keratan sulfateor hyaluronic acid, inhibited IGF-I binding of IGFBP-3 and -5, and also separate preformed complexes, but had no effect on IGF binding by IGFBP-1, -2, or -4 (Arai et al., 1994b). Heparin, heparan sulfate and dermatan sulfate also inhibited IGFBP-5 degradation in fibroblast cultures (Arai et al., 1994a). Apart from glycosaminoglycan/proteoglycans, the ECM contains structural proteins such as collagen and elastin, and specialized proteins such as fibrillin, fibronectin, and laminin. ECM binding of IGFBP-5 influences cellular response to IGF-I, since porcine smooth-muscle cells expressing the weak ECM-binding mutants are less responsive to IGF-I (Parker et al., 1998).

Mutagenesis studies have identified several basic residues in the C-domain of IGFBP-5 that are important for binding to heparin (Arai et al., 1996b; Song et al., 2000; Allan et al., 2006) or ECM (Parker et al., 1998), but the contributions of other basic residues in the 10 basic residue region and the three-dimensional location of the binding site were unclear. Interestingly, the residues important for ECM binding were overlapped but not identical to those for heparin binding. This may have been due to the heterogeneous nature of the ECM of different cells and under different circumstances.

Mechanisms underlying the decrease of IGF binding affinity of IGFBP-3 and -5 upon binding to heparin or ECM have been proposed, but remains unclear. One possibility is that the binding of heparin or ECM may result in conformational changes of the IGFBP molecule that influence IGF binding. On the other hand, as the heparin binding site is adjacent to, or even overlaps, with the IGF binding site, heparin and ECM binding may generate a direct spatial occlusion of the IGF binding site (Bramani et al., 1999; Song et al., 2000; Headey et al., 2004a). Beattie and co-workers reported that IGF and heparin binding to IGFBP-3 and -5 are partly competitive in nature, and the IGF-IGFBP-heparin ternary complexes are either unable to form or very unstable (Beattie et al., 2005). Since IGF-I and IGF-II can displace endogenous IGFBP-5 from

monolayer mouse mammary epithelial cells, the authors argued that this competitive process may be biologically significant (Beattie et al., 2005).

Interestingly, whether IGFBP-2 binds heparin, other glycosaminoglycans or ECM has been somewhat controversial; in the reports that heparin binding was shown, its effects on IGF binding were contradictory. Arai and co-workers reported that IGFBP-2 did not bind heparin but acquired heparin binding ability upon binding to IGFs (Arai et al., 1996b), while others found heparin binding by IGFBP-2 in the absence of IGFs (Russo et al., 1997) and indicated that its heparin binding site may be located in the linker-domain (Russo et al., 2005). However, the absence of heparin binding by IGFBP-2 either with or without IGF-I has also been shown (Song et al., 2001; Beattie et al., 2005).

1.5.4 Integrin binding

The C-domains of IGFBP-1 and -2 contain RGD motifs that can bind integrins, leading to altered cell migration. (Jones et al., 1993b; Schutt et al., 2004; Wang et al., 2006). Both IGFBP-1 (Jones et al., 1993b) and IGFBP-2 (Schutt et al., 2004) bind $\alpha5\beta1$ integrin through their RGD motifs. IGFBP-1 stimulated Chinese hamster ovary cell migration and bound to $\alpha5\beta1$ integrin, both in an RGD-dependent manner since WGD mutant did not have these effects (Jones et al., 1993b). The interaction of IGFBP-2 with integrin $\alpha5\beta1$ decreased sarcoma cell adhesion and cell proliferation (Schutt et al., 2004). In addition, it was shown recently that IGFBP-2 enhanced glioma cell migration through its RGD-mediated interaction with integrin $\alpha5$ (Wang et al., 2006). Nevertheless, Hoeflich and co-workers reported that cell surface association of IGFBP-2 in the IGFBP-2 transgenic mice is RGD-independent (Hoeflich et al., 2002). Furthermore, interaction of IGFBP-2 with $\alpha v\beta3$ integrin negatively modulated IGF-mediated breast tumour cell migration, and was correlated with reduced tumour size of cancer cells, but the interaction between IGFBP-2 and $\alpha v\beta3$ integrin may be RGD-independent (Pereira et al., 2004). Interestingly, crosstalk seems to exist between the IGF-1R and the integrin signalling pathways, and this may lead to more complicate cellular actions by IGFBP-1 and -2 (Clemmons and Maile, 2005).

Cell surface IGFBP receptors have been proposed but further analysis is lacking (Ricort, 2004). The type V transforming growth factor- receptor was identified as a

receptor for IGFBP-3, -4, and -5, and mediates a growth-inhibitory response (Leal et al., 1997; Leal et al., 1999). A potential receptor for IGFBP-3 on platelet membranes was also suggested, but its nature is unclear (Taylor and Spencer, 2001).

1.5.5 Other molecular interactions

Some other molecular interactions have been reported for IGFBPs. The great majority of these interactions were reported for IGFBP-3 and -5, and most of the interactions are mediated by their C-domains (Bach et al., 2005). Interestingly, the heparin-binding 10 basic residue regions in IGFBP-3 and -5 were found to contribute to many of these interactions.

In circulation, IGFBP-3 and IGFBP-5 form IGF-IGFBP-acid-labile subunit (ALS) ternary complexes, and the majority of the IGFs are transported in these complexes. Firth and co-workers showed that IGFBP-3 residues Lys228-Arg232 (Firth et al., 1998) and IGFBP-5 residues Lys211/Arg214/Lys217/Arg218 (Firth et al., 2001), which are within the 10 basic residue regions of their C-domains, are essential for ALS binding. IGFBP-5 linker domain residues Lys134/Arg136 or Lys138/Lys139 also contribute to the binding (Twigg et al., 2000; Firth et al., 2001).

IGFBP-3 binds plasminogen (Campbell et al., 1998), fibrinogen and fibrin (Campbell et al., 1999), and type Iα collagen (Liu et al., 2003), whereas IGFBP-5 binds plasminogen activator inhibitor-1 (PAI-1) (Nam et al., 1997), thrombospondin and osteopontin (Nam et al., 2000), vitronectin (Nam et al., 2002), and the 10th and 11th type I repeats of fibronectin (Xu et al., 2004). Using IGF-II affinity chromatography, Oesterreicher and co-workers have found that IGF-II interacts with multiple plasma proteins, including IGFBP-3 and -5, transferrin, plasminogen, prekallikrein, and antithrombin III, indicating that IGFs, IGFBPs and IGFBP-associated proteins may interact to form complexes that are capable of modulating the IGF binding and the stability of IGFBPs (Oesterreicher et al., 2005). In fact, it was suggested that lowered IGF-I affinity of the fibrin-bound IGFBP-3 at wound sites can increase IGF-I release (Campbell et al., 1999). Binding of thrombospondin, osteospontin and vitronectin to IGFBP-5 enhanced IGF-I-stimulated smooth muscle cell growth (Nam et al., 2000) (Nam et al., 2002). IGFB-IGFBP-5-vitronectin complex also enhanced breast tumour cell migration (Kricker et al., 2003). Interestingly, binding of fibronectin to IGFBP-5

negatively regulated the IGF-dependent action of IGFBP-5 by increasing IGFBP-5 proteolysis (Xu et al., 2004), whereas PAI-1 binding to IGFBP-5 partially protects IGFBP-5 from proteolysis (Nam et al., 1997). Recently, Sorrell and co-workers found that IGFBP-5 directly activates tissue plasminogen activator (tPA) to activate plasminogen, in addition to inhibition of PAI-1, resulting in cell death induced by cleavage of focal adhesions (Sorrell et al., 2006).

The C-domains of IGFBP-1 and IGFBP-3 were found to bind metal ions. In the crystal structure of C-BP-1, a Fe^{2+} ion was found to bind to His213 and Ser214 of C-BP-1 (equivalent to Asn232 and Leu233 of C-BP-2) and this interaction decreased IGF-II binding affinity (Sala et al., 2005). On the other hand, Singh and co-workers reported that the C-domain of IGFBP-3 contains a short region that can bind different ions including Ni^{2+}, Co^{2+}, Zn^{2+} and Fe^{3+}, and the binding may influence the cellular/nuclear internalisation of IGFBP-3 (Singh et al., 2004). IGFBP-3 has been shown to bind humanin , an Alzheimer's survival peptide (Ikonen et al., 2003), and the cell surface retention sequence binding protein-1 (CRSBP-1) (Huang et al., 2003)

IGFBP-3 and -5 induce apoptosis in some cell types (Firth and Baxter, 2002; Lee and Cohen, 2002). A number of IGFBP-3 and -5 binding partners are involved in this IGF-independent function. IGFBP-3 can bind transferrin via the C-domain, forming a IGFBP-3-transferrin-transferrin receptor (TfR1) ternary complex, and internalise into the cells (Weinzimer et al., 2001; Lee et al., 2004). In T47D human carcinoma cells, IGFBP-3 and -5 share a common nuclear transport pathway (Schedlich et al., 1998). A basic sequence in their C-domains is essential for the nuclear localization of IGFBP-3 and -5 by interaction with importin-β subunit (Schedlich et al., 2000). In the nucleus, IGFBP-3 interacts with transcription factor retinoid X receptor-α (RXRα) (Liu et al., 2000), whereas IGFBP-5 interacts with a four and a half LIM protein 2 (FHL2) (Amaar et al., 2002).

Furthermore, it is recently shown that IGFBP-5 interacts with the nuclear histone/DNA complex, and several conserved acidic residues in the its N-domain are critical for its transactivation activity (Zhao et al., 2006)

1.6 Introduction to Protein Structure and Dynamics by NMR Spectroscopy

In this thesis, high-resolution NMR spectroscopy solution was used to determine the structure of C-BP-2 and analyse its backbone dynamics, as well as to study its interactions with IGFs, N-BP-2, and heparin molecules. Some basic background of NMR and its application in protein structure and dynamic studies are presented briefly in this section, whereas the NMR methods used in this thesis to analyse molecular interactions will be described in Chapter 5. Some excellent textbooks (Wuthrich, 1986) and review articles (Hinds and Norton, 1997; Wider, 2000; Kanelis et al., 2001; Palmer, 2001) are available, and they were referenced in preparing this section.

The basis of NMR is a quantum mechanical property of the nucleus: the spin. The NMR-active nuclei in biological macromolecules are either naturally abundant isotopes (^{1}H, ^{31}P), or naturally non-abundant isotopes (^{15}N, ^{13}C) that can be incorporated into the molecules through isotope labelling. These nuclei carry a spin of 1/2, which can be viewed as a magnetic dipole. When an external magnetic field exists, such as that provided by the superconductive magnet in the NMR spectrometer, the small atomic dipoles can only orient either parallel or antiparallel to the external magnetic field. These two possible orientations correspond to slightly different energy states, and the spins absorb or emit the energy difference in the form of electromagnetic radiation. Because the difference in spin numbers between these two states is extremely small, and because only this small difference gives rise to an NMR signal, NMR is fundamentally an insensitive technique.

In NMR experiments, after the equilibrium state of the spins is perturbed by suitable irradiation of electromagnetic waves, often in the form of radio-frequency (*rf*) pulses, an NMR signal can be observed. Since the system tends to return to its equilibrium state, the signal will decay (in a process known as relaxation) over time, and the NMR signal (free induction decay, FID), is recorded during this time. The NMR frequency spectrum is obtained from these data by Fourier transformation. In an NMR spectrum, the nuclei are represented by resonance frequencies. The exact resonance frequency depends on the chemical environment of each nucleus, and the slight differences in their frequencies are called chemical shift. The chemical shift is

very sensitive to a combination of structural, electronic, magnetic and dynamic variables.

NMR spectroscopy is one of the two main techniques (the other being X-ray crystallography) that can currently provide structures of macromolecules such as proteins and nucleic acids and their complexes at atomic resolution. Whereas X-ray crystallography requires crystals of these molecules, which are sometimes difficult to grow, NMR measurements are carried out in solution under near physiological conditions. The strategy of NMR protein structure determination is summarized in Figure 1.9 (a).

The first step is the preparation of the protein solution sample. Since NMR spectroscopy is an insensitive technique, relatively high (~1 mM) protein concentration is required. Nevertheless, with the very recent application of cryoprobes, lower protein concentrations (a few hundred micromolar) can be used. This step also includes the optimisation of experimental conditions such as the sample pH and the temperature, so that spectra are recorded as close as possible physiological conditions while highest spectral quality is achieved. In fact, the importance of this step is evident in Chapter 4 and Chapter 6.

NMR spectra are then acquired. They contain a wealth of information regarding the NMR active atoms (spins) in the protein (see below). One-dimensional (1D) NMR spectra of proteins often contain hundreds or even thousands of resonance lines that cannot be resolved. Multidimensional NMR spectra not only provided increased resolution, but also provide the information on the correlations (through-bond or through-space) between different nuclei, which is required for a detailed analysis of the spectra. A most commonly acquired 2D spectrum, the ^{1}H-^{15}N heteronuclear single quantum coherence (HSQC) spectrum, is illustrated in Figure 1.9 (c). The spectrum exhibits correlations between amide nitrogens and the attached protons, and there are two different chemical shift axes $\omega 1$ and $\omega 2$: one for nitrogens (^{15}N) and the other for protons (^{1}H). Consequently, each cross-peak in the spectrum represents one N-H group in the protein. Since all amino acid residues except prolines have an amide group in the backbone, each cross-peak in the backbone amide region in this spectrum can represent one residue. ^{1}H-^{15}N HSQC spectra are used as building blocks for many 3D spectra, and are most frequently used in monitoring protein-protein and protein-ligand interactions (Chapter 5).

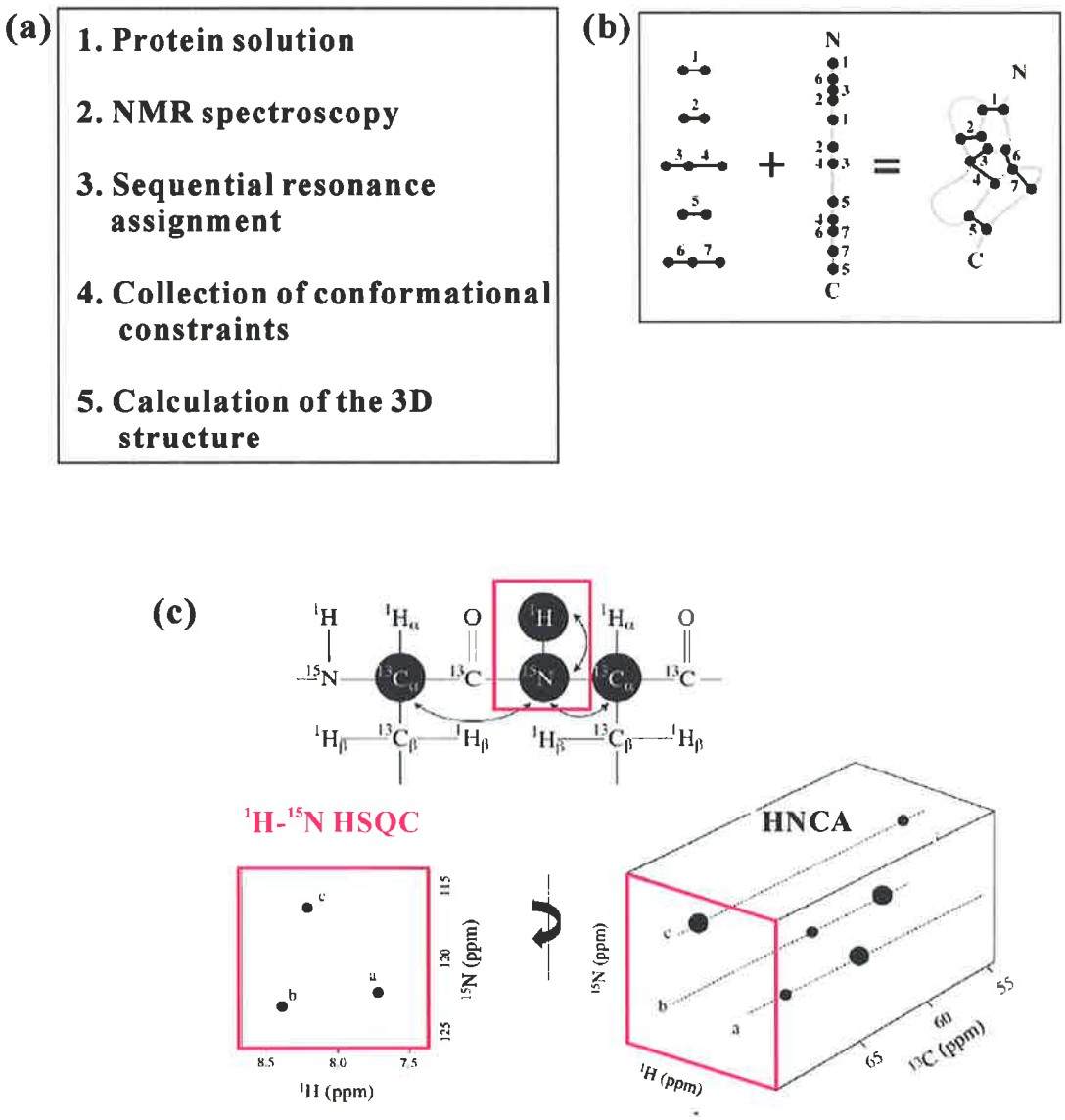


Figure 1.9 Introduction to protein structure determination by NMR

(a) Strategy of structure determination by NMR. Outline of the general steps used to solve the protein solution structures by NMR. (b) Principle of the structure determination by NMR. Adding the information on the distance between protons (left) and their sequential position in the polypeptide (middle) allows construction of the possible three-dimensional structures of the peptide (right). (c) Correlations in the 2D ^{1}H-^{15}N HSQC and 3D HNCA spectra. (a) and (b) were adapted from (Wider, 2000). (c) was adapted from the teaching material by Dr Jeff Urbauer (2004).

The NMR signals (chemical shifts) then have to be assigned to individual nuclei in the protein. Spectra showing the "through-bond" correlations of the atoms are used to perform such sequential resonance assignments. Chemical shifts of the backbone atoms, especially those of N and H^N, are often assigned before obtaining the assignments of the sidechain atoms. The most commonly used 3D spectrum for this purpose is the HNCA spectrum, which is illustrated in Figure 1.9 (c). This spectrum is built up on the 2D ^1H-^{15}N HSQC spectrum by recording the ^{13}C chemical shifts of the intra-residue and preceding C^α atoms in as a third dimension. Therefore, by connecting the C^α_{i-1} and C^α_i resonances, in conjunction with a comparison of the C^α and C^β (provided by other spectra) chemical shifts with the average values for 20 different amino acid residues, these ^1H, ^{15}N, and ^{13}C chemical shift are assigned to individual residues in the protein. Spectra showing the correlations of the sidechain atoms are then used to assign their chemical shifts. Computer programs for automatic resonance assignment have been developed recently. However, in the cases when there is serious overlap or missing resonances in the spectra due to the intrinsic nature of the protein, such as in C-BP-2, the relatively time-consuming manual assignment procedure is still necessary.

The chemical shifts of H^N, H^α, C^α, C^β, and N are sensitive to the secondary structure of the region of the protein from which they arise. Therefore, secondary structure information can be obtained by comparing the chemical shifts with the average values of residues in the random coil conformation, and this information can be taken into account in structure calculations.

Once the chemical shifts are assigned, the conformational information for structure calculation is derived from a variety of NMR spectra. The most important information is the distance constraints defining the "through-space" distance between the atoms (normally protons). In fact, as demonstrated in Figure 1.9 (b), NMR-derived protein structures are usually determined by using the distances between proton pairs and their location in the polypeptide sequence to construct the possible three-dimensional arrangements in which all distance (and angle, see below) constraints are fulfilled. The distance constraints are derived from the nuclear Overhauser effect (NOE) spectra, which show the interactions between protons separated by less the 5-6 Å in space. Dihedral angle constraints, such as for ϕ, ψ, and $\chi 1$ angles, are often used to supplement the NOEs. Dihedral constraints are derived from through-bond correlation

structure information provided by the C^{α}, C^{β}, C', H^{α}, and N chemical shifts, using computer programs such as TALOS (Cornilescu et al., 1999). The distance and angle constraints usually cannot be determined exactly, rather, they are used as ranges in the structure calculation to define the peptide backbone and sidechain conformations. Therefore, it is highly desirable to have a large number of both types of constraints to better define the structure. Other constraints often used include hydrogen bond constraints and those derived from residual dipolar couplings (RDCs).

These constraints are then used as input in the structure calculations using computer programs, supplemented by some other information, such as the amino acid sequence, bond lengths, bond angles, chiralities, and planar groups, and the steric repulsion between non-bonded atom pairs. A family of structures that satisfy the experimental data is usually reported. The quality of the NMR structures is assessed by various parameters. The root mean square deviation (RMSD) among structures of the NMR ensemble indicates the coordinate precision. The RMSD from experimental constraints should be small. Most ϕ, ψ angles should be within the most favoured or allowed regions in the Ramachandran plot.

In addition to structure determination, NMR is very powerful for the characterizing the overall tumbling and internal dynamics of the proteins in solution. While NMR can provide protein dynamic information on a much larger time scale, one most widely carried out analysis is for the protein dynamics on the nanosecond and picosecond time scale via the NMR relaxation measurements. Relaxation is caused by time-dependent fluctuations of interaction energies, and as such it is sensitive to both overall and local motions. Relaxation measurements are often carried out using ^{15}N-labelled proteins, and the motions of the backbone N-H vector are used to represent the motions of the polypeptide chain. Although the so-called 'model-free' approach is often used to analyse the relaxation data and obtain the amplitudes and frequencies of internal motions, the model-free formalism requires choice of a correct mathematical model for the fitting, otherwise the results can be biased. On the other hand, the method of reduced spectral density mapping does not require such a mathematical model. The spectral density functions, $J(\omega)$, reflect the intensity of internuclear vector motion at the NMR frequencies ω. Because three relaxation parameters are relatively easy to measure, i.e., ^{15}N R_1, R_2 and steady-state ^{15}N-{1H} NOE, $J(\omega)$ values are calculated for three frequencies in the reduced spectral density mapping. Furthermore, the translational

diffusion of molecules can be measured using pulsed field gradient NMR spectroscopy. In brief, the molecules are marked according to their positions in the sample tube by a gradient. After a period of diffusion time, their new positions are decoded by a second gradient, and their diffusion coefficient (D) can be derived.

1.7 Aims and Strategy of the Present Work

IGFBP-1 to -6, the high affinity IGF binding proteins, play important roles in regulating IGF actions as they can both inhibit and potentiate IGF activity. Previous studies established that both the N- and C-terminal domains are required for high affinity IGF binding. In addition to IGFs, IGFBPs interact with a rapidly expanding list of other proteins, such as integrins, and some serum and extracellular matrix molecules, such as heparin and other glycosaminoglycans. These molecular interactions can modulate the IGF binding affinity and/or mediate IGF-independent functions. Different IGFBPs have overlapping but different interactions of this kind. Importantly, the C-domain of IGFBPs not only contributes to the IGF binding differences among IGFBPs, but is also a major contributor to the diversity of other molecular interactions.

While the three-dimensional structures of a truncated N-domain fragment (mini-N-BP-5) and its interaction with IGFs have been studied using NMR spectroscopy (Kalus et al., 1998) and crystallography (Zeslawski et al., 2001), there was limited structure and function relationship information regarding the C-domain of IGFBPs at the time this project commenced.

The high-resolution three-dimensional structure of either a full-length IGFBP or an IGFBP C-domain has not been determined. It was therefore of desirable to undertake such a structural study. The solution structure of the IGFBP-6 C-domain (C-BP-6) was solved during the course of this project (Headey et al., 2004a), and the crystal structure of the IGFBP-1 C-domain (C-BP-1) (Sala et al., 2005) and the IGF-I·N-BP-4·C-BP-4 (Sitar et al., 2006) have been reported more recently. A comparison of the C-BP-2 structure obtained in this study with the IGFBP C-domain structures is of value to get insight into their functional differences.

Secondly, the IGFBP-2 C-domain and IGF binding interfaces had not been determined using structural approaches, although some IGF-IGFBP binding residues

had been identified using various techniques, mostly truncations and point mutations. During the course of this project, Headey and co-workers revealed the binding sites of C-BP-6-IGF-II interaction using NMR. However, because of the very low affinity of C-BP-6 for IGF-I (Headey et al., 2004c), interaction between C-BP-6 and IGF-I was not studied by NMR. It is of interest to investigate both IGF-I and IGF-II interactions of C-BP-2 in the current study.

Furthermore, glycosaminoglycan binding by IGFBPs is believed to be important in modulation of IGF actions, but previous data regarding the heparin binding ability of IGFBP-2, as well as the binding sites and the effects on IGF binding, have been contradictory. Having assigned the resonances and solved the structure of C-BP-2, the next aim was to study the interactions between C-BP-2 and IGFs as well as heparin molecules and map the binding sites.

More importantly, although the C-domain of IGFBPs is indispensable for high affinity IGF binding as naturally occurring or recombinant N-domain fragments have significantly reduced IGF binding affinities, the IGF binding affinities of C-domain fragments reported in the literature vary greatly. In some reports such binding affinities were very low or even undetectable. A more in-depth understanding of the interaction between IGFBP C-domain and IGFs may require structural studies in the absence and presence of IGFBP N-domain. Previously, few studies investigated a possible cooperativity between N- and C-domains in IGF binding and the results were controversial (Carrick et al., 2001; Payet et al., 2003; Headey et al., 2004c).

In summary, the following aims were set out for the thesis.

(1) To prepare sufficient amounts of unlabeled and isotope-labeled IGFBP-2 N- and C-domain fragments for structural studies;
(2) To assess the IGF binding ability of these IGFBP-2 domain fragments using cross-linking and BIAcore;
(3) To determine the three-dimensional solution structure and the backbone dynamic properties of the C-domain of IGFBP-2 using NMR;
(4) To map the binding sites among IGFBP-2 N- and C-domains and IGFs.
(5) To investigate heparin binding by IGFBP-2 C-domain, and, if any, to map the heparin binding site and investigate the effects on IGF binding;
(6) To investigate the possible cooperativity between N- and C-domain in IGF binding and the underlying structural mechanism;

To fulfill these aims, expression vectors for IGFBP-2 N- and C-domain fragments were generated (Chapter 2). These fragments were expressed in E. *coli* cells and purified, and their IGF binding abilities were assessed (Chapter 3). The solution structure of C-BP-2, which is the largest C-domain in the IGFBP family, was solved and its backbone dynamics analyzed using NMR (Chapter 4). Heparin binding by C-BP-2 was determined and the binding site was identified (Chapter 5). NMR experimental conditions for ^{15}N-labelled IGF-I, IGF-II, and N-BP-2 were optimised (Chapter 6). The The IGF binding sites on C-BP-2, as well as the IGFBP-2 N-domain binding site on C-BP-2 were investigated, and possible structural mechanisms underlying the cooperativity between the N- and C-domains in IGF binding were identified (Chapter 7).

Chapter 2

Generation of Vectors for Recombinant Protein Expression

2.1 Introduction

As reviewed in Chapter 1, IGFBPs play important roles in modulating IGF actions, in addition to having IGF-independent functions. Both N- and C-domains of IGFBP-2 are important for IGF binding by IGFBP-2 (Carrick et al., 2001). Furthermore, the C-domain of IGFBPs is a major contributor to the IGF binding differences and other diverse functions among IGFBPs. Whilst the structure of the C-domain of IGFBP-6 (Headey et al., 2004a), and more recently those of IGFBP-1 (Sala et al., 2005) and IGFBP-4 (Sitar et al., 2006) were reported during the course of this thesis, the structure of the C-domain of IGFBP-2 (C-BP-2) has not been reported. C-BP-2 is the largest C-domain in IGFBP family, containing an insertion and a unique extension in the C-terminus compared with other C-domains of IGFBPs. It is also of great interest to investigate the IGF binding of C-BP-2 in the absence and presence of the N-domain (N-BP-2), as well as the possible heparin binding by C-BP-2. NMR spectroscopy is a major approach in the present study for these purposes.

For NMR studies, large amounts (~10 mg) of IGFBP-2 N- and C-domain proteins are required. Recombinant bovine IGFBP-2 and its domain fragments were expressed in *E. coli* cells using the pET32a(+) expression vector by Carrick and co-workers (Carrick et al., 2001). Using this vector, the target proteins were expressed with an N-terminal (thioredoxin-$(His)_6$-tag-S-tag) fusion partner. An enterokinase cleavage site was present between the fusion partner and the protein of interest. After initial purification of the fusion protein using Ni-IDA chromatography, enterokinase was used to cleave the fusion partner. One limitation these authors found was in the enterokinase cleavage step, during which non-specific proteolysis occurred, resulting in a low yield of the IGFBP-2 protein (Carrick, 2001, PhD thesis, The University of Adelaide).

It was therefore desirable to improve the purification procedure by replacing enterokinase with other proteases that have higher cleavage specificities. 3C protease

and TEV protease are reported to have high specificities, and recently have been used extensively in recombinant protein preparation systems. In this Chapter, expression vectors for full-length IGFBP-2, as well as N- and C-domain fragments were generated, with either a 3C protease or a TEV protease cleavage site incorporated into the N-termini of the proteins of interest during PCR amplification of the cDNA.

As shown in Figure 2.1, three full-length IGFBP-2 constructs were generated. Firstly, wild-type IGFBP-2 was expressed in the fusion protein form with either a 3C protease or a TEV protease cleavage site between the fusion tag and IGFBP-2. Specificities of the 3C and TEV protease cleavages were subsequently tested on the products of these constructs. In another construct, a TEV cleavage site was placed between the fusion partner and IGFBP-2, while a 3C cleavage site was incorporated into the linker domain immediately N-terminal to the C-BP-2 fragment (see below) to be used in structure determination. This construct allowed the production of either a full-length IGFBP-2 protein, carrying a 3C cleavage site in the linker domain, or the C-BP-2 fragment. The binding affinity of the full-length protein was compared to that of a mammalian source IGFBP-2 standard sample, whereas the C-domain cleaved from this full-length IGFBP-2 was compared to that expressed as C-domain alone (thioredoxin fusion protein).

A second group of expression vectors were generated for expression of individual IGFBP-2 N- or C-domains in *E. coli* (Figure 2.1). The N-domain construct, $^{1\text{-}138}$IGFBP-2 (N-BP-2) contained approximately the first half of the linker domain. Two C-domain constructs were generated: $^{141\text{-}289}$IGFBP-2 (Large-C-BP-2) contained the C-domain plus approximately the second half of the linker domain, whereas $^{183\text{-}289}$IGFBP-2 (C-BP-2) contained the C-domain and only a few additional residues that belong to the linker domain.

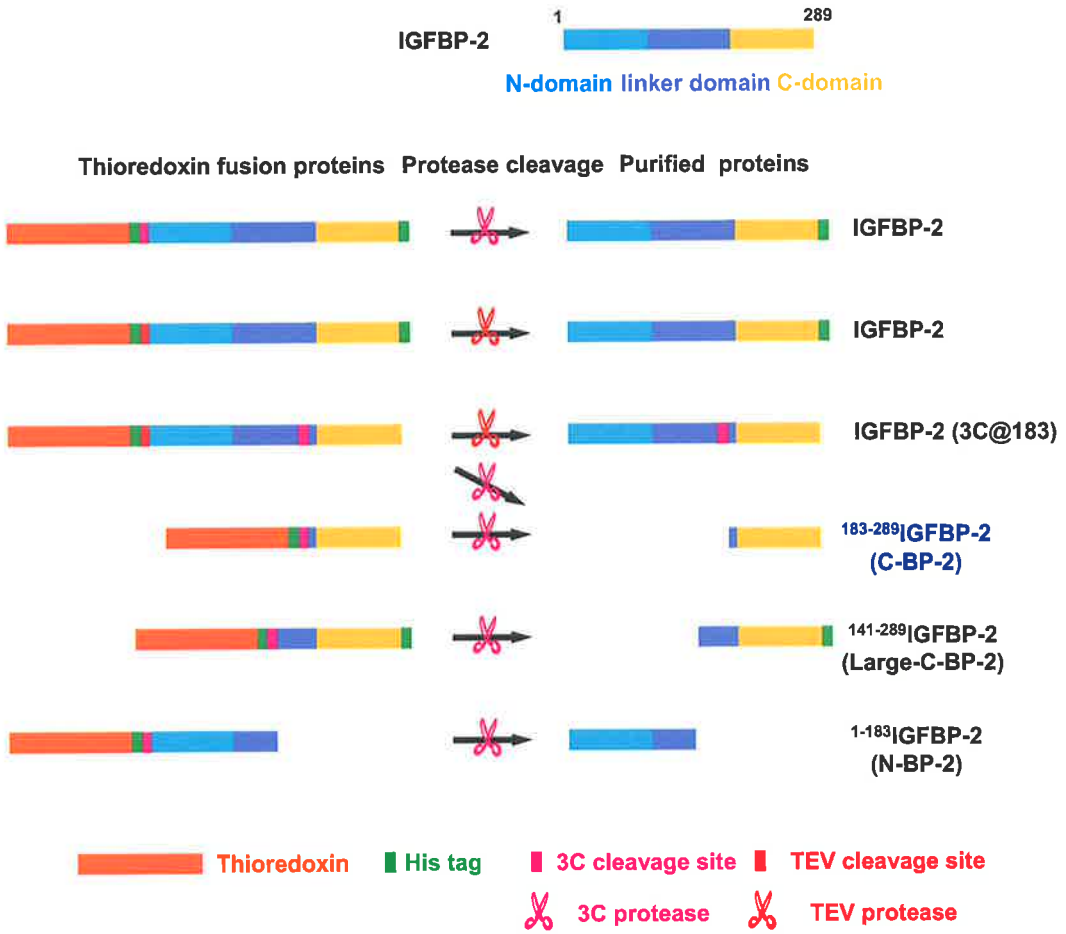


Figure 2.1 Schematic representation of IGFBP-2 constructs in this study

Wild-type IGFBP-2 and N- and C-domain fragments are represented as boxes and the N-domain, C-domain, and linker domain are indicated by different colours. The incorporated protease cleavage sites and the strategies for preparation of these domain fragments are shown.

2.2 Materials

2.2.1 Reagents and kits

Pfu DNA polymerase, *Nco* I, *EcoR* I, *Xho* I, and *Nhe* I endonucleases, and T4 DNA ligase were purchased from New England Biolabs (Beverly, MA, USA). pET32a(+) vector was from Novagen (Madison WI, USA). SPP-1/*EcoR* I and pUC19/*Hpa* II DNA markers were purchased from Geneworks (Thebarton, SA, Australia). Shrimp alkaline phosphatase was purchased from USB Corporation (Cleveland, OH, USA). UltraClean GelSpin agarose gel DNA purification kit was purchased from MO BIO Laboratiories Inc (Solana Beach, CA, USA). BigDye terminator sequencing kit was purchased from Applied Biosystems (Foster city, CA, USA). DyeEx Spin kit was purchased from QIAGEN (Hombrechtikon, Switzerland). All other reagents used were of analytical grade or higher.

2.2.2 Bacterial strains

***E. coli* DH5-α:** supE44, ΔlacU169, hsdR17, recA1, endA1, gyrA96, λ^-, relA1, [ϕ80lacZΔM15]

2.2.3 Bacterial culture media

Bacterial culture media were prepared by the Central Service Unit of the School of Molecular Biomedical Science, The University of Adelaide, using methods described elsewhere (Sambrook et al., 1989). Solutions were made in H_2O generated by the Milli-Q Ultra Pure Water System (Millipore Pty Ltd, North Ryde, NSW, Australia). Solutions and media were sterilised by autoclave.

Luria-Bertani medium (LB)

1 % (w/v) tryptone, 0.5 % (w/v) yeast extract, 0.17 M NaCl pH 7.0

Luria-Bertani plates (L plates)

LB + 1.5 % (w/v) agar

2YT

1.6 % (w/v) tryptone, 1 % (w/v) yeast extract, 0.5 % (w/v) NaCl. Adjusted to pH 7.0 with NaOH

2.2.4 Solutions for DNA procedures

PCR reaction buffer
50 mM Tris-HCl pH 9.0, 20 mM $(NH_4)_2SO_4$, 0.005 % BSA, 0.25 mM dNTPs, 3 mM $MgCl_2$

Tris borate EDTA (TBE)

0.09 M Tris, 0.09 M boric acid, 2.5 mM EDTA

Ethidium bromide staining solution stock

0.1 mg/ml ethidium bromide prepared in distilled water

10 x DNA loading dye

50 % (v/v) glycerol, 0.25 % (w/v) bromophenol blue, 0.25 % (w/v) xylene cyanol FF

Tris EDTA (TE)

10 mM Tris, pH 7.5, 0.1 mM EDTA

Transformation buffer 1 (TB1)

30 mMKOAc 100 mM RbCl, 10 mM $CaCl_2$ ($2H_2O$), 50 mM $MnCl_2$ ($4H_2O$), 15% glycerol; adjusted to pH 5.8 and filter sterilised

Transformation buffer 2 (TB2)

10 mM 3-N-morpholinopropanesulfonic acid (MOPS), 10 mM RbCl, 75 mM $CaCl_2$ ($2H_2O$), 15 % glycerol; adjusted to pH 6.5 and filter sterilised

2.2.5 PCR primers

PCR primers are listed in Table 2.1, and were synthesised by Geneworks (Thebarton, SA, Australia). Primers were dissolved in Milli-Q H_2O to 100 ng μl^{-1} before use.

2.3 Methods

2.3.1 Amplification of DNA by PCR

cDNA segments encoding full-length IGFBP-2, as well as its N- and C-domains, were amplified using polymerase chain reactions (PCR). The DM003M plasmid vector containing the full-length human IGFBP-2 sequence in the pXMT-2 vector (Forbes et al., 1998) was used as PCR templates. PCR reactions were performed in 40 µl volume, containing 1 X PCR buffer (100 mM Tris-HCl, pH 8.85, 250 mM KCl, 50 mM $(NH_4)_2SO_4$, 20 mM $MgSO_4$), 250 µM deoxynucleotide triphosphates, 10 % DMSO, 2.5 units of Pfu DNA polymerase, 200 ng of oligonucleotide primers and 20 ng of template plasmid DNA. PCR reactions were performed in a Perkin Elmer Cetus Thermal Cycler (Norwalk, CT, USA) using the following program: initial denaturation at 92 °C for 3 min; followed by 30 cycles, each of which contained denaturation at 92 °C for 1 min; annealing at 60 °C for 1 min and extension at 72 °C for 1.5 min. A final extension at 72 °C for 10 min was performed before terminating the reaction by lowering the temperature to 4 °C. PCR products were identified on 1-2 % agarose gels against molecular weight markers as described in 2.3.4.

The amplified DNA was precipitated to remove the unincorporated primers and dNTPs. Thus 25 µl of the PCR product was mixed with 2.5 µl of 3 MNaOAc(pH 5.2) and 63 µl of -20 °C cold ethanol, and incubated at -20 °C for 10 min. The sample was then centrifuged at 13000 x g for 10 min. The pellet was washed with 100 µl of 75% ethanol, and lyophilised under vacuum (Speed-Vac, Savant). The DNA was dissolved in 20 µl Milli-Q H_2O before restriction endonuclease digestion.

Table 2.1 PCR primers

bBP2NCO3Cfw	**5' ttttt tcc atg gca ctg gaa gtt ctg ttc cag ggg ccc gag gtg ctg ttc cgc tgc cc 3'** *Up-stream primer introducing an Nco I resstriction site and a 3C protease cleveage site.*
bBP2NCOTEVfw	**5' ta tta tcc atg gca gaa aac ctg tac ttc cag ggg gag gtg ctg ttc cgc tgc cc 3'** *Up-stream primer introducing an Nco I restriction site and a TEV protease cleveage site.*
bBP-2HISRev	**5' ttt tac gaa ttc tta gtg gtg gtg gtg gtg gtg ctg cat ccg ctg ggt gtg cac 3'** *Down-stream primer introducing a $(His)_6$ tag, a stop codon, and an Eco RI restriction site*
3C183hBP2Rev	**5' ttt ttt ctc gag cag gcc aag gtg atg ctt gc 3'** *Down-stream primer introducing a Xho I restriction site.*
3C183hBP2Fw	**5' ttt ttt gcc atg gcg ctc gag gtg ctg ttt cag ggc cca ccc cct gcc agg act cc 3'** *Up-stream primer introducing Nco I and Xho I restriction sites and a 3C protease cleveage site.*
IGFBP2Rev	**5' ttt tac gaa ttc tta ctg cat ccg ctg ggt gtg cac 3'** *Down-stream primer introducing a stop codon and an Eco RI restriction site.*
bBP2C1363Cf	**5' ttt ttt gcc atg gca ctg gaa gtt ctg ttc cag ggg ccc ggt cgg aag ccc ctc aag tc 3'** *Up-stream primer introducing an Nco I restriction sites and a 3C protease cleveage site.*
138hBP2Rev	**5' ttt tac gaa ttc tta gcc tcc ccc gcc caa cat gtt c 3'** *Down-stream primerintroducing a stop codon and an Eco RI restriction site.*

Restriction enzyme sites are underlined. Sequences complementary to the cDNA of IGFBP-2 are coloured blue. 3C and TEV protease cleavage sites are coloured pink and red, respectively.

2.3.2 Mini-preparation of plasmid DNA

Plasmid DNA was propagated in the *E. coli* strain DH5-α. 5 ml of LB (100 μg/ml ampicillin) was inoculated with a single colony containing the plasmid of interest from a freshly streaked selective plate. After overnight growth at 37 °C with shaking, the cells were harvested by centrifugation at 3200 x g for 5 min at room temperature.

Plasmid DNA was purified using the Spin Miniprep kit (QIAGEN) according to the manufacturer's instruction. Briefly, cell pellets were resuspended in 250 μl P1 buffer (Tris-EDTA buffer) and transferred to a microcentrifuge tube. 250 μl P2 buffer (200 mM NaOH, 1% SDS) was added and mixed thoroughly to lyse the cells. 350 μl N3 buffer (neutralization buffer) was added and the tube was inverted 6 times and immediately centrifuged at 17900 x g for 10 min. the supernatant was passed through a column containing a silica membrane. The column was washed with 500 μl PB buffer and 750 μl PE buffer. The DNA was eluted in 50 μl Milli-Q H_2O.

2.3.3 Restriction endonuclease digestion of DNA

Both the PCR products and the vector plasmids were treated by specific restriction endonucleases, under conditions recommended by the supplier of the enzyme.

For PCR products, reactions were performed in 40 μl, containing 1X buffer, 5 units of each endonuclease, and 20 μl of PCR product (resuspended in H_2O after precipitation). The reactions were incubated at 37 °C for 3 h. Digestion was terminated by inactivation of the endonuclease at 65 °C for 20 min.

2.3.4 Agarose gel electrophoresis of DNA

DNA samples were analysed by separation on agarose gels using TBE running buffer (100 mM Tris-borate, pH 8.3, 2.5 mM EDTA). The concentration of the agarose gels was 1-2 % depending on the size of the DNA sample being electrophoresed. Samples were loaded into the wells with the loading buffer (1X: 2 % glycerol, 0.04 % bromophenol blue, 0.04 % xylene cyanol in 1 x TBE). Electrophoresis was usually performed at 100 V. The gels were then stained with 1 μg ml^{-1} ethidium bromide in

TBE for 10 min followed by destaining in H_2O for 5 min. The DNA was visualised under long- or short-wave UV light as appropriate, and photos were recorded.

2.3.5 Isolation and purification of DNA from agarose gel

After restriction endonuclease digestion, both the cDNA segments and the plasmid vector DNA were purified from the reactions using agarose gel electrophoresis. Thus, the digestion products were electrophoresed at 80 V on an agarose gel. Gel slices containing the desired DNA fragments were excised from the gel under long wave UV light.

The DNA was then isolated from the agarose and purified using the UltraClean GelSpin DNA purification kit (MO BIO Laboratories) according to the manufacturer's instructions. In brief, the gel band was incubated in the GelBind buffer ($NaClO_4$ solution) at 65 °C for 5 min or until the gel was melted. The solution was then passed through a spin filter provided by the kit. The flow through was reloaded and passed through the spin field. The spin filter was washed with GelWash buffer (Tris/Ethanol solution). DNA was eluted from the spin filter using 50 µl H_2O.

2.3.6 Dephosphorylation

After endonuclease digestion and purification from agarose gel, the plasmid vector DNA was 5'-dephosphoryated using the shrimp alkaline phosphatase (USB Co). Thus, 50 µl of plasmid DNA was treated by 1 unit of phosphatase in 1X reaction buffer (20 mM Tris-HCl, pH 8.0, 10 mM $MgCl_2$) at 37 °C for 1 h, followed by incubation at 65 °C for 15 min to inactivate the enzyme.

2.3.7 DNA Ligation

The endonuclease digested cDNA fragments and the digested and 5'-dephosphorylated plasmid vector DNA were ligated at a molar ratio ≥ 3:1 in ligase buffer (66 mM Tris-HCl pH 7.6, 6.6 mM $MgCl_2$, 50 µM ATP) with 1 unit of T4 DNA

ligase at 16 °C overnight in a total volume of 10 µl. Negative controls included no insert DNA ligation and no ligase ligation.

2.3.8 Preparation of competent cells

Competent *E. coli* cells were prepared as described by CLONTECH Laboratories, Inc (Palo Alto, CA, USA). Fresh LB (5 ml) was inoculated with a single colony and grown overnight at 37 °C with shaking. A fresh 10 ml of LB was inoculated with the overnight culture (330 µl) and grown to an A_{600} of approximately 0.6 before subculturing 5 ml into 100 ml fresh LB. This culture was grown to an A_{600} of 0.6 then incubated on ice for 5 min. The cells were pelleted by 4000 x g centrifugation at 4°C for 5 min, and resuspended in 4 ml of ice cold TB1 solution on ice for 5 min. The cells were then pelleted by 4000 x g centrifugation at 4 °C for 5 min, and resuspended in 4 ml of ice cold TB2 solution. The cells were incubated on ice for 15 min prior to use or aliquoted and stored at -80 °C in 100 µl aliquots.

2.3.9 Transformation of bacterial cells

5 µl of the ligation reaction was added to 50 µl competent cells and incubated on ice for 30 min. The cells were heat-shocked at 42 °C for 90 sec, followed by recovering on ice for 5 min and at room temperature for 5 min. The mixture was subsequently incubated in 1 ml LB at 37 °C for 30 min with shaking. The cells were then centrifuged at 6500 x g for 2 min, resuspended in 100 µl LB and spread on LB plates (100 µg/ml Ampicillin), and incubated overnight at 37 °C.

2.3.10 Screening of clones

Three to six colonies of each transformation were screened for plasmids that contained inserted fragment of desired size. Thus, 2 ml LB (100 µg/ml Ampicillin) was inoculated with a single colony from the selective plate and grown overnight at 37 °C with shaking. Plasmid DNA was prepared from 1.9 ml of the overnight culture as described in 2.3.2 and dissolved in 50 µl Milli-Q H_2O. 10 µl of the plasmid DNA was

digested by appropriate restriction endonucleases as described in 2.3.3. Digestion products were separated on agarose gel against molecular weight markers as described in 2.3.4.

2.3.11 DNA sequencing

Selective plasmids that contain inserted fragments of the desired size were subject to DNA sequencing. The sequencing reaction contained 11 μl of the plasmid DNA, 1 μl of sequencing primer and 8 μl of ABI PRISM BigDye mixture (Applied Biosystems), which included A/C/G/T-Dye terminators, dATP, dCTP, dGTP, dTTP, Tris-HCl pH 9.0, $MgCl_2$, thermal stable pyrophosphatase and AmpliTaq polymerase. The sequencing primers were up- and down-stream universal primers that complement to the T7 or Sp6 promoter regions in the pET32a(+) vector. In the cases when one reaction was unable to determine the whole inserted sequence, more reactions were performed using different primers to generate overlapping sequences. The forward and reverse primers for IGFBP-2 cDNA amplification listed in Table 2.1 were also used in sequencing some constructs. The sequencing reactions were carried out using 30 cycles of PCR amplifications, each of which included denaturation at 96 °C for 30 sec, annealing at 50 °C for 15 sec, and extension at 72 °C for 4 min. Unincorporated dye terminators were then removed using the DyeEx Spin kit (QIAGEN) according to manufacturer's instructions. In brief, 20 μl of sequencing PCR reaction was passed through the spin column containing gel-filtration resins. The flow through containing the DNA was dried in a speed vac at room temperature for 25 min.

The resultant DNA samples were analysed by the DNA sequencing service at the Institute of Medical and Veterinary Science (IMVS), Adelaide, Australia, using a model 377 ABI sequencing machine, and the DNA sequences were reported.

2.4 Results and discussion

2.4.1 Wild-type IGFBP-2 constructs

Schematic representations of the pET32a(+) vector and the incorporated 3C and TEV protease cleavage sites are shown in Figure 2.2. The strategy for generating wild-type full-length IGFBP-2 constructs is illustrated in Figure 2.3.

2.4.2 IGFBP-2 (3C protease site at 183) construct

A construct was generated for expressing the full-length IGFBP-2 protein that contains a TEV protease cleavage site at the N-terminus of IGFBP-2 and a 3C protease cleavage site in the linker domain preceding Gly183 (Figure 2.4).

2.4.3 IGFBP-2 N-domain constructs

Similar to the preparation of full-length IGFBP-2 constructs, a construct for expressing $^{1\text{-}138}$IGFBP-2 (N-BP-2) was generated (Figure 2.5).

2.4.4 IGFBP-2 C-domain constructs

Constructs for expressing $^{183\text{-}289}$IGFBP-2 (C-BP-2) and $^{141\text{-}289}$IGFBP-2 (Large-C-BP-2) using the pET32a(+) vector were generated as demonstrated in Figure 2.6.

2.4.5 Results of DNA work

The results of PCR amplifications are shown in Figure 2.7. PCR reactions were performed routinely with 10 % DMSO as described (Sun et al., 1993) to overcome possible secondary structure formation within the GC rich regions in the IGFBP-2 cDNA. Previously Carrick and co-workers found that 10 % DMSO assisted successful amplification of IGFBP-2, especially the N-terminal domain region (Carrick, 2001, PhD thesis, The University of Adelaide).

Both the PCR amplified fragment and the pET32a(+) vector were treated by appropriate endonucleases. The desired fragments were then isolated and purified using agarose gel electrophoresis. The IGFBP-2 cDNA fragments were then ligated into the pET32a(+) vector. The resultant plasmid DNA was used to transformed *E. coli* DH5α competent cells. 100 to 200 colonies were usually observed on the plates, while 0-10 colonies can be found on the "no insert DNA" or "no ligase ligation" plate, resulting from the self-ligated or un-cleaved vectors. Diagnostic restriction digestion was used to check the inserts of 3-6 colonies of each plate (Figure 2.8). Selected clones that had inserts of the correct size were subjected to DNA sequencing, which confirmed the correct sequence of IGFBP-2 cDNA inserts.

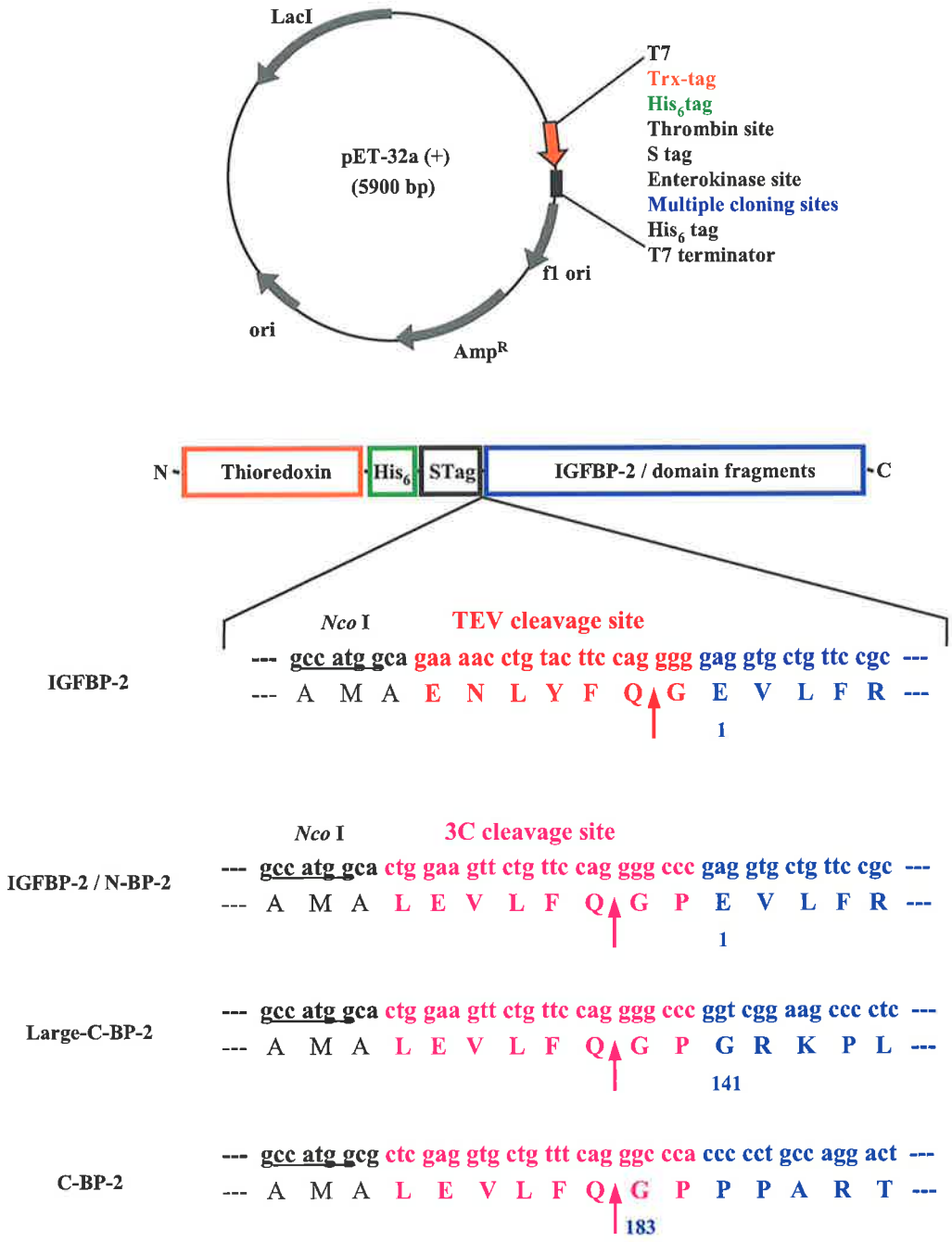


Figure 2.2 pET32a (+) vector and the incorporated 3C and TEV cleavage sites

The pET32a (+) expression vector (Novagen, Madison, USA) is designed for high-level expression of proteins fused with a 109 amino acid thioredoxin tag. The resultant protein also contains His_6-tag and S-tag for detection and purification. The pET32a (+) plasmid is under T7 transcriptional regulation and contains the β-lactamase ampicillin resistance gene (AmpR). The incorporated 3C and TEV cleavage sites are shown.

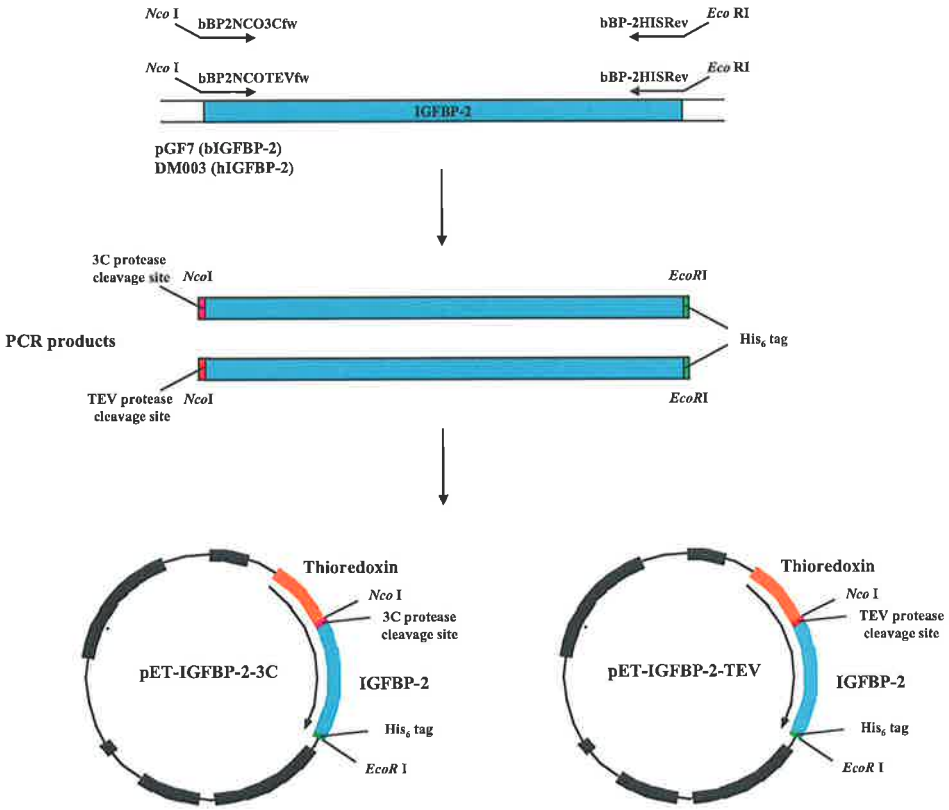


Figure 2.3 Cloning of IGFBP-2 into pET32a expression vector

Full-length bovine and human IGFBP-2 cDNA were amplified by PCR using the pGF7 and DM003M plasmids as templates, respectively. The PCR primers are represented as arrows above the cDNA. Forward primers bBP2NCO3Cfw (for incorporation of a 3C cleavage site) and bBP2NCOTEVfw (for incorporation of a TEV cleavage site), and reverse primer bBP-2HISRev were used. The fragments were digested using *Nco* I and *Eco* RI, and cloned into the pET32a (+) vector.

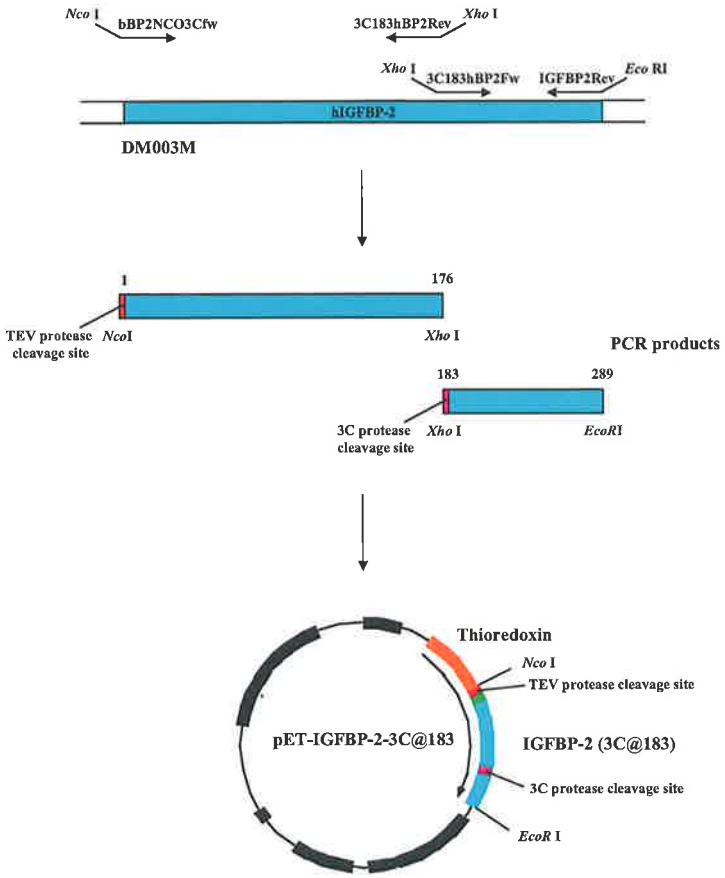


Figure 2.4 Cloning of IGFBP-2 (3C at 183) into pET32a expression vector

IGFBP-2(3C@183) cDNA was amplified by PCR using the DM003M plasmid as template. Forward primer bBP2NCO3Cfw and reverse primer 3C183hBP2Rev were used to amplify the $^{1\text{-}176}$IGFBP-2 fragment, while forward primer 3C183hBP2Fw and reverse primer IGFBP2Rev were used to amplify the $^{177\text{-}289}$IGFBP-2 fragment containing amino acid mutations (177-183) to a 3C protease cleavage site. These two fragments were digested using *Nco* I, *Xho* I, and *Eco* RI, and then ligated into one pET32a (+) vector.

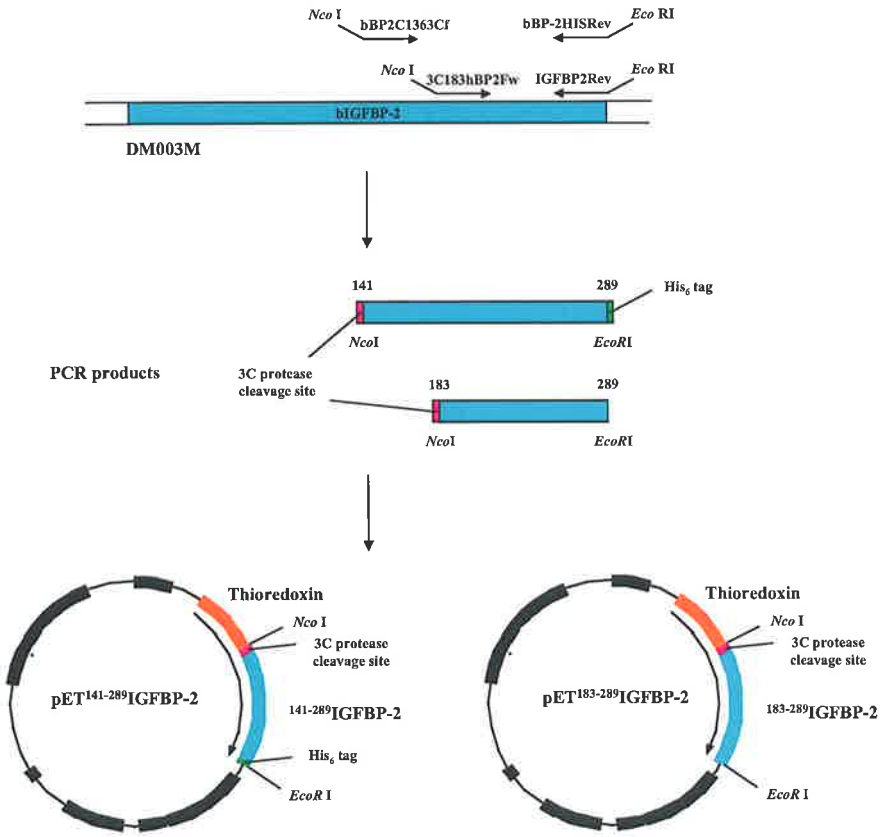


Figure 2.5 Cloning of IGFBP-2 C-domain into pET32a expression vector

cDNAs of two C-domain fragments were amplified by PCR using the DM003M plasmid as template. Forward primer bBP2C136Cf and reverse primer bBP2-HISRev were used to amplify the $^{141\text{-}289}$IGFBP-2 fragment, while forward primer 3C183hBP2Fw and reverse primer IGFBP2Rev were used to amplify the $^{183\text{-}289}$IGFBP-2 fragment. The fragments were digested using *Nco* I and *Eco* RI, and cloned into the pET32a (+) vector.

