Technische Universität München

Max-Planck-Institut für Biochemie

Abteilung Strukturforschung

Biologische NMR-Arbeitsgruppe

Structural basis for the inhibition of insulin-like growth factors by insulin-like growth factor-binding proteins and structural and biochemical characterization of formins – the actin nucleating factors

Tomasz Sitar

Vollständiger Abdruck der von der Fakultät für Chemie der Technischen Universität München zur Erlangung des akademischen Grades eines

Doktors der Naturwissenschaften

genehmigten Dissertation.

Vorsitzender:	Univ.-Prof. Dr. St. J. Glaser
Prüfer der Dissertation:	1. apl. Prof. Dr. Dr. h.c. R. Huber, i. R.
	2. Univ.-Prof. Dr. Dr. A. Bacher

Die Dissertation wurde am 14.12.2006 bei der Technischen Universität München eingereicht und durch die Fakultät für Chemie am 03.05.2007 angenommen.

To Ania....

Acknowledgements

I would like to thank everyone who contributed to this work.

First of all, I would like to thank my supervisor Dr. Tad A. Holak, for his support, discussions and encouragement.

I am grateful to Professor Robert Huber for giving me the opportunity to work in his department and for being my 'Doktorvater'.

I would like to thank all the labmates: Ania Ducka; Grzegorz Popowicz, Ulli Rothweiler, Ola Mikolajka, Marcin Krajewski, Przemyslaw Ozdowy, Mahavir Singh, Kinga Brongel, Igor Siwanowicz, Sudipta Majumdar, Joma Joy, Loy D'Silva, Anna Czarny, Ola Szwagierczak, for creating a friendly atmosphere in the lab.

My special thanks go to Ania Ducka, Grzegorz Popowicz, Mahavir Singh and Ulli Rothweiler for their help in all matters, interesting discussions, and friendship.

Publications

Parts of this thesis have been or will be published in due course:

Tomasz Sitar, Grzegorz M. Popowicz, Igor Siwanowicz; Robert Huber, and Tad A. Holak
Structural basis for the inhibition of insulin-like growth factors by insulin-like growth factor-binding proteins
Proceedings of the National Academy of Science of the United States of America, 2006, 103: 13028-13033

Kinga Brongel, **Tomasz Sitar**, Igor Siwanowicz, Loy D'Silva, Joma Joy, Sue M. Firth, Robert Baxter, Robert Huber, and Tad A. Holak
Molecular architecture of the insulin-like growth factor-binding proteins
Manuscript submitted to Biochemistry

Marcin Krajewski, **Tomasz Sitar**, Shravan Kumar Mischra, Stefan Jentsch, and Tad A. Holak.
Identification of chemical shift changes in NMR spectra of the slowly exchanging ligand-protein interactions
Manuscript under preparation

Contents

1 Introduction

1.1 The structure and function of the IGF system

The insulin-like growth factor (IGF) system is a conserved signaling pathway that is composed of two IGF ligands, two IGF receptors, and six IGF high-affinity binding proteins (Figure 1.1.1). The IGF-1 and IGF-2 bind to the insulin/IGF family of cell surface receptors and activate their intrinsic tyrosine kinase domain. The family of high affinity IGF binding proteins (IGFBPs) modulate the availability of IGF-1 and -2 to bind the receptors. All three components of the IGF system act together to control a number of biological processes including cellular growth, proliferation, differentiation, survival against apoptosis and migration. These processes are involved in tissue formation and remodeling, bone growth, brain development, and regulation of metabolism.

1.1.1 Insulin-like growth factors (IGF-1 and IGF-2)

Insulin like-growth factors IGF-1 and IGF-2, are evolutionarily conserved polypeptides (Duan, 1997, 1998). The mature IGF-1and IGF-2 are, respectively, 70 and 67 amino acid single chain peptides, which consist of A, B, C, and D domains. The IGF A and B domains are homologous to insulin A and B chains (50% sequence similarity), respectively. Several three-dimensional structures of IGFs by both NMR and X-ray crystallography have been resolved (Cooke et al., 1991; Sato et al., 1993; Schaffer et a., 2003; Vajdos et al., 2001). The overall structure of IGF-1 and IGF-2 within the A and B domains is similar to the crystal structure of insulin (Bentley et al., 1976; Baker et al., 1988), and the NMR structure of proinsulin (Weiss et al., 1990). The major secondary structural elements of IGF-1, IGF-2, and insulin are α-helical. The A domain contains helix 2 (Ile43–Cys47 of IGF-1; Glu44–Phe48 of IGF-2) and helix 3 (Leu54–Glu58 of IGF-1; Ala54–Tyr59 of IGF-2) whereas the B domain is built of helix 1 (Gly8–Cys18 of IGF-1; Gly10–Val20 of IGF-2). The IGF C and D domains are unstructured and highly flexible in solution.

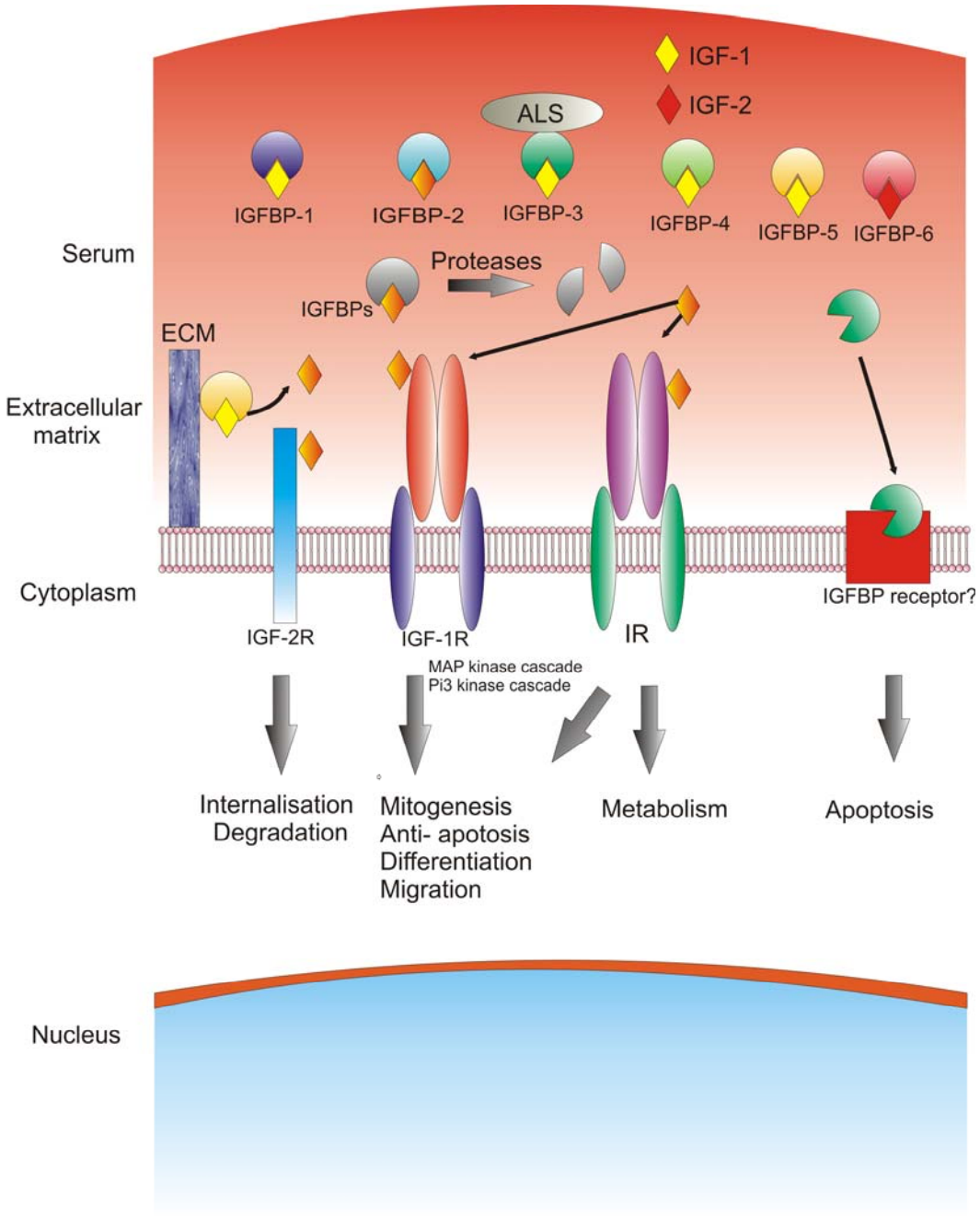


Figure 1.1.1. Schematic representation of the IGF system. IGFs circulate mainly in an IGFBP-3:IGF:ALS complex. Release of IGFs from IGFBPs occurs upon IGFBP proteolysis or extracellular matrix binding. The IGFBP-5 can act independently of IGF entering the cell via undefined receptor.

The three-dimensional fold is stabilized by three disulphide bonds (Cys6–Cys48; Cys18–Cys61; Cys47–Cys52 for IGF-1). The truncated form of IGF-1, known as DES(1-3)IGF-1, has been found in fetal and adult human brain (Carlsson–Skwirut et al., 1986; Sara et al., 1986; Humbel, 1990). The DES(1-3)IGF-1 is the product of differential processing of pro-IGF-1 lacking the first three residues at the amino terminus: Gly-Pro-Glu. The biological potency of this truncated form is 10 times higher than that of the full-length form and is explained by reduced binding to IGF-binding proteins (Francis et al., 1988; Beck et al., 1993; Carlsson–Skwirut et al., 1989; Ballard et al., 1996). The DES-(1-3)IGF-1 binds the IGFBP-3 with several times lower affinity than full-length IGF-1 and shows greatly reduced binding to other IGFBPs (Forbes et al., 1988). Mutational analysis showed that Glu3 is an important determinant of binding, because variants like IGF-1Glu3Arg and IGF-1Glu3Gln Thr4Ala show considerably reduced affinity to IGFBPs. Alanine scanning mutagenesis of IGF-1 identified also Gly7, Leu10; Val17 and Phe25 as residues important for IGFBPs binding (Dubaquie et al., 1999). Other important IGFBP-binding determinants of IGF-1, as revealed by mutagenesis experiments, include Gln15 and Phe16 in the B domain of IGF-1 and the A domain residues Phe49, Arg50, and Ser51. Substitution of these amino acids to the corresponding residues in insulin considerably reduces the IGFBP binding (Clemmons et al., 1990). In IGF-2, mutation of Phe26 in the B domain has a pronounced effect on binding to all six of the IGFBPs, most notably IGFBP-1, and, as with the corresponding residues of IGF-1, the residues Phe48, Arg49, and Ser50 are also important (Bach et al., 1993).

In mammals, IGFs are widely expressed during fetal and prenatal stages. In postnatal stages, hepatic production of IGF-1 under the regulation of growth hormone (GH) becomes the major source of circulating IGF-1, however, both IGF-1 and IGF-2 are expressed in many non-hepatic tissue (LeRoith et al., 2001). Despite the high structural similarity between IGF-1, IGF-2, and insulin each ligand result in unique signaling outcomes. At the cellular level, IGFs stimulate cell proliferation, differentiation, migration, survival, metabolism, and contractility (Jones and Clemons, 1995; LeRoith et al., 2001).

1.1.2 Insulin-like growth factor binding proteins (IGFBPs)

In the extracellular environment most, if not all, IGFs are bound to IGF-binding proteins. The IGFBP family comprises six soluble proteins (IGFBP-1 to -6) containing 216-289 residues that bind to IGFs with nM affinities (Firth et al., 2002; Clemmons, 2001; Bach et al., 2005; Bunn and Fowlkes, 2003). Because of their sequence homology, IGFBPs are assumed to share a common overall fold and are expected to have closely related IGF binding determinants. The IGFBPs, with apparent molecular mass 24-45 kDa, share a common domain organization (Figure1.1.2). Each IGFBP can be divided into three distinct domains of approximately equal lengths: highly conserved cysteine-rich N- and C-domains and a central linker domain unique to each IGFBP species. This domain structure is highly conserved among this gene family and across species (Duan et al., 1999; Maures and Duan, 2002). The N-domains of IGFBPs 1-5 contain six disulphide bridges and share conserved GCGCC motif. The IGFBP-6 does not contain the two adjacent cysteines in this motif, therefore the first three N-terminal disulphides bonds differ from those of the other IGFBPs (Neumann et al., 1999). C-domains of all IGFBPs have six conserved cysteines, which form three disulphide linkages. Both the N- and C-domains participate in the binding to IGFs, although the specific roles of each of these domains in IGF binding have not been decisively determined (Firth et al., 2002; Clemmons, 2001; Bach et al., 2005; Bunn and Fowlkes, 2003; Payet et al., 2003; Allan et al., 2006; Carrick et al., 2005; Kibbey et al., 2006; Fernandez-Tornero et al., 2005; Siwanowicz et al., 2005; Headey et al., 2004). The carboxyl-terminal domain may be responsible for preferences of IGFBPs for one species of IGF over the other. The IGF binding is the only one clearly identified function of the N-domains, whereas C-domains are implicated in a wide range of functions and interact with a large number of proteins (Table 1.1.1) and other biologically significant molecules that can either modulate IGF-dependent actions or mediate IGF-independent actions (Bach et al., 2005). Numerous IGF-independent effects mediated by C-domains have been described *in vitro* (Firth and Baxter, 2005; Ricort, 2004), and *in vivo (*Miyakoshi et al., 2001), including growth inhibition, promotion of apoptosis, and

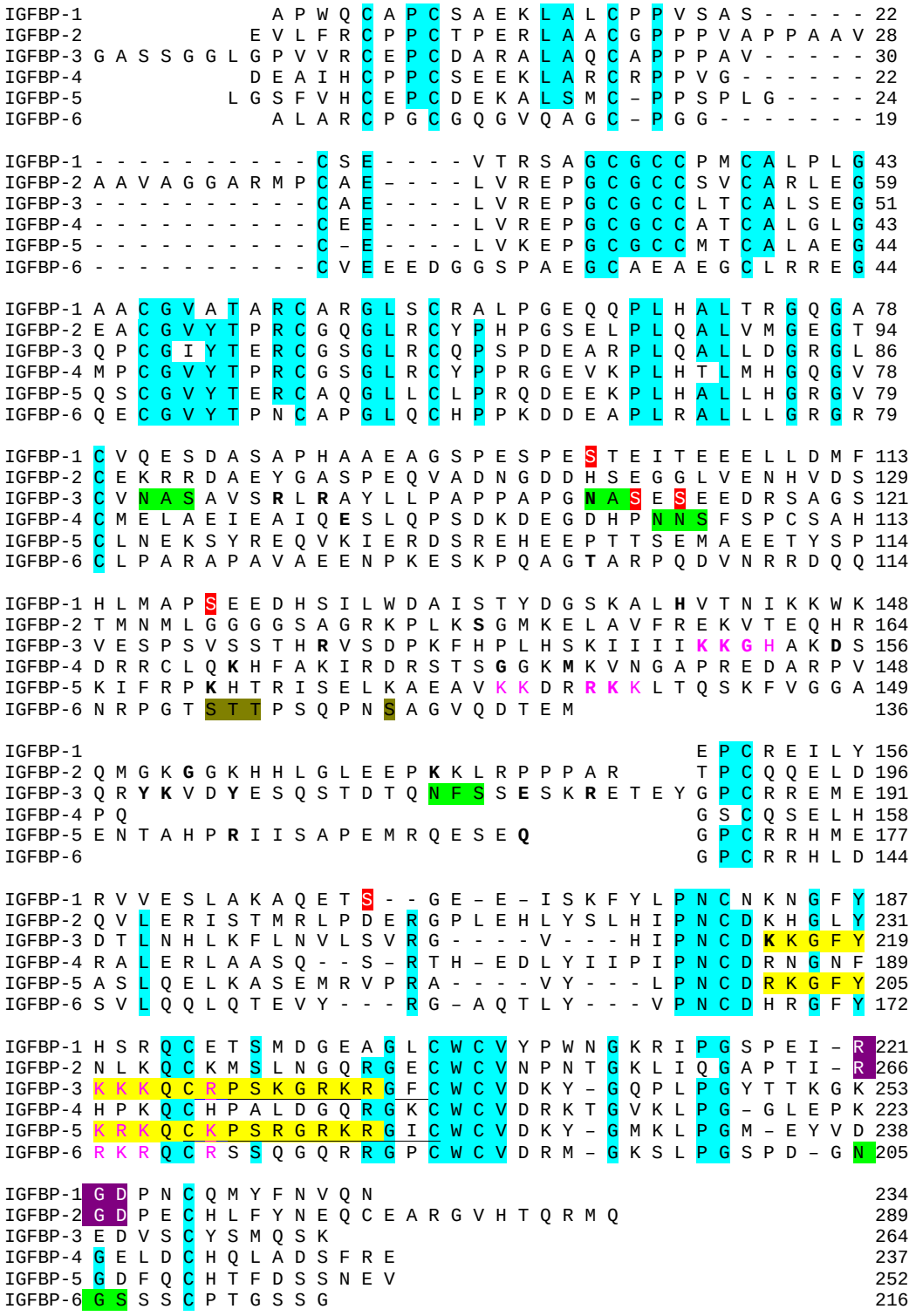


Figure 1.1.2. Sequence alignment of human IGFBP-1 to -6. Conserved residues are indicated by cyan shading; yellow – nuclear localization sequences (NLS) (Schedlich et al., 1998); purple – potential integrin-binding sequences; green –

potential N-glycosylation sites; dark yellow – identified O-glycosylation sites; red – phosphorylated serines (Jones et al., 1995; Hoeck et al., 1994), lettering in magenta - heparin-binding domains (Booth et al., 1995; Knudtson et al., 2001); underlined – metal binding domain (MBD) (Singh et al., 2004). In bold – amino acids, after which cleavage occurs (Binoux et al., 1999).

modulation of cell adhesion and migration. C-domains of human IGFBP-1 and IGFBP-2 contain Arg-Gly-Asp sequence which binds to the $\alpha_5\beta_1$integrin and stimulate cell migration (Jones et al., 1993). C-domains of IGFBP-3 and IGFBP-5 have nuclear localization sequences (NLS) that interact with the importin β-subunit (Schedlich et al., 2000). Although the role of the NLS of IGFBPs is not understood, it might have a role in the IGF-independent promoting apoptosis and facilitating IGF transport to the nucleus. Most C-domains of IGFBPs contain heparin-binding domains (XBBXBX or XXBBBXXBX, where B represents a basic amino acids and X other amino acids). The basic regions of IGFBP-3, -5, -6 are implicated in many non-IGF interactions like binding of glycosaminoglycans (GAGs). This interaction is important because it can result in cell association of IGFBPs, which has been linked to potentiation of IGF actions. Binding by GAGs reduces the IGF-binding affinity to IGFBPs (Firth and Baxter, 2002; Mohan and Baylink, 2002). The C-domain of IGFBP-5 binds several extracellular matrix (ECM) proteins, including PAI-1, thrombospondin and osteopontin (Nam et al., 2000) IGFBP-3 binds and modulates the retinoid X receptor-α , interacts with TGFβ signaling through Smad proteins, and influences other signaling pathways (Fanayan et al., 2002).

The central linker domain is the least conserved region and has never been cited as part of the IGF-binding site for any IGFBP. This domain is the site of post-translational modifications, specific proteolysis (Bunn and Fowlkes, 2003), and the acid-labile subunit (ALS) (Firth and Baxter, 2002) and ECM associations (Firth and Baxter, 2002; Clemmons, 2001; Xu et al., 2004) known for IGFBPs. Proteolytic cleavage in this domain is believed to produce low affinity N- and C-terminal fragments that cannot compete with IGF receptors for IGFs and thus the proteolysis is assumed to be the predominant mechanism for IGF release from

IGFBPs (Bunn and Fowlkes, 2003; Fernandez-Tornero et al., 2005). However recent studies indicate that the resulting N- and C-terminal fragments still can inhibit IGF activity and have functional properties that differ from those of the intact proteins (Firth and Baxter, 2002; ,Bach et al., 2005; Allan et al., 2006; Fernandez-Tornero et al., 2005).

Table 1.1.1. Binding partners of the C-domains of IGFBPs

Molecule	IGFBP	Location of molecule	Reference
IGF dependent			
IGF-1, -2	1-6	Serum and EC	(Firth and Baxter, 2002)
Acid-labile subunit	3,5	Serum	(Firth et al., 1998)
Glycosaminoglycans	3,5,6	EC and cell membrane	(Firth and Baxter, 2002
IGF independent			
Plasminogen	3	Serum	(Campbell et al., 1998)
Transferin	3	Serum	(Lee et al., 2004)
Ilp45	2	EC	(Song et al., 2004)
Fibrinogen-fibrin	3	EC	(Campbell et al., 1999)
Humanin	3	EC	(Ikonen et al., 2003)
Metal ions	3	EC	(Singh et al., 2004)
Collagen 1α	3	ECM	(Liu et al., 2003)
Fibronectin	5	ECM	(Xu et al., 2004)
Osteopontin	5	ECM	(Schedlich et al., 2000)
PAI-1	5	ECM	(Nam et al., 1997)
Thrombospondin	5	ECM	(Nam et al., 2000)
Vitronectin	5	ECM	(Nam et al., 2002)
Caveolin	3	Cell membrane	(Lee et al., 2004)
CRS binding protein-1	3	Cell membrane	(Huang et al., 2003)
Integrins	1,2	Cell membrane	(Jones et al., 1993)
RXR-α	3	Nucleus	(Liu et al., 2004)
Importin-β	3,5	Cytosol	(Schedlich et al., 2000)

EC – extracellular; ECM – extracellular matrix; RXR – retinoid X receptor; CRS – cell surface retention binding protein

The structure of the N-terminal domain of IGFBP-5, free (Kalus et al., 1998) and complexed to IGF-1 (Zeslawski et al., 2001), was solved some time ago.

Most recently the structure of the binary complex (NBP-4(3-82)/IGF-1) has been resolved (Siwanowicz et al., 2005). The global folds of NBP-4 (residues Ala39–Glu82) and miniNBP-5 (residues Ala40–Glu83) are almost identical in both structures. NBP-4 has L-like shape and covers both the N- and C-terminal parts of IGF-1. The core of the NBP4(3–38) subdomain is stabilized by a short two-stranded β sheet and four disulphide bridges forming a disulphide bond ladder-like structure. This structure is connected to the miniNBP fragment by a short stretch of amino acids (Ala39, Leu40, Gly41). The miniNBP is globular, whereas in NBP-4(3–38), the β sheet and disulphide bridges are all in one plane. The two subdomains are perpendicular to each other, creating the 'L' shape for the whole N-terminal domain (Figure 1.1.3).

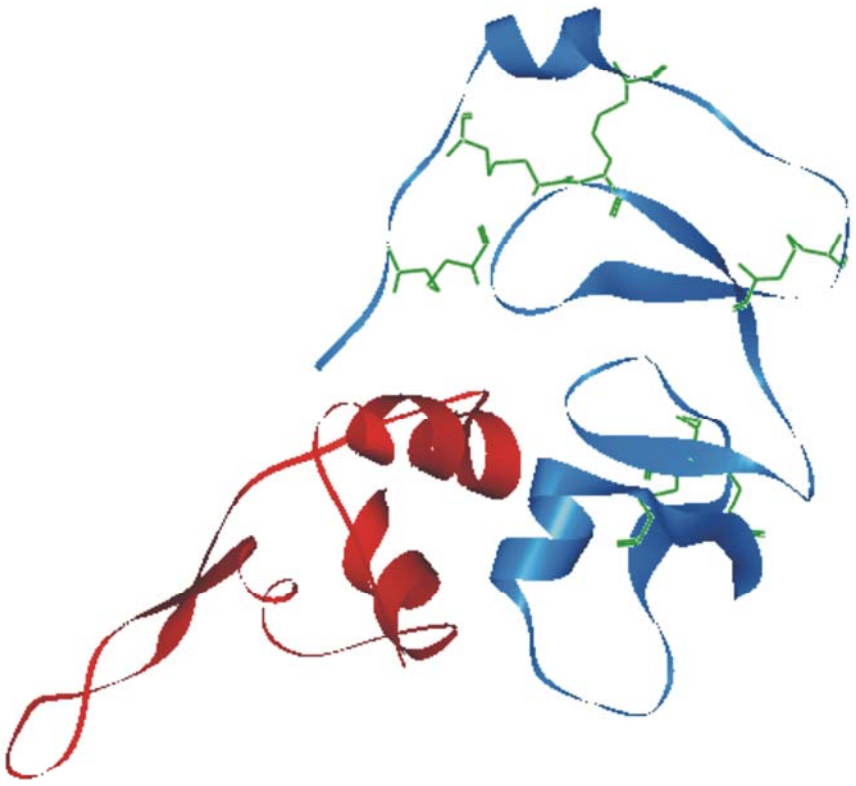


Figure 1.1.3. Overall structure and folding of the binary complex NBP-4(3-82) and IGF-1 (adapted from pdb file 1WQJ). NBP-4(3-82) is shown in blue; IGF-1 in red; disulphide bridges in NBP-4 are shown as green sticks .