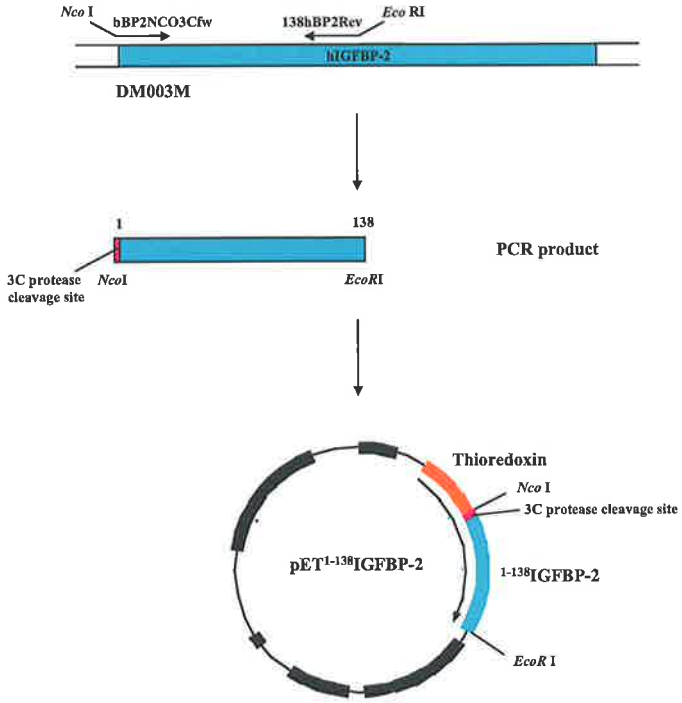


Figure 2.6 Cloning of IGFBP-2 N-domain into pET32a expression vector

cDNA of $^{1\text{-}183}$IGFBP-2 was amplified by PCR using the DM003M plasmid as template. Forward primer bBP2NCO3Cfw and reverse primer 138hBP2Rev were used. The fragment was digested using *Nco* I and *Eco*RI, and cloned into the pET32a (+) vector.

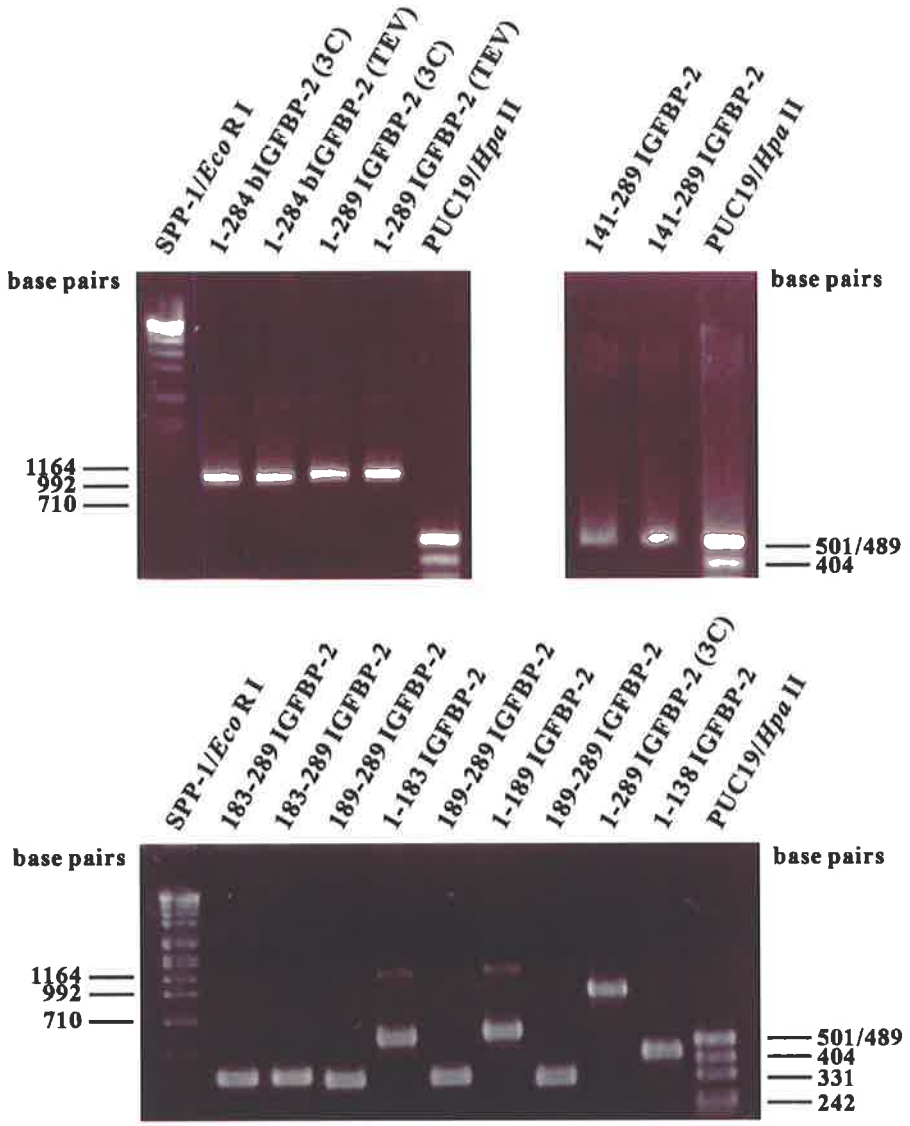


Figure 2.7 Polymerase chain reaction (PCR) of cDNA encoding IGFBP-2

PCR reactions were analysed on 1.5 % agarose gels as described in section 2.3.4. 4 µl of the 40 µl reaction were loaded onto the gel. 500 ng of molecular weight markers (SPP-1/*Eco* RI and pUC19/*Hpa* II) were also loaded. Amino acid regions of the amplified fragment are labelled on the top of each lane. Molecular weight sizes (base pairs) are indicated.

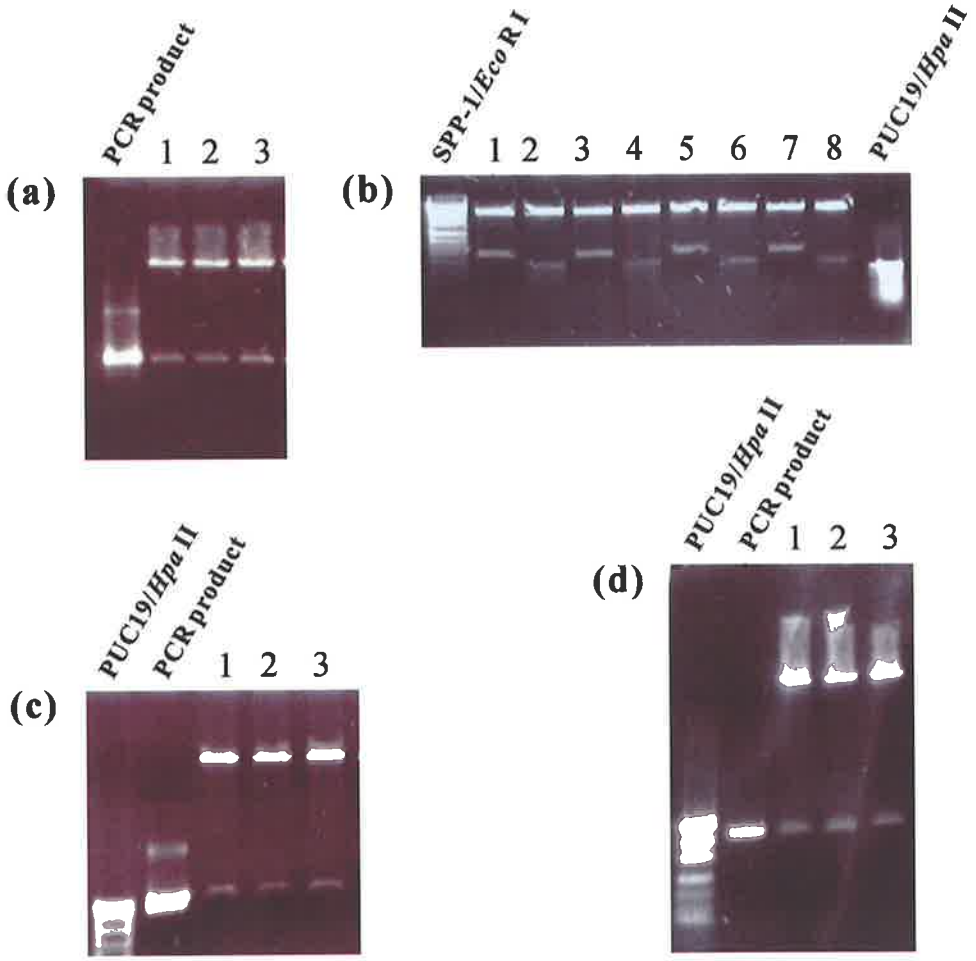


Figure 2.8 Screening of clones

To screen the clones containing target cDNA inserts, digested plasmid DNA samples were analyzed on 1.2 % agarose gel against DNA markers and the target PCR products. Lanes 1, 2, and 3 in (a), (c), and (d), and lanes 1, 3, 5, 7 in (b) were samples of three or four clones digested using *Nco* I and *Eco* RI. (a) pET-IGFBP-2-3C construct. (b) pET-IGFBP-2(3C@183) construct. lanes 2, 4, 6, 8 were the samples of 1, 3, 5, 7, respectively, treated by *Xho* I showing the incorporated 3C protease site at 183. (c) pET-$^{141\text{-}289}$IGFBP-2 construct. (d) pET-$^{1\text{-}138}$IGFBP-2 construct.

Chapter 3

Preparation and Characterization of IGFBP-2 N- and C-terminal Domains

3.1 Introduction

Previously, IGFBP-2 N- and C-domain fragments have been expressed in both COS-1 monkey kidney cells and *E. coli* cells (Carrick et al., 2001). These authors found that the enterokinase cleavage, which was performed to remove the fusion partner, was not sufficiently specific (Carrick, 2002, PhD thesis, The University of Adelaide). Thus, either a 3C protease or a TEV protease cleavage sequence was incorporated into the expression vector to replace that for enterokinase as described in the previous Chapter.

In this Chapter, N- and C-domain fragments, as well as full-length IGFBP-2 proteins were expressed in E. *coli* cells using the expression vectors described in Chapter 2. Cleavage of the thioredoxin-IGFBP-2 fusion proteins by 3C and TEV proteases was investigated and the cleavage conditions were optimised. Large amount of unlabelled N- and C-domain proteins were prepared using the improved purification protocol. Mass spectroscopy and N-terminal sequencing were performed to determine the integrity and purity of the proteins.

There has been a marked discrepancy in the literature regarding the IGF binding affinity of IGFBP C-domain fragments, which may have reflected a combination of genuine binding differences together with the technical differences, such as the different specificity and sensitivity of the methods used. Therefore, in this Chapter, binding of the recombinant N-BP-2 and C-BP-2 fragments were assessed using two methods, i.e., cross-linking and BIAcore analysis.

NMR studies of proteins with more than 70 amino acid residues often require labelling of the protein with NMR-active nitrogen and carbon isotopes, ^{15}N and ^{13}C, respectively. Therefore, expression of the N- and C-domain fragments in minimal media was investigated. ^{15}N-labelled and $^{15}N/^{13}C$-labelled C-BP-2 proteins, as well as ^{15}N-labelled N-BP-2 protein were prepared.

3.2 Materials

3.2.1 Reagents and kits

$^{15}NH_4Cl$ and $^{13}C_6$-glucose were purchased form Cambridge Isotope laboratories Inc. (Andover, MA, USA). Tricine, dithiothreitol (DTT), Coomassie brilliant blue R250, imidazole, nickel sulphate, polyoxethylene-sorbitan monolaurate (Tween 20), sodium dodecyl sulphate (SDS), N,N,N'N'-tetramethyl ethylene-diamine (TEMED), tris-HCl-(hydroxymethyl) aminomethane (Tris-HCl), and protease inhibitory cocktail were purchased from Sigma (St. Louis, MO, USA). Isopropyl-β-D-thiogalactopyranosid (IPTG), acetic acid and HPLC-grade methanol and acetonitrile were obtained from BDH Chemicals Australia (Kilsyth, VIC, Australia). Sequencing Grade trifluoroacetic acid (TFA) was obtained from Perkin Elmer, Applied Biosystems Division (Warrington, Great Britain). Affi-Gel 10, acrylamide, bis-acrylamide, and the Mini-Protean II Gel system for casting and running protein gels were purchased from Bio-Rad (North Ryde, NSW, Australia). Nitrocellulose membranes (0.45 μm) were from Schleicher and Schuell (Dassel, Germany), while membranes for filtering were obtained from Millipore Corp (Bedford, MA, USA). 3C protease (PreScission™ protease) was purchased from Amersham Biosciences (Piscataway, NJ, USA). TEV protease (AcTEV™ protease) and Mark12™ unstained standard protein marker were purchased from Invitrogen life technologies (Carlsbad, CA, USA). Prestained protein marker was purchased from New England BioLabs (Beverly, MA, USA). Zymatrix nickel-iminodiacetate (Ni-IDA) affinity resin was purchased from Bioserve (Heidleberg West, VIC, Australia). HPLC was performed using two Waters 510 HPLC pumps, an online Waters 486 Tunable Absorbance Detector and Millennium Software from Waters (Sydney, NSW, Australia). The PE Brownlee™ Aquapore C4 (100 x 10 mm) BU-300 column with 20 μm particles was purchased from Rainin LC and Supplies (Mulgrave, Victoria, Australia). The Delta-Pak PrepLC (25 x 100 mm) column segments and guard packs were purchased from Waters (Sydney, NSW, Australia).

The BIAcore™-2000 instrument, CM5 sensor chips, surfactant P20 and amine coupling reagents, N-ethyl-N'-[(dimethylamino) propy]] carbodi-imide (EDC), N-hydroxysuccin-imide (NHS) and ethanolamine were purchased from BIAcore AB

(Uppsala, Sweden). N-2-Hydroxyethylpiperazine-N-2-ethanesulphonic acid (HEPES) was obtained from BDH Laboratory Supplies (Poole, England). ^{125}I-IGF-II was from ProSearch$_t$ (Australia). All other chemicals were of analytical reagent grade. IGFBP-2 standard, IGF-I and IGF-II for BIAcore and cross-linking analysis were purchased from GroPep (Adelaide, Australia).

3.2.2 Bacterial strain

E. Coli BL21: *E. coli* B F$^-$dam opt hods ($b_o{}^m b^-$) gal (DE3)

3.2.3 Minimal media for isotope-labelled protein expression

Rich media for unlabelled protein expression are described in Chapter 2.

Modified M9 minimal medium

To 965 ml of 3 % M9 salt, add the following filter-sterilised solutions under sterile conditions: 20 ml of 5 % ^{15}N-NH_4Cl, 10 ml of 20 % glucose (use $^{13}C_6$-glucose for $^{15}N/^{13}C$-labelling), 2 ml of 1 M $MgSO_4$, 0.1 ml of 1 M $CaCl_2$, 2 ml of 500 X Trace element stock, 1 ml of 4 % thiamine, and ampicillin (100 μg/ml).

3 % M9 salt (1 L)

Na_2HPO_4	6 g
KH_2PO_4	3 g
NaCl	0.5 g

Adjust to pH 7.4 and autoclave

500 X Trace element stock (100 ml)

$FeSO_4.7H_2O$	1.0 g
$MnSO_4.H_2O$	0.255 g
$ZnSO_4.7H_2O$	0.43 g
$CuSO_4.5H_2O$	0.0375 g
Na_3citrate	4.4 g
HCl (conc.)	2 ml

MOPS minimal medium

For 1 L of MOPS minimal medium, the following filter-sterilised solutions were mixed under sterile conditions: 95 ml of MOPS concentrate, 10 ml of 0.132 M K_2HPO_4, 5 ml of 1.9 M ^{15}N-NH_4Cl, 17.5 ml of 20 % glucose (use $^{13}C_6$-glucose for $^{15}N/^{13}C$-labelling), 1 ml of 100mg/ml Ampicillin, and 871.5 ml of Milli-Q H_2O (Neidhardt et al., 1974).

MOPS concentrate

The following solutions were mixed in this order to yield 950 ml of MOPS concentrate: 400 ml of 1 M MOPS (pH 7.4), 40 ml of 1 M Tricine (pH7.4), 10 ml of 0.01 M $FeSO_4$, 10 ml of 0.276 M K_2SO_4, 10 ml of 0.5 mM $CaCl_2$, 10 ml of 0.528 M $MgCl_2$, 125 ml of 4 M NaCl, 10 ml of micro-nutrient stock, and 335 ml of Milli-Q H_2O.

Micronutrient stock

3 μM $(NH_4)_{6Mo7O24}$, 0.4 mM H_2BO_4, 30 μM $CoCl_2$, 10 μM $CuSO_4$, 80 μM $MnCl_2$, 10 μM $ZnSO_4$

3.2.4 Nickel-IDA Chromatography buffers

$NiSO_4$ charge buffer

50 mM $NiSO_4$; adjusted to pH 7.9 and filter sterilised (0.45 μm)

Binding buffer

5 mM imidazole, 0.5 M NaCl, 20 mM Tris; adjusted to pH 7.9 and filter sterilised (0.45 μm)

Wash buffer

10 mM imidazole, 0.5 M NaCl, 20 mM Tris; adjusted to pH 7.9 and filter sterilised (0.45 μm)

Elute buffer

0.5 M imidazole, 0.5 M NaCl, 20 mM Tris; adjusted to pH 7.9 and filter sterilised (0.45 μm)

Strip buffer

0.1 M EDTA, 0.5 M NaCl, 20 mM Tris; adjusted to pH 7.9 and filter sterilised (0.45 μm).

3.2.5 IGF affinity column buffer

50 mM Tris pH 7.5, 150 mM NaCl, 1 mM EDTA, 0.05 % (v/v) Tween-20

3.2.6 HPLC buffers

HPLC buffer A

0.1 % (v/v) TFA; filtered (0.2 μm)

HPLC buffer B

80 % acetonitrile, 0.088 % (v/v) TFA; filtered (0.2 μm)

3.2.7 3C and TEV protease cleavage buffer

Although recommended by the manufacturer, DTT was not supplemented in these buffers (see below).

1X 3C protease cleavage buffer

50 mM Tris-HCl pH 7.0, 150 mM NaCl, 1mM EDTA

1X TEV protease cleavage buffer

50 mM Tris-HCl pH 8.0, 0.5 mM EDTA

3.2.8 Buffers for protein electrophoresis

Lysis buffer
2% (v/v) NaOH, 10% (v/v) β-mercaptoethanol

2 x Protein loading buffer
125 mM Tris-HCl, 4 % (w/v) SDS, 10 % (v/v) glycerol, 0.1 % Bromophenol Blue pH 6.8

Polyacrylamide tricine gel running buffer
250 mM tricine, 25 mM Tris-HCl pH 8.3, 0.1 % (w/v) SDS

Gel fixing solution
50 % (v/v) ethanol, 10 % (v/v) acetic acid

Coomassie blue staining solution

0.5 % (w/v) Coomassie blue R-250, 10 % (v/v) acetic acid

Gel destain solution
10 % acetic acid

Gel washing solution
30 % (v/v) ethanol, 5 % (v/v) glycerol

Gel Drying Solution
5 % (v/v) glycerol, 30 % (v/v) ethanol

3.2.9 Buffers for cross-linking and BIAcore analysis

Phosphate buffered saline (PBS)

140 mM NaCl, 2.5 mM KCl, 1.6 mM KH_2PO_4, 15 mM Na_2HPO_4, pH 7.2

HEPES buffered saline (HBS)

10 mM HEPES, 150 mM NaCl, 3.4 mM EDTA, 0.05 % (v/v) surfactant P20; prepared fresh from a 10 x concentration of stock solution; adjusted to pH 7.4 and filter sterilised (0.2 μm)

Regeneration buffer

100 mM HCl, 500 mM NaCl

3.3 Methods

3.3.1 Expression of proteins in *E. coli*

Full-length IGFBP-2 and the N- and C-domains of IGFBP-2 were expressed in *E. coli* BL21 (DE3) cells using the expression vectors described in Chapter 2. For unlabelled protein expression, cells were grown in LB medium. Briefly, LB medium (100 μg/ml ampicillin) was inoculated with 1/10 volume of an overnight LB culture, grown at 37 °C with shaking. The $O.D._{600}$ of the culture was monitored. When $O.D._{600}$ reached ~0.6, IPTG was added to the culture to final concentration of 0.1 mM to induce

the expression of the recombinant protein. The proteins were expressed for a minimum of 4 h prior to pelleting the cells (5000 x g, 5 min, 4 °C). For ^{15}N-labelled or $^{15}N/^{13}C$-labelled protein expression, cells were grown in M9 medium, and $^{15}NH_4Cl$ (1 g L^{-1}), or $^{15}NH_4Cl$ (1 g L^{-1}) and $^{13}C_6$-glucose (2 g L^{-1}) were used as the nitrogen and carbon sources.

Cell culture samples (1 ml) were taken prior to induction of protein expression and upon cell harvest to check the expression of the recombinant proteins using SDS polyacrylamide gel electrophoresis as described in 3.3.5. In initial experiments, cell culture samples were also taken at different time points during the expression period.

3.3.2 Nickel-IDA chromatography

Thioredoxin fusion proteins of N- and C-domain fragments were purified using Ni-IDA column. The cell pellets were resuspended in the ice-cold Ni-IDA column binding buffer (1/10 volume of original culture) and protease inhibitor cocktail (Sigma) was added according to manufacturer's instructions. The cells were then homogenised by two passes through the French Press (1000 psi). 200 µl of the lysate was centrifuged at 10000 x g for 10 min; the supernatant and the pellets, washed thoroughly using the binding buffer, were analysed as the soluble fraction and insoluble fraction, respectively, in SDS polyacrylamide gel electrophoresis as described in 3.3.5. The rest of the lysate was centrifuged at 10000 x g for 10 min at 4 °C. The supernatant was filtered (0.8 µm) before loading onto a Ni-IDA column.

The Ni-IDA column contained approximately 10 ml of resin. Ni-IDA chromatography was performed using a peristaltic pump with an online Pharmacia UV-1 monitor set at 280 nm. The resin was charged with four column volumes of $NiSO_4$ charge buffer, and equilibrated with three column volumes of binding buffer prior to loading of the protein sample. The sample was loaded at a flow rate of 2 ml min^{-1}. Non-specifically bound proteins were washed from the column with firstly ten column volumes of binding buffer followed by six column volumes of wash buffer. Specifically bound His_6-tagged proteins were eluted using the elute buffer. The column was then stripped with five column volumes of strip buffer and stored at 4 °C.

3.3.3 IGF affinity chromatography

Thioredoxin fusion proteins of full-length IGFBP-2 were purified using IGF affinity column. Long IGF-I (pGH(1-11)-VN-IGF-I) was prepared by Ms Kerrie McNeil (The University of Adelaide) using the previously described protocol (King et al., 1992). Long IGF-I was coupled to Affi-Gel 10 (Bio-Rad) as recommended by the supplier. The column was equilibrated with IGF affinity column buffer prior to loading of the protein sample. The *E. coli* cell pellets were resuspended in IGF affinity column buffer and were homogenised as above. The supernatant was filtered (0.8 μm), and then was loaded onto the column using a peristaltic pump at a flow rate of 1 ml min^{-1}. Unbound proteins were washed from the column using IGF affinity column buffer until a baseline absorbance at 280 nm was obtained. Specifically bound protein was eluted in 0.5 M acetic acid. The fractions containing protein were pooled for further purification by reverse phase HPLC as described in 3.3.6.

3.3.4 Initial trials of 3C and TEV protease cleavage

Thioredoxin-IGFBP-2 fusion proteins, eluted from the IGF affinity column, were further purified by reverse phase HPLC as described in 3.3.6. 3C and TEV protease cleavage trials were then carried out on these fusion proteins. Initial cleavages were performed using the conditions recommended by the suppliers, including 1 mM DTT in the cleavage buffers. It was found that the presence of DTT changed the HPLC profile of the target proteins compared to standard protein samples (see below). 3C and TEV protease cleavages without DTT or with 0.1 mM or 1 mM DTT at varying pH and temperature were then compared in detail to determine the optimum conditions.

3.3.5 3C protease cleavage

For large scale purification of the unlabelled and isotope-labelled $^{183-289}$hIGFBP-2 (C-BP-2), the $^{1-138}$hIGFBP-2 (N-BP-2), and the unlabelled $^{141-289}$hIGFBP-2 (large C domain) proteins, the Ni-IDA column elute that contained the fusion protein was dialysed against 1 X 3C protease cleavage buffer at 4 °C for 24 h. Typically, from a 1 L culture fermentation, the Ni-IDA column elute was about 30 to 50 ml. An initial 5 L of

1 X 3C protease cleavage buffer was used in the dialysis, and it was changed to fresh buffer twice during the dialysis.

Aliquots of the protein samples were checked and quantified as described in 3.3.8 before 3C protease cleavage was performed. Concentration of the fusion protein in the cleavage reaction was about 0.5 to 1.0 mg ml^{-1}. Ten units of 3C protease per mg of fusion protein were added to the samples and the samples were then incubated at 4 °C. The cleavage reactions were monitored by running small aliquots of the samples on reverse phase HPLC as described in 3.3.8.

3.3.6 Preparative HPLC

After the cleavage reaction was completed, the protein sample was purified using reverse phase HPLC. Firstly, 0.2 ml of 1 % TFA per ml of the cleavage reaction was added to reduce the pH of the sample to 2.0-3.0. The sample was then filtered (0.2 μm), and loaded onto a Delta-Pak PrepLC column at a flow rate of 2 ml min^{-1} using a Waters 510 pump. The recombinant protein was isolated from the cleavage product using a 24-56% acetonitrile gradient over 160 min in the presence of 0.1% TFA at a flow rate of 2 ml min^{-1}. Fractions containing the target protein were pooled, quantified and lyophilised.

3.3.7 SDS polyacrylamide gel electrophoresis and staining

Samples from the protein expression and purification procedures, as well as purified protein samples were analysed using SDS polyacrylamide tricine gel (12%) electrophoresis under non-reducing conditions unless otherwise noted.

Un-induced and induced *E. coli* cell cultures (1 ml) were pelleted and the cells resuspended in lysis buffer to a final concentration equivalent to 50 μl per OD unit and boiled for 10 min. Other protein samples were heated to 100 °C for 5 min in 2X protein loading buffer prior to electrophoresis. The whole cell lysates (10 μl) and other protein samples were then loaded onto the gels and electrophoresed at 40 mA per gel until adequate separation was achieved.

Protein bands were visualised by Coomassie blue staining. Gels were fixed in the gel fixing solution for 20 min, stained in Coomassie blue staining solution for 60 min and then washed in gel destain solution until the background staining was negligible.

The gel was then washed in gel wash solution for 10 min and dried between cellophanes at room temperature.

3.3.8 Analytical HPLC and quantification of the protein

The protein purification procedures, as well as the quantity and purity of the purified protein samples were also routinely assessed using a C4 BU-300 column on an Agilent model 1100 HPLC system. Absorbance at 215 nm was measured and concentration determined based on absorbance ratios relative to accurately quantified IGF-I standards (GroPep).

3.3.9 Mass Spectroscopy and N-terminal Sequencing

Electrospray mass spectroscopy and N-terminal sequencing, both performed by Mr Chris Cursaro (School of Molecular and Biomedical Science, The University of Adelaide), were used to confirm the identity of the proteins. Lyophilised protein samples were submitted for electrospray mass spectroscopy on a Micromass Q-TOF mass spectrometer. N-terminal amino acid sequences were determined by Edman degradation using an Applied Biosystems Procise protein sequencer.

3.3.10 Cross-linking

Cross-linking was performed by Mr Julian Thompson (Department of Medicine, Monash University) to assess the binding of full-length IGFBP-2, N-BP-2, and C-BP-2 to IGF-II. 5 μg of C-BP-2, 1 μg of N-BP-2, or 20 ng of IGFBP-2 were incubated with 1.4 ng ^{125}I-IGF-II (200,000 cpm; ProSearch), with and without excess unlabelled IGF-II, at 4 °C overnight in 40 μL of PBS. Disuccinimidyl suberate was then added to the samples to a final concentration of 0.25 mM and incubated at 15 °C for 30 min. Cross-linking was terminated by the addition of Tris-HCl buffer (pH 7.5) to a concentration of 100 mM, followed by 5 min incubation at 15 °C. Samples were then boiled for 5 min with electrophoresis loading buffer and separated on a 15% SDS polyacrylamide gel, followed by autoradiography.

3.3.11 Preparation of BIAcore biosensor surface

IGF binding by full-length IGFBP-2 and N- and C-domain fragments was analysed using BIAcore 2000 (BIAcore AB, Uppsala, Sweden). Ligands were covalently immobilised on CM5 gold surfaces using the amine coupling method as described previously (Carrick et al., 2001). Firstly, the CM5 gold surfaces in the sensor chip were activated by injection of a fresh mixture of EDC and NHS solutions as per the manufacturer's instructions (BIAcore). The peptides to be immobilised (12.5 μg/ml in 50 mM sodium acetate pH 4.6) were injected onto the activated CM5 surface at 5 μl/minute. The surfaces were then deactivated by passage of 1 M ethanolamine. For IGF-I and IGF-II surfaces, negative control surfaces included a blank surface (no protein) and an insulin surface. Biosensor surfaces were coupled to final resonance values of approximately 100 response units (RU) for IGF/insulin surfaces, 200 RU for C-BP-2 surface, and 600 RU for Large-C-BP-2 and IGFBP-2 surfaces.

3.3.12 Acquisition and analysis of BIAcore data

Various concentrations of analytes were injected in the HBS running buffer at 25 °C during the association phase for 2 min at a flow rate of 40 μL min^{-1}, followed by a 10 min dissociation phase with running buffer alone. The flow rate was previously optimised by Lucic and co-workers to minimise the mass transport limitation effects (Lucic, 2000, PhD thesis, The University of Adelaide). In each experiment, duplicate samples were used for any given concentration, and samples were injected in random order. Between each sample, the biosensor surfaces were regenerated by a 200 sec injection of regeneration buffer (section 3.2.9).

Kinetic data for the interactions were derived by fitting experimental association and dissociation curves (sensorgrams) using the BIAevaluation program version 3.0 (BIAcore AB, Uppsala, Sweden). The sensorgrams were corrected for the non-specific response and bulk refractive index contribution by subtraction of the sensorgram of the reference flow cell. For the interaction between IGFs and full-length IGFBP-2, a two-state reaction (conformational change) model, i.e., a 1:1 binding of a analyte to a immobilised ligand followed by a conformational change, was used as this model was previously found to fit these sensorgrams better by Lucic and co-workers (Lucic, 2000,

PhD thesis, The University of Adelaide). For the interactions between IGFs and the N- and C-domains of IGFBP-2, the 1:1 Langmuir binding interaction describing 1:1 binding between an analyte and ligand was used to fit the data. The 1:1 stoichiometry for the interaction between IGFs and IGFBPs was previously established (Bourner et al., 1992). Models were fitted globally across the data sets, i.e, simultaneously across the entire concentration series.

3.4 Results

3.4.1 3C and TEV protease cleavage

3C and TEV protease cleavage trials were carried out on the thioredoxin-IGFBP-2 fusion proteins, which have either a 3C or a TEV protease cleavage site incorporated between the thioredoxin tag and IGFBP-2 protein. Both 3C and TEV cleavages were found to be highly specific in cleaving these fusion proteins, because using excess proteases and prolonged incubation times did not alter the product patterns as monitored by SDS polyacrylamide gel electrophoresis and analytical HPLC. The correctness of the cleavage was also confirmed by N-terminal protein sequencing (see below).

However, the reducing reagent DTT, which is recommended by the manufacturers to be a component (1 mM) in both 3C and TEV cleavage buffers, was found to change the HPLC profiles of full-length IGFBP-2, most likely due to the reduction of certain disulfide bonds in the molecule. As shown in Figure 3.1 (a and b), in the absence of DTT, TEV protease cleaves the fusion protein and yields a peak for bovine IGFBP-2 at 17.6 min using a 20 to 60 % acetonitrile gradient on analytical HPLC; the samples were then mixed with 1 mM DTT, as recommended for the cleavage reaction, and were analysed either immediately or after incubation at 30 °C for 30 min. It can been seen that, incubation in the cleavage buffer with DTT changed the 17.6 min peak to 20.7 min and 18.5 min peaks. This effect was still significant at 4 °C and with less DTT (0.1 mM), as can be seen in the 3C protease cleavage experiments (Figure 3.1(c-f)). The additional IGFBP-2 peaks generated during cleavage with DTT were probably partially reduced forms of IGFBP-2 but were not due to non-specific cleavages, because

cleavage of the fusion proteins without DTT did not produce these peaks. Furthermore, incubation of bovine IGFBP-2 sample (Kindly supplied by Dr Briony Forbes from a mammalian source) with DTT but no protease resulted in identical peak change pattern (Figure 3.1(g-h)). As would be expected for the disulfide bond reduction, the effect of DTT on IGFBP-2 was found to be more significant at higher temperature and higher pH.

To avoid its reducing effect on IGFBP-2, DTT was not supplemented in the cleavage buffers, although diluted (~1-5 μM) DTT was expected to originate from the protease storage buffers (1 mM in 3C protease and 5 mM in the TEV protease). Unfortunately, the absence of DTT resulted in reduced activities of 3C and TEV proteases, both of which are cysteine proteases, and thus larger amounts of the proteases and much longer cleavage time were necessary to complete the cleavages. Since the recommended cleavage conditions of the 3C protease are at 5 °C and pH 7.0, while those of TEV protease are at 30 °C and pH 8.0, 3C protease was chosen for large scale preparations of N- and C-domain fragments, and the cleavage was performed at 4 °C to minimise potential proteolytic degradation by other contaminating proteases during long incubation times.

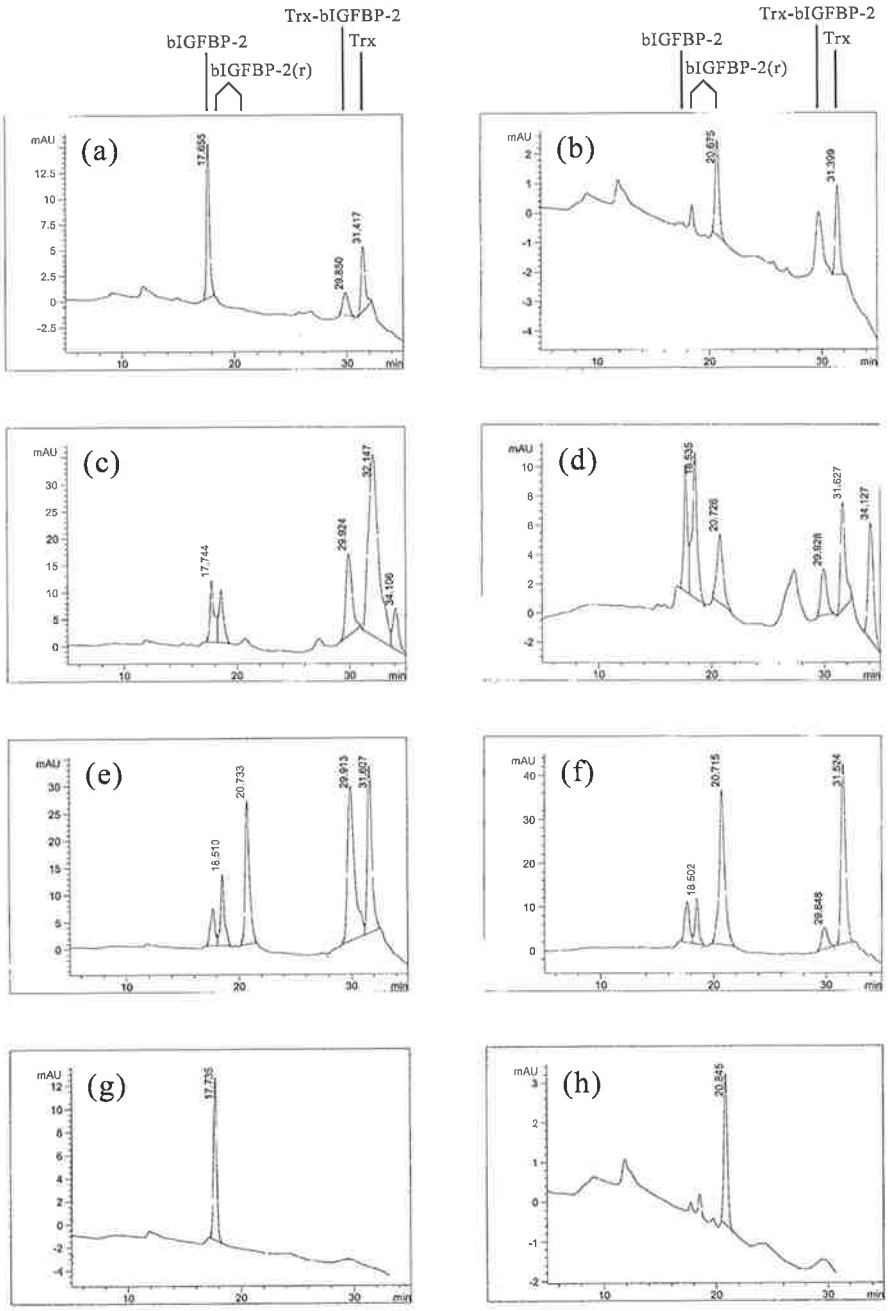


Figure 3.1 HPLC profiles showing the effects of DTT on IGFBP-2

Thioredoxin (Trx) fusion protein of bovine IGFBP-2 was cleaved by TEV protease without DTT, then mixed with DTT (1 mM) and was analysed on HPLC either immediately (a) or after incubation at 30 °C for 30 min (b); Trx-bIGFBP-2 was cleaved by 3C protease with either 0.1 mM (c and d) or 1 mM (e and f) DTT at 4 °C for 30 min (c and e) or 60 min (d and f); Mammalian cell sourced bovine IGFBP-2 standard sample (g) or after incubation with 1 mM DTT in 1 X TEV buffer at 30 °C for 30 min. Using this gradient, bIGFBP-2 has a peak at 17.7 min, whereas the partially reduced products (bIGFBP-2(r)) have peaks at 18.5 and 20.7 min.

3.4.2 Expression and purification of C-BP-2 and N-BP-2

In *E. coli* BL21 (DE3) cells, the majority of the full-length IGFBP-2 fusion proteins, expressed using the pET-IGFBP-2 or pET-IGFBP-2(3C@183) vectors, were found in the insoluble fraction (data not shown). Nevertheless, the fusion protein in the soluble fraction was purified under native conditions. Either a full-length IGFBP-2 protein (with a 3C cleavage site at 183) or the C-domain fragment was then cleaved from the thioredoxin-IGFBP-2(3C@183) protein using TEV protease or 3C protease, respectively (Figure 2.1). The IGF binding properties of these full-length or C-domain fragments were assessed using BIAcore, and compared to the native full-length IGFBP-2 or the C-domain fragment produced using the second vector (see below).

Large-scale expressions of unlabelled and isotope-labelled C-BP-2 fragments were performed using the pET-$^{183\text{-}289}$IGFBP-2 expression vector. The thioredoxin-C-BP-2 fusion protein was in the soluble fraction when expressed in the *E. coli* BL21 (DE3) cells. The fusion proteins were initially purified using Ni-IDA column under native conditions (Figure 3.2). The Ni-IDA column elute was dialysed against 3C protease cleavage buffer to remove the Ni^{2+}, and the fusion protein was subsequently cleaved using 3C protease. The volume of the cleavage reaction was 30 to 50 ml, containing ~ 0.5 mg ml^{-1} fusion protein. Because DTT was not added to the cleavage buffer, it was desirable to eliminate O_2 in the sample to prevent oxidization of the catalytic cysteine residue of the 3C protease. Thus, the samples were bubbled very gently with N_2 gas for 5 min before the 3C protease was added; care was taken to avoid foaming during this step. In addition, a greater amount of the protease was required to complete the cleavage reactions, compared to the cleavage conditions with DTT, as recommended by the manufacturer. It was found that the C-domain fragment product remained intact in this cleavage reaction at 4 °C up to 4 days, as monitored by analytical HPLC. Therefore, the cleavage reactions were allowed to proceed for 24 to 36 h, and at different time points aliquots were taken and analysed by HPLC. Additional 3C protease was added occasionally during the cleavage period. Final cleavage products were separated using HPLC (Figure 3.3). The C-BP-2 protein was quantified using HPLC and lyophilised.

12 mg of unlabelled C-BP-2 was purified from 2 L of LB medium culture, whereas 8 mg of ^{15}N-labelled C-BP-2, and 2.5 mg of ^{15}N/^{13}C-labelled C-BP-2 proteins

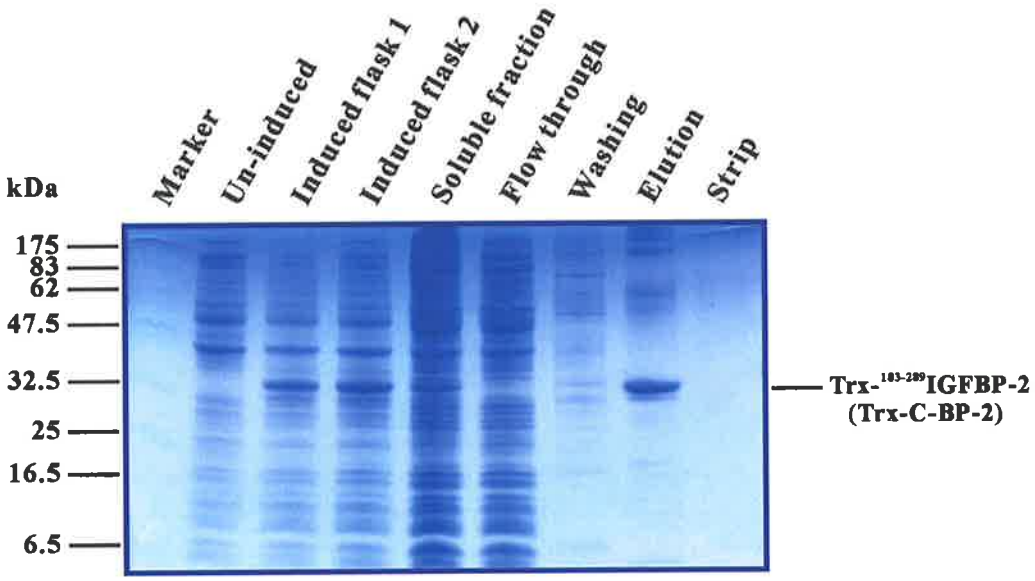


Figure 3.2 Expression and purification of unlabelled C-BP-2 protein

Uninduced and induced cell culture samples from the *E. coli* expression of Trx-$^{183-289}$IGFBP-2, and the protein samples at each state during the purification using nickel affinity chromatography (as described in section 3.3.2) were separated by electrophoresis on 12% polyacrylamide tricine gel and stained with Coomassie Blue-R250 (as described in section 3.3.7). Molecular weight markers are indicated on the left.

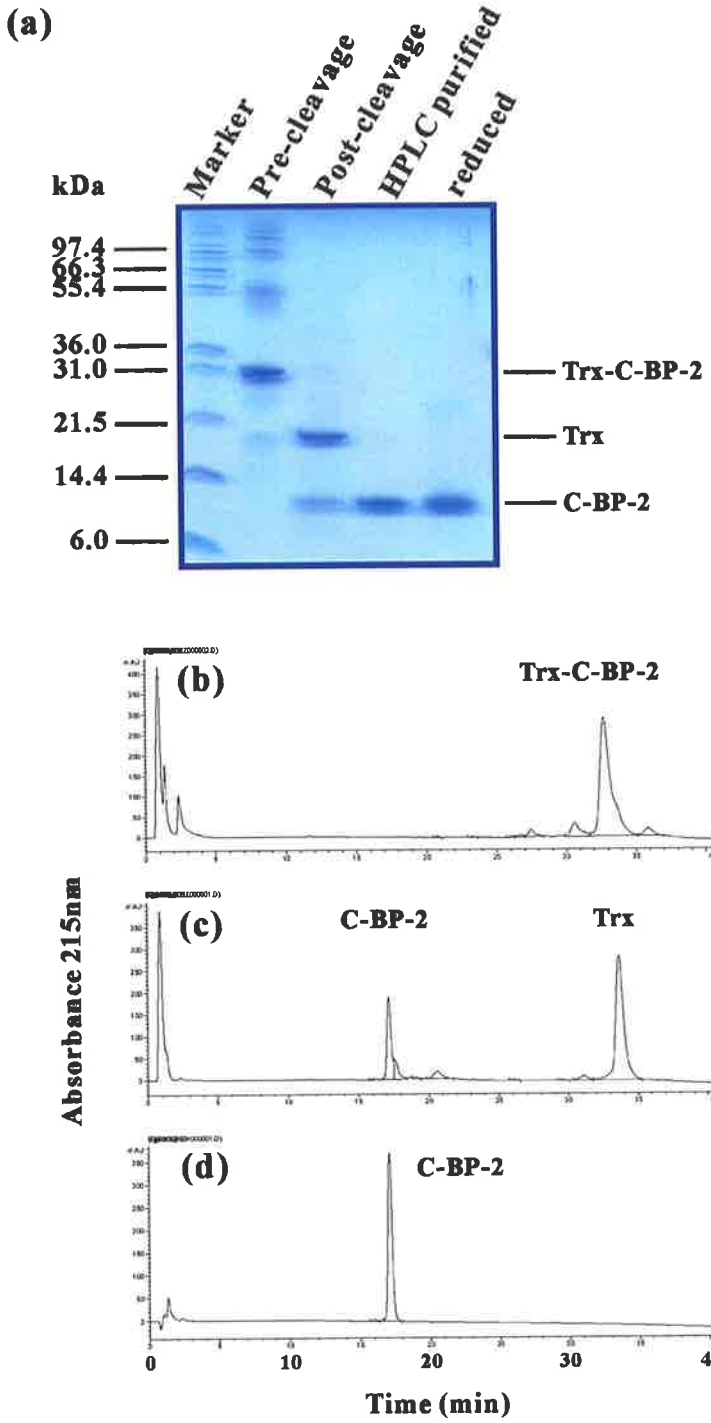


Figure 3.3 3C protease cleavage of thioredoxin-C-BP-2 fusion protein

Protein samples of Trx-$^{183\text{-}289}$IGFBP-2 before and after 3C protease cleavage (section 3.3.5), and the $^{183\text{-}289}$IGFBP-2 (C-BP-2) sample purified by preparative HPLC (section 3.3.6) were analysed by (a) electrophoresis on 12% polyacrylamide tricine gel and stained with Coomassie Blue-R250 (section 3.3.7), or (b-d) by analytical HPLC using a C4 BU-300 column on an Agilent model 1100 HPLC system (section 3.3.8). Absorbance at 215 nm was measured.

were purified from 2 L and 1 L of M9 minimal medium cultures, respectively, using the appropriate isotopes.

Another IGFBP-2 C-domain fragment, $^{141\text{-}289}$hIGFBP-2 (Large-C-BP-2), was expressed using the pET-$^{141\text{-}289}$IGFBP-2 vector. Interestingly, although the thioredoxin-Large-C-BP-2 fusion protein was expressed in the soluble fraction using LB medium (Figure 3.4), and 6 mg of the Large-C-BP-2 protein was purified from 1 L of LB medium culture (Figure 3.5), this fusion protein was found in the insoluble fraction when using either M9 minimal medium (Figure 3.6) or MOPS minimal medium. Therefore isotopic labelling of this fragment was not performed.

An IGFBP-2 N-domain fragment, $^{1\text{-}138}$hIGFBP-2 (N-BP-2), was expressed using the pET-$^{1\text{-}138}$IGFBP-2 vector (Figure 3.7), and purified using a similar purification approach (Figure 3.8). 6 mg of unlabelled N-BP-2 protein was purified from 2 L LB medium culture. 4 mg and 5 mg of ^{15}N-labelled N-BP-2 protein was purified from 2 L M9 medium and 2 L of MOPS medium culture, respectively.

3.4.3 Mass Spectroscopy and N-terminal Sequencing

Electrospray mass spectrometry data are summarized in Table 3.1. The detected mass data for unlabelled C-BP-2, Large-C-BP-2, and N-BP-2 were in excellent agreement with their theoretical molecular masses, confirming that three disulfide bonds were formed. Furthermore, it can be estimated from the data on the ^{15}N-labelled proteins and the ^{15}N/^{13}C-labelled protein that the incorporation of ^{15}N and ^{13}C was about 98 %. Amino acid sequencing confirmed the correct N-termini and showed the high purity of the samples (Figure 3.9).

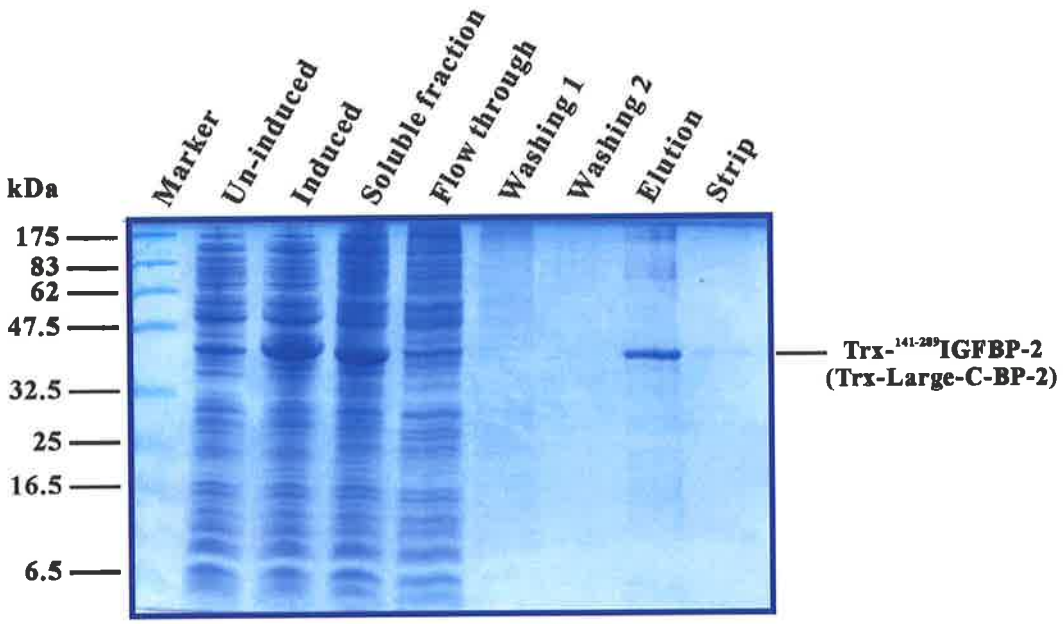


Figure 3.4 Expression and purification of unlabelled Large-C-BP-2 protein

Uninduced and induced cell culture samples from the *E. coli* expression of Trx-$^{141-289}$IGFBP-2, and the protein samples at each state during the purification using nickel affinity chromatography (as described in section 3.3.2) were separated by electrophoresis on 12% polyacrylamide tricine gel and stained with Coomassie Blue-R250 (as described in section 3.3.7). Molecular weight markers are indicated on the left.

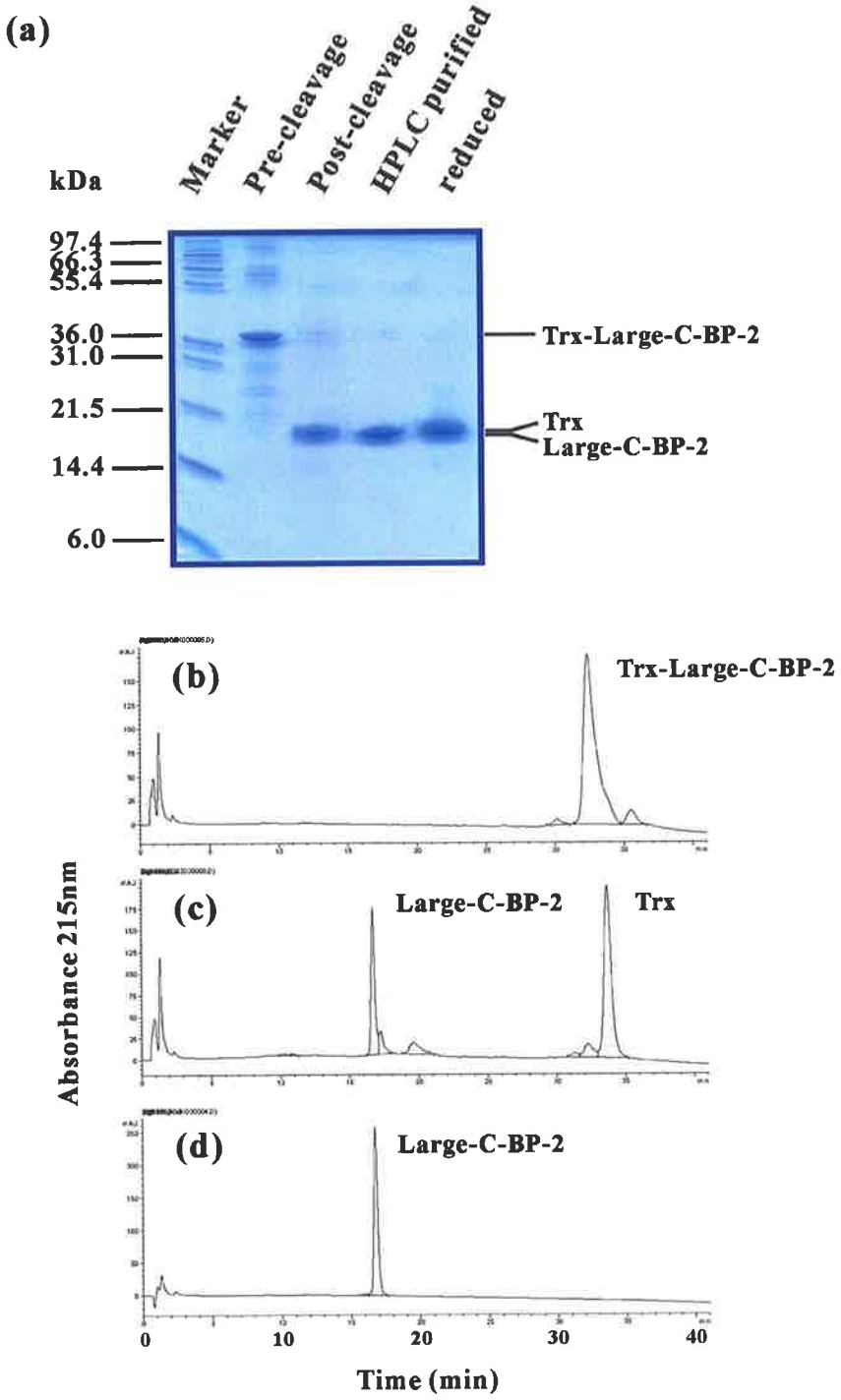


Figure 3.5 3C protease cleavage of thioredoxin-Large-C-BP-2 fusion protein

Protein samples of Trx-$^{141\text{-}289}$IGFBP-2 before and after 3C protease cleavage (section 3.3.5), and the $^{141\text{-}289}$IGFBP-2 (Large-C-BP-2) sample purified by preparative HPLC (section 3.3.6) were analysed by (a) electrophoresis on 12% polyacrylamide tricine gel and stained with Coomassie Blue-R250 (section 3.3.7), or (b-d) by analytical HPLC using a C4 BU-300 column on an Agilent model 1100 HPLC system (section 3.3.8). Absorbance at 215 nm was measured.

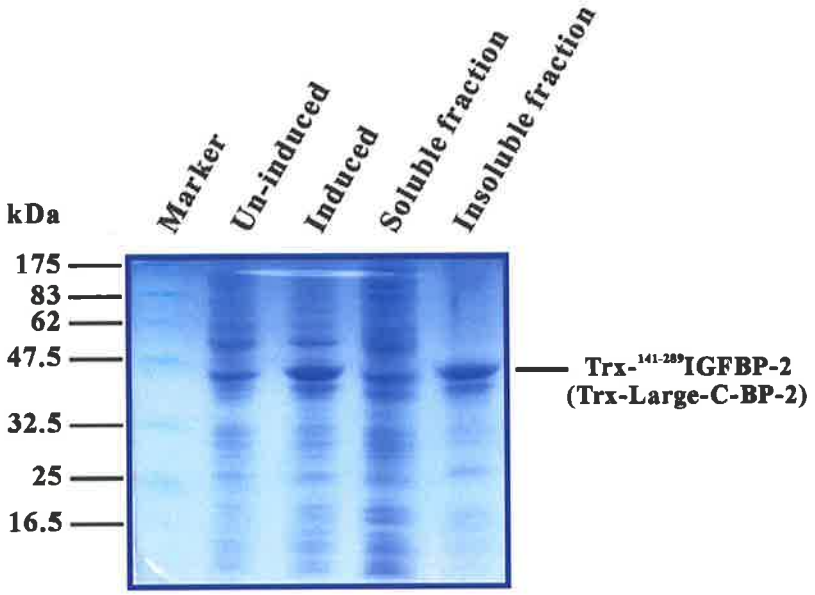


Figure 3.6 Expression of Large-C-BP-2 protein in M9 minimal medium

Uninduced and induced cell culture samples from the *E. coli* expression of Trx-$^{141-289}$IGFBP-2 in M9 minimal medium, and the samples of soluble and insoluble fractions prepared as described in section 3.3.2 were separated by electrophoresis on 12% polyacrylamide tricine gel and stained with Coomassie Blue-R250 (section 3.3.7). Molecular weight markers are indicated on the left.

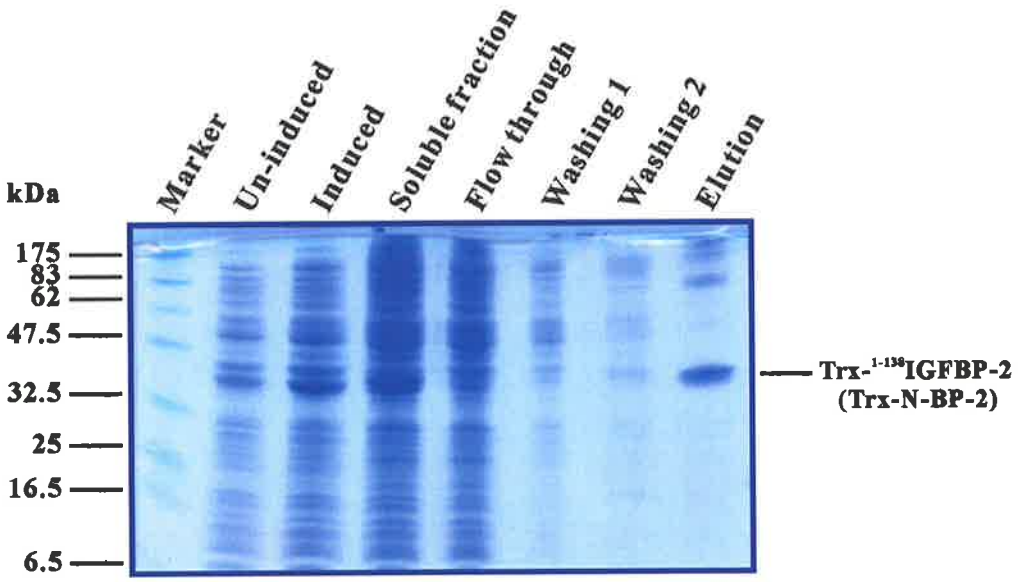


Figure 3.7 Expression and purification of unlabelled N-BP-2 protein

Uninduced and induced cell culture samples from the *E. coli* expression of Trx-$^{1-138}$IGFBP-2, and the protein samples at each state during the purification using nickel affinity chromatography (as described in section 3.3.2) were separated by electrophoresis on 12% polyacrylamide tricine gel and stained with Coomassie Blue-R250 (as described in section 3.3.7). Molecular weight markers are indicated on the left.

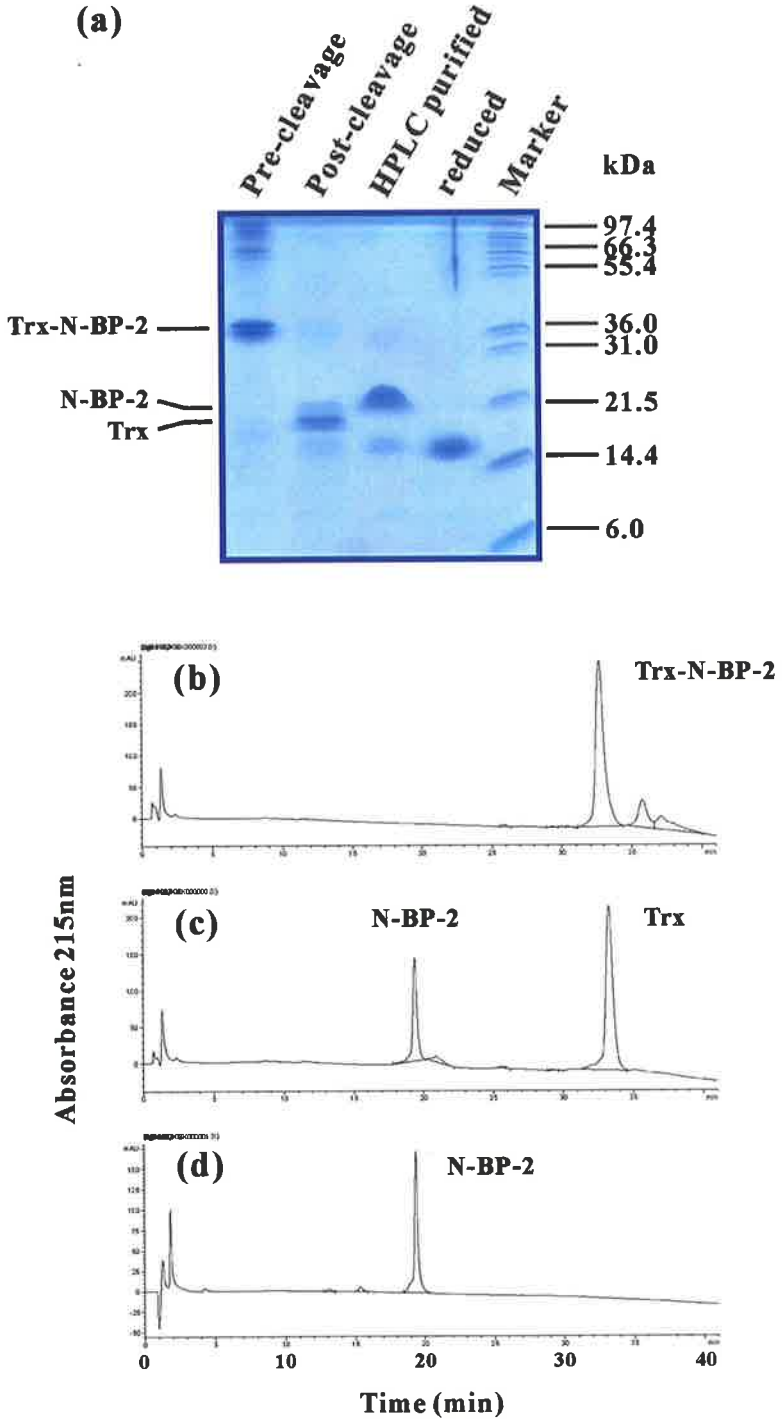


Figure 3.8 3C protease cleavage of thioredoxin-N-BP-2 fusion protein

Protein samples of Trx-$^{1\text{-}138}$IGFBP-2 before and after 3C protease cleavage (section 3.3.5), and the $^{1\text{-}138}$IGFBP-2 (N-BP-2) sample purified by preparative HPLC (section 3.3.6) were analysed by (a) electrophoresis on 12% polyacrylamide tricine gel and stained with Coomassie Blue-R250 (section 3.3.7), or (b-d) by analytical HPLC using a C4 BU-300 column on an Agilent model 1100 HPLC system (section 3.3.8). Absorbance at 215 nm was measured.

Table 3.1 Mass spectroscopy results of purified IGFBP-2 domain fragments

	Theoretical mass (Da) [a]	Determined mass (Da)
C-BP-2		
unlabelled	12213.9	12213.3 ± 0.1
^{15}N-labelled	12372.6 [b]	12373.4±0.3
^{15}N/^{13}C-labelled	12885.1 [c]	12882.5 ±0.1
Large-C-BP-2		
unlabelled	18066.8	18066.2±0.1
N-BP-2		
unlabelled	14282.2	14282.0±0.1
^{15}N-labelled	14458.5 [b]	14457.8±0.9

[a] As all cysteines are disulfide bonded. [b] Assumption of 98% ^{15}N incorporation. [c] Assumption of 98% ^{15}N and 98% ^{13}C incorporation.

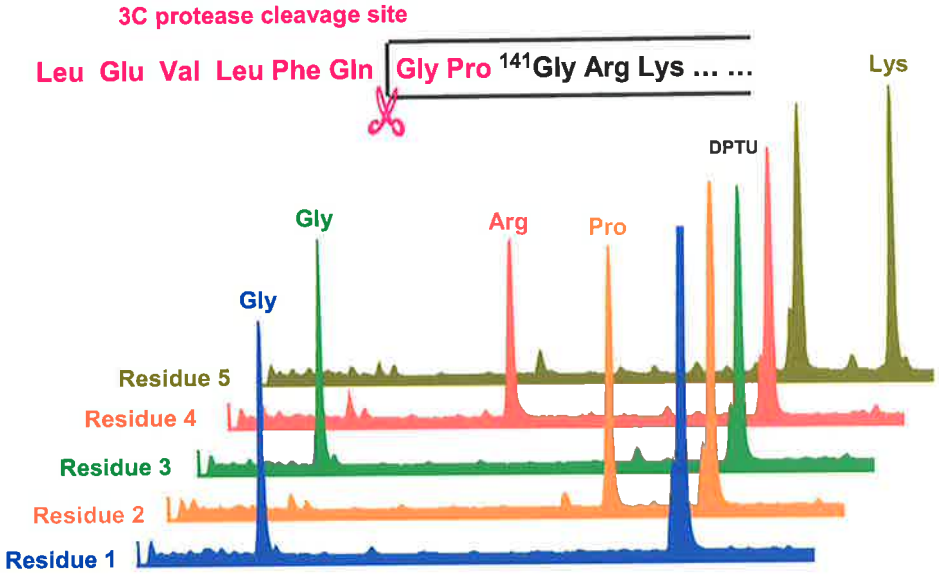


Figure 3.9 N-terminal sequencing result showing the 3C protease cleavage

The N-terminal amino acid sequence (five residues) of $^{141\text{-}289}$IGFBP-2 was determined by Edman degradation using an Applied Biosystems Procise protein sequencer. The predicted sequence is shown on the top.

3.4.4 Cross-linking

In the cross-linking experiments, ^{125}I-labelled IGF-II bound to C-BP-2, N-BP-2 and full-length IGFBP-2 (Figure 3.10). The amount of C-BP-2 required was higher than that of N-BP-2, and much higher than that of IGFBP-2. Noticeably, the IGF-II·C-BP-2 complex band was very weak compared to the IGF-II·IGFBP-2 or IGF-II·N-BP-2 bands, even though 5 μg C-BP-2, which was approximately five-fold more than N-BP-2 and 700-fold more than IGFBP-2, was used in the interaction. The results indicated the IGF-II binding affinity of C-BP-2 was significantly lower (> 10-fold) than that of N-BP-2, and much lower (> 1000-fold) than that of IGFBP-2.

3.4.5 BIAcore analysis

The IGF binding properties of the IGFBP-2 N- and C-domain fragments generated in this study were further assessed using BIAcore.

It was necessary firstly to confirm that the C-BP-2 fragment for structure determination was folded properly and had identical IGF binding ability with the C-domain within the full-length protein. This was particularly important because C-BP-2 showed weak IGF binding in the cross-linking experiments and also a low response in the BIAcore analysis (see below). To this end, as described in previous Chapter, a full-length IGFBP-2 mutant, IGFBP-2(3C@183) was generated, which had 6-residue substitutions in the linker domain preceding Gly183 forming a 3C protease cleavage site. In BIAcore experiments, the IGFBP-2(3C@183) protein showed essentially identical binding sensorgrams compared to wild-type IGFBP-2 (Figure 3.11), while the C-domain fragment cleaved from the IGFBP-2(3C@183) protein had essentially identical BIAcore sensorgrams compared to the C-BP-2 fragment generated from the thioredoxin-C-BP-2 fusion protein (Figure 3.12).

The sensorgrams of IGFBP-2 and the N- and C-domain fragments binding to immobilized IGF-I and IGF-II ligands are shown in Figure 3.13. Kinetic data derived from these experiments by globally fitting the sensorgrams using either 1:1 Langmuir binding model or steady-state affinity determination (for N- and C-domain fragments), and a two-state reaction model (for full-length IGFBP-2) are summarized in Table 3.2.

The binding of IGFBP-2 and domain fragments to IGF-I and IGF-II surfaces was specific, as no binding of C-BP-2 to the insulin surface was observed (Figure 3.14).

Full-length IGFBP-2 exhibited high IGF binding affinities resulting from slow dissociation rates and relatively fast association rates. N-BP-2 showed markedly increased dissociation and more rapid association compared to the full-length protein, resulting in a 1032-fold and 300-fold reduced IGF-I and IGF-II binding affinities, respectively.

The larger C-domain fragment containing the second half of the linker domain, $^{141-289}$hIGFBP-2 (Large-C-BP-2), showed similar sized responses (RU) compared to those for N-BP-2. Surprisingly, the IGF binding affinities of Large-C-BP-2 derived from these data were close to full-length IGFBP-2. The smaller C-domain fragment, $^{183-289}$hIGFBP-2 (C-BP-2), exhibited much smaller responses over the IGF-I and IGF-II surfaces compared to N-BP-2 and Large-C-BP-2, when similar concentrations were used, but much slower off-rates. Therefore, C-BP-2 was also analysed using a set of higher concentration samples, which showed larger responses, and the kinetic data from this set of concentrations were also derived. Interestingly, while the binding affinities derived from low and high concentration runs were very close for the IGF-I surface, they were different for the IGF-II surface. Overall, C-BP-2 has approximately 250-fold lower IGF-I binding and about 40 to 400-fold lower IGF-II binding.

To further analyse the binding between IGFs and C-domain fragments, BIAcore experiments of the reverse orientation were performed, i.e., IGF-I and IGF-II were passed over the surfaces immobilized Large-C-BP-2, C-BP-2 and IGFBP-2 proteins. Surprisingly, while IGFs bind to the immobilized IGFBP-2, there was no binding response in either the large-C-BP-2 or the C-BP-2 surface after subtracting the sensorgram of the reference blank surface (not shown).

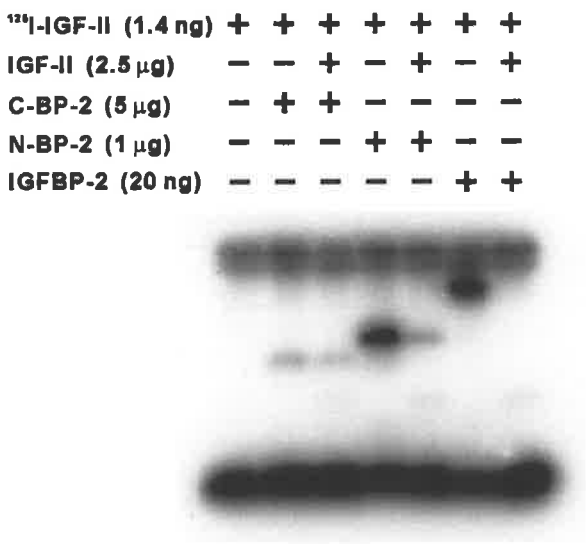


Figure 3.10 Cross-linking of IGFBP-2 N- and C-domains to IGF-II

Binding of C-BP-2, N-BP-2, and IGFBP-2 to IGFs to IGF-II was demonstrated by cross-linking of these proteins to ^{125}I-labelled IGF-II as described in section 3.3.10.

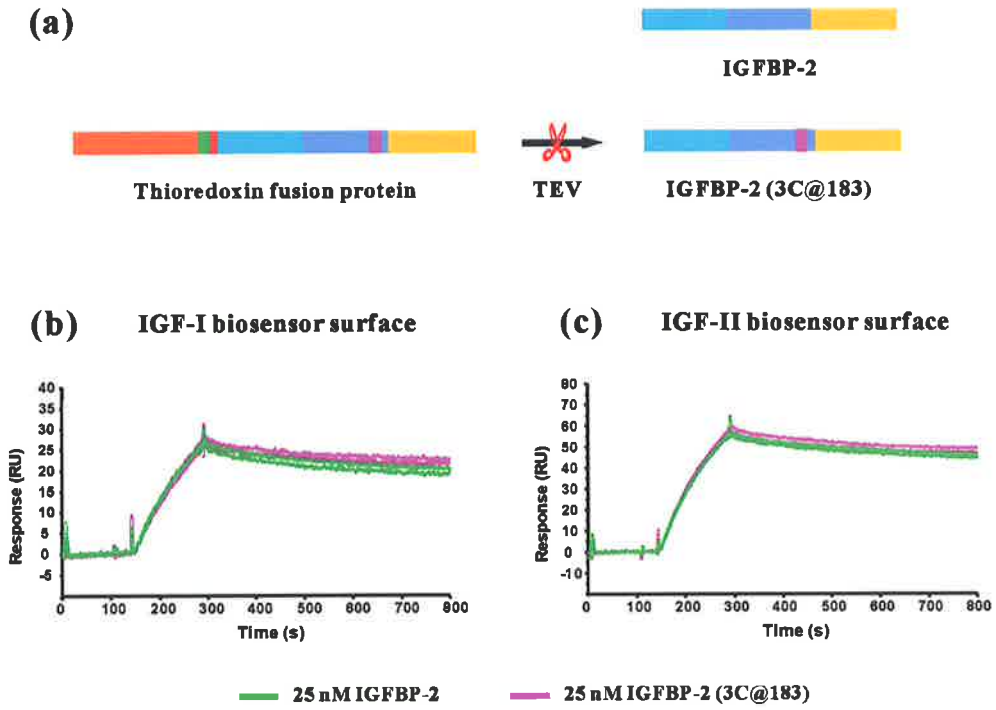


Figure 3.11 IGF binding of IGFBP-2 and IGFBP-2(3C@183)

(a) Schematic representation of the preparation of IGFBP-2(3C@183) protein. This mutant full-length IGFBP-2 protein had a 3C cleavage site preceding Gly183. It was expressed in *E. coli* (as described in section 3.3.1) and purified using IGF affinity chromatography (as described in section 3.3.3). (b and c) Comparison of the IGF binding by wild-type IGFBP-2 and IGFBP-2(3C@183) using BIAcore.

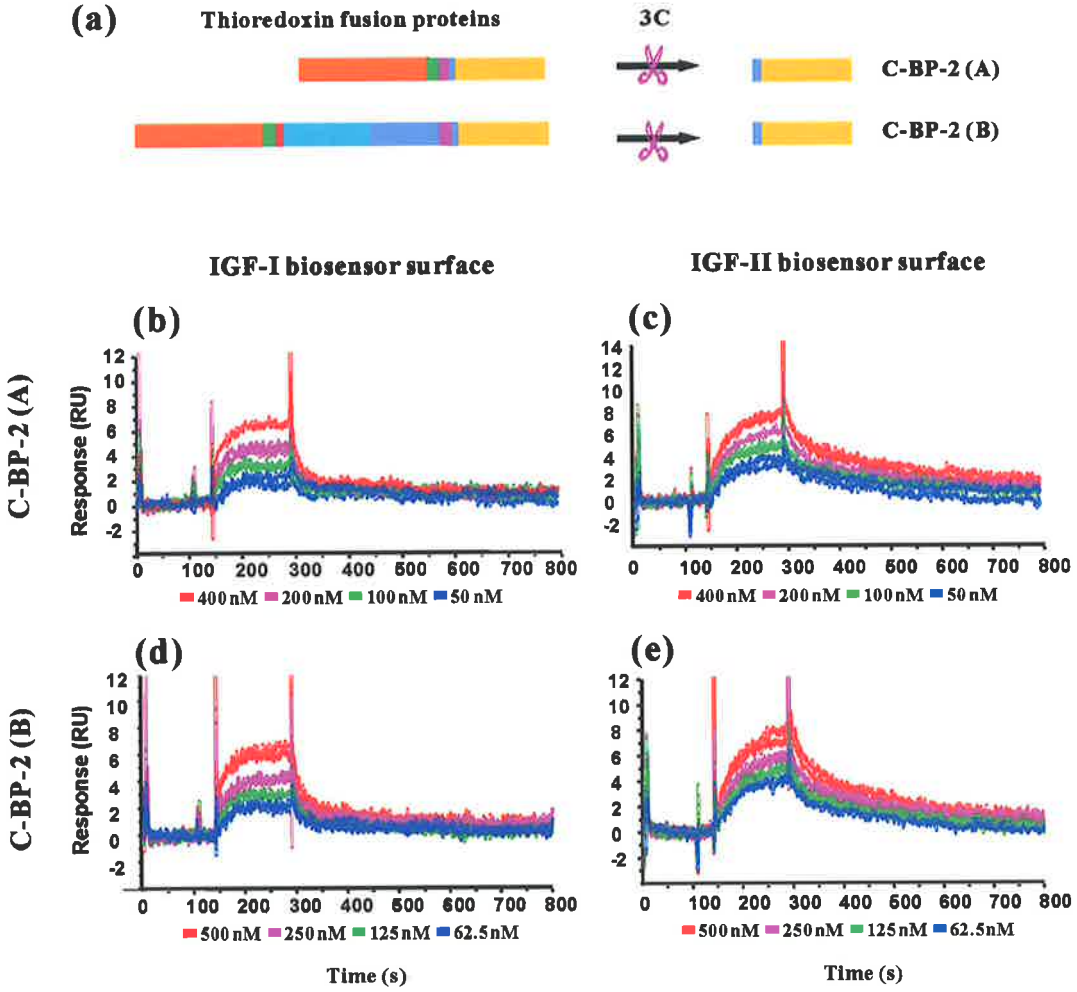


Figure 3.12 IGF binding of C-BP-2 fragments prepared using two approaches

(a) Schematic representation of the preparation of C-BP-2 fragments in two different ways: cleaved by 3C protease from Trx-$^{183\text{-}289}$IGFBP-2 or cleaved by 3C protease from Trx-IGFBP-2(3C@183) protein. (b-e) Comparison of the IGF binding by C-BP-2 fragments prepared in two different ways.

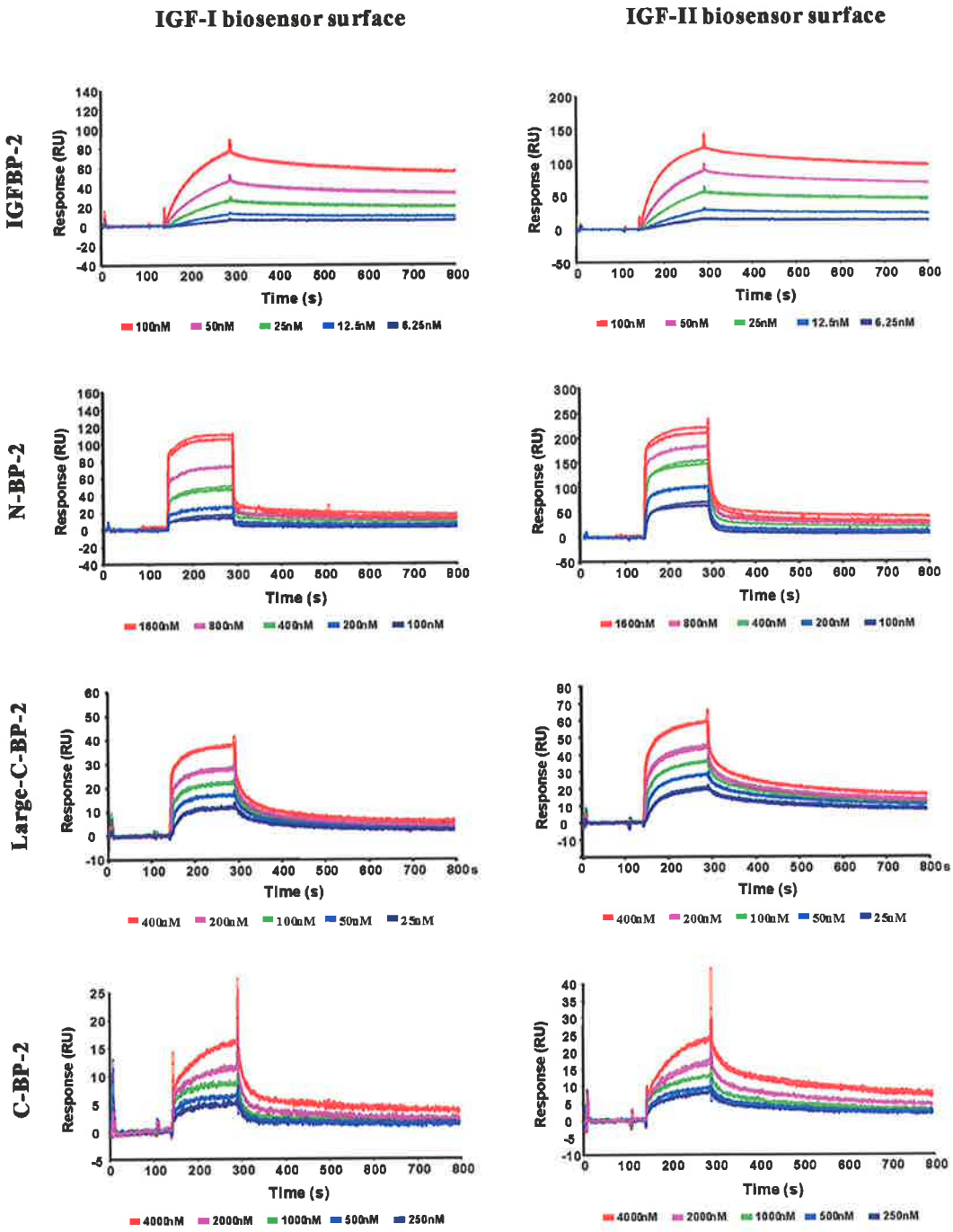


Figure 3.13 BIAcore analysis of IGFBP-2 N- and C-domain fragments over IGF surfaces

Biosensorgrams are presented for the interactions of IGFBP-2, $^{1\text{-}138}$IGFBP-2 (N-BP-2), $^{141\text{-}289}$IGFBP-2 (Large-C-BP-2), and $^{183\text{-}289}$IGFBP-2 (C-BP-2) with IGF-I and IGF-II biosensor surfaces. BIAcore analysis was carried out as described in sections 3.3.11 and 3.3.12. The concentrations of the analytes are indicated.

Table 3.2 IGF binding affinities of IGFBP-2 domain fragments determined using BIAcore.

	Concentrations $\times 10^{-9}$ M	K_D $\times 10^{-9}$ M	Relative K_D	Fitting method	χ^2
IGF-I surface					
IGFBP-2	100, 50, 25, 12.5, 6.25	1.25	1	Two states	2.49
N-BP-2	1600, 800, 400, 200, 100	1290	1032	Steady state	6.37
C-BP-2	400, 200, 100, 50, 25	309	247	Steady state	0.14
C-BP-2	4000, 2000, 1000, 500, 250	317	254	1:1 Langmuir	0.40
Large-C-BP-2	400, 200, 100, 50, 25	1.19	0.95	1:1 Langmuir	1.15
IGF-II surface					
IGFBP-2	100, 50, 25, 12.5, 6.25	0.96	1	Two states	5.26
N-BP-2	1600, 800, 400, 200, 100	288	300	Steady state	34
C-BP-2	400, 200, 100, 50, 25	40.5	42.2	1:1 Langmuir	0.29
C-BP-2	4000, 2000, 1000, 500, 250	379	394	1:1 Langmuir	1.16
Large-C-BP-2	400, 200, 100, 50, 25	2.7	2.82	1:1 Langmuir	0.59

Data were derived from experiments where IGFBP-2 and domains were injected over immobilized IGF-I or IGF-II. All fittings were globably. A two-state reaction (conformational change) model was used for IGFBP-2. The 1:1 Langmuir binding model or steady state affinity determination were used for N- and C-domain fragments dependent on which method resulted in better fits. Values relative to IGFBP-2 are also given.

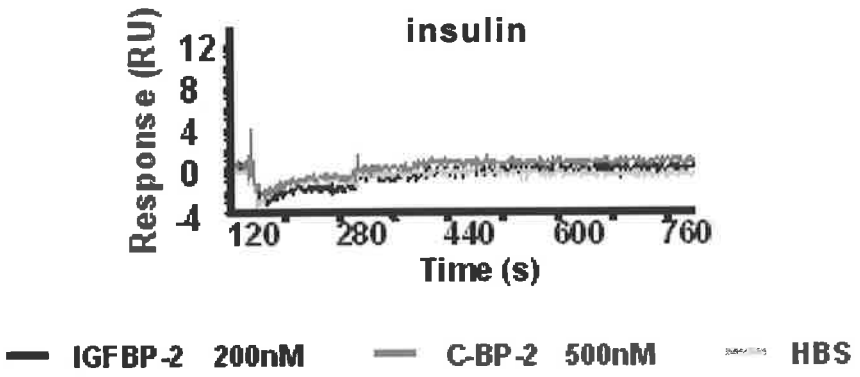


Figure 3.14 Binding of IGFBP-2 and C-BP-2 to insulin surface

In BIAcore experiments, no binding of IGFBP-2, C-BP-2, and N-BP-2 (not shown) to the control insulin sensor surface was detected. HBS, HEPES buffered saline.

3.5 Discussion

3.5.1 Recombinant expression of IGFBP-2 fragments: the solubility issue

In this study, a protocol previously established by Carrick and co-workers (Carrick et al., 2001) in the lab was further modified for the expression and purification of IGFBP-2 N- and C-domain fragments for structural studies.

IGFBP molecules have conserved cysteines in both the N- and C-domains and these cysteines are disulfide-bonded. The N- and C-domain of IGFBP-2 contain six and three intra-domain disulfide bonds, respectively. The disulfide bonds are believed to be important for the integrity of the three-dimensional structures of IGFBPs. A limitation of the production of disulfide bond containing protein in *E. coli* has been the relatively high reducing potential of the cytoplasmic compartment.

In this study, it was found that majority of the thioredoxin fused N- and C-domain fragments expressed in *E. coli* BL21 (DE3) cells were in the soluble fraction of the cell lysates. However, the thioredoxin partner did not seem to be efficient enough to maintain the full-length IGFBP-2 soluble since the thioredoxin-IGFBP-2 fusion proteins were largely found in the insoluble fraction, although a small amount of this fusion protein was purified from the soluble fractions and the IGFBP-2 protein obtained from the fusion protein had identical IGF binding ability compared to the mammalian source IGFBP-2 sample. It was also noted that the smaller C-domain fragment (C-BP-2) was present in higher amounts in the soluble fraction compared to the ~40-residue larger fragments, Large-C-BP-2 and N-BP-2. Although the amounts of the fusion proteins in the insoluble fractions were not quantified, it appeared on the SDS polyacrylamide gels that the larger the size of the fusion protein, the lower amount of total protein was synthesised and the smaller the percentage of the proteins was in the soluble fraction. The disulfide linkages of the full-length proteins in the insoluble fractions were not analysed, and re-folding of the IGFBP-2 from the insoluble factions was not attempted.

More interestingly, while the Large-C-BP-2 construct was expressed in the soluble fraction using LB medium, it was mainly found in the insoluble fraction using M9 minimal medium. This was also true when using the MOPS minimal medium (Neidhardt et al., 1974), which was found to provide a higher yield for some proteins

(Dr Steven Polyak and Dr Anne Chapman-Smith, personal communications). Although the exact explanation is unknown, it is possible that the molecular machinery involved in maintaining this fragment in the soluble fraction is significantly suppressed when growing in minimal media. Isotope-labelling and refolding of the Large-C-BP-2 fragment using minimal medium was not pursued. In fact, this did not hamper the current studies because subsequent NMR experiments using the unlabelled Large-C-BP-2 suggested that the ~40 residue additional linker domain residues do not directly contribute to IGF binding, and are very likely unstructured in both the free form and the complex forms (Chapter 7).

Although sufficient amounts of unlabelled and isotope-labelled IGFBP-2 N- and C-domain fragments was successfully prepared using the current *E. coli* expression system, substantial future work is required to improve the system for the large-scale preparation of full-length IGFBP-2 for structural studies or for biotechnological purposes. This would see either the establishment of a refolding approach to prepare IGFBP-2 from the insoluble fraction, in which case the thioredoxin fusion partner may be unnecessary, or, optimisation to increase the proportion of the fusion protein in the soluble fraction.

Current results also provided hints for the latter direction. Because there is a rough inverse correlation between the size of the fragments and the yields in the soluble fraction, the full-length IGFBP-2 constructs can be further optimised so that it is smallest while fully functional. The S-tag and enterokinase regions between the thioredoxin tag and IGFBP-2 may be removed. The 19-residue C-terminal tail following the last cysteine is not required for IGF binding (Forbes et al., 1998), and is unstructured as revealed by NMR (Chapter 4). A 15-residue alanine rich insertion in the N-domain is also likely to be unstructured and not involved in IGF binding based on homology modelling and NMR studies (Chapter 7). Therefore, these regions could be deleted without affecting IGF binding. Since partial deletions of the linker domain did not affect IGF binding of IGFBP-4 (Qin et al., 1998), ti is worth determining the minimal length of the linker domain in IGFBP-2 that is required to maintain its gross structure and IGF binding ability. The linker domain in the constructs may be shortened accordingly. On the other hand, since the differences in the culture media also affected the portion of the protein in the soluble fraction, both the *E. coli* cell type and the medium for the expression may be further optimised. Alternatively, some other fusion tags that have been reported to significantly enhance the solubility of the recombinant

proteins, such as maltose-binding protein (MBP) and E. coli host factor protein NusA, may be tested for expression of the full-length IGFBP-2 protein.

3.5.2 Purification of IGFBP-2 domains: the use of 3C and TEV proteases

A key change from the previous method used by Carrick and co-workers was the replacement of the enterokinase by the 3C protease or TEV protease in the fusion tag removal step.

The enterokinase cleavage site is provided in the pET32a expression vector for possible removal of the thioredoxin tag. However, Carrick and co-workers found that using this protease caused non-specific proteolysis of IGFBP-2 and its domain fragments, resulting in low yield of the proteins (Carrick, 2002, PhD thesis, The University of Adelaide). This was not too surprising because the ~100-residue non-conserved linker domains of IGFBPs are known to possess multiple protease cleavage sites and are believed to be largely unstructured. While an enterokinase cleavage site (DDDDK) is not present in the IGFBP-2 sequence, the non-specific cleavages of IGFBP-2 may have been caused by other protease contaminants in the commercial preparations (Carrick, 2002, PhD thesis, The University of Adelaide) or by the low fidelity of enterokinase over the cleavage site of other proteases. It was therefore desirable to replace the enterokinase by another highly specific protease for the tag removal step to achieve higher yields of proteins for NMR studies.

3C protease and TEV protease were chosen in current study since they have been recently used in recombinant protein purification very widely and have shown high specificity. It was found in this study that both 3C protease and TEV protease were able to cleave at the incorporated cleavage site between the thioredoxin fusion tag and the target protein. The cleavages were indeed of high specificity, as revealed by the analytical HPLC and SDS polyacrylamide gel electrophoresis. N-terminal sequencing of the purified proteins also showed the expected N-termini resulted from the cleavages.

Nevertheless, a conflict between the requirement of reducing reagent DTT by both 3C and TEV proteases for high activities and the susceptibility of IGFBP-2 and its domain fragments to DTT was found. Even 0.1 mM DTT changed the retention time of the IGFBP-2 product peak on HPLC. It is believed that this was due to the reduction of some relatively more susceptible disulfide bonds in IGFBP-2 molecules and one or

more of these disulfide bonds is/are located in the C-domain since a similar effect was found with the C-domain fragments. Higher molecular weight aggregates were also found in the samples prepared using DTT in the cleavage step, and these samples showed abnormal and non-reproducible sensorgrams in BIAcore experiments. The reducing conditions required by 3C and TEV, both cysteine proteases, may compromise their future use in large-scale production of recombinant IGFBP-2 proteins in a biotechnological context.

3.5.3 IGF-binding by C-domain fragments

Before investigating the interactions between IGFs and IGFBP domain fragments using NMR, the IGF binding activities of the recombinant N- and C-domain fragment were assessed using two methods, cross-linking and BIAcore.

Although it is well accepted that the C-domain is an indispensable part of IGFBPs for high affinity IGF binding, the affinities of IGFs for both naturally occurring and recombinant C-domain fragments reported in the literature vary greatly. From as small as 10-fold to as large as 2000-fold loss of affinity compared to their full-length protein has been reported by different authors for different C-domain fragments (Ho and Baxter, 1997; Carrick et al., 2001; Vorwerk et al., 2002; Payet et al., 2003; Headey et al., 2004c; Mark et al., 2005). In some reports the IGF binding affinity of C-domains was not detectable (Kalus et al., 1998; Qin et al., 1998; Fernandez-Tornero et al., 2005; Siwanowicz et al., 2005). These discrepancies may reflect a combination of genuine binding differences together with the technical differences, such as the different specificity and sensitivity of the methods used.

In the present study, the cross-linking experiments showed that C-BP-2 bound to ^{125}I-labelled IGF-II very weakly. Because the band of IGF-II·C-BP-2 complex was much weaker than those of IGF-II·IGFBP-2 or IGF-II·N-BP-2 complexes, the differences in IGF binding between C-BP-2 and N-BP-2 or IGFBP-2 were expected to be larger than the differences in the amount of proteins used in the interaction. Thus, the IGF-II binding affinity of C-BP-2 was significantly lower than that of N-BP-2, and much lower than that of IGFBP-2.

The BIAcore results of the C-domain fragments were complex. There was a noticeable difference in terms of the size of responses between C-BP-2 and Large-C-

BP-2 over both IGF-I and IGF-II surfaces, when these fragments were run at the same concentrations (Figure 3.13). Kinetic data were therefore also derived from a set of higher C-BP-2 concentration runs, which resulted in relatively larger size of responses. The kinetic data for C-BP-2 (both low and high concentrations) dramatically differed from those of the Large-C-BP-2. In contrast to the cross-linking results, the BIAcore data showed that the C-BP-2 has comparable or higher affinities for IGFs compared to N-BP-2. There were also discrepancies between the current results and previous data for IGFBP-2 domain fragments (Carrick et al., 2001). Furthermore, in the reverse orientation experiment (not shown), no specific binding of IGF-I or IGF-II to the immobilized C-BP-2 was detected. Surprisingly, this was also true for the immobilized Large-C-BP-2, although this fragment showed appreciable binding to the immobilized IGFs.

The BIAcore data must be, therefore, interpreted with care and in conjunctions with the results of the cross-linking and the NMR interaction studies (Chapter 7). Importantly, because the C-BP-2 fragment expressed and cleaved from the thioredoxin-C-BP-2 fusion protein had identical IGF binding properties compared to the corresponding fragment cleaved from the full-length IGFBP-2(3C@183) protein that in turn had identical IGF binding to wild-type protein, its low responses in binding to IGFs were not due to mis-folding of the fragment. This was further confirmed by subsequent NMR experiments.

The larger responses of Large-C-BP-2 compared to C-BP-2 over the same IGF surfaces seem not to be explained solely by the mass difference between C-BP-2 (12.2 kDa) and Large-C-BP-2 (18.0 kDa). Because the presence of the linker domain (Large-C-BP-2) did not cause significant additional chemical shift perturbations of N-BP-2 or IGFs upon binary and ternary complex formation (Chapter 7), it is unlikely that the linker domain part in Large-C-BP-2 interacts with IGFs and increases its binding affinities. Therefore, the discrepancy between the BIAcore and cross-linking or NMR (Chapter 7) results in comparing the N- and C-domains, as well as the differences in the kinetic data for Large-C-BP-2 and C-BP-2, may have arisen from other factors contributing to the sensorgrams that are not yet revealed. One possibility is a preferential attachment to the CM5 chip via the extended N-terminal linker region leading to greater mobility of Large-C-BP-2.

Moreover, the reason why IGFs did not bind to either the C-BP-2 or Large-C-BP-2 surface is unclear. Interestingly, Galanis and co-workers reported a very similar

phenomenon previously (Galanis et al., 2001). In that study, although an IGFBP-3 C-domain fragment, $^{165\text{-}264}$IGFBP-3, exhibited binding to the IGF surfaces, it did not bind IGFs when it was immobilized on the chip (Galanis et al., 2001). The reason was unclear, but these authors suggested that the immobilised C-domain fragment may have an orientation such that is was no longer functional (Galanis et al., 2001). In the present study, since subsequent NMR results suggested that interaction between C-BP-2 and IGFs may require conformational changes of part of the C-BP-2 molecule (Chapters 4 and 7), and such conformational changes might be hampered when C-BP-2 is immobilized. It will be of interest to prepare a site-specifically immobilized C-BP-2 surface, such as by biotinylating the N- or C-terminus of the proteins and immobilizing it on a streptavidin chip. A similar strategy has been applied to preparation of IGF chips to improve the accuracy of the kinetic data (Dr Briony Forbes, The University of Adelaide).

Taken together, the cross-linking and BIAcore results indicated that C-BP-2 fragment binds IGFs to a limited extent, although the C-domain is clearly very important for high affinity IGF binding of full-length IGFBP-2, as N-BP-2 has reduced binding affinity. In particular, IGF·N-BP-2 complexes have much faster dissociation rates compared to IGF·IGFBP-2 complexes. Thus, it appears that the C-domain contributes significantly to stabilizing the IGF·IGFBP-2 complex, but the mechanism for such an interaction may not be (fully) revealed by analysing the binding of C-BP-2 alone to IGFs. In general, these results are in consistent with the majority evidence in the literature, which has shown weak, or even undetectable, binding affinities of C-domain fragments for IGFs using different techniques and reduced binding affinity of N-domain fragments for IGFs. The structural mechanisms underlying the role of C-domain in IGF binding are further analysed in detail in Chapter 7.

Chapter 4

Solution Structure and Backbone Dynamics of the IGFBP-2 C-terminal Domain

4.1 Introduction

As reviewed in Chapter 1, the N- and C-domains of IGFBPs are both required for high-affinity IGF binding, because N- and C-domain fragments bind IGFs with lower affinities than full-length IGFBPs (Hwa et al., 1999; Baxter, 2000; Firth and Baxter, 2002; Bach et al., 2005). Deletion of the C-domain residues 222-284 of bovine IGFBP-2 dramatically reduced its IGF binding affinity (Forbes et al., 1998). Site-directed mutagenesis (Bramani et al., 1999; Song et al., 2000; Shand et al., 2003; Yan et al., 2004; Allan et al., 2006), antibody binding (Schuller et al., 1993), photoaffinity labelling (Horney et al., 2001) and subsequent studies (Kibbey et al., 2006) have identified residues or regions in the C-domain thought to be involved in IGF binding. Recently, Headey and co-workers identified the IGF-II binding site on the C-terminal domain of IGFBP-6 (C-BP-6) using NMR spectroscopy and mutagenesis (Headey et al., 2004a).

In addition to both inhibiting and enhancing IGF actions, IGFBPs have many IGF-independent actions, and both IGF-dependent and IGF-independent activities vary according to the binding protein and system under examination (Hoeflich et al., 2001b; Firth and Baxter, 2002). Importantly, the C-domain of IGFBPs is a major contributor to these diverse functions. The C-domain of IGFBP-6 is the principal determinant of the 20 to 100-fold IGF-II binding preference of the full-length protein, which thus is a relatively specific inhibitor of IGF-II actions (Bach, 2005). The C-domains of IGFBP-1 and -2 contain RGD motifs that bind integrins, leading to altered cell migration (Jones et al., 1993b; Schutt et al., 2004; Wang et al., 2006), whereas those of IGFBP-3 or -5 are not only implicated in IGFBP-IGF-acid-labile subunit ternary complex formation, but also cellular internalisation via transferrin, and nuclear localization via importin-β (Firth and Baxter, 2002). IGFBP-2 and a C-domain fragment have also been found in both the cytoplasm and nucleus under certain circumstances (Hoeflich et al., 2004;

Terrien et al., 2005). Binding of the C-domain of some IGFBPs to a rapidly expanding list of serum and extracellular matrix molecules, such as heparin and other glycosaminoglycans, plasminogen, fibrinogen-fibrin, vitronectin, fibronectin, and metal ions, may alter the IGF-IGFBP binding affinity or the spatial distribution of these molecules, but their structural mechanisms and the biological relevance are incompletely understood (Bach et al., 2005).

A comparison of the three-dimensional structures of IGFBP C-domains, together with other structurally similar proteins, is an important step towards understanding their roles in both IGF binding and IGF-independent functions. The three-dimensional structure of a full-length IGFBP has not yet been determined, but structures have been reported for a truncated N-domain of IGFBP-5 (mini-N-BP-5) (Kalus et al., 1998; Zeslawski et al., 2001) and the complete N-domain of IGFBP-4 (Siwanowicz et al., 2005). During the course of the work described in this thesis, Headey and co-workers solved the solution structure of C-BP-6, and mapped the IGF-II binding sites on C-BP-6 and the C-BP-6 binding site on IGF-II (Headey et al., 2004a; Headey et al., 2004b), while Sala and co-workers reported the crystal structure of C-BP-1 (Sala et al., 2005). As current study was nearing completion, Sitar and co-workers reported the crystal structure of the IGF-I·N-BP-4·C-BP-4 ternary complex (Sitar et al., 2006). Crystallization of C-BP-4 alone was also reported (Fernandez-Tornero et al., 2005). Backbone amide dynamics showed that C-BP-6 residues undergoing sub-millisecond to millisecond timescale conformational or chemical exchange correlate with the IGF-II binding surface (Yao et al., 2004).

This Chapter described the determination of the solution structure of the IGFBP-2 C-domain using NMR spectroscopy. Compared to the C-BP-6 (Headey et al., 2004a), C-BP-1 (Sala et al., 2005) and C-BP-4 (Sitar et al., 2006) structures, structural differences that may affect IGF binding have and implications for other functional diversities were identified. The backbone dynamics of C-BP-2 were analysed using the reduced spectral density mapping approach based on the measured backbone amide ^{15}N relaxation parameters (R_1, R_2 and steady-state ^{15}N-{^1H} NOE), and the results were compared to the dynamics properties of C-BP-6 (Yao et al., 2004).

4.2 Materials and Methods

4.2.1 Materials

Unlabelled, ^{15}N-labelled and $^{15}N/^{13}C$-labelled $^{183-289}$IGFBP-2 (C-BP-2) fragments were prepared as described in Chapter 3. $^{2}H_2O$ was purchased form Cambridge Isotope laboratories Inc. (Andover, MA, USA). Shigemi NMR tubes were purchased from Shigemi Inc. (Allison Park, PA, USA).

4.2.2 NMR spectroscopy experiments

NMR experiments were performed on C-BP-2 samples in 95% H_2O/5% $2H_2O$ containing 10 mM sodium acetate and 0.02% (w/v) sodium azide, using either a Bruker Avance 500 spectrometer equipped with a cryoprobe or a Bruker DRX-600 spectrometer with a triple-resonance probe equipped and triple-axis gradients. ^{1}H-^{15}N HSQC spectra of a 0.06 mM ^{15}N-C-BP-2 sample were acquired at 298 K and pH 4.0, 5.0, 6.0 and 7.0. The following NMR experiments were performed on ^{15}N-C-BP-2 (0.8 mM) and $^{15}N/^{13}C$-C-BP-2 (0.5 mM) samples at 298 K and pH 6.0, or 293 K and pH 4.2: ^{1}H-^{15}N HSQC, ^{1}H-^{13}C HSQC, HNCA, HN(CO)CA, HNCO, CBCA(CO)NH, HNCACB, HNHA, HBHA(CO)NH, HCCH-TOCSY, CNH-NOESY-HSQC, ^{15}N-edited NOESY-HSQC (τ_m = 50 and 150 ms) and 13C-edited NOESY-HSQC (τ_m = 120 ms, ^{13}C carrier frequencies of 39 ppm and 125 ppm). Two-dimensional homonuclear TOCSY and NOESY (τ_m = 150 ms) spectra were acquired on unlabelled C-BP-2 (0.8 mM) in $^{2}H_2O$. Spectra were referenced to an internal impurity peak at 0.15 ppm in the ^{1}H dimension and indirectly using the experimentally determined frequency ratios for ^{15}N and ^{13}C (Wishart et al., 1995b). Spectra were processed using XWINNMR version 3.5 (Bruker Biospin) and analysed using XEASY, version 1.3 (Bartels et al., 1995). Ms Christina Wang (Walter and Eliza Hall Institute of Medical Research) acquired 2D and 3D spectra of C-BP-2 at pH 4.2 and 283 K, and also contributed to initial resonance assignments. Details of protein samples used and NMR spectra recorded in this study are summarized in Tables 4.1 and 4.2.

Table 4.1 C-BP-2 NMR samples for structure and dynamics study

ID	Protein	Concentration (mM)	Buffer [a]	pH	Purpose [b]
[1]	Unlabelled C-BP-2	1.0	10 mM NaAc, 95% H_2O / 5% $2H_2O$		Initial 1D spectra Variable temperature and pH
[2]	^{15}N-labelled C-BP-2	0.06	10 mM NaAc, 95% H_2O / 5% $2H_2O$	4.0, 5.0, 6.0, 7.0	pH effects
[3]	^{15}N-labelled C-BP-2	0.8	10 mM NaAc, 95% H_2O / 5% $2H_2O$	4.2, 6.0	Assignments, constraints
[4]	^{15}N/^{13}C-labelled C-BP-2	0.5	10 mM NaAc, 95% H_2O / 5% $2H_2O$	4.2, 6.0	Assignments, constraints
[5]	Unlabelled C-BP-2	0.8	10 mM NaAc, 100% $2H_2O$	4.2	Assignments
[6]	^{15}N-labelled C-BP-2	0.3	10 mM NaAc, 95% H_2O / 5% $2H_2O$	6.0	^{15}N relaxation, diffusion
[7]	^{15}N-labelled C-BP-2	0.03	10 mM NaAc, 95% H_2O / 5% $2H_2O$	6.0	^{15}N relaxation,[c] diffusion

[a] All buffers also contained 0.02% (w/v) NaN_3. [b] NMR experiments performed on these samples are listed in Table 4.2. [c] First-increment FIDs.

Table 4.2 C-BP-2 NMR experiments for structure and dynamics study

Experiment	Sample a	Field	Major parameters						Time
		(MHz)	TD	SW (ppm)	Offset (ppm)	NS	D1	τ_m	
			F3/F2/F1	F3/F2/F1	1H/15N/13C		(s)	(ms)	
1D spectra	[1]	600 b	8192	12.0	4.7	256	1.8		0.5 h
1H-15N HSQC	[2]	500 c	2048/128	12.0/30.0	4.7/117.5	128	1.2		6.5 h
1H-15N HSQC	[3]	500	2048/128	12.0/30.0	4.7/117.5	8	1.2		0.5 h
1H-13C HSQC	[4]	600	2048/128	12.0/180	4.7/80.0	8	1.5		0.5 h
1H-13C HSQC	[4]	600	2048/128	12.0/65.0	4.7/39.0	8	1.5		0.5 h
HNCA	[4]	500	2048/40/92	10.5/25.0/30.0	4.7/117.5/54.0	16	1.0		1 d
HN(CO)CA	[4]	500	2048/40/92	10.5/25.0/30.0	4.7/117.5/54.0	16	1.0		1 d
HNCO	[4]	500	2048/44/128	10.5/25.0/25.0	4.7/117.5/176.0	8	1.0		16 h
CBCA(CO)NH	[4]	500	2048/44/128	10.5/25.0/75.0	4.7/117.5/41.0	32	1.0		3 d
HNCACB	[4]	500	2048/44/128	10.5/25.0/75.0	4.7/117.5/41.0	32	1.0		3 d

HNHA	[3]	500	2048/44/128	10.5/10.5/30.0	4.7/117.5	32	1.0		3 d
HBHA(CO)NH	[3]	500	2048/44/96	10.5/25.0/10.5	4.7/117.5	48	1.0		3 d
HCCH-TOCSY	[4]	600	2048/64/128	10.5/75.0/10.5	4.7/117.5/39.0	24	1.0	12.5	3 d
^{15}N-edited NOESY-HSQC	[3]	500	2048/44/148	11.5/25.0/11.5	4.7/117.5	16	1.5	150	3 d
^{15}N-edited NOESY-HSQC	[3]	500	2048/44/148	11.5/25.0/11.5	4.7/117.5	16	1.5	50	3 d
^{13}C-edited NOESY-HSQC	[4]	600	2048/72/144	11.5/62.0/11.5	4.7/39.0	24	1.5	120	5 d
13C-edited NOESY-HSQC	[4]	500	2048/64/128	11.5/30.0/11.5	4.7/125.0	32	1.0	120	3 d
CNH-NOESY-HSQC	[4]	500	2048/48/154	10.5/25.0/10.5	4.7/117.5/41.0	24	1.0	120	3 d
2D homonuclear TOCSY	[5]	600	4096/512	10.5/10.5	4.7	64	1.8	65	1 d

2D homonuclear NOESY	[5]	600	4096/512	10.5/10.5	4.7	64	1.8	150	1 d
Translational diffusion	[6]	600	8192/12	12.0	4.7	512	2.0		4 h
15N R_1 measurement	[6]	600	2048/128/12	11.5/30.0	4.7/117.5	64	2.8		3 d
15N R_2 measurement	[6]	600	2048/128/12	11.5/30.0	4.7/117.5	64	2.8		3 d
15N NOE measurement	[6]	600	2048/256	11.5/30.0	4.7/117.5	256	3.8		2.5 d
CLEANEX	[6]	600	2048/128	10.5/25.0	4.7/117.5	256	1.5	26	1 d
Translational diffusion	[7]	600	8192/12	12.0	4.7	1024	2.0		1 d
15N R_1 first FIDs	[7]	600	2048/1/12	11.5/30.0	4.7/117.5	2048	2.8		1 d

15N R_2 first FIDs	[7]	600	2048/1/12	11.5/30.0	4.7/117.5	2048	2.8	1 d

[a] Details of the samples are listed in Table S1. NMR experiments were performed at 298 K for pH 6.0 samples, and at 293 K for pH 4.2 samples. [b] Bruker DRX-600 spectrometer with a triple-resonance probe equipped and triple-axis gradients. c Bruker Avance 500 spectrometer equipped with a cryoprobe.

4.2.3 Assignments

A ^{1}H-^{15}N HSQC spectrum was used to pick the ^{1}H-^{15}N correlated peaks. The peak list containing the frequencies was loaded into HNCA, HN(CO)CA, CBCA(CO)NH and HNCACB spectra and used to identify the frequencies of correlated $C^{\alpha(i)}$, $C^{\alpha(i-1)}$, C^{β} and $C^{\beta(i-1)}$ peaks. Sequential connections between residues were established using these frequencies in the HNCA and HNCACB spectra. Residue-specific assignments were made by comparing the observed C^{α} and C^{β} chemical shift values for sequentially connected residues with the expected average values for different amino acid residues. The assigned ^{1}H-^{15}N correlated frequencies were used to assign the H^{α} and H^{β} chemical shifts using HNHA and HBHA(CO)NH spectra, which also confirm the sequential assignments. Other side-chain ^{1}H, ^{15}N and ^{13}C frequencies were assigned using ^{15}N NOESY-HSQC, HCCH-TOCSY, ^{13}C NOESY-HSQC, and CNH-NOESY-HSQC spectra. A ^{13}C NOESY-HSQC with the carrier frequency centred on the aromatic region, as well as two-dimensional homonuclear TOCSY and NOESY spectra acquired on unlabelled C-BP-2 in $2H_2O$, were used to identify additional aromatic resonances.

4.2.4 Distance constraints

NOE cross peaks in the ^{15}N NOESY-HSQC, ^{13}C NOESY-HSQC, and 2D NOESY (unlabelled sample in $2H_2O$) spectra were manually assigned, and peak intensities were converted into upper bound constraints using CYANA version 1.0.3 (Guntert et al., 1997).

4.2.5 Dihedral constraints

ϕ angle constraints were based on either $3J_{HNHA}$ coupling constants derived from intensity ratios in the HNHA spectrum (Kuboniwa et al., 1994) or an analysis of chemical shifts using TALOS (Cornilescu et al., 1999). For the remaining residues, ϕ angles were restricted to $-90\pm90^{\circ}$ where positive values could be excluded based on NOE data (Ludvigsen and Poulsen, 1992). ψ angles were determined with TALOS

(Cornilescu et al., 1999) and only applied to well-defined secondary structure. $\chi 1$ angles were determined based on NOESY spectra (τ_m = 50 ms).

4.2.6 Hydrogen bond constraints

Hydrogen bond constraints were assigned by analysis of the solvent exchange of amide protons. ^{1}H-^{15}N HSQC spectra were acquired at 10 min, 30 min and 4 h after dissolving the ^{15}N-labelled protein in $^{2}H_2O$ at 288 K. Rapid exchange with the solvent was also analysed using the CLEANEX experiment, which detects H^N magnetisation transferred by chemical exchange from magnetically excited water protons (Hwang et al., 1998). Hydrogen bonds were only included in later rounds of structure calculations in the helix or strand regions that had initially converged without hydrogen bond constraints.

4.2.7 Structure calculations and refinement

Structure calculations were started from 200 randomised conformers in the program CYANA version 1.0.3 (Guntert et al., 1997), using a torsion angle dynamics algorithm. The 20 conformers with the lowest final CYANA target function values were subjected to restrained energy minimization in a water shell of 8 Å thickness with the program OPALp (Koradi et al., 2000), using the AMBER force field (Cornell et al., 1995). Dr Peter Güntert (RIKEN Genomic Sciences Centre, Japan) performed the energy minimization. PROCHECK-NMR (Laskowski et al., 1996) and MOLMOL (Koradi et al., 1996) were used to analyse and visualize the structures. Structure figures were prepared using MOLMOL and PyMOL (DeLano Scientific).

4.2.8 Backbone ^{15}N relaxation parameters

Laboratory-frame ^{15}N relaxation rates, R_1, R_2, and steady-state ^{15}N-{1H} NOE were measured at pH 6.0 and 298 K using a 0.3 mM sample on a Bruker DRX-600 spectrometer. Pulse sequences similar to those published previously were used (Farrow et al., 1994), with a data matrix size of 2048×128 and 64 (for R_1 and R_2) or 128 (for 15N-{1H} NOE) scans per t_1 increment. A total of 14 relaxation durations, including 2

duplicated durations, was used for R_1 (ranging from 10 to 2400 ms) and R_2 (from 15.5 to 526.6 ms), respectively. The recycle time for R_1 and R_2 measurements was 2.8 s and that for steady-state ^{15}N-{1H} NOE measurement was 3.8 s. All data were recorded in an interleaved manner.

^{15}N relaxation rates of C-BP-2 were obtained by fitting peak intensities of the 15N-1H HSQC spectra at a series of relaxation durations to two-parameter (R_2) and three-parameter single-exponential decays (R_1) using the *Curvefit* program (A.G. Palmer, Columbia University). Errors for the relaxation rates were determined by *Curvefit* using a Monte Carlo simulation. Uncertainties of peak intensities were derived from duplicated spectra. Steady-state ^{15}N-{^{1}H} NOE values were calculated from peak intensity ratios obtained from spectra acquired in the presence and absence of proton saturation. Uncertainties for the peaks were estimated from the background noise of the spectra (Farrow et al., 1994).

To compare the ^{15}N relaxation rates of C-BP-2 at different concentrations, average ^{15}N R_1 and R_2 values of C-BP-2 at 0.03 mM were obtained from the first-increment FIDs of ^{15}N T_1 and T_2 (R_1=1/ T_1, R_2=1/ T_2) weighted ^{15}N-HSQC spectra, by fitting the integrals of all peaks across the backbone amide region to a three- and two-parameter exponential decay, respectively. Corresponding values for C-BP-2 at 0.3 mM were also derived similarly by reprocessing the relaxation data using the first-increment FIDs.

4.2.9 Translational diffusion

Translational diffusion coefficients of C-BP-2 were measured, as previously described (Yao et al., 2000), on C-BP-2 samples at 0.3 mM (as used for ^{15}N relaxation measurements) and 0.03 mM, in both cases under the same conditions as for relaxation measurements. A series of 12 diffusion-weighted spectra was recorded in a two-dimensional manner using gradient pulses of 6 ms duration with a separation of 46.8 ms and gradient strengths ranging from 3.4 to 36.5 G cm^{-1}.

Diffusion coefficients, D, were obtained by fitting peak intensities to the following equation using the program *Simfit* (Bruker Biospin):

$$I = I_0 \exp(-\gamma 2 g 2 \delta 2(\Delta - \delta/3)D) \qquad (1)$$

where γ is the gyromagnetic ratio and g, δ and Δ are the amplitude, duration and separation of the single pair of gradient pulses, respectively (Stejskal and Tanner, 1965).

4.2.10 Rotational motions

Rotational motions of C-BP-2 were analysed in conjunction with Dr Shenggen Yao (Walter and Eliza Hall Institute of Medical Research) as described previously (Yao et al., 2004). Initial estimation of the global reorientation time of C-BP-2 was obtained by analysis of R_2/R_1 ratios from residues satisfying the criteria of both steady-state ^{15}N-{^{1}H} NOE $\geq$ 0.6 and $|T_1/T_2 - \langle T_1/T_2 \rangle| \leq$ SD (Kay et al., 1989), which assumes the global reorientation to be isotropic. The ratio of three principal moments of the inertial tensor of C-BP-2 was obtained from its closest-to-average structure using pdbinertial (A.G. Palmer, Columbia University). *Modelfree* analysis of ^{15}N relaxation data was attempted by Dr Shenggen Yao using program *Modelfree* (version 4.0, A.G. Palmer, Columbia University), as described previously (Yao et al., 2004).

4.2.11 Reduced spectral density mapping

Dr Shenggen Yao performed the reduced spectral density mapping analysis according to equations (2-5):

$$\sigma_{NH} = R_1(NOE-1)\gamma_N/\gamma_H \qquad (2)$$

$$J(0) = (6R_2-3R_1-2.72\ \sigma_{NH})/(3d2+4c2) \qquad (3)$$

$$J(\omega_N) = (4R_1-5\sigma_{NH})/(3d2+4c2) \qquad (4)$$

$$J(0.87\omega_H) = 4\sigma_{NH}/(5d2) \qquad (5)$$

where $d=(\mu_0 h\gamma_N\gamma_H/(8\pi 2))/\langle r3_{NH}\rangle$, $c=\omega_N(\sigma_\parallel-\sigma_\perp)/31/2$, μ_0 is the permeability of free space, h is Plank's constant, γ_H and γ_N are the gyromagnetic ratio of ^{1}H and ^{15}N respectively, r_{NH} = 1.02 is the average amide bond length, $(\sigma_\parallel-\sigma_\perp)$ = -170 ppm is the chemical shift anisotropy for ^{15}N nuclei (Farrow et al., 1995).

4.3 Results

4.3.1 NMR assignments and constraints

Figure 4.1 shows the cross-peak assignments of the ^{1}H-^{15}N HSQC spectrum of a 0.05 mM C-BP-2 sample at 298 K and pH 6.0, recorded on an Avance 500 spectrometer equipped with a cryoprobe. Overlays of ^{1}H-^{15}N HSQC spectra recorded at pH 4.0, 5.0, 6.0 and 7.0 are shown in Figures 4.2, and ^{1}H and ^{15}N chemical shift differences between pH 6.0 and 4.0 are plotted in Figure 4.3. Some backbone amide resonances are better resolved at pH 6.0 and pH 7.0 than pH 4.0. The most affected residues are located in the distal part of loop III and the beginning of the C-terminal tail (see below). This region is relatively flexible (see below) and the chemical shift differences are believed to arise from the different protonation states of the His272 imidazolium ring. For resonances from other regions of the structure, the differences between pH 4.0 and either 6.0 or 7.0 are relatively small and the chemical shift differences of all residues between pH 6.0 and 7.0 are small, although some resonances are weaker at pH 6.0, and much weaker at pH 7.0, compared with pH 4.0, presumably because of more rapid exchange with H_2O. Therefore, structures were calculated based on the NMR constraints derived at pH 6.0 to represent the structure under physiological conditions, although spectra at pH 4.2 were also used to confirm backbone and side-chain assignments.

The final assignments of backbone ^{1}H and ^{15}N resonances were 97% complete excluding 10 proline residues, ^{13}C backbone assignments were 92% complete, and assignments of side-chain 1H and ^{13}C resonances were 81% and 79% complete, respectively. Difficulties in assignments were encountered as a result of spectral overlap and missing cross-peaks in ^{13}C-related spectra. C-BP-2 has larger transverse relaxation rate R_2 than expected for a protein of this size, as revealed by ^{15}N relaxation analysis (see below), giving rise to broad linewidths and some missing NMR peaks. Chemical shift assignments have been deposited in the BioMagResBank database (Seavey et al., 1991) with accession number 7137. H^{N}, H^{α}, C^{α}, and N chemical shift deviations from random coil values are plotted in Figure 4.4.

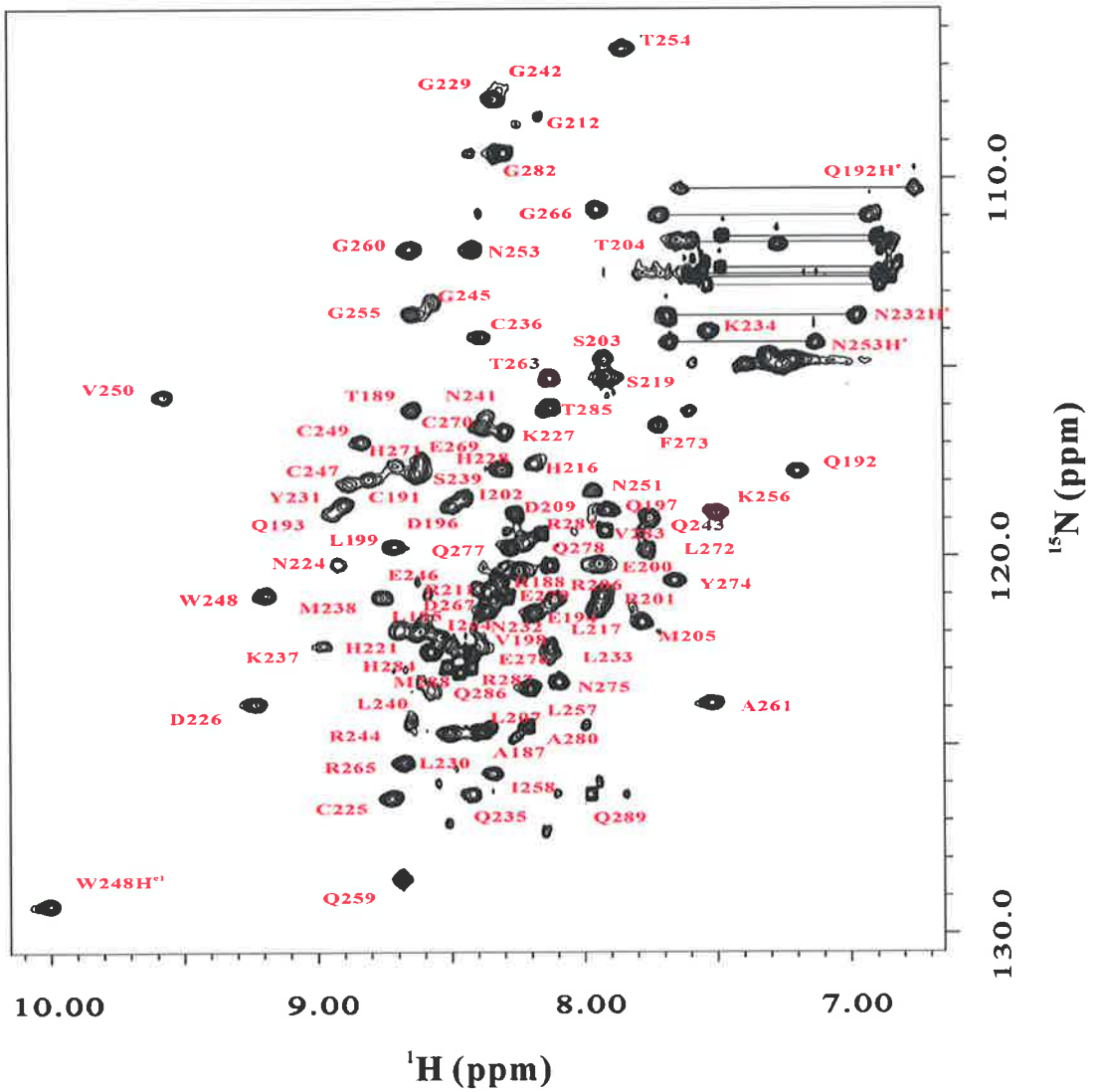


Figure 4.1 ^{1}H-^{15}N HSQC spectrum of C-BP-2

^{1}H-^{15}N HSQC spectrum of C-BP-2 of a 0.05 mM C-BP-2 sample at 298 K and pH 6.0, recorded on an Avance 500 spectrometer equipped with a cryoprobe. Chemical shift assignments are labelled.

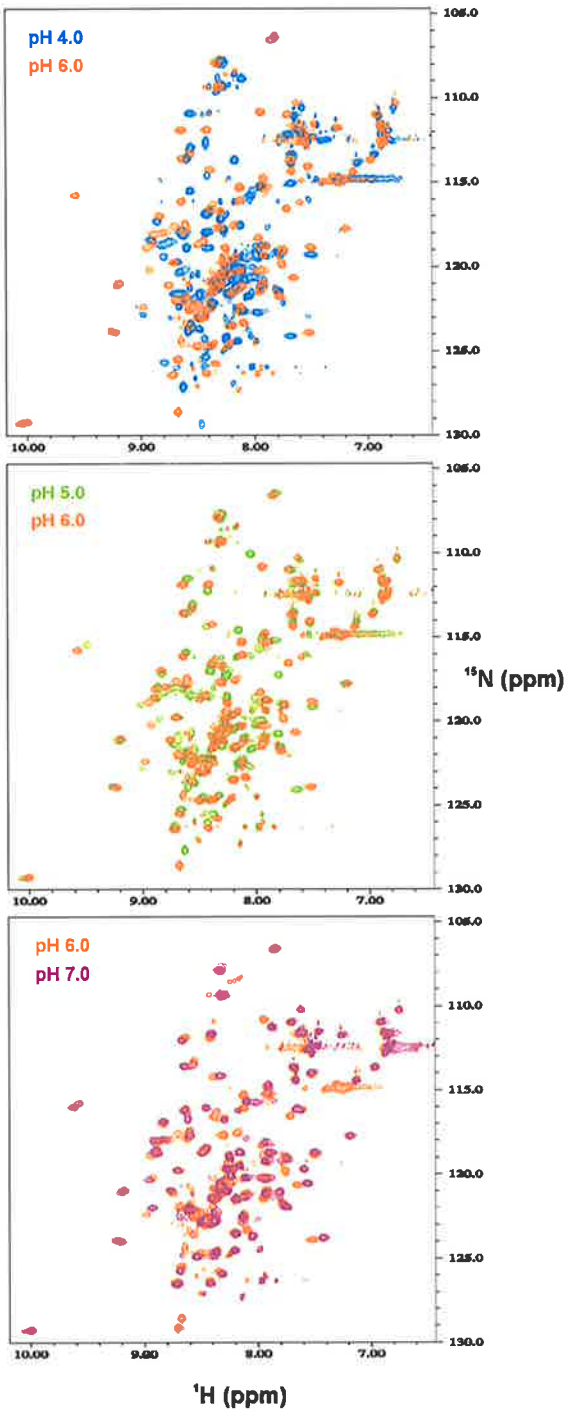


Figure 4.2 Overlay of C-BP-2 ^{1}H-^{15}N HSQC spectra at different pH values

These spectra were acquired using a 0.06 mM ^{15}N-labelled C-BP-2 sample at 298 K on an Avance 500 spectrometer equipped with a cryoprobe.