The low-resolution structures of the C-terminal domain of IGFBP-6 (Headey et al., 2004) and its binding surface on IGF-2 (Bach et al., 2005; Headey et al., 2004) have been determined with NMR spectroscopy, and also recently the X-ray structure of the isolated C-terminal fragment of IGFBP-1 has been solved (Figure 1.1.4) (Sala et al., 2005). So far, there was no crystal structure of a ternary complex of the C-terminal domain of any IGFBPs bound to both the N-terminal domain and IGF. In this thesis, the X-ray structures of the ternary complex of the N- and C-terminal domains of IGFBP-4 bound to IGF-1, and the ternary complex of the N-terminal domain of IGFBP-4, IGF-1, and C-terminal domain of IGFBP-1 are reported (see *Results and discussion*).

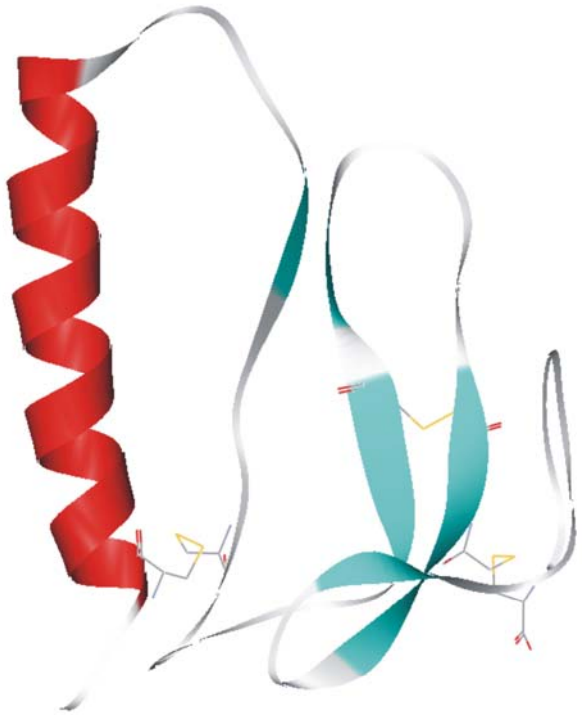


Figure 1.1.4. Overall structure and folding of the C-terminal domain of IGFBP-1 isolated from the human amniotic fluid (adapted from pdb file 1ZT3). The α-helix is red, β-strands are shown in blue. The three disulphide bonds are shown as sticks (Sala et al., 2005).

1.1.3 Insulin-like growth factor receptors (IGF-1R and IGF-2R)

IGF-1, IGF-2, and insulin interact with specific cell surface receptors (IGF-1R, IGF-2R, IR). The physiological functions of insulin and IGFs are very different, while their receptor structures are similar. The IGF-1R is a membrane glycoprotein of 400-450 kDa, consisting of two α-subunits (135 kDa each) and two β-subunits (90 kDa each) (Rechler, 1985; Yamasaki et al., 1993). Disulphide bonds connect both α- and β-subunits to form a functional heterotetrameric receptor complex. In analogy with the insulin receptor, IGF-1 receptor subunits are encoded within a single 180 kDa polypeptide precursor that is glycosylated, dimerised and proteolytically processed to yield the mature $\alpha_2\beta_2$ form of the receptor. The α-subunit is entirely extracellular and contains the ligand-binding site, a cysteine-rich domain. The β-subunit contains the hydrophobic transmembrane domain with a short extracellular region, and a tyrosine kinase domain in its cytoplasmic portion. The structure of the first three domains of the extracellular portion of IGF-1R (L1-CR-L2, residues 1-462) has been determined to a 2.6 Å resolution (Garrett et al., 1998). The L domains each adopt a compact shape consisting of a single-stranded right-handed β-helix. The Cys-rich region is composed of eight disulphide-bonded modules, seven of which form a rod-shaped domain. This fragment of the receptor is not active in terms of IGF binding. The crystal structure of the insulin receptor ectodomain has been resoled recently (McKern et al., 2006). The structure reveals the domain arrangement in the disulphide-linked ectodomain dimer, showing that the insulin receptor adopts a folded-over conformation that places the ligand-binding regions in a juxtaposition. Each ectodomain monomer contains a leucine-rich repeat (L1) domain, a cysteine-rich (CR) region and a second leucine-rich repeat (L2) domain, followed by three fibronectin type III domains (FnIII-1 to FnIII-3). FnIII-2 contains an insert domain (ID) of 120 residues, within which lies the α-β cleavage site. Each α-β monomer has en inverted 'V' layout with respect to the cell membrane. One leg of the 'V' is formed by the L1, CR and L2 domains; the other is formed by an extended linear arrangement of the three FnIII domains. In the ectodomain homodimer, the L2 domain of the first monomer contacts the FnIII-1

domain of the second monomer at the apex of the inverted 'V', whereas the L1 domain of the first monomer contacts the FnIII-2 domain of the second at the approximate midpoint of one of the two legs of the inverted 'V'. This arrangement which is significantly different from previous models (De Meyts and Whittaker, 2002), shows that the L1domains are on opposite sides of the dimer, too far apart to allow insulin to bind both L1 domains simultaneously (Figure 1.1.5).

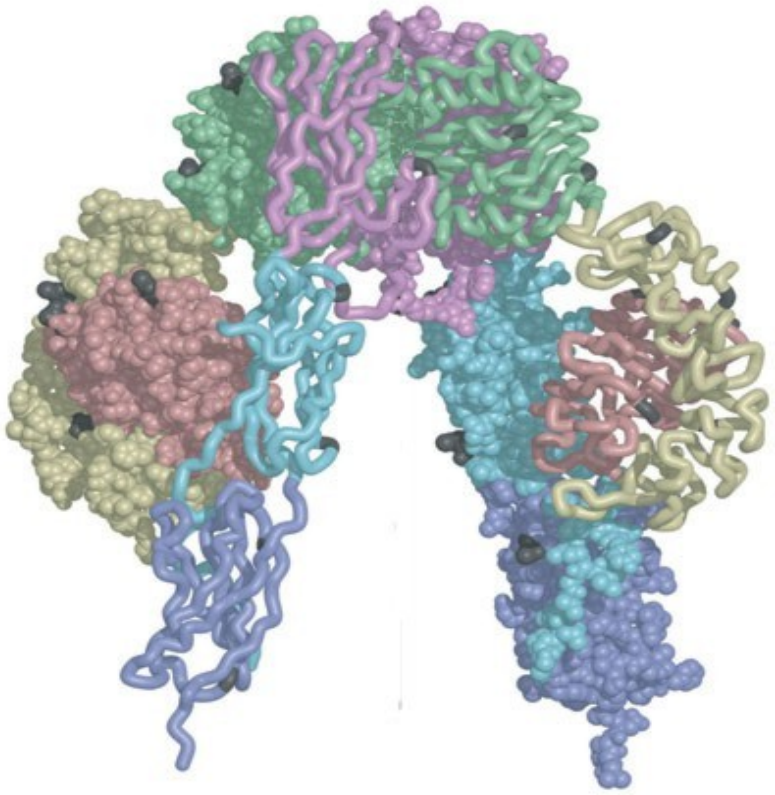


Figure 1.1.5. The IR ectodomain homodimer, showing juxtaposition of domains between the monomers. One monomer is shown in tube representation, the other in atomic sphere representation. L1 domain is in brown; CR in yellow; L2 in green; FnIII-1 in magenta; FnIII-2 in cyan; FnIII-3 in blue. The potential N-linked glycosylation sites are shown in black (McKern et al., 2006).

IGF-1 binding to extracellular α-subunit of IGF-1R causes autophosphorylation of three tyrosines in the activation loop of the tyrosine kinase domain in the cytoplasmic portion of the β-subunit, which results in

amplification of tyrosine kinase activity and further autophosphorylation of additional tyrosine residues. These phosphotyrosine-containing motifs are binding sites for adaptor and effector molecules in receptor signaling pathways, including insulin receptor substrates and Src homology/collagen (Shc), which are subsequently phosphorylated on their tyrosines (White and Kahn, 1998; Kim et al., 1998). The insulin receptor substrates (IRSs) are known as “docking” proteins and constitute a family of four structurally related adaptor proteins that can link the IGF-1 receptor to downstream signal transduction mediators regulating cellular growth. IRS 1 is the most extensively studied, and has multiple tyrosines, which associates with SH2 domain-containing proteins including the growth factor receptor bound-2 protein (Grb2) and the p85 regulatory subunit of phosphoinositol-3 (PI-3) kinase, p110. Phosphorylation of IRS 1 and 2 leads to activation of two downstream signaling cascades: the mitogen-activated protein kinase (MAPK) and the phosphatidylinositol 3-kinase (P3K) cascades.

1.1.4 The IGF system and diseases

Insulin like-growth factors are implicated in many common diseases including cancer, atherosclerosis, and diabetic complications. Epidemiological studies show that increased risks of breast, prostate, colorectal, and lung carcinomas are associated with increasing serum concentrations of IGF-1 (Furstenberger and Senn, 2002). These findings were confirmed in animal models, where reduced circulating IGF-1 levels result in significant reductions in cancer development, growth, and metastases, whereas increased circulating IGF-1 levels are associated with enhanced tumor growth (Wu et al., 2003). In humans, a homozygous partial deletion of the IGF-1 gene is associated with mental retardation and sensorineural deafness, in addition to fetal and postnatal growth retardation (Woods et al., 1996). Reduced circulating IGF-1 levels are associated with type I diabetes, and the IGF-1 treatment improves glucose and protein metabolism and attenuates diabetic cardiomyophaty (Carrol et al., 2000; Norby et al., 2002). In the central nervous system, expression of IGF-1 and the

IGF-1R are induced by brain injury, and exogenous administration of IGF-1 after injury ameliorates the damage (Guan et al., 2003). Accumulation of β-amyloids in the brain during aging is associated with decreased levels of IGF-1 in serum. IGF-1 treatment delays progression of amyotrophic lateral sclerosis (ALS) in a mouse model (Kaspar et al., 2003). Recent studies suggests that insulin/IGF signaling is required for male sex determination: XY mice deficient in the insulin receptor (IR), IGF-1R receptor and IR-related receptors (IRR) have a completely female phenotype (Nef et al., 2003). Other diseases, like hepatocellular and cutaneous carcinomas, and breast carcinomas, are linked to the overexpression of IGF-1R. IGF-1R is overexpressed in many diverse tumor types and is a critical signaling molecule for tumor cell proliferation and survival. Therapeutic strategies targeting the IGF-1R by eliminating it from the cell membrane, blocking the interaction with IGFs, or interrupting the signal transduction pathway downstream of IGF-1R, may therefore be an effective broad-spectrum anticancer action (Yu and Rohan, 2000; LeRoith et al., 1995). Inhibition of IGF/IGF-receptor binding interferes with cell growth and represents a strategy for the development of IGFBPs and their variants as natural IGF antagonists in many common diseases that arise from disregulation of the IGF system (Firth and Baxter, 2002; Kibbey et al., 2006; Pollak et al., 2004).

1.2 Formins

1.2.1 Domain organization

Eucaryotic cells rely on de novo nucleation of actin filaments in order to elicit temporal and spatial remodeling of the actin cytoskeleton. New filaments are created by different nucleating proteins (Figure 1.2.1) and nucleation is the most limiting step in actin filament polymerization. Three various classes of actin nucleators are known until now:

- The Arp2/3 complex
- Spire
- The formin-homology protein

These three groups of proteins employ different mechanisms of nucleation. The Arp2/3 complex binds to the side of a mother filament and initiates a branch by mimicking the pointed end of a new daughter filament. Spire binds and stabilizes the pointed end of an actin filament, leaving the barbed end free to elongate. Both mechanisms of nucleation allow to create only short filaments because the free barbed ends are rapidly bound by capping proteins, which blocks monomer addition. Long unbranched filaments are formed by formins, which remain processively (continuously) associated with the elongating barbed ends of actin filaments. This mechanism protects the growing barbed end from capping proteins and allows the FH1FH2-profilin-mediated actin assembly.

Formins are large, multi-domain proteins, usually more than 1000 amino acids long. The defining feature of formins is the formin homology 2 domain (FH2), an ~400 amino acid conserved sequence, and the adjacent variable-length proline-rich FH1 domain. The N-terminal regulatory region contains a GTPase-binding domain (GBD) followed by an adjacent Diaphanous-inhibitory domain (DID) and a dimerization domain (DD). A structurally less defined region, following the GBD and containing both the DID and DD, previously referred as a FH3 domain is believed to be implicated in subcellular localization. The C-terminal Diaphanous auto-regulatory domain (DAD), which is composed of only a small stretch of amino acid residues, is involved in autoregulation. The C-terminal

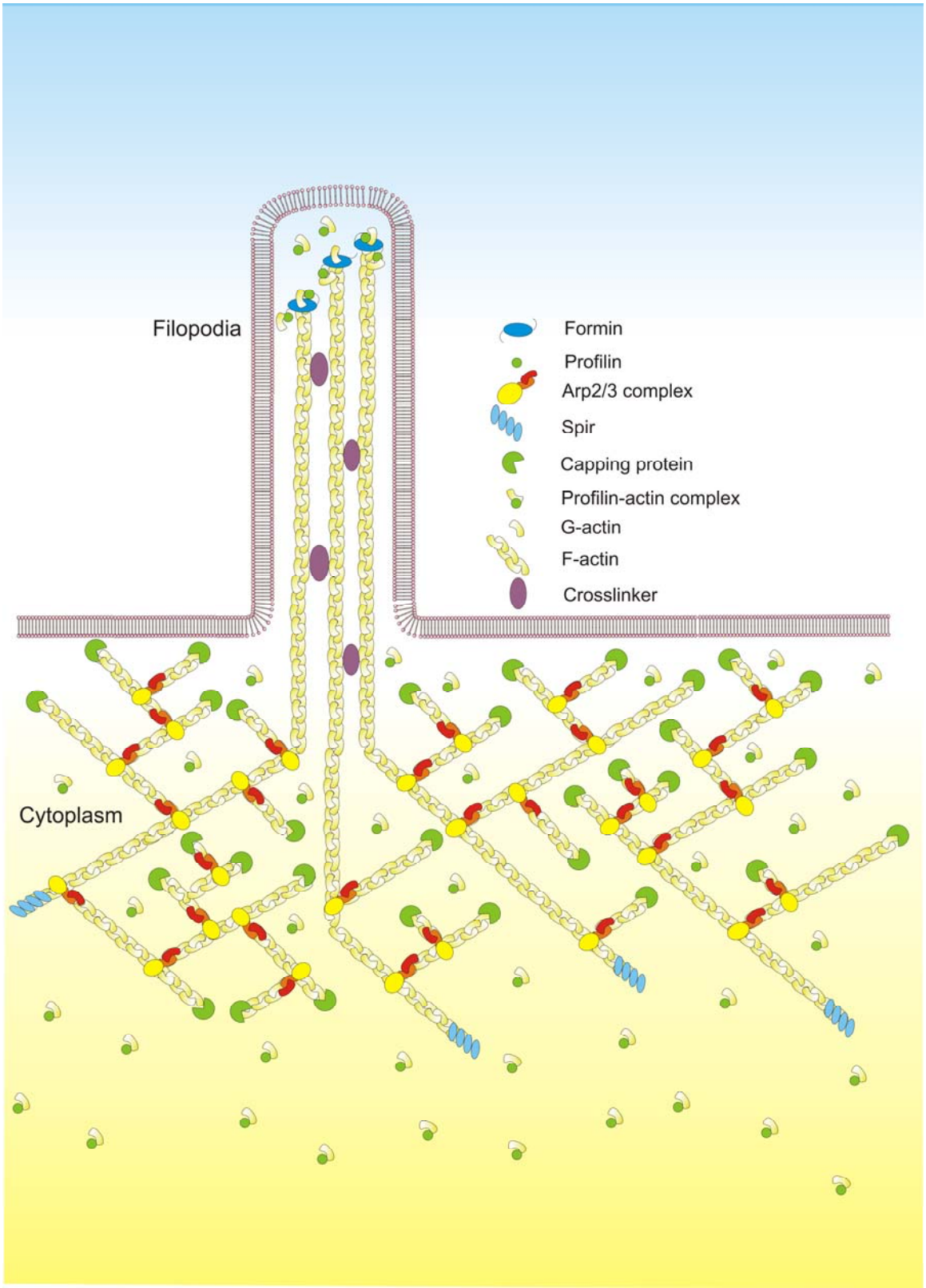


Figure 1.2.1. Actin network in the leading edge of the motile cell. Filaments are nucleated by three different nucleating proteins: formins, spire, and Arp2/3.

Diaphanous auto-regulatory domain (DAD),which is composed of only a small stretch of amino acid residues, is involved in autoregulation. A schematic domain organization and roughly estimated lengths of various formin subfamilies are presented in Figure 1.2.2.

Figure 1.2.2. Domain organization of metazoan formins. Domains depicted are GTPase binding domain (GBD), diaphanous inhibitory domain (DID), dimerization domain (DD), coiled-coil (CC), formin homology 1 domain (FH1), Formin homology 2 domain (FH2), diaphanous-autoregulatory domain (DAD), PDZ (Psd/Dlg/Zo-1). *-FHOD proteins might contain auto-inhibitory regions that behave similarly to DIDs but have diverged significantly in primary structure.

Metazoan formins fall into seven groups:

- Dia (diaphanous)
- DAAM (dishevelled-associated activator of morphogenesis)
- FRL (formin-related gene in leukocytes)

- FHOD (formin homology domain-containing protein)
- INF (inverted formin)
- FMN (formin)
- Delphilin

The subfamilies with the general structure GBD/FH3-FH1-FH2-DAD are called conventional formins. The formin homology 2 (FH2) domain is present in all formins while other domains might be absent in particular members of the formin family or in alternatively spliced variants. GBD/FH3, DAD or both can be identified, with a very few exceptions, in almost all conventional formins, including fungal and Dictyostelium formins.

Dictyostelium discoideum genome comprises ten formin genes, which belong to conventional formins. Although Dictyostelium formins vary considerably in length, with two exceptions, they have in common a core of about 1100 residues that harbors a GBD/FH3-FH1-FH2-DAD structure. Features of Dictyostelium formins are shown in Table 1.2.1.

Table 1.2.1. Features of Dictyostelium discoideum formins (Rivero et al., 2005).

Formin	Dictybase ID	Chromosome	No. of XPPPPP motifs inFH1	No. of residues
forA	DDB0214996	3	8	1218
forB	DDB0215000	3	4	1126
forC	DDB0191362	5	0	1158
forD	DDB0205290	3	1	1214
forE	DDB0190413	1	4	1561
forF	DDB0188569	5	5	1220
forG	DDB0169087	2	2	1074
forH(dDia2)	DDB0186588	4	2	1087
forI*	DDB0186053	4	2	935
forJ	DDB0183855	6	4	2546

*- forI does not contain GBD/FH3 and DAD domains

A phylogenetic tree based on the alignment of complete sets of sequences of FH2 domains from various organisms shows high conservation of this domain. On average, Dictyostelium formins are 45,5% similar and 23,8% identical to each other, except forC, which is more divergent (40,0% similarity, 20,4% identity). A comparable degree of similarity (identity) is observed to members of metazoan formins, and ranged between 40%(20%) and 48%(24%) (Rivero et al., 2005; Higgs, 2005).

1.2.2 Molecular regulation of formins

Conventional formins (Dia, DAAM, FRL) are regulated by autoinhibition through the interaction between the N- and C-termini (Figure 1.2.3). The FH1FH2 domains are flanked on either side by regulatory domains. Important for this interaction is the diaphanous auto-inhibitory domain (DAD), a stretch of 20-30 amino acids, located C-terminally to the FH2 domain. Residues at the N terminus of DAD are relatively conserved with a consensus sequence: G-X-X-M-D-X-L-L-X-X-L, while the C terminus is less well defined and is often basic. The N-terminal GBD-DID-DD region is defined more precisely and binds DAD with sub-μM affinity what is sufficient for autoinhibition. In a basal state, formins exist as auto-inhibited proteins through intramolecular interactions between DID and DAD. Upon binding of an activated Rho GTPase, to the GBD domain, the DAD is released and the ability of formin to elongate unbranched actin filaments is induced. The crystal structure of mDia1 revealed that the DID-DD region forms a stable homodimer and associates together with GBD into a joined folding unit containing armadillo repeats (Otomo et al., 2005). The other recently solved crystal structure of Rho GTPase in complex with the regulatory N-terminus of mDia1 containing the GBD-DID-DAD region shows that Rho uses its switch I and II regions for the interaction with both portion of GBD and DID domains (Nezami et al., 2006; Rose et al., 2005). Although the binding of Rho and DAD at the N -terminus of mDia1 is mutually exclusive, the binding sites are only partially overlapping. The DID-DAD auto-inhibitory interaction is released after the Rho-

induced restructuring of GBD, which interferes with binding of DAD to the neighboring DID domain.

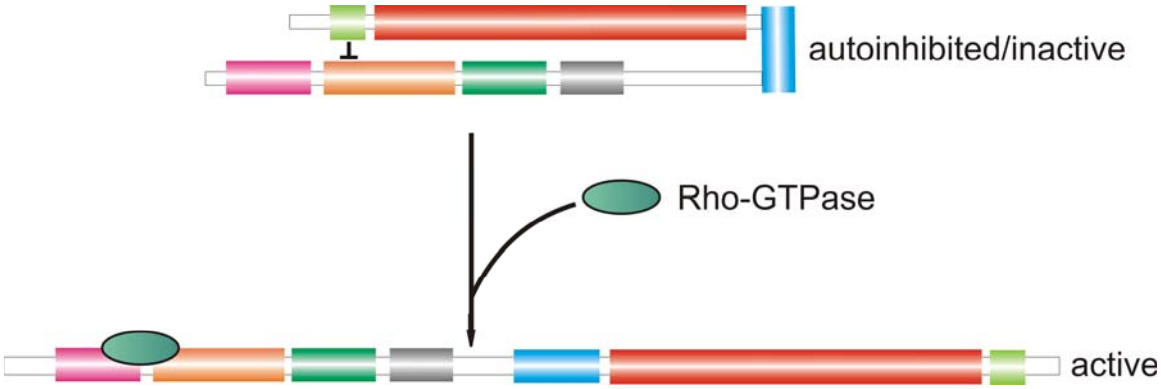


Figure 1.2.3. Molecular regulation of Diaphanous-related formins. Autoinhibition of DRFs, caused by the interaction of DAD with DID, is relived by association of an active, GTP bound Rho GTPase to the GBD, what induces the release of DAD leading to the activation of DRF. The interacting domains are: the diaphanous inhibitory domain (DID) shown in green, the diaphanous-autoregulatory domain (DAD) shown in orange, and Rho-GTPase (dark green ellipse).

Addition of a GTP bound Rho usually does not fully activate formins in vitro, suggesting that either the autoinhibited state might be formed also by different interactions, or that additional signals are required for full activation (Li and Higgs, 2003; 2005). A regulatory mechanism of other formins is still unknown and might be different. The FHOD formins possess the DAD domain but not a clear DID sequence, although some auto-inhibitory regions behave similarly to DID but have diverged significantly in their primary structures (Gasteier et al., 2003; Westendorf et al., 2001). Other metazoan formins (FMN, INF, delphilin) contain neither DAD or DID. Delphilin contain a N-terminal PDZ domain, which binds to the cytoplasmatic region of the ionotropic glutamete receptor but its role, if any, is unknown (Miyagi et al., 2002).

1.2.3 Biochemical and structural properties of formin homology 1 and 2 domains

The FH2 domain is the best-conserved domain that is present in all subfamilies of formins. Usually the FH2 domain is necessary and sufficient to nucleate actin polymerization from G-actin in vitro (Kovar at al., 2003). Two crystal structures of the FH2 domain have been solved recently. The three-dimensional structure of the Bni1pFH2 domain revealed a flexible, tethered dimer architecture, in which two elongated actin binding heads are tied together at either end to form a doughnut-shaped circular structure with large central hole (Xu et al., 2004). The FH2 domain fold is almost entirely α-helical (Figure 1.2.4).

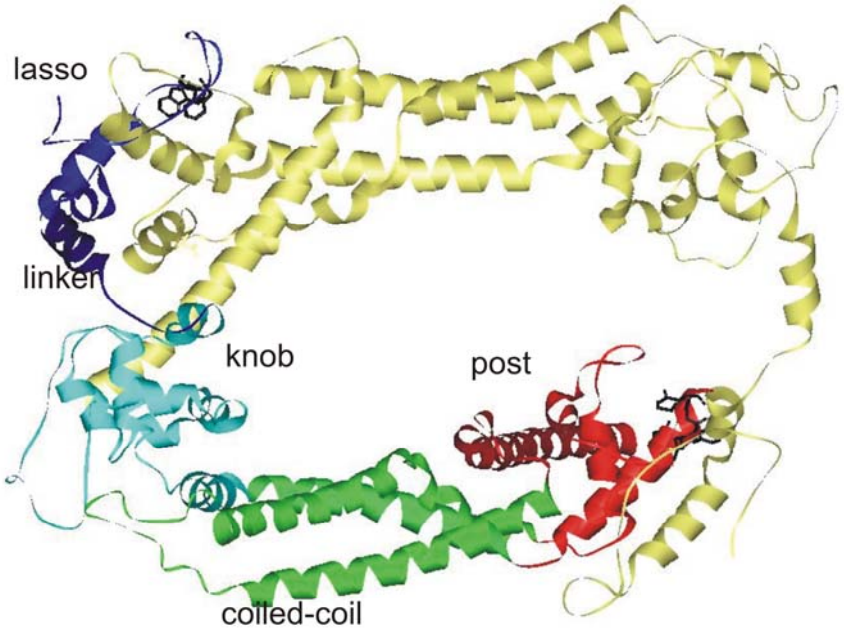


Figure 1.2.4. Ribbon diagram showing the three-dimensional structure of the Bni1FH2 domain dimer (adapted from pdb file 1UX5). One monomer is shown with colored subdomains (lasso in blue; linker in dark blue; knob in azure; coiled-coil in green; post region in red), the second monomer is colored tan. Residues involved in dimerization are shown as a stick models.

The structure of this domain can be subdivided into five subdomains. At the N-terminal region a so-called lasso domain is connected to a globular knob (helices αF to αI) by a 17-residue linker segment. The knob is followed by a three-helix bundle with a coiled-coil structure (helices αJ to αL). The C-terminal subdomain (αM to αT) forms the so-called post region. The lasso subdomain of one subunit encircles the post subdomain of the other subunit in a dimer. The N-terminal fragment of the lasso subdomain lacks the regular secondary structure, and two highly conserved tryptophan residues in this region insert into hydrophobic pockets in the post subdomain. The G-N-Y/F-M-N sequence motif that originally defined FH2 domain lies in the post region and all residues in this motif participate in dimerization. This motif is also highly conserved in other formins. Extensive contacts in the lasso-post interface are believed to be responsible for a very stable dimer. The crystal structure of a partial, monomeric FH2 domain from mDia1 has also been solved (Shimada et al., 2004). This construct lacks the lasso region but its structure is similar to the corresponding regions of Bni1. Neither the monomeric mDia1FH2 nor truncated mutants of Bni1FH2 are able to accelerate actin polymerization, suggesting that only the full length dimeric FH2 remains active.

Most formins contain a proline-rich FH1 domain adjacent to the N-terminal part of the FH2 domain. FH1 domains are highly variable in length, proline content, and number of potential profilin binding sites (0-16). The high proline content (35-100%) suggests that FH1 domains are most probably unstructured.

1.2.4 Formin-mediated actin assembly

Actin filaments are highly dynamic structures formed in a multi-step process, involving a slow and thermodynamically unfavorable nucleation phase, followed by an elongation phase. Cells have developed specialized machineries to accelerate the process of actin nucleation. Interactions of formins with actin are mediated through the FH2 domain, which alters actin polymerization dynamics in the following ways:

- accelerate the de novo filament nucleation
- alter a filament elongation/depolymerization rate
- prevent the filament barbed-end capping by capping proteins and gelsolin

The FH2 domain nucleates new actin filaments, most likely through stabilizing an actin dimer, and remains tightly bound with low nanomolar affinity to the barbed ends of the filaments (Mosley et al., 2004). The FH2 domain is able to move processively with an elongating actin filament barbed end and does not have to dissociate and reassociate when new actin monomer is added. There are several models of the FH2 domain processivity. In these models each subunit of the FH2 dimer binds one actin subunit at the barbed end, and the FH2 dimer 'stair-steps' with the elongating filament. Some models predict that only one FH2 subunit is bound at a time, with the other subunit free to accept a new monomer. Another model suggests that both FH2 subunits bind simultaneously to the two barbed end subunits and addition of an actin monomer to the filament causes one FH2 subunit release its previous actin and rebind to the newly added actin (Moseley et al., 2004; Harris et al., 2004; Higgs, 2005). The next model based on the crystal structure of Bni1FH2 suggests that interior residues of the FH2 dimer interact with the barbed end, and that the anti-cooperative binding of FH2 subunits to the actin subunits enables processive movement (Xu et al., 2004). The recently solved crystal structure of Bni1FH2 complexed with tetramethyl-rhodamine labeled actin (TMR-actin) shows that flexibility of the FH2 dimer permits the actin filament to sit inside the central hole of the FH2 dimer, allowing both halves of the FH2 dimer to interact with actin subunits at the end of the filament. Each half of FH2 makes two contacts with actin subunits and only one of them needs to be released to allow addition of a new monomer. In the proposed model, the FH2 dimer exists in an equilibrium between a closed state, when all four binding sites are engaged, and an open state, in which three actin binding sites are occupied and one is exposed in solution (Figure 1.2.5) (Otomo et al., 2005). Transition from closed to open states involves movement of the lagging unit of the FH2 dimer towards the barbed end, causing the two subunits

of the FH2 dimer to exchange roles. This mechanism would imply a considerable rotational movement of the formin around the helical actin filament but there is no evidence of supercoiling or rotation, suggesting that the FH2 dimer slips on the barbed end during elongation.

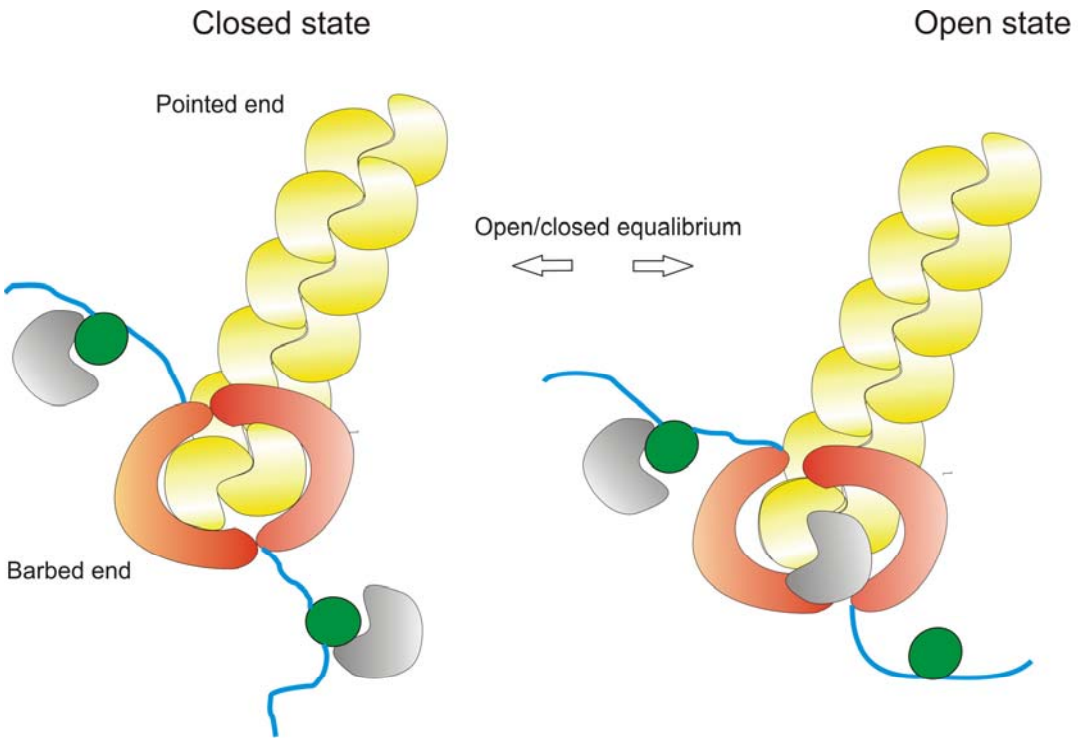


Figure 1.2.5. Formin mediated actin elongation. Continual attachment of formin with the elongating barbed end of actin filament is dependent on the FH2 dimer, which encircles the filament. The FH2 dimer is in rapid equilibrium between 'closed' and 'open' states. Although profilin and profilin-actin can bind to the proline-rich FH1 domain in both states, actin or profilin-actin can be added to the barbed end only when the FH2 dimer is in the 'open' state. Proteins depicted are: F-actin (yellow), G-actin (gray), the FH2 domain of formin (red), the FH1domain of formin (blue), profilin (green).

1.2.5 Cellular and organismal roles of formins

Formins play important role in many cellular events as well as in physiological and pathophysiological processes in organisms. Present understanding of molecular functions of formins has grown over the past few

years derived from detailed structural insights to complex and diverse cellular roles. Formins are required for cytokinesis (Glotzer, 2005; Tominaga et al., 2000; Evangelista et al., 2002), filopodium formation (Schirenbeck et al. 2005; Faix and Rottner, 2006), cell adhesion and motility (Wantanabe et al., 1999; Kobielak et al., 2004), endocitosis (Gasman et al., 2003; Fernandez-Borja et al., 2005), cell polarity (Evangelista et al., 1997; Sagot et al., 2002; Moseley and Goode, 2005), morphogenesis (Habas et al., 2001; Grosshans et al., 2005; Aspenstroen et al., 2006), microtubule stabilization (Gundersen et al., 2004; Wen et al., 2004), serum response factor activity (Miralles et al., 2003; Sun et al., 2006). In the future, formins may represent a family of attractive drug targets and may provide novel possibilities for the treatment of actin-dependent processes such as inflammation, metastasis and invasive diseases.

2 Goals of the study

The main goals of this thesis were to determine molecular bases of biological actions of two groups of proteins: the insulin-like growth factor-binding proteins and formins.

The IGF system is essential for normal embryonic and postnatal growth, and plays an important role in the function of the immune system, lymphopoiesis, myogenesis, and other physiological functions. Disregulation of the IGF system leads, for example, to growth and stimulation of cancer cells. In order to manipulate the IGF system in the treatment of diseases of IGF disregulation, protein-protein interactions of the components of the IGF system at the molecular level must be understood. The three-dimensional structures of the IGFBPs/IGF complexes presented in this thesis provide the molecular basis for our understanding of the structure-function relationships among the elements of the IGF system.

The formin proteins are regulators of actin filament assembly and polymerization dynamics. Analysis of three-dimensional structures of various formins, or formins in complexes with actin and profilins, would shed light on the molecular details of the action of this diverse family of proteins in a large number of actin-dependent processes.

3 Materials and laboratory methods

3.1 Materials

3.1.1 *E. coli* strains and plasmids

Cloning strains

XL1-Blue	Stratagene (USA)
One Shot TOP10	Invitrogen (Holland)
DH5α	Novagen (Canada)
GigaSingles	Novagen (Canada)

Protein expression strains

BL21 Star	Invitrogen (Holland)
BL21 Star(DE3)	Invitrogen (Holland)
BL21 Star(DE3) pLysS	Invitrogen (Holland)
Rosetta(DE3)	Novagen (Canada)
RosettaBlue(DE3)	Novagen (Canada)

Plasmids

pET 28a(+)	Novagen (Canada)
pET 22a(+)	Novagen (Canada)
pET 41 Ek/LIC	Novagen (Canada)
pET 46 Ek/LIC	Novagen (Canada)
pGEX 4T-2	Amersham Pharmacia (Sweden)
pGEX 6P-1	Amersham Pharmacia (Sweden)
pQE 30	Qiagen (Germany)
pQE 80	Qiagen (Germany)

3.1.2 Cell growth media and stocks

Media

For 1 liter LB medium:
10 g bacto tryptone
5 g bacto yeast extract
10 g sodium chloride

pH was adjusted to 7.0. For the preparation of agar plates the medium was supplemented with 15 g agar.

For 1 liter TB medium
12 g bacto tryptone
24 g bacto yeast extract
10 g sodium chloride
4 ml glycerol
900 ml deionized water

The medium was autoclaved, cooled; 100 ml sterile K phosphate and glucose were added. The final concentration of glucose was 0.5%

For 1 liter K-phosphate, pH 7.1:
23.1 g $KH_2 PO_4$
125.4 g K_2HPO_4

Minimal medium (MM) for uniform enrichment with ^{15}N:

For 1 liter MM:
0.5 g NaCl
1.3 ml trace elements solution
1 g citric acid monohydrate
36 mg ferrous citrate
4.02 g KH_2PO_4
7.82 g K_2HPO_4 x $3H_2O$
1 ml Zn-EDTA solution
1 g NH_4Cl or $^{15}NH_4Cl$

pH was adjusted to 7.0 with NaOH, the mixture was autoclaved, upon cooling separately sterilized solutions were added: 25 ml glucose, 560 µl thiamin, antibiotics, 2 ml $MgSO_4$ stock.

Defined medium for selective ^{15}N labeling of proteins

For 1 litre of medium:

400 mg Ala, Gln, Glu, Arg, Gly
255 mg Asp, Met
125 mg cytosine, guanosine, uracil
100 mg Asn, Leu, His, Lys, Pro, Thr
100 mg Try
400 mg Ile, Val
50 mg Phe, thymine, thymidine
1.6 g Ser
10 mg $CaCl_2$
2 g NaAc
10 g K_2HPO_4
1 g citric acid
1.3 ml trace element solution
36 mg ferrous citrate
1 ml Zn-EDTA solution
1g NH_4Cl

pH was adjusted to 7.0 with NaOH, the mixture was autoclaved. To the cooled medium, separately sterilized solutions were added: 25 ml glucose, 560 μl thiamin, antibiotics, 2 ml 1 M $MgSO_4$, sterile filtered:

50 mg Cys, Trp, nicotinic acid
0.1 mg biotin
X mg ^{15}N-amino acid

Another portion of the ^{15}N-amino acid is added at the time of induction as well (same amount as added before, 0.22 μm filtered).

Stock solutions

Ampicillin: 100 mg/ml of ampicillin in deionised H_2O, sterilized by filtration, stored in aliquots at -20°C until used. Working concentration: 150 μg/ml.

Chloramphenicol was dissolved in ethanol (0.34 g/10 ml) to the end concentration of 34 mg/ml. Working concentration: 34 μg/ml.

Kanamycin: 100 mg/ml of kanamycin in deionised H_2O, sterile filtrated and stored in aliquots at -20°C until used. Working concentration: 100 μg/ml.
IPTG: A sterile filtered 1 M stock of IPTG in distilled water was prepared and stored in aliquots at -20°C until used.
Glucose: 20% (w/v) in deionised H_2O, autoclaved.
Thiamin, 1%, in deionised H_2O, sterilized by filtration.
$MgSO_4$, 1 M, in deionised H_2O, sterilized by filtration.

Zn-EDTA solution:	5 mg/ml EDTA
	8.4 mg/ml $Zn(Ac)_2$
Trace elements solution:	2.5 g/l H_3BO_3
	2.0 g/l $CoCl_2$ x H_2O
	1.13 g/l $CuCl_2$ x H_2O
	9.8 g/l $MnCl_2$ x $2H_2O$
	2.0 g/l Na_2MoO_4 x $2H_2O$
	pH lowered with citric acid or HCl.

3.1.3 Solutions for making chemically competent *E. coli* cells

Solutions for making competent cells

Buffer A	100 mM $MgCl_2$ x $6H_2O$
Buffer B	100 mM $CaCl_2$ –15% glycerol

3.1.4 Protein purification – buffers

Ion exchange and gel filtration chromatography buffers

Buffer P(0)	8 mM KH_2PO_4
	16 mM Na_2HPO_4
	0.05% NaN_3
	pH 7.2

Buffer P(1000)	8 mM KH_2PO_4 16 mM Na_2HPO_4 1 M NaCl 0.05% NaN_3 pH 7.2
PBS	140 mM NaCl 2.7 mM KCl 10 mM Na_2HPO_4 1.8 mM KH_2PO_4 0.05% NaN_3 pH 7.3

Affinity chromatography buffers - Sepharose glutathione

Binding buffer	PBS
Wash buffer	PBS
Elution buffer	50 mM Tris-HCl 150 mM NaCl 30 mM glutathione pH 8.0

Buffers for immobilized metal-chelate chromatography (IMAC) under native conditions

Binding buffer	50 mM NaH_2PO_4 300 mM NaCl 10 mM imidazole pH 8.0
Wash buffer	50 mM NaH_2PO_4 300 mM NaCl

	20 mM imidazole pH 8.0
Elution buffer	50 mM NaH_2PO_4 300 mM NaCl 250 mM imidazole pH 8.0

Buffers for IMAC under denaturing conditions

Buffer A (binding buffer)	6 M guanidinium chloride 100 mM NaH_2PO_4 x H_2O 10 mM Tris 10 mM β-mercaptoethanol pH 8.0
Buffer B (wash buffer)	6 M guanidinium chloride 100 mM NaH_2PO_4 x H_2O 10 mM Tris 10 mM β-mercaptoethanol pH 6.5
Buffer C (elution buffer)	6 M guanidinium chloride 100 mM NaAc x $3H_2O$ 10 mM β-mercaptoethanol pH 4.0
Buffer D (dialysis buffer)	6 M guanidinium chloride pH 3.0
Buffer E (refolding buffer)	200 mM arginine HCl

	1 mM EDTA
	100 mM Tris
	2 mM red GSH
	2 mM ox GSH
	10% (v/v) glycerol
	0.05% NaN_3
	pH 8.4
Protease buffers	
Buffer X (Factor Xa cleavage buffer)	50 mM Tris
	100 mM NaCl
	4 mM $CaCl_2$
	0.05% NaN_3
	pH 8.0
Buffer T (thrombin cleavage buffer)	50 mM Tris
	60 mM NaCl
	60 mM KCl
	2.5 mM $CaCl_2$
	0.05% NaN_3
	pH 8.0
Buffer EK (Enterokinase cleavage buffer)	20 mM Tris
	100 mM NaCl
	2 mM $CaCl_2$
	0.01% NaN_3
	pH 7.5
Buffer PP	50 mM Tris
(PreScission Protease cleavage buffer)	150 mM NaCl
	1 mM EDTA
	1 mM DTT

pH 7.0

3.1.5 Buffer for DNA agarose gel electrophoresis

50X TAE buffer (for 1 l)

40 mM Tris-acetate	242 g of Tris base
1 mM EDTA	100 ml of 0.5 M EDTA (pH 8.0)
Glacial acetic acid	57.1 ml

3.1.6 Reagents and buffers for the SDS-PAGE

Anode buffer (+):	200 mM Tris pH 8.9
Cathode buffer (-):	100 mM Tris pH 8.25 100 mM tricine 0.1% SDS
Separation buffer:	1 M Tris pH 8.8 0.3% SDS
Stacking buffer:	1 M Tris pH 6.8 0.3% SDS
Separation acrylamide:	48% acrylamide 1.5% bis-acrylamide
Stacking acrylamide:	30% acrylamide 0.8% bis-acrylamide

Pouring polyacrylamide gels

Separation gel:	1.675 ml H_2O

2.5 ml separation buffer
2.5 ml separation acrylamide
0.8 ml glycerol
25 µl APS
2.5 µl TEMED

Intermediate gel:
1.725 ml H_2O
1.25 ml separation buffer
0.75 ml separation acrylamide
12.5 µl APS
1.25 µl TEMED

Stacking gel:
2.575 ml H_2O
0.475 ml stacking buffer
0.625 ml stacking acrylamide
12.5 µl 0.5 M EDTA, pH 8.0
37.5 µl APS
1.9 µl TEMED

Protein visualization

Coomassie-blue solution:
45% ethanol
10% acetic acid

Destaining solution:
5% ethanol
10% acetic acid

3.1.7 Reagents and buffers for western blots

Transfer buffer
25 mM Tris
192 mM glycine
pH 8.3

To make the final working solution mix 80 ml of the transfer buffer with 20 ml of methanol

Alkaline phosphatase buffer	100 mM Tris 100 mM NaCl 5 mM $MgCl_2$ pH 9.5
Wash buffer	10 mM Tris 150 mM NaCl 0.05% Tween20 pH 8.0
1 st antibody solution	1:2000 diluted in the wash buffer
2 nd antibody solution (linked to alkaline phosphatase)	1:2000 diluted in alkaline phosphatase buffer
Substrate for alkaline phosphatase	BCIP (Sigma); dissolve 1 tablet in 10 ml of water

3.1.8 Enzymes and other proteins

BSA	New England BioLabs (USA)
CIP	New England BioLabs (USA)
BamH I	New England BioLabs (USA)
EcoR I	New England BioLabs (USA)
Hind III	New England BioLabs (USA)
Nde I	New England BioLabs (USA)
Sal I	New England BioLabs (USA)

Xho I	New England BioLabs (USA)
Pfu turbo DNA Polymerase	Stratagene (USA)
Pfu DNA Polymerase	Fermentas (Lithuania)
Phusion HF DNA Polymerase	BioCat (Germany)
Vent^{R} polymerase	New England BioLabs (USA)
T4 DNA Ligase	New England BioLabs (USA)
Dpn I	Stratagene (USA)
Xa Factor	Novagen (Canada)
Enterokinase	Novagen (Canada)
PreScissionProtease	Amersham Biosciences(Sweden)
TAGzyme	Qiagen (Germany)
Thrombin	Sigma (USA)
hIGF-I receptor grade	GroPep (Australia)
Anti His antibodies (mouse)	Santa Cruz biotech (USA)
Goat anti mouse antibodies	Santa Cruz biotech (USA)

3.1.9 Kits and reagents

Advantage 2 PCR kit	BD biosciences (USA)
QIAquick PCR Purification Kit	Qiagen (Germany)
QIAprep Spin Miniprep Kit	Qiagen (Germany)
QIAGEN Plasmid Midi Kit	Qiagen (Germany)
QIAGEN Plasmid Maxi Kit	Qiagen (Germany)
QuikChange Site-Directed Mutagenesis Kit	Stratagene (USA)
Pre-Crystallization Test (PCT)	Hampton Research (USA)
Rapid Ligation Kit	Roche (Germany)
Complete Protease Inhibitor Cocktail	Roche (Germany)
pET LIC cloning Kits	Novagen (Canada)

3.1.10 Protein and nucleic acids markers

Prestained Protein Marker	New England BioLabs (USA)
100 BP DNA marker	New England BioLabs (USA)
1Kb DNA marker	New England BioLabs (USA)
Broad Range (6-175 kDa) 1 kb DNA-Leiter	Peqlab (Germany)

3.1.11 Chromatography equipment, columns and media

ÄKTA explorer 10	Amersham Pharmacia (Sweden)
Peristaltic pump P-1	Amersham Pharmacia (Sweden)
Fraction collector RediFrac	Amersham Pharmacia (Sweden)
Recorder REC-1	Amersham Pharmacia (Sweden)
UV flow through detector UV-1	Amersham Pharmacia (Sweden)
BioloLogic LP System	Biorad (USA)
HiLoad 26/60 Superdex S75pg	Amersham Pharmacia (Sweden)
HiLoad 16/60 Superdex S75pg	Amersham Pharmacia (Sweden)
HiLoad 16/60 Superdex S200pg	Amersham Pharmacia (Sweden)
HiLoad 10/30 Superdex S75pg	Amersham Pharmacia (Sweden)
HiLoad 10/30 Superdex S200pg	Amersham Pharmacia (Sweden
Mono Q HR 5/5, 10/10	Amersham Pharmacia (Sweden)
Mono S HR 5/5, 10/10	Amersham Pharmacia (Sweden)
NiNTA-agarose	Qiagen (Germany)
GST Sepharose FF	Amersham Pharmacia (Sweden)

3.2 Laboratory methods and principles

3.2.1 Construct design and choice of the expressions system

Optimization of protein constructs is essential for X-ray crystallography and NMR studies. Unstructured and flexible fragments of proteins or loop regions usually inhibit crystallization or result in crystals of low quality. The idea behind designing the protein constructs is to have well defined, folded and stable

domains, and at the same time to have them biologically active. Determination of stable and folded constructs require employing various techniques like for example: limited proteolysis, protein sequencing, mass spectrometry, and NMR spectroscopy (Rehm et al., 2004). In some cases designing of the constructs may be based on previously published literature and secondary structure prediction with the help of bioinformatics tools.