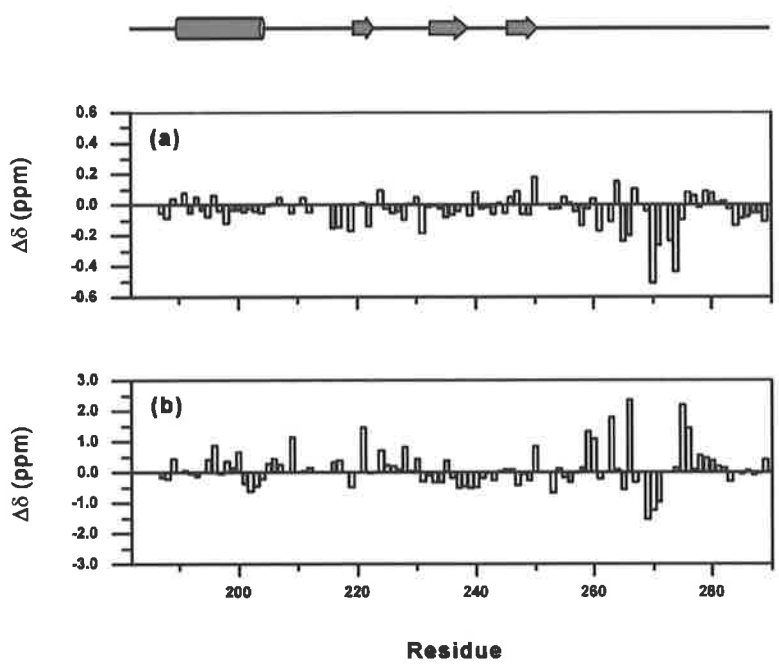


Figure 4.3 Chemical shift differences of (a) H^N and (b) N between pH 6.0 and pH 4.0.

Chemical shifts of a 0.06 mM 15N-C-BP-2 sample at 298 K. $\Delta\delta=\delta_{(pH\ 6.0)}-\delta_{(pH\ 4.0)}$.

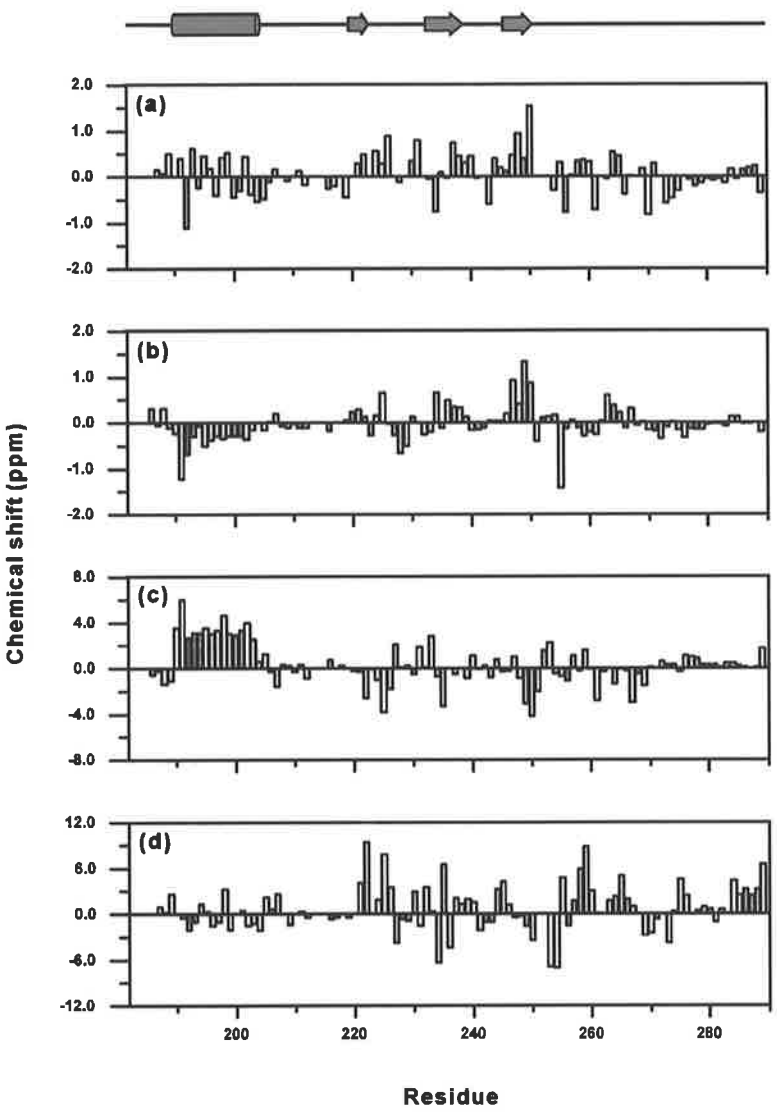


Figure 4.4 C-BP-2 chemical shift deviations from random coil values

(a) H^N, (b) H^α, (c) C^α, and (d) N. Chemical shifts of ^{15}N-C-BP-2 (0.8 mM) and $^{15}N/^{13}$C-C-BP-2 (0.5 mM) samples at pH 6.0 and 298 K. Random coil chemical shift values were taken from (Wishart et al., 1995a).

4.3.2 Solution Structure of C-BP-2

Table 4.3 summarizes the input data for structure calculations, as well as structure statistics for the final family of 20 structures. The structures fit the experimental restraints, with no inter-proton distance and torsion angle violations greater than 0.2 Å and 5°, respectively.

Superimposition of the 20 backbone structures and ribbon views of the closest-to-average structure are shown in Figure 4.5. Surface models showing the residues corresponding to the IGF-II binding site on C-BP-6 and the electrostatic potentials at neutral pH are shown in Figure 4.6. Angular order parameters for ϕ, ψ and $\chi 1$ dihedral angles are summarized in Figure 4.7. The coordinates of the 20 final structures have been deposited in the Protein Data Bank (Berman et al., 2002) with accession code 2H7T.

At the N-terminus, residues Gly183 to Pro186 are not restrained because of a lack of assignments for the first residue and the following triplet of prolines, whereas Ala187 to Pro190 form a turn-like structure. The first major secondary structure element of C-BP-2 is an α-helix spanning residues Pro190 to Thr204. The α-helix is followed by the first loop, which encompasses residues Met205 to Ser219 and is disordered. Pro208 and Pro213 are located in this loop. The presence of a weak doublet of Gly212 amide cross peaks in the ^{1}H-^{15}N-HSQC spectrum implies that the Gly212-Pro213 peptide bond undergoes *cis-trans* isomerization. This could not be confirmed in the 13C-NOESY-HSQC spectrum because of peak overlap. Loop I is followed by the first strand, Leu220 to Ile222, of a three-stranded anti-parallel β-sheet. Residues Pro223 to Asp226 form an extended strand-like structure with β-strand ϕ and ψ angles, but inter-strand $d_{\alpha\alpha}$ or $d_{\alpha N}$ NOEs with the second strand are not observed. In addition to hydrophobic interactions, the helix and the first extended strand are further stabilized by the Cys190-Cys225 disulfide bond. A type I β-turn is formed by Asp226 to Gly229, between the first and second strands. The second strand, Lys234 to Met238, and the third strand, Glu246 to Val250, are stabilized by the Cys236-Cys247 disulfide bond. Residues Ser239 to Gly245 between the second and third strands form the second flexible loop. The third strand is followed by the third loop, from Asn251 to Cys270. The first part of this loop, Asn251 to Gly255, is stabilized against a hydrophobic core formed by the aromatic ring of Tyr231. In the second part of loop III, an RGD motif

and the flanking resides form a distorted inverse γ-turn (Figure 4.8). The end of loop III is fixed by the Cys249-Cys270 disulfide bond. Residues Leu271-Glu275 are restricted by medium-range ($1< |i - j| < 5$) NOEs, but no long-range NOEs to other parts of the protein were observed. The last 14 residues, 276-289, are unstructured. Structures were also calculated using XPLOR-NIH (Schwieters et al., 2003). The CYANA structures seem to be more extended for this unstructured C-terminal tail (Figure 4.9). Only the CYANA structures, after energy minimization using OPALp (Koradi et al., 2000), were used for further analysis.

Table 4.3 Structural statistics for C-BP-2

Distance constraints	
Total	856
Intra-residue (i = j)	320
Sequential (|i – j| = 1)	300
Medium-range (1< |i – j| < 5)	111
Long-range (|i – j| ⩾ 5)	125
Dihedral constraints	
Total	135
ϕ angles	88
ψ angles	38
$\chi 1$ angles	9
Constrained hydrogen bonds	13
CYANA target function value ($Å^2$) [b]	0.09 ± 0.01
AMBER energies (kcal mol^{-1})	
Total	-3969 ± 118
Van der Waals	-192 ± 15
Electrostatic	-4742 ± 109
Deviations from ideal geometry	
Bond length (Å)	0.0127 ± 0.0001
Bond angles (°)	2.35 ± 0.03
RMS deviations (Å) [c]	
Backbone heavy atoms (N, C^{α}, C')	1.42 ± 0.38
All heavy atoms	2.16 ± 0.39
Ramachandran plot [d]	
Most favored (%)	71.3
Allowed (%)	26.6
Additionally allowed (%)	2.1
Disallowed (%)	0

G-factors [d]	
Overall	-0.42 ± 0.04
All dihedral angles	-0.80 ± 0.05
ϕ and ψ angles	-0.97 ± 0.10

Except for the number of constraints and the CYANA target function, the average values for the 20 energy-minimized conformers with the lowest CYANA target function values are listed. [b] The average target function value for the 20 CYANA conformers before energy-minimization. [c] Mean pairwise RMSD over residues 190-204, 220-238, 246-250. RMS deviations from the mean structure over residues 190-204, 220-238, 246-250 are 0.97 ±0.29 Å for backbone heavy atoms and 1.49 ±0.28 Å for all heavy atoms. [d] As determined by the program PROCHECK-NMR (Laskowski et al., 1996).

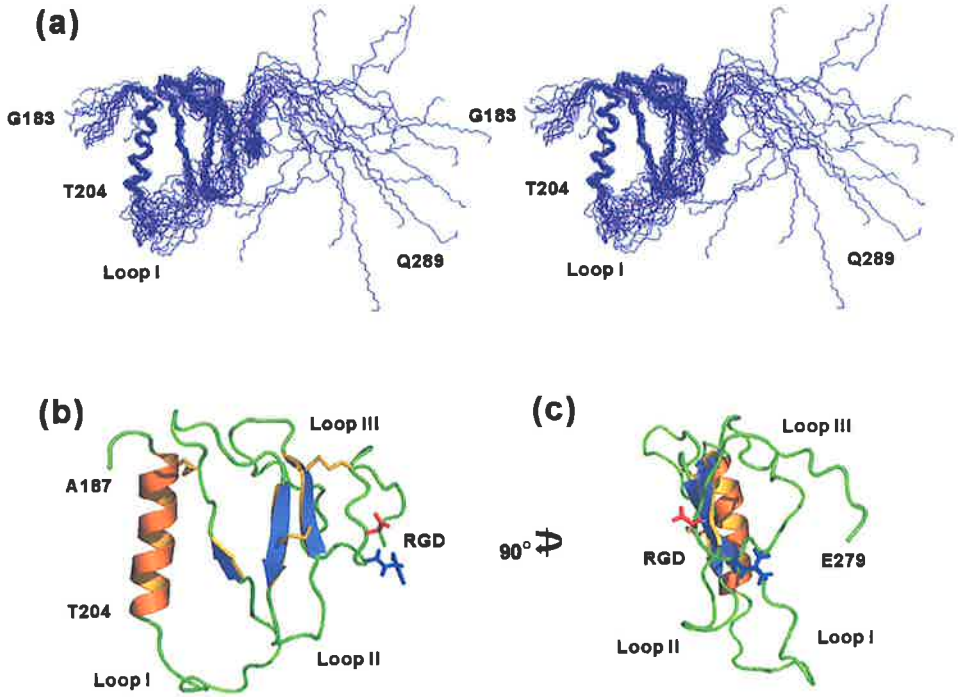


Figure 4.5 C-BP-2 structure

(a) Stereo view of the family of 20 final structures, superimposed over the backbone heavy atoms of residues 190-204, 220-238, 246-250. (b) Ribbon view of the closest-to-average structure highlighting the α-helix, β-sheet, three disulfide bonds and the RGD motif (residues 265-267). Parts of the unstructured regions (residues 183-186 and 280-289) are excluded for clarity. (c) Structure in B, rotated by 90° about the vertical axis. This figure was prepared using MOLMOL (Koradi et al., 1996) and PyMOL (DeLano Scientific).

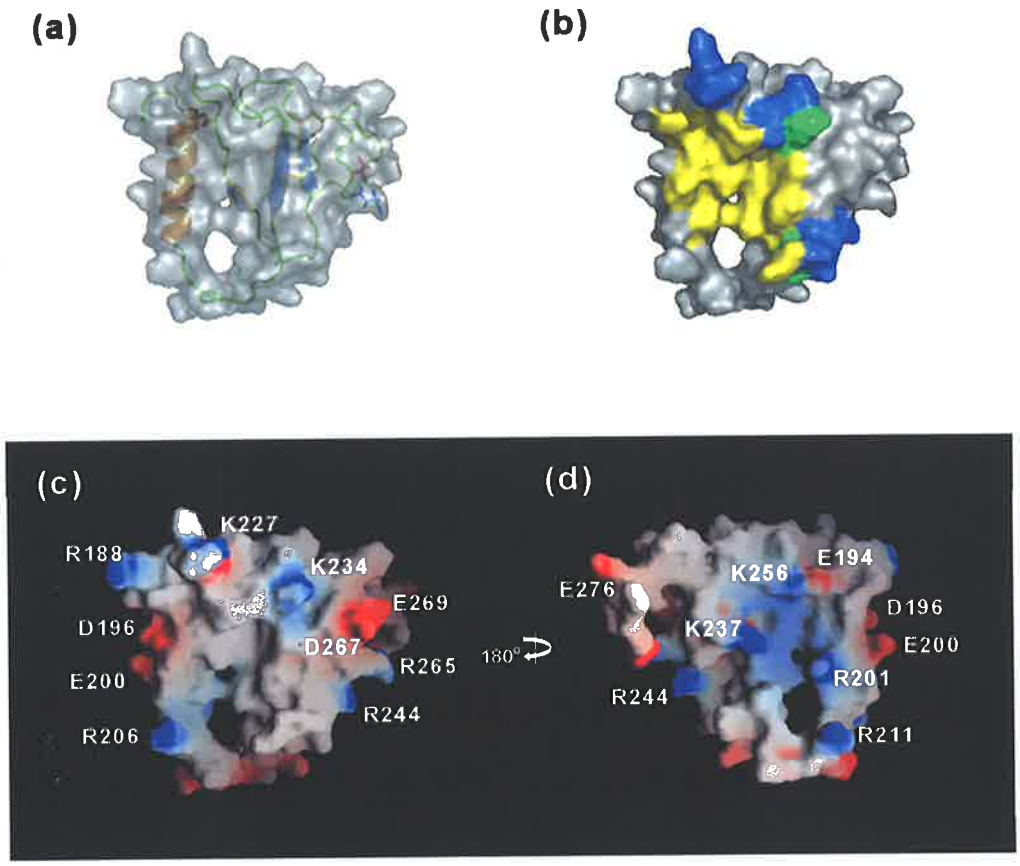


Figure 4.6 Surface models of C-BP-2

(a) Transparent surface model with ribbon representation inside. (b) Surface model with equivalent residues to those involved in IGF-II binding by C-BP-6 (Headey et al., 2004a) coloured yellow, equivalent residues of the clustered basic residues of IGFBP-3 and IGFBP-5 coloured blue, and common residues coloured green. (c and d) Electrostatic potential of C-BP-2 at neural pH. Basic and acidic residues are labelled. (a), (b), and (c) have same orientations as Figure 4.6 (b). This figure was prepared using PyMOL (DeLano Scientific) and GRASP (Nicholls et al., 1991).

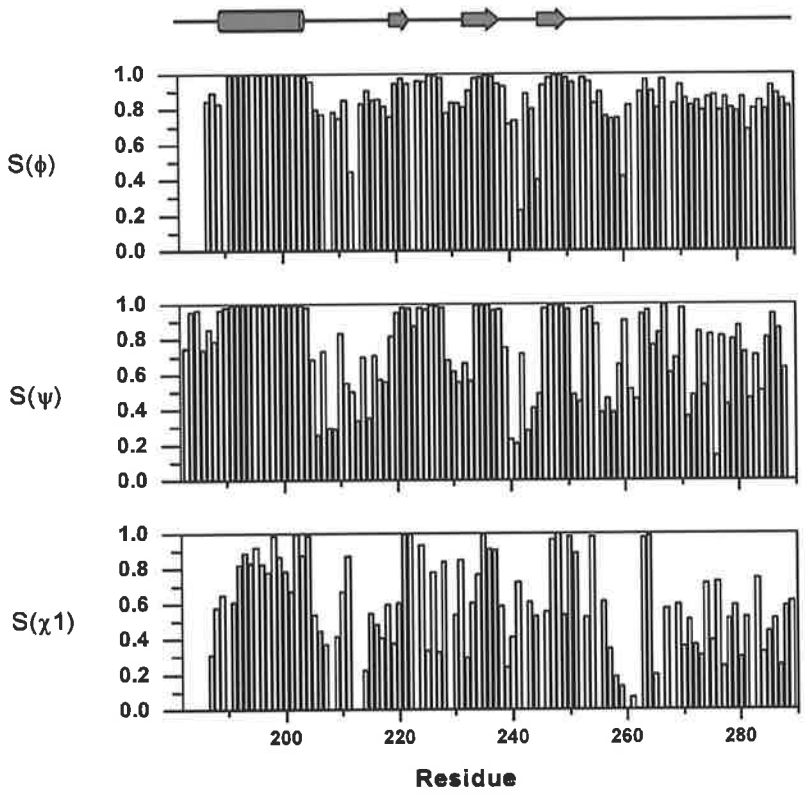


Figure 4.7 Angular order parameters of C-BP-2

Angular order parameters were calculated from 20 final structures using MOLMOL (Koradi et al., 1996).

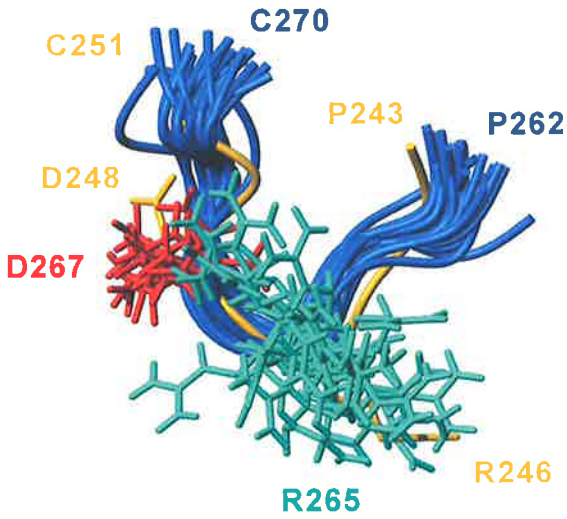


Figure 4.8 Structures of the RGD motifs of C-BP-2 and C-BP-1

Ribbon views of the RGD motifs of C-BP-2 (blue, residues 262-270 of 20 structures) and C-BP-1 (gold, residues 243-251) (Sala et al., 2005) are shown. Side-chains of RGD are shown and labelled (cyan and red for C-BP-2; gold for C-BP-1). This figure was prepared using MOLMOL (Koradi et al., 1996).

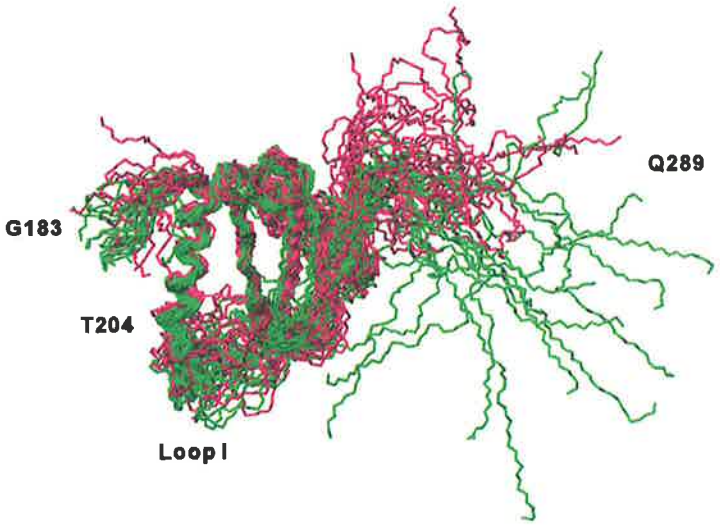
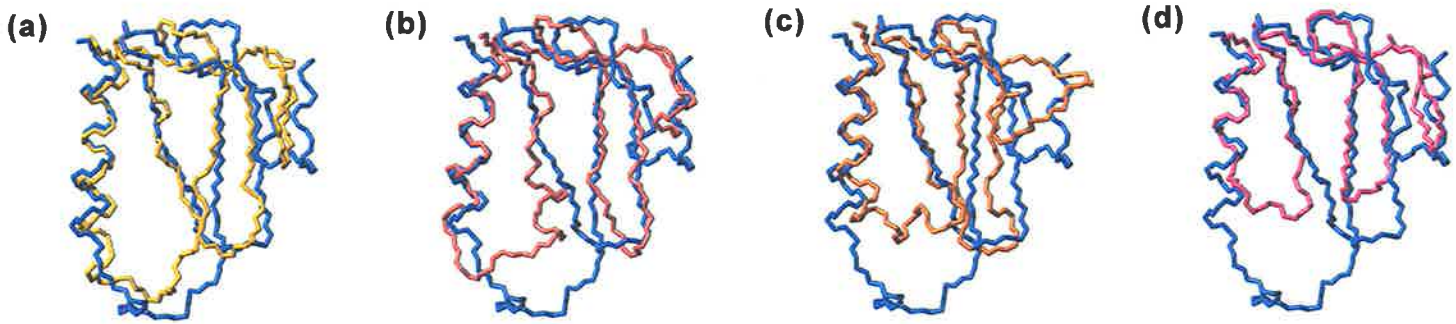


Figure 4.9 Comparison of the CYANA and XPLOR structures of C-BP-2

The family of 20 CYANA structures (green) and the family of 20 XPLOR structures (pink) are superimposed over residues 190-204, 220-238, and 246-250. This figure was prepared using MOLMOL (Koradi et al., 1996).

4.3.3 Comparison with C-BP-1, C-BP-4, and C-BP-6

Figure 4.10 shows the superimpositions of the backbone heavy atoms of C-BP-2 with C-BP-1 (PDB 1ZT3) (Sala et al., 2005), C-BP-4 (PDB 2DSR) (Sitar et al., 2006), C-BP-6 (PDB 1RMJ) (Headey et al., 2004a), and the MHC class II-associated invariant chain fragment p41 Ii (1ICF) (Guncar et al., 1999). Sequence alignments of IGFBP C-domains and several representative non-IGFBP thyroglobulin type 1 domains, including p41 Ii, are also shown in Figure 4.10. C-BP-2 shares 39%, 38%, and 33% sequence identity with C-BP-1, C-BP-4, and C-BP-6, respectively. Over the equivalents of the C-BP-2 well-defined regions 190-204, 220-238 and 246-250, the average RMSD values calculated using MOLMOL (Koradi et al., 1996) for backbone heavy atoms between C-BP-2 and C-BP-1, C-BP-4, and C-BP-6 are 2.49 Å, 1.74 Å and 1.84 Å, respectively, and between C-BP-2 and p41Ii over C-BP-2 residues 190-197, 220-238 and 246-250 it is 1.18 Å. This indicates that the spatial organization of the helix, the β-sheet and the β-turn between the first and second strands is well conserved among these domains. Interestingly, C-BP-4 residues His172-Ile177 (equivalent to C-BP-2 Leu214-Ser219) appear to have a quite different structure compared to other C-BPs; these residues form the second α-helix in the C-BP-4 structure (Figure 4.10 (b)). Equivalent residues in the C-BP-2 and C-BP-6 solution structures are part of the flexible loop I, whereas in the C-BP-1 crystal structure they form a short β-strand. C-BP-2 has an insertion of four to seven residues between the first and second cysteines compared to other IGFBP C-domains (Figure 4.10 (e)); this insertion results in a longer disordered loop I (15 residues) after the α-helix, compared with those of C-BP-6 (8 residues) and C-BP-1 (4 residues). The α-helix (15 residues) is similar to those of C-BP-6 (15 residues) and C-BP-4 (13 residues) but shorter than that of C-BP-1 (19 residues). In contrast, the lengths of loop II are similar. Another noticeable feature of the C-BP-2 sequence relative to other IGFBP C-domains is its 11-13 residue extended C-terminal tail (Figure 4.10 (e)). The solution structure of C-BP-2 shows clearly that this extended C-terminal tail is unstructured (Figure 4.5).



(e)

```
                    190       200       210       220       230       240       250       260       270       280
IGFBP-2      183 RPPPARTPCQQELDQVLERISTMRLPDERGPLEHLYSLHIPNCDKHGLYNLKQCKMSLNGQRGECWCVNPNTGKLIQGAPTIRGD--PECHLFYNEQQEARGVHTQRMQ 289
IGFBP-1      148 -----KEPCRIELYRVVESLAKA----QETSGEEISKFYLPNCNKNGFYHSRQCETSMDGEAGLCWCVYPWNGKRIPGSPEIRGD--PNCQIYFNVQN---------- 234
IGFBP-3      183 -----YGPCRREMEDTLNHLKFL-------NVLSPRGVHIPNCDKKGFYKKKQCRPSKGRKRGFCWCVDKY-GQPLPG-YTTKGKEDVHCYSMQSK---------- 264
IGFBP-4      150 -----QGSCQSELHRALERLAAS----QSRTHEDLYIIPIPNCDRNGNFHPKQCHPALDGQRGKCWCVDRKTGVKLPGGLEPKGE--LDCHQLADSFRE---------- 237
IGFBP-5      169 -----QGPCRRHMEASLQELKAS-------PRMVPRAVYLPNCDRKGFYKRKQCKPSRGRKRGICWCVDKY-GMKLPGMEYVDGD--FQCHTFDSSNVE---------- 252
IGFBP-6      133 -----MGPCRRHLDSVLQQLQTE-------VYRGAQTLYVPNCDHRGFYRKRQCRSSQGQRRGPCWCVDRM-GKSLPGSPDGNGS--SSCPTGSSG---------- 213
p41 Ii       210 -----LTKCQEEVS-HIPAVHPG-------------SFRPKCDENGNYLPLQCY----GSIGYCWCVFPN-GTEVPNTRSR-GH--HNCSES---------------- 274
Thyroglobulin 12 -----LRPCELQRETAFLK---------------QADYVPQCAEDGSFQTVQCQ----NDGRSCWCVGAN-GSEVLGSRQP-GRP-VACLSF---------------- 76
Nidogen-1    818 -----KTRCQHEREHILGAAGA-----TDPQRPIPPGLFVPECDAHGHYAPTQCH----GSTGYCWCVDRD-GREVEGTRTRPGMT-PPCLST---------------- 894
SMOC-1        66 -----QSKCRLERAQALEQAK----------KPQEAVFVPECGEDGSFTQVQCH----TYTGYCWCVTPD-GKPISGSSVQ-NKT-PVCSGS---------------- 135
Testican-1   289 -----GLPCQNEMNRIQKLS----------KGKSLLGAFIPRCNEEGYYKATQCH----GSTGQCWCVDKY-GNELAGSRKQ-GAV-S-CEEE---------------- 358
TACD1         40 -----AAKCLVMKAEMNGSKLG-RRAKPEGALQNNDGLYDPDCDESGLFKAKQCN-----GTSTCWCVNTA-GVRRTDKDTE-----ITCSER---------------- 115
Saxiphilin    88 -----LQKCLKERQQALAKK----------MIGHYIPQCDEKGNYQPQQCH----GSTGHCWCVNAM-GEKISGTNTPPGQTRATCERH---------------- 156
Equistatin     2 -----LSKCQQLQASAN---------------SGLIGAYVPQCKETGEFEEKQCW----GSTGYCWCVDED-GKEILGTKIR-GS--PDCSRR---------------- 66
```

Figure 4.10 Structural comparisons and sequence alignments of C-BP-2 and other thyroglobulin type 1 domains

Comparisons of C-BP-2 (blue, residues 189-270 of the closest to average structure) and (a) C-BP-1 (PDB 1ZT3, gold, residues 149-226) (Sala et al., 2005), (b) C-BP-4 (PDB 2DSR, coral, residue 151-228) (Sitar et al., 2006), (c) C-BP-6 (PDB 1WQJ, orange, residue 134-207) (Headey et al.,

2004a), and (d) p41 Ii fragment (PDB 1ICF, pink, residues 211-271) (Guncar et al., 1999). (e) Sequence alignments of IGFBP C-domains and several representative non-IGFBP thyroglobulin type 1 domains. Except for saxiphilin (*Rana catesbeiana*) and equistatin (*Actinia equina*), sequences are of human proteins. For proteins containing multiple thyroglobulin type 1 domains (thyroglobulin, SMOC1, saxiphilin, and equistatin), sequences of the first thyroglobulin type 1 domains are shown. Numberings correspond to full-length proteins without secretory signal sequences. Conserved cysteines are coloured red. IGFBP-2 (Swiss-Prot P18065): numbering is given above the sequence; primary heparin binding residues are coloured pink; loop II residues are coloured green; the RGD motif is coloured orange. IGFBP-1 (P08833): the RGD motif is coloured orange. IGFBP-3 (P17936) and -5 (P24593): the 10 basic residues in the 18-residue regions are coloured blue. IGFBP-6 (P24592): residues involved in IGF-II binding are coloured yellow. IGFBP-4 (P22692); p41 Ii (P04233); thyroglobulin (P01266); nidogen-1 (P14543); SMOC1 (secreted modular calcium binding protein 1, Q9H4F8); testican-1 (Q08629); TACD1 (tumour-associated calcium signal transducer 1, P16422); saxiphilin (P31226); equistatin (P81439). Comprehensive sequence alignments of thyroglobulin type 1 domains can be found elsewhere (Novinec et al., 2006).

4.3.4 15N relaxation parameters, overall rotational and translational motions

Representative ^{15}N relaxation decay curves for C-BP-2 are shown in Figure 4.11, and backbone ^{15}N R_1, R_2 and steady-state ^{15}N-{^1H} NOE values are summarized in Figure 4.12 and compared with those of C-BP-6 (Yao et al., 2004). After residues heavily overlapped in the spectra were excluded, all three R_1, R_2 and steady-state ^{15}N-{^1H} NOE values were measured for 68 residues, giving averaged values of 1.22 ± 0.23 s^{-1}, 16.54 ± 6.99 s^{-1} and 0.46 ± 0.57, respectively. While the average R_1 rate is modestly lower than that of C-BP-6 (1.56 ± 0.18 s^{-1}), the average R_2 value is considerably higher (C-BP-6, 10.35 ± 3.17 s^{-1}) (Yao et al., 2004).

Of the 68 residues for which R_1, R_2 and steady-state ^{15}N-{^1H} NOE values were measured, 25 satisfy the criteria of steady-state ^{15}N-{^1H} NOE ≥ 0.6 and $|T_1/T_2 - \langle T_1/T_2 \rangle| \leq$ SD, indicating that those spins are unlikely to participate either in slow internal motion or in chemical/conformational exchange processes (Kay et al., 1989). An isotropic global correlation time, τ_c, of 13.8 ± 1.3 ns was thus obtained from the ratios of R_2 and R_1 for these 25 residues. This correlation time is significantly longer than 8.3 ± 0.5 ns of C-BP-6 (Yao et al., 2004), as well as τ_c values of molecules with similar molecular masses (Maciejewski et al., 2000). The principal moments of the inertial tensor of C-BP-2 for the closest-to-average structure are 1.00:0.69:0.61. Fitting ^{15}N relaxation parameters to an axially symmetric rotational diffusion tensor model, however, was unsuccessful. We were therefore unable to perform a full analysis of these measured relaxation parameters using the conventional *Modelfree* approach, as performed previously for C-BP-6 (Yao et al., 2004).

To probe possible self-association, ^{15}N relaxation rates and translational diffusion coefficients of C-BP-2 at different concentrations were compared. Average ^{15}N R_1 and R_2 values measured using the first-increment FIDs of ^{15}N T_1- and T_2-weighted ^{15}N-HSQC spectra of C-BP-2 at 0.03 mM are 1.28 s^{-1} and 10.6 s^{-1}, respectively. When possible differences in viscosity are taken into account, these values agree with the R_1 (1.23 s^{-1}) and R_2 (11.2 s^{-1}) values at 0.3 mM derived in the same manner. It is worth noting, however, that although this method provided estimates of the R_1 and R_2 for the diluted sample, these "average" values, which differ from the average values calculated above from individual residues, seem to be significantly biased by residues in the flexible regions. The translational diffusion coefficients of C-BP-2 were 1.22 and 1.24

$\times 10^{-10}$ m^2 s^{-1} at 298 K and concentrations of 0.3 and 0.03 mM, respectively (Figure 4.13). These results agree very well with 1.20 × 10-10 m^2 s^{-1} for C-BP-6 under similar conditions (Yao et al., 2004). Although C-BP-6 is smaller than C-BP-2, the C-BP-6 construct had a 27-residue leader sequence and the total residue number, 107, is the same as for C-BP-2.

4.3.5 Backbone dynamics from reduced spectral density mapping

As an alternative approach, reduced spectral density mapping analysis, which makes no assumptions about the nature of molecular global rotational diffusion (Farrow et al., 1995), was performed for the ^{15}N relaxation parameters of C-BP-2 and C-BP-6, and the results are summarized in Figure 4.14.

The reduced spectral density functions at different frequencies, $J(0)$, $J(\omega_N)$, and $J(0.87\omega_H)$, describe the motions of the polypeptide chain. $J(0)$ is sensitive to both internal motions on the nanosecond time scale and slow motions on the millisecond to microsecond time scale, whereas $J(\omega_N)$ and $J(0.87\omega_H)$ are sensitive to fast internal motions on the time scales of $1/\omega_N$ and $1/\omega_H$, respectively (Viles et al., 2001). High $J(0)$ values of C-BP-2 may reflect contributions from chemical or conformational exchange, as well as the influence of the disordered regions (see below). On the other hand, the good agreement of $J(0.87\omega_H)$ between C-BP-2 and C-BP-6 is consistent with the fact that both domains are folded and adopt similar structures (Farrow et al., 1995). C-BP-2 residues in the helix, the β-strands and the first β-turn exhibit $J(0.87\omega_H) < 7.5$ ps rad^{-1} (Viles et al., 2001), indicating a relative lack of internal flexibility. Residues that exhibit $J(0.87\omega_H)$ between 7.5 and 15 ps rad^{-1} possess significant internal flexibility (Viles et al., 2001); these include residues in loops I and II, and the second part of loop III (Gly260-Asp267) including the RGD motif (~10 ps rad^{-1}). Residues in the C-terminal tail beyond Gln277 exhibit $J(0.87\omega_H) > 15$ ps rad^{-1} and thus have greater flexibility.

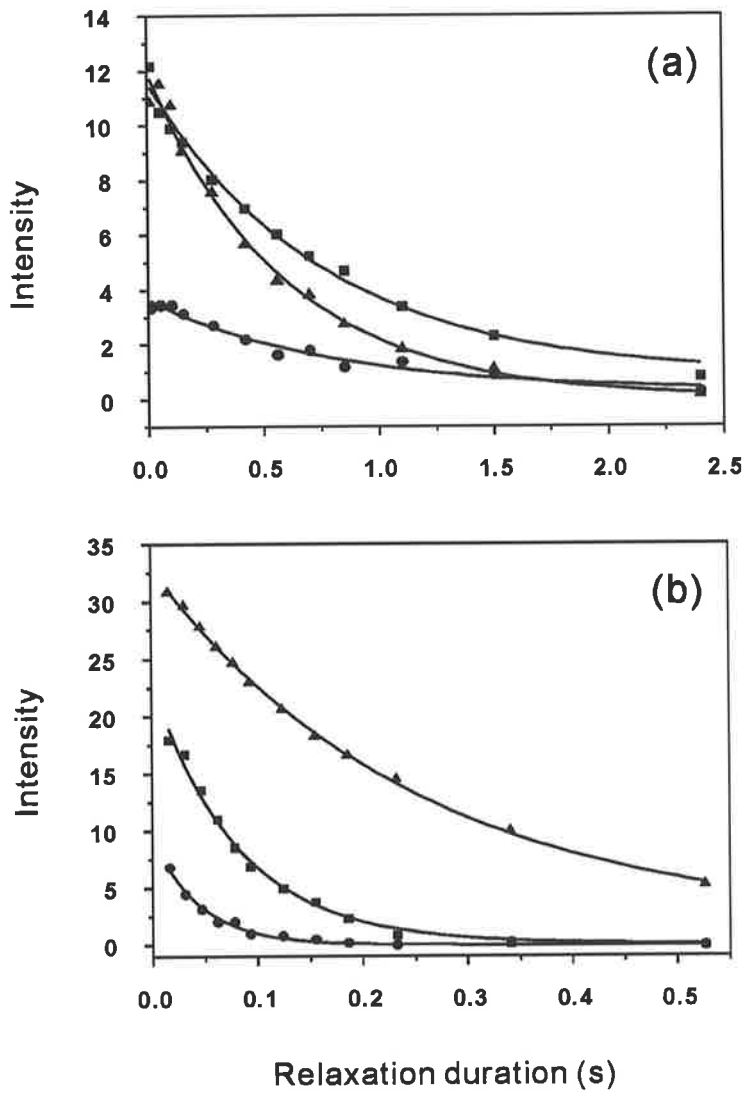


Figure 4.11 Representative 15N relaxation decay curves of C-BP-2

(a) R_1 and (b) R_2 for residues R188 (■), E194 (●), and T285 (▲). ^{15}N relaxation experiments were performed using a 0.3 mM 15N-C-BP-2 sample at pH 6.0 and 298 K on a DRX-600 spectrometer.

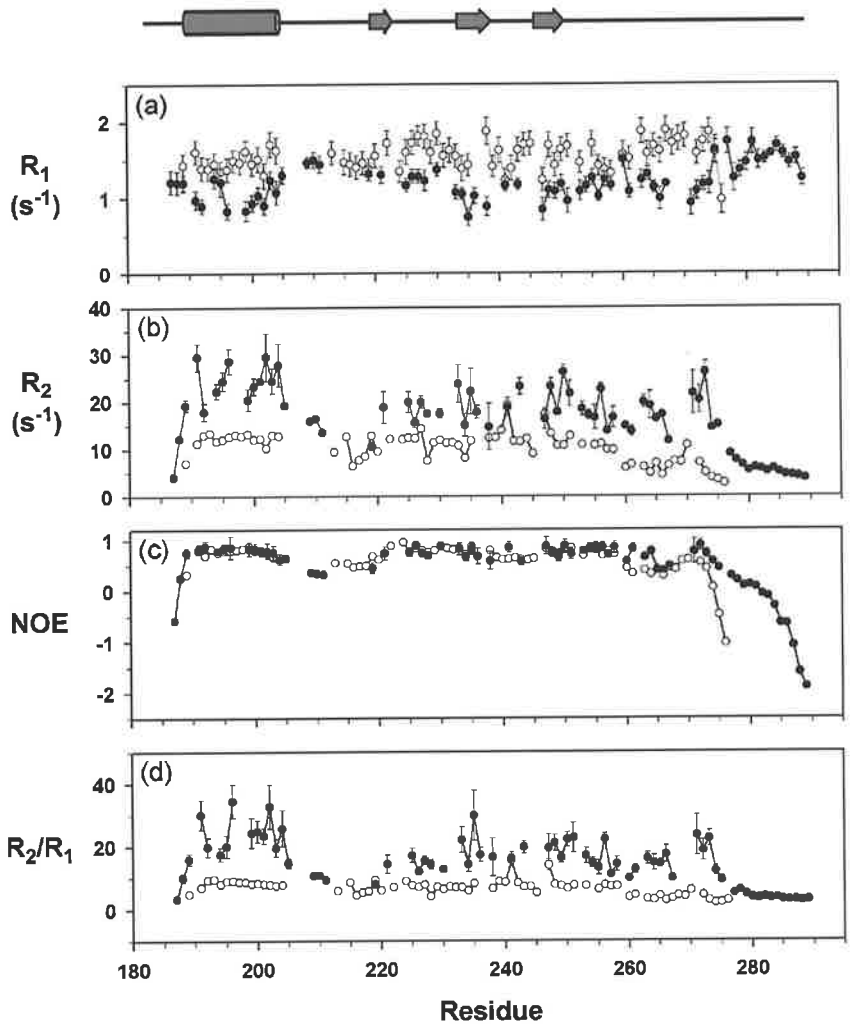


Figure 4.12 Summary of backbone 15N relaxation parameters

(a) R_1, (b) R_2, (c) steady-state 15N-{1H} NOE and (d) R_2/R_1 ratios for C-BP-2 (●), measured using a 0.3 mM sample at pH 6.0, 298 K and 15N frequency of 60.81 MHz. Corresponding values for C-BP-6 (○) measured under similar experimental conditions (Yao et al., 2004) are also depicted for comparison using the sequence alignment of C-BP-2 and C-BP-6 given in Figure 4.10 (e). Residue numbers correspond to C-BP-2. Schematic representation of the secondary structure elements of C-BP-2 is also shown.

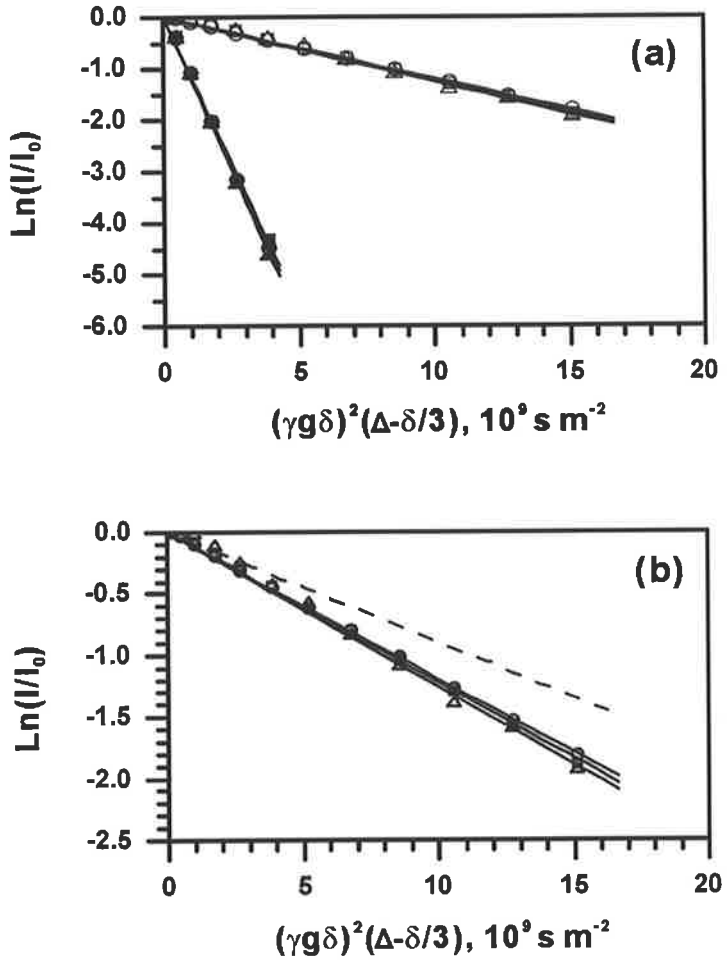


Figure 4.13 Translational diffusion coefficients of C-BP-2

(a) Logarithmic (normalised) intensities of C-BP-2 at 0.3 mM (○) and 0.03 mM (△) and C-BP-6 at 1.0 mM (□) versus the strength of diffusion encoding, $\gamma 2g2\delta 2(\Delta-\delta/3)$. Corresponding intensities of NaAc signals of the 0.3 mM C-BP-2 sample (●), the 0.03 mM C-BP-2 sample (▲), and the C-BP-6 sample (■) are also shown. As expected they attenuated much faster than the protein signals. Lines represent the results of nonlinear regression to equation (1). (b) A partial expansion of (a) showing only the signal attenuations of C-BP-2 and C-B-6. The dashed line represents the calculated attenuation of a 25% increase in translational diffusion coefficient, which corresponds to the change of translational diffusion coefficient for a spherical molecule upon dimerization in the ideal case (Teller et al., 1979).

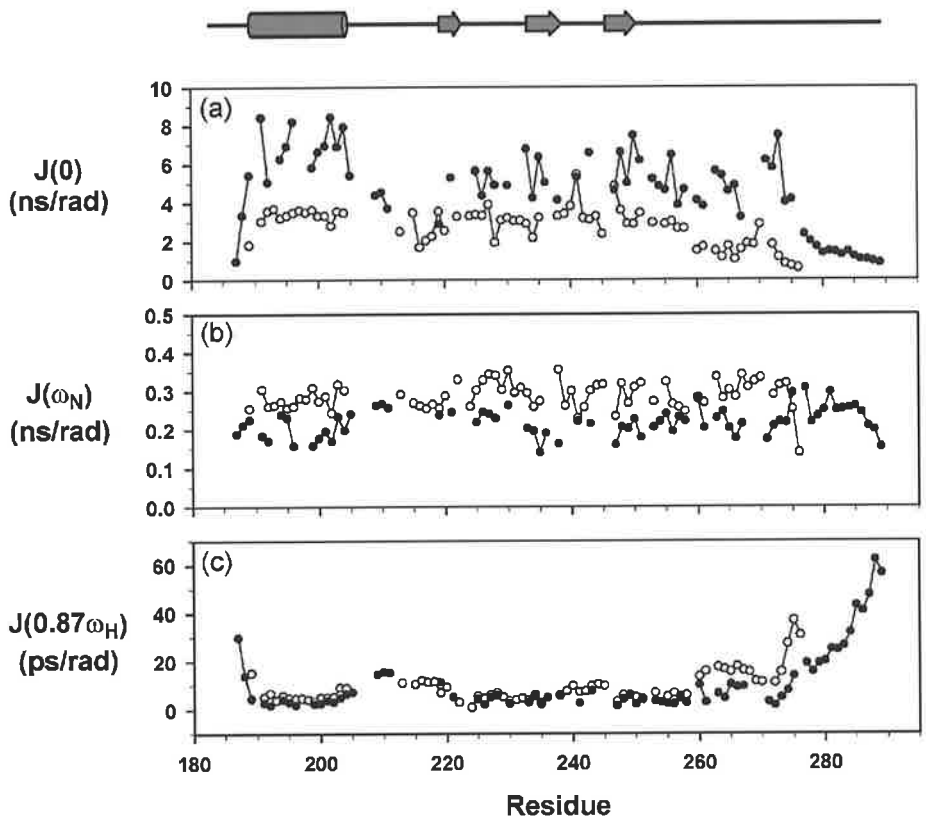


Figure 4.14 Reduced spectral density functions of C-BP-2

(a) $J(0)$, (b) $J(\omega_N)$ and (c) $J(0.87\omega_H)$ calculated from the data shown in Figure 4.13 using Equations (2-5) for C-BP-2 (●) and C-BP-6 (○) are compared according to the sequence alignment of C-BP-2 and C-BP-6 given in Figure 4.10 (e). Residue numbers correspond to C-BP-2. Schematic representation of the secondary structure elements of C-BP-2 is also shown.

4.4 Discussion

4.4.1 Structural differences and implications for IGF binding and functional diversity

The C-BP-2 structure has the same thyroglobulin type 1 fold as IGFBP C-domain structures reported during the course of this thesis, i.e., C-BP-6 (Headey et al., 2004a), C-BP-1 (Sala et al., 2005), and C-BP-4 (Sitar et al., 2006). However, differences among these structures that may be relevant to the functional differences are evident. Based on sequence homology, single or multiple thyroglobulin type 1 domains have been found in many proteins with different origins and functions (Novinec et al., 2006), but, apart from IGFBP C-domains, the only other thyroglobulin type 1 structure published is of the p41 Ii fragment (Guncar et al., 1999; Chiva et al., 2003). C-BP-2 is not only the largest C-domain of any IGFBP but also the largest thyroglobulin type 1 domain structure reported thus far.

The largest differences among the known structures are in the length of the α-helix and the length and position of the flexible loops. Equivalent residues of C-BP-2 Leu214-Ser219 in C-BP structures adopt very different conformations, i.e., α-helix in C-BP-4 but β-strand in C-BP-1 and are part of the flexible loop I in both C-BP-2 and C-BP-6 NMR structures. It is worth noting that the C-BP-4 structure is in a IGF-I·N-BP-4·C-BP-4 ternary complex, whereas other C-BP structures are in the absence of the ligand or the N-domain. It is therefore possible that this region, as well as the loop II region (see below), in IGFBP C-domains undergoes significant conformational change or is stabilised in one pre-existing conformation upon ligand binding.

The α-helices in all C-BPs are longer than in p41 Ii and the helix in C-BP-1 is longer than those in C-BP-2, C-BP-4, and C-BP-6. The end of the α-helix in C-BP-2 at this position is confirmed by the pattern of deviations of C^{α} and H^{α} chemical shifts from their random coil values (Figure 4.4). In C-BP-6, the end of the α-helix is involved in IGF-II binding, as the ^{1}H-^{15}N cross peak of Thr152 (equivalent to Thr204 in C-BP-2) disappeared and Gln151 (equivalent to C-BP-2 Ser203) showed a large increase in ^{15}N transverse relaxation rate upon IGF-II binding (Headey et al., 2004a). Gln151, among some other residues, was found to undergo conformational or chemical exchange that may correlate with the IGF-II binding (Yao et al., 2004). Recently, Mark

and co-workers identified from human haemofiltrate an IGFBP-2 fragment (Gly167-Glu279) that has a cleaved peptide bond between Thr204 and Met205 but is apparently stabilized by disulfide bonds (Mark et al., 2005), which is consistent with our results showing a less well ordered structure in this region. This hydrolytic fragment has four-fold higher IGF-I binding affinity but three-fold lower IGF-II binding affinity compared to the equivalent 167-279 fragment without cleavage (Mark et al., 2005), indicating that this region influences IGF binding.

C-BP-2 residues between Cys236 and Cys247 form the flexible loop II and part of the second and third β-strands flanking the loop II. Equivalent residues in IGFBP-6 contribute to IGF-II binding (Headey et al., 2004a), whereas in IGFBP-3 and -5 they are not only the major acid-labile subunit binding sites (Firth et al., 1998), but are also essential for nuclear localization (Schedlich et al., 2000), and in IGFBP-3 can form a metal binding site (Singh et al., 2004). Although the composition of the loop II in C-BP-2 is very different from that of C-BP-6 (Figure 4.10 (e)), the length of the loop and its dynamic properties, as measured by ^{15}N relaxation data, are similar. They both have considerable flexibility (see below) and this may be required for high-affinity IGF binding, with loop II sampling a range of conformations in the absence of ligand, but including those that are most complementary to the bound IGF structures. In fact, in subsequent NMR binding experiments, it is found that loop II residues not only interact with IGFs, consistent with previous data of C-BP-6, but also interact with N-BP-2 (Chapter 6). These results are also supported by the recently published crystal structure of IGF-I·N-BP-4·C-BP-4 ternary complex (Sitar et al., 2006). Loop II in several non-IGFBP thyroglobulin type 1 domains is shorter by four residues than in IGFBP C-domains (Figure 4.10 (e)). In the crystal structure of p41 Ii and cathepsin L complex, the loop II residues are inserted into a cleft formed by cathepsin L (Guncar et al., 1999), which is a different interaction mode compared to the IGF binding by C-BPs. The variety and flexibility of loop II and its flanking residues are thus important both for modulating IGF binding affinity and for other molecular interactions in thyroglobulin type 1 domains.

4.4.2 Structure and function of the RGD motif

IGFBPs interact with several other molecules through their C-domains, and these interactions can modulate IGF actions and/or mediate IGF-independent effects (Firth and Baxter, 2002; Bach et al., 2005). IGFBP-1 and 2, but not IGFBP-3 to -6, have RGD motifs in their C-domains.

In C-BP-2, this RGD motif is located in loop III and is solvent exposed; it forms a distorted inverse γ-turn turn as defined by short and medium-range NOEs. Superposition of the backbone heavy atoms of this region with the corresponding region of C-BP-1 showed that the conformations of the RGD turns are very similar, although some flanking residues are different (Figures 4.8 and 4.10 (e)). Nevertheless, backbone amide dynamics showed that this part of the loop III has considerable mobility (section 4.3.5). Its structure and dynamic properties are thus consistent with functional RGD motifs in many other proteins (Carr et al., 1997).

Jones and co-workers demonstrated that IGFBP-1 stimulated Chinese hamster ovary cell migration and bound to α5β1 integrin, both in an RGD-dependent manner since WGD mutant did not have these effects (Jones et al., 1993b). The effect of IGFBP-1 on migration was independent of IGF-I and may be mediated by the α5β1 integrin (Jones et al., 1993b). Similarly, it has been shown that IGFBP-1 can stimulate extravillous trophoblast cells, by binding of its RGD motif to the α5β1 integrin, and activate focal adhesion kinase (FAK) and mitogen-activated protein kinase (MAPK) pathway (Gleeson et al., 2005).

IGFBP-2 bound to α5β1 integrin on the cell surface through its RGD motif; the interaction caused dephosphorylation of FAK and MAPK, and may correlate with the enhanced cell de-adhesion and decreased cell proliferation by IGFBP-2 (Schutt et al., 2004). Recently, Wang and co-workers reported that overexpression of IGFBP-2 in glioma cells up-regulated expression of integrin α5, and that IGFBP-2 interacts with integrin α5 through the RGD motif in the IGFBP-2 C-domain (Wang et al., 2006). Overexpression of IGFBP-2 resulted in extensive cell surface lamellipodia and increased migration, which could be abrogated by integrin α5 siRNA (Wang et al., 2006). Previously, these authors discovered a novel protein, invasion inhibitory protein 45 (IIp45), which bound to the C-domain of IGFBP-2 and inhibited glioma cell invasion (Song et al., 2003). Although detailed support was lacking in that paper, the

authors proposed that IIp45 may inhibit IGFBP-2 activation of the integrin pathway by interacting with the RGD motif of IGFBP-2 (Song et al., 2003).

Although the above studies showed that the RGD motif of IGFBP-2 mediates cell surface association via integrin in some tumour cells, Hoeflich and co-workers reported that cell surface association of IGFBP-2 in IGFBP-2 transgenic mice is RGD-independent (Hoeflich et al., 2002). In addition, interaction of IGFBP-2 with αvβ3 integrin negatively modulate IGF-mediated breast tumour cell migration, and correlated with reduced tumour size of cancer cells that overexpressed the αvβ3 integrin; but the interaction between IGFBP-2 and αvβ3 integrin may be RGD-independent, since it was not inhibited by RGD-containing peptides (Pereira et al., 2004).

4.4.3 Backbone flexibility of C-BP-2

NMR relaxation methods have been applied to probe the dynamics of IGFBP domains and reveal significant correlations between their dynamics and IGF binding (Renner and Holak, 2001; Yao et al., 2004). Previous results from *Modelfree* analysis and other relaxation dispersion data indicated that conformational exchange of a substantial number of C-BP-6 residues on sub-millisecond to millisecond time scales correlated with IGF binding (Yao et al., 2004). In this study we have compared the backbone dynamics of C-BP-2 and -6. C-BP-2 has considerably higher backbone ^{15}N R_2 values than C-BP-6 (Figure 4.12), and the reduced spectral density mapping analysis showed significantly higher $J(0)$ values for C-BP-2, although there is good agreement of $J(0.87\omega_H)$ between C-BP-2 and C-BP-6 (Figure 4.14), indicating they have very similar picosecond time scale motions in corresponding regions. Since $J(0)$ is sensitive to both fast motions on the nanosecond time scale and slow motions on the microsecond to millisecond time scales, care must be taken in interpreting these results. After considering the fast motion information provided by $J(0.87\omega_H)$ and $J(\omega_N)$, increased $J(0)$ values for some residues in both C-BP-2 and C-BP-6 are likely to have arisen from conformational or chemical exchange, which is consistent with our previous C-BP-6 data (Yao et al., 2004). Residues that may have moderate $J(0)$ increases of this kind seem to be located in equivalent regions in C-BP-2 and C-BP-6, i.e. around the disulfide bonds, around the end of the helix, and in loop II, although a detailed residue-by-residue comparison is compromised by the relatively large error

bars in the C-BP-2 data, especially for R_2, and by missing data for quite a few residues because of peak overlap. Interestingly, the residues that undergo exchange broadening are distributed quite widely in C-BP-6 (Yao et al., 2004) and C-BP-2, while fewer residues were found by Renner and Holak (2001) to have such exchange in a truncated N-domain of IGFBP-5 (mini-N-BP-5) and those residues were converged in a loop involved in IGF binding (Renner and Holak, 2001). This probably reflects the fact that C-domains use a larger surface for IGF binding than mini-N-BP-5, as well as mediating other molecular interactions.

The conformational exchange itself, however, does not appear to be the major contributor to the marked R_2 increase in C-BP-2 compared to C-BP-6, as almost all residues in the structured core are affected. The lack of significant changes in measured average ^{15}N R_1 and R_2 values (from the first-increment FIDs) and translational diffusion coefficients of C-BP-2 at concentrations of 0.3 and 0.03 mM, as well as the excellent agreement between the translational diffusion coefficients of C-BP-2 and C-BP-6, indicate that self-association of C-BP-2 under the experimental conditions is unlikely. A possible explanation for the anomalously higher R_2 values and the failure in fitting the relaxation parameters is that the motions of the core of C-BP-2 cannot be described by a simple anisotropic tumbling model with a unique diffusion tensor. This is probably a consequence of the presence of a longer and very flexible C-terminal tail together with a longer disordered loop compared to C-BP-6. In fact, similar anomalously high R_2 values and consequent failure of conventional *Modelfree* analysis have been reported and discussed in detail by Viles and co-workers for other proteins with long unstructured termini (Viles et al., 2001). Using computer simulations these authors found that even a modest distribution of correlation times, which resulted from the disordered regions, could lead to a large increase in $J(0)$ and a corresponding increase in R_2, while the R_1 and the ^{15}N-$\{^1H\}$ NOE were much less affected (Viles et al., 2001).

Collectively, C-BP-2 possesses significant fast motions in the loops and termini, and may also have slow conformational or chemical exchange in the structured domain core and loop II. In fact, flexibility seems to be a significant feature of IGFBP molecules. In particular, the structure and dynamics of the ~100 residue non-conserved linker domains in IGFBPs, in relation to their function, have not been established, although they are known to contain multiple protease cleavage sites and are generally believed not to have ordered structures. IGFBP-2 is at one extreme in this respect, since it has insertions in both the N-domain (a ~15-residue alanine-rich region) and C-domain,

and a longer C-terminal tail compared to other IGFBPs, making it the largest member in the family and probably the most flexible.

Chapter 5

Heparin Binding by the IGFBP-2 C-terminal Domain

5.1 Introduction

Binding of IGFBPs to glycosaminoglycans and other components of the extracellular matrix is believed to be important in modulation of IGF actions, but previous data regarding the glycosaminoglycan binding by IGFBP-2 have been contradictory.

Glycosaminoglycans are long unbranched polysaccharides containing a repeating disaccharide unit. As shown in Figure 5.1, the disaccharide units contain either of two modified sugars, N-acetylgalactosamine or N-acetylglucosamine and an uronic acid such as glucuronate or iduronate. Heparin, heparan sulfate, keratan sulfate, dermatan surfate, chondroitin sulfate, and hyaluronic acid are commonly found glycosaminoglycans of physiological significance. Although each of these glycosaminoglycans has a predominant disaccharide component (Figure 5.1), heterogeneity exists in terms of the sugar composition and the percentage of the sulfation in any given class. Glycosaminoglycans are highly negatively charged molecules, and are located primarily on cell surface or in the ECM. Over the last two decades heparin and other glycosaminoglycans have been shown to interact with an increasing number of proteins, thereby playing an important role in the regulation of various physiological processes (Capila and Linhardt, 2002; Raman et al., 2005). These include the regulation of proteases such as the interaction of heparin with antithrombin III, modulation of cell growth, differentiation, and inflammation via the interactions of heparin with growth factors and chemokines, and mediation of cell-cell and cell-matrix adhesions, as well as pathogen invasions.

Heparin binding by IGFBPs and its effects on IGF binding were investigated initially with respect to an 18-residue region that contains ten arginine or lysine residues in the C-domains of IGFBP-3 and -5, equivalent to C-BP-2 Lys227-Arg244 (Figure 4.10 (e)). Heparin binding of IGFBP-3 and -5 reduces IGF binding affinity, and

is thought to be important in controlling the availability of unbound IGFs in the ECM (Arai et al., 1994b). Although mutagenesis studies have identified several basic residues in IGFBP-5 that are important for heparin or ECM binding (Arai et al., 1996b; Parker et al., 1998; Song et al., 2000; Allan et al., 2006), the contributions of other basic residues in this region and the three-dimensional location of the binding site are unclear. There is also some debate as to whether IGFBP-2 binds heparin and other glycosaminoglycans. Arai and co-workers reported that IGFBP-2 did not bind heparin but acquired heparin binding ability upon binding to IGFs (Arai et al., 1996b), while Russo and co-workers found heparin binding by IGFBP-2 in the absence of IGFs (Russo et al., 1997), and suggested that the heparin binding site is in the linker domain (Russo et al., 2005). However, others have failed to detect heparin binding by IGFBP-2 either with or without IGF-I (Song et al., 2001; Beattie et al., 2005).

NMR spectroscopy is a very powerful technique for studying molecular interactions, and, while many NMR methods have been developed recently for this purpose, chemical shift perturbation mapping is the most widely used (Zuiderweg, 2002). In brief, the protein whose binding site is to be mapped is ^{15}N-labelled, and its ^{1}H-^{15}N HSQC spectrum is monitored as the unlabelled interaction partner(s) is (are) titrated in. Chemical shift perturbations of a specific set of resonances occur, indicating the changes in the local chemical environment of these nuclei caused by direct contact with the interaction partner and/or a conformational change resulting from the interaction.

In addition to mapping the interface, titration with NMR also allows a good estimation of the affinity and stoichiometry, as well as the kinetics of binding, which determines how the chemical shifts change during the titration (Zuiderweg, 2002). If the complex dissociation is fast, i.e., the inverse of the lifetime ($1/\tau$) is greater than the chemical shift differences ($\Delta\delta$) between free and bound states, only a single set of resonances whose chemical shifts are the fractionally weighted average of the free and bound chemical shifts will be observed. Thus the resonances of the perturbed nuclei move in a continuous manner during the titration (Figure 5.2). This regime is referred to as fast chemical exchange and is often observed for weak interactions with equilibrium constants (K_D) greater than 10^{-5} M. In contrast, if the dissociation is very slow, i.e., $1/\tau$ is smaller than $\Delta\delta$, then one set of resonances for the free protein and another set for the bound resonance are observed. Thus, during the titration, the "free

set" gradually disappears and the "bound set" appears simultaneously (Figure 5.2). This regime is referred to as slow chemical exchange and is often observed for high affinity interactions with K_D at 10^{-8} M or smaller. Interactions with K_D in the 10^{-6}-10^{-7} M range often fall into the intermediate chemical exchange regime, where $1/\tau$ is close to $\Delta\delta$. In this case, extensive broadening of the changing peaks occurs, and when the lines become broad enough, the resonances may disappear from the NMR spectrum (Figure 5.2).

In this Chapter, chemical shift perturbation mapping was used to identify the heparin binding site on C-BP-2. Interactions among IGFs, N-BP-2, and C-BP-2 were also investigated by this approach in following Chapters.