Successful expression of heterologous proteins requires proper expression system (Makrides, 1996). In this work pET or pGEX series of vectors were used for the expression of proteins in *E. coli* (pET system manual, Novagen, 2003; Amersham Pharmacia, 2003). Large fusion tags, such as glutathione S-transferase, thioredoxin or maltose binding protein often increase solubility and promote proper folding of recombinant fusion partners. For proteins directed into inclusion bodies, presence of a 6-histidine fusion peptide (His-Tag) makes use of immobilized metal chromatography (IMAC) possible under strong denaturing/reducing conditions, enabling rapid purification.

Removal of the fusion tag is often required for functional studies of a recombinant protein and for crystallization. This is usually achieved with an aid of a specific restriction protease. The enzymes most commonly used are: 1) factor Xa - cleaves after R in an IEGRX sequence; 2) thrombin - cleaves after R in a LVPRGS sequence; 3) enterokinase - cleaves after K in a DDDDK sequence; 4) PreScission Protease - cleaves after Q in a LEVLFQGP sequence; 5) exopeptidases (dipeptidase) - cleave dipeptides from the N-terminus of a fusion protein, and stops before Arg, Lys or 1 or 2 residues before Pro. The enzyme is active in the presence of PMSF and EDTA, making it ideal for work with degradation susceptible proteins. However, this enzyme is used for removal of only short N-terminal His-tags.

Vectors with specific protease sites are commercially available otherwise they could be inserted using PCR. All constructs of IGFBPs and a few constructs of formins were designed as a His-tag fusion protein. Most of formin and profilin constructs were expressed as a GST-fusion protein with enterokinase, PreScission protease or thrombin cleavage sites.

3.2.2 DNA techniques

3.2.2.1 Preparation of plasmid DNA

The isolation of plasmid DNA from *E. coli* was carried out using dedicated plasmid purification kits from Qiagen. The kits employ a standard alkalic lysis of the precipitated bacteria in the presence of RNAse and a strong ionic detergent, SDS, followed by neutralization/DNA renaturation with acetate. For purification, a crude cell lysate is loaded onto a silica gel column, washed with an ethanol-containing buffer, and eluted in a small volume, yielding up to 20 μg of the plasmid DNA.

3.2.2.2 PCR

A polymerase chain reaction was employed to amplify desired DNA fragments and genes, introduce restriction sites, STOP codons and sequences encoding restriction protease cleavage sites. The primers were prepared according to standardized principles regarding the length, GC-content, melting temperature and occurrence of secondary structures of the hairpin type. All primers used for cloning and mutagenesis are listed in Table 3.1. Three different kinds of recombinant thermostable DNA polymerases were used, each operating at slightly different conditions:

	Melting temp.	Annealing temp.	Synthesis temp.
Phusion HF	98°C	55°C	72°C
Pfu Turbo	95°C	55°C	68, 72°C
Vent^{R}	95°C	55°C	72°C

The stock solution of the primer was always 0.1 nM. The working solution was 0.01 nM. Usually 2 µl of the working solution was used per PCR reaction for each primer

Table 3.1. Primers used in this work

No	Name	Nucleotide sequence
1	NBP4D1Xa**F**	CGC**GGATCC**ATTGAGGGTCGCGACGAAGCCATCCACTGCCCGCCC
2	NBP4A3Xa**F**	CGC**GGATCC**ATTGAGGGTCGCGCCATCCACTGCCCGCCC
3	NBP4L82**R**	CCC**AAGCTT***TCATTA*CAGCTCCATGCACACGCCTTGCCCG
4	NBP4L92**R**	CCC**CTCGAG***TCATTA*CAGGCTTTCCTGGATGGCCTCGATCTC
5	CBP1XaF	CGC**GAATTC**ATCGAAGGTCGTGTCACCAACATCAAAAAATGGAAGG
6	CBP1ThF	CGC**GAATTC**CTGGTTCCGCGTGGATCCGTCACCACCATCAAAAAATGGAAGGAGCCC
7	CBP1R	CGC**CTCCGAG***TTA*GTTTTGTACATTAAAATATATCTGGC
8	CBP4XaF	CGC**GGATCC**ATTGAGGGTCGCTGCCAGAGCGAGCTGCACCGGG
9	CBP4R	CCG**CTCGAG***TCATTA*CAGCTCCATGCACACGCCTTGCCCG
10	CBP4inWKF	CGC**GGATCC**CTGGTTCCGCGTGGATCCCAGTGGAAGGGCTCCTGCCAGAGCGA GCTGCACCGGGCG
11	CBP4N189FF	CGCAACGGCTTCTTCCACCCCAAGCAGTGTCACCC
12	CBP4N189FR	GGGGTGGAAGAAGCCGTTGCGGTCGCAGTTGGGG
13	CBP4K211WF	GTGGACCGGTGGACGGGGGTGAAGCTTCCGGGGGGC
14	CBP4K211WR	CACCCCCGTCCACCGGACCACACACCAGCACTTGCC
15	DIAPH1553F	CGC**GGATCC**ATGGCTTCCCTCTCTGCGGCAGCTATTACT
16	DIAPH1614F	CGC**GGATCC**GGAGGTACTGCTATCTCTCCACCCCCTCCT
17	DIAPH11134R	CCG**CTCGAG***TTATCA*TGTCTTCCGCCGCTTCTGGTTCTCCTT
18	DIAPH11200R	CCG**CTCGAG***TTATCA*CTTCCTGTTGGCTTGACGGGGCCCG
19	DIAPH1748R	CCG**CTCGAGAAGCTT***TTA*GGGGGTTAATCCAAATGGCAGAACTGGGGC
20	DAAM1594F	CCG**GGATCC**GATGACGACGACAAGATGGGCCTAGCACTGAAGAAGAAAAGCATTCC
21	DAAM1527FLic	GACGACGACAAGATCCCAGGTGGACCCTCGCCTGGAGCACCAGG
22	DAAM1542F	CCG**GGATCC**GATGACGACGACAAGTCCTCTGTGCCTGGATCTCTCCTTCCT CCC
23	DAAM11030R	CCG**GTCGAC***TTA*ACTATTCTCTTTAGCTTTTCTCATTTTACG
24	DAAM11059RLic	GAGGAGAAGCCCGGT*TTA*TCTCTCTCTGCTGCTGTCAGTCATCTGGTTGG
25	DAAM11078R	CCG**GTCGAC***TTA*GAAATTAAGTTTTGTGATTGGTCTCTC
26	DAAM2588F	CCG**GGATCC**AGGAAAAAGCGTGTCCCCCAGCCTTCTCACCC
27	DAAM21020R	CCG**GTCGAC***TTA* GCAGCCAGGACCTTCCGCTGCCGCTG

28	dDia2585F	CGC**GGATCC**ACTGAACCAATTTTAGGT
29	dDia2602F	CCA**GGATCC**GGAGGAGGAGGACCACCACCACCA
30	dDia2616F	CCG**GGATCC**GGTGGAAAGAGTAATAAACCTGCTAAACC
31	dDia2619F	CCG**GGATCC**AGTAATAAACCTGCTAAACCAATTATTAAACC
32	dDia2636F	CGC**GGATCC**TTCATTTGGATTACAATTCCAGCACTTAAA
33	dDia21053R	CCG**GTCGAC***TTA*ATTTTCTTAAACTATCAACATTTTGCATTTGAGCAGC
34	dDia21004R	CCG**GTCGAC**TTAAGCCTTTCTAATCATAGCTTGATATTCACCG
35	dDia2670R	CCG**GTCGAC***TTA*ACTCTCTAATCCCACCTTATCC
36	dDia2745R	CCC**GTCGAC***TTA*ATCCTCTTTGGTTGGTGCAAATTGTAAGA
37	dDia2 F R688G K689N	CAATTAACAGGAAATGTGGTTGTTACAGTTATCGAT
38	dDia2R R688G K689N	CAACCACATTTCCTGTTAATTGTTTACTTTCTACC
39	dDia2F K618G K621A K624N	CCTGGTGGAGGGAGTAATGCACCTGCTAATCCAATTATTAAACCATCAG
40	dDia2R K618G K621A K624N	TGGATTAGCAGGTGCATTACTCCCTCCACCAGGTGGAGGTGG
41	dDia2 F K632AR634G	CCATCAGTTGCGATGGGAAATTTCAATTGGATTAC
42	dDia2R K632AR634G	GAAATTTCCCATCGCAACTGATGGTTTAATAATTGG
43	dDia2 F L644KQ647K	CTTTAACTTTTTTTGCTGGAATTGTAATCCAATTGAAATTTCTC
44	dDia2 R L644KQ647K	CCAGCAAAAAAAGTTAAAGGTACATTTTGGGATAAATTGGATG
45	dDia2 F W638K	AATTTCAATAAAATTACAATTCCAGCAAAAAAAGTGAAAGG
46	dDia2 R W638K	AATTGTAATTTTATTGAAATTTCTCATCTTAACTGATGG
47	dProf2fusF	CCC**GGATCC**ATGACTTGGCAAGCATACGTCGATAACAAC
48	dProf2fusR	GGC**GAATTC**CGGCACAGTTGTTGTCAATTAAATAATCGGC
49	dDia2FH2fusF	CCG**GAATTC**TTAGTAATAAACCTGCTAAACCAATTATTAAACC
50	hProfilin1F	CGC**GGATCC**ATGGCCGGGTGGAACGCCTACATCGACAACCTCATG
51	hProfilin1R	CCG**CTCGAG***TTA*GTACTGGGAACGCCGAAGGTGGGAGGCCATTTC
52	hProfilin2F	CGC**GGATCC**ATGGCCGGTTGGCAGAGCTACGTGGATAACCTGATGTGC
53	hProfilin2R	CCG**CTCGAG***TTA*GAACCCAGAGTCTCTCAAGTATTTTGCCATTGAGTATGCC

F – forward; R – reverse; restriction sites in bold; stop codons in italics; mutations are coloured red; encoded protease cleavage sites are underlined

3.2.2.3 Digestion with restriction enzymes

Usually, 1-2 units of each restriction enzyme were used per 1 μg of plasmid DNA to be digested. The digestion was performed in a buffer specified

by the manufacturer at the optimal temperature (37°C) for 5-16 h. The fragments ends that occurred after digestion were cohesive. To eliminate possibility of plasmid recirculation (possible when double-digestion does not occur with 100% efficiency), 5'-ends of a vector were dephosphorylated using calf intestine phosphatase (CIP). CIP treatment was performed with 1 unit of enzyme per 3 μg of plasmid DNA, at 37 °C for 1 h.

3.2.2.4 Purification of PCR and restriction digestion products

DNA obtained from restriction digestion, phosphatase treatment or PCR was purified from primers, nucleotides, enzymes, buffering substances, mineral oil, salts, agarose, ethidium bromide, and other impurities, using a silica-gel column (QIAquick PCR Purification Kit, Qiagen). The QIAquick system uses a simple bind-wash-elute procedure. A binding buffer was added directly to the PCR sample or other enzymatic reaction, and the mixture was applied to the spin column. Nucleic acids adsorbed to the silica-gel membrane in the high-salt conditions provided by the buffer. Impurities and short fragments of single or double-stranded DNAs were washed away and pure DNA was eluted with a small volume of 10 mM Tris pH 8.0 or water.

3.2.2.5 Ligation

The ligation of digested and purified inserts and vectors was performed according to the protocol described in the T4 DNA Ligase instruction (New England BioLabs).

The ligation mixture contained (20 μl):

insert	16 μl (0.15 μM)
10 x T4 DNA ligase reaction buffer	2 μl
vector	1 μl
T4 DNA Ligase	1 μl

3.2.2.6 Ligation independent cloning

The ligation independent cloning requires the PCR amplification of the gene of interest with the compatible overhang at both ends (LIC cloning kit). Following are the sequences of the overhangs:

For pET41 and 46 LIC/Ek
Forward primer 5' – GACGACGACAAGAT – 3'
Reverse primer 5' – GAGGAGAAGCCCGGT – 3'

Rest of the procedure was followed as per the instructions of the manufacturer (pET system:manual, Novagen, 2003).

3.2.2.7 Mutagenesis

Site directed mutagenesis of CBP-4 and dDia2 were performed with PCR, using enzymes and instructions supplied in the QuikChange Site-Directed Mutagenesis Kit (Stratagene). The mutagenic oligonucleotide primers were designed according to suggestions provided by the manufacturer. The desired mutation was in the middle of the primer with ~ 10-15 bases of a correct sequence on both sides (Table 3.1, primers:). Vectors pET 28a and pGEX 4T-2 containing copies of a gene encoding CBP-4 and dDia2, respectively, were used as DNA templates. High concentration of the template DNA and low number of PCR cycles, combined with high accuracy and fidelity of highly processive DNA polymerase *Pfu* Turbo, minimizes the occurrence of unwanted mutations.

The mutagenic PCR reaction mixture contained:

10 × reaction buffer	5 μl
dNTP mix	1 μl
plasmid (5 ng/μl)	1 μl (10 ng)
oligonucleotide primer F	2 μl (125 ng)
oligonucleotide control primerF	2 μl (125 ng)

PfuTurbo DNA polymerase (2.5 U/μl)	1 μl
milli-Q to a final volume of 50 μl	38 μl

PCR cycling parameters:	
denaturation:	95°C, 1’
denaturation:	95°C, 30’’
annealing:	55°C, 1’
synthesis (1 min per 1000 base pairs):	68°C, 6’

Following the temperature cycling, the product was treated with Dpn I (10 U, 37°C, for 2 h). The Dpn I endonuclease (target sequence: 5´-Gm^{6}ATC-3´) is specific for methylated and hemimethylated DNA and is used to digest the parental DNA template and to select for mutation-containing synthesized DNA. 2 μl of the mixture were used to transform XL1-Blue or Top10 chemically competent cells. Plasmid DNA was isolated using QIAprep Spin Miniprep Kit (Qiagen) and was subjected to verification by automated sequencing.

3.2.2.8 Agarose gel electrophoresis of DNA

For verification of the presence and length of PCR or restriction digestion products, agarose gel electrophoresis was performed. For this purpose 1% agarose in a TAE buffer supplemented with ethidium bromide was prepared. The DNA samples were mixed with the 6x sample buffer prior to loading. DNA samples were run along with the 100bp and/or 1kb DNA ladder (NEB or pEQ lab) at 100-120 V DC. Results were vizualized using UV illumination.

3.2.3 Transformation of *E. coli*

3.2.3.1 Making chemically competent cells

A single colony of overnight grown bacteria from a LB agar plate was inoculated into 100 ml of LB media in a 500 ml flask. Culture was incubated at

37°C with vigorous agitation, monitoring the growth of cells. Cells were grown till the OD_{600} reach ~0.6. The bacterial culture was transferred to sterile, disposable, ice-cold 50 ml polypropylene tubes and cooled down to 4°C on ice for 10 min. Cells were recovered by centrifugation at 3000 g for 10 min at 4°C. Supernatant media was decanted and tubes were kept in an inverted position on a pad of paper towel for 1 min to allow the last traces of media to drain away. Pellets were resuspended by gentle vortexing in 30 ml of the ice-cold $MgCl_2$ solution. Again, cells were recovered by centrifugation at 3000 g for 10 min at 4°C. Supernatant solution was decanted and tubes were kept in an inverted position on a pad of a paper towel for 1 min to allow the last traces of solution to drain away. Pellet of the cells was recovered by gentle vortexing in 2 ml of ice-cold 0.1 M $CaCl_2$ containing 15% glycerol, for each 50 ml of original culture. After this cells were dispensed into aliquots of 50 µl, flash frozen in liquid nitrogen and stored at -70°C.

3.2.3.2 Transformation of chemically competent cells

3 µl of a ligation mix or ca. 50 µg of plasmid DNA was added to 50 µl of chemically competent cells. The mixture was incubated on ice for 30 min followed by a heat shock of 45 s at 42°C, 2 min cooling on ice, and the addition of a 250 µl of glucose and magnesium containing medium. After 1 h of incubation at 37°C, 20-50 µl of the mixture was spread out on LB agar plates (supplemented with selective antibiotic) and incubated overnight at 37°C. A number of factors have been elucidated that produced an increase in transformation efficiency. Such factors include: prolonged incubation of bacteria with $CaCl_2$, addition of multiple cations, such as Mg^{2+}or Cs^{2+} into the transformation mixture and treatment of bacteria with dimethyl sulfoxide (DMSO), polyethylene glycol, hexaminecobalt, and dithiothreitol in the presence of both monovalent and divalent cations (Chung et al., 1989). After incubation with DNA, in order to make the cells retain the plasmid and to be certain that they survive, the cells were heat shocked for several seconds to induce heat shock genes, which aid in

survival and recovery. The cells were than incubated at 37°C without selective pressure; sufficient time was given for expression of antibiotic resistance genes. Plating on selective media enabled recovery of those cells that actually received the DNA.

3.2.3.3 Transformation by electroporation

1 μl of an aqueous solution of plasmid DNA (ca. 100 ng/μl) was added to 50 μl of electrocompetent cells, the mixture was pipetted into a 2 mm electroporation cuvette. The electroporation was performed in an electroporation vessel (Gene pulser) at 1650 V. Then the suspension was transferred into an Eppendorf tube and mixed with 1 ml LB medium. After 1 h of incubation at 37°C, cells were plated as described above.

3.2.4 Protein chemistry methods & techniques

3.2.4.1 Protein expression

Optimization of the conditions is important for expressing a protein. The aim is to get the maximum amount of a protein in the soluble fraction of the lysate. Often the recombinant protein is expressed as insoluble inclusion bodies. The protein might be solubilized and then refolded back to the native form by employing various refolding strategies. In this work both methods of protein expression were applied to obtain soluble product. A number of parameters were checked to get the maximum yields of the protein, which include optimization of a type of culture media, temperature, induction duration, induction OD, concentration of inducer (IPTG), and cell type. All the proteins were cloned and expressed in pET and pGEX vectors, which are IPTG inducible. The general strategy of protein expression and purification is shown in Figure 3.1.

3.2.4.1.1 Expression and purification of IGFBPs

The overnight cultures of 10 ml were used as an inoculum for the 1 l culture. The cells were induced at OD_{600} = 0.8-1.0 by addition of IPTG to the final concentration of 1 mM and grown for the next 3 h with vigorous shaking at 37°C. The cells were harvested by centrifugation (4000 x G, 25 min, 4°C), bacterial pellets were resuspended in PBS and sonicated 5 x 3 min using a macrotip (output control 8, 70%). Lysate was centrifuged at 60000 x g for 30' at 4 °C. Pellet was solubilized overnight in buffer A. After centrifugation, the resulting supernatant was incubated with a Ni-NTA slurry (Qiagen) for 2 h at room temperature with gentle agitation. Next, the mixture was loaded onto an empty column and washed with buffer A and B. The protein was eluted with buffer C. The fractions containing the desirable protein were pooled, concentrated and dialysed against buffer D for removal of reducing agents. Subsequently, refolding of the protein was performed. For this purpose, the protein sample was stepwise diluted in buffer E in a 1:50 volume proportion. The refolding mixture was left with stirring at 4°C. After 3 days, the mixture was concentrated and dialysed against a low salt phosphate buffer, and then was loaded onto a cation exchange column (MonoS). Bound proteins were fractionated with a linear NaCl gradient. Fractions containing the protein of interest were pooled and dialysed against thrombin or Factor Xa cleavage buffers. Factor Xa digestion was performed for 4 days in 50 mM Tris buffer pH 8.0, 150 mM NaCl, 4 mM $CaCl_2$ at 4°C using 1 U of the activated protease (Novagen) per 1 mg of the fusion protein. Thrombin cleavage was carried out in the same conditions; 2 U of thrombin (Sigma) was used per 1 mg of protein. The proteins of interest were separated from the HisTag and protease by gel filtration on the Superdex 75 prep grade 16/60 XR column. The buffer used contained 140 mM NaCl, 2.7 mM KCl, 10 mM Na_2HPO_4, 1.8 mM KH_2PO_4, 0.05% NaN_3, pH 7.3. The purity of the protein was checked by mass spectrometry and NMR.

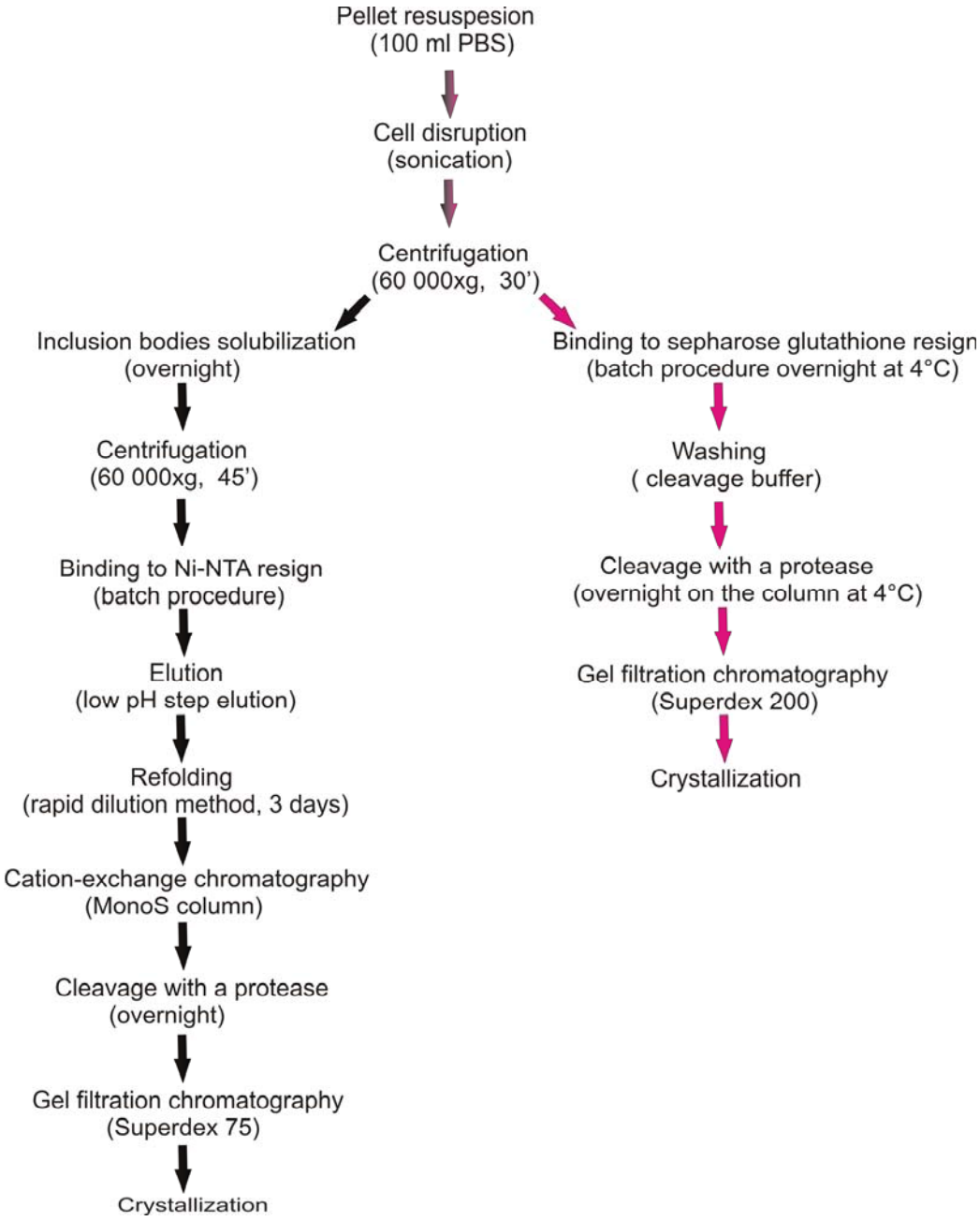


Figure 3.1. Flow chart of the purification scheme for all proteins studied in this work (denaturating conditions in black; native conditions in pink).

3.2.4.1.2 Expression and purification of formins and profilins

1 l LB media with appropriate antibiotic were inoculated with 10 ml of preculture. The cells were induced at OD_{600} = 0.8-1.0 by addition of IPTG to the

final concentration of 0.5 mM and grown for the next 20 h with vigorous shaking at 20°C. The cells were harvested by centrifugation (4000 x G, 25 min, 4°C), bacterial pellets were resuspended in PBS supplemented with protease inhibitors and sonicated 5 x 3 min using a macrotip (output control 8, 70%). Lysate was centrifuged at 60000 x g for 30' at 4°C. The supernatant was incubated with a sepharose glutathione slurry (AmershamBiosciences) for 3–4 h at 4°C with gentle agitation. After loading onto a column and washing with appropriate buffer the protein was eluted with 30 mM of the reduced glutathione, 150 mM NaCl, and 50 mM Tris, or cleaved on the column with thrombin, enterokinase or the PreScission Protease. The fractions containing protein of interest were pooled, concentrated and purified by gel filtration in the crystallization buffer.

3.2.4.2 Sonication

Pulsed mode of operation was applied (output control 8, 60% duty cycle) and sonication was carried out on ice, in 3 steps of 5 min each, with 5 min intervals between steps, to avoid overheating of the sample.

3.2.4.3 SDS polyacrylamide gel electrophoresis (SDS-PAGE)

The SDS polyacrylamide gel electrophoresis was performed at various stages of purification to check the purity and identity of the eluted proteins. For all expressed proteins, tricine gels were applied (Schagger and von Jagow, 1987). The protein samples were prepared by mixing 20 μl of protein solution with 5 μl of sample buffer (SB) followed by 5 min incubation at 100 °C. Due to rapid precipitation of SDS in contact with guanidine, the samples (after Ni-NTA chromatography under denaturing conditions) were prepared as follows: 20 μl of the protein solution in a denaturing buffer was diluted with 400 μl 20% trichloroacetic acid (TCA). The sample was incubated for 5 min on ice followed by centrifugation for 5 min at 20 000 x g. Supernatant was discarded by suction, the precipitated protein pellet was washed once by vortexing with 400 μl ethanol.

After centrifugation and ethanol removal, the protein pellet was resuspended in 20 μl of 2x SB and the sample was boiled for 5 min.

3.2.4.4 Visualization of separated proteins

For visualization of the protein bands, the gels were stained in a Coomassie-blue solution. Background was cleared by incubation of the gel in a destaining solution. Both processes were greatly accelerated by brief heating with microwaves of the gel submerged in an appropriate solution.

3.2.4.5 Western blot

Western blot is a functional assay to check the identity of proteins or identify the protein out of a number of proteins. The semi-dry Western blot was applied. The Western blot assay starts with running the desired sample on the SDS-PAGE. The nitrocellulose membrane and six Watmann paper of the size of the SDS-PAGE gel were cut and soaked in the transfer buffer. Watmann paper, SDS-PAGE (gel), and nitrocellulose membrane were arranged in the following order over the electroblot: three wet Watmann paper, wet nitrocellulose paper, and SDS-PAGE gel followed by three Watmann wet papers. The apparatus was closed and run (transfer) at constant voltage of 100 V for 1-1:30 h. After the transfer, the nitrocellulose membrane is taken and kept in the blocking solution for 2 h with constant shaking. The SDS-PAGE is stained by Commassie blue solution to check the success of transfer. After blocking, the membrane is washed three times with the wash buffer and incubated for 1.5 h at room temperature with the 1 st antibody solution (the procedure can be stopped at this point by keeping the blot in the 1 st antibody solution at 4°C for overnight). The membrane was washed with the wash buffer and incubated in the 2 nd antibody solution for 1.5 h at room temperature. After this the membrane was washed three times with the wash buffer and the blot was developed by incubating it in the substrate (BCIP) solution for 10 min.

3.2.4.6 Determination of protein concentration

The concentration of proteins in solution was estimated by means of the Bradford colorimetric assay. 5 µl of the protein sample was added to 1 ml (10 x diluted stock) of Bradford reagent (BioRad) in a plastic cuvette. After gentle mixing, A_{595} was measured and converted to the protein concentration on the basis of a calibration curve prepared for known concentrations of BSA.

Determination of protein concentration was performed spectrophotometrically. Absorption at 280 nm was measured and converted to a protein concentration on the basis of theoretical extinction coefficients. It has been shown that it is possible to estimate the molar extinction coefficient E_λ(Prot) of a protein from knowledge of its amino acid composition (Gill and Hippel, 1989). From the molar extinction coefficient of tyrosine, tryptophan and cystine (cysteine residues do not absorb appreciably at wavelenghts >260 nm, while cystine does) at a given wavelength λ the extinction coefficient of a protein can be computed using the equation:

$$E_\lambda(\text{Prot}) = \text{Numb(Y)x}\text{Ext}_\lambda\text{(Y)} + \text{Numb(W)x}\text{Ext}_\lambda\text{(W)} + \text{Numb(C)x}\text{Ext}_\lambda\text{(C)}$$

Protein concentration (C_p) can be calculated using the following formula:

$$A_\lambda\ (\text{Prot}) = E_\lambda(\text{Prot}) \times C_p(\text{Prot}) \times (\text{cuvette path length in cm})$$

3.2.5 NMR spectroscopy

All NMR experiments were carried out at 300 K on a Bruker DRX 600 spectrometer equipped with a triple resonance, triple gradient 5 mm probehead. The samples contained typically 0.1-0.5 mM protein in the PBS buffer supplemented with 10% ^2H_2O. All 1D ^1H NMR spectra were recorded with a time domain of 32 K complex points and a sweep-width of 10,000 Hz. The 2D ^1H-^{15}N-HSQC spectra were recorded with a time domain of 1K complex data points with

128 complex increments with a sweep width of 8 kHz in the ^{1}H dimension and 2 kHz in the ^{15}N dimension.

3.2.6 X-ray crystallography

3.2.6.1 Protein crystallization

The ternary complexes of NBP-4(1-92)/IGF-1/CBP-1(141-234), NBP-4/(3-82)/IGF-1/CBP-4(151-232) were prepared by mixing equimolar amounts of the components. The complexes were separated from any excess of free proteins by gel filtration chromatography on the Superdex 75 16/60XR column in buffer containing 5 mM Tris, 50 mM NaCl, pH 8.0. Screening for crystallization conditions was performed at 4ºC and 20ºC using Hampton Research Screens. Crystallization of the complexes was carried out with the sitting drop vapor diffusion method by mixing equal volumes (2 µl) of protein (10 mg/ml) and reservoir solution. The complexes have crystallized in several conditions.

The FH2 and FH1FH2 domains of dDia2 and DAAM 1 were purified by gel filtration chromatography on the Superdex 200 16/60XR column in buffer containing 5 mM Tris, 150 mM NaCl, pH 7.2. After initial screening seven crystallization conditions for FH2 DAAM1 and one for FH2 dDia2 were found and optimised using both sitting- and hanging-drop vapour diffusion techniques. The complex of dDia2 FH2 domain and nonpolymerizable actin mutant was prepared by mixing equimolar amounts of both proteins.

3.2.6.2 Data collection and structure analysis

The data for all crystals were collected from shock frozen crystals at a rotating anode laboratory source. Prior to freezing, the crystals were soaked for 30 s in a drop of a reservoir solution containing 15% v/v ethylene glycol or 15'% glycerol as cryoprotectant. The high-resolution datasets were collected on the MPG/GBF beamline BW6 at DESY, Hamburg, Germany. Collected data were integrated, scaled and merged by XDS and XSCALE programs (Kabsch, 1993).

The structure was determined by molecular replacement using the Molrep program from the CCP4 suite (CCP4, 1994).

3.2.7 Isothermal titration calorimetry

Isothermal titration calorimetry (ITC) is a thermodynamic technique for monitoring any chemical reaction initiated by the addition of a binding component, and has become the method of choice for characterizing biomolecular interactions. When substances bind, heat is either generated or absorbed. Measurement of this heat allows accurate determination of dissociation constants (K_D), reaction stoichiometry (n), enthalpy (ΔH) and entropy (ΔS), thereby providing a completethermodynamic profile of the molecular interaction in a single experiment. All ITC experiments were carried out according to references provided by the manufacturer. Proteins (N-terminal and C-terminal of IGFBPs and IGF-1) were used at 0.4 mM in PBS and titrated from a 300 µl syringe into a sample chamber holding 1.43 ml of 0.04 mM of a respective binding partner. All solutions were degassed prior to measurements. Heat generated by protein dilution was determined in separate experiments by injecting a protein solution into PBS filled sample chamber. All data were corrected for the heat of protein dilution. Data were fitted using χ^2 minimization for a model assuming a single set of sites to calculate the binding affinity K_D. All steps of the data analysis were performed using ORIGIN (V5.0) software provided by the manufacturer.

The details of the experimental and injection parameters are described below.

Parameters of the experiments*:*

total number of injections:	58
volume of a single injection [µl]:	5
duration of an injection [s]:	10
intervals between injections [s]:	400

filter period [s]:	2
equilibrium cell temperature [°C]:	20
initial delay [s]:	60
reference power [μCal/s]:	15
stirring speed [RPM]:	270

3.2.8 Pyrene actin assays

Pyrenyliodoacetamide-labeled actin monomers (pyrene-actin) provide a fluorescent readout of actin filament polymerization because a 30-fold increase in fluorescence occurs on incorporation of a labeled actin subunit into the polymer. Only low levels (5–10%) of pyrene labeled actin are required for a strong signal. For actin assembly, a final concentration of 1.8-2 μM actin was routinely used. The reaction (800 μl) was started by addition of actin.

3.2.8.1 The nucleating activity of formins

To test the nucleating activity of formins, the chosen concentrations of the actin/pyrene-actin was so low (usually 1-2 μM total actin) that an extended lag phase occurred in the control sample. For actin assembly under physiological salt conditions the reaction mixture was supplemented with 100 mM KCl.

4 Results and discussion

This section of the thesis contains the description of protein cloning, expression, purification, crystallization, and functional studies of IGFBPs and formins

4.1 Cloning, purification, crystallization and structure determination of IGFBPs domains

4.1.1 Construct design and cloning

All constructs of N- and C-terminal domains of IGFBP-4 and IGFBP-1 were designed, expressed and purified based on previously published literature (Kalus et al., 1998; Siwanowicz et al., 2005; Sala et al., 2005). Domain organization of IGFBP-5 was determined by various methods, including NMR spectroscopy and limited proteolysis (Kalus et al., 1998). Because of the high degree of sequence homology among IGFBPs, the estimates of the IGFBP-5 studies could be extended and applied to the remaining members of the family. Various constructs of the N-terminal and C-terminal domains of IGFBP-4 were designed and cloned by our group during the last few years (Siwanowicz et al.,2005; Zeslawski et al., 2001). The C-terminal domain of IGFBP-1 used in this work was similar to that crystallized recently by Sala et al. (2005). The gene of IGFBP-1 was amplified from the cDNA purchased in the BD Biosciences Clontech (USA). Mutants of the C-terminal domain of IGFBP-4 were designed in order to secure crystals of better quality. The domain organization of studied IGFBPs is shown in Figure 4.1.1. The constructs design, choice of expression vectors, fusion tags and means of their removal is described in a detail in Chapter 3. The list of the constructs and mutants of IGFBPs generated in this work, with references to the primers used, is enclosed in Table 4.1.1.

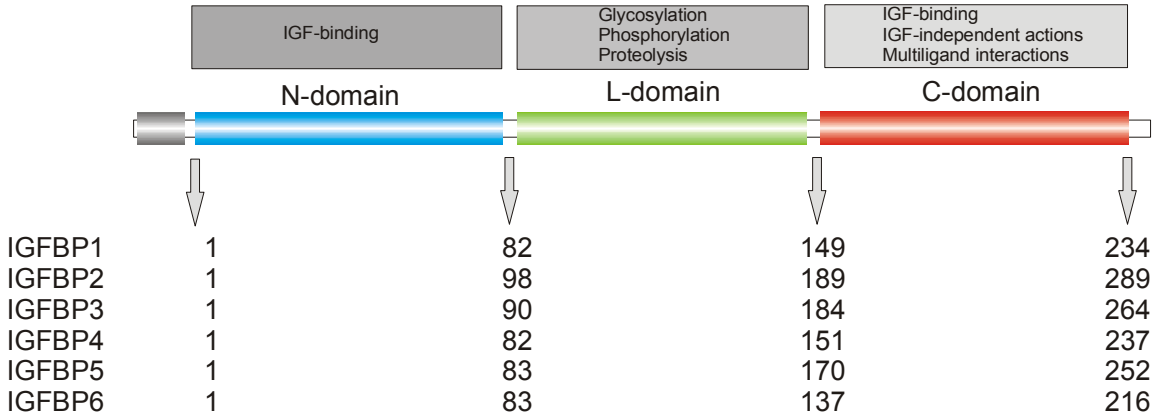


Figure 4.1.1. Domain organization of IGFBP-1-6. Domains and their lengths are shown in blue (the N-terminal domain), green (the central linker domain), and red (the C-terminal domain).

Table 4.1.1. List of IGFBP constructs used in this thesis

No.	Construct name	Primer	Vector	Expression in *E. coli*
1	NBP-4 (1-92)		pET28a	high, insoluble, refold.
2	NBP-4 (3-82)		pET28a	high, insoluble, refold.
3	CBP-1 (141-234)	5,6’	pET28a	moderate, insoluble, refold.
4	CBP-4 (151-232)		pET28a	high, insoluble, refold.
5	CBP-4 (151-232) in150WK151	10’	pET28a	high, insoluble, refold.
6	CBP-4 (151-232) in150WK151 N189F	11,12’	pET28a	high, insoluble, refold.
7	CBP-4 (151-232) in150WK151 N189F K211W	13,14’	pET28a	high, insoluble, refold.

refold. – protein obtained after refolding
‘ – primers are enclosed in Table 3.1

Properties of the designed constructs, such as theirs theoretical isoelectric points (pI), molecular masses (MW) and extinction coefficients (E_{280}) are shown in Table 4.1.2. The information was deduced on the basis of the amino acid

composition with the aid of the program ProtParam on www.expasy.org. Examination of the already existing methods of purification of similar or homologous proteins was used in designing purification protocols.

Table 4.1.2. Physicochemical properties of N- and C-terminal constructs of IGFBP-1 and -4 and IGF-1.

No.	Construct name	MW(Da)	E_{280} (M^{-1} cm^{-1})	pI
1	IGF-1	7649	4200	7.76
2	IGFBP-1	25270	37565	5.09
3	IGFBP-4	25954	10730	6.87
4	NBP-4 (1-92)	9803	3280	5.28
5	NBP-4 (3-82)	8475	3280	7.65
6	CBP-1 (141-234)	10865	24325	7.72
7	CBP-4 (151-232)	9096	7330	7.81
8	CBP-4 (151-232) in150WK151	9416	12865	8.31
9	C-BP4 (151-232) in150WK151 N189F	9449	12865	8.31
10	CBP-4 (151-232) in150WK151 N189F K211W	9507	18365	7.81

4.1.2 Expression and purification

Purification of all IGFBP constructs followed a similar protocol, involving three different liquid chromatography steps: immobilized metal affinity chromatography on Ni-NTA in denaturating conditions, cation exchange followed

by the gel filtration chromatography. Major variations of the procedure considered the conditions of the His- or His/T7-tag cleavages (determined by the use of different enzymes).

4.1.2.1 Solubilization of inclusion bodies

Expression of recombinant proteins in *E. coli* often leads to production of aggregated and insoluble inclusion bodies (IBs). This has been the case for all IGFBP domains investigated in this thesis. In order to obtain soluble, native proteins, the IBs had to be solubilized and refolded. Solubilization was achieved by a 12 h incubation, accompanied by vigorous stirring of the previously disrupted cells in solutions containing 6 M guanidinium hydrochloride, under strong reducing conditions (20 mM β-ME). Any insolubilities were removed and samples were clarified by centrifugation.

4.1.2.2 Affinity chromatography (Ni-NTA)

All expressed proteins were 6-His-tagged what enabled the use of nickel affinity chromatography under denaturating and reducing conditions. The pH of the supernatant was adjusted to 8.0 with 1 M NaOH, which is optimal for the binding to the Ni-NTA resin. A Ni-NTA slurry was added and binding was performed for 1-2 h with gentle agitation. The ratio of the Ni-NTA matrix used to the amount of the His-tagged protein is crucial for purity of the protein. It is more efficient to use less resin and perform a stepwise elution, obtaining a pure, concentrated protein in a shorter time.

4.1.2.3 Refolding

Refolding is initiated by reducing the concentration of a denaturant, what can be performed by dialyses, dilution or buffer exchange by gel filtration

(Tsumoto et al., 2003). The rapid dilution technique was applied for refolding of the IGFBP constructs. The volume ratio of the denaturant containing sample to the buffer was kept at 1:50, resulting in the final Gu-HCl concentration of less than 0.15 M. The presence of low concentrations of denaturant or other low molecular compounds like arginine-HCl or glycerol suppresses aggregation. The GSH-GSSG redox system was added to enable shuffling of the disulfide bridges. This disulfide shuffling mimics the action of disulfide isomerase enzymes *in vivo*, preventing the structure from being prematurely rigidified by forming disulfide linkages. Sufficient time is thus given for the protein to collapse into a global free energy minimum state. The recovery of the protein of interest varied between 50-70%. In general, the refolding of C-terminal constructs of IGFBPs were less efficient the N-terminal constructs.

4.1.2.4 Ion exchange chromatography

The second purification step was performed using anion exchange chromatography. The theoretical pI informs of the net charge of a protein placed in a solvent of a given pH. Thus a proper buffer and ion-exchanger (cation- or anion-exchanger) medium combination can be chosen. All constructs of N- and C-terminal domains of IGFBPs were purified on a 8-ml MonoS column. Proteins were fractionated with a linear NaCl gradient, the length of which was optimized for each construct (Figure 4.1.2.). It is common that a protein of an acidic pI would bind to a cation-exchanger resin at pH > pI. This can be observed for the N-terminal domains of IGFBP-4, which bind the MonoS medium in a buffer of pH 7.3 even though their theoretical isoelectric points are low (pI 6.3). Polar arrangement of charged residues resulting in positively charged patches, can explain this phenomenon.

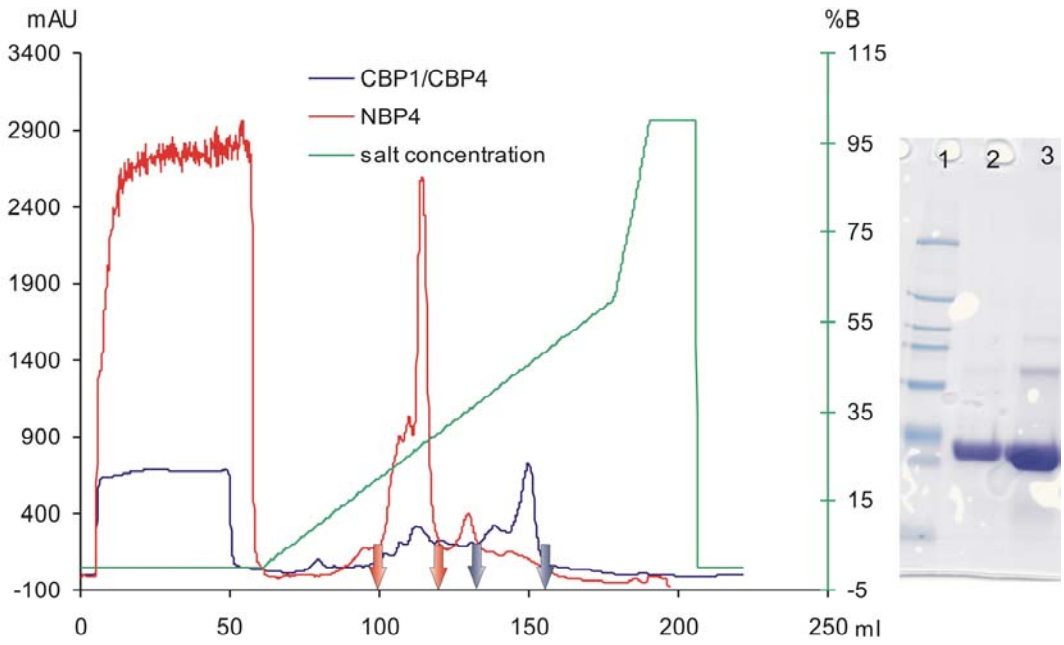


Figure 4.1.2. A typical chromatogram and the SDS-PAGE analysis of the N- and C-terminal domains of IGFBPs after MonoS ion exchange chromatography. The fractions pooled are contained between the two arrows (NBP-4 in red, CBP-1/CBP-4 in blue); SDS-PAGE (line 1, marker; line 2, CBP-1; line 3, NBP-4).

4.1.2.5 Gel filtration chromatography

Gel filtration was the final step of purification enabling separation of the proteins from possible aggregates, digested tag peptides, and the restriction protease used for the tag cleavage. Complete exchange of the buffer was also possible, removing any low molecular weight substances that could interfere with NMR measurements. The column used, a 120-ml Superdex 75 prep grade, is characterized by good resolution in the range 10-75 kDa. To make the best of its capability, low flow rates (0.8 ml/min) were used and the samples not larger than 6 ml were loaded onto the column. A typical chromatogram and the SDS-PAGE of the N- and C-terminal domains of IGFBPs purification is shown in Figure 4.1.3.

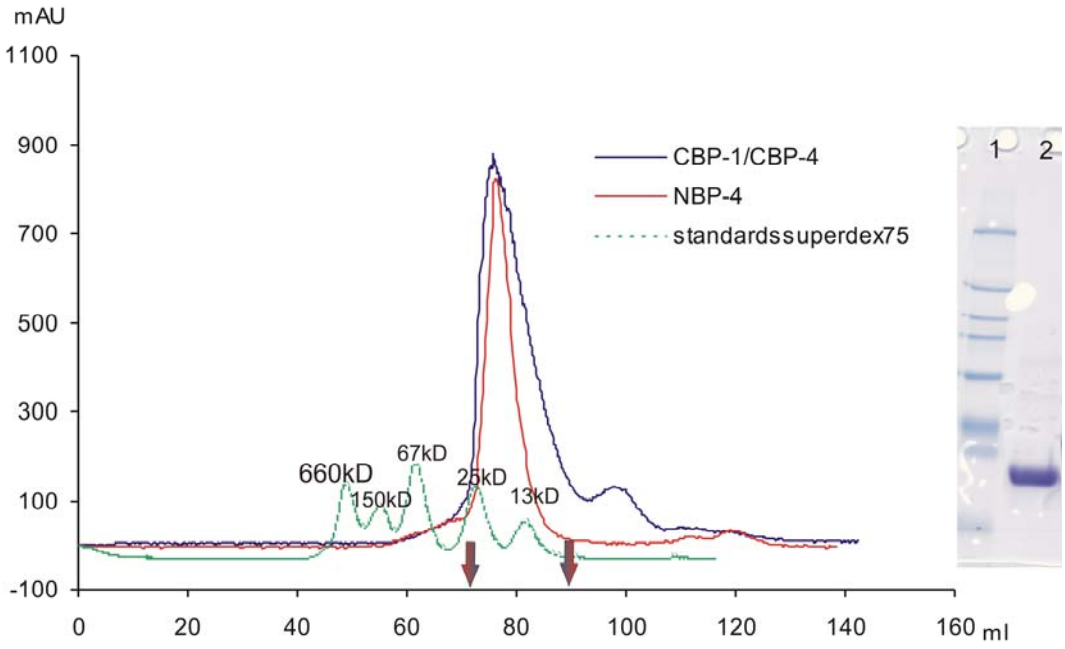


Figure 4.1.3. A typical chromatogram and the SDS-PAGE analysis of the N- and C-terminal domains of IGFBPs after gel filtration. The fractions pooled are contained between the two arrows; SDS-PAGE (line 1, marker; line 2, CBP-1).