The heparin binding site was identified by titrating ^{15}N-labelled C-BP-2 with sucrose octasulfate (SOS) and low molecular weight heparin oligosaccharides in the presence of 150 mM NaCl, mimicking physiological ionic strength, while monitoring chemical shift changes in the ^{1}H-^{15}N HSQC spectra. SOS has a simpler and better defined chemical structure than heparin (Figure 5.1), and has been used as a functional substitute for heparin in studies of other proteins using NMR (Zhou et al., 1999; Hung et al., 2005).

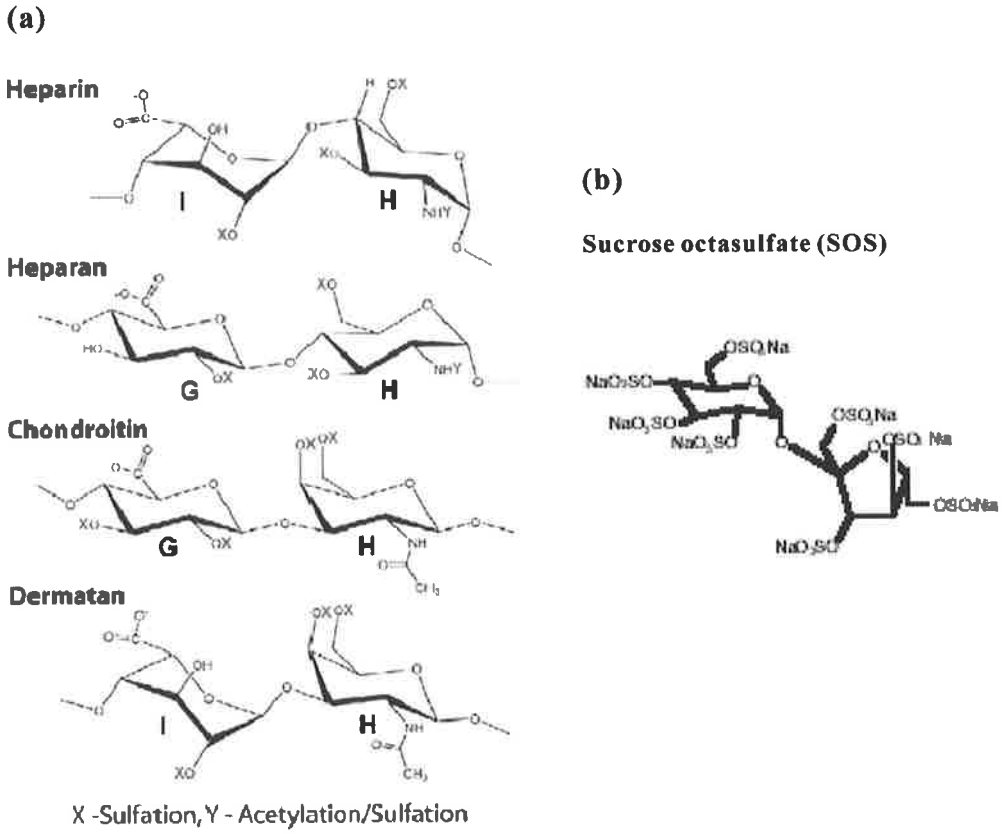


Figure 5.1 Structures of glycosaminalglycans and SOS

Structures of (a) glycosaminalglycans and (b) sucrose octasulfate (SOS, sodium salt). Figure (a) was from (Raman et al., 2005).

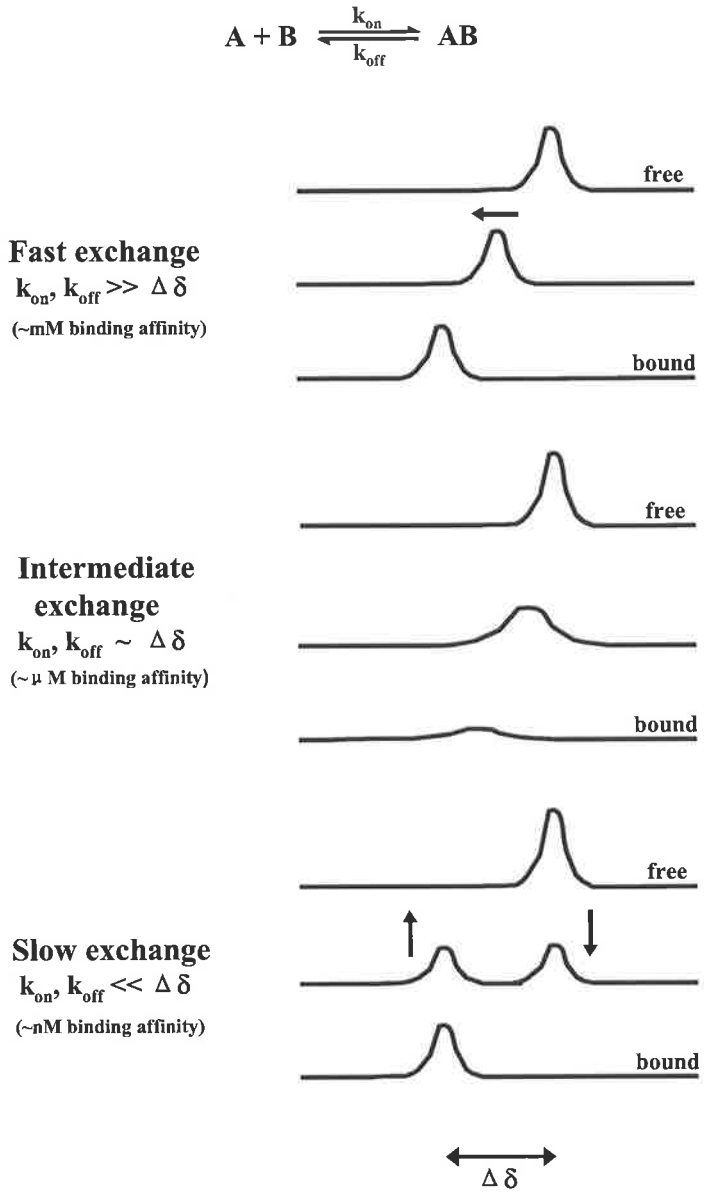


Figure 5.2 Spectral characteristics in different exchange regimes of interaction

In fast exchange, only a single set of resonances whose chemical shifts are the fractionally weighted average of the free and bound chemical shifts will be observed, and the resonances of the perturbed nuclei move in a continuous manner; In intermediate exchange, extensive broadening of the changing peaks occurs, and the resonances may disappear from the NMR spectrum; In slow exchange, one set of resonances for the free protein and another set for the bound resonance are observed (Zuiderweg, 2002).

5.2 Materials and Methods

5.2.1 Materials

^{15}N-labelled C-BP-2 was prepared as described in Chapter 3. SOS was purchased from Toronto Research Chemicals Inc (Toronto, ON, Canada). Low molecular weight heparin, with an average molecular mass of ~3 kDa (~12-15 monosaccharide units), was purchased from Sigma (Saint Louis, MO, USA).

5.2.2 Methods

Titration experiments were carried out by adding concentrated SOS or heparin solution to 0.03 mM ^{15}N-C-BP-2 samples in 95% H_2O/5% $2H_2O$ containing 10 mM sodium acetate, 150 mM NaCl and 0.02% (w/v) sodium azide at pH 6.0. ^{1}H-^{15}N HSQC spectra were recorded at SOS concentrations of 0, 0.2, 0.5, 1.0, 1.5, 2.0, 2.5 and 3.0 mM, or heparin concentrations of 0, 0.5, 1.0, 1.5 and 2.0 mM; changes in resonance position were examined. The spectra were recorded with a data matrix size of 2048×128 and 256 scans per t_1 increment. Spectrum widths were 11.5 ppm for ^{1}H and 30.0 ppm for ^{15}N. Spectra were processed as described in Chapter 4.

^{1}H-^{15}N HSQC spectra of the samples with 0 and 3.0 mM SOS, or 0 and 2.0 mM heparin were also recorded and compared at pH 5.5 and 7.4. Titration experiments with SOS were also performed at pH 5.5.

5.3 Results

5.3.1 Chemical shift perturbations

As shown in Figures 5.3 and 5.4, addition of increasing amounts (0.2-2.5 mM) of SOS to a 0.03 mM ^{15}N-C-BP-2 sample at pH 6.0 induced continuous movements of specific cross-peaks, while 3.0 mM of SOS did not cause further changes compared to 2.5 mM. This indicated that the binding was on the fast-exchange time scale relative to

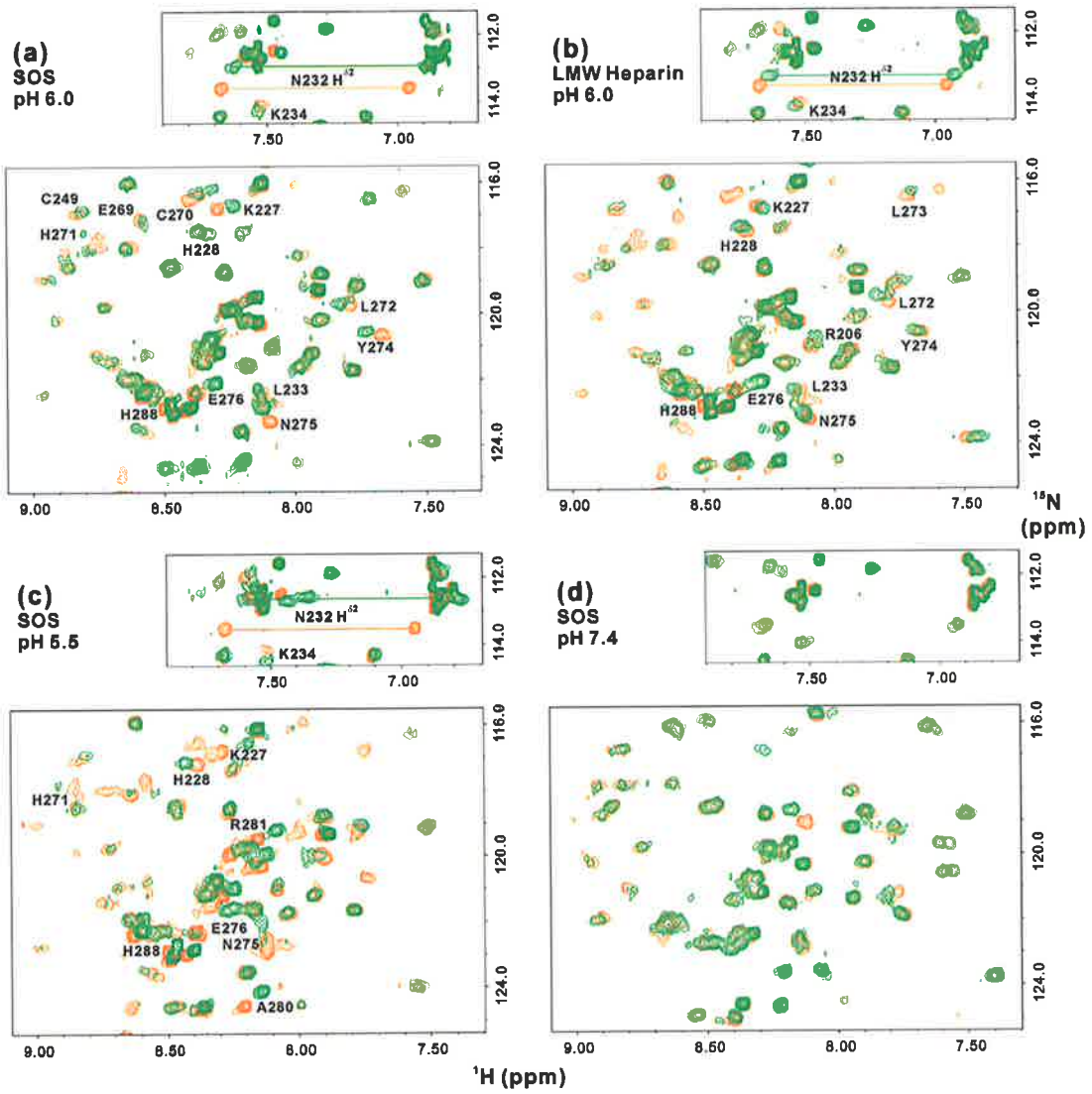


Figure 5.3 SOS or LMW heparin binding by C-BP-2

Overlay of selected regions of the ^{1}H-^{15}N HSQC spectra of 0.03 mM C-BP-2 in the absence (orange) and presence (green) of 3.0 mM SOS (a, c and d), or 2.0 mM low molecular weight heparin (b), at pH 6.0 (a and b), pH 5.5 (c) and pH 7.4 (d). Spectra were recorded at 500 MHz and 298 K. The full spectra are shown in Figure 5.4.

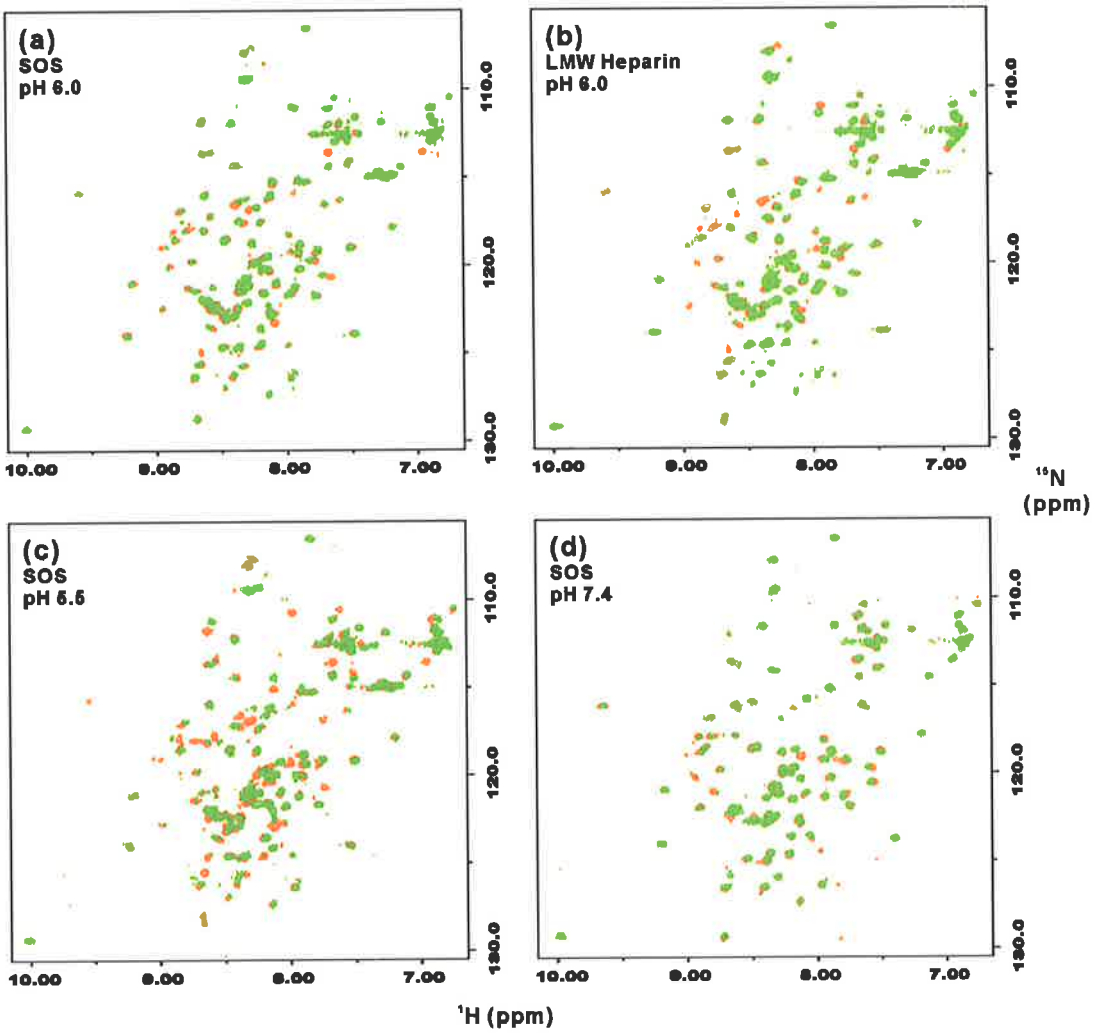


Figure 5.4 SOS or LMW heparin binding by C-BP-2 (full spectra)

Spectra presented in Figure 5.3, but with all peaks shown. Overlay of selected regions of the ^{1}H-^{15}N HSQC spectra of 0.03 mM C-BP-2 in the absence (orange) and presence (green) of 3.0 mM SOS (a, c and d), or 2.0 mM low molecular weight heparin (b), at pH 6.0 (a and b), pH 5.5 (c) and pH 7.4 (d). Spectra were recorded at 500 MHz and 298 K.

the resonance frequency differences between free and bound C-BP-2 protein. Addition of low molecular weight heparin to ^{15}N-C-BP-2 to saturation caused movements of a similar set of peaks together with significant peak broadening (Figures 5.3 and 5.4). The residues that exhibited relatively larger ^{1}H and ^{15}N chemical shift changes upon SOS or heparin binding (Figure 5.5) were mapped to the surface of C-BP-2 (Figure 5.6 (a and b)). These residues are mainly located at the β-turn connecting the first and second strands, part of the third strand, and the beginning of the C-terminal tail. Two lysine residues (Lys227 and Lys234) and two histidine residues (His228 and His271) are within this binding site.

5.3.2 Effects of pH on the binding

To check whether a change of pH and thus the protonation states of the histidines would affect binding, ^{1}H-^{15}N HSQC spectra of ^{15}N-C-BP-2 with 0 and 3.0 mM SOS, or with 0 and 2.0 mM low molecular weight heparin at pH 5.5 and pH 7.4 were compared. At pH 5.5, SOS induced larger chemical shift changes of a similar set of peaks, and residues in the loop III and the C-terminal tail, which are close to the binding site and are relatively flexible, were also perturbed (Figures 5.3 and 5.4), whereas low molecular weight heparin caused a drastic broadening of the peaks (not shown). Furthermore, in titration experiments at pH 5.5, less SOS was required to saturate the protein compared to pH 6.0. These results indicate that binding was enhanced at pH 5.5. At pH 7.4, chemical shift differences were very small and the line-widths were restored (Figures 5.3 and 5.4), indicating that the binding was largely suppressed at neutral pH. Although pH titration alone of ^{15}N-C-BP-2 without ligands can also cause certain chemical shift changes of residues in this binding site because of partial deprotonation of the imidazolium rings of His271 and His228 (Chapter 4), the displacement of the side-chain ^{15}N resonance of Asn232 (Figures 5.3 and 5.4) occurred only when SOS or heparin were added, not by altering the pH (5.0-7.0) alone (Figures 4.2 and 5.3). Moreover, the pH of the samples was maintained very carefully, and duplicate experiments produced identical results. Thus, we conclude that the chemical shift changes seen in the titrations are due to the binding of SOS or heparin.

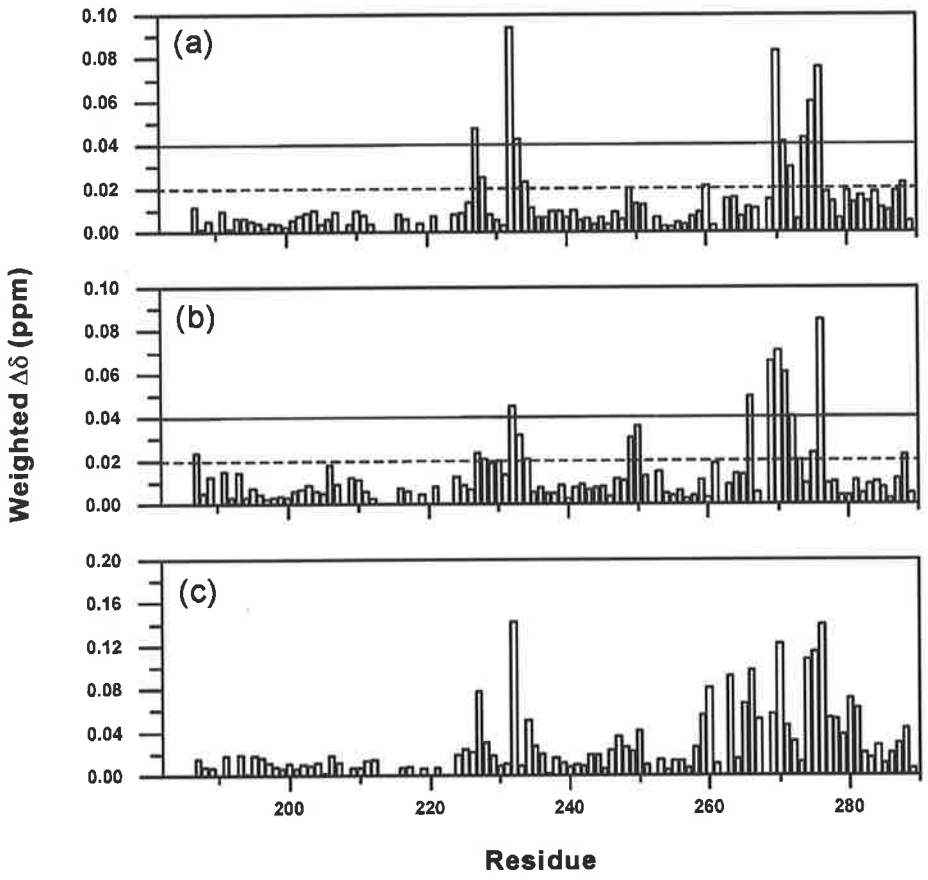


Figure 5.5 Chemical shift changes induced by SOS or LMW heparin

Weighted chemical shift changes of C-BP-2 induced by SOS (a and c) or low molecular weight heparin (b) binding, at pH 6.0 (a and b) or pH 5.5 (c), are plotted as a function of residue number. Weighted backbone amide chemical shift changes Δδ (ppm) were calculated using the chemical shift differences derived from the spectra of free and bound C-BP-2 according to the equation $\Delta\delta=[((\Delta\delta 1HN)2+(0.2\Delta\delta 15N)2)/2]^{1/2}$. The backbone amide cross-peak of Asn232 is seriously overlapped, so the weighted side-chain amide chemical shift changes are used for this residue.

5.3.3 The heparin binding site on C-BP-2

Based on these results and an analysis of the sequence alignments (see below), C-BP-2 residues Lys227, His228, Asn232, Leu233, Lys234 and His271 are proposed to be the primary heparin binding residues, and adjacent residues 229-231, 249-251, 269-270 and 272-276 are also involved. A schematic representation of the heparin binding by C-BP-2 is shown in (Figure 5.6 (c and d)).

5.4 Discussion

5.4.1 Heparin binding by IGFBP-2 C-domain

In this study, NMR chemical shift perturbation experiments showed clearly that the C-domain of IGFBP-2 binds heparin and the strength of the binding appeared to be modulated by pH. These new findings are of significant interest, because previous data in the literature regarding the heparin binding ability of IGFBP-2 have been rather contradictory.

Arai and co-workers reported that IGFBP-2 did not bind heparin but acquired heparin binding in the presence of IGFs, in contrast to IGFBP-1, -3, -4 or -5, which did not have increased heparin binding in the presence of IGFs (Arai et al., 1996b). The binding seemed to be physiologically significant, since IGFBP-2 also bound to human fibroblast ECM if IGF-I or IGF-II was added (Arai et al., 1996b). IGF-II was more potent than IGF-I, and des(1-3)-IGF-I or insulin had no effect, suggesting that binding of IGFs to IGFBP-2 is required to induce a conformational change of IGFBP-2 that exposes the heparin binding site (Arai et al., 1996b). Similarly, IGF-enhanced binding of IGFBP-2 to ECM was reported by Conover and Khosla (2003); the binding was inhibited by heparin, heparan sulfate, but not chondroitin sulfate, and ECM-associated IGFBP-2 had unchanged affinity for IGF-I and -II compared with soluble IGFBP-2 (Conover and Khosla, 2003). In the presence of IGF-II, IGFBP-2 binds heparin, but not collagen I, fibronectin, or laminin (Conover and Khosla, 2003).

In contrast, Russo and co-workers found that IGFBP-2 binds to glycosaminoglycans including heparin, keratan sulfate, chondroitin-4-sulfate, and

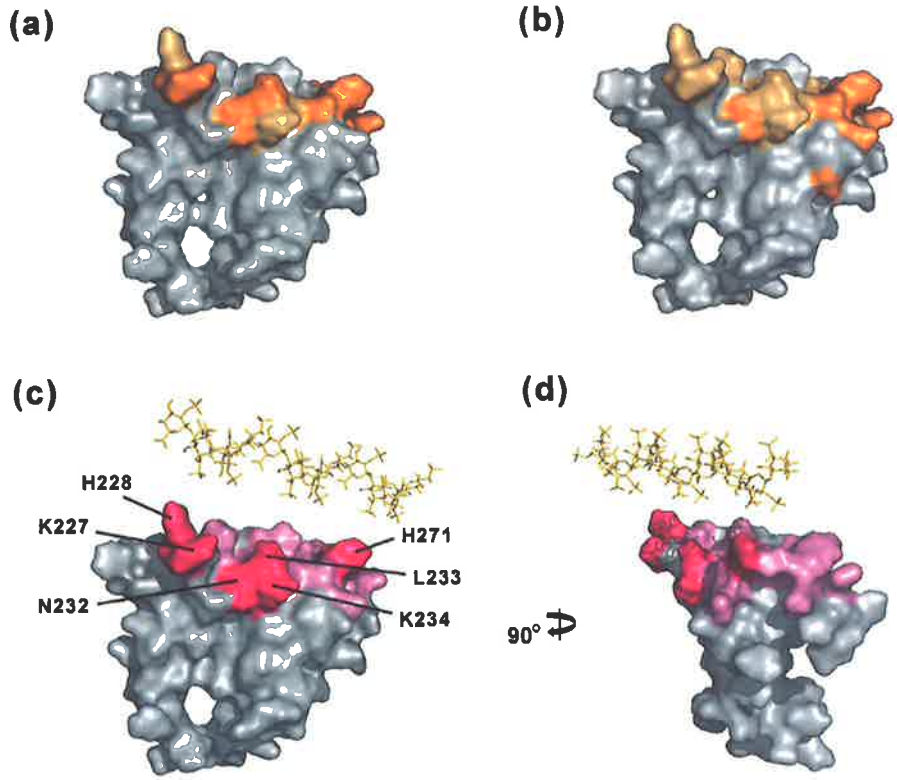


Figure 5.6 The heparin binding site on C-BP-2 identified by NMR

Residues with $\Delta\delta>0.04$ ppm (orange) and $0.02<\Delta\delta<0.04$ (light orange) upon SOS (a) or low molecular weight heparin (b) binding at pH 6.0 are shown. (c) Schematic representation of the proposed heparin binding site. Primary heparin binding residues are coloured pink and labelled; adjacent residues that also exhibited significant chemical shift changes are coloured light pink. A 12-mer heparin model (PDB 1HPN) (Mulloy et al., 1993) is shown as a stick model (gold). (d) Structures in (c), rotated by 90° about the vertical axis. Parts of the unstructured regions (residues 183-186 and 280-289) are excluded for clarity. Figures have the same orientation as Figure 4.6 (b). These figures were prepared using PyMOL (DeLano Scientific).

chondroitin-6-sulfate, and proteoglycan aggrecan, in the absence of IGFs; binding of IGFBP-2 to chondrointin-6-sulfate decreased the IGF-I binding affinity of IGFBP-2 approximately three-fold (Russo et al., 1997). The binding appeared to be biologically relevant, since IGFBP-2 binds to cell surface in the rat brain olfactory bulb (Russo et al., 1997), and IGFBP-2 and a 14-kDa proteolytic fragment bind to the neuroblastoma cell surface (Russo et al., 1999). Recently, these authors proposed that the heparin binding site of IGFBP-2 may be located in the linker-domain, because IGFBP-2 mutant of the predicted heparin binding site in the linker domain had significant reduction (~70-80%) in binding to ECM components including heparin, vitronectin, laminin, fibronectin, and collagen IV (Russo et al., 2005). Furthermore, IGFBP-2 enhanced proliferation and metastasis of neuroblastoma cells and this function was dependent on the proposed linker domain heparin binding site (Russo et al., 2005).

Nevertheless, the absence of heparin binding by IGFBP-2 either with or without IGFs has also been described. Song and co-workers reported that neither IGFBP-2 nor an chimeric IGFBP containing the N-domain of IGFBP-5 and the linker domain and C-domain of IGFBP-2 bind heparin (Song et al., 2001). In a recent paper, these authors used IGFBP-2 as a negative control in the study of IGF and heparin binding of IGFBP-3 and -5; no binding of IGFBP-2 to heparin was seen in heparin ligand blot or BIAcore analysis, either in the absence or presence of IGF-I (Beattie et al., 2005).

Recently, the C-domain of IGFBP-4 (C-BP-4), which, like C-BP-2, lacks the basic residue cluster, was unexpectedly found to bind heparin and acquire IGF binding ability upon heparin binding (Fernandez-Tornero et al., 2005). The results presented here may provide a plausible explanation for their observations (see below).

The heparin binding by full-length IGFBP-2 was not studied in this thesis. However, the heparin binding of the C-domain likely contributes to the heparin binding by the full-length protein. A previously proposed heparin binding sequence (^{179}PKKLRP, in the linker domain) was not included in the C-BP-2 construct, and its significance in heparin binding is to be analyzed in future. Furthermore, the effects of heparin binding by C-BP-2 on its IGF binding are investigated in Chapter 7.

5.4.2 The heparin binding site on C-BP-2

In this Chapter, the heparin binding site of C-BP-2 was mapped onto its three-dimensional structure using the chemical shift perturbation method. Previously, the heparin binding site on IGFBP-2 C-domain was unknown, although mutagenesis studies have been used to probe the heparin and ECM binding sites on the C-domain of IGFBP-5 (Arai et al., 1996b; Parker et al., 1998; Song et al., 2000; Allan et al., 2006). Those studies have targeted the 18-residue region that contains 10 basic residues. While several basic residues were found to be important for heparin binding, there were discrepancies among the results, and the contributions of other basic residues in the region as well as the three-dimensional location of the binding site remain unclear.

Arai and co-workers found that IGFBP-5 residues Arg201, Lys202, Lys206, Lys208, and Arg214 (equivalent to IGFBP-2 Lys227, His228, Asn232, Lys234, and Leu240, respectively) are important for heparin binding, whereas Lys211, Lys217, Arg218, and probably also Arg207 are not (Arai et al., 1996b). While the Arg201Leu/Lys202Glu/Arg214Ala mutant had reduced heparin binding (Song et al., 2000), mutation of Arg214 to Ala alone only slightly decreased heparin binding, and Lys202Ala/Arg208Ala mutant had a very sharp decrease in binding and the net contribution of other basic residues was unclear (Allan et al., 2006). Interestingly, a heparin binding site (Arg136 and Arg137) in the linker domain of IGFBP-5 that is active only in the absence of the C-domain has also been proposed (Song et al., 2001).

Parker and co-workers reported that two basic regions of IGFBP-5, Ala131-Thr141 and Arg201-Arg218, contribute to fibroblast ECM binding, and the 10 basic residue region (Arg201-Arg218) has a greater contribution than the Ala131-Thr141 region that is located in the linker domain. The Lys202Asn/Lys206Asn/Arg207Ala IGFBP-5 mutant had a 79% reduction and multiple substitutions (Lys211Asn/Arg214Ala/Lys217Ala/Arg218Asn) resulted in an 88% decrease in ECM binding (Parker et al., 1996). These authors subsequently showed that Arg207 and Arg214 are the most important for ECM binding, whereas alterations of Lys211 and Arg218 only had minimal effects (Parker et al., 1998).

The heparin binding site of C-BP-2 identified in this thesis is a patch containing the β-turn connecting the first and second strands, part of the third strand, and the beginning of the C-terminal tail (Figure 5.6). Lys227, His228, Asn232, Leu233, Lys234

and His271 are proposed to be the primary heparin binding residues, and the positive charges of His228 and His271 seem to be required for the interaction. The sequence alignments (Figure 4.10 (e)) reveal interesting residue substitution patterns among IGFBPs within equivalent regions of C-BP-2 Lys227-Lys234. Both IGFBP-3 and -5 have five basic residues, BBGXYBBB (B: basic residues; X: non-basic residues), while other IGFBPs tend to have basic residues at the first and fifth B position, and histidine and asparagine are frequent substitutions at other B positions. It is possible that glycosaminoglycan binding to this site is a common function shared by IGFBP C-domains. While occupation of all B positions by basic residues seems to provide strong heparin binding ability, substitutions by histidines may confer weaker pH-dependent binding. The structural requirement of the glycosaminoglycans for the binding was not explicitly determined in this study. Whether the distinct residue substitutions in the binding patches among IGFBPs also lead to different glycosaminoglycan preferences is an interesting question that requires future investigation. Another histidine, His271, is distant from these residues in the primary sequence, but is located within the heparin binding site, a fact that could not be identified previously due to the lack of a three-dimensional structure. Histidines are also found at equivalent positions in IGFBP-4 and 5, and mouse and rat IGFBP-1, but not in several non-IGFBP thyroglobulin type 1 domains (Figure 4.10 (e)). Because C-BP-4 has a very similar patch including two basic residues, two histidines and one asparagine, it is likely that its heparin binding at pH 6.0, observed recently using analytical ultracentrifugation (Fernandez-Tornero et al., 2005) has a similar binding site to that found here for C-BP-2 by NMR.

It will be of great interest to study the heparin binding by other IGFBP C-domains using NMR, especially the heparin binding by C-BP-3 and -5, which have the greatest number (ten) of basic residues in the 18-residue region compared with other C-BPs, and that of IGFBP-6, which has more basic residues than C-BP-2 but fewer than C-BP-3, and -5. A comparison with the results presented here for C-BP-2 will reveal whether the sequence differences, especially the basic residues in the loop II region in those C-domains lead to different heparin binding properties and different orientations of the heparin binding site on these domains.

5.4.3 Implication for the role of C-BP-2 in cancer

IGFBP C-domain fragments are biologically active in their own right. For example, C-BP-1 stimulated cultured fibroblast migration (Sala et al., 2005), whereas C-BP-4 promoted bone formation in calvarial cultures (Fernandez-Tornero et al., 2005). This is expected to be the case for the IGFBP-2 C-domain as well, and, while the RGD motif points to one such function, pH-dependent heparin binding may be another. Histidine-mediated pH-dependent heparin binding and its biological significance have been demonstrated for granulocyte-macrophage colony-stimulating factor (Sebollela et al., 2005) and mast cell tryptase (Hallgren et al., 2004). Several physiological and disease states are known to have slightly acidic extracellular microenvironment, for example cancer and inflammation. Tumours have acidic extracellular matrix pH values, resulting from a constitutive increase in glycolysis (Gatenby and Gillies, 2004). When IGFBP-2 is over-expressed in some tumours, its C-domain might preferentially bind to glycosaminoglycans in acidic tumour ECM, but not in neutral normal tissue ECM. IGFBP-2 and/or IGFBP-2-IGF complexes and/or C-domain fragments may therefore be concentrated in tumours and modulate IGF release from IGFBP-2-IGF complexes. Subsequently, the accumulated IGFs and the RGD-containing IGFBP-2 C-domain, in the full-length protein or fragment forms, could exert their actions via IGF receptors and integrins, resulting in cancer cell proliferation and migration.

Chapter 6

NMR Spectroscopy of IGFs and IGFBP-2 N-terminal Domain

6.1 Introduction

In previous Chapters, experimental conditions for NMR studies of C-BP-2 have been optimised, and heparin binding by C-BP-2 was studied by monitoring its ^{1}H-^{15}N HSQC spectra while titrating SOS or low molecular weight heparin into the sample. In the next Chapter, molecular interactions among C-BP-2, N-BP-2, and IGFs will be described in detail using the same chemical shift mapping approach, and titration experiments performed using different orientations. Thus, ^{15}N-labelled C-BP-2, IGF-I, IGF-II, and N-BP-2 were monitored in separate experiments, while other unlabelled interaction partner(s) were titrated into the samples. This Chapter is devoted to the necessary optimisation of the experimental conditions for good quality ^{1}H-^{15}N HSQC spectra of IGF-I, IGF-II, and N-BP-2, before carrying out those titration experiments.

Both IGF-I (Cooke et al., 1991; Sato et al., 1993; Jansson et al., 1998; Schaffer et al., 2003; Carrick et al., 2005) and IGF-II (Terasawa et al., 1994; Torres et al., 1995; Headey et al., 2004b; Carrick et al., 2005) were known to give poor NMR spectra, containing many broad peaks. This is thought to arise from a combination of aggregation at the protein concentrations required for NMR experiments, together with the internal dynamics of the molecules. These previous reports also suggested that the peak linewidths are dependent on temperature, pH, and protein concentration. The recent use of cryoprobes for signal detection in NMR spectrometers has significantly increased the sensitivity, such that acquiring 2D ^{1}H-^{15}N HSQC spectra at protein (^{15}N-labelled) concentrations down to 0.05-0.1 mM has become routinely practical, whereas without a cryoprobe it was necessary to have 0.2-1.0 mM protein. It was of interest to test whether lowering protein concentration would improve the spectral quality of IGFs. Thus, effects of protein concentration, pH, and temperature on the ^{1}H-^{15}N HSQC spectra of IGF-I, IGF-II, and N-BP-2 were investigated in this Chapter.

In addition, although the cross-peaks in the ^{1}H-^{15}N HSQC spectra of IGF-I were assigned by Dr Francine Carrick (The University of Adelaide) and Ms Christina Wang (Walter and Eliza Hall Institute of Medical Research) (Carrick et al., 2005), significant difficulties were encountered, and thus there was a considerable number of missing assignments. Here, 3D spectra of $^{15}N/^{13}C$-IGF-I under the optimised experimental conditions were used to achieve a more complete set of assignments for the IGF-I ^{1}H-^{15}N HSQC spectra.

Having identified the heparin binding property of C-BP-2 using NMR, it was interest to determine whether N-BP-2 also bound heparin. Many growth factors and chemokines are heparin-binding proteins, and these heparin-protein interactions modulate cell growth, differentiation, and inflammation (Capila and Linhardt, 2002). Heparin binding by IGF-I was not detected in one study (Fernandez-Tornero et al., 2005). In this Chapter, preliminary NMR chemical shift perturbation experiments were performed to test the possibility of heparin binding by N-BP-2 or IGF-I. ^{1}H-^{15}N HSQC spectra in the absence and presence of heparin were also required for future interpretation of the effects of heparin on the binary and ternary complexes of IGF-I, N-BP-2, and C-BP-2.

6.2 Materials and Methods

6.2.1 Materials

$^{15}N/^{2}H$-labelled IGF-I and $^{15}N/^{2}H$-labelled IGF-II were kindly provided by Dr Francine Carrick (The University of Adelaide) (Carrick et al., 2005). The percentage of ^{2}H incorporation in these samples was about 50 %. $^{15}N/^{13}C$-labelled IGF-I was kindly prepared by Ms Kerrie McNeil (The University of Adelaide). These isotope-labelled IGF-I and IGF-II samples were prepared as described previously (King et al., 1992). ^{15}N-labelled $^{1\text{-}138}$IGFBP-2 (N-BP-2) was prepared by the author as described in Chapter 3.

6.2.2 NMR experiments and assignments for IGF-I

^{1}H-^{15}N HSQC spectra of a 0.07 mM $^{15}N/^{2}H$-IGF-I sample in 95% H_2O/5% $2H_2O$ containing 10 mM sodium acetate and 0.02% (w/v) sodium azide were acquired at 298 K and pH 4.0, 5.0, 6.0 and 7.0. The spectra were recorded with a data matrix size of 2048×128 and 128 scans per t_1 increment on a Bruker Avance 500 spectrometer equipped with a cryoprobe. The spectral widths were 12.0 ppm for ^{1}H and 35.0 ppm for ^{15}N. Carrier frequencies were 4.7 ppm for ^{1}H and 117.5 ppm for ^{15}N. The spectra were processed as described in Chapter 3.

^{1}H-^{15}N HSQC spectra of a 0.7 mM sample of $^{15}N/^{13}C$-IGF-I were acquired at pH 6.0 and at 298 K, 303 K, 308 K, 313 K, and 315 K at 500 MHz. The buffer condition was the same as above. Four scans per t_1 increment were used and ^{13}C decoupling was applied. HNCA and CBCA(CO)NH spectra were acquired on this sample at 298 K and 315 K. HNCA experiments were recorded with a data matrix size of 2048×44×64 and 32 scans per t_1 increment. The spectral widths were 11.5 ppm for ^{1}H, 20.0 ppm for ^{15}N, and 35.0 ppm for ^{13}C. Carrier frequencies were 4.7 ppm for ^{1}H, 118.0 ppm for ^{15}N, and 53.0 ppm for ^{13}C. CBCA(CO)NH experiments were recorded with a data matrix size of 2048×44×44 and 64 scans per t_1 increment. The spectral widths were 11.5 ppm for ^{1}H, 20.0 ppm for ^{15}N, and 60.0 ppm for ^{13}C. Carrier frequencies were 4.7 ppm for ^{1}H, 118.0 ppm for ^{15}N, and 44.0 ppm for ^{13}C. H^N, N, C^α, and C^β chemical shifts of IGF-I at 298 K and 315 K were obtained using the HNCA and CBCA(CO)NH spectra as described in Chapter 4, and the ^{1}H-^{15}N HSQC spectra were assigned.

^{1}H-^{15}N HSQC spectra of 0.04 mM $^{15}N/^{13}C$-IGF-I in the absence and presence of 2.0 mM low molecular weight heparin were recorded, in 95% H_2O/5% $^{2}H_2O$ containing 10 mM sodium acetate, 150 mM NaCl, and 0.02% (w/v) sodium azide at pH 6.0 and 298 K. The number of scans per t_1 increment was 128.

6.2.3 NMR experiments on IGF-II

^{1}H-^{15}N HSQC spectra of a 0.05 mM sample of $^{15}N/^{2}H$-IGF-II in 95% H_2O/5% $2H_2O$ containing 10 mM sodium acetate and 0.02% (w/v) sodium azide were acquired at 298 K and pH 4.0 and 5.0. 1D ^{1}H spectra were also recorded at pH 4.0 and at 293 K, 303 K, and 308 K, and a ^{1}H-^{15}N HSQC spectrum was also acquired at pH 4.0 and 308 K. ^{1}H-^{15}N HSQC spectrum was also recorded in the presence of 50 mM arginine plus 50 mM glutamate (Golovanov et al., 2004) at pH 4.0 and 298 K. Spectra were recorded

with a data matrix size of 2048×128 and 128 scans per t_1 increment on a Bruker Avance 500 spectrometer equipped with a cryoprobe. The spectral widths were 12.0 ppm for ^{1}H and 35.0 ppm for ^{15}N. Carrier frequencies were 4.7 ppm for ^{1}H and 117.5 ppm for ^{15}N.

6.2.4 NMR experiments on N-BP-2

^{1}H-^{15}N HSQC spectra of a 0.07 mM sample of ^{15}N-N-BP-2 in 95% H_2O/5% $2H_2O$ containing 10 mM sodium acetate and 0.02% (w/v) sodium azide were acquired at 298 K and pH 7.0. The spectra were recorded with a data matrix size of 2048×128 and 128 scans per t_1 increment on a Bruker Avance 500 spectrometer equipped with a cryoprobe. The spectral widths were 12.0 ppm for ^{1}H and 35.0 ppm for ^{15}N. Carrier frequencies were 4.7 ppm for ^{1}H and 118.0 ppm for ^{15}N.

^{1}H-^{15}N HSQC spectra of a 0.7 mM sample of $^{13}C/^{15}N$-IGF-I in 95% H_2O/5% $^{2}H_2O$ containing 10 mM sodium acetate, 150 mM NaCl and 0.02% (w/v) sodium azide at pH 6.0 and 298 K, in the absence and presence of 2.0 mM low molecular weight heparin were recorded and compared. The number of scans per t_1 increment was 128.

6.3 Results

6.3.1 NMR spectra of IGF-I

^{1}H-^{15}N HSQC spectra of 0.07 mM $^{15}N/^{2}H$-IGF-I at 298 K and pH 4.0, 5.0, 6.0 and 7.0 are shown in Figure 6.1. Compared to the spectra previously recorded by Carrick and co-workers on a 0.5 mM sample (Carrick et al., 2005), lowering the protein concentration appeared to narrow the peak linewidths. In addition, it can be seen that several peaks were visible at pH 6.0 that were too broad to see at pH 4.0, whereas at pH 7.0 some peaks were weaker probably due to more rapid chemical exchange with H_2O. Peaks were of a more uniform linewidth at pH 6.0 than pH 4.0.

6.3.2 Assignments of the ^{1}H-^{15}N HSQC spectra of IGF-I

Backbone ^{1}H-^{15}N correlation peaks on the HSQC spectrum were assigned using 3D HNCA and CBCA(CO)NH spectra acquired on a 0.7 mM $^{13}C/^{15}N$-IGF-I sample at pH 6.0. ^{1}H-^{15}N HSQC spectra of this sample at different temperatures are shown in Figure 6.2. Interestingly, at 298 K, several weak peaks that were present in the ^{1}H-^{15}N

(Please turn to page 168)

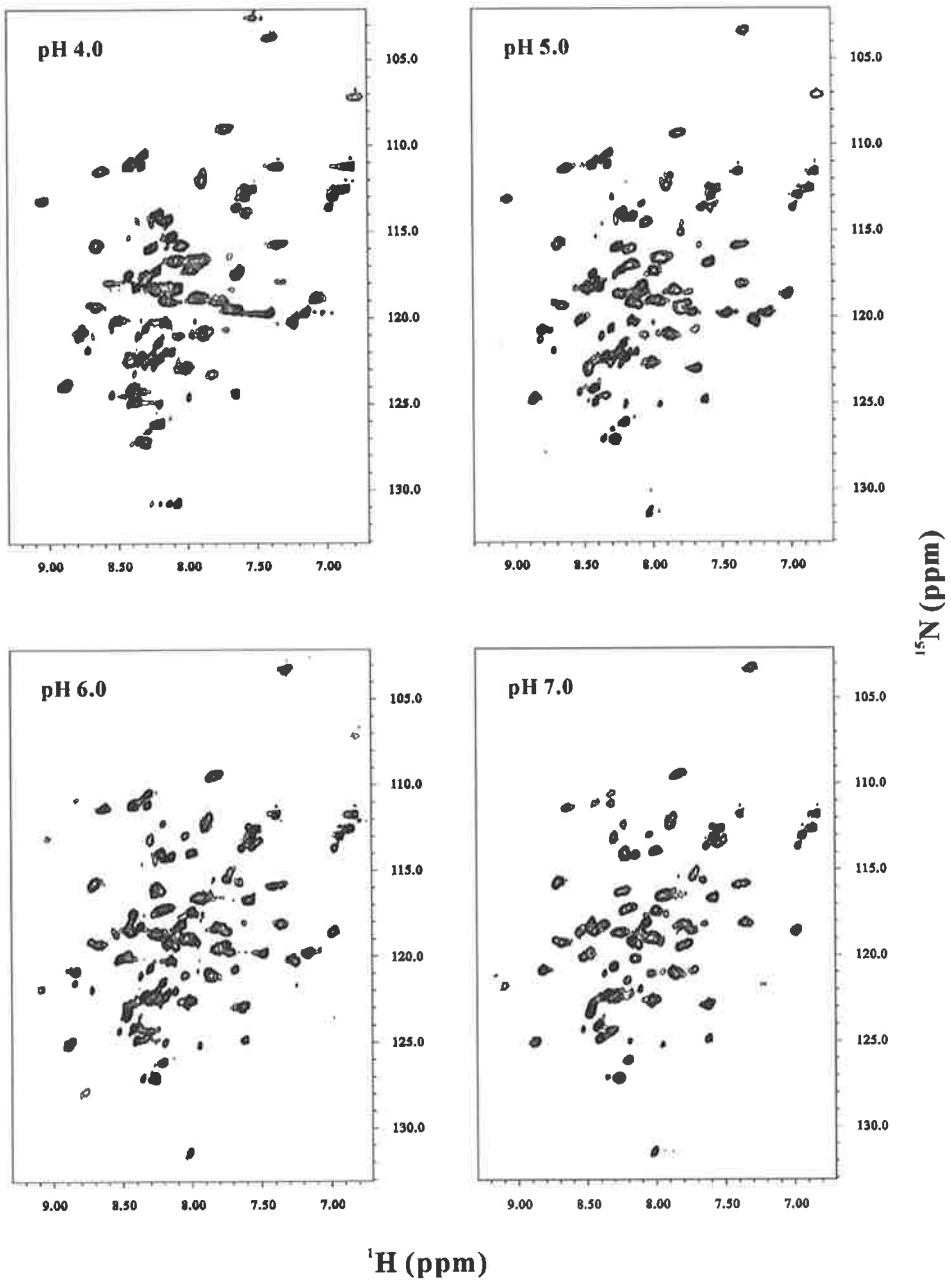


Figure 6.1 ^{1}H-^{15}N HSQC spectra of IGF-I at different pH values

The spectra were acquired using a 0.07 mM ^{15}N/^{2}H-labelled IGF-II sample at 500 MHz and 298 K.

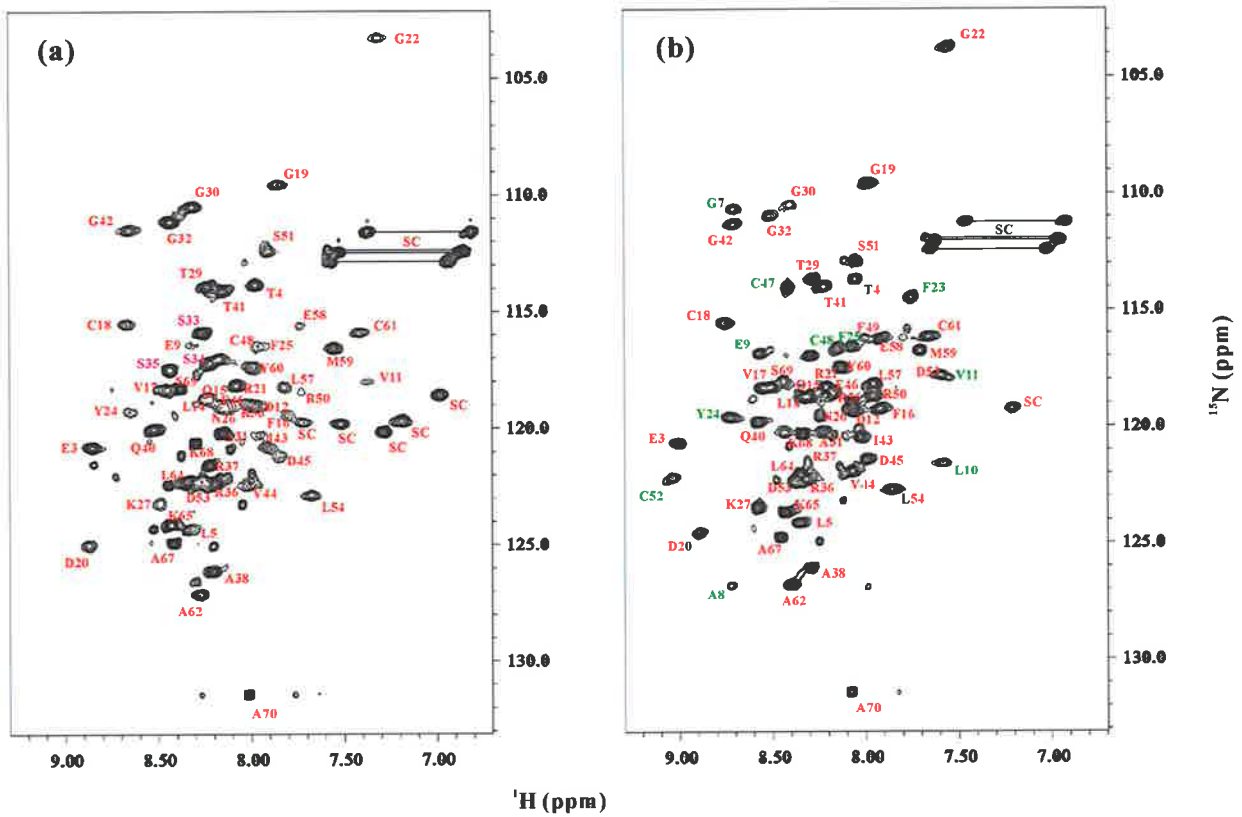


Figure 6.2 Assignments of the ^{1}H-^{15}N HSQC spectra of IGF-I at different temperatures

^{1}H-^{15}N HSQC spectrum of 0.7 mM ^{13}C/^{15}N-labelled IGF-I at pH 6.0 and 298 K (a) or 315 K (b). Spectra were recorded at 500 MHz. Assignments for backbone amide HN-N cross-peaks are shown. Peaks with significantly narrowed linewidth and/or were more intense at 315 K compared to 298 K are labelled green in (b).

HSQC spectrum of the 0.07 mM $^{15}N/^{2}H$-IGF-I sample (Figure 6.1) were invisible or very weak in this spectrum (Figure 6.2 (a)). Corresponding peaks were also very broad and weak in the HNCA and CBCA(CO)NH spectra at 298 K, and assignments of these peaks were therefore difficult to obtain. Nevertheless, these peaks were much sharper in the ^{1}H-^{15}N HSQC spectra at 315 K (Figure 6.2 (b)), and their assignments were obtained using the HNCA and CBCA(CO)NH spectra at 315 K. Assignments of the 315 K spectrum were then used to confirm or facilitate the assignments of the 298 K spectrum. In fact, backbone amide cross-peaks of Gly7, Ala8, Leu10, Val11, Phe23, Tyr24, Phe25, Cys47, and Cys48 had significantly improved signal to noise ratio at 315 K compared with 298 K, while several other resonances, in particular Ser33, Ser34, and Ser35, were weaker or even disappeared at 315 K (Figure 6.2 (b)) compared with 298 K. In general, peaks were of a more uniform linewidth at 315 K compared to 298 K. Backbone H^N and N assignments were obtained for 61 out of 64 residues (excluding the first residue and five prolines). The assignments are summarized in Table 6.1.

6.3.3 Spectra of IGF-I in the absence and presence of heparin

An overlay of the ^{1}H-^{15}N HSQC spectra of 0.04 mM $^{13}C/^{15}N$-IGF-I in the absence and presence of 2.0 mM low molecular weight heparin, in buffer containing 150 mM NaCl at pH 6.0, is shown in Figure 6.3 (a). Interestingly, while addition of heparin caused significant chemical shift changes of the backbone amide cross-peaks for Arg37 and Ala38, perturbations of other cross-peaks were very small. Although the backbone amide cross-peak of Arg36 did not change significantly, it is possible that the changes of Arg37 and Ala38 cross-peaks were induced by the binding of Arg36 and Arg37 to heparin. This is because engagement of the sidechain in heparin binding may disrupt its original packing with subsequent residues, resulting in larger chemical shift changes to the backbone amide of subsequent residues than to the backbone amide of its own. A similar phenomenon was also evident in C-BP-2 heparin binding experiments (Chapter 5). Arg36, Arg37, and Ala38 belong to the C domain of IGF-I and are located in a flexible loop (Figure 6.3 (b)). Whether the results represent a physiologically relevant heparin binding by IGF-I is an interesting question that requires future investigation (see below).

Table 6.1 Backbone H^N and N chemical shift assignments of IGF-I

Residue	298 K		315 K	
	H^N (ppm)	N (ppm)	H^N (ppm)	N (ppm)
Gly1				
Pro2				
Glu3	8.85	120.7	9.00	120.7
Thr4	7.97	113.8	8.05	113.7
Leu5	8.35	124.3	8.35	124.1
Cys6				
Gly7	8.70	110.7	8.70	110.7
Ala8	8.74	127.3	8.72	126.8
Glu9	8.32	116.4	8.56	116.8
Leu10			7.59	121.5
Val11	7.37	118.0	7.59	117.8
Asp12	7.99	119.0	8.06	119.3
Ala13				
Leu14	8.30	119.0	8.33	118.7
Gln15	8.22	118.7	8.30	118.7
Phe16	7.81	119.4	7.92	119.3
Val17	8.45	118.3	8.54	118.3
Cys18	8.67	115.5	8.76	115.6
Gly19	7.86	109.5	7.99	109.6
Asp20	8.87	124.9	8.89	124.5
Arg21	8.08	118.2	8.21	118.3
Gly22	7.32	103.2	7.56	103.7
Phe23	7.64	113.7	7.76	114.5
Tyr24	8.65	119.3	8.72	119.6
Phe25	7.91	116.4	8.07	116.6
Asn26	8.15	119.1	8.24	119.5
Lys27	8.48	123.1	8.58	123.5
Pro28				
Thr29	8.24	113.9	8.28	113.7
Gly30	8.31	110.5	8.40	110.5
Tyr31	8.15	120.2	8.23	120.2
Gly32	8.43	111.1	8.51	111.0
Ser33	8.26	115.8	8.31	115.8
Ser34	8.22	117.2		
Ser35	8.43	117.5		
Arg36	8.19	122.3	8.32	122.3
Arg37	8.22	121.5	8.31	121.5

Residue	298 K		315 K	
	H^N (ppm)	N (ppm)	H^N (ppm)	N (ppm)
Ala38	8. 20	126. 1	8. 30	126. 1
Pro39				
Gln40	8. 52	120. 0	8. 58	119. 8
Thr41	8. 15	114. 1	8. 22	114. 0
Gly42	8. 65	111. 4	8. 71	111. 4
Ile43	7. 91	120. 8	8. 02	120. 5
Val44	8. 01	122. 4	8. 07	121. 9
Asp45	7. 84	121. 2	7. 99	121. 4
Glu46	8. 20	118. 9	8. 18	118. 6
Cys47	8. 28	113. 1	8. 41	114. 1
Cys48	7. 97	116. 5	8. 15	116. 7
Phe49	7. 85	116. 1	7. 99	116. 2
Arg50	7. 73	118. 4	7. 96	118. 7
Ser51	7. 90	112. 4	8. 05	112. 9
Cys52	9. 07	121. 9	9. 03	122. 2
Asp53	8. 26	122. 2	8. 37	122. 4
Leu54	7. 68	122. 9	7. 86	122. 7
Arg55				
Arg56	8. 05	118. 9	8. 07	119. 0
Leu57	7. 82	118. 2	7. 96	118. 2
Glu58	7. 74	115. 5	7. 92	116. 3
Met59	7. 55	116. 6	7. 71	116. 8
Tyr60	8. 00	117. 4	8. 13	117. 5
Cys61	7. 42	115. 9	7. 67	116. 2
Ala62	8. 27	127. 0	8. 40	126. 8
Pro63				
Leu64	8. 34	122. 3	8. 35	122. 0
Lys65	8. 43	124. 1	8. 42	123. 7
Pro66				
Ala67	8. 42	124. 9	8. 46	124. 7
Lys68	8. 30	120. 5	8. 34	120. 3
Ser69	8. 37	118. 3	8. 44	118. 0
Ala70	8. 01	131. 4	8. 08	131. 4

Chemical shifts were of a 0.7 mM $^{15}N/^{13}C$-labelled IGF-I sample at pH 6.0.

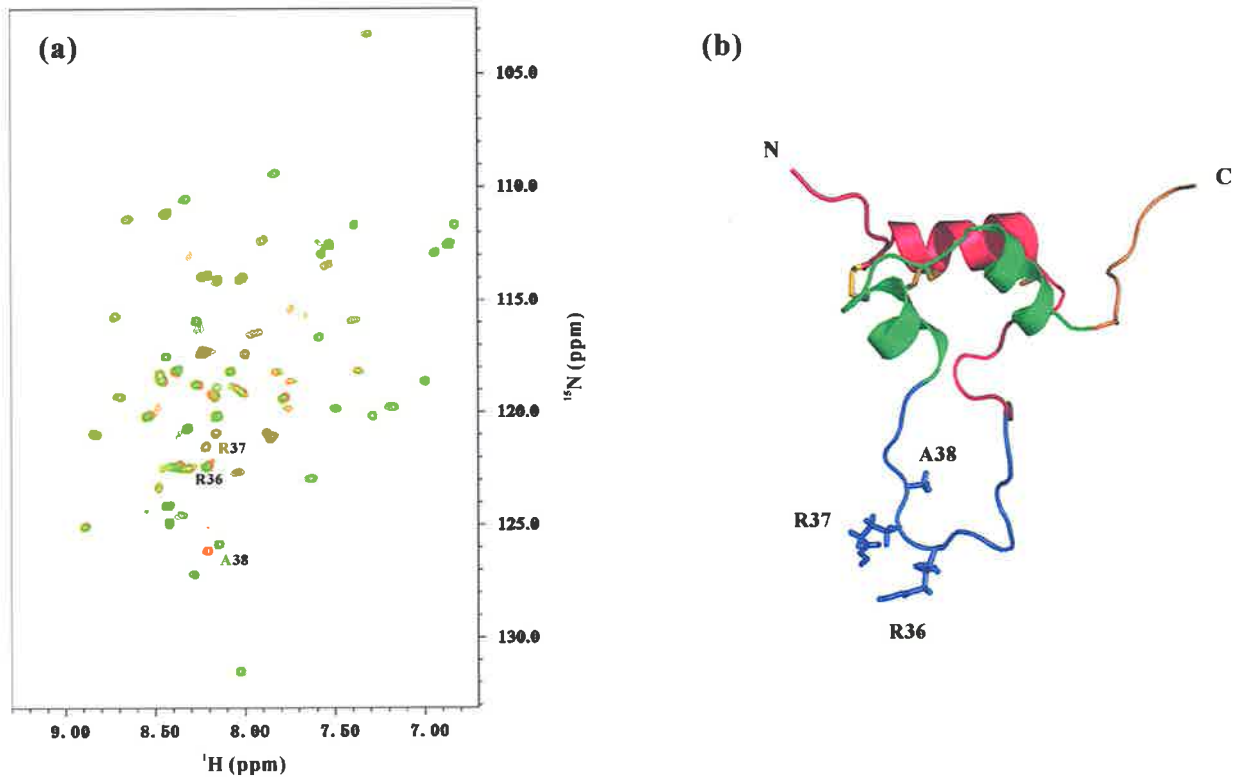


Figure 6.3 ^{1}H-^{15}N HSQC spectra of IGF-I in the absence and presence of heparin

(a) Overlay of the ^{1}H-^{15}N HSQC spectra of 0.04 mM ^{13}C/^{15}N-labelled IGF-I in the absence (orange) and presence (green) of 2.0 mM low molecular weight heparin. The buffer contained 150 mM NaCl and was at pH 6.0. Spectra were recorded at 500 MHz and 298 K. (b) Ribbon view of IGF-I (PDB 1PMX, (Schaffer et al., 2003)) showing the sidechains of Arg36, Arg37, and Ala38.

6.3.4 NMR spectra of IGF-II

The ^{1}H-^{15}N HSQC spectrum of a 0.05 mM $^{15}N/^{2}H$-IGF-II sample at 298 K and pH 4.0 is shown in Figure 6.4. It can be seen that the cross-peaks had very broad linewidths and there was serious overlap. Compared to the spectra previously recorded by Headey and co-workers on a 0.6 mM sample (Headey et al., 2004b), the linewidths were narrower at lower protein concentration but the differences were relatively small.

The peak broadening was more serious at pH 5.0 (not shown), and when the pH was further increased (~5.5), there was reversible precipitation in the sample despite the relatively low concentration (for NMR studies). Unlike the $^{15}N/^{13}C$-IGF-I sample, increasing the temperature to 308 K did not significantly narrow the linewidths. Furthermore, addition of 50 mM arginine plus 50 mM glutamate (Golovanov et al., 2004) did not improve the spectrum. Assignments of IGF-II spectra have been published by others (Torres et al., 1995).

6.3.5 NMR spectra of N-BP-2

During preparation of the N-BP-2 fragment (Chapter 3), it was found that this fragment was prone to precipitate at buffer pH of about 4.5-5.5, which was close to its theoretical pI (4.8). Therefore, ^{1}H-^{15}N HSQC spectra of N-BP-2 were recorded at pH 7.0 (Figure 6.5 (a)) to assist solubility and stability while avoiding chemical exchange with H_2O. The N-BP-2 fragment contained 140 residues. It can be seen that, while some cross-peaks were well dispersed, indicating the corresponding residues were in structured regions of the protein, there was a considerable number of cross-peaks clustered within the random coil region in the centre of the spectrum, i.e., ^{1}H ~8.0-8.5 ppm. The residues giving rise to those resonances are presumably located in disordered regions of the protein.