4.1.3 Functional and structural studies

4.1.3.1 A gel filtration mobility shift assay

The analytical gel filtration binding assays were carried out to estimate the IGF-1 binding activity of purified constructs of IGFBPs. To observe the formation of a stable protein complexes in a gel filtration experiment, the binding must occur with at least low micromolar affinity. Higher K_D values cause the complex to dissociate during the purification. The proteins were mixed in equimolar ratios and separated on an analytical Superdex 75 column. In order to examine all possible interactions between IGFBPs constructs and IGF-1, the different combinations of the proteins were prepared. Figure 4.1.4 summarizes the results of the assays performed for NBP-4, CBP-1, CBP-4, and IGF-1.

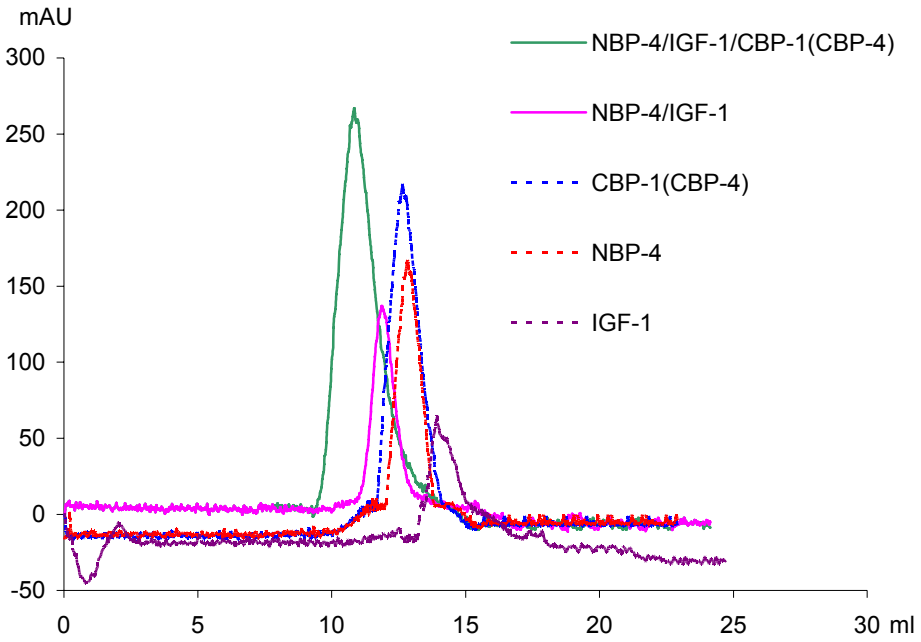


Figure 4.1.4. The gel filtration mobility shift assay on the analytical Superdex 75 column. The chromatogram shows the profiles of ternary complexes (NBP-4/IGF-1/CBP-4 and NBP-4/IGF-1/CBP-1) in green, the binary complex (NBP-4/IGF-1) in pink, CBP-1 and CBP-4 in blue, NBP-4 in red, IGF-1 in violet.

The experiments showed that all tested N-terminal constructs formed binary complexes with IGF-1, while the C-terminal domains required the presence of the N-terminal domains for the ternary complexes formation with IGF-1. Thus, formation of the NBP-4/IGF-1/CBP-4(also mutants) complex and the hybrid ternary complexes of NBP-4/IGF-1/CBP-1 was observed (Figure 4.1.4). The interaction of various constructs of N-terminal domains with C-terminal domains of IGFBP-1 and IGFBP-4 was not detected in gel filtration binding assays. This is in agreement with previous NMR studies, which excluded weak interactions between the C-terminal domains and IGF-1 or between the C- and N-terminal domains (Siwanowicz et al., 2005).

4.1.3.2 NMR studies of the folding and domain organization of IGFBPs

The one-dimensional proton NMR spectra of various constructs of N- and C-domains of IGFBP-1 and IGFBP-4 are shown in Figure 4.1.5.

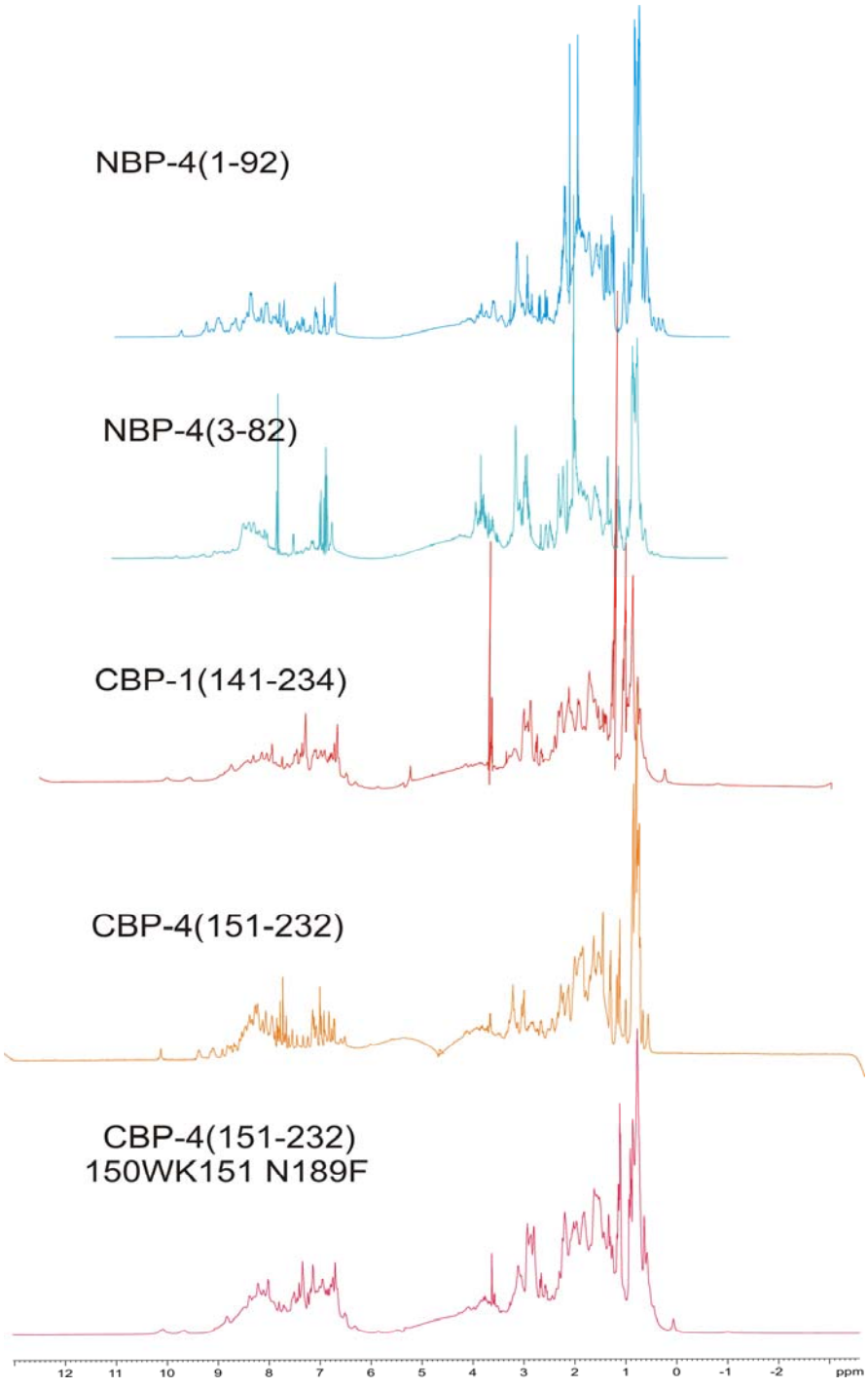


Figure 4.1.5. Characterization of the structural integrity of N- and C-domains of IGFBP-1 and IGFBP-4 by one-dimensional proton NMR spectrum.

Inspection of such spectra yields semi-quantitative information on the extent of the folding in partially structured proteins or their domains (Rehm et al, 2002). The appearance of intensities at chemical shifts near ~8.3 ppm is an indicator for a disordered protein, as this is a region characteristic of backbone amides in a random coil configuration (Wüthrich, 1986). On the other hand large signal dispersion beyond 8.3 ppm proves a protein to be folded. Due to the different chemical environment and thus the varying NMR shielding effects the NMR resonances of single protons will be distributed over a wide range of frequencies. The spectra of the previously studied full-length IGFBP-3, IGFBP-4 and IGFBP-5 exhibit a substantial peak at 8.3 ppm, suggesting that the full-length protein is only 50%-60% folded (Kalus et al., 1998; Siwanowicz et al., 2005).

The N-terminal domain of IGFBP-4(3-82 and 1-92) gives spectra with a typical intensity pattern of a folded protein. The spectra of the C-terminal fragments, CBP-4(151-232) and CBP-1(141-234), reveal that there are some unstructured regions located in these C-terminal fragments. The peak-width in a spectrum of the CBP-1(141-234) suggests that this construct might dimerize in solution. Based on these spectra it can be concluded that the whole central variable domain of IGFBP-4 is in a random coil conformation. Because of the high homology between IGFBPs, this observation can be extended also to IGFBP-1

4.1.3.3 ITC measurements

The isothermal titration calorimetry allows to characterize biomolecular interactions by monitoring chemical reaction initiated by the addition of a binding component. Measurement of the heat generated during the interaction allows to determine thermodynamic parameters like dissociation constants or reaction stoichiometry. The isothermal titration calorimetry was used to study the interaction between NBP-4, IGF-1 and CBP-1/CBP-4 during formation of the ternary complex (Figure 4.1.6). Table 4.1.3 summarizes the results of ITC measurements for the formation of two different ternary complexes. Data

obtained from ITC measurements suggest that in the 'hybrid' ternary complex the interaction between CBP-1 and NBP-4/IGF-1 is weaker than in the NBP-4/IGF-1/CBP-4 complex.

Table 4.1.3. ITC data. Dissociation constants of ternary complexes

Complex	Reservoir	Titrant	K_D [nM]
NBP-4(1-92)/IGF-1/CBP-1(141-234)	CBP-1 + IGF-1	NBP-4	750±190
	NBP-4 + IGF-1	CBP-1	1186±190
NBP-4(3-82)/IGF-1/CBP-4(151-232)	CBP-1 + IGF-1	NBP-4	266±28*
	NBP-4 + IGF-1	CBP-1	450±100*
* - data from Siwanowicz et. al. (2005)			

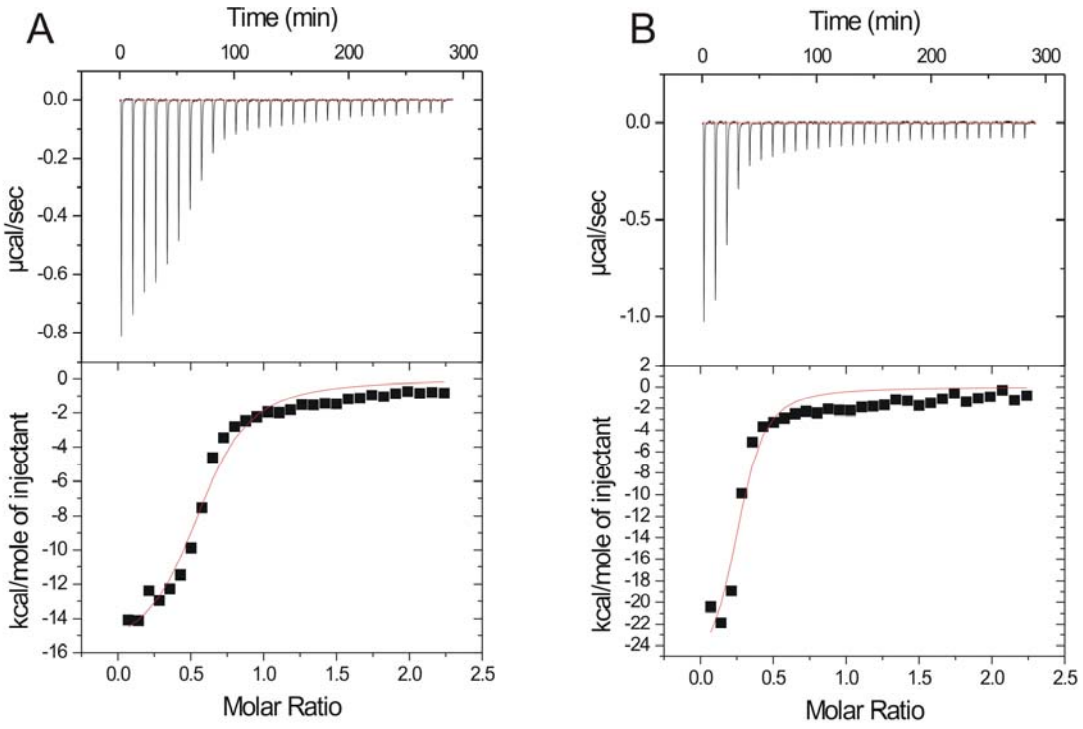


Figure 4.1.6. Examples of ITC data for the formation of the 'hybrid' ternary complex (NBP-4(1-92)/IGF-1/CBP-1(141-234). (A) IGF-1 + CBP-1 titrated with NBP-4; (B) IGF-1 + NBP-4 titrated with CBP-1.

Studies of Siwanowicz et al. (2005) indicate that NBP-4(3-82) and NBP-4(1-92) bind IGF-1 with K_Ds of ca. 0.8 μM and 0.3 μM, respectively. The presence of the C-terminal domain increases affinities of both NBP-4s to IGF-1, resulting in much tighter binding.

4.1.4 Structure of IGFBPs/IGF-1 complexes

4.1.4.1 Crystallization of the ternary and binary complexes

The ternary complexes of NBP-4(1-92)/IGF-1/CBP-1(141-234), NBP-4(3-82)/IGF-1/CBP-4(151-232) and NBP-4(1-92)/IGF-1/CBP-4(151-232) (mutants) were prepared by mixing equimolar amounts of the components. The complexes were separated from any excess of free proteins by gel filtration chromatography (Figure 4.1.7).

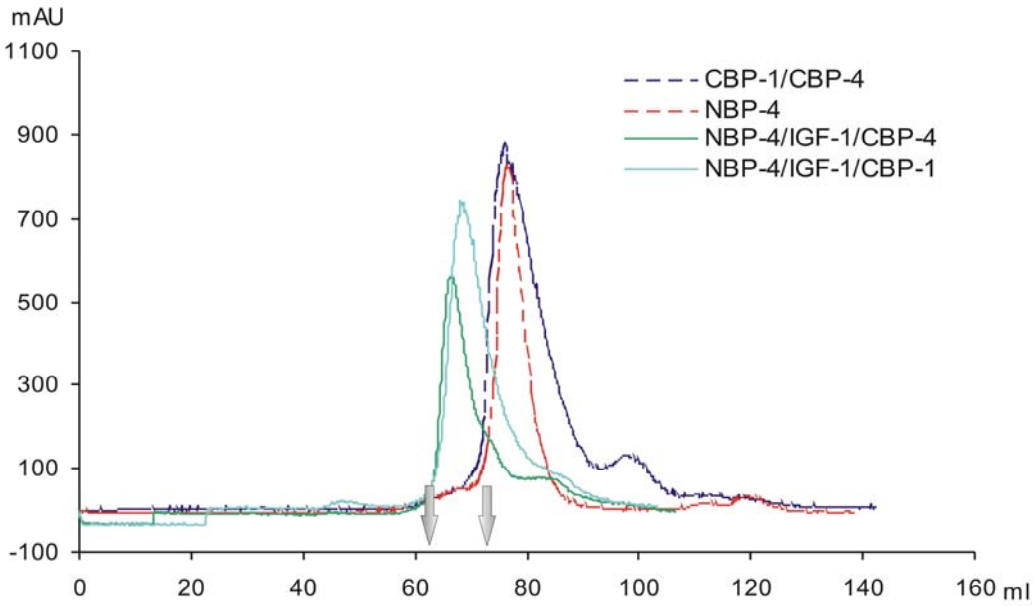


Figure 4.1.7. Purification of ternary complexes. The elution profile from the preparative 120 ml Superdex 75 column. The fractions pooled are contained between the two arrows.

The presence of uncomplexed proteins might have negatively affected crystallization. For the ternary complex, addition of a slight molar excess of the C-domain removed any remaining binary complex species in the mixture. Any surplus proteins were thus easily separated from the ternary complex. Gel filtration enabled also a thorough buffer exchange. The solvent used contained low salt (50 mM NaCl) and buffering substance (5 mM Tris, pH 8.0) to minimize any possible interference with crystallization reagents. Purity, homogeneity and folding of the proteins to be used in crystallization trials were evaluated using mass spectrometry, the amino-terminal sequencing and NMR.

The complexes have crystallized in several conditions, which are shown in Table 4.1.4. The best diffracting crystals of NBP-4/(1-92)/IGF-1/CBP-1(141-234) were obtained from 20% PEG 3350, 0.2 M lithium acetate pH 7.3 after 2 days at 20ºC. The crystals had a space group P21, with unit cell parameters a=71.3 Å, b=43.7 Å, c=81.1 Å β=91.7° There are two complexes per asymmetric unit. Crystals of NBP-4/(3-82)/IGF-1/CBP-4(151-232) (space group C2, a=74.4 Å, b=50.2 Å, c=64.3 Å β=115.3°) grew after 4 months at 4ºC in 1 M lithium sulfate monohydrate and 2% PEG 8000.

The crystals of the binary complex NBP-4 (1-92)/IGF-1 were obtained from 23% PEG 1500, 25 mM Tris pH 7 after 3 weeks in a form of plates measuring ca. 0.5 x 0.3 x 0.1 mm. The crystals belong to the space group P21 (a=32.3 Å, b=39.0 Å, c=61.3 Å β=99.9°) and contained one complex per an asymmetric unit. Prior to plunge freezing, all crystals were soaked for ca. 30 s in a drop of a reservoir solution containing 20% v/v glycerol or 20% ethylene glycol as cryoprotectant.

Table 4.1.4. Crystallization conditions for ternary complexes

No	Protein	Crystallization conditions
1	NBP-4(1-92)/IGF-1/CBP-1(141-234)	**Crystal Screen I 10** 0.2 M Ammonium acetate, 0.1M Sodium acetate trihydrate pH 4.6, 30% w/v Polyethylene glycol 4000 **Peg/Ion 21** 0.2 M Sodium formate pH 7.2, 20% w/v Polyethylne glycol 3350 **Peg/Ion 22** 0.2 M Potassium formate pH 7.3, 20% w/v Polyethylne glycol 3350 **Peg/Ion 23** 0.2 M Ammonium formate pH 6.6, 20% w/v Polyethylne glycol 3350 **Peg/Ion 24** 0.2 M Lithium acetate dihydrate pH 7.9, 20% w/v Polyethylne glycol 3350 **Peg/Ion 30** 0.2 M Ammonium acetate pH 7.1, 20% w/v Polyethylne glycol 3350 **Peg/Ion 31** 0.2 M Lithium sulfate monohydrate pH 6.0, 20% w/v Polyethylne glycol 3350 **Crystal Screen Lite 10** 0.2 M Ammonium acetate, 0.1 M sodium acetate trihydrate pH 4.6, 15% w/v Polyethylene glycol 4000 **Grid Screen PEG 6000 B3** 0.1 M MES pH 6.0, 20% w/v Polyethylene glycol 6000
2	NBP-4/(3-82)/IGF-1/CBP-4(151-232)	**Crystal Screen I 49** 1.0 M Lithium sulfate monohydrate, 15% w/v PEG 8000 **Crystal Screen I 50** 1.0 M Lithium sulfate monohydrate, 15% w/v PEG 8000 **In-house factorial 30** 0.1 M Ammonium sulfate, 0.1 M Hepes/NaOH pH 7.3,

		20% PEG 4000
3	NBP-4(1-92)/IGF-1/CBP-4(mutants)	**Peg/Ion 21** 0.2 M Sodium formate pH 7.2, 20% w/v Polyethylne glycol 3350 **Peg/Ion 22** 0.2 M Potassium formate pH 7.3, 20% w/v Polyethylne glycol 3350 **Peg/Ion 23** 0.2 M Ammonium formate pH 6.6, 20% w/v Polyethylne glycol 3350 **Peg/Ion 24** 0.2 M Lithium acetate dihydrate pH 7.9, 20% w/v Polyethylne glycol 3350 **Peg/Ion 30** 0.2 M Ammonium acetate pH 7.1, 20% w/v Polyethylne glycol 3350 **Grid Screen PEG 6000 B3** 0.1 M MES pH 6.0, 20% w/v Polyethylene glycol 6000

4.1.4.2 Structure determination

The data for the NBP-4(1-92)/IGF-1 crystals were collected from shock frozen crystals at a rotating anode laboratory source. The structure was determined by molecular replacement (Molrep program of the CCP4 suite). The structure of the complex of IGF-1 and a fragment of the N-terminal domain of IGFBP-4 (residues 3-82) (PDB entry 1WQJ; (Siwanowicz et al., 2005)) was used as a probe structure. Rotation search in the Patterson space yielded one peak of height 12.11 σ over the highest noise peak of 4.21 Å. Translation search gave a 14.47 σ peak over the noise height of 4.49 σ. The initial R-factor of the model was 0.47. The model was completed and revised manually using Xfit software (McRee, 1999). Arp/wArp was used to add solvent atoms (Lamzin et al., 1993). The structure was finally refined by the Refmac5 program. Final electron density maps were of good quality; there were however no interpretable densities for residue Pro63 and side chains of residues Glu11, Glu12, Lys13, Arg16, Thr37,

Leu42, Glu66, His70, Gln76, Met80, Glu81 and Leu82 in NPB-4(1-92) model. The IGF-1 model had no interpretable electron density for the region Gly30-Pro39 and side chains of Arg50 and Glu58. These parts were removed from the model. The final R crystallographic factor was 0.23 and R_{free} 0.27. The structure of the NBP-4(1-92)/IGF-1 complex was then used as a molecular replacement probe for the data of the NBP-4(1-92)/IGF-1/CBP-1(141-234) crystals. Rotation search in the Patterson space yielded two peaks of heights 8.37 and 7.1 σ over the highest noise peak of 4.3 σ. Translation search gave a 6.67 σ peak over the noise height of 3.82 σ for the first complex and 10.92 σ over 4.95 σ for the second one. The initial R-factor of the model was 0.47. Phases calculated at this point allowed the building of a partial model of the missing CBP-1 part, the structure was then refined by a subsequent use of Refmac5 and the manual model building. Non-crystallographic symmetry was used to improve the process as the asymmetric unit contains two complexes. Due to limited quality of the experimental data it was not possible to refine the model of the whole complex below the R-factor of 25.3 an R-free of 34.7 with acceptable stereochemistry. The regions Gln166-Ile173 and Asp197-Gly198 in CBP-1, and Gly30-Gln40 in IGF-1 had no interpretable electron density and were removed from the model as well as flexible sidechains invisible on the electron density map.

The refined model of NBP-4/(1-92)/IGF-1/CBP-1 was then used for molecular replacement with the diffraction data of the NBP-4/(3-82)/IGF-1/CBP-4(151-232) crystals. The data, although good quality up to 2.1 Å, did not allow building a structure using only the binary complex as a search model. The molecular replacement using NBP-4/(1-92)/IGF-1/CBP-1 was however very clear. Rotation search in the Patterson space yielded peak of height 8.59 σ over the highest noise peak of 4.77 σ. Translation search gave a 14.75 σ peak over the noise height of 4.74 σ. The initial R-factor of the model was 0.48 and dropped rapidly to 0.43 after rigid body refinement. At this stage, the calculated phases were improved by the DM program and the phases obtained were used for the automatic model building in Arp/wArp. About 80% of the model was built automatically and it was further completed and revised manually using the Xfit

Table 4.1.5. Data collection and refinement statistics

Data collection	NBP-4(1-92)/IGF-1/CBP-1	NBP-4(3-82)/IGF-1/CBP-4
X-ray source	BW6, DESY, Hamburg	ID29, ESRF, Grenoble
Space group	$P2_1$	C2
Cell constants (Å)	a=71.28	a=74.4
Resolution range (Å)	b=43.66 β =91.67	b=50.25 β =115.3
Wavelength (Å)	c=81.15	c=64.3
Observed reflections	20-2.8	30-2.1
Unique reflections	1.05	0.97
Whole range	91185	83038
Completeness (%)	13980	12370
R_{merge}	83.4	92.6
$I/\sigma(I)$	10.0	3.3
Last shell	11.9	25.4
Resolution range (Å)	2.8-2.9	2.1-2.2
Completeness (%)	60.1	68.8
R_{merge}	37	9.4
$I/\sigma(I)$	4.5	8.4
Refinement		
No. of reflections	12605	12315
R-factor (%)	25.3	19.9
R_{free} (%)	34.7	25.6
Average B ($Å^2$)	51.8	20.5
R.m.s bond lenght (Å)	0.03	0.007
R.m.s. angles (°)	3.18	1.08
Content of asymmetric unit		
No. of protein complexes	2	1
No. of protein residues/atoms	407/3044	460/1845
No. of solvent atoms		254

software. Arp/wArp was used to add solvent atoms. The final model encompasses residues Gly151-His229 of the C-terminal domain of IGFBP-4, Ala3-Leu82 of the N-terminal domain of IGFBP-4 and Pro2-Leu64 of IGF-1. The region between Ser34 and Ala38 is missing in the model; IGF-1 was cleaved in this region since the distance between Ser34 and Ala38 is over 17 Å. This fact explains the difficulties in repeating the crystallization of the complex since this region is involved in many crystal contacts. Some solvent exposed sidechains are also missing in the model and were removed from the structure. The final R crystallographic factor was 0.20 and R_{free} 0.26. Data collection and refinement statistics are shown in Table 4.1.5.

4.1.4.3 Overall structures of the NBP-4(3-82)/IGF-1/CBP-4 and NBP-4(1-92)/IGF-1/CBP-1 ternary complexes

Fold of the C-terminal domains of IGFBP-4 and IGFBP-1

The C-terminal domain of IGFBP-4 is a flat molecule without a pronounced hydrophobic core (Figure 4.1.8 and 4.1.10). The secondary structure elements comprise two antiparallel helices α1 (Cys153(C)-Ala165(C)) and α2 (His172(C)-Ile177(C)), followed by four β strands: β1 (Asn182(C)-Asp184(C)), β2 (Asn188(C)-His195(C)), β3 (Lys204(C)-Arg210(C)), β4 (Val214(C)-Leu216(C)) that form a single, twisted β-sheet (Figure 4.1.8). Strands β1 and β2, as well as β3 and β4, are connected by type I turns, while a hairpin-like elongated loop between Cys194(C)-Gly203(C) links β2 and β3. The loop has a conformation close to β, with a regular type I' turn at the top of the loop, but lacks a proper H-bonding pattern. The C-terminal part of the molecule, starting from Pro217(C), does not have any regular secondary structure, but is instead stabilized by the disulfide bridges Cys207(C)-Cys228(C), Cys153(C)-Cys183(C), Cys194(C)-Cys205(C), and two H-bonds between main chains of Cys194(C) and Glu221(C).

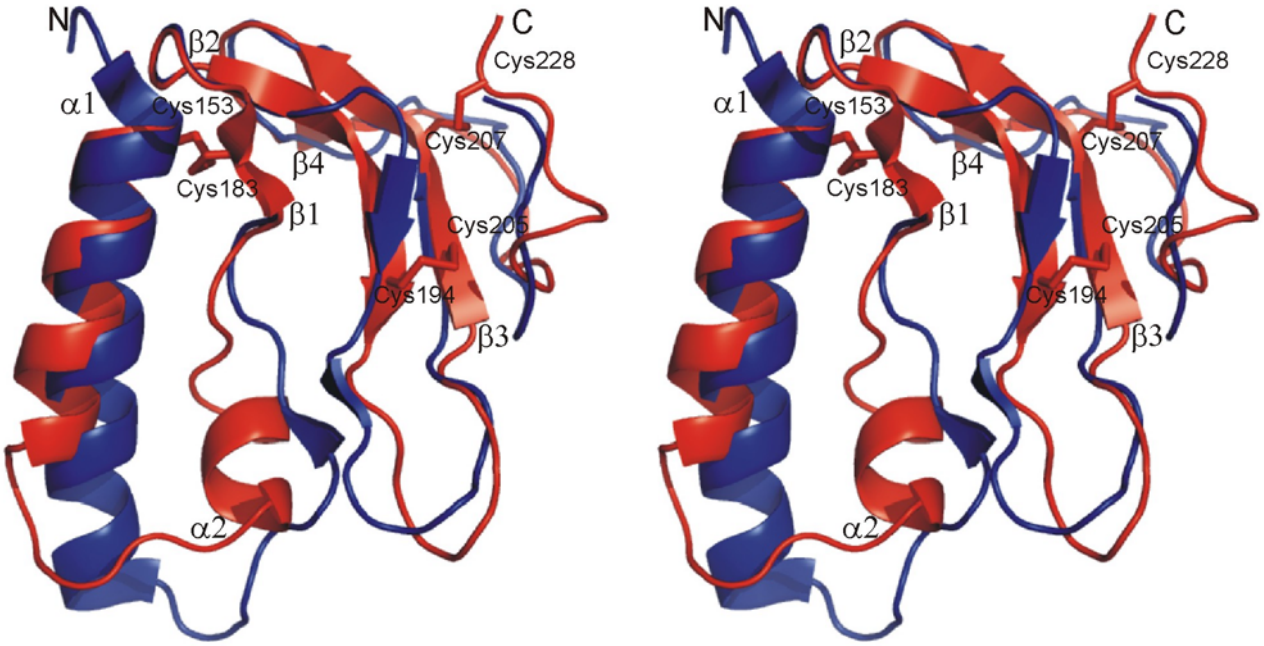
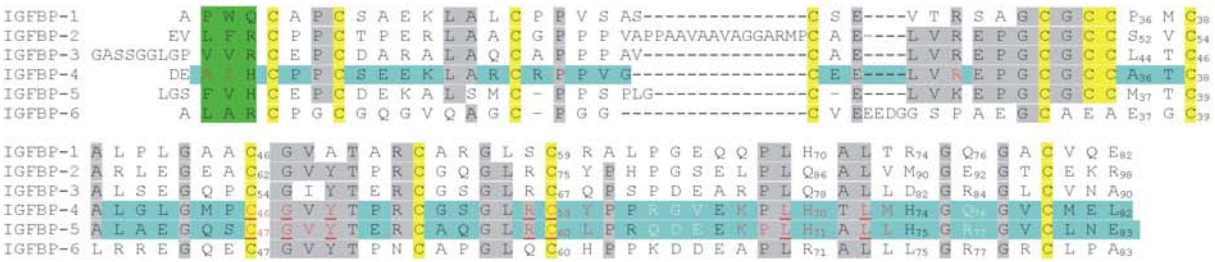


Figure 4.1.8. Comparison of structures of the C-terminal domains of IGFBP-1 (blue) (PDB code 1ZT5, (Sala et al., 2005) and IGFBP-4 (red). Disulfide bonds are shown as sticks.

The structure of CBP-1(141-234), in complex with NBP-4(1-92) and IGF-1, is less well-ordered than that of an isolated CBP-1(147-221) recently published by Sala et al. (2005). In addition, regions Gln166(C)-Ile173(C) and Asp197(C)-Gly198(C) in our CBP-1 are fully disordered. I therefore used the structure of Sala et al. (2005) in comparing CBP-1 to CBP-4. The defined parts of our CBP-1 are nearly identical to those of free CBP-1, with of the backbone rmsd 0.96 Å. The structures of the C-terminal domains of isolated IGFBP-1 (PDB code 1ZT5, (Sala et al. 2005)) and complexed IGFBP-4 have a very similar fold (Figure 4.1.8), as expected from their sequence identities (Figure 4.1.9). The rmsd for backbone atoms is 1.3 Å. A major difference is the region Ala166(C)-Pro179(C) of CBP-4 that includes the end of helix α1 and the whole α2 helix.

Human IGFBPs

N-domains:



L-domains:

C-domains:

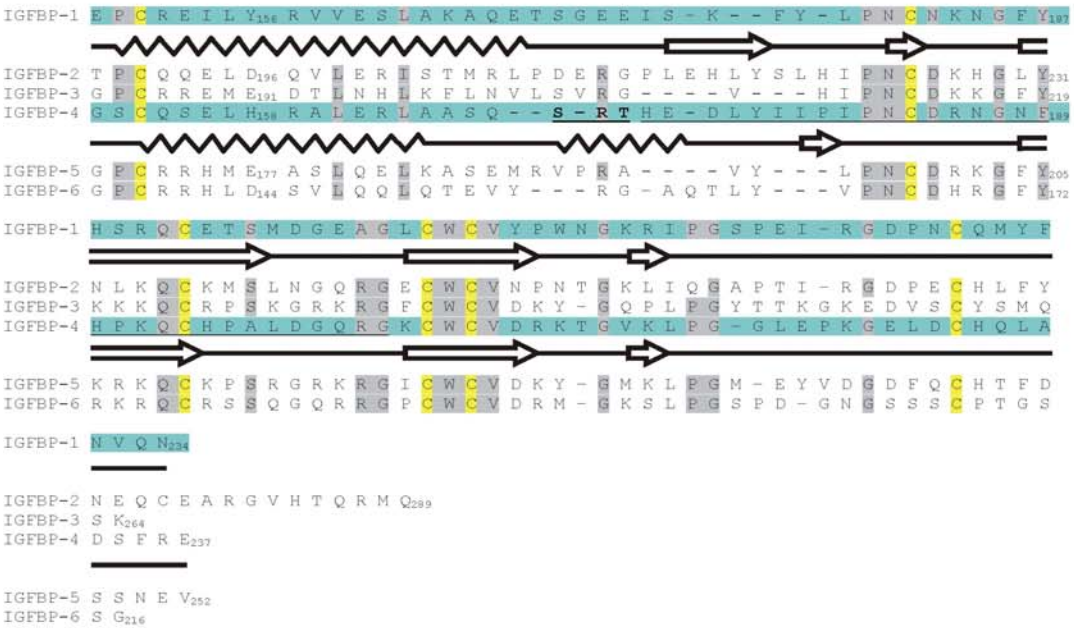


Figure 4.1.9. Sequence and structure alignment of human IGFBP-1 to –6. The N- and C-domains studied are marked by light blue. Conserved residues are indicated by gray shading and cysteines are in yellow. N-domains: Residues shown in white have no amino acid and structural homology between N-domains of IGFBP-4 and IGFBP-5. Residues that interact with IGF-1 are highlighted in red (with primary sites underlined) and hydrophobic residues of the "thumb" segments are boxed in green. For L-domains: In magenta - protease cleavage sites reported until now, amino acids labeled in blue show the calpain cleavage site. Cleavage occurs after the marked amino acid. C-domains: residues that interact with IGF-1 or NBP-4 are underlined with residues that interact with the thumb in bold.

In contrast to CBP-4, corresponding residues in CBP-1 (Lys164(C)-Tyr177(C)) are not helical but form a short β-strand. Also in CBP-4, the α1 helix is shorter by one turn and the loop following it is bent to the opposite direction. The presence of a short β-strand after the α1 helix relocates the β3 strand of CBP-1 which is longer by three residues and by about 2.8 Å away from the binding surface of IGF compared to CBP-4. The CBP-4 region Pro222(C)-Glu225(C) has a 3_{10} turn absent in the CBP-1 structure. This element is however involved in crystal contacts and its conformation might be artificial.

The C-terminal domains of IGFBPs in the ternary complexes NBP-4(3-82)/IGF-1/CBP-4(151-232) and hybrid NBP-4(1-92)/IGF-1/CBP-1(141-234)

CBP-4 interacts with both IGF-1 and NBP-4 in the complex (Figure 4.1.11). The interaction surface encompasses a side of the CBP-4 molecule that is built up by helices α1, α2, segment Ile178(C)-Arg185(C) and a hairpin-like loop between residues Cys194(C)-Gly203(C). I divided this CBP-4 interface into three parts based on its interacting partners: the part that makes contacts with IGF-1 only, the segment that interacts both with NBP-4 and IGF-1, and the stretch of residues Glu168(C)-Glu173(C) of CBP-4 which interacts with the so-called "thumb" region of NBP-4. The thumb region consists of a short stretch of the very first N-terminal residues of IGFBPs that precede the first N-terminal cysteine

(amino acids 1-5 in IGFBP-4) (Siwanowicz et al., 2005). It is worth noting that both CBP-4 and NBP-4 occupy one side, a dome part, of a pear-like IGF-1 structure (Figure 4.1.11 and 4.1.12). Gln154(C), Leu157(C), Leu161(C) of one face of helix α1, most of the residues in α2, the Ile178(C)-Arg185(C) fragment, and Cys194(C), Pro196(C), Ala197(C), Arg202(C), Gly203(C) of the Cys194(C)-Gly203(C) segment of CBP-4 make direct contacts with IGF-1. Several residues of segment Cys194(C)-Gly203(C) also make contacts to NBP-4. Most of these amino acids are hydrophobic and the interaction of CBP-4 and IGF-1 is based principally on the hydrophobic contact. None of the CBP-4 residues inserts deeply into IGF-1, except Ile180(C) whose sidechain is in a hydrophobic cleft made up by the aromatic ring of Phe25(I), the backbone of Cys6(I) and Gly7(I) on the sides, and at the bottom, by Leu10(I). Phe25(I) is believed to be involved in binding to the IGF-1 receptor (Firth and Baxter, 2002; Clemmons, 2001; Bach et al., 2005; Carrick et al., 2005).

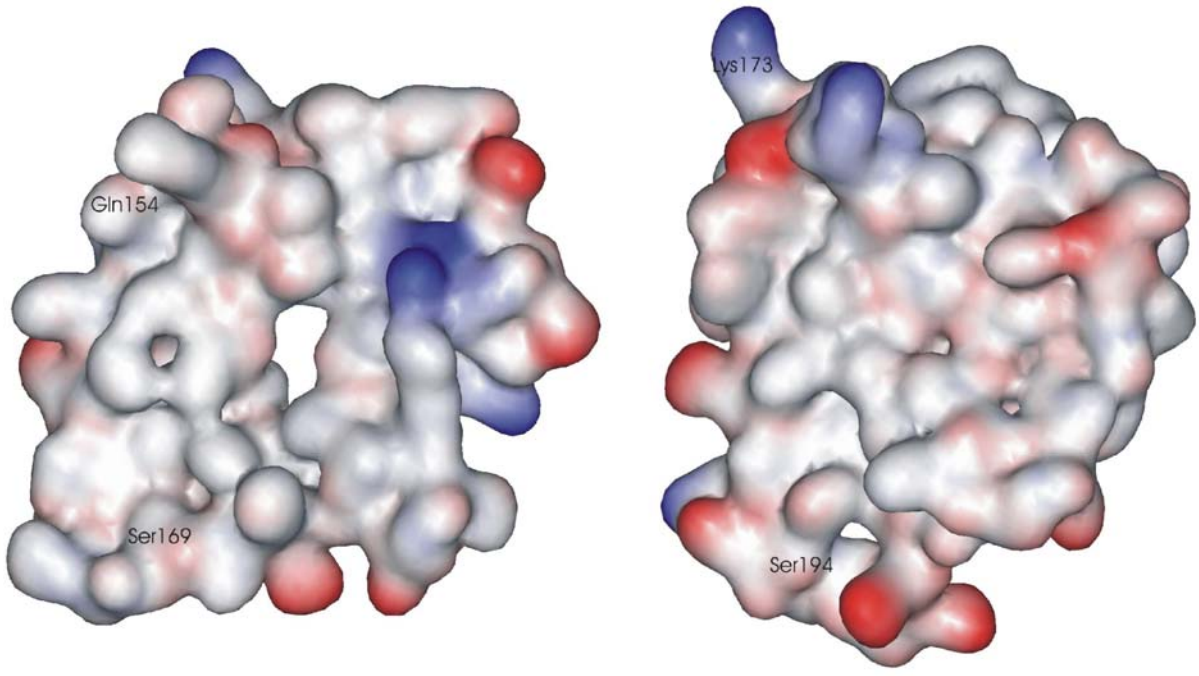


Figure 4.1.10. Surface representations of the CBP-4 (the left side) and CBP-1 structures (the right side).

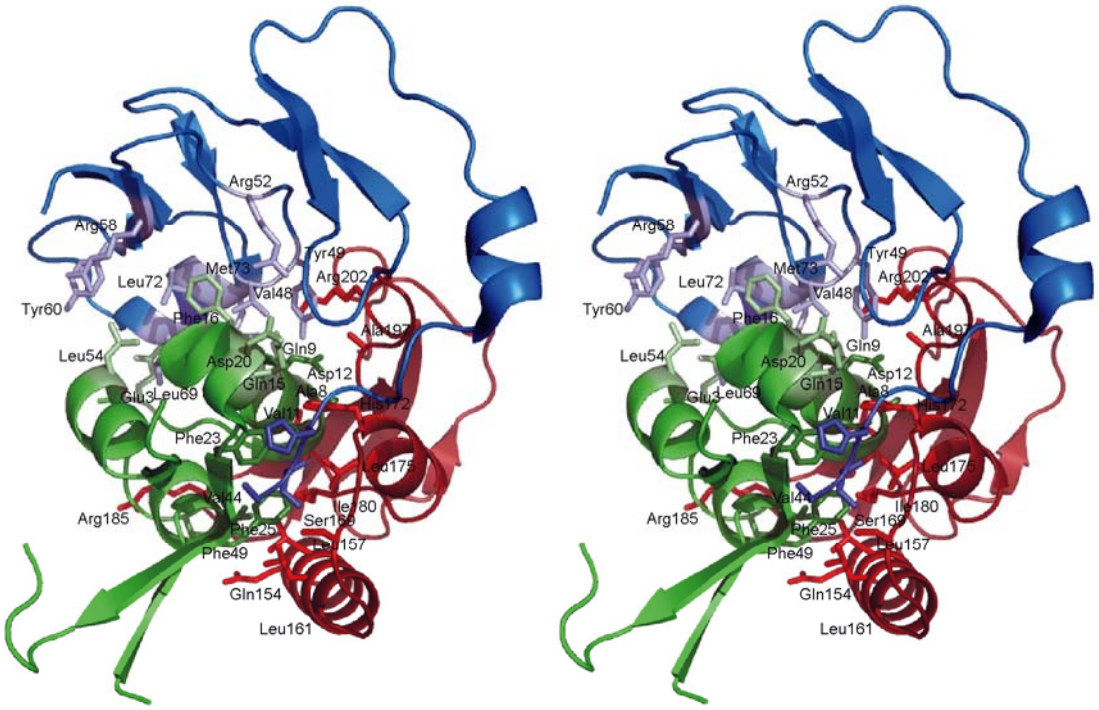


Figure 4.1.11. Overall structure of the NBP-4(3-82)/IGF-1/CBP-4(151-232) ternary complex. The NBP-4 domain is in blue, IGF-1 in green, and CBP-4 in red. The residues shown are important for binding among protein components.

This hydrophobic interaction is enhanced by Leu161(C), close to Ile180(C), which makes also contacts to Phe25(I), and it is further extended by sidechain interactions of Leu157(C)/Val44(I), His172(C)/Val11(I)/Glu15(I), Leu175(C)/Phe25(I)/Ala8(I), and Pro196(C)/Cys45(I). Residues His172(C), Leu175(C), Ile180(C), Pro181(C), Asn182(C), Pro196(C), and Ala197(C) constitute the rim of the hole present in the CBP-4 structure (Figure 4.1.10). This hole is filled by the N-terminus of the IGF α-helix between Cys6(I) and Gly19(I) (i.e. Cys6(I), Gly7(I), Ala8(I), Glu9(I), and Leu10(I)), with primary "inserting" residues being Gly7(I) and Ala8(I) (Figure 4.1.12).

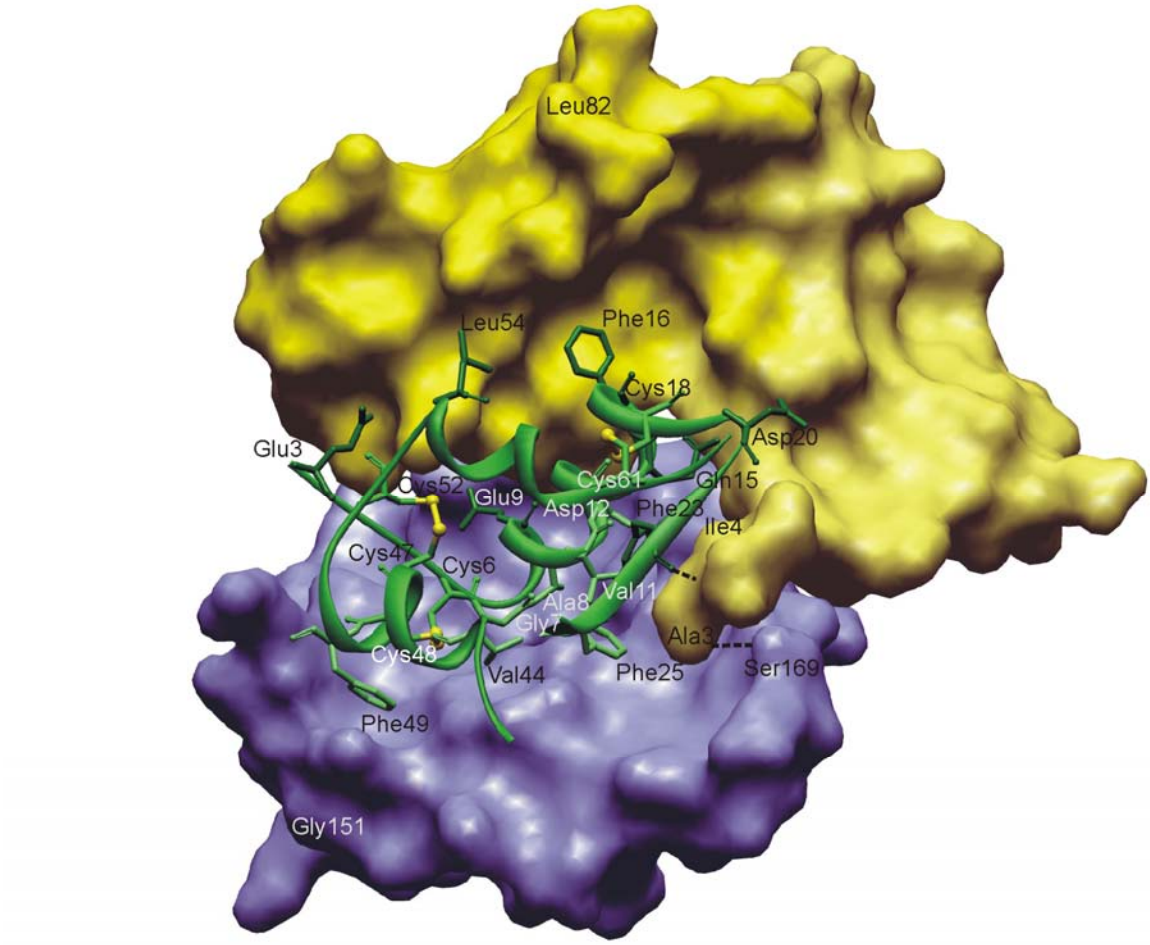


Figure. 4.1.12. The interaction of IGF-1 (green) with NBP-4(3-82) (the yellow surface) and CBP-4(151-232) (the blue surface). Residues discussed in the text are labeled and shown in stick representation; two hydrogen bonds from the thumb to IGF-1 and CBP-4 are also shown.

The other main interactions on the IGF side are made by residues of a short helix Val44(I) to Arg50(I) of IGF-1. Together with Phe25(I), these two helical segments constitute the major binding sites of CBP-4 on IGF-1 (Figure 4.1.11 and 4.1.12). Asp199(C), Gly200(C), Gln201(C), Arg202(C), which are part of a hairpin-like loop between Cys194(C)-Gly203(C), constitute direct contacts of CBP-4 to NBP-4. On the NBP-4 side, the N-terminal portion of NBP-4, from Ala3(N) up to Tyr49(N), makes contacts with CBP-4. Notable is the stacking of

the sidechain of Arg202(C) with the aromatic ring of Tyr49(N) (Figure 4.1.11 and 4.1.12). A CBP-4 segment between residues 168-173 (Glu168(C), Ser169(C), backbone Arg 170(C), Thr171(C), His172(C), Glue173(C)) adopts mostly an extended conformation and interact with the thumb region of the NBP-4 (residues 1-5 of NBP-4) (Figure 4.1.11 and 4.1.12). The two fragments form a short parallel-like β-sheet (hydrogen bonds are formed between His172(C)-N and His5(N)-O, and between Ser169(C)-OG and Ala3(N)-N)). The core sidechain interactions are purely hydrophobic even for charged residues; for example, Pro7(N) makes contact to the β-atoms of Glu173(C).