^{1}H-^{15}N HSQC spectra of N-BP-2 in the absence and presence of low molecular heparin were compared at pH 7.0 (Figure 6.5 (a)) and pH 6.0 (not shown). Unlike C-BP-2 and IGF-I, addition of heparin did not cause any significant chemical shift perturbation to N-BP-2 ^{1}H-^{15}N HSQC cross-peaks at either pH 7.0 or pH 6.0.

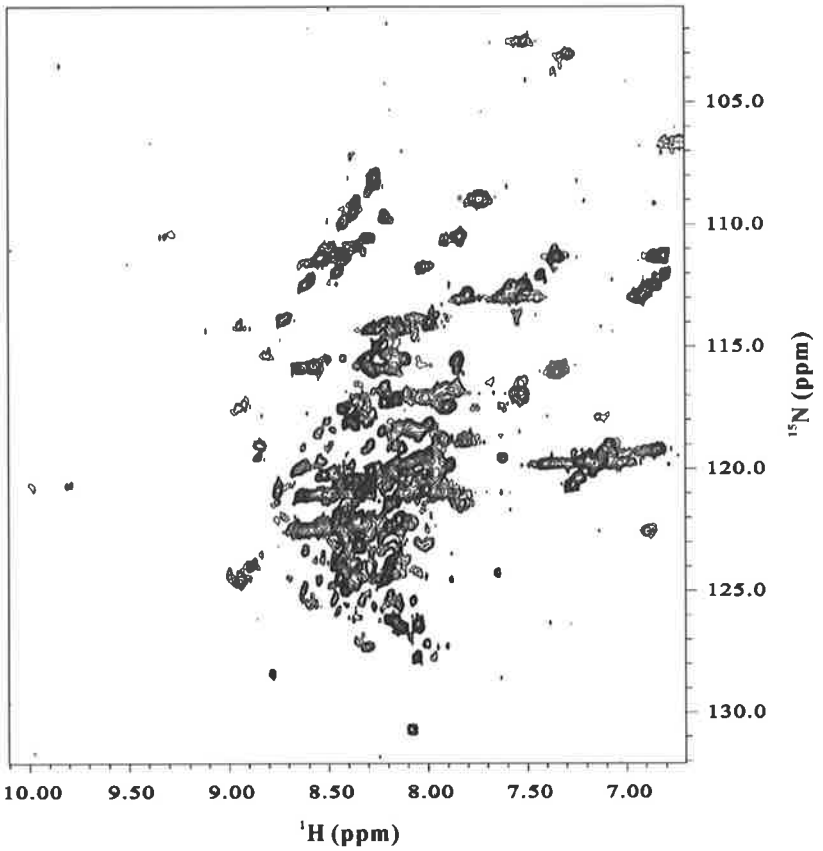


Figure 6.4 ^{1}H-^{15}N HSQC spectrum of IGF-II

^{1}H-^{15}N HSQC spectrum of 0.05 mM ^{15}N/^{2}H-labelled IGF-II at 298 K and pH 4.0, recorded at 500 MHz.

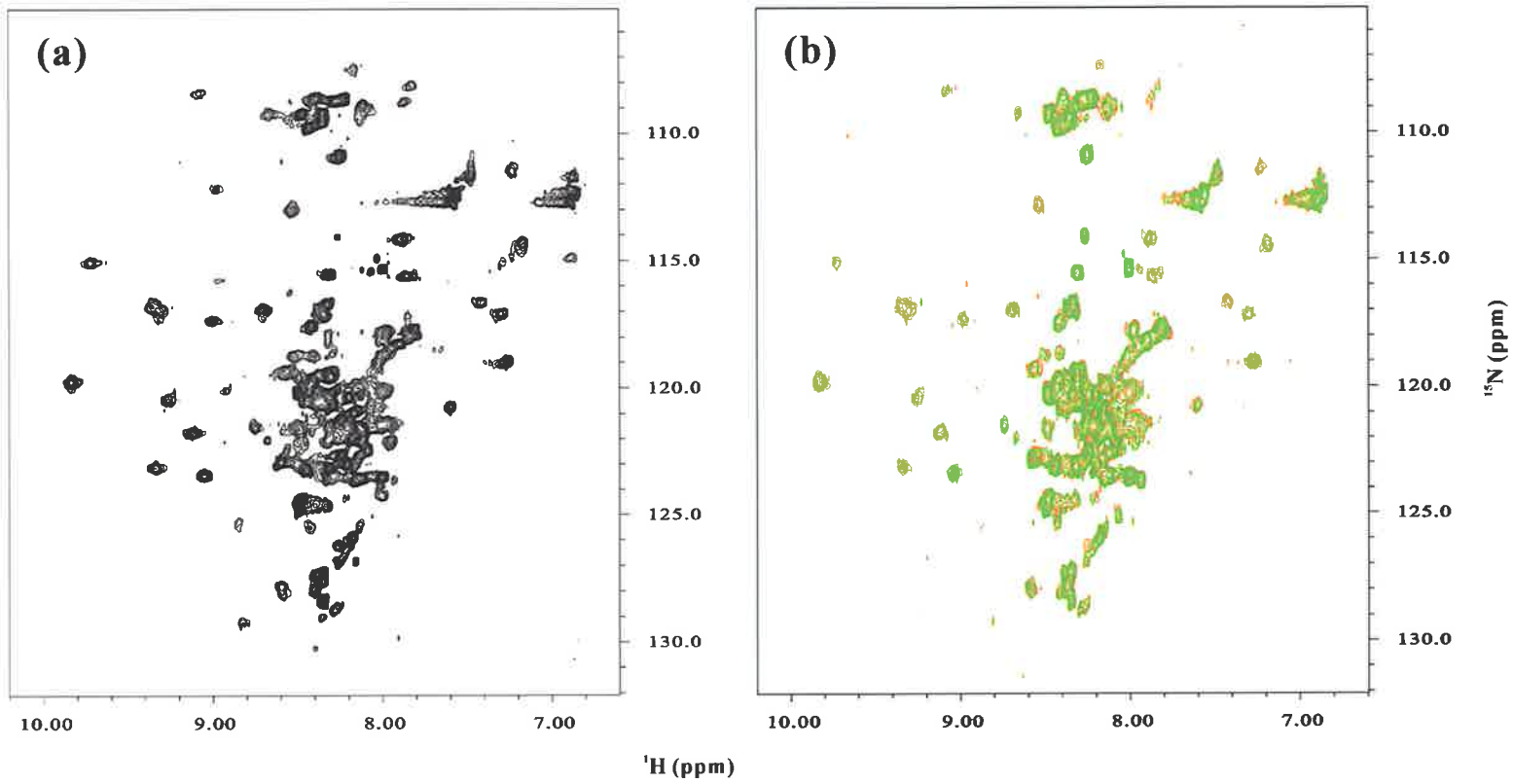


Figure 6.5 ^{1}H-^{15}N HSQC spectra of N-BP-2

(a) ^{1}H-^{15}N HSQC spectrum of 0.05 mM ^{15}N-labelled N-BP-2. (b) Overlay of the ^{1}H-^{15}N HSQC spectra of 0.05 mM ^{15}N-labelled N-BP-2 in the absence (orange) and presence (green) of 2.0 mM low molecular weight heparin, in a buffer containing 150 mM NaCl. Spectra were recorded at 298 K and pH 7.0 at 500 MHz.

6.4 Discussion

6.4.1 NMR spectroscopy of IGF-I and IGF-II at lower protein concentration

In this Chapter, experimental conditions for ^{1}H-^{15}N HSQC spectra of IGF-I, IGF-II, and N-BP-2 were established. These conditions were then used in titration experiments described in the following Chapter. Compared to the conditions used by others in analysing IGF-I (Jansson et al., 1998; Carrick et al., 2005) and IGF-II (Headey et al., 2004b; Carrick et al., 2005) interactions using NMR, current conditions have gained the advantage of a recent development in NMR spectrometry, i.e., the use of cryoprobes, and have resulted in improved spectral quality.

Before the use of a cryoprobe, to record a 2D ^{1}H-^{15}N HSQC spectrum in a reasonable time period often requires the protein concentration to as higher as 0.2-1.0 mM. As described in the Introduction, at these concentrations, both IGF-I and IGF-II appeared to aggregate and give poor NMR spectra. In previous studies, low pH (~3-3.5) conditions were often used to suppress the aggregation (Cooke et al., 1991; Sato et al., 1993; Terasawa et al., 1994; Torres et al., 1995; Jansson et al., 1998 ; Headey et al., 2004b). Although Carrick and co-workers found that the IGF·IGFBP-2 complex did not dissociate at low pH down to 3.5 (Carrick et al., 2005), it is desirable to study the interactions at a pH closer to the physiological pH. The use of a cryoprobe has provided an opportunity to acquire ^{1}H-^{15}N HSQC spectra in a short period of time at protein concentration down to ~0.05-0.1 mM. Here, a ^{1}H-^{15}N HSQC spectrum with 128 scans per t_1 increment on protein samples of 0.04-0.07 mM took ~6.5 h to record, using a recycle time of 1.2 s, and generally had a satisfactory signal to noise ratio. Lower protein concentrations of IGF-I and IGF-II do indeed reduce the aggregation and improve the spectral quality at higher pH (see below).

6.4.2 NMR experimental conditions and assignments for IGF-I

The ^{1}H-^{15}N HSQC spectra of $^{15}N/^{2}H$-IGF-I at 0.07 mM had significantly narrower linewidths compared to those previously recorded by Carrick and co-workers on the same sample at 0.5 mM (Carrick et al., 2005). Interestingly, cross-peaks were of much more uniform and slightly narrower linewidths at pH 6.0 and 7.0 than at pH 4.0.

Therefore, subsequent titration experiments of labelled IGF-I were carried out at similar concentration at pH 6.0 (Chapter 7).

To assign the ^{1}H-^{15}N HSQC spectra of IGF-I, a $^{15}N/^{13}C$-IGF-I ample was used to acquire a 3D HNCA spectrum was acquired, which showed the sequential connectivity of the C^{α} cross-peaks. In addition, a 3D CBCA(CO)NH spectrum was used to provide C^{β} information. Although ~0.05-0.1 mM protein was sufficient for 2D experiments, 3D experiments have lower sensitivity and thus ~0.5-1.0 mM protein was necessary even using a cryoprobe. Therefore these 3D spectra were recorded on a 0.7 mM $^{15}N/^{13}C$-IGF-I sample. Several cross-peaks were very broad or even invisible in the ^{1}H-^{15}N HSQC spectrum of this sample at pH 6.0 at 298 K, and were therefore too weak to establish connectivities in corresponding 3D spectra. This was likely to be due to the enhancement of aggregation at higher concentration, together with a faster relaxation of magnetization in the presence of ^{13}C nuclei. Elevated temperature (315 K) and thus a faster tumbling of the protein in solution was then found to significantly narrow the linewidths and permit assignments of most ^{1}H-^{15}N cross-peaks. While elevated temperature was used by others previously, its effect appeared to be more significant in the current study. Schaffer and co-workers reported that ^{1}H-^{15}N HSQC spectra of 1.0 mM IGF-I at pH 5.0 and 40 °C contained only 18 intense and 15 broad and weak backbone cross-peaks out of an expected total of 64 (Schaffer et al., 2003). Addition of a phage-derived IGF-I binding peptide led to a dramatic improvement in the spectral quality (Schaffer et al., 2003).

6.4.3 NMR experimental conditions for IGF-II

Lowering the IGF-II concentration had less significant effects than IGF-I. Most of the previous studies were carried out at pH 3.0-3.5, which was the highest pH where reasonable quality IGF-II spectra could be obtained (Terasawa et al., 1994; Torres et al., 1995; Headey et al., 2004b). Here, ^{1}H-^{15}N HSQC spectra of 0.05 mM IGF-II at pH 4.0 were comparable to those reported for 0.5-2.0 mM IGF-II at pH 3.0-3.5. However, at higher pH (~5.0-5.5) the spectra were much worse and IGF-II still precipitated at this concentration. Thus, lowering the protein concentration appeared to shift the aggregation regime by ~0.5-1.0 pH unit. Interestingly, the spectra were not significantly improved either by higher temperature (308 K), or by the addition of 50 mM arginine

plus 50 mM glutamate, a method reported recently for improving protein solubility in NMR samples (Golovanov et al., 2004). Previously, addition of acetic acid (~5 % (v/v)) was found to prevent aggregation of an IGF-II NMR sample in one study (Terasawa et al., 1994), but was found to produce the opposite effect in another (Torres et al., 1995).

6.4.4 NMR spectroscopy of N-BP-2

^{1}H-^{15}N HSQC spectra of N-BP-2 were recorded at pH 7.0, a pH chosen to be distant from its theoretical pI (4.8) while avoiding serious chemical exchange with H_2O. The presence of widely dispersed cross-peaks together with some sharp cross-peaks in the random coil region strongly indicated that N-BP-2 contained structured regions as well as a significant percentage of disordered region. Most of the sharp cross-peaks overlapping in the centre of the spectra were believed to correspond to the ~40 residues belonging to the IGFBP-2 linker domain, and, are likely to be disordered. In addition, the considerable number of cross-peaks clustering in the lower region of the spectra, i.e., ^{15}N ~125-130 ppm, is consistent with the fact that N-BP-2 contained 16 alanine residues. In particular, the N-domain of IGFBP-2 has a unique ~15-residue alanine-rich (seven alanines) insertion that is not present in other IGFBPs. Some of the cross-peaks in the ^{15}N ~125-130 ppm region were expected to correspond to these alanines, and, because these cross-peaks were sharp and had narrow ^{1}H chemical shift dispersion, these residues are likely to be in a region of flexible structure.

Usually 3D spectra using a $^{15}N/^{13}C$-labelled sample are required to assign the ^{1}H-^{15}N HSQC spectrum of a protein of this size, especially when there is significant overlap. $^{15}N/^{13}C$-labelled N-BP-2 was not prepared in this thesis, and hence assignments of the ^{1}H-^{15}N HSQC spectra were not performed. Since many sharp cross-peaks in the random coil region were not perturbed upon binary and ternary complex formation with IGF and C-BP-2 (Chapter 7), indicating these residues, very likely in the linker domain, were not involved in the interactions, the linker domain region can be removed from the N-BP-2 construct to simplify its spectra in the future.

6.4.5 Effects of heparin on IGF-I and N-BP-2 spectra

In Chapter 5, it was showed that C-BP-2 bound heparin in a pH-dependent manner and the binding site was mapped. The effects of heparin on the binary and ternary complexes are to be studied in future studies by the chemical shift perturbation method using ^{15}N-labelled C-BP-2, IGF-I, and N-BP-2 (Chapter 7). Before these experiments are carried out, it was necessary to determine whether heparin binds to IGF-I or N-BP-2 alone. While no heparin binding of IGF-I was detected by others (Fernandez-Tornero et al., 2005), some other growth factors, including fibroblast growth factor (FGF), hepatocyte growth factor (HGF), vascular endothelial growth factor (VEGF), platelet derived growth factor (PDGF), and transforming growth factor-β1 (TGF-β1), and many chemokines have been shown to be heparin-binding proteins, and these heparin-protein interactions modulate cell growth, differentiation, and inflammation (Capila and Linhardt, 2002).

In this study, addition of low molecular weight heparin to IGF-I induced significant chemical shift changes of two C domain residues located in a flexible loop, but chemical shift changes of other resonances were very small. The results indicated that Arg36 and Arg37 may interact with the heparin molecules, but IGF-I does not have a moderately large heparin binding interface as do C-BP-2, some heparin-binding growth factors and chemokines (Capila and Linhardt, 2002). Whether such weak heparin binding, as detected by NMR, is biologically relevant requires future investigation. It will be also of interest to perform a similar experiment on IGF-II, which has equivalent doublet of arginine residues in its C domain. On the other hand, N-BP-2 did not bind heparin at either pH 7.0 or pH 6.0, and therefore, heparin binding by IGFBP-2 is mediated by the C-domain and the linker domain.

Chapter 7

Binding Sites and Cooperativity of IGFBP-2 N- and C-terminal Domains in IGF Binding

7.1 Introduction

As reviewed in Chapter 1, both the N- and C-domains of IGFBPs are required for high-affinity IGF binding. However, while the IGF binding sites on an N-domain fragment (mini-N-BP-5) had been studied using NMR spectroscopy (Kalus et al., 1998) and crystallography (Zeslawski et al., 2001), there was limited structural information regarding the IGF binding sites on IGFBP C-domains at the time this project commenced. Some residues or regions in the C-domain were found to be important in IGF binding in previous studies, by means of deletions (Brinkman et al., 1991; Forbes et al., 1998), site-directed mutagenesis (Bramani et al., 1999; Song et al., 2000; Shand et al., 2003; Yan et al., 2004; Allan et al., 2006), antibody binding (Schuller et al., 1993), photoaffinity labelling (Horney et al., 2001) and subsequent studies (Kibbey et al., 2006). Structural studies were required to reconcile these findings.

Only recently have structures of IGFBP C-domains (Headey et al., 2004a; Sala et al., 2005; Sitar et al., 2006) and IGF·C-domain binding interfaces (Headey et al., 2004a; Headey et al., 2004b; Sitar et al., 2006) been reported. Headey and co-workers mapped the IGF-II binding site on C-BP-6 and C-BP-6 binding site on IGF-II using NMR spectroscopy (Headey et al., 2004a; Headey et al., 2004b). Because of the very low affinity of C-BP-6 for IGF-I (Headey et al., 2004c), interaction between C-BP-6 and IGF-I was not studied by NMR.

C-BP-2 has similar affinities for IGF-I and IGF-II (Chapter 3). Therefore, it was of interest to investigate both IGF-I and IGF-II binding of C-BP-2 using NMR spectroscopy. The unlabelled and ^{15}N-labelled C-BP-2 and N-BP-2 fragments prepared in this thesis (Chapter 3), as well as the ^{15}N-labelled IGF-I and IGF-II prepared previously by Dr Francine Carrick and Ms Kerrie McNeil (The University of Adelaide), enabled a detailed analysis of the molecular interactions among IGFBP-2 domains and IGFs. Since the NMR assignments of C-BP-2 (Chapter 4) and IGF-I (Chapter 6), as

well as the three-dimensional structures of C-BP-2 (Chapter 4) and IGFs were available, their binding interfaces could be identified using the chemical shift perturbation mapping method as reviewed in Chapter 5. Although assignments for N-BP-2 were not pursued in this study, the perturbations of ^{15}N-labelled N-BP-2 upon binding to unlabelled IGFs or C-BP-2 were valuable to supplement the results of ^{15}N-labelled C-BP-2 and IGFs.

Using chemical shift perturbation mapping, one can not only identify residues that directly contact the binding partner and/or undergo conformational changes upon binding, but also estimate the strength of binding based on the changes in resonance (Zuiderweg, 2002). As such, the titrations with NMR in this Chapter provided a third measurement of the IGF binding by N-BP-2 and C-BP-2, in addition to the cross-linking and BIAcore analysis in Chapter 3.

With both N-BP-2 and C-BP-2, it was also of interest to determine whether co-incubation of these domain fragments would enhance their IGF binding affinities using NMR. Previously, Payet and co-workers reported that cooperativity existed between the IGFBP-3 N- and C-domain fragments (Payet et al., 2003), but others failed to detect such enhancement using BIAcore (Carrick et al., 2001; Headey et al., 2004c).

Carrick and co-workers studied IGF binding by full-length IGFBP-2 using ^{15}N-labelled IGFs (Carrick et al., 2005), and more recently Sitar and co-workers solved the crystal structure of the IGF-I·N-BP-4·C-BP-4 ternary complex (Sitar et al., 2006). The results presented in this Chapter, when taken together with other recent reports (Payet et al., 2003; Headey et al., 2004a; Headey et al., 2004b; Carrick et al., 2005; Sitar et al., 2006), provide valuable insights into the IGF-IGFBP binding interaction.

7.2 Materials and Methods

7.2.1 Materials

^{15}N/^{2}H-labelled IGF-I and ^{15}N/^{2}H-labelled IGF-II were kindly provided by Dr Francine Carrick (The University of Adelaide) (Carrick et al., 2005). ^{15}N/^{13}C-labelled IGF-I was kindly prepared by Ms Kerrie McNeil (The University of Adelaide). Unlabelled IGF-I was kindly provided by Dr Adam Denley (The University of

Adelaide). Unlabelled IGF-II was purchased from GroPep (Adelaide, Australia). ^{15}N-labelled C-BP-2 and N-BP-2, as well as unlabelled C-BP-2, N-BP-2, Large-C-BP-2 were prepared by the author as described in Chapter 3.

7.2.2 NMR experiments on ^{15}N-labelled IGF-I and IGF-II

Unlabelled C-BP-2, N-BP-2, and Large-C-BP-2 were titrated into the ^{15}N-labelled (^{15}N/^{2}H or ^{15}N/^{13}C) IGF-I samples in separate experiments. Unlabelled C-BP-2 and Large-C-BP-2 were then titrated into the ^{15}N/^{2}H-IGF-I + N-BP-2 (1:1.5) samples. ^{1}H-^{15}N HSQC spectra of IGF-I were recorded as described in 6.2.2 at particular protein molar ratios. Details of experimental conditions are also given in figure captions in the Results section. 3D HNCA and CBCA(CO)NH spectra were recorded on a ^{15}N/^{13}C-IGF-I (0.2 mM) + N-BP-2 (1:1.2) sample at 298 K and 313 K. Experimental parameters of the 3D spectra were similar to those used in 6.2.2

Similarly, titration experiments were carried out using ^{15}N/^{2}H-IGF-II samples, and ^{1}H-^{15}N HSQC spectra were recorded as described in 6.2.3.

^{1}H-^{15}N HSQC spectra of IGF-I were also recorded on a ^{15}N/^{2}H-IGF-I (0.07 mM) + N-BP-2 + C-BP-2 (1:1.5:1.5) sample in a buffer containing 150 mM NaCl at pH 6.0, and in the absence and presence of 2.0 mM low molecular weight heparin.

7.2.3 NMR experiments on ^{15}N-labelled C-BP-2

Unlabelled IGF-I, IGF-II, and N-BP-2 were titrated into the ^{15}N-labelled C-BP-2 samples in separate experiments. Unlabelled N-BP-2 was then titrated into the ^{15}N-C-BP-2 + IGF-I (1:1.5) or ^{15}N-C-BP-2 + IGF-II (1:1.5) samples. ^{1}H-^{15}N HSQC spectra of C-BP-2 were recorded as described in 4.2.2.

7.2.4 NMR experiments on ^{15}N-labelled N-BP-2

Unlabelled IGF-I, or C-BP-2 were titrated into the ^{15}N-labelled C-BP-2 samples in separate experiments. Unlabelled C-BP-2 and Large-C-BP-2 were then titrated into the ^{15}N-N-BP-2 + IGF-I (1:1.5) samples. ^{1}H-^{15}N HSQC spectra of N-BP-2 were recorded as described in 6.2.4.

7.3 Results

7.3.1 Binding of C-BP-2 to IGFs

Binding interactions between C-BP-2 and IGFs were studied using two orientations in separate experiments. In one set of experiments unlabelled C-BP-2 was titrated into ^{15}N-labelled IGF-I or ^{15}N-labelled IGF-II, while in the other unlabelled IGF-I and IGF-II were titrated into ^{15}N-labelled C-BP-2.

^{1}H-^{15}N HSQC spectra of free ^{15}N/^{13}C-IGF-I and at IGF-I:C-BP-2 ratios of 1:0.5, 1:1, and 1:1.5 are shown in Figure 7.1. Titration of C-BP-2 into ^{15}N-labelled IGF-I caused gradual broadening of the IGF-I cross-peaks, indicating that the binding interaction was in the intermediate exchange regime (described in Figure 5.2). Additions of unlabelled C-BP-2 to ^{15}N-labelled IGF-II caused a similar broadening of specific IGF-II cross-peaks (not shown). In addition, titrations of Large-C-BP-2 into ^{15}N-labelled IGF-I and ^{15}N-labelled IGF-II samples induced basically identical peak broadening patterns compared to those induced by C-BP-2 (not shown).

^{1}H-^{15}N HSQC spectra of free ^{15}N-labelled C-BP-2 and at C-BP-2:IGF-I ratios of 1:0.5, 1:1, and 1:1.5 are shown in Figures 7.2 and 7.3, and those of ^{15}N-C-BP-2 at C-BP-2:IGF-II ratios of 1:1.5 are shown in Figure 7.3. It can be seen that IGF-I and IGF-II induced very similar broadening of the C-BP-2 cross-peaks. C-BP-2 backbone amide cross-peaks of His221, Asn224, Cys225, Cys236, Lys237, Met238, Ser239, Leu240, Asn241, and Gly242, and the sidechain amide cross-peak of Gln192 disappeared or broadened significantly at C-BP-2:IGF-I or C-BP-2:IGF-II ratios of 1:1.5. The locations of these residues on the C-BP-2 surface are discussed in 7.4.4.

7.3.2 Binding of N-BP-2 to IGFs

^{1}H-^{15}N HSQC spectra of free IGF-I and at IGF-I:N-BP-2 ratios of 1:0.5, 1:1, and 1:1.5 are shown in Figure 7.4. In contrast to the effects seen upon addition of C-BP-2, titrations of N-BP-2 caused a gradual disappearance of the "free" set of IGF-I cross-peaks and the simultaneous appearance of a "bound" set of cross-peaks. The results indicated that the binding interaction between N-BP-2 and IGF-I was in the slow exchange regime (described in Figure 5.2). Apart from the cross-peaks that did not shift

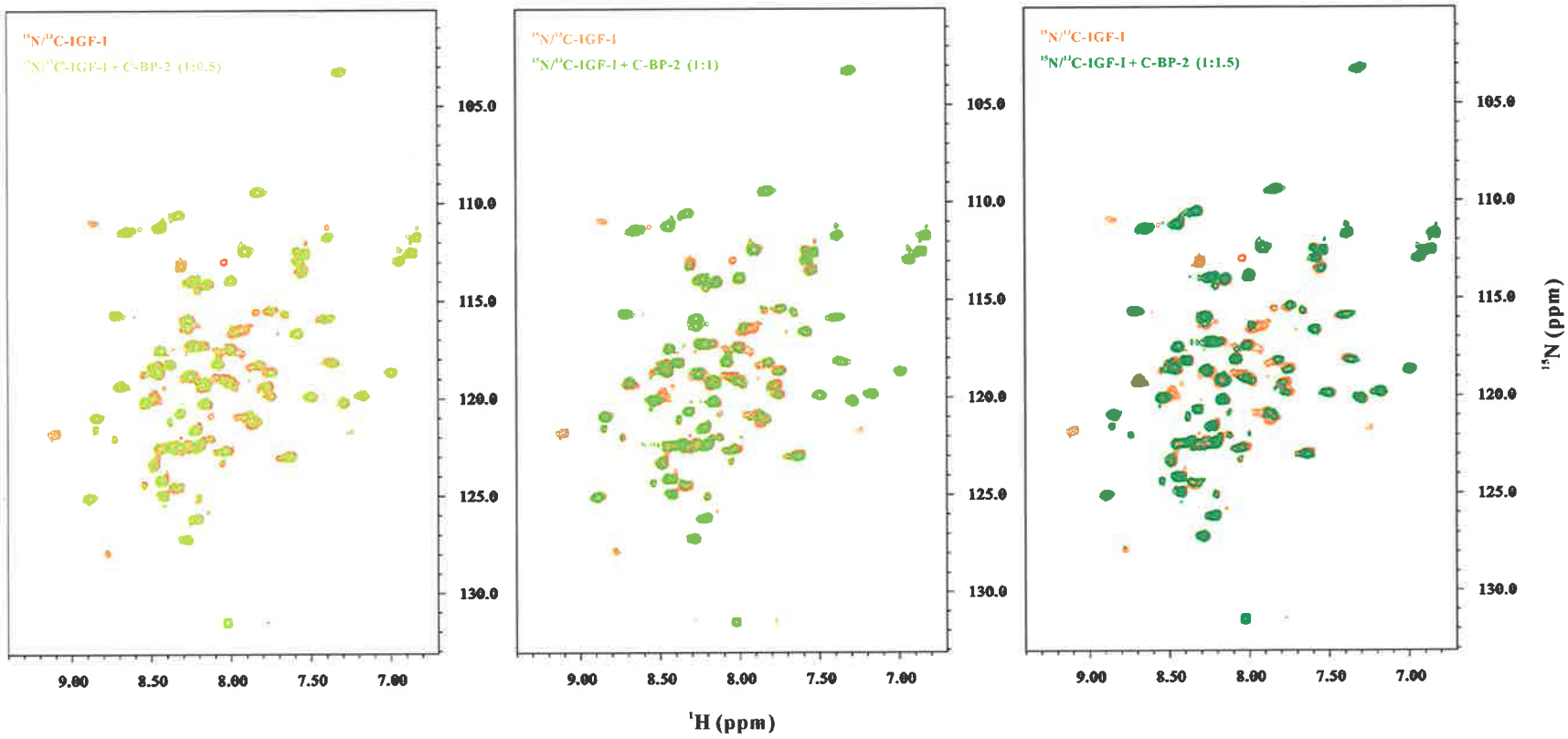


Figure 7.1 Titration of unlabelled C-BP-2 into ^{15}N/^{13}C-labelled IGF-I

Overlay of the ^{1}H-^{15}N HSQC spectra of 0.04 mM ^{15}N/^{13}C-labelled IGF-I in the absence and presence of unlabelled C-BP-2 at molar ratios of 1:0.5, 1:1, and 1:1.5. The samples were in 95% H_2O/5% $2H_2O$ containing 10 mM sodium acetate, 150 mM NaCl, and 0.02% (w/v) sodium azide at pH 6.0. Spectra were recorded at 500 MHz and 298 K.

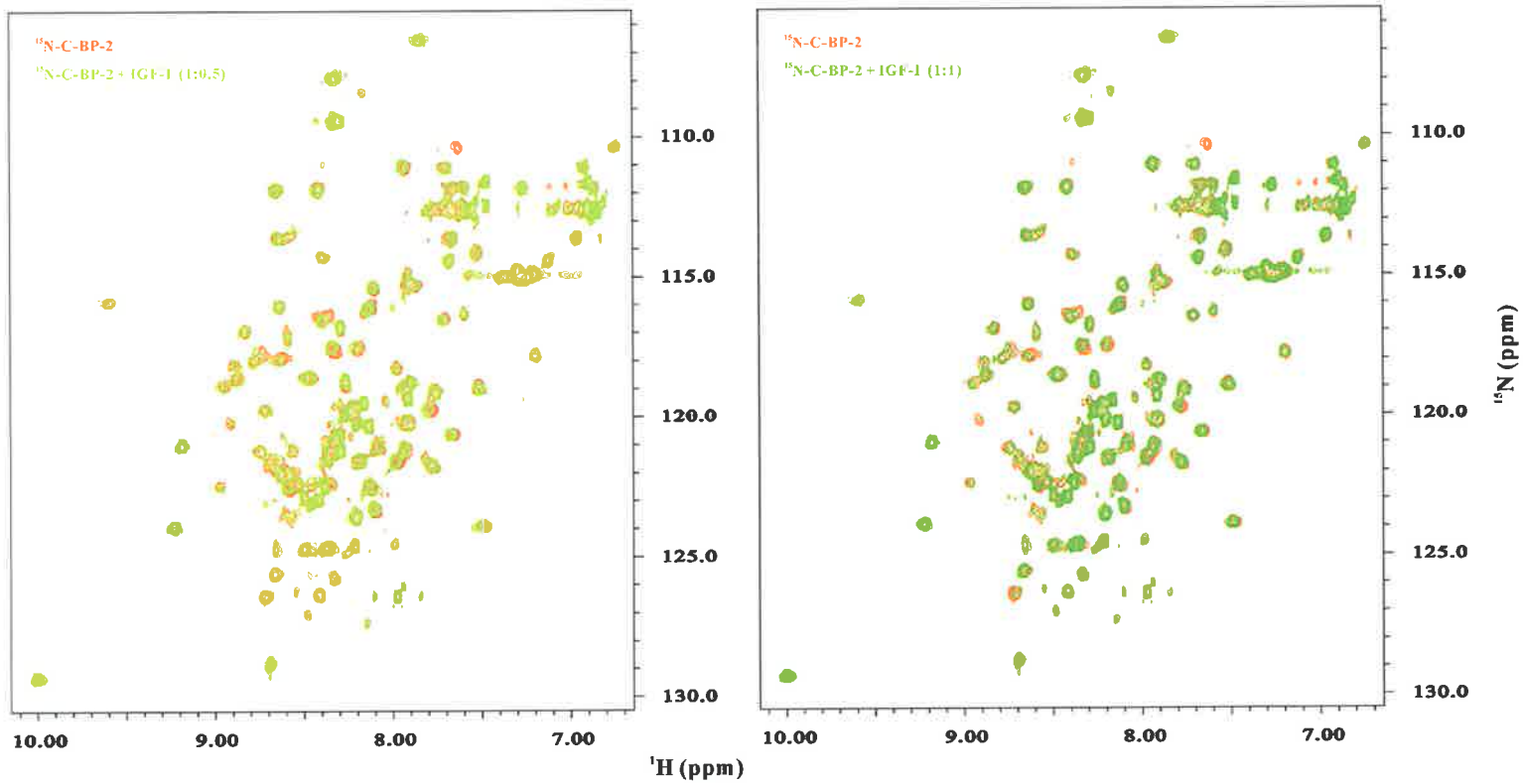


Figure 7.2 Titration of unlabelled IGF-I into ^{15}N-labelled C-BP-2

Overlay of the ^{1}H-^{15}N HSQC spectra of 0.04 mM ^{15}N-labelled C-BP-2 in the absence and presence of unlabelled IGF-I at molar ratios of 1:0.5 and 1:1. The samples were in 95% H_2O/5% $2H_2O$ containing 10 mM sodium acetate, 150 mM NaCl, and 0.02% (w/v) sodium azide at pH 6.0. Spectra were recorded at 500 MHz and 298 K.

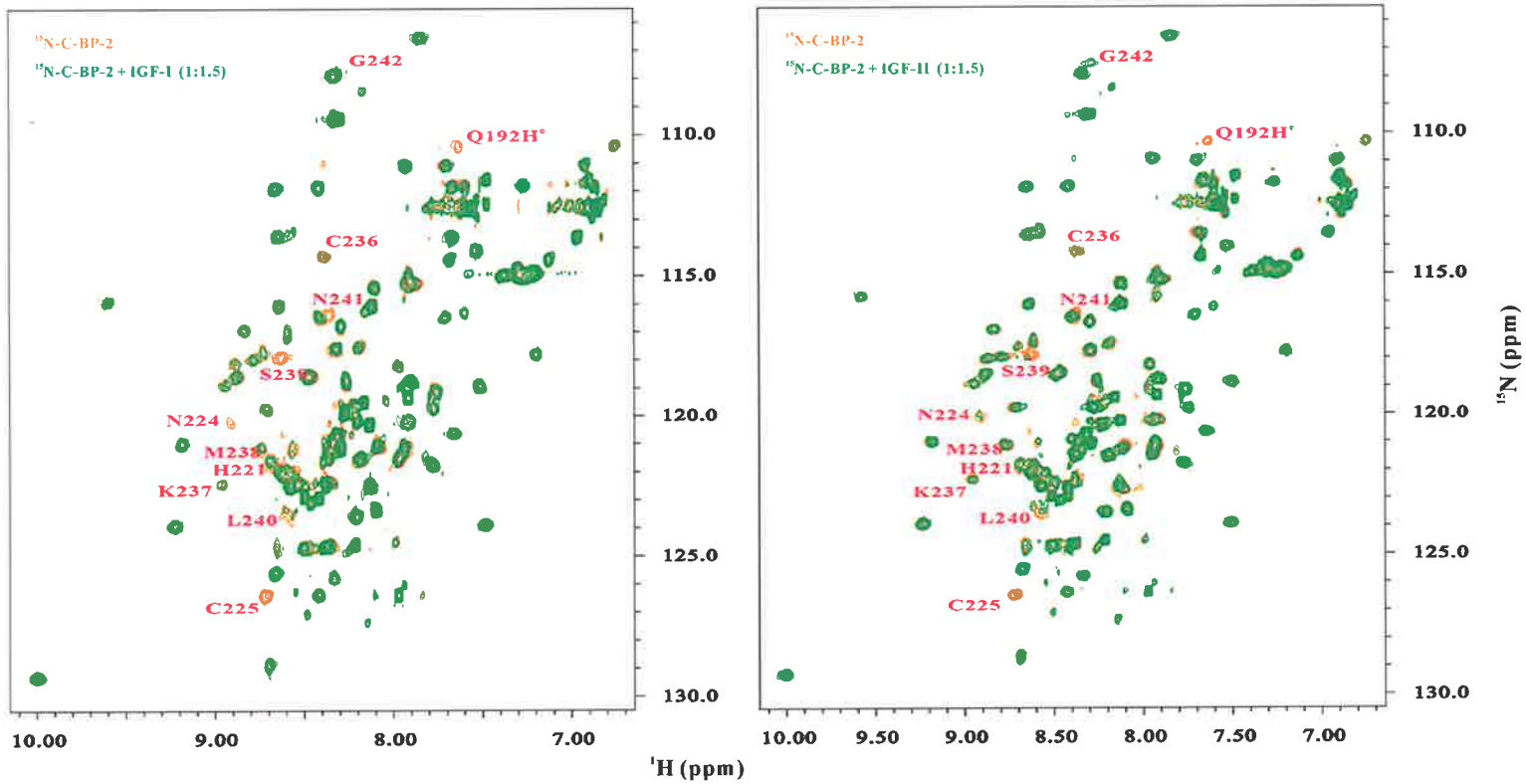


Figure 7.3 Comparison of ^{15}N-labelled C-BP-2 in the absence and presence of unlabelled IGF-I or unlabelled IGF-II

Overlay of the ^{1}H-^{15}N HSQC spectra of 0.04 mM ^{15}N-labelled C-BP-2 in the absence and presence of unlabelled IGF-I (left) or unlabelled IGF-II (right) at C-BP-2:IGF molar ratios of 1:1.5. The samples were in 95% H_2O/5% $2H_2O$ containing 10 mM sodium acetate, 150 mM NaCl, and 0.02% (w/v) sodium azide at pH 6.0. Spectra were recorded at 500 MHz and 298 K. The most affected residues are labelled.

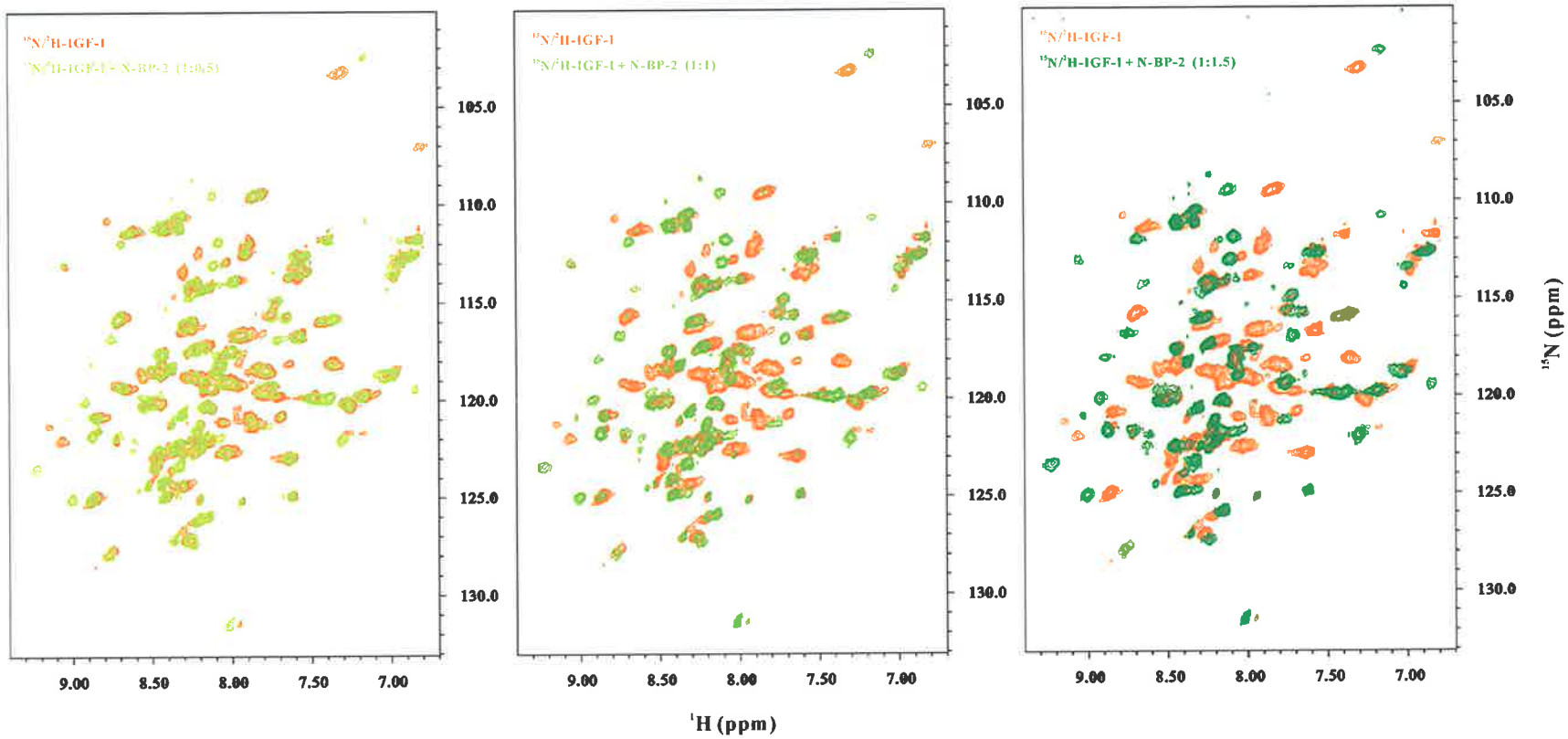


Figure 7.4 Titration of unlabelled N-BP-2 into the ^{15}N/^{2}H-labelled IGF-I

Overlay of the ^{1}H-^{15}N HSQC spectra of 0.07 mM ^{15}N/^{2}H-labelled IGF-I in the absence and presence of unlabelled N-BP-2 at molar ratios of 1:0.5, 1:1, and 1:1.5. The samples were in 95% H_2O/5% $2H_2O$ containing 10 mM sodium acetate and 0.02% (w/v) sodium azide at pH 6.0. Spectra were recorded at 500 MHz and 298 K.

upon binding to N-BP-2, chemical shift assignments of the "bound" set of cross-peaks were not obtained using 3D HNCA and CBCA(CO)NH spectra recorded on a $^{15}N/^{13}C$-IGF-I + N-BP-2 sample at 298 K and 313 K because peaks were very broad or even absent in the 3D spectra and connectivities could not be established. In the reverse experiments, addition of unlabelled IGF-I into ^{15}N-labelled N-BP-2 also resulted in chemical shift changes (Figure 7.5 (left)).

7.3.3 Binding of C-BP-2 to the IGF·N-BP-2 binary complexes

When unlabelled C-BP-2 was titrated into the $^{15}N/^{2}H$-IGF-I:N-BP-2 binary complex described in 7.3.2 (IGF-I:N-BP-2 1:1.5), some IGF-I cross-peaks underwent further chemical shift changes in a slow exchange manner (Figure 7.6). Thus, the results indicated that the binding affinity of C-BP-2 for the IGF-I·N-BP-2 binary complex was higher than the binding affinity of C-BP-2 for IGF-I alone. Similar results were also seen in the titrations of C-BP-2 to the IGF-II·N-BP-2 binary complex (not shown).

A comparison of the ^{1}H-^{15}N HSQC spectra of IGF-I in the IGF-I·N-BP-2 binary complex and the IGF-I·N-BP-2·C-BP-2 ternary complex with that of IGF-I in the IGF-I·IGFBP-2 complex (Figure 7.7) (Carrick et al., 2005) showed that IGF-I resonances in the IGF-I·N-BP-2·C-BP-2 ternary complex had equivalent chemical shifts compared to the IGF-I·IGFBP-2 complex. (Note that the spectra of IGFs in complex with IGFBP-2 were TROSY spectra (Carrick et al., 2005), which have backbone amide chemical shifts differences of approximately +0.7-0.8 ppm in the ^{15}N dimension and -0.08 ppm in the ^{1}H dimension compared to the conventional HSQC spectra used here (Pervushin et al., 1997). Small chemical shift differences also arose from slightly different temperature and pH conditions.)

Titrations of Large-C-BP-2 into the IGF-I·N-BP-2 binary complex induced essentially identical chemical shift changes in IGF-I compared with those caused by C-BP-2 (not shown). The differences were that the peaks of IGF-I in the IGF-I·N-BP-2·Large-C-BP-2 complex were broader than those in the IGF-I·N-BP-2·C-BP-2 complex, presumably because of the larger molecular size of the complex formed by Large-C-BP-2.

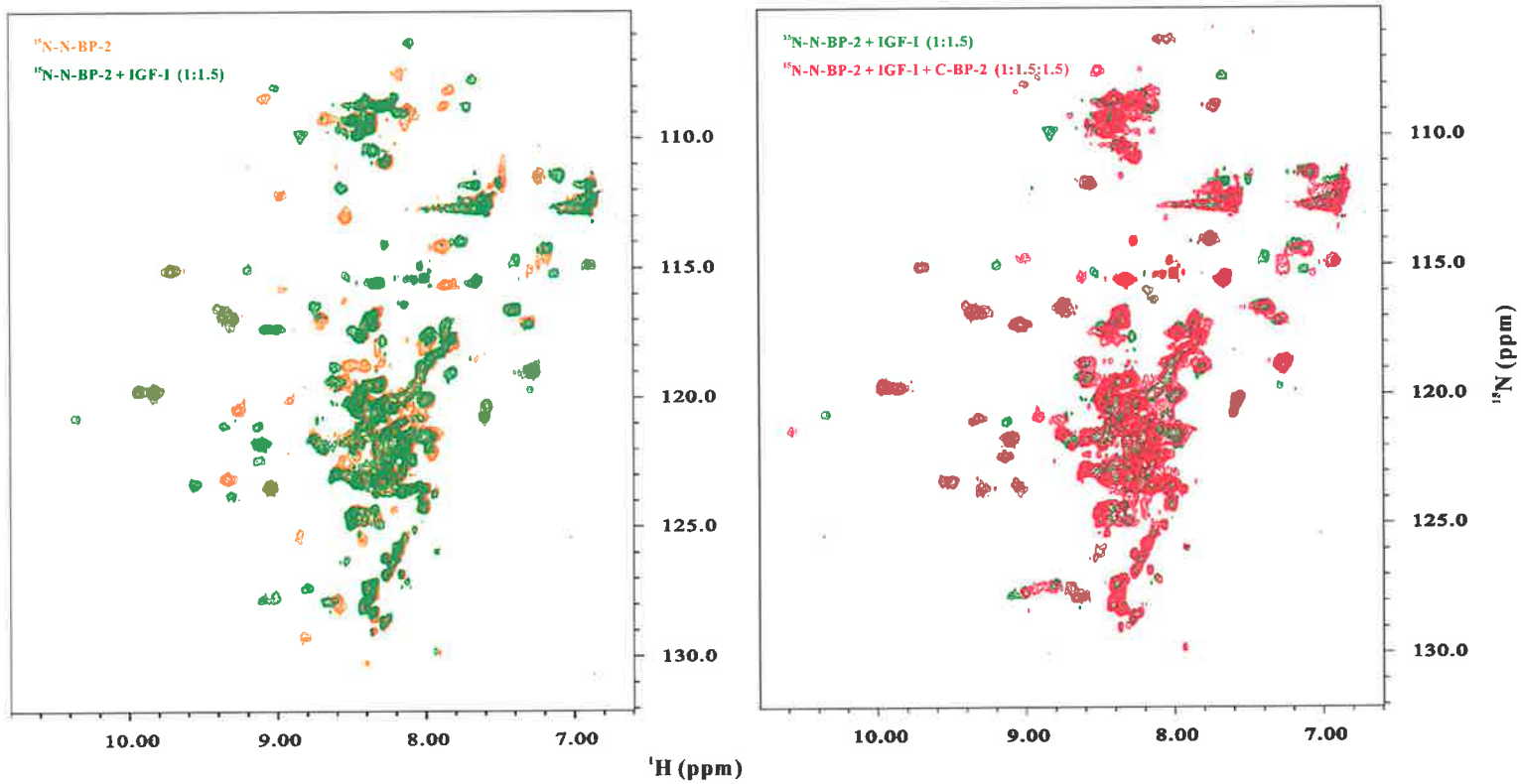


Figure 7.5 Comparison of ^{15}N-labelled N-BP-2 free and in IGF-I·N-BP-2 or IGF-I·N-BP-2·C-BP-2 complexes

Overlay of the ^{1}H-^{15}N HSQC spectra of 0.07 mM ^{15}N-labelled N-BP-2, free or in IGF-I·N-BP-2 complex (Left), and in the IGF-I·N-BP-2 complex or in the IGF-I·N-BP-2·C-BP-2 complex (right). The samples were in 95% H_2O/5% $2H_2O$ containing 10 mM sodium acetate and 0.02% (w/v) sodium azide at pH 7.0. Spectra were recorded at 500 MHz and 298 K.

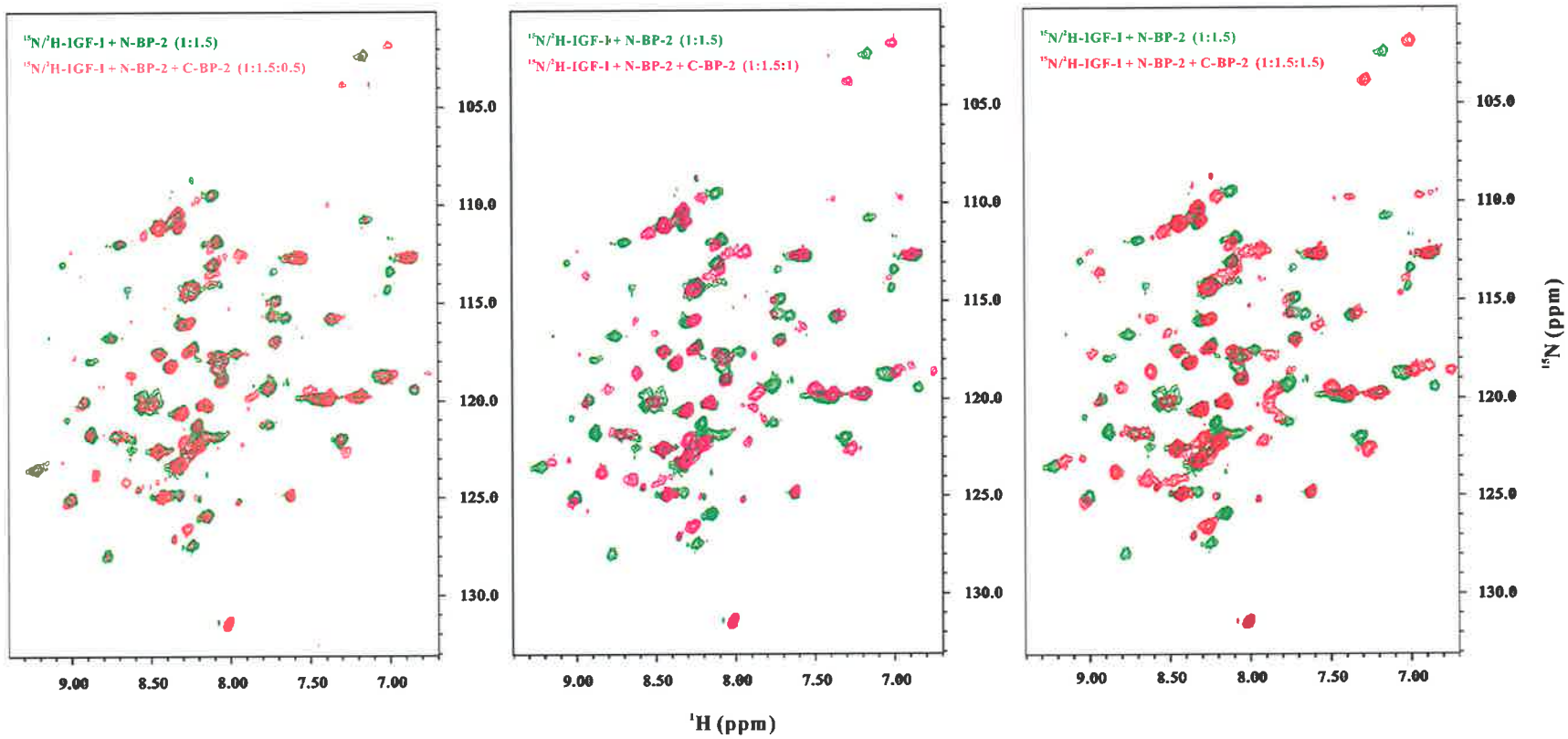


Figure 7.6 Titration of unlabelled C-BP-2 into the $^{15}N/^{2}H$-IGF-I·N-BP-2 binary complex

Overlay of the ^{1}H-^{15}N HSQC spectra of $^{15}N/^{2}H$-IGF-I·N-BP-2 binary complex ($^{15}N/^{2}H$-IGF-I 0.07 mM) in the absence and presence of unlabelled C-BP-2 at IGF-I:N-BP-2:C-BP-2 molar ratios of 1:1.5:0.5, 1:1.5:1, and 1:1.5:1.5. The samples were in 95% H_2O/5% $2H_2O$ containing 10 mM sodium acetate and 0.02% (w/v) sodium azide at pH 6.0. Spectra were recorded at 500 MHz and 298 K.

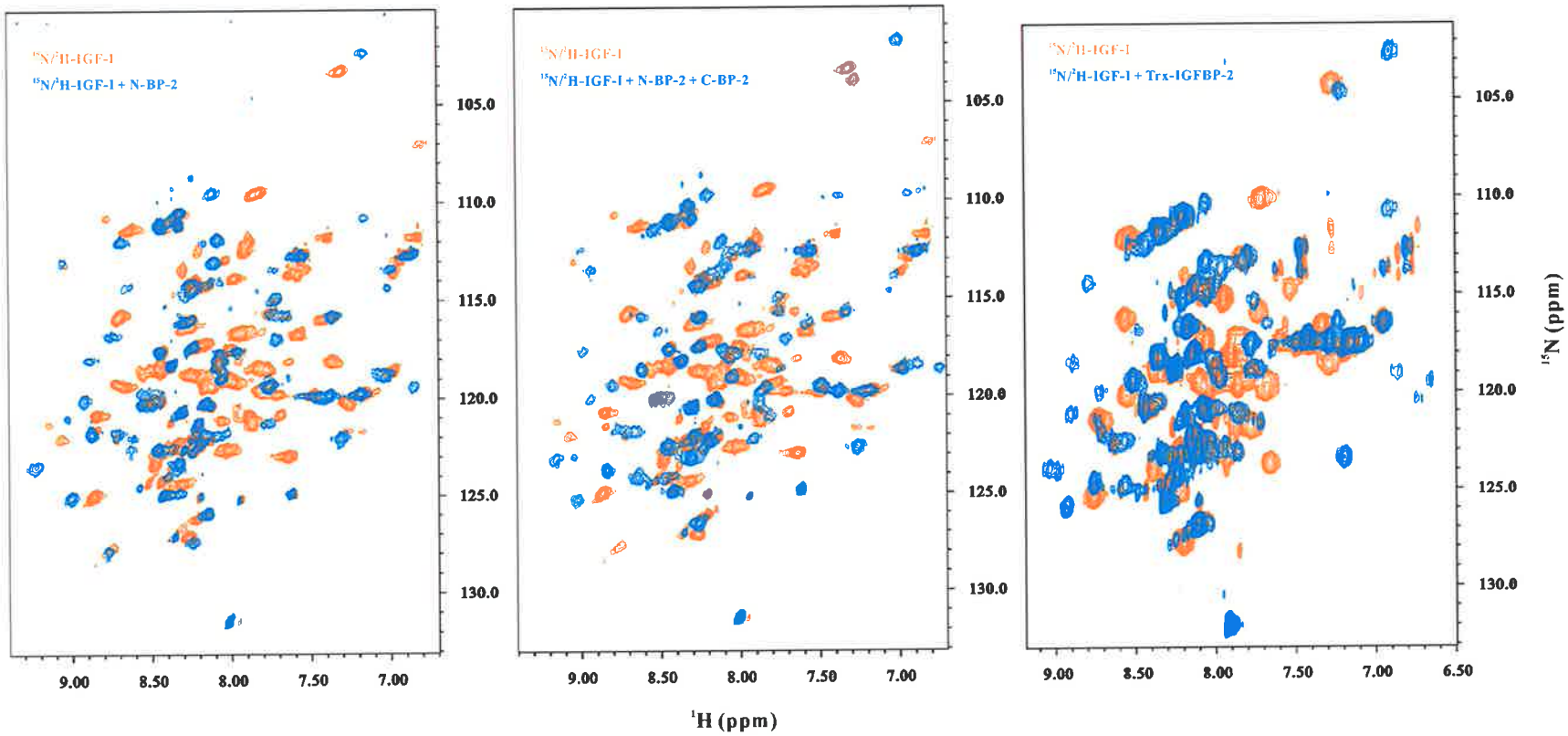


Figure 7.7 Comparison of ^{15}N/^{2}H-IGF-I in IGF-I·N-BP-2·C-BP-2 ternary complex and in IGF-I·Trx-IGFBP-2 complex

Overlay of the ^{1}H-^{15}N HSQC spectra of ^{15}N/^{2}H-IGF in free and in IGF-I·N-BP-2 binary complex (left), in IGF-I·N-BP-2·C-BP-2 ternary complex (middle), and in IGF-I·thioredoxin-IGFBP-2 complex (Carrick et al., 2005) (right). Left and middle spectra were recorded on 0.07 mM ^{15}N/^{2}H-IGF-I samples at pH 6.0, 298 K and 500 MHz. Right spectra were TROSY spectra, and were recorded on 0.5 mM ^{15}N/^{2}H-IGF-I samples at pH 4.8-5.1, 37 °C and 600 MHz.

The formation of ternary complexes was also monitored using ^{15}N-labelled N-BP-2. Addition of unlabelled C-BP-2 into the ^{15}N-N-BP-2·IGF-I binary complex described in 7.3.2 induced chemical shift perturbations of several N-BP-2 cross-peaks (Figure 7.5 (right)). Addition of Large-C-BP-2 into the ^{15}N-N-BP-2·IGF-I binary complex formed in 7.3.2 induced essentially identical chemical shift changes compared with those induced by C-BP-2 (not shown), indicating that in the ternary complex the additional ~40 residues (approximately the second half of the linker domain) in Large-C-BP-2 did not make additional stable contacts with N-BP-2, which contained approximately the first half of the linker domain.

The formation of the ternary complex was also monitored using ^{15}N-labelled C-BP-2. In separate experiments, unlabelled N-BP-2 was added into the ^{15}N-C-BP-2·IGF samples formed in 7.3.1, whereas unlabelled IGFs were added into the ^{15}N-C-BP-2·N-BP-2 samples described in 7.3.4 (see below). ^{15}N-C-BP-2 in the ternary complexes gave very broad cross-peaks that were beyond the point of detection using the experimental parameters employed for its free form and binary complexes. Nevertheless, ^{1}H-^{15}N HSQC spectra were recorded using a significantly increased number of scans per increment (Figure 7.8). Superposition of the ^{1}H-^{15}N HSQC spectra of free C-BP-2 and in the IGF-I·N-BP-2·C-BP-2 ternary complex (Figure 7.8 (left)) showed that most of C-BP-2 backbone amide cross-peaks, except Ala187 at the N-terminus and those at the C-terminus beyond Gln278, shifted to new chemical shift positions or disappeared. Furthermore, the majority of C-BP-2 cross-peaks had the same chemical shifts in the ternary complexes containing either IGF-I or IGF-II, although several cross-peaks had slightly different chemical shifts in these two ternary complexes (Figure 7.8 (right)).

7.3.4 Interaction between N-BP-2 and C-BP-2 in the absence of IGFs

Interaction between N-BP-2 and C-BP-2 in the absence of IGFs was investigated using both ^{15}N-labelled C-BP-2 and ^{15}N-labelled N-BP-2. Titration of unlabelled N-BP-2 into ^{15}N-labelled C-BP-2 sample caused significant broadening of many cross-peaks (Figure 7.9). Some cross-peaks disappeared at C-BP-2:N-BP-2 ratio of 1:0.5, and more cross-peaks were invisible at a molar ratio of 1:1, although small chemical shift changes for some cross-peaks were also evident (Figure 7.9). The locations of the most

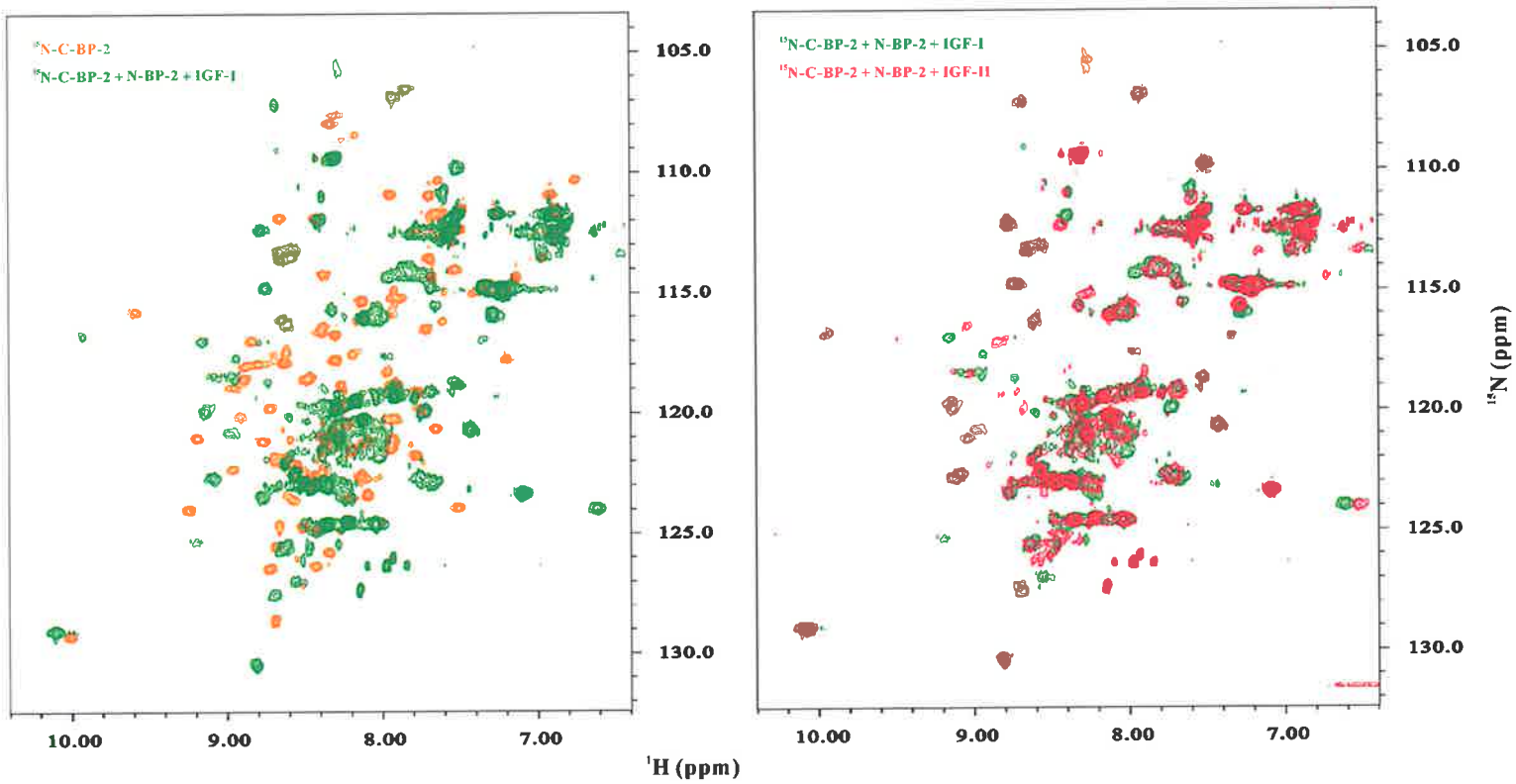


Figure 7.8 Comparison of ^{15}N-labelled C-BP-2 in free and in IGF·N-BP-2·C-BP-2 ternary complexes

Overlay of the ^{1}H-^{15}N HSQC spectra of 0.05 mM ^{15}N-labelled C-BP-2 in free and in IGF-I·N-BP-2·C-BP-2 ternary complex (left), or in IGF-I·N-BP-2·C-BP-2 and IGF-II·N-BP-2·C-BP-2 ternary complexes (right). The samples were in 95% H_2O/5% $2H_2O$ containing 10 mM sodium acetate and 0.02% (w/v) sodium azide at pH 6.0. Spectra were recorded at 500 MHz and 298 K.

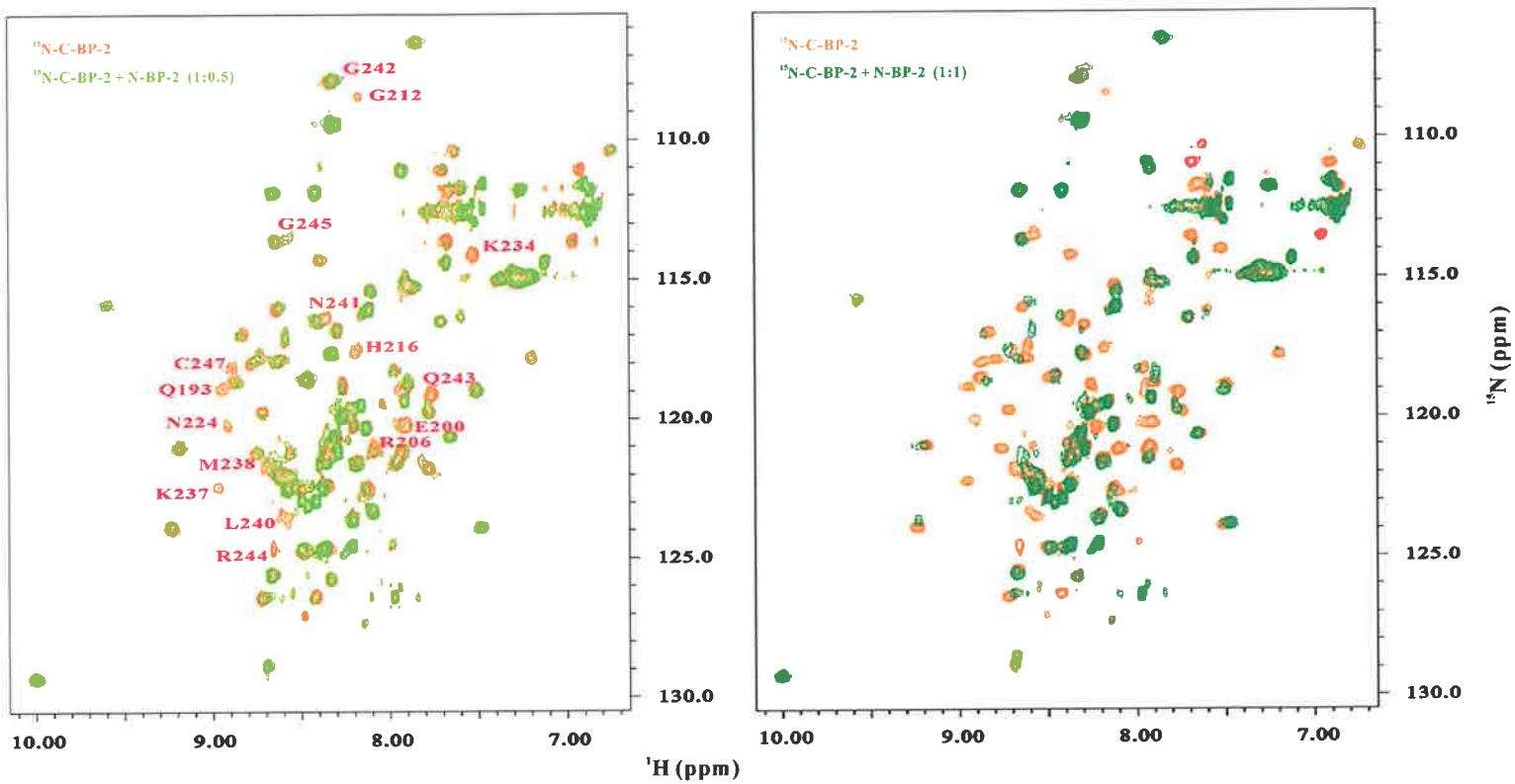


Figure 7.9 Titration of unlabelled N-BP-2 into ^{15}N-labelled C-BP-2

Overlay of the ^{1}H-^{15}N HSQC spectra of 0.04 mM ^{15}N-labelled C-BP-2 in the absence and presence of unlabelled N-BP-2 at C-BP-2:N-BP-2 molar ratios of 1:0.5 and 1:1. The samples were in 95% H_2O/5% $2H_2O$ containing 10 mM sodium acetate, 150 mM NaCl, and 0.02% (w/v) sodium azide at pH 6.0. Spectra were recorded at 500 MHz and 298 K. Resonances that disappeared or significantly broadened at 1:0.5 are labelled.

affected residues on C-BP-2 surface are discussed in 7.4.4. Interestingly, although more C-BP-2 cross-peaks were broadened upon binding to N-BP-2 compared with the changes induced by IGFs (Figure 7.3), and some cross-peaks were broadened in both interactions (see below), the broadening patterns were different, which indicated some specificity. For example, while backbone amide cross-peaks of Cys225 and Cys236 of C-BP-2 were among the most affected residues in IGF interactions (Figure 7.3), they were less affected compared to other residues in N-BP-2 interaction (Figure 7.9). Addition of unlabelled C-BP-2 into ^{15}N-labelled N-BP-2 induced chemical shift changes and peak broadening of some cross-peaks (Figure 7.10), which further confirmed the existence of an interaction between N-BP-2 and C-BP-2 even in the absence of IGFs.

7.3.5 Effects of heparin binding on IGF binding by N- and C-BP-2

Addition of low molecular weight heparin into the ^{15}N/^{2}H-IGF-I·N-BP-2·C-BP-2 ternary complex (section 7.3.1) at pH 6.0 caused significant broadening or even disappearance of IGF-I cross-peaks (Figure 7.11). Nevertheless, several remaining cross-peaks had chemical shifts representing IGF-I in the IGF-I·N-BP-2·C-BP-2 ternary complex, but not in the IGF-I·N-BP-2 binary complex or in the free form (Figure 7.11). When the sample pH was decreased to 5.5, more cross-peaks disappeared and only a few cross-peaks for residues in the flexible regions remained in the spectra. Thus, the results suggested that the IGF-I·N-BP-2·C-BP-2 ternary complex can bind heparin, probably mediated by the C-BP-2 heparin binding site (Chapter 5), and binding to heparin did not significantly dissociate the ternary complex.

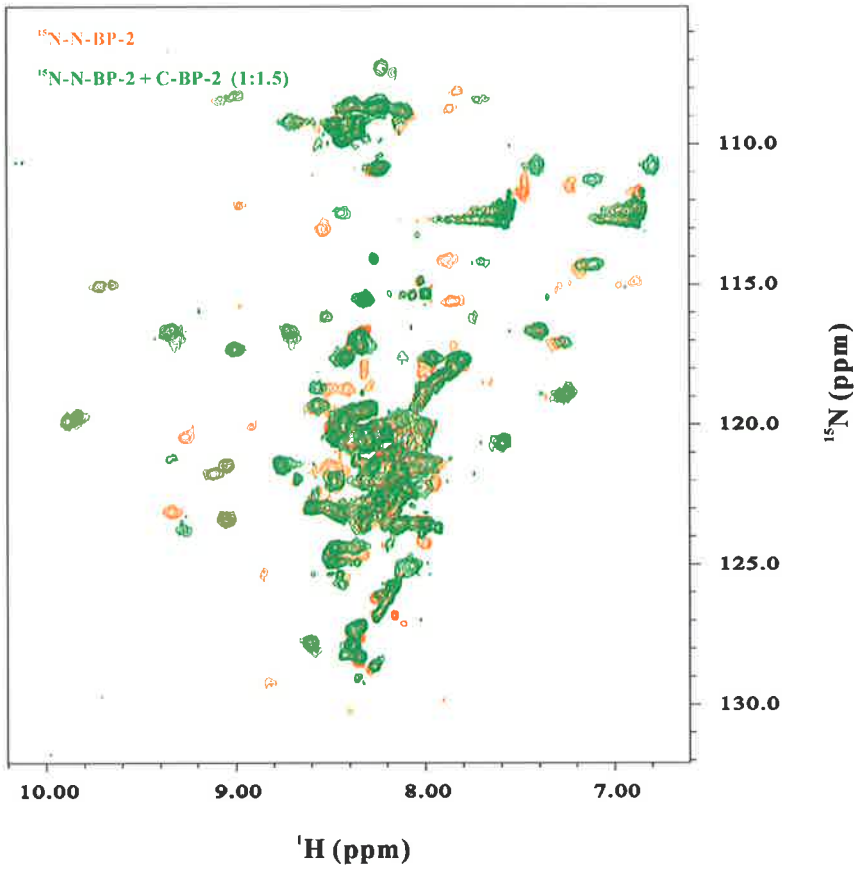


Figure 7.10 Comparison of ^{15}N-labelled N-BP-2 in the absence and presence of C-BP-2

Overlay of the ^{1}H-^{15}N HSQC spectra of 0.07 mM ^{15}N-labelled N-BP-2 in the absence and presence of unlabelled C-BP-2 at N-BP-2:C-BP-2 molar ratios of 1:1.5. The samples were in 95% H_2O/5% $2H_2O$ containing 10 mM sodium acetate and 0.02% (w/v) sodium azide at pH 7.0. Spectra were recorded at 500 MHz and 298 K. Addition of C-BP-2 to ^{15}N-labelled N-BP-2 in the presence of 150 mM NaCl at pH 6.0 or 7.4 produced similar results (not shown).

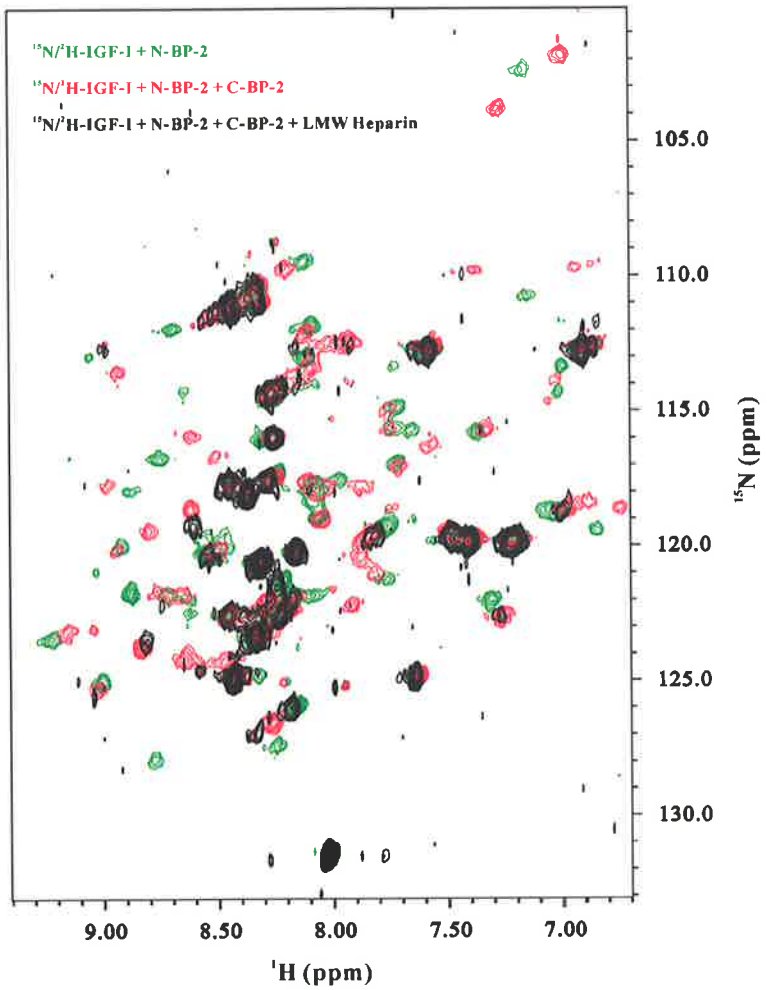


Figure 7.11 Effect of LMW heparin on IGF-I·N-BP-2·C-BP-2 ternary complex

Overlay of the ^{1}H-^{15}N HSQC spectra of ^{15}N/^{2}H-IGF (0.07 mM) in IGF-I·N-BP-2 binary complex (red), IGF-I·N-BP-2·C-BP-2 ternary complex (green), and IGF-I·N-BP-2·C-BP-2 in the presence of 2 mM LMW heparin (black). The samples were in 95% H_2O/5% $2H_2O$ containing 10 mM sodium acetate and 0.02% (w/v) sodium azide at pH 6.0. The sample with LMW heparin also contained 150 mM NaCl. Spectra were recorded at 500 MHz and 298 K.

7.4 Discussion

7.4.1 Cooperativity between N- and C-BP-2 in IGF binding

In this Chapter, the NMR chemical shift perturbation mapping was used to analyse the molecular interactions among N-BP-2, C-BP-2, and IGFs. Whilst others have carried out similar investigations on IGF binding by mini-N-BP-5 (Kalus et al., 1998), C-BP-6 (Headey et al., 2004a; Headey et al., 2004b) and full-length IGFBP-2 (Carrick et al., 2005), the availability of unlabelled and ^{15}N-labelled N- and C-domain fragments and IGF ligands in this thesis enabled a more comprehensive study, and indeed provided some novel information.

In addition to the identification of residues involved in binding interactions, titration with NMR provides an estimation of the binding affinity (Zuiderweg, 2002). As reviewed in Chapter 5, based on how the resonances change the interaction can be referred to as being in slow, intermediate, or fast chemical exchange (from high affinity to low affinity). Therefore, the results in this Chapter provided a third measurement of IGF binding by N-BP-2 and C-BP-2, in addition to the cross-linking and BIAcore analyses described in Chapter 3. This was particularly valuable, because there were some discrepancies between the cross-linking and BIAcore results (Chapter 3). The cross-linking results indicated the IGF-II binding affinity of C-BP-2 was significantly lower (> 10-fold) than that of N-BP-2, and much lower (> 1000-fold) than that of IGFBP-2. In contrast, the IGF-II binding affinity of C-BP-2 derived from BIAcore data was close to that of N-BP-2, whereas its IGF-I binding affinity was higher than that of N-BP-2. As discussed in 3.5.3, technical differences might have also contributed to a similar discrepancy in the literature. IGFBP C-domain fragments were reported to have as little as 10-fold to as large as 2000-fold lower IGF binding affinity compared to their full-length counterparts (Ho and Baxter, 1997; Carrick et al., 2001; Vorwerk et al., 2002; Payet et al., 2003; Headey et al., 2004c; Mark et al., 2005). Noticeably, in some reports the IGF binding affinity of C-domains was not detectable (Kalus et al., 1998; Qin et al., 1998; Fernandez-Tornero et al., 2005; Siwanowicz et al., 2005).

Titrations monitored by NMR showed that the binding of C-BP-2 or Large-C-BP-2 to both IGF-I and IGF-II was in intermediate exchange, and was weaker than the binding of N-BP-2 to IGFs, which was in slow exchange. Experiments using both ^{15}N-

labelled C-BP-2 and ^{15}N-labelled IGFs supported this conclusion. The results appeared to be more consistent with the cross-linking (Chapter 3), and were very similar to the IGF-II·C-BP-6 interaction reported by Headey and co-workers (Headey et al., 2004a). Therefore, C-domain fragments appear to bind IGFs to a limited extent. An interesting question then arises: given that C-domain fragments have very weak IGF binding, why does loss of the C-domain cause a dramatic reduction in IGF binding by N-domain fragments? The results in this Chapter provide a possible answer to this question as follows: because the binding of C-BP-2 to the IGF·N-BP-2 binary complexes is significantly stronger than its binding to free IGF, there is cooperativity between N-BP-2 and C-BP-2, and therefore the "true" IGF binding capacity of the C-domain in a full-length protein may not be fully revealed by studying the IGF binding of the C-domain alone.

Previous studies of whether there is cooperativity between the N- and C-domains of IGFBPs in IGF binding were limited and controversial (Bach et al., 2005). Payet and co-workers demonstrated that co-incubation of the IGFBP-3 N- and C-domain fragments resulted in high IGF binding affinities that were only 4-fold (for IGF-II) and 13-fold (for IGF-I) lower than that of full-length IGFBP-3 (Payet et al., 2003). Since in that study the N-BP-3 fragment bound IGFs poorly and the C-BP-3 fragment had 100-fold (IGF-I) and 50-fold (IGF-II) lower affinities than IGFBP-3, the high affinities of the (N-BP-3 + C-BP-3) clearly indicated a strong cooperative effect of the two fragments in IGF binding (Payet et al., 2003). Furthermore, (N-BP-3 + C-BP-3) inhibited IGF-stimulated DNA synthesis and IGF-II binding to IGF-2R far more effectively than individual fragments (Payet et al., 2003). Interestingly, no interaction between N-BP-3 and C-BP-3 was detected unless IGF was also present (Payet et al., 2003). A similar synergistic effect was also seen in a study of N- and C-domain fragments of IGFBP-4 (Siwanowicz et al., 2005). In that case, C-BP-4 binding to IGF-I was not detectable using isothermal titration calorimetry (ITC) or NMR, but its presence increased the affinities of N-BP-4 to IGF-I by ITC experiments (Siwanowicz et al., 2005). On the other hand, BIAcore analysis did not detect enhancement of IGF binding by co-incubation of the N- and C-domains of IGFBP-2 (Carrick et al., 2001) or those of IGFBP-6 (Headey et al., 2004c), and enhanced inhibition of IGF action was not seen in the latter study (Headey et al., 2004c).

Here, such cooperativity was clearly evident. Based on the NMR data, it appeared that the presence of the N- or C-domain increased the IGF binding by the other domain

through a conformational change of the IGF molecule, whereas direct interaction between N- and C-domains may further enhance the interaction (see below). IGF residues had essentially identical chemical environments in the IGF·N-BP-2·C-BP-2 ternary complexes and in the IGF·IGFBP-2 complexes as seen when comparing the spectra of IGFs in the presence of both individual domains with the IGF spectra when complexed with intact IGFBP-2 (Carrick et al., 2005). The NMR data showed that both IGFBP-2 and (N-BP-2 + C-BP-2) bound IGFs with high affinities that were in the slow exchange regime. However, the binding constants of these strong interactions could not be determined in a quantitative way by NMR (Zuiderweg, 2002), because, while quantitative measures can be obtained if the *K*d is within an order of magnitude of the concentration of the studied species, the protein concentrations required for the experiments often exceed this range (Zuiderweg, 2002). It is generally believed that covalently linking the N- and C-domain binding sites generates higher IGF binding affinity of IGFBPs than (N-BP + C-BP), whereas limited proteolytic cleavages in the linker domain reduce binding affinity and release bound IGFs. While the IGF binding affinity difference between IGFBP-2 and N-BP-2 + C-BP-2 could not be measured in the current NMR experiments, Payet and co-workers showed that N-BP-3 + C-BP-3 bound IGFs with high affinities, which were 4- and 13-fold lower than those of IGFBP-3 for IGF-I and IGF-II, respectively, and retained the ability to block IGF·IGF-1R interactions (Payet et al., 2003).

7.4.2 Conformational change of IGF-I

One possible mechanism underlying the cooperativity is that the IGF molecules and/or the N-and C-domains may undergo conformational changes upon binding to one partner and adopt a higher affinity conformation for the other partner. The NMR data presented here, in conjunction with the crystal structures of IGF-I·mini-N-BP-5 (Zeslawski et al., 2001), IGF-I·N-BP-4 (Siwanowicz et al., 2005), and IGF-I·N-BP-4·C-BP-4 (Sitar et al., 2006) complexes published recently, provide evidence for such conformational changes.

In chemical shift perturbation mapping, both a direct contact with the binding partner and a conformational change induced by the binding can cause perturbation of the resonances (Zuiderweg, 2002). When unlabelled N-BP-2 was titrated into ^{15}N-

labelled IGF-I, it caused chemical shift changes of a large number of IGF-I cross-peaks (Figure 7.4). Although the chemical shifts of IGF-I in the IGF-I·N-BP-2 complex form could not be completely assigned because of the very broad resonances in 3D NHCA spectra on the $^{15}N/^{13}C$-IGF-I·N-BP-2 sample, it was clear that cross-peaks that did not change were Gly30-Arg36 and Lys65-Ala70, which are located in either the C or D domain flexible regions. It was also evident that cross-peaks of some residues in the A1 helix (Gly42-Phe49) were perturbed significantly. This was interesting, because in the IGF-I·N-BP structures (Zeslawski et al., 2001; Siwanowicz et al., 2005; Sitar et al., 2006) IGF-I residues Gly42-Phe49 were not close to the IGF-I·N-BP binding interfaces. In Figure 7.12, crystal structures of IGF-I in the "free" forms (bound to detergent molecules) (Vajdos et al., 2001; Brzozowski et al., 2002), and in the IGF-I·N-BP complexes (Zeslawski et al., 2001; Siwanowicz et al., 2005) are superimposed over backbone heavy atoms in three helices. It can be seen that the spatial organization of the helices and the structure of residues Gly42-Glu46 did not significantly change upon binding to N-BP-4. However, IGF-I residues Phe49-Leu54 as well as the Phe49 aromatic ring adopted a significantly different conformation in the IGF-I·N-BP-4 and IGF-I·N-BP-5 complexes compared to that in the free form. We could therefore speculate that the conformational changes in this region might be induced by the binding of the Leu54 sidechain into a hydrophobic pocket of the N-BPs (Zeslawski et al., 2001; Siwanowicz et al., 2005), leading to a conformational change in the neighbouring Phe49 sidechain.

It is possible that the significant chemical shift perturbations of IGF-I residues Gly42-Glu46 upon binding to N-BP-2 were caused by the conformational change of the Phe49 aromatic ring through the so-called "ring current effect" (Haigh and Mallion, 1980). An aromatic ring produces a small magnetic field and thus affects the atoms close to it. It will "shield" the atoms from the outside magnetic field if the atoms are above or below the ring, resulting in upfield chemical shifts, and will "de-shield" the atoms if they are within the ring plane, resulting in downfield chemical shifts (Haigh and Mallion, 1980). Therefore, residues can have significant chemical shift perturbations if a nearby aromatic ring changes its orientation. Previous NMR studies have shown significant internal flexibility of IGF-I (Cooke et al., 1991; Sato et al., 1993; Laajoki et al., 2000) and IGF-II (Terasawa et al., 1994; Torres et al., 1995). The conformations of Phe49-Leu54 and the Phe49 sidechain in IGF-I, as well as the equivalent residues in IGF-II in these NMR structures were not well defined, probably

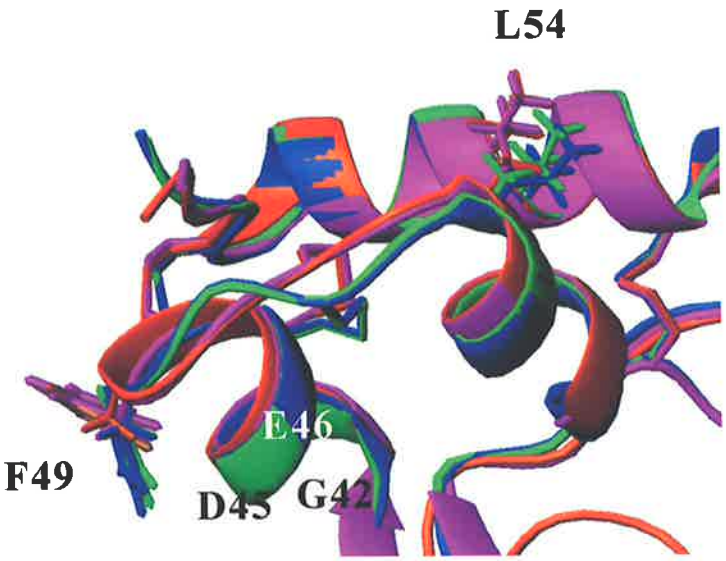


Figure 7.12 Superimposition of IGF-I crystal structures showing conformational changes upon binding to N-domain

Crystal structures of IGF-I in the "free" form (bound to detergent molecules, PDB 1GZR (blue) (Brzozowski et al., 2002) and PDB 1IMX (green) (Vajdos et al., 2001)), and in the IGF-I·mini-N-BP-5 complex (PDB 1H59 (orange) (Zeslawski et al., 2001)), and in the IGF-I·N-BP-4 complex (PDB 1WQJ (pink) (Siwanowicz et al., 2005)) are superimposed over backbone heavy atoms in three helices (Ala8-Cys18, Gly42-Cys48, Leu54-Cys61). Sidechains of Phe49 and Leu54 are shown. Using the IGF-I structure in the IGF-I·N-BP-4·C-BP-4 complex (Sitar et al., 2006) produced similar results (not shown).

resulting from flexibility in this region together with low NMR spectral quality due to aggregation. In NMR binding experiments, the "free" set and "bound" set of cross-peaks represented the average conformations of these two states in solution. Thus, the NMR data suggest that the conformational differences around Phe49-Leu45 between free and N-BP-bound IGF-I seen in the crystal structures may exist in solution.

If IGF-I has multiple interconverting conformations in solution, and the most favourable or average conformation around residues Phe49-Leu54 in the free form differs significantly from that in the IGFBP-bound form, this would be energetically unfavourable for the free IGF-I. The binding interaction between IGFBP N-domain and IGF-I is strong, and is sufficient to either induce a conformational change of IGF-I or selectively stabilize the IGF-I in the bound conformation. On the other hand, interaction between IGFBP C-domain and the free IGF-I is weak, which might be due to low complementarity between the free form of Phe49 sitechain and the IGF binding site on the C-domain. It is known that IGF-I residue Phe49 and the IGF-II equivalent Phe48 are very important for IGFBP binding (Clemmons et al., 1992; Bach et al., 1993; Oh et al., 1993; Dubaquie and Lowman, 1999). Thus, IGF·C-BP interactions may be too weak to either induce the conformational change or stabilize the IGF in the bound conformation. When IGF-I is bound to the N-domain, its C-domain binding site, and especially the Phe49 sidechain, is stabilized in a conformation that binds C-domain with high affinity. In return, binding of C-domain to this site may further stabilize the bound conformation of IGF-I and thus reciprocally enhance the interaction between IGF-I and the N-domain. Interestingly, it seems that the Phe49-Leu54 region of IGF-I mediates the cooperativity between the N- and C-domains, with the N-domain interacting at the Leu54 end while the C-domain interacting at the Phe49 end. This may also apply to IGF-II, since similar NMR results were observed.

7.4.3 Inter-domain interaction between N-BP-2 and C-BP-2

If C-BP-2 binds N-BP-2, this interaction could also contribute to the observed higher binding affinity of C-BP-2 for IGF·N-BP-2 complex than that for IGFs alone. While N-BP-2 and C-BP-2 may undergo local conformational change upon binding to IGF that enhance their binding to one another, the NMR results using both ^{15}N-labelled C-BP-2 and ^{15}N-labelled N-BP-2 indicated that N-BP-2 and C-BP-2 interact with each

other in the absence of IGFs. The inter-domain interaction between IGFBP N- and C-domains in the absence of IGF has not been detected previously.

Addition of unlabelled N-BP-2 caused broadening of a considerable number of C-BP-2 ^{1}H-^{15}N cross-peaks. The most affected residues are located in the end of the helix (Arg206), loop I (Gly212, His216), the second strand (Lys234, Lys237, Met238), and the second half of loop II (L240-Gly245) (Figure 7.13 (b)). The N-BP-2 binding site on C-BP-2 identified by NMR largely correlated with the C-BP-4 surface contacting N-BP-4 in the IGF-I· N-BP-4·C-BP-4 ternary complex structure reported recently (Sitar et al., 2006) (Figure 7.13 (c)). In this structure, N-BP-4 interacts with the loop I and loop II regions of C-BP-4 (Sitar et al., 2006). Therefore, the direct interaction between N-BP-2 and C-BP-2, as well as the N-BP-2 binding site on C-BP-2 revealed in this thesis appear to be biologically relevant.

While previous studies suggested that the N- and C-domains were close to or each other or even in contact in the IGF·IGFBP complex as the N- and C-domain binding sites on IGF-II are adjacent (Headey et al., 2004b), and this was supported by the recent ternary complex structure (Siwanowicz et al., 2005; Sitar et al., 2006), little was known regarding the spatial organization of the N- and C-domains in the free form of IGFBPs. Since inter-domain interaction was detected in the absence of IGFs, the N- and C-domains may be contacting each other in the free IGFBP-2 form, rather than being loosely connected merely by the linker domain. However, this interaction does not seem to be strong or stable, as was not detected in other assays (Payet et al., 2003), and would be unlikely to hinder IGFs access to both the N- and C-domain binding sites. The significant flexibility in the inter-domain binding sites, especially loop I and loop II of the C-domain, would enable large conformational changes and domain movements upon IGF binding. Future structural and biophysical studies will be very valuable.

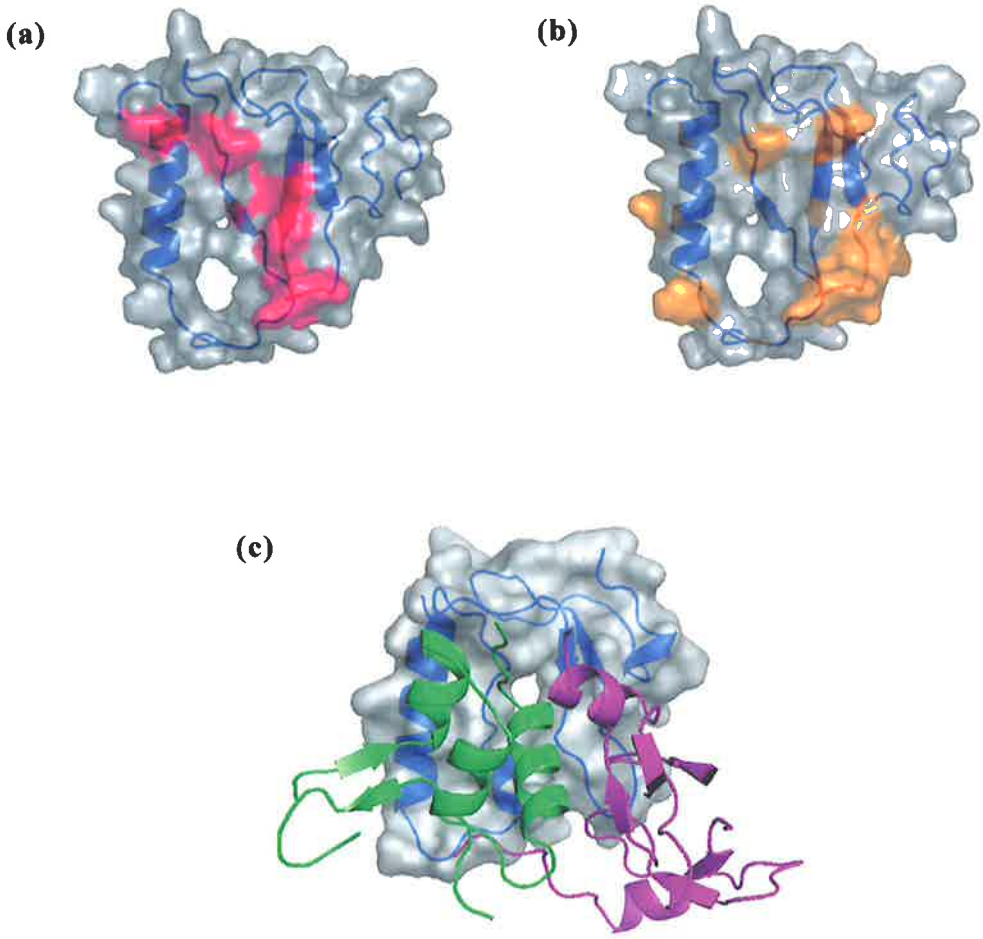


Figure 7.13 IGF and N-BP-2 binding sites on C-BP-2

(a) and (b) Transparent surface models of C-BP-2 showing IGF (a) or N-BP-2 (b) binding sites, and with ribbon representation (blue) inside. Residues most affected upon titration of IGF-I or IGF-II (Figure 7.3) are coloured pink (a). Residues most affected upon titration of N-BP-2 (Figure 7.9) are coloured orange (b). Parts of the unstructured regions (residues 183-186 and 280-289) are excluded for clarity. (c) Crystal structure of IGF-I·N-BP-4·C-BP-4 ternary complex (PDB 2DSR) (Sitar et al., 2006). C-BP-4 is shown as transparent surface model with ribbon representation (blue) inside. N-BP-4 (pink) and IGF-I (green) are shown as ribbon models. This figure was prepared using PyMOL (DeLano Scientific). C-BP-2 and C-BP-4 structures are shown in equivalent orientations.

7.4.4 Binding sites and effect of heparin binding

In addition to the N-BP-2 binding site, the IGF binding sites on C-BP-2 were identified in this Chapter. In the absence of N-BP-2, addition of IGF-I and IGF-II induced very similar peak broadening of ^{15}N-labelled C-BP-2 (Figure 7.3). ^{1}H-^{15}N cross-peaks of C-BP-2 residues that disappeared or were significantly broadened at C-BP-2:IGF-I or C-BP-2:IGF-II ratios of 1:1.5 are mapped onto C-BP-2 surface in Figure 7.13 (a). These residues are located in the helix (Gln192), the first (His221, Asn224, Cys225), and the second (Cys236-Met238) strands, and the first half of loop II (Ser239-Gly242) (Figure 7.13 (a)). Indeed, these residues correlated very well with the C-BP-6 residues that disappeared upon binding to IGF-II (Headey et al., 2004a), although NMR experiments for the IGF-I·C-BP-6 interaction were not carried out because the binding was too weak. Using mutagenesis, Headey and co-workers confirmed that IGFBP-6 Asn163 (equivalent to C-BP-2 Asn224) contributes to both IGF-I and IGF-II binding, whereas Ser176, Ser177, and Gln178 (equivalent to C-BP-2 Met238-Leu40) contribute to the IGF-II binding preference (Headey et al., 2004a). Previously, progressive deletions by Forbes and co-workers highlighted the importance of bovine IGFBP-2 residues Lys220-Asn236 (equivalent to C-BP-2 Lys227-Asn241) in IGF binding (Forbes et al., 1998), which was consistent with the NMR results. Several mutagenesis studies have reported that IGFBP-3 and IGFBP-5 residues equivalent to C-BP-2 Gly229 and Gln235 are important in IGF binding (Bramani et al., 1999; Song et al., 2000; Shand et al., 2003; Yan et al., 2004; Allan et al., 2006). However, since the cross-peaks of Gly229 and Gln235 were not significantly broadened upon binding to IGFs, and the sidechain of Gln235 is on the opposite face of C-BP-2 to the IGF binding site, it appears that moderate reductions in IGF binding by these mutations might have been due to structural perturbations. In the IGF-I·N-BP-4·C-BP-4 ternary complex structure (Sitar et al., 2006), IGF-I bound to a C-BP-4 surface formed by one face of the first helix, loop I, the second helix (equivalent to a region of loop I in C-BP-2), the second strand, and first half of the loop II, which was thus consistent with the NMR data (Figure 7.X (c)).

Binding of IGF·N-BP-2 binary complexes to ^{15}N-labelled C-BP-2 had very profound effects on its ^{1}H-^{15}N HSQC spectra, as cross-peaks of most C-BP-2 residues except the N- and C-termini shifted to some extent. These results also suggested that C-

BP-2 has very similar structure and chemical environments in both the IGF-I·N-BP-2·C-BP-2 and IGF-II·N-BP-2·C-BP-2 ternary complexes. However, a detailed analysis of the chemical shift changes was hampered because C-BP-2 in the ternary complexes gave very broad ^{1}H-^{15}N cross-peaks and resonances could not be assigned. Likewise, while the ^{15}N-labelled IGF experiments revealed the cooperativity between N- and C-domains and suggested a conformational change of IGF, future work is required to completely assign the chemical shift changes, and thereby to further analyse the binding events, as well as to identify IGF residues that were perturbed upon binding of the IGF·N-BP-2 binary complexes to C-BP-2. Furthermore, in the NMR experiments Large-C-BP-2 fragment did not show any extra interaction compared to C-BP-2, indicating the lack of stable structure formed by the additional linker domain region.

Having established that (N-BP-2 + C-BP-2) mimicked IGFBP-2 in IGF binding, it was of interest to study the effect of heparin binding on IGF·IGFBP interaction using these domain fragments. Preliminary results indicated that the IGF-I·N-BP-2·C-BP-2 ternary complex bound to low molecular weight heparin, and the interaction appeared to result in the formation of larger complexes rather than dissociation of the IGF-I·N-BP-2·C-BP-2 complex, as IGF-I ^{1}H-^{15}N cross-peaks were significantly broadened but yet still represented the ternary complex form. Since the peak broadening was more serious at pH 5.5 than 6.0, it is likely that heparin bound to the pH-dependent heparin binding site on C-BP-2 identified in Chapter 5. Based on the locations of IGF and N-BP-2 binding sites and the IGF-I·N-BP-4·C-BP-C structure (Sitar et al., 2006), the C-BP-2 heparin binding site is likely accessible in the IGF-I·N-BP-2·C-BP-2 ternary complex, and heparin binding to this site may not significantly interfere with IGF binding. As reviewed in Chapter 5, previous data regarding the heparin binding by IGFBP-2 and its effect on IGF binding were controversial. Arai and co-workers reported that heparin did not inhibit IGF·IGFBP-2 binding (Arai et al., 1994b), rather, heparin only bound IGFBP-2 when it was in IGF·IGFBP-2 complex form (Arai et al., 1996a), whereas Russo and co-workers reported that binding of IGFBP-2 to chondroitin-6-sulfate decreased the binding affinity of IGFBP-2 for IGF-I approximately three-fold (Russo et al., 1997). It is worth noting that small to moderate change in binding constants may not be detected in the current NMR experiments, if the binding interaction remained in the same chemical exchange regime. More investigations are warranted to establish whether IGF binding and heparin binding by IGFBP-2 enhance or (slightly) compete with each other. In Chapter 5, it was suggested

that IGFBP-2 or IGF·IGFBP-2 complexes might preferentially bind to glycosaminoglycans in acidic tumour ECM, but not in neutral, normal tissue ECM. Results in this Chapter suggest that binding of IGF·IGFBP-2 complexes to glycosaminoglycans may not automatically result in IGF release. Nevertheless, many tumours overexpress IGFBP-cleaving proteases (Bunn and Fowlkes, 2003), thus unbound IGFs and RGD-containing IGFBP-2 C-domain fragments may be accumulated in tumour ECM by enhanced glycosaminoglycans binding and limited proteolysis. In contrast, heparin binding by IGFBP-5 reduced its IGF-I binding by 17-fold (Arai et al., 1994b), and strong competition between IGF binding and heparin binding by IGFBP-3 and -5 were demonstrated by both BIAcore and solution binding assays (Beattie et al., 2005). It will be of interest to further study the heparin binding using ^{15}N-labelled C-BP-2 and N-BP-2, as well as to compare the heparin binding by different IGFBPs using NMR.

Blank page

Chapter 8

Final Discussion

8.1 Summary of Findings

As discussed in Chapter 1, IGFBPs play important roles in regulating IGF actions, as well as having IGF-independent functions. The C-domains of IGFBPs are indispensable for high affinity IGF binding, and are responsible for a wide range of other molecular interactions. However, at the commencement of this study there was limited structural and functional information on these important domains.

The first goal of this thesis was to prepare unlabelled and isotope-labelled IGFBP-2 N- and C-domain fragments in sufficient amounts (~10 mg) and of sufficient purity (>95 %) to be used in NMR studies and other analyses. In Chapter 2, expression vectors were generated for the production of full-length IGFBP-2 and its N- and C-domain fragments in *E. coli* cells. 3C protease and TEV protease cleavage sites were incorporated into these constructs so that these proteases could be used in the fusion tag removal step. In Chapter 3, unlabelled, ^{15}N-labelled and ^{15}N/^{13}C-labelled $^{183\text{-}289}$IGFBP-2 (C-BP-2) fragments, unlabelled $^{141\text{-}289}$IGFBP-2 (Large-C-BP-2) fragment, and unlabelled and ^{15}N-labelled $^{1\text{-}138}$IGFBP-2 (N-BP-2) fragments were prepared. 3C protease and TEV protease were used to remove the thioredoxin-His_6 tag fusion partner. Unlike the enterokinase used previously (Carrick, 2002, PhD thesis, The University of Adelaide), 3C protease and TEV protease cleavages were highly specific, and therefore the loss of recombinant protein due to non-specific proteolytic degradation was avoided. However, the requirement for the reducing agent DTT to ensure high activities may compromise the future use of 3C and TEV proteases in large-scale preparation of recombinant IGFBP-2, which has some disulfide bonds that are sensitive to DTT. Mass spectroscopy and N-terminal sequencing data confirmed the identity and integrity of these fragments.

High-resolution three-dimensional structures of either a full-length IGFBP or an IGFBP C-domain had not been determined previously. In Chapter 4, the solution structure of C-BP-2 was determined using heteronuclear multidimensional NMR spectroscopy. C-BP-2 is the largest thyroglobulin type 1 structure determined so far,

consisting of an α-helix, a three-stranded anti-parallel β-sheet and three flexible loops. Compared with other IGFBP C-domain structures (C-BP-6 (Headey et al., 2004a), C-BP-1 (Sala et al., 2005), and C-BP-4 (Sitar et al., 2006)) reported during the course of this thesis, differences that may affect IGF binding and underlie other functional differences were found. C-BP-2 has a longer disordered loop I, and an extended C-terminal tail, which is unstructured and very mobile. The length of the helix is identical to that of C-BP-6 but shorter than that of C-BP-1. On the other hand, the length of loop II and the dynamic properties of C-BP-2 are similar to those of C-BP-6, although their compositions are rather different. The considerable flexibility of loop II may be required for high-affinity IGF binding. The variety and flexibility of loop II and its flanking residues seem to influence the IGF-binding affinities of the six IGFBPs, and also bestow other molecular interaction abilities on thyroglobulin type 1 domains. An RGD motif is located in a solvent-exposed turn in the loop III and has considerable mobility, as revealed by backbone amide dynamic measurements. Its structure and dynamic properties are thus consistent with functional RGD motifs in many other proteins.

There has been increasing evidence that protein dynamics in solution may be related to protein function, especially protein-protein interactions. In Chapter 4, the backbone dynamics of C-BP-2 were analysed based on ^{15}N relaxation parameters (R_1, R_2 and steady-state ^{15}N-{^1H} NOE). C-BP-2 has considerably higher backbone ^{15}N R_2 values than C-BP-6 (Yao et al., 2004), and fitting ^{15}N relaxation parameters to an axially symmetric rotational diffusion tensor model was unsuccessful. Thus, a full analysis using the conventional *Modelfree* approach was hampered. Current ^{15}N relaxation and translational diffusion data of C-BP-2 at different concentrations indicated that self-association under the experimental conditions is unlikely. The anomalously higher R_2 values and the failure in fitting the relaxation parameters are probably a consequence of motions of the core of C-BP-2 that cannot be described by a simple anisotropic tumbling model with a unique diffusion tensor, as a result of the presence of a longer and very flexible C-terminal tail together with a longer disordered loop compared to C-BP-6. Reduced spectral density mapping analysis showed that C-BP-2 possesses significant rapid motion in the loops and termini, and may also undergo slower conformational or chemical exchange in the structured core and loop II.

The IGF binding abilities of the N- and C-domain fragments were investigated using different methods. In Chapter 3, BIAcore and cross-linking experiments were

carried out. N-BP-2 has a lower IGF binding affinity than full-length IGFBP-2, resulting from a significantly faster dissociation rate, which was consistent with previous findings (Carrick et al., 2001).While both cross-linking and BIAcore results indicated that the isolated C-BP-2 binds IGFs to only a limited degree, the differences in IGF binding affinities between C-BP-2 and N-BP-2 or full-length IGFBP-2 appeared to be larger in the cross-linking results than in the BIAcore results. On the other hand, Large-C-BP-2 had a significantly higher BIAcore response than C-BP-2, and its IGF binding affinities derived from BIAcore data were higher than N-BP-2 and C-BP-2.

Molecular interactions among N- and C-domains and IGF, as well as heparin binding by C-BP-2, were analysed in detail using NMR chemical shift perturbation mapping. The resonance assignments and the solution structure of C-BP-2 described in Chapter 4 provided the basis for analysis of its interactions using ^{15}N-labelled C-BP-2, whereas in Chapter 6, experimental conditions for ^{15}N-labelled IGF-I, IGF-II, and N-BP-2 were optimised. Acquiring 2D ^{1}H-^{15}N HSQC spectra at lower protein concentration reduced the aggregation of IGF-I, and to a lesser extent, of IGF-II, and thus improved the spectral quality at pH conditions closer to physiological pH. A more complete set of assignments for IGF-I ^{1}H-^{15}N HSQC spectra was thus obtained.

In Chapter 7, titrations with NMR showed that binding of C-BP-2 or Large-C-BP-2 to IGFs was in the intermediate exchange regime and was of lower affinity than the binding of N-BP-2 to IGFs, which was in the slow exchange regime. The NMR results were more consistent with the cross-linking results than the BIAcore results. NMR experiments did not detect extra interactions by Large-C-BP-2 compared with C-BP-2 that would be contributed by the additional linker domain region. Therefore, the discrepancy between the BIAcore and cross-linking or NMR results, and that between two binding orientations in BIAcore experiments, suggested that other factors contributed to the apparent results, and the kinetic data derived using current models may not be able to accurately describe the interactions between the IGFBP-2 C-domain and IGFs.

Binding of C-BP-2 or Large-C-BP-2 to the IGF·N-BP-2 binary complexes was significantly stronger than the binding of C-BP-2 to IGFs alone, switching from intermediate exchange to slow exchange, indicating that there is cooperativity between N-BP-2 and C-BP-2 in IGF binding. A comparison with the spectra of IGFs in complex with IGFBP-2 (Carrick et al., 2005) showed that IGF residues had essentially identical chemical environments in IGF·N-BP-2·C-BP-2 ternary complexes compared to the

IGF·IGFBP-2 complexes. Moreover, two possible mechanisms underlying the cooperativity between the N- and C-domains were identified. The NMR data, in conjunction with the recently reported crystal structures (Zeslawski et al., 2001; Siwanowicz et al., 2005; Sitar et al., 2006) suggested that a significant conformational change of the Phe49-Leu54 region and the sidechain aromatic ring of Phe49 in IGF-I exists in solution upon its binding to the N-domain, which may increase the binding of this surface to the C-domain IGF binding site. The binding of C-domain to this site may then further stabilize the bound conformation of IGF-I and thus may enhance the interaction between IGF-I and the N-domain. Furthermore, the NMR results using both ^{15}N-labelled C-BP-2 and ^{15}N-labelled N-BP-2 provided evidence for a direct interaction between N-BP-2 and C-BP-2 in the absence of IGFs, which has not been detected previously. Such inter-domain interaction may further enhance the IGF binding by (N-BP-2 + C-BP-2) compared with that by N-BP-2 or C-BP-2 alone.

The IGF and N-BP-2 binding sites on C-BP-2 were identified. The most affected C-BP-2 residues upon binding to IGFs were located in the helix (Gln192), the first (His221, Asn224, Cys225), and the second (Cys236-Met238) strands, and the first half of loop II (Ser239-Gly242). No significant difference was observed between the IGF-I and IGF-II binding sites on C-BP-2. The most affected C-BP-2 residues upon binding to N-BP-2 were located in the end of the helix (Arg206), loop I (Gly212, His216), the second strand (Lys234, Lys237, Met238), and the second half of loop II (L240-Gly245). These results were consistent with the IGF-II binding site on C-BP-6 (Headey et al., 2004a) and the recent IGF-I·N-BP-4·C-BP-4 complex structure (Sitar et al., 2006).

Binding of IGFBPs to glycosaminoglycans and other ECM components is believed to be important in modulation of IGF actions, but previous data regarding the glycosaminoglycan binding of IGFBP-2 were contradictory. In Chapter 5, a heparin binding site of C-BP-2 was identified by NMR. The heparin binding site is a patch containing the β-turn connecting the first and second strands, part of the third strand, and the beginning of the C-terminal tail. Lys227, His228, Asn232, Leu233, Lys234 and His271 are proposed to be the primary heparin binding residues. Protonation of His271 and His228 seems to be important for the binding, which occurs at slightly acidic pH (6.0) and is more significant at pH 5.5, but is largely suppressed at pH 7.4. Possible preferential binding of IGFBP-2 and its C-domain fragments to glycosaminoglycans in the acidic extracellular matrix of tumours may be related to their roles in cancer. Similar experiments were carried out in Chapter 7 to check whether IGF-I and N-BP-2

interact with heparin. IGF-I might bind to heparin via two arginine residues (Arg36 and Arg37) in its flexible C-domain loop, but N-BP-2 did not bind heparin. A preliminary investigation of the effect of heparin binding on IGF binding was described in Chapter 7. Consistent with the locations of the heparin, IGF and N-BP-2 binding sites on C-BP-2, heparin bound to the IGF-I·N-BP-2·C-BP-2 ternary complex, and heparin binding did not significantly reduce IGF binding. These results provided a valuable basis on which to reconcile some of the discrepancies in the literature.

8.2 Future Directions

More IGF·IGFBP (domain) complex structures are to be solved by crystallography or by NMR. The recent IGF-I·N-BP-4·C-BP-4 crystal structure (Sitar et al., 2006), together with a series of NMR studies on C-BP-6 (Headey et al., 2004a; Headey et al., 2004b; Yao et al., 2004) and C-BP-2 (this thesis), as well as the C-BP-1 structure (Sala et al., 2005), have greatly advanced our knowledge on the structure and function of IGFBP C-domains. This new information, when combined with the earlier findings on the IGFBP N-domains (Kalus et al., 1998; Zeslawski et al., 2001; Carrick et al., 2005; Siwanowicz et al., 2005), provides a more detailed picture of IGF·IGFBP interactions that was unavailable a few years ago. In particular, the relationship between the N- and C-domains in IGF binding has become clearer on the basis of results in this thesis and those of others (Payet et al., 2003). However, crystal or solution structures of the remaining IGFBP domains, especially in the IGF-complexed form, will be of great value for further understanding the structural and IGF binding differences among IGFBPs. One example is IGFBP-6, which has a marked IGF-II binding preference compared to other IGFBPs. To date, no three-dimensional structure of a full-length IGFBP either in the free form or in complex with IGFs has been reported. The expected disordered nature of the linker domains might have made crystallization of a full-length IGFBP difficult. While the long flexible regions may also be disadvantageous for NMR studies, IGFBPs, even in the IGF-complexed form, are of a size amenable to structure determination using state-of-the-art heteronuclear multidimensional NMR techniques. Nevertheless, since the results in this thesis did not suggest additional stable structure or interactions formed by the linker domain regions in N-BP-2 and C-BP-2 even in the

ternary complex form, future crystallography or NMR studies may be carried out using IGFBPs with shorter linker domains or even using separate N- and C-domain fragments without the linker regions. Importantly, the significant affinity differences of C-BP-2 for IGFs alone or for IGF·N-BP-2 binary complex may explain why successful crystallization of a IGF·C-BP binary complex has not been reported. IGF·C-BP binary complex structures will probably not be solved by NMR either, because the resonance of the interacting residues disappear form the spectra as a result of intermediate chemical exchange. However, the more stable ternary complexes are clearly suitable for NMR structure determination. Exclusions of the C-terminal tail from the C-BP-2 construct and the linker domain region from the N-BP-2 construct will significantly simplify their NMR spectra, and improve spectral quality by decreasing their tumbling correlation time in solution, resulting in sharper and stronger NMR peaks. The IGF·N-BP-2·C-BP-2 ternary complex would then be ready to solve. Nevertheless, IGFBP linker domains are involved in interactions with other proteins. For example, the linker domain of IGFBP-5 contains an ALS binding site (Twigg et al., 2000; Firth et al., 2001), in addition to the ALS binding site in the C-domain (Firth et al., 1998; Firth et al., 2001). In this case, co-crystallization of IGF + IGFBP-3 (or IGFBP-5) + ALS may stabilize the linker domain and yield a ~150 kDa IGF·IGFBP-3/5·ALS ternary complex that is present *in vivo*.

The binding interaction between IGFs and IGFBPs should also be analysed further. Whilst residues in the IGF·IGFBP binding sites are revealed by the recent complex structures and the NMR studies, more mutagenesis studies are needed to identify the residues that are most important for the binding. Since the results in this thesis suggest that the IGF interacting residues and the N-domain interacting residues on C-BP-2 both contribute to the IGF binding of full-length IGFBP-2, residues within both binding sites are to be analysed. On the basis of this thesis, IGFBP-2 residues suggested for site-directed mutagenesis are Gln192, Thr204-Arg206, His216, His220-Cys225, Lys234-Glu246, and adjacent residues. Alternatively, a phage-display approach that was previously carried out by Lucic and co-workers on IGFBP-2 (Lucic et al., 1998) can be applied to these regions. In addition to providing conclusive evidence for contributions of these residues, the mutagenesis studies may generate mutant IGFBP molecules that have increased IGF binding affinities. Structure-guided deletion studies can be conducted to obtain a smaller IGFBP molecule that retains high IGF binding affinity. Such modified IGFBP molecules are potentially very valuable for pharmaceutical

purposes. As discussed in Chapter 1, recombinant IGFBP proteins can be developed to block IGF·IGF-1R interactions in cancer therapy.

The cooperativity between IGFBP N- and C-domains in IGF binding is clearly worthy of future investigation. Experiments are to be performed to see whether the conformational change of IGF molecules or the interaction between N- and C-domains is a major contributor to this cooperativity. Because the ternary complex structure showed that the C-domain mainly interacts with the N-terminal "thumb" region of the N-domain (Sitar et al., 2006), using a smaller N-domain fragment equivalent to mini-N-BP-5 in NMR studies similar to this thesis might be able to dissect the contributions of these two mechanisms. Importantly, many protein-protein interactions are mediated by two or more domains in one protein. In particular, insulin and IGFs have multiple binding sites on different domains of their receptors. Therefore, future structural and biophysical studies on how IGFs bind to IGFBP N- and C-domains, which are smaller and easier to study compared to domains of insulin/IGF receptors, will assist in understanding other similar molecular interactions. Interestingly, conformational changes of the insulin molecule upon binding to its receptor have been reported (Hua and Weiss, 1991; Dong et al., 2003; Hua et al., 2006). Future analysis of the protein dynamics of IGF and IGFBP molecules and their functional relevance will be very valuable.

Finally, as described in Chapter 1, IGFBPs have been found recently to interact with many other molecules in addition to IGFs, and these interactions modulate IGF binding and/or confer IGF-independent actions. Most of these interactions are mediated by the C-domain of IGFBPs (Firth and Baxter, 2002; Bach et al., 2005). While others have used NMR spectroscopy to investigate the IGF·IGFBP interactions before or during the period in which the work described in this thesis was undertaken, the use of NMR to study other molecular interactions of IGFBPs and their effects on IGF binding has not been reported. The results presented here highlight its successful applications, and therefore this approach will be of great value for studying yet other interactions of IGFBP C-domains in the dramatically expanding list.

In conclusion, the findings described in this thesis make a substantial contribution to better understanding the structure and function of IGFBP-2 C-domain, and provide a valuable insight into the roles of IGFBP C-domains in both IGF binding and IGF-independent actions.

References

Allan, G. J., Tonner, E., Szymanowska, M., Shand, J. H., Kelly, S. M., Phillips, K., Clegg, R. A., Gow, I. F., Beattie, J., and Flint, D. J. (2006). Cumulative mutagenesis of the basic residues in the 201-218 region of insulin-like growth factor (IGF)-binding protein-5 results in progressive loss of both IGF-I binding and inhibition of IGF-I biological action. Endocrinology *147*, 338-349.

Amaar, Y. G., Thompson, G. R., Linkhart, T. A., Chen, S. T., Baylink, D. J., and Mohan, S. (2002). Insulin-like growth factor-binding protein 5 (IGFBP-5) interacts with a four and a half LIM protein 2 (FHL2). J Biol Chem *277*, 12053-12060.

Andress, D. L., Loop, S. M., Zapf, J., and Kiefer, M. C. (1993). Carboxy-truncated insulin-like growth factor binding protein-5 stimulates mitogenesis in osteoblast-like cells. Biochem Biophys Res Commun *195*, 25-30.

Arai, T., Arai, A., Busby, W. H., Jr., and Clemmons, D. R. (1994a). Glycosaminoglycans inhibit degradation of insulin-like growth factor-binding protein-5. Endocrinology *135*, 2358-2363.

Arai, T., Busby, W., Jr., and Clemmons, D. R. (1996a). Binding of insulin-like growth factor (IGF) I or II to IGF-binding protein-2 enables it to bind to heparin and extracellular matrix. Endocrinology *137*, 4571-4575.

Arai, T., Clarke, J., Parker, A., Busby, W., Jr., Nam, T., and Clemmons, D. R. (1996b). Substitution of specific amino acids in insulin-like growth factor (IGF) binding protein 5 alters heparin binding and its change in affinity for IGF-I response to heparin. J Biol Chem *271*, 6099-6106.

Arai, T., Parker, A., Busby, W., Jr., and Clemmons, D. R. (1994b). Heparin, heparan sulfate, and dermatan sulfate regulate formation of the insulin-like growth factor-I and insulin-like growth factor-binding protein complexes. J Biol Chem *269*, 20388-20393.

Bach, L. A. (2005). IGFBP-6 five years on; not so 'forgotten'? Growth Horm IGF Res *15*, 185-192.

Bach, L. A., Headey, S. J., and Norton, R. S. (2005). IGF-binding proteins--the pieces are falling into place. Trends Endocrinol Metab *16*, 228-234.

Bach, L. A., Hsieh, S., Sakano, K., Fujiwara, H., Perdue, J. F., and Rechler, M. M. (1993). Binding of mutants of human insulin-like growth factor II to insulin-like growth factor binding proteins 1-6. J Biol Chem *268*, 9246-9254.

Bach, L. A., and Parker, N. J. (2003). IGFBPs-gene and protein structure. In Insulin-Like Growth Factors, D. LeRoith, W. Zumkeller, and R. Baxter, eds. (Kluwer Academic), pp. 48-63.

Bach, L. A., Thotakura, N. R., and Rechler, M. M. (1992). Human insulin-like growth factor binding protein-6 is O-glycosylated. Biochem Biophys Res Commun *186*, 301-307.

Badinga, L., Song, S., Simmen, R. C., Clarke, J. B., Clemmons, D. R., and Simmen, F. A. (1999). Complex mediation of uterine endometrial epithelial cell growth by insulin-like growth factor-II (IGF-II) and IGF-binding protein-2. J Mol Endocrinol *23*, 277-285.

Bagley, C. J., May, B. L., Szabo, L., McNamara, P. J., Ross, M., Francis, G. L., Ballard, F. J., and Wallace, J. C. (1989). A key functional role for the insulin-like growth factor 1 N-terminal pentapeptide. Biochem J *259*, 665-671.

Baker, J., Liu, J. P., Robertson, E. J., and Efstratiadis, A. (1993). Role of insulin-like growth factors in embryonic and postnatal growth. Cell *75*, 73-82.

Ballard, F. J., Francis, G. L., Ross, M., Bagley, C. J., May, B., and Wallace, J. C. (1987). Natural and synthetic forms of insulin-like growth factor-1 (IGF-1) and the potent derivative, destripeptide IGF-1: biological activities and receptor binding. Biochem Biophys Res Commun *149*, 398-404.

Bartels, C., Xia, T. H., Billeter, M., Guntert, P., and Wuthrich, K. (1995). The program XEASY for computer-supported NMR spectral-analysis of biological macromolecules. J Biomol NMR *6*, 1-10.

Bartke, A. (2005). Minireview: role of the growth hormone/insulin-like growth factor system in mammalian aging. Endocrinology *146*, 3718-3723.

Baxter, R. C. (2000). Insulin-like growth factor (IGF)-binding proteins: interactions with IGFs and intrinsic bioactivities. Am J Physiol Endocrinol Metab *278*, E967-976.

Bayne, M. L., Applebaum, J., Chicchi, G. G., Miller, R. E., and Cascieri, M. A. (1990). The roles of tyrosines 24, 31, and 60 in the high affinity binding of insulin-like growth factor-I to the type 1 insulin-like growth factor receptor. J Biol Chem *265*, 15648-15652.

Bayne, M. L., Applebaum, J., Underwood, D., Chicchi, G. G., Green, B. G., Hayes, N. S., and Cascieri, M. A. (1989). The C region of human insulin-like growth factor (IGF) I is required for high affinity binding to the type 1 IGF receptor. J Biol Chem *264*, 11004-11008.

Beattie, J., Phillips, K., Shand, J. H., Szymanowska, M., Flint, D. J., and Allan, G. J. (2005). Molecular recognition characteristics in the insulin-like growth factor (IGF)-insulin-like growth factor binding protein -3/5 (IGFBP-3/5) heparin axis. J Mol Endocrinol *34*, 163-175.

Berman, H. M., Battistuz, T., Bhat, T. N., Bluhm, W. F., Bourne, P. E., Burkhardt, K., Feng, Z., Gilliland, G. L., Iype, L., Jain, S., *et al.* (2002). The Protein Data Bank. Acta Crystallogr D Biol Crystallogr *58*, 899-907.

Blackburn, A., Dressendorfer, R. A., Blum, W. F., Erhard, M., Brem, G., Strasburger, C. J., and 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*, 701-708.

Blundell, T. L., Bedarkar, S., and Humbel, R. E. (1983). Tertiary structures, receptor binding, and antigenicity of insulinlike growth factors. Fed Proc *42*, 2592-2597.

Blundell, T. L., Bedarkar, S., Rinderknecht, E., and Humbel, R. E. (1978). Insulin-like growth factor: a model for tertiary structure accounting for immunoreactivity and receptor binding. Proc Natl Acad Sci U S A *75*, 180-184.

Boisclair, Y. R., Brown, A. L., Casola, S., and Rechler, M. M. (1993). Three clustered Sp1 sites are required for efficient transcription of the TATA-less promoter of the gene for insulin-like growth factor-binding protein-2 from the rat. J Biol Chem *268*, 24892-24901.

Boisclair, Y. R., Rhoads, R. P., Ueki, I., Wang, J., and Ooi, G. T. (2001). The acid-labile subunit (ALS) of the 150 kDa IGF-binding protein complex: an important but forgotten component of the circulating IGF system. J Endocrinol *170*, 63-70.

Bourner, M. J., Busby, W. H., Jr., Siegel, N. R., Krivi, G. G., McCusker, R. H., and Clemmons, D. R. (1992). Cloning and sequence determination of bovine insulin-like growth factor binding protein-2 (IGFBP-2): comparison of its structural and functional properties with IGFBP-1. J Cell Biochem *48*, 215-226.

Bramani, S., Song, H., Beattie, J., Tonner, E., Flint, D. J., and Allan, G. J. (1999). Amino acids within the extracellular matrix (ECM) binding region (201-218) of rat insulin-like growth factor binding protein (IGFBP)-5 are important determinants in binding IGF-I. J Mol Endocrinol *23*, 117-123.

Braulke, T. (1999). Type-2 IGF receptor: a multi-ligand binding protein. Horm Metab Res *31*, 242-246.

Brinkman, A., Kortleve, D. J., Zwarthoff, E. C., and Drop, S. L. (1991). Mutations in the C-terminal part of insulin-like growth factor (IGF)-binding protein-1 result in dimer formation and loss of IGF binding capacity. Mol Endocrinol *5*, 987-994.

Brzozowski, A. M., Dodson, E. J., Dodson, G. G., Murshudov, G. N., Verma, C., Turkenburg, J. P., de Bree, F. M., and Dauter, Z. (2002). Structural origins of the functional divergence of human insulin-like growth factor-I and insulin. Biochemistry *41*, 9389-9397.

Buckway, C. K., Wilson, E. M., Ahlsen, M., Bang, P., Oh, Y., and Rosenfeld, R. G. (2001). Mutation of three critical amino acids of the N-terminal domain of IGF-binding protein-3 essential for high affinity IGF binding. J Clin Endocrinol Metab *86*, 4943-4950.

Bunn, R. C., and Fowlkes, J. L. (2003). Insulin-like growth factor binding protein proteolysis. Trends Endocrinol Metab *14*, 176-181.

Butt, A. J., Firth, S. M., King, M. A., and Baxter, R. C. (2000). Insulin-like growth factor-binding protein-3 modulates expression of Bax and Bcl-2 and potentiates p53-independent radiation-induced apoptosis in human breast cancer cells. J Biol Chem *275*, 39174-39181.

Byun, D., Mohan, S., Baylink, D. J., and Qin, X. (2001). Localization of the IGF binding domain and evaluation of the role of cysteine residues in IGF binding in IGF binding protein-4. J Endocrinol *169*, 135-143.

Byun, D., Mohan, S., Kim, C., Suh, K., Yoo, M., Lee, H., Baylink, D. J., and Qin, X. (2000). Studies on human pregnancy-induced insulin-like growth factor (IGF)-binding protein-4 proteases in serum: determination of IGF-II dependency and localization of cleavage site. J Clin Endocrinol Metab *85*, 373-381.

Campbell, P. G., Durham, S. K., Hayes, J. D., Suwanichkul, A., and Powell, D. R. (1999). Insulin-like growth factor-binding protein-3 binds fibrinogen and fibrin. J Biol Chem *274*, 30215-30221.

Campbell, P. G., Durham, S. K., Suwanichkul, A., Hayes, J. D., and Powell, D. R. (1998). Plasminogen binds the heparin-binding domain of insulin-like growth factor-binding protein-3. Am J Physiol *275*, E321-331.

Capila, I., and Linhardt, R. J. (2002). Heparin-protein interactions. Angew Chem Int Ed Engl *41*, 391-412.

Carr, P. A., Erickson, H. P., and Palmer, A. G., 3rd (1997). Backbone dynamics of homologous fibronectin type III cell adhesion domains from fibronectin and tenascin. Structure *5*, 949-959.

Carrick, F. E., Forbes, B. E., and Wallace, J. C. (2001). BIAcore analysis of bovine insulin-like growth factor (IGF)-binding protein-2 identifies major IGF binding site determinants in both the amino- and carboxyl-terminal domains. J Biol Chem *276*, 27120-27128.

Carrick, F. E., Hinds, M. G., McNeil, K. A., Wallace, J. C., Forbes, B. E., and Norton, R. S. (2005). Interaction of insulin-like growth factor (IGF)-I and -II with IGF binding protein-2: mapping the binding surfaces by nuclear magnetic resonance. J Mol Endocrinol *34*, 685-698.

Carrick, F. E., Wallace, J. C., and Forbes, B. E. (2002). The interaction of insulin-like growth factor (IGFs) with insulin-like growth factor binding proteins (IGFBPs): a review. Lett Pept Sci *8*, 147-153.

Cascieri, M. A., Chicchi, G. G., Applebaum, J., Green, B. G., Hayes, N. S., and Bayne, M. L. (1989). Structural analogs of human insulin-like growth factor (IGF) I with altered affinity for type 2 IGF receptors. J Biol Chem *264*, 2199-2202.

Cascieri, M. A., Chicchi, G. G., Applebaum, J., Hayes, N. S., Green, B. G., and Bayne, M. L. (1988). Mutants of human insulin-like growth factor I with reduced affinity for the type 1 insulin-like growth factor receptor. Biochemistry *27*, 3229-3233.

Cazals, V., Nabeyrat, E., Corroyer, S., de Keyzer, Y., and Clement, A. (1999). Role for NF-kappa B in mediating the effects of hyperoxia on IGF-binding protein 2 promoter activity in lung alveolar epithelial cells. Biochim Biophys Acta *1448*, 349-362.

Ceda, G. P., Fielder, P. J., Henzel, W. J., Louie, A., Donovan, S. M., Hoffman, A. R., and Rosenfeld, R. G. (1991). Differential effects of insulin-like growth factor (IGF)-I and IGF-II on the expression of IGF binding proteins (IGFBPs) in a rat neuroblastoma cell line: isolation and characterization of two forms of IGFBP-4. Endocrinology *128*, 2815-2824.

Chelius, D., Baldwin, M. A., Lu, X., and Spencer, E. M. (2001). Expression, purification and characterization of the structure and disulfide linkages of insulin-like growth factor binding protein-4. J Endocrinol *168*, 283-296.

Chernausek, S. D., Smith, C. E., Duffin, K. L., Busby, W. H., Wright, G., and Clemmons, D. R. (1995). Proteolytic cleavage of insulin-like growth factor binding protein 4 (IGFBP-4). Localization of cleavage site to non-homologous region of native IGFBP-4. J Biol Chem *270*, 11377-11382.

Chiva, C., Barthe, P., Codina, A., Gairi, M., Molina, F., Granier, C., Pugniere, M., Inui, T., Nishio, H., Nishiuchi, Y., *et al.* (2003). Synthesis and NMR structure of p41icf, a potent inhibitor of human cathepsin L. J Am Chem Soc *125*, 1508-1517.

Christofori, G., Naik, P., and Hanahan, D. (1995). Deregulation of both imprinted and expressed alleles of the insulin-like growth factor 2 gene during beta-cell tumorigenesis. Nat Genet *10*, 196-201.

Ciszak, E., Beals, J. M., Frank, B. H., Baker, J. C., Carter, N. D., and Smith, G. D. (1995). Role of C-terminal B-chain residues in insulin assembly: the structure of hexameric LysB28ProB29-human insulin. Structure *3*, 615-622.

Clemmons, D. R., Dehoff, M. L., Busby, W. H., Bayne, M. L., and Cascieri, M. A. (1992). Competition for binding to insulin-like growth factor (IGF) binding protein-2, 3, 4, and 5 by the IGFs and IGF analogs. Endocrinology *131*, 890-895.

Clemmons, D. R., and Maile, L. A. (2005). Interaction between insulin-like growth factor-I receptor and alphaVbeta3 integrin linked signaling pathways: cellular responses to changes in multiple signaling inputs. Mol Endocrinol *19*, 1-11.

Cohick, W. S., and Clemmons, D. R. (1993). The insulin-like growth factors. Annu Rev Physiol *55*, 131-153.

Collett-Solberg, P. F., and Cohen, P. (2000). Genetics, chemistry, and function of the IGF/IGFBP system. Endocrine *12*, 121-136.

Conover, C. A., Faessen, G. F., Ilg, K. E., Chandrasekher, Y. A., Christiansen, M., Overgaard, M. T., Oxvig, C., and Giudice, L. C. (2001). Pregnancy-associated plasma protein-a is the insulin-like growth factor binding protein-4 protease secreted by human ovarian granulosa cells and is a marker of dominant follicle selection and the corpus luteum. Endocrinology *142*, 2155.

Conover, C. A., and Khosla, S. (2003). Role of extracellular matrix in insulin-like growth factor (IGF) binding protein-2 regulation of IGF-II action in normal human osteoblasts. Growth Horm IGF Res *13*, 328-335.

Conover, C. A., and Kiefer, M. C. (1993). Regulation and biological effect of endogenous insulin-like growth factor binding protein-5 in human osteoblastic cells. J Clin Endocrinol Metab *76*, 1153-1159.

Cooke, R. M., Harvey, T. S., and Campbell, I. D. (1991). Solution structure of human insulin-like growth factor-I: a nuclear magnetic resonance and restrained molecular dynamics study. Biochemistry *30*, 5484-5491.

Cornell, W. D., Cieplak, P., Bayly, C. I., Gould, I. R., Merz, K. M., Ferguson, D. M., Spellmeyer, D. C., Fox, T., Caldwell, J. W., and Kollman, P. A. (1995). A second generation force field for the simulation of proteins, nucleic acids, and organic molecules. J Am Chem Soc *117*, 5179-5197.

Cornilescu, G., Delaglio, F., and Bax, A. (1999). Protein backbone angle restraints from searching a database for chemical shift and sequence homology. J Biomol NMR *13*, 289-302.

Coverley, J. A., and Baxter, R. C. (1997). Phosphorylation of insulin-like growth factor binding proteins. Mol Cell Endocrinol *128*, 1-5.

Cui, H., Cruz-Correa, M., Giardiello, F. M., Hutcheon, D. F., Kafonek, D. R., Brandenburg, S., Wu, Y., He, X., Powe, N. R., and Feinberg, A. P. (2003). Loss of IGF2 imprinting: a potential marker of colorectal cancer risk. Science *299*, 1753-1755.

de Pagter-Holthuizen, P., van Schaik, F. M., Verduijn, G. M., van Ommen, G. J., Bouma, B. N., Jansen, M., and Sussenbach, J. S. (1986). Organization of the human genes for insulin-like growth factors I and II. FEBS Lett *195*, 179-184.

De Souza, A. T., Hankins, G. R., Washington, M. K., Orton, T. C., and Jirtle, R. L. (1995). M6P/IGF2R gene is mutated in human hepatocellular carcinomas with loss of heterozygosity. Nat Genet *11*, 447-449.

De Vroede, M. A., Rechler, M. M., Nissley, S. P., Joshi, S., Burke, G. T., and Katsoyannis, P. G. (1985). Hybrid molecules containing the B-domain of insulin-like growth factor I are recognized by carrier proteins of the growth factor. Proc Natl Acad Sci U S A *82*, 3010-3014.

Denley, A., Bonython, E. R., Booker, G. W., Cosgrove, L. J., Forbes, B. E., Ward, C. W., and Wallace, J. C. (2004). Structural determinants for high-affinity binding of insulin-like growth factor II to insulin receptor (IR)-A, the exon 11 minus isoform of the IR. Mol Endocrinol *18*, 2502-2512.

Denley, A., Cosgrove, L. J., Booker, G. W., Wallace, J. C., and Forbes, B. E. (2005a). Molecular interactions of the IGF system. Cytokine Growth Factor Rev *16*, 421-439.

Denley, A., Wallace, J. C., Cosgrove, L. J., and Forbes, B. E. (2003). The insulin receptor isoform exon 11- (IR-A) in cancer and other diseases: a review. Horm Metab Res *35*, 778-785.

Denley, A., Wang, C. C., McNeil, K. A., Walenkamp, M. J., van Duyvenvoorde, H., Wit, J. M., Wallace, J. C., Norton, R. S., Karperien, M., and Forbes, B. E. (2005b). Structural and functional characteristics of the Val44Met insulin-like growth factor I missense mutation: correlation with effects on growth and development. Mol Endocrinol *19*, 711-721.

Devi, G. R., Yang, D. H., Rosenfeld, R. G., and Oh, Y. (2000). Differential effects of insulin-like growth factor (IGF)-binding protein-3 and its proteolytic fragments on ligand binding, cell surface association, and IGF-I receptor signaling. Endocrinology *141*, 4171-4179.

Dong, J., Wan, Z., Popov, M., Carey, P. R., and Weiss, M. A. (2003). Insulin assembly damps conformational fluctuations: Raman analysis of amide I linewidths in native states and fibrils. J Mol Biol *330*, 431-442.

Duan, C. (2002). Specifying the cellular responses to IGF signals: roles of IGF-binding proteins. J Endocrinol *175*, 41-54.

Dubaquie, Y., and Lowman, H. B. (1999). Total alanine-scanning mutagenesis of insulin-like growth factor I (IGF-I) identifies differential binding epitopes for IGFBP-1 and IGFBP-3. Biochemistry *38*, 6386-6396.

Ehring, H. (1999). Hydrogen exchange/electrospray ionization mass spectrometry studies of structural features of proteins and protein/protein interactions. Anal Biochem *267*, 252-259.

Farrow, N. A., Muhandiram, R., Singer, A. U., Pascal, S. M., Kay, C. M., Gish, G., Shoelson, S. E., Pawson, T., Forman-Kay, J. D., and Kay, L. E. (1994). Backbone dynamics of a free and phosphopeptide-complexed Src homology 2 domain studied by 15N NMR relaxation. Biochemistry *33*, 5984-6003.

Farrow, N. A., Zhang, O., Szabo, A., Torchia, D. A., and Kay, L. E. (1995). Spectral density function mapping using 15N relaxation data exclusively. J Biomol NMR *6*, 153-162.

Fernandez-Tornero, C., Lozano, R. M., Rivas, G., Jimenez, M. A., Standker, L., Diaz-Gonzalez, D., Forssmann, W. G., Cuevas, P., Romero, A., and Gimenez-Gallego, G. (2005). Synthesis of the blood circulating C-terminal fragment of insulin-like growth factor (IGF)-binding protein-4 in its native conformation. Crystallization, heparin and IGF binding, and osteogenic activity. J Biol Chem *280*, 18899-18907.

Firth, S. M., and Baxter, R. C. (1999). Characterisation of recombinant glycosylation variants of insulin-like growth factor binding protein-3. J Endocrinol *160*, 379-387.

Firth, S. M., and Baxter, R. C. (2002). Cellular actions of the insulin-like growth factor binding proteins. Endocr Rev *23*, 824-854.

Firth, S. M., Clemmons, D. R., and Baxter, R. C. (2001). Mutagenesis of basic amino acids in the carboxyl-terminal region of insulin-like growth factor binding protein-5 affects acid-labile subunit binding. Endocrinology *142*, 2147.

Firth, S. M., Ganeshprasad, U., and Baxter, R. C. (1998). Structural determinants of ligand and cell surface binding of insulin-like growth factor-binding protein-3. J Biol Chem *273*, 2631-2638.

Forbes, B. E., Hartfield, P. J., McNeil, K. A., Surinya, K. H., Milner, S. J., Cosgrove, L. J., and 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. Eur J Biochem *269*, 961-968.

Forbes, B. E., McNeil, K. A., Scott, C. D., Surinya, K. H., Cosgrove, L. J., and Wallace, J. C. (2001). Contribution of residues A54 and L55 of the human insulin-like growth

factor-II (IGF-II) A domain to Type 2 IGF receptor binding specificity. Growth Factors *19*, 163-173.

Forbes, B. E., Turner, D., Hodge, S. J., McNeil, K. A., Forsberg, G., and Wallace, J. C. (1998). Localization of an insulin-like growth factor (IGF) binding site of bovine IGF binding protein-2 using disulfide mapping and deletion mutation analysis of the C-terminal domain. J Biol Chem *273*, 4647-4652.

Foulstone, E., Prince, S., Zaccheo, O., Burns, J. L., Harper, J., Jacobs, C., Church, D., and Hassan, A. B. (2005). Insulin-like growth factor ligands, receptors, and binding proteins in cancer. J Pathol *205*, 145-153.

Francis, G. L., Aplin, S. E., Milner, S. J., McNeil, K. A., Ballard, F. J., and Wallace, J. C. (1993). Insulin-like growth factor (IGF)-II binding to IGF-binding proteins and IGF receptors is modified by deletion of the N-terminal hexapeptide or substitution of arginine for glutamate-6 in IGF-II. Biochem J *293 (Pt 3)*, 713-719.

Francis, G. L., Upton, F. M., Ballard, F. J., McNeil, K. A., and Wallace, J. C. (1988). Insulin-like growth factors 1 and 2 in bovine colostrum. Sequences and biological activities compared with those of a potent truncated form. Biochem J *251*, 95-103.

Furstenberger, G., and Senn, H. J. (2002). Insulin-like growth factors and cancer. Lancet Oncol *3*, 298-302.

Galanis, M., Firth, S. M., Bond, J., Nathanielsz, A., Kortt, A. A., Hudson, P. J., and Baxter, R. C. (2001). Ligand-binding characteristics of recombinant amino- and carboxyl-terminal fragments of human insulin-like growth factor-binding protein-3. J Endocrinol *169*, 123-133.

Gatenby, R. A., and Gillies, R. J. (2004). Why do cancers have high aerobic glycolysis? Nat Rev Cancer *4*, 891-899.

Gleeson, H. K., Lissett, C. A., and Shalet, S. M. (2005). Insulin-like growth factor-I response to a single bolus of growth hormone is increased in obesity. J Clin Endocrinol Metab *90*, 1061-1067.

Gockerman, A., and Clemmons, D. R. (1995). Porcine aortic smooth muscle cells secrete a serine protease for insulin-like growth factor binding protein-2. Circ Res *76*, 514-521.

Golovanov, A. P., Hautbergue, G. M., Wilson, S. A., and Lian, L. Y. (2004). A simple method for improving protein solubility and long-term stability. J Am Chem Soc *126*, 8933-8939.

Gulve, E. A., and Dice, J. F. (1989). Regulation of protein synthesis and degradation in L8 myotubes. Effects of serum, insulin and insulin-like growth factors. Biochem J *260*, 377-387.

Guncar, G., Pungercic, G., Klemencic, I., Turk, V., and Turk, D. (1999). Crystal structure of MHC class II-associated p41 Ii fragment bound to cathepsin L reveals the structural basis for differentiation between cathepsins L and S. Embo J *18*, 793-803.

Guntert, P., Mumenthaler, C., and Wuthrich, K. (1997). Torsion angle dynamics for NMR structure calculation with the new program DYANA. J Mol Biol *273*, 283-298.

Haigh, C. W., and Mallion, R. B. (1980). Ring current theories in nuclear magnetic resonance. Progress in Nuclear Magnetic Resonance Spectroscopy *13*, 303-314.

Hallgren, J., Backstrom, S., Estrada, S., Thuveson, M., and Pejler, G. (2004). Histidines are critical for heparin-dependent activation of mast cell tryptase. J Immunol *173*, 1868-1875.

Hashimoto, R., Fujiwara, H., Higashihashi, N., Enjoh-Kimura, T., Terasawa, H., Fujita-Yamaguchi, Y., Inagaki, F., Perdue, J. F., and Sakano, K. (1995). N-terminal deletion mutants of insulin-like growth factor-II (IGF-II) show Thr7 and Leu8 important for binding to insulin and IGF-I receptors and Leu8 critical for all IGF-II functions. J Biol Chem *270*, 18013-18018.

Hashimoto, R., Ono, M., Fujiwara, H., Higashihashi, N., Yoshida, M., Enjoh-Kimura, T., and Sakano, K. (1997). Binding sites and binding properties of binary and ternary complexes of insulin-like growth factor-II (IGF-II), IGF-binding protein-3, and acid-labile subunit. J Biol Chem *272*, 27936-27942.

Headey, S. J., Keizer, D. W., Yao, S., Brasier, G., Kantharidis, P., Bach, L. A., and Norton, R. S. (2004a). C-terminal domain of insulin-like growth factor (IGF) binding protein-6: structure and interaction with IGF-II. Mol Endocrinol *18*, 2740-2750.

Headey, S. J., Keizer, D. W., Yao, S., Wallace, J. C., Bach, L. A., and Norton, R. S. (2004b). Binding site for the C-domain of insulin-like growth factor (IGF) binding protein-6 on IGF-II; implications for inhibition of IGF actions. FEBS Lett *568*, 19-22.

Headey, S. J., Leeding, K. S., Norton, R. S., and Bach, L. A. (2004c). Contributions of the N- and C-terminal domains of insulin-like growth factor (IGF) binding protein-6 to IGF binding. J Mol Endocrinol *33*, 377-386.

Hinds, M. G., and Norton, R. S. (1997). NMR spectroscopy of peptides and proteins. Practical considerations. Mol Biotechnol *7*, 315-331.

Ho, P. J., and Baxter, R. C. (1997). Characterization of truncated insulin-like growth factor-binding protein-2 in human milk. Endocrinology *138*, 3811-3818.

Hobba, G. D., Forbes, B. E., Parkinson, E. J., Francis, G. L., and Wallace, J. C. (1996). The insulin-like growth factor (IGF) binding site of bovine insulin-like growth factor binding protein-2 (bIGFBP-2) probed by iodination. J Biol Chem *271*, 30529-30536.

Hobba, G. D., Lothgren, A., Holmberg, E., Forbes, B. E., Francis, G. L., and Wallace, J. C. (1998). Alanine screening mutagenesis establishes tyrosine 60 of bovine insulin-like growth factor binding protein-2 as a determinant of insulin-like growth factor binding. J Biol Chem *273*, 19691-19698.

Hodgson, D. R., May, F. E., and Westley, B. R. (1995). Mutations at positions 11 and 60 of insulin-like growth factor 1 reveal differences between its interactions with the type I insulin-like-growth-factor receptor and the insulin receptor. Eur J Biochem *233*, 299-309.

Hoeck, W. G., and Mukku, V. R. (1994). Identification of the major sites of phosphorylation in IGF binding protein-3. J Cell Biochem *56*, 262-273.

Hoeflich, A., Fettscher, O., Lahm, H., Blum, W. F., Kolb, H. J., Engelhardt, D., Wolf, E., and Weber, M. M. (2000). Overexpression of insulin-like growth factor-binding protein-2 results in increased tumorigenic potential in Y-1 adrenocortical tumor cells. Cancer Res *60*, 834-838.

Hoeflich, A., Nedbal, S., Blum, W. F., Erhard, M., Lahm, H., Brem, G., Kolb, H. J., Wanke, R., and Wolf, E. (2001a). Growth inhibition in giant growth hormone transgenic mice by overexpression of insulin-like growth factor-binding protein-2. Endocrinology *142*, 1889-1898.

Hoeflich, A., Reisinger, R., Lahm, H., Kiess, W., Blum, W. F., Kolb, H. J., Weber, M. M., and Wolf, E. (2001b). Insulin-like growth factor-binding protein 2 in tumorigenesis: protector or promoter? Cancer Res *61*, 8601-8610.

Hoeflich, A., Reisinger, R., Schuett, B. S., Elmlinger, M. W., Russo, V. C., Vargas, G. A., Jehle, P. M., Lahm, H., Renner-Muller, I., and Wolf, E. (2004). Peri/nuclear localization of intact insulin-like growth factor binding protein-2 and a distinct carboxyl-terminal IGFBP-2 fragment in vivo. Biochem Biophys Res Commun *324*, 705-710.

Hoeflich, A., Reisinger, R., Vargas, G. A., Elmlinger, M. W., Schuett, B., Jehle, P. M., Renner-Muller, I., Lahm, H., Russo, V. C., and Wolf, E. (2002). Mutation of the RGD sequence does not affect plasma membrane association and growth inhibitory effects of elevated IGFBP-2 in vivo. FEBS Lett *523*, 63-67.

Hoflich, A., Lahm, H., Blum, W., Kolb, H., and Wolf, E. (1998). Insulin-like growth factor-binding protein-2 inhibits proliferation of human embryonic kidney fibroblasts and of IGF-responsive colon carcinoma cell lines. FEBS Lett *434*, 329-334.

Hong, J., Zhang, G., Dong, F., and Rechler, M. M. (2002). Insulin-like growth factor (IGF)-binding protein-3 mutants that do not bind IGF-I or IGF-II stimulate apoptosis in human prostate cancer cells. J Biol Chem *277*, 10489-10497.

Horney, M. J., Evangelista, C. A., and Rosenzweig, S. A. (2001). Synthesis and characterization of insulin-like growth factor (IGF)-1 photoprobes selective for the IGF-binding proteins (IGFBPS). photoaffinity labeling of the IGF-binding domain on IGFBP-2. J Biol Chem *276*, 2880-2889.

Hua, Q. X., Nakagawa, S., Hu, S. Q., Jia, W., Wang, S., and Weiss, M. A. (2006). Toward the active conformation of insulin: stereospecific modulation of a structural switch in the B chain. J Biol Chem *281*, 24900-24909.

Hua, Q. X., and Weiss, M. A. (1991). Comparative 2D NMR studies of human insulin and des-pentapeptide insulin: sequential resonance assignment and implications for protein dynamics and receptor recognition. Biochemistry *30*, 5505-5515.

Huang, S. S., Tang, F. M., Huang, Y. H., Liu, I. H., Hsu, S. C., Chen, S. T., and Huang, J. S. (2003). Cloning, expression, characterization, and role in autocrine cell growth of cell surface retention sequence binding protein-1. J Biol Chem *278*, 43855-43869.

Huhtala, M. L., Koistinen, R., Palomaki, P., Partanen, P., Bohn, H., and Seppala, M. (1986). Biologically active domain in somatomedin-binding protein. Biochem Biophys Res Commun *141*, 263-270.

Hung, K. W., Kumar, T. K., Kathir, K. M., Xu, P., Ni, F., Ji, H. H., Chen, M. C., Yang, C. C., Lin, F. P., Chiu, I. M., and Yu, C. (2005). Solution structure of the ligand binding domain of the fibroblast growth factor receptor: role of heparin in the activation of the receptor. Biochemistry *44*, 15787-15798.

Hwa, V., Oh, Y., and Rosenfeld, R. G. (1999). The insulin-like growth factor-binding protein (IGFBP) superfamily. Endocr Rev *20*, 761-787.

Hwang, T. L., van Zijl, P. C., and Mori, S. (1998). Accurate quantitation of water-amide proton exchange rates using the phase-modulated CLEAN chemical EXchange (CLEANEX-PM) approach with a Fast-HSQC (FHSQC) detection scheme. J Biomol NMR *11*, 221-226.

Ikonen, M., Liu, B., Hashimoto, Y., Ma, L., Lee, K. W., Niikura, T., Nishimoto, I., and Cohen, P. (2003). Interaction between the Alzheimer's survival peptide humanin and insulin-like growth factor-binding protein 3 regulates cell survival and apoptosis. Proc Natl Acad Sci U S A *100*, 13042-13047.

Imai, Y., Busby, W. H., Jr., Smith, C. E., Clarke, J. B., Garmong, A. J., Horwitz, G. D., Rees, C., and Clemmons, D. R. (1997). Protease-resistant form of insulin-like growth factor-binding protein 5 is an inhibitor of insulin-like growth factor-I actions on porcine smooth muscle cells in culture. J Clin Invest *100*, 2596-2605.

Imai, Y., Moralez, A., Andag, U., Clarke, J. B., Busby, W. H., Jr., and Clemmons, D. R. (2000). Substitutions for hydrophobic amino acids in the N-terminal domains of IGFBP-3 and -5 markedly reduce IGF-I binding and alter their biologic actions. J Biol Chem *275*, 18188-18194.

Janssen, J. A., and Lamberts, S. W. (2004). Igf-I and longevity. Horm Res *62 Suppl 3*, 104-109.

Jansson, M., Andersson, G., Uhlen, M., Nilsson, B., and Kordel, J. (1998). The insulin-like growth factor (IGF)binding protein 1 binding epitope on IGF-I probed by heteronuclear NMR spectroscopy and mutational analysis. J Biol Chem *273*, 24701-24707.

Jansson, M., Uhlen, M., and 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*, 4108-4117.

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

Jones, J. I., D'Ercole, A. J., Camacho-Hubner, C., and Clemmons, D. R. (1991). Phosphorylation of insulin-like growth factor (IGF)-binding protein 1 in cell culture and in vivo: effects on affinity for IGF-I. Proc Natl Acad Sci U S A *88*, 7481-7485.

Jones, J. I., Gockerman, A., Busby, W. H., Jr., Camacho-Hubner, C., and Clemmons, D. R. (1993a). Extracellular matrix contains insulin-like growth factor binding protein-5: potentiation of the effects of IGF-I. J Cell Biol *121*, 679-687.

Jones, J. I., Gockerman, A., Busby, W. H., Jr., Wright, G., and Clemmons, D. R. (1993b). Insulin-like growth factor binding protein 1 stimulates cell migration and binds to the alpha 5 beta 1 integrin by means of its Arg-Gly-Asp sequence. Proc Natl Acad Sci U S A *90*, 10553-10557.

Kalus, W., Zweckstetter, M., Renner, C., Sanchez, Y., Georgescu, J., Grol, M., Demuth, D., Schumacher, R., Dony, C., Lang, K., and Holak, T. A. (1998). Structure of the IGF-binding domain of the insulin-like growth factor-binding protein-5 (IGFBP-5): implications for IGF and IGF-I receptor interactions. EMBO J *17*, 6558-6572.

Kanelis, V., Forman-Kay, J. D., and Kay, L. E. (2001). Multidimensional NMR methods for protein structure determination. IUBMB Life *52*, 291-302.

Katic, M., and Kahn, C. R. (2005). The role of insulin and IGF-1 signaling in longevity. Cell Mol Life Sci *62*, 320-343.

Kay, L. E., Torchia, D. A., and Bax, A. (1989). Backbone dynamics of proteins as studied by 15N inverse detected heteronuclear NMR spectroscopy: application to staphylococcal nuclease. Biochemistry *28*, 8972-8979.

Kibbey, M. M., Jameson, M. J., Eaton, E. M., and Rosenzweig, S. A. (2006). Insulin-like growth factor binding protein-2: contributions of the C-terminal domain to IGF-1 binding. Mol Pharmacol *69*, 833-845.

King, R., Wells, J. R., Krieg, P., Snoswell, M., Brazier, J., Bagley, C. J., Wallace, J. C., Ballard, F. J., Ross, M., and Francis, G. L. (1992). Production and characterization of recombinant insulin-like growth factor-I (IGF-I) and potent analogues of IGF-I, with Gly or Arg substituted for Glu3, following their expression in Escherichia coli as fusion proteins. J Mol Endocrinol *8*, 29-41.

Koradi, R., Billeter, M., and Guntert, P. (2000). Point-centered domain decomposition for parallel molecular dynamics simulation. Comput Phys Commun *124*, 139-147.

Koradi, R., Billeter, M., and Wuthrich, K. (1996). MOLMOL: a program for display and analysis of macromolecular structures. J Mol Graph *14*, 51-55.

Kornfeld, S. (1992). Structure and function of the mannose 6-phosphate/insulinlike growth factor II receptors. Annu Rev Biochem *61*, 307-330.

Kricker, J. A., Towne, C. L., Firth, S. M., Herington, A. C., and Upton, Z. (2003). Structural and functional evidence for the interaction of insulin-like growth factors (IGFs) and IGF binding proteins with vitronectin. Endocrinology *144*, 2807-2815.

Kubler, B., Cowell, S., Zapf, J., and Braulke, T. (1998). Proteolysis of insulin-like growth factor binding proteins by a novel 50-kilodalton metalloproteinase in human pregnancy serum. Endocrinology *139*, 1556-1563.

Kuboniwa, H., Grzesiek, S., Delaglio, F., and Bax, A. (1994). Measurement of HN-H alpha J couplings in calcium-free calmodulin using new 2D and 3D water-flip-back methods. J Biomol NMR *4*, 871-878.

Laajoki, L. G., Francis, G. L., Wallace, J. C., Carver, J. A., and Keniry, M. A. (2000). Solution structure and backbone dynamics of long-[Arg(3)]insulin-like growth factor-I. J Biol Chem *275*, 10009-10015.

Lalou, C., Sawamura, S., Segovia, B., Ogawa, Y., and Binoux, M. (1997). Proteolytic fragments of insulin-like growth factor binding protein-3: N-terminal sequences and relationships between structure and biological activity. C R Acad Sci III *320*, 621-628.

Laskowski, R. A., Rullmannn, J. A., MacArthur, M. W., Kaptein, R., and Thornton, J. M. (1996). AQUA and PROCHECK-NMR: programs for checking the quality of protein structures solved by NMR. J Biomol NMR *8*, 477-486.

Laursen, L. S., Overgaard, M. T., Soe, R., Boldt, H. B., Sottrup-Jensen, L., Giudice, L. C., Conover, C. A., and Oxvig, C. (2001). Pregnancy-associated plasma protein-A (PAPP-A) cleaves insulin-like growth factor binding protein (IGFBP)-5 independent of IGF: implications for the mechanism of IGFBP-4 proteolysis by PAPP-A. FEBS Lett *504*, 36-40.

Leal, S. M., Huang, S. S., and Huang, J. S. (1999). Interactions of high affinity insulin-like growth factor-binding proteins with the type V transforming growth factor-beta receptor in mink lung epithelial cells. J Biol Chem *274*, 6711-6717.

Leal, S. M., Liu, Q., Huang, S. S., and Huang, J. S. (1997). The type V transforming growth factor beta receptor is the putative insulin-like growth factor-binding protein 3 receptor. J Biol Chem *272*, 20572-20576.

Lee, K. W., and Cohen, P. (2002). Nuclear effects: unexpected intracellular actions of insulin-like growth factor binding protein-3. J Endocrinol *175*, 33-40.

Lee, K. W., Liu, B., Ma, L., Li, H., Bang, P., Koeffler, H. P., and Cohen, P. (2004). Cellular internalization of insulin-like growth factor binding protein-3: distinct endocytic pathways facilitate re-uptake and nuclear localization. J Biol Chem *279*, 469-476.

LeRoith, D., and Roberts, C. T., Jr. (2003). The insulin-like growth factor system and cancer. Cancer Lett *195*, 127-137.

Liu, B., Lee, H. Y., Weinzimer, S. A., Powell, D. R., Clifford, J. L., Kurie, J. M., and Cohen, P. (2000). Direct functional interactions between insulin-like growth factor-binding protein-3 and retinoid X receptor-alpha regulate transcriptional signaling and apoptosis. J Biol Chem *275*, 33607-33613.

Liu, B., Weinzimer, S. A., Gibson, T. B., Mascarenhas, D., and Cohen, P. (2003). Type Ialpha collagen is an IGFBP-3 binding protein. Growth Horm IGF Res *13*, 89-97.

Liu, J. P., Baker, J., Perkins, A. S., Robertson, E. J., and Efstratiadis, A. (1993). Mice carrying null mutations of the genes encoding insulin-like growth factor-I (*Igf-1*) and type 1 IGF receptor (*Igf1r*). Cell *75*, 59-72.

Loddick, S. A., Liu, X. J., Lu, Z. X., Liu, C., Behan, D. P., Chalmers, D. C., Foster, A. C., Vale, W. W., Ling, N., and De Souza, E. B. (1998). Displacement of insulin-like growth factors from their binding proteins as a potential treatment for stroke. Proc Natl Acad Sci U S A *95*, 1894-1898.

Lucic, M. R., Forbes, B. E., Grosvenor, S. E., Carr, J. M., Wallace, J. C., and Forsberg, G. (1998). Secretion in Escherichia coli and phage-display of recombinant insulin-like growth factor binding protein-2. J Biotechnol *61*, 95-108.

Ludvigsen, S., and Poulsen, F. M. (1992). Positive theta-angles in proteins by nuclear magnetic resonance spectroscopy. J Biomol NMR *2*, 227-233.

Maciejewski, M. W., Liu, D., Prasad, R., Wilson, S. H., and Mullen, G. P. (2000). Backbone dynamics and refined solution structure of the N-terminal domain of DNA polymerase beta. Correlation with DNA binding and dRP lyase activity. J Mol Biol *296*, 229-253.

Magee, B. A., Shooter, G. K., Wallace, J. C., and Francis, G. L. (1999). Insulin-like growth factor-I and its binding proteins: a study of the binding interface using B-domain mutants. Biochemistry *38*, 15863-15870.

Marinaro, J. A., Neumann, G. M., Russo, V. C., Leeding, K. S., and Bach, L. A. (2000). O-glycosylation of insulin-like growth factor (IGF) binding protein-6 maintains high IGF-II binding affinity by decreasing binding to glycosaminoglycans and susceptibility to proteolysis. Eur J Biochem *267*, 5378-5386.

Mark, S., Kubler, B., Honing, S., Oesterreicher, S., John, H., Braulke, T., Forssmann, W. G., and Standker, L. (2005). Diversity of human insulin-like growth factor (IGF) binding protein-2 fragments in plasma: primary structure, IGF-binding properties, and disulfide bonding pattern. Biochemistry *44*, 3644-3652.

Mathews, L. S., Hammer, R. E., Behringer, R. R., D'Ercole, A. J., Bell, G. I., Brinster, R. L., and Palmiter, R. D. (1988). Growth enhancement of transgenic mice expressing human insulin-like growth factor-I. Endocrinology *123*, 2827-2833.

Miyakoshi, N., Richman, C., Kasukawa, Y., Linkhart, T. A., Baylink, D. J., and Mohan, S. (2001). Evidence that IGF-binding protein-5 functions as a growth factor. J Clin Invest *107*, 73-81.

Mohan, S., and Baylink, D. J. (2002). IGF-binding proteins are multifunctional and act via IGF-dependent and -independent mechanisms. J Endocrinol *175*, 19-31.

Monget, P., Mazerbourg, S., Delpuech, T., Maurel, M. C., Maniere, S., Zapf, J., Lalmanach, G., Oxvig, C., and Overgaard, M. T. (2003). Pregnancy-associated plasma protein-A is involved in insulin-like growth factor binding protein-2 (IGFBP-2) proteolytic degradation in bovine and porcine preovulatory follicles: identification of cleavage site and characterization of IGFBP-2 degradation. Biol Reprod *68*, 77-86.

Monzavi, R., and Cohen, P. (2002). IGFs and IGFBPs: role in health and disease. Best Pract Res Clin Endocrinol Metab *16*, 433-447.

Moss, J. A., Francis, G. L., Ross, M., Wallace, J. C., and Ballard, F. J. (1991). Insulin-like growth factor (IGF)-I and IGF-II binding to an IGF binding protein. An investigation using chemical modification of tyrosine residues as a structural probe for the sites of interaction. J Biol Chem *266*, 909-914.

Mulloy, B., Forster, M. J., Jones, C., and Davies, D. B. (1993). N.m.r. and molecular-modelling studies of the solution conformation of heparin. Biochem J *293*, 849-858.

Nam, T., Moralez, A., and Clemmons, D. (2002). Vitronectin binding to IGF binding protein-5 (IGFBP-5) alters IGFBP-5 modulation of IGF-I actions. Endocrinology *143*, 30-36.

Nam, T. J., Busby, W., Jr., and Clemmons, D. R. (1997). Insulin-like growth factor binding protein-5 binds to plasminogen activator inhibitor-I. Endocrinology *138*, 2972-2978.

Nam, T. J., Busby, W. H., Jr., Rees, C., and Clemmons, D. R. (2000). Thrombospondin and osteopontin bind to insulin-like growth factor (IGF)-binding protein-5 leading to an alteration in IGF-I-stimulated cell growth. Endocrinology *141*, 1100-1106.

Neidhardt, F. C., Bloch, P. L., and Smith, D. F. (1974). Culture medium for enterobacteria. J Bacteriol *119*, 736-747.

Neumann, G. M., and Bach, L. A. (1999). The N-terminal disulfide linkages of human insulin-like growth factor-binding protein-6 (hIGFBP-6) and hIGFBP-1 are different as determined by mass spectrometry. J Biol Chem *274*, 14587-14594.

Neumann, G. M., Marinaro, J. A., and Bach, L. A. (1998). Identification of O-glycosylation sites and partial characterization of carbohydrate structure and disulfide linkages of human insulin-like growth factor binding protein-6. Biochemistry *37*, 6572-6585.

Nicholls, A., Sharp, K. A., and Honig, B. (1991). Protein folding and association: insights from the interfacial and thermodynamic properties of hydrocarbons. Proteins *11*, 281-296.

Novinec, M., Kordis, D., Turk, V., and Lenarcic, B. (2006). Diversity and Evolution of the Thyroglobulin Type-1 Domain Superfamily. Mol Biol Evol *23*, 744-755.

Oesterreicher, S., Blum, W. F., Schmidt, B., Braulke, T., and Kubler, B. (2005). Interaction of insulin-like growth factor II (IGF-II) with multiple plasma proteins: high affinity binding of plasminogen to IGF-II and IGF-binding protein-3. J Biol Chem *280*, 9994-10000.

Oh, Y., Mueller, H. L., Lee, D. Y., Fielder, P. J., and Rosenfeld, R. G. (1993). Characterization of the affinities of insulin-like growth factor (IGF)-binding protein -1 to -4 for IGF-I, IGF-II, IGF-I/insulin hybrid, and IGF-I analogs. Endocrinology *132*, 1337-1344.

Okamoto, K., Morison, I. M., Taniguchi, T., and Reeve, A. E. (1997). Epigenetic changes at the insulin-like growth factor II/H19 locus in developing kidney is an early event in Wilms tumorigenesis. Proc Natl Acad Sci U S A *94*, 5367-5371.

Overgaard, M. T., Boldt, H. B., Laursen, L. S., Sottrup-Jensen, L., Conover, C. A., and Oxvig, C. (2001). Pregnancy-associated plasma protein-A2 (PAPP-A2), a novel insulin-like growth factor-binding protein-5 proteinase. J Biol Chem *276*, 21849-21853.

Palmer, A. G., 3rd (2001). Nmr probes of molecular dynamics: overview and comparison with other techniques. Annu Rev Biophys Biomol Struct *30*, 129-155.

Parker, A., Clarke, J. B., Busby, W. H., Jr., and Clemmons, D. R. (1996). Identification of the extracellular matrix binding sites for insulin-like growth factor-binding protein 5. J Biol Chem *271*, 13523-13529.

Parker, A., Rees, C., Clarke, J., Busby, W. H., Jr., and Clemmons, D. R. (1998). Binding of insulin-like growth factor (IGF)-binding protein-5 to smooth-muscle cell extracellular matrix is a major determinant of the cellular response to IGF-I. Mol Biol Cell *9*, 2383-2392.

Payet, L. D., Wang, X. H., Baxter, R. C., and Firth, S. M. (2003). Amino- and carboxyl-terminal fragments of insulin-like growth factor (IGF) binding protein-3 cooperate to bind IGFs with high affinity and inhibit IGF receptor interactions. Endocrinology *144*, 2797-2806.

Pereira, J. J., Meyer, T., Docherty, S. E., Reid, H. H., Marshall, J., Thompson, E. W., Rossjohn, J., and Price, J. T. (2004). Bimolecular interaction of insulin-like growth factor (IGF) binding protein-2 with alphavbeta3 negatively modulates IGF-I-mediated migration and tumor growth. Cancer Res *64*, 977-984.

Perks, C., and Holly, J. (2003). Actions of IGFBP on epithelial cancer cells: potential for new therapeutic targets. Horm Metab Res *35*, 828-835.

Pervushin, K., Riek, R., Wider, G., and Wuthrich, K. (1997). Attenuated T2 relaxation by mutual cancellation of dipole-dipole coupling and chemical shift anisotropy indicates an avenue to NMR structures of very large biological macromolecules in solution. Proc Natl Acad Sci U S A *94*, 12366-12371.

Peterkofsky, B., Gosiewska, A., Wilson, S., and Kim, Y. R. (1998). Phosphorylation of rat insulin-like growth factor binding protein-1 does not affect its biological properties. Arch Biochem Biophys *357*, 101-110.

Pintar, J. E., Cerro, J. A., and Wood, T. L. (1996). Genetic approaches to the function of insulin-like growth factor-binding proteins during rodent development. Horm Res *45*, 172-177.

Pintar, J. E., Schuller, A., and Bradshaw, S. (1999). Combinatorial KOs of IGFBPs. Growth Horm IGF Res *9*, 308.

Pollak, M. N., Schernhammer, E. S., and Hankinson, S. E. (2004). Insulin-like growth factors and neoplasia. Nat Rev Cancer *4*, 505-518.

Qin, X., Strong, D. D., Baylink, D. J., and Mohan, S. (1998). Structure-function analysis of the human insulin-like growth factor binding protein-4. J Biol Chem *273*, 23509-23516.

Rajaram, S., Baylink, D. J., and Mohan, S. (1997). Insulin-like growth factor-binding proteins in serum and other biological fluids: regulation and functions. Endocr Rev *18*, 801-831.

Raman, R., Sasisekharan, V., and Sasisekharan, R. (2005). Structural insights into biological roles of protein-glycosaminoglycan interactions. Chem Biol *12*, 267-277.

Rees, C., Clemmons, D. R., Horvitz, G. D., Clarke, J. B., and Busby, W. H. (1998). A protease-resistant form of insulin-like growth factor (IGF) binding protein-4 inhibits IGF-1 actions. Endocrinology *139*, 4182-4188.

Reeve, J. G., Morgan, J., Schwander, J., and Bleehen, N. M. (1993). Role for membrane and secreted insulin-like growth factor-binding protein-2 in the regulation of insulin-like growth factor action in lung tumors. Cancer Res *53*, 4680-4685.

Renner, C., and Holak, T. (2001). NMR 15N relaxation of the insulin-like growth factor (IGF)-binding domain of IGF binding protein-5 (IGFBP-5) determined free in solution and in complex with IGF-II. Eur J Biochem *268*, 1058-1065.

Ricort, J. M. (2004). Insulin-like growth factor binding protein (IGFBP) signalling. Growth Horm IGF Res *14*, 277-286.

Robinson, S. A., and Rosenzweig, S. A. (2004). Synthesis and characterization of biotinylated forms of insulin-like growth factor-1: topographical evaluation of the IGF-1/IGFBP-2 AND IGFBP-3 interface. Biochemistry *43*, 11533-11545.

Rosenfeld, R. G. (2005). The IGF system: new developments relevant to pediatric practice. Endocr Dev *9*, 1-10.

Rosenzweig, S. A. (2004). What's new in the IGF-binding proteins? Growth Horm IGF Res *14*, 329-336.

Roth, B. V., Burgisser, D. M., Luthi, C., and Humbel, R. E. (1991). Mutants of human insulin-like growth factor II: expression and characterization of analogs with a substitution of TYR27 and/or a deletion of residues 62-67. Biochem Biophys Res Commun *181*, 907-914.

Rotwein, P. (1991). Structure, evolution, expression and regulation of insulin-like growth factors I and II. Growth Factors *5*, 3-18.

Russo, V. C., Bach, L. A., Fosang, A. J., Baker, N. L., and Werther, G. A. (1997). Insulin-like growth factor binding protein-2 binds to cell surface proteoglycans in the rat brain olfactory bulb. Endocrinology *138*, 4858-4867.

Russo, V. C., Rekaris, G., Baker, N. L., Bach, L. A., and Werther, G. A. (1999). Basic fibroblast growth factor induces proteolysis of secreted and cell membrane-associated insulin-like growth factor binding protein-2 in human neuroblastoma cells. Endocrinology *140*, 3082-3090.

Russo, V. C., Schutt, B. S., Andaloro, E., Ymer, S. I., Hoeflich, A., Ranke, M. B., Bach, L. A., and Werther, G. A. (2005). Igfbp-2 binding to extracellular matrix plays a critical

role in neuroblastoma cell proliferation, migration and invasion. Endocrinology *146*, 4445-4455.

Sakano, K., Enjoh, T., Numata, F., Fujiwara, H., Marumoto, Y., Higashihashi, N., Sato, Y., Perdue, J. F., and Fujita-Yamaguchi, Y. (1991). The design, expression, and characterization of human insulin-like growth factor II (IGF-II) mutants specific for either the IGF-II/cation-independent mannose 6-phosphate receptor or IGF-I receptor. J Biol Chem *266*, 20626-20635.

Sala, A., Capaldi, S., Campagnoli, M., Faggion, B., Labo, S., Perduca, M., Romano, A., Carrizo, M. E., Valli, M., Visai, L., *et al.* (2005). Structure and Properties of the C-terminal Domain of Insulin-like Growth Factor-binding Protein-1 Isolated from Human Amniotic Fluid. J Biol Chem *280*, 29812-29819.

Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989). Molecular cloning: a laboratory manual, 2nd edn (New York, Cold Spring Harbor Press).

Santer, F. R., Bacher, N., Moser, B., Morandell, D., Ressler, S., Firth, S. M., Spoden, G. A., Sergi, C., Baxter, R. C., Jansen-Durr, P., and Zwerschke, W. (2006). Nuclear insulin-like growth factor binding protein-3 induces apoptosis and is targeted to ubiquitin/proteasome-dependent proteolysis. Cancer Res *66*, 3024-3033.

Sato, A., Nishimura, S., Ohkubo, T., Kyogoku, Y., Koyama, S., Kobayashi, M., Yasuda, T., and Kobayashi, Y. (1993). Three-dimensional structure of human insulin-like growth factor-I (IGF-I) determined by 1H-NMR and distance geometry. Int J Pept Protein Res *41*, 433-440.

Savage, M. O., Camacho-Hubner, C., and Dunger, D. B. (2004). Therapeutic applications of the insulin-like growth factors. Growth Horm IGF Res *14*, 301-308.

Schaffer, M. L., Deshayes, K., Nakamura, G., Sidhu, S., and Skelton, N. J. (2003). Complex with a Phage Display-Derived Peptide Provides Insight into the Function of Insulin-like Growth Factor I. Biochemistry *42*, 9324-9334.

Schedlich, L. J., Le Page, S. L., Firth, S. M., Briggs, L. J., Jans, D. A., and Baxter, R. C. (2000). Nuclear import of insulin-like growth factor-binding protein-3 and -5 is mediated by the importin beta subunit. J Biol Chem *275*, 23462-23470.

Schedlich, L. J., Nilsen, T., John, A. P., Jans, D. A., and Baxter, R. C. (2003). Phosphorylation of insulin-like growth factor binding protein-3 by deoxyribonucleic acid-dependent protein kinase reduces ligand binding and enhances nuclear accumulation. Endocrinology *144*, 1984-1993.

Schedlich, L. J., Young, T. F., Firth, S. M., and Baxter, R. C. (1998). Insulin-like growth factor-binding protein (IGFBP)-3 and IGFBP-5 share a common nuclear transport pathway in T47D human breast carcinoma cells. J Biol Chem *273*, 18347-18352.

Schneider, M. R., Lahm, H., Wu, M., Hoeflich, A., and Wolf, E. (2000). Transgenic mouse models for studying the functions of insulin-like growth factor-binding proteins. FASEB J *14*, 629-640.

Schuller, A. G., Lindenbergh-Kortleve, D. J., de Boer, W. I., Zwarthoff, E. C., and Drop, S. L. (1993). Localization of the epitope of a monoclonal antibody against human insulin-like growth factor binding protein-1, functionally interfering with insulin-like growth factor binding. Growth Regul *3*, 32-34.

Schutt, B. S., Langkamp, M., Rauschnabel, U., Ranke, M. B., and Elmlinger, M. W. (2004). Integrin-mediated action of insulin-like growth factor binding protein-2 in tumor cells. J Mol Endocrinol *32*, 859-868.

Schwieters, C. D., Kuszewski, J. J., Tjandra, N., and Clore, G. M. (2003). The Xplor-NIH NMR molecular structure determination package. J Magn Reson *160*, 65-73.

Seavey, B. R., Farr, E. A., Westler, W. M., and Markley, J. L. (1991). A relational database for sequence-specific protein NMR data. J Biomol NMR *1*, 217-236.

Sebollela, A., Cagliari, T. C., Limaverde, G. S., Chapeaurouge, A., Sorgine, M. H., Coelho-Sampaio, T., Ramos, C. H., and Ferreira, S. T. (2005). Heparin-binding sites in granulocyte-macrophage colony-stimulating factor. Localization and regulation by histidine ionization. J Biol Chem *280*, 31949-31956.

Shand, J. H., Beattie, J., Song, H., Phillips, K., Kelly, S. M., Flint, D. J., and Allan, G. J. (2003). Specific amino acid substitutions determine the differential contribution of the N- and C-terminal domains of insulin-like growth factor (IGF)-binding protein-5 in binding IGF-I. J Biol Chem *278*, 17859-17866.

Shooter, G. K., Magee, B., Soos, M. A., Francis, G. L., Siddle, K., and 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*, 237-246.

Silha, J. V., and Murphy, L. J. (2002). Insights from insulin-like growth factor binding protein transgenic mice. Endocrinology *143*, 3711-3714.

Singh, B., Charkowicz, D., and Mascarenhas, D. (2004). Insulin-like growth factor-independent effects mediated by a C-terminal metal-binding domain of insulin-like growth factor binding protein-3. J Biol Chem *279*, 477-487.

Sitar, T., Popowicz, G. M., Siwanowicz, I., Huber, R., and Holak, T. A. (2006). Structural basis for the inhibition of insulin-like growth factors by insulin-like growth factor-binding proteins. Proc Natl Acad Sci U S A *103*, 13028-13033.

Siwanowicz, I., Popowicz, G. M., Wisniewska, M., Huber, R., Kuenkele, K. P., Lang, K., Engh, R. A., and Holak, T. A. (2005). Structural basis for the regulation of insulin-like growth factors by IGF binding proteins. Structure (Camb) *13*, 155-167.

Song, H., Beattie, J., Campbell, I. W., and Allan, G. J. (2000). Overlap of IGF- and heparin-binding sites in rat IGF-binding protein-5. J Mol Endocrinol *24*, 43-51.

Song, H., Shand, J. H., Beattie, J., Flint, D. J., and Allan, G. J. (2001). The carboxy-terminal domain of IGF-binding protein-5 inhibits heparin binding to a site in the central domain. J Mol Endocrinol *26*, 229-239.

Song, S. W., Fuller, G. N., Khan, A., Kong, S., Shen, W., Taylor, E., Ramdas, L., Lang, F. F., and Zhang, W. (2003). IIp45, an insulin-like growth factor binding protein 2 (IGFBP-2) binding protein, antagonizes IGFBP-2 stimulation of glioma cell invasion. Proc Natl Acad Sci U S A *100*, 13970-13975.

Sorrell, A. M., Shand, J. H., Tonner, E., Gamberoni, M., Accorsi, P. A., Beattie, J., Allan, G. J., and Flint, D. J. (2006). Insulin-like growth factor-binding protein-5 activates plasminogen by interaction with tissue plasminogen activator, independently of its ability to bind to plasminogen activator inhibitor-1, insulin-like growth factor-I, or heparin. J Biol Chem *281*, 10883-10889.

Spencer, E. M., and Chan, K. (1995). A 3-dimensional model for the insulin-like growth factor binding proteins (IGFBPs); supporting evidence using the structural determinants of the IGF binding site on IGFBP-3. Prog Growth Factor Res *6*, 209-214.

Standker, L., Braulke, T., Mark, S., Mostafavi, H., Meyer, M., Honing, S., Gimenez-Gallego, G., and Forssmann, W. G. (2000). Partial IGF affinity of circulating N- and C-terminal fragments of human insulin-like growth factor binding protein-4 (IGFBP-4) and the disulfide bonding pattern of the C-terminal IGFBP-4 domain. Biochemistry *39*, 5082-5088.

Standker, L., Wobst, P., Mark, S., and Forssmann, W. G. (1998). Isolation and characterization of circulating 13-kDa C-terminal fragments of human insulin-like growth factor binding protein-5. FEBS Lett *441*, 281-286.

Stejskal, E. O., and Tanner, J. E. (1965). Spin diffusion measurements: spin echoes in the presence of a time-dependent field gradient. J Chem Phys *42*, 288-292.

Stewart, C. E., and Rotwein, P. (1996). Growth, differentiation, and survival: multiple physiological functions for insulin-like growth factors. Physiol Rev *76*, 1005-1026.

Sun, Y., Hegamyer, G., and Colburn, N. H. (1993). PCR-direct sequencing of GC-rich region by inclusion of 10% DMSO: Application to C-jun. Biotechniques *15*, 372-374.

Taylor, V. L., and Spencer, E. M. (2001). Characterisation of insulin-like growth factor-binding protein-3 binding to a novel receptor on human platelet membranes. J Endocrinol *168*, 307-315.

Teller, D. C., Swanson, E., and de Haen, C. (1979). The translational friction coefficient of proteins. Methods Enzymol *61*, 103-124.

Terasawa, H., Kohda, D., Hatanaka, H., Nagata, K., Higashihashi, N., Fujiwara, H., Sakano, K., and Inagaki, F. (1994). Solution structure of human insulin-like growth factor II; recognition sites for receptors and binding proteins. Embo J *13*, 5590-5597.

Terrien, X., Bonvin, E., Corroyer, S., Tabary, O., Clement, A., and Henrion Caude, A. (2005). Intracellular colocalization and interaction of IGF-binding protein-2 with the cyclin-dependent kinase inhibitor p21CIP1/WAF1 during growth inhibition. Biochem J *392*, 457-465.

Torres, A. M., Forbes, B. E., Aplin, S. E., Wallace, J. C., Francis, G. L., and Norton, R. S. (1995). Solution structure of human insulin-like growth factor II. Relationship to receptor and binding protein interactions. J Mol Biol *248*, 385-401.

Twigg, S. M., Kiefer, M. C., Zapf, J., and Baxter, R. C. (2000). A central domain binding site in insulin-like growth factor binding protein-5 for the acid-labile subunit. Endocrinology *141*, 454-457.

Vajdos, F. F., Ultsch, M., Schaffer, M. L., Deshayes, K. D., Liu, J., Skelton, N. J., and de Vos, A. M. (2001). Crystal structure of human insulin-like growth factor-1: detergent binding inhibits binding protein interactions. Biochemistry *40*, 11022-11029.

Valentinis, B., Bhala, A., DeAngelis, T., Baserga, R., and Cohen, P. (1995). The human insulin-like growth factor (IGF) binding protein-3 inhibits the growth of fibroblasts with a targeted disruption of the IGF-I receptor gene. Mol Endocrinol *9*, 361-367.

Viles, J. H., Donne, D., Kroon, G., Prusiner, S. B., Cohen, F. E., Dyson, H. J., and Wright, P. E. (2001). Local structural plasticity of the prion protein. Analysis of NMR relaxation dynamics. Biochemistry *40*, 2743-2753.

Vorwerk, P., Hohmann, B., Oh, Y., Rosenfeld, R. G., and Shymko, R. M. (2002). Binding properties of insulin-like growth factor binding protein-3 (IGFBP-3), IGFBP-3 N- and C-terminal fragments, and structurally related proteins mac25 and connective tissue growth factor measured using a biosensor. Endocrinology *143*, 1677-1685.

Vorwerk, P., Mohnike, K., Wex, H., Rohl, F. W., Zimmermann, M., Blum, W. F., and Mittler, U. (2005). Insulin-like growth factor binding protein-2 at diagnosis of childhood acute lymphoblastic leukemia and the prediction of relapse risk. J Clin Endocrinol Metab *90*, 3022-3027.

Walenkamp, M. J., Karperien, M., Pereira, A. M., Hilhorst-Hofstee, Y., van Doorn, J., Chen, J. W., Mohan, S., Denley, A., Forbes, B., van Duyvenvoorde, H. A., *et al.* (2005). Homozygous and heterozygous expression of a novel insulin-like growth factor-I mutation. J Clin Endocrinol Metab *90*, 2855-2864.

Wang, G. K., Hu, L., Fuller, G. N., and Zhang, W. (2006). An interaction between insulin-like growth factor binding protein 2 and integrin alpha 5 is essential for IGFBP2-induced cell mobility. J Biol Chem *281*, 14085-14091.

Wang, J. F., Hampton, B., Mehlman, T., Burgess, W. H., and Rechler, M. M. (1988). Isolation of a biologically active fragment from the carboxy terminus of the fetal rat binding protein for insulin-like growth factors. Biochem Biophys Res Commun *157*, 718-726.

Weinzimer, S. A., Gibson, T. B., Collett-Solberg, P. F., Khare, A., Liu, B., and Cohen, P. (2001). Transferrin is an insulin-like growth factor-binding protein-3 binding protein. J Clin Endocrinol Metab *86*, 1806-1813.

Wider, G. (2000). Structure determination of biological macromolecules in solution using NMR spectroscopy. Biotechniques *29*, 1278-1294.

Wishart, D. S., Bigam, C. G., Holm, A., Hodges, R. S., and Sykes, B. D. (1995a). 1H, 13C and 15N random coil NMR chemical shifts of the common amino acids. I. Investigations of nearest-neighbor effects. J Biomol NMR *5*, 67-81.

Wishart, D. S., Bigam, C. G., Yao, J., Abildgaard, F., Dyson, H. J., Oldfield, E., Markley, J. L., and Sykes, B. D. (1995b). 1H, 13C and 15N chemical shift referencing in biomolecular NMR. J Biomol NMR *6*, 135-140.

Wolf, E., Kramer, R., Blum, W. F., Foll, J., and Brem, G. (1994). Consequences of postnatally elevated insulin-like growth factor-II in transgenic mice: endocrine changes and effects on body and organ growth. Endocrinology *135*, 1877-1886.

Wuthrich, K. (1986). NMR of proteins and nucleic acids).

Xu, Q., Yan, B., Li, S., and Duan, C. (2004). Fibronectin binds insulin-like growth factor-binding protein 5 and abolishes Its ligand-dependent action on cell migration. J Biol Chem *279*, 4269-4277.

Yakar, S., Leroith, D., and Brodt, P. (2005). The role of the growth hormone/insulin-like growth factor axis in tumor growth and progression: Lessons from animal models. Cytokine Growth Factor Rev *16*, 407-420.

Yan, X., Forbes, B. E., McNeil, K. A., Baxter, R. C., and Firth, S. M. (2004). Role of N- and C-terminal residues of insulin-like growth factor (IGF)-binding protein-3 in regulating IGF complex formation and receptor activation. J Biol Chem *279*, 53232-53240.

Yao, S., Headey, S. J., Keizer, D. W., Bach, L. A., and Norton, R. S. (2004). C-terminal domain of insulin-like growth factor (IGF) binding protein 6: conformational exchange and its correlation with IGF-II binding. Biochemistry *43*, 11187-11195.

Yao, S., Howlett, G. J., and Norton, R. S. (2000). Peptide self-association in aqueous trifluoroethanol monitored by pulsed field gradient NMR diffusion measurements. J Biomol NMR *16*, 109-119.

Yee, D. (2006). Targeting insulin-like growth factor pathways. Br J Cancer *94*, 465-468.

Zapf, J., Born, W., Chang, J. Y., James, P., Froesch, E. R., and Fischer, J. A. (1988). Isolation and NH2-terminal amino acid sequences of rat serum carrier proteins for insulin-like growth factors. Biochem Biophys Res Commun *156*, 1187-1194.

Zeslawski, W., Beisel, H. G., Kamionka, M., Kalus, W., Engh, R. A., Huber, R., Lang, K., and Holak, T. A. (2001). The interaction of insulin-like growth factor-I with the N-terminal domain of IGFBP-5. EMBO J *20*, 3638-3644.

Zhang, R. W., Simmons, D. J., Crowther, R. S., Mohan, S., and Baylink, D. J. (1991). Contribution of marrow stromal cells to the regulation of osteoblast proliferation in rats: evidence for the involvement of insulin-like growth factors. Bone Miner *13*, 201-215.

Zhang, W., Wang, H., Song, S. W., and Fuller, G. N. (2002). Insulin-like growth factor binding protein 2: gene expression microarrays and the hypothesis-generation paradigm. Brain Pathol *12*, 87-94.

Zhao, Y., Yin, P., Bach, L. A., and Duan, C. (2006). Several acidic amino acids in the N-domain of insulin-like growth factor-binding protein-5 are important for its transactivation activity. J Biol Chem *281*, 14184-14191.

Zhou, H., Casas-Finet, J. R., Heath Coats, R., Kaufman, J. D., Stahl, S. J., Wingfield, P. T., Rubin, J. S., Bottaro, D. P., and Byrd, R. A. (1999). Identification and dynamics of a heparin-binding site in hepatocyte growth factor. Biochemistry *38*, 14793-14802.

Zuiderweg, E. R. (2002). Mapping protein-protein interactions in solution by NMR spectroscopy. Biochemistry *41*, 1-7.