The refined model of NBP-4/(1-92)/IGF-1/CBP-1, after removing part of NBP-4 residues, was used for molecular replacement with the experimental data of NBP-4/(3-82)/IGF-1/CBP-4(151-232). The crystals of NBP-4(1-92)/IGF-1/CBP-1(141-234) diffracted to 3.2 Å and the model built is lacking several residues at the interface of CBP-1 and NBP-4 or IGF-1. I therefore did not use this structure for a detailed analysis of the CBP complexes but describe NBP-4(3-82)/IGF-1/CBP-4(151-232) which is identical with the rmsd of 0.96 Å for backbone atoms.

4.2 Discussion

The C-terminal domains of all IGFBPs show sequence homology with thyroglobulin type-1 domains (Bach et al., 2005;, Headey et al., 2004;,Sala et al., 2005; Kiefer et al., 1991; Novinec et al., 2006) and share common elements of secondary structure: an α-helix and a 3-4 β-stranded β-sheet. The core of the molecule is connected by the consensus three disulfide pairings, possesses conserved Tyr/Phe amino acids and have the QC, CWCV motifs (Novinec et al., 2006). These essential features are preserved in CBP-1, CBP-4, and CBP-6, the structures of C-domains solved so far, although there are significant variations in detail. For example, CBP-4 has helix α2, whereas the corresponding residues in CBP-1 form a short β-strand seen in other structures of the thyroglobulin type-1 domain superfamily (Sala et al., 2005;, Novinec et al., 2006). This particular

region of CBPs has high sequence diversity and is involved in the IGF complex formation and thus may perform a role of an affinity regulator.

The ternary complexes of NBP-4(3-82)/IGF-1/CBP-4(151-232) and NBP-4(1-92)/IGF-1/CBP-1(141-234) provide an understanding of the roles of N- and C-terminal IGFBP domains in modulating IGF actions and show that the N- and C-terminal domains come into close proximity mutually in the complex to enhance or stabilize IGF binding. There has been a considerable body of work to delineate the determinants of IGFBPs binding to IGFs and vice versa (Firth and Baxter, 2002; Clemmons, 2001; Payet et al., 2003; Allan et al., 2006; Carrick et al., 2005; Kibbey et al., 2006; Siwanowicz et al., 2005; Headey et al., 2004; Xu et al., 2004; Kalus et al., 1998; Zeslawski et al., 2001; Galanis et al., 2001; Vorwerk et al., 2002; Qin et al., 1998; Shand et al., 2003). The structural information presented in this work is broadly in agreement with these data, but disagree with reports of a critical role of the completely conserved Gly187(C) and Gln193(C) (in the IGFBP-4 sequence) for binding of C-domains to IGFs (Clemmons, 2001; Allan et al., 2006, Shand et al., 2003, and reference cited therein): these residues are not in contact with IGF-1 although they are close to the IGF-1/CBP-4 interface surface. A second example is provided by the recent NMR mapping of the binding surfaces of IGFBP-2 on IGF-1 (Carrick et al., 2005). These data, combined with previous mutational analyses (Clemmons, 2001), would suggest that the Gly22(I)-Phe25(I) region of IGF-1 directly interacts with the C-domain of IGFBP-2. In our structure, this segment of IGF-1 (with the exception of Phe25(I)) clearly binds to the thumb region of NBP-4. The thumb masks the IGF residues responsible for the type 1 IGF receptor (IGF-1R) binding and in turn interacts with residues 168(C)-173(C) of the C-domain. Thus although our isolated CBP-4 does not bind individually to either IGF-1 or NBP-4, in the ternary complex CBP-4 contacts both and contributes to blocking of the IGF-1R binding region of IGF-1.

Both N- and C-terminal domains of IGFBPs were reported to bind to IGFs (Firth and Baxter, 2002; Clemmons, 2001; Bach et al., 2005; Bunn and Fowlkes, 2003; Payet et al., 2003; Allan et al., 2006; Carrick et al., 2005; Kibbey et al., 2006; Fernandez-Tornero et al., 2005; Siwanowicz et al., 2005; Carric et al.,

2001; Headey et al., 2004; Xu et al., 2004; Kalus et al., 1998; Zeslawski et al., 2001; Galanis et al., 2001; Vorwerk et al., 2002; Qin et al., 1998; Shand et al., 2003). Isolated N-terminal fragments of IGFBPs bind to IGF with 10-1000-fold lower affinities than full-length IGFBPs. There also is no doubt that the C-terminal domains of IGFBPs increase the affinity of IGFBPs for IGFs (Table 4.1.3). However, there are inconsistent reports in the literature regarding the strength of direct binding between isolated C-terminal domains and IGFs (Firth and Baxter, 2002; Bach et al., 2005; Allan et al., 2006; Siwanowicz et al., 2005; Carrick et al., 2001; Galanis et al., 2001; Vorwerk et al., 2002; Shand et al., 2003). The C-domain of bovine IGFBP-2(136-279) was reported to bind IGF-1 with K_D of 23.8 nM, whereas affinity of the N-domain (residues 1-185) was 78.1 nM (Carrick et al., 2001); e.g. C-domain had higher affinity than the N-domain for IGF-1. A similar trend was observed for IGFBP-3 in one report (Galanis et al., 2001), but Vorwerk et al. (2002) found that their C-domain had a 3-fold weaker binding to IGF-1 than the N-domain, which had a K_D of 160 nM. The C-domains of IGFBP-4 (Fernandez-Tornero et al., 2005; Siwanowicz et al., 2005), IGFBP-5 (Kalus et al., 1998), and IGFBP-6 (Headay et al., 2004) showed lower binding affinities than their respective N-counterpartners. Features of the IGF-binding regions of CBP-1 and CBP-4 seen in current structures would support the latter trend. Although contact areas of NBP-4 and CBP-4 to IGF are about equal (758 $Å^2$ and 670 $Å^2$, respectively), a direct access of C-domains to IGF is obscured in two sites by N-domain residues (Figure 4.1.12, for example, the C-domain interacts through the N-terminal thumb residues with IGF). In addition, the hydrophobic interaction between IGF-1 and the miniNBP subdomain of NBP-4, consisting of interlaced protruding side chains of IGF-1 and solvent-exposed hydrophobic side chains of the NBP-4, seems to be more extensive than that seen in the CBP-4/IGF binding. It is expected that the IGF-binding structure of our complexes will be shared among six IGFBPs but the binding affinities of isolated N- and C-domains may differ among IGFBPs and between IGF-1 and IGF-2, due to sequence differences.

The NMR spectrum of the full-length IGFBP-4 indicated that the central variable domain of IGFBP-4 (Figure 4.1.9) is unstructured and flexible both, when free and when in complex with IGF-1 (Siwanowicz et al., 2005). Positions of the last C-terminal residue of NBP-4 (Leu82(N)) and the first N-terminal residue of CBP-4 (Gly151(C)) shown in Figure 4.1.12 suggest the location of the central L-domain of IGFBP-4 in the full-length IGFBP. The central domain might act as a "mechanical flap" that covers the IGF not yet wrapped by N- and C-terminal domains. Proteolysis of the IGF/IGFBP capsule would first remove the central domain by degradation. This partial removal of the capsule exposes IGF but still maintains IGF inhibition towards IGF-1R as long as the N-terminal thumb and/or CBP fragments of IGFBPs are not removed.

The first 92 residues of IGFBP-4 are 59% identical to the corresponding N-terminal residues of IGFBP-5, and the remaining residues are mostly functionally conserved. For miniNBP-5 (residues 40-92), the last 9 amino acids showed no electron density in its IGF complex structure (Zeslawski et al., 2001) and were unstructured as determined by NMR (Kalus et al., 1998). Equivalent residues therefore were not expressed in the construct NBP-4(3-82) to aid crystallization of the complex. However, residues Glu90 and Ser91 of IGFBP-4 were reported to be important for high affinity binding with IGFs (Qin et al., 1998), therefore we decided to include these residues in our extended N-terminal construct NBP-4(1-92). The two first N-terminal residues were also added because the previous IGFBP-4(3-82)/IGF-1 structure (Siwanowicz et al., 2005) indicated the importance of the N-terminal hydrophobic residues conserved among IGFBPs. A possibility that eliminating the two first negatively charged residues, Asp1 and Glu2, at the N-terminus in the IGFBP-4 could have changed the properties of this amino terminal part, also existed. These residues were therefore added to the refined N-terminal construct, generating NBP-4(1-92). It is evident from the structure of the current NBP-4(1-92)/IGF-1 binary complex that the sequence Ala83-Leu92, of which the fragment Glu84-Glu90 forms a short helix, does not contact IGF directly. In the study of Qin et al. (1998) deletion of Glu90 and Ser91 led to the reduced IGF-1 and -2 binding activity, suggesting

functional importance of these residues. The molecular structure, however, shows no direct contribution of these two residues to the formation of the IGF binding site. The presence of the 10-amino acid-long fragment may, however, have an indirect influence on IGF binding: side chains of Ile85(N), Ile88(N), and Gln89(N) shield the Tyr60(N) side chain from the solvent and constrain its conformation that otherwise would point away from the IGF surface, as can be seen in the NBP-4(3-82)/IGF-1 complex structure. Tyr60(N), along with Pro61(N), forms a small hydrophobic cleft, in which Leu54(I) of IGF-1 is inserted, thus extending a hydrophobic contact area of the two proteins.

The N-terminal domain of IGFBPs can be viewed as consisting of a globular base, corresponding to the miniBP-5 that contains an important IGF binding site (Siwanowicz et al., 2005; Kalus et al., 1998; Zeslawski et al., 2001), and an extended “palm” followed by a short hydrophobic thumb (Ala3(N), Ile4(N)) (Figure 5.1.11; Figure 5.1.12). The thumb interacts with IGF residues Phe23(I), Tyr24(I) and Phe25(I) upon complex formation. The palm is rigid because of four disulfide bonds arranged in a ladder-like plane and several inter-subdomain H-bonds. Rigidity of the N-terminal domain of IGFBPs may be of significance when the competition with IGF-1R for IGF binding is concerned. Previous studies revealed that Phe23, Tyr24 and Phe25 of IGF-1, and corresponding Phe26, Tyr27 and Phe28 of IGF-2, are important for binding to insulin and IGF-type 1 receptors (Firth and Baxter, 2002; Clemmons, 2001; Bach et al., 2005; Sakano et al., 1991; Hodgson et al., 1996; Cascieri et al., 1988; Bayne et al., 1990; Perdue et al., 1994). To displace the hydrophobic thumb that covers the primary IGF-1R binding site of IGFs (IGF-1, Phe23(I)-Phe25(I)), the receptor also has to remove the rest of the N-terminal domain, which is bound to the opposite side of the IGF-1 molecule, and does not prevent receptor binding on its own (Kalus et al., 1998). Given the conserved arrangement of the N-terminal cysteine residues and the consistent presence of two hydrophobic residues at positions -2 and -3 with respect to the first N-terminal Cys residue, we expect that this mechanism is shared by all IGFBPs.

The very first amino terminal residues of IGFBPs have been neglected to date in mutagenesis studies that aimed at delineation of these proteins' structure/function relationships (reviewed in Clemmons, 2001). Truncation of the thumb (residues 1-5) reduces the IGF-1 binding of NBP-4 to that of miniNBP-4 (residues 39-82) (Siwanowicz et al., 2005), suggesting that the palm (residues 6-39) does not contribute directly to IGF binding. The segment appears therefore to serve solely a mechanical purpose as a rigid linker between primary binding sites - the base and the thumb residues. The central element of the palm consists of a GCGCCXXC consensus motif, around which the polypeptide chain (residues Cys6-Cys23) is bent forming a disulfide ladder and assuring a proper spatial relationship between the base and the thumb (Figure 4.1.12). Substitution of the IGFBP-4 thumb with a corresponding region from two other IGFBPs does not influence the strength of IGF-1 binding markedly. Interestingly, available sequences of IGFBPs from different species show a remarkably high degree of inter-species conservation of the thumb residues within a single type of IGFBP implying that the sequence of the thumb may confer unique properties to each IGFBP.

The structural studies carried out here provide important information to aid in the design of IGFBP-based therapeutics (Firth and Baxter, 2002; Kibbey et al., 2006; Yee, 2006). The involvement of the IGF system in tumor cell growth and survival makes it an excellent target for anticancer treatment; especially in view that IGF-1R is not absolutely necessary for normal growth (Firth and Baxter, 2002, Kibbey et al., 2006; Pollak et al., 2004; Yee, 2006). Indeed, recent data has shown that targeting the IGF system produces impressive antineoplastic activity in many in vitro and in vivo models of human cancers. The most advanced are the strategies based on the direct disruption of receptor function that are based on small-molecule inhibitors of the tyrosine kinase domain of IGF-1R and on antibodies directed against IGF-1R (Firth and Baxter, 2002; Pollak et al., 2004; Yee, 2006). The possibility that these therapies will lack sufficient specificity to avoid co-targeting the insulin receptor must be however considered, and therefore carefully designed clinical trails will be required to assess the effect

on host glucose metabolism (Pollak et al., 2004). Neutralization of IGF ligands through IGFBPs (which only target the IGFs) would avoid this problem, as insulin action would be mostly unaffected (Firth and Baxter, 2002, Kibbey et al., 2006; Pollak et al., 2004; Yee, 2006). The design of the therapeutic IGFBPs or their fragments would have to take into account the building characteristics of the IGFBP/IGF structure presented here. For example, our structures revealed the importance of the N-terminal thumb hydrophobic residues for blocking the IGF/IGF-1R interaction and thus IGFBP constructs should include "long" N-terminal thumb segments.

4.3 Formins and profilins

The objective of obtaining FH1FH2 and FH2 domains of various formins was to study their structural and biochemical properties. Two recently resolved crystal structures of Bni1pFH2 alone and in complex with TMR-actin have provided new insights into the molecular details of the formin-mediated actin assembly but a detailed mechanism of actin nucleation remains not clear. Structural analysis of other members of formin family is important for understanding their precise molecular mode of action. Characterization of the formin, actin and profilin interactions using various biochemical techniques, NMR spectroscopy, x-ray crystallography and pyrene-actin assays, was the main goal of the project. Studies presented in this subsection were carried out on the three human formins (DIAPH1, DAAM1, DAAM2), one Dictyostelium discoideum formin (dDia2) and four profilins (profilin 1 and profilin 2 from Dictyostelium discoideum and human).

4.3.1 Construct design and cloning

The domain organization of studied formins is shown in Figure 4.3.1. Constructs of FH1FH2 and FH2 domains of DIAPH1, DAAM1, DAAM2, and dDia2 were designed after alignment to the amino acid sequence of the yeast formin Bni1, which was extensively characterized in the published literature (Moseley et al., 2004; Sagot et al., 2002; Xu et al., 2004). Because of difficulties with crystallization a large number of constructs and mutants were cloned, expressed, and purified. Limited proteolysis of dDia2 helped to find the regions that were stable to degradation and allowed the design of mutations that increased the stability of constructs.

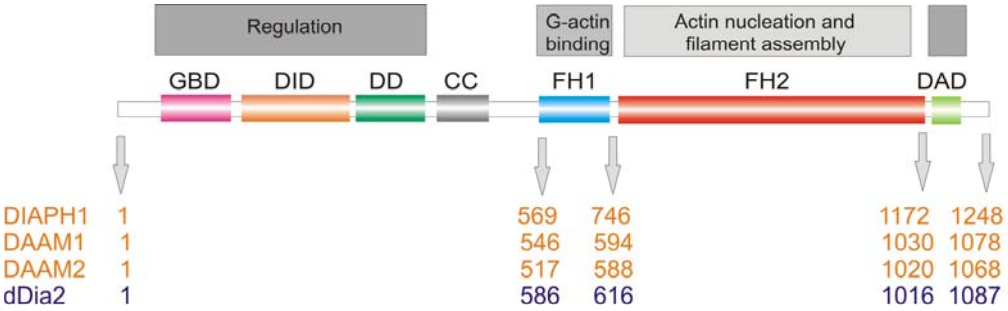


Figure 4.3.1. Schematic representation of the domain organization of formins. GBD, GTPase binding domain; DID, diaphanous-inhibitory domain; DD, dimerization domain; CC, coiled coil; FH1, formin homology 1 domain; FH2, formin homology 2 domain; DAD, diaphanous auto-regulatory domain; DIAPH1, human diaphanous-related formin-1; DAAM1, human disheveled-associated activator of morphogenesis 1; DAAM2, human disheveled-associated activator of morphogenesis 2; dDia2, Dictyostelium discoideum diaphanous-related formin-2.

Table 4.3.1. Results of cloning and expression of various DIAPH1 constructs

No.	Construct name	Primer	Vector	Expression in *E. coli*
1	DIAPH1(614-1134)	16,17	pGEX-4T-1	high, agg., deg., prec.
2	DIAPH1(553-1134)	15,17	pGEX-4T-1	high, agg., deg., prec.
3	DIAPH1(553-1200)	15,18	pGEX-4T-1	low, agg., deg., prec.
4	DIAPH1(614-748)	16,19	pGEX-4T-1	high
5	DIAPH1(553-748)	15,19	pGEX-4T-1	high
agg. – aggregation; deg. – degradation, prec. - precipitation				

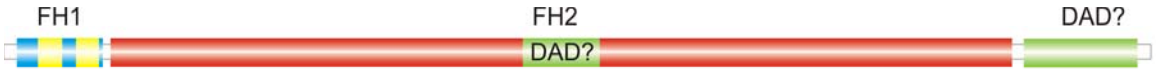
All human formins: DIAPH1, DAAM1 and DAAM2 were cloned from the cDNA purchased in the BD Biosciences Clontech (USA). Most constructs were cloned into the pGEX vector series and expressed in *E. coli* as GST-tagged proteins with the thrombin, PreScission protease, or enterokinase cleavage site. All constructs were cloned using standard protocols described in section 3.2.2. The list of the constructs of human DIAPH1, DAAM1 and DAAM2 is shown in Table 4.3.1 and 4.3.2, respectively.

Table 4.3.2. Results of cloning and expression of various DAAM1 and DAAM2 constructs

No.	Construct name	Primer	Vector	Expression in *E. coli*
1	DAAM1(594-1030)	20,23	pGEX-6P-1	high, very soluble
2	DAAM1(594-1078)	20,25	pGEX-6P-1	n.t.
3	DAAM1(527-1059)	21,59	pET41E/Lic	n.t.
4	DAAM1(542-1030)	22,23	pGEX-6P-1	n.t.
5	DAAM2(588-1020)	26,27	pGEX-6P-1	high, deg.
n.t. – not tested, deg. - degradation				

The Dictyostelium discoideum diaphanous-related formin-2 shows a high overall degree of similarity to DAAM1 (30% identity, 42% similarity) and DIAPH1 (28% identity, 40% similarity) (Schirenbeck et al., 2005). The general domain organization of dDia2 is similar to other formins, except for the localization of the DAD domain, which is located in the middle of the FH2 domain, between residues 770-815 (Schirenbeck et al., 2005). The FH1 domain is much shorter in comparison to other formins, and contains only two poly-proline stretches.

The schematic organization of domains responsible for actin nucleation and filament assembly is depicted in Figure 4.3.2. The constructs of dDia2 were cloned taking the dDia2(572-1087) as a template. Because of the problems with degradation and aggregation several mutants of the FH1FH2 or FH2 domains were made in order to find a construct suitable for crystallization and binding studies. All constructs and mutants of dDia2 are listed in Table 4.3.3.



NSTEPILGSPPPPPPPPMSGGGGPPPPPPPPGGKSNKPAKPIIKPSVKMRNFNWITIPALKVQGTFWDKLDE
TSFIQSLDKVELESLFSAKAPTVKVESKQLTRKVVVTVIDMKKANNCAIMLQHFKIPNEQLKKMQIMLDEKHF
SQENAIYLLQFAPTKEDIEAIKEYQGDQMQLGAAEQYMLTVMDIPKLDSRLKAFIFKQKFEGLVEDLVPDIKAI
KAASLELKKSKRLSDILKFILAIGNYVNGSTTRGGAFGFKVLETLPKMRDARSNDNKLSLLHFLAKTLQDRIPE
IWNIGAELPHIEHASEVSLNNIISDSSEIKRSIDLIERDFVPMINDPLFAHDKHWIHKITEFQKIAKVQYQRIEKEI
DEMNKAFEEITSYFGEPKSTQPDVFFSTINNFLEDLEKAYGEYQAMIRKAELENSKMEDPEKGGLQDLSSQI
RSGQLFKDRRVGDSVIAQMQNVDSLRKNLKSTSTTTPNTPPTIKIELPSQSILKPSGQLKK

Figure 4.3.2. Schematic domain organization and the sequence of dDia2. Domains and their sequences are shown in blue and yellow (FH1), red (FH2), and green (DAD). The N-terminal fragment of the FH2 domain responsible for solubility is marked by a box. Mutated residues of this fragment are shown in blue. The residues mutated in order to avoid degradation are in violet.

Table 4.3.3..Expressed constructs and mutants of dDia2

No.	Construct name	Primer	Vector	Expression in *E. coli*
1	dDia2 (572-1087)		pQE30, pGEX-4T-1	low, agg., deg., prec.
2	dDia2 (585-1053)	28,33	pQE30, pGEX-6P-1	low, agg., deg., prec.
3	dDia2 (602-1053)	29,33	pGEX-6P-1	low, agg., deg., prec.
4	dDia2 (585-1004)	28,34	pGEX-6P-1	high, deg.
5	dDia2 (602-1004)	29,34	pGEX-6P-1	high, deg.
6	dDia2 (616-1053)	30,33	pGEX-6P-1	high, deg.
7	dDia2 (616-1004)	30,34	pGEX-6P-1	high, deg.
8	dDia2 (619-1053)	31,33	pGEX-6P-1	high, deg.
9	dDia2 (619-1004)	31,34	pGEX-6P-1	high, deg.
10	dDia2 (636-1004)	32,34	pGEX-6P-1	high, insoluble
11	dDia2 (585-670)	28,35	pGEX-6P-1	high
12	dDia2 (602-670)	29,35	pGEX-6P-1	high
13	dDia2 (585-745)	28,36	pGEX-6P-1	high, agg.
14	dDia2 (602-745)	29,36	pGEX-6P-1	high, agg.

15	dDia2 (585-1004) R688G K689N	37,38	pGEX-6P-1	high
16	dDia2 (616-1004) R688G K689N	37,38	pGEX-6P-1	high
17	dDia2 (619-1004) R688G K689N	37,38	pGEX-6P-1	high
18	dDia2 (585-1004) R688G K689N K618G K621A K624N	37,38 39,40	pGEX-6P-1	high, low solubility
19	dDia2 (585-1004) R688G K689N K618G K621A K624N K632A R634G	37,38 39,40 41,42	pGEX-6P-1	high, insoluble
20	dDia2 (616-1004) R688G K689N L644K Q647K	37,38 43,44	pGEX-6P-1	high
21	dDia2 (619-1004) R688G K689N L644K Q647K	37,38 43,44	pGEX-6P-1	high
22	dDia2 (616-1004) R688G K689N L644K Q647K W638K	37,38 43,44 45,46	pGEX-6P-1	high, monomeric
23	dProf2dDia2 (619-1004)	47,48 49,34	pGEX-6P-1	high, deg.

agg. – aggregation; deg. – degradation, prec. - precipitation

Table 4.3.4. Results of cloning and expression of profilins

No.	Construct name	Primer	Vector	Expression in E. coli
1	human Profilin1	50,51	pGEX-4T-1	high
2	human Profilin2	52,53	pGEX-4T-1	high
3	dicty Profilin1		pGEX-4T-1 pET 22a	high
4	dicty Profilin1		pGEX-4T-1	high

Human profilin 1 and profilin 2 were cloned using the human cDNA as a template (BD Biosciences Clontech). Both the Dictyostelium discoideum profilin constructs were a gift from Dr Jan Faix. All expressed profilin constructs are listed in Table 4.3.4.

4.3.2 Expression and purification

Purification of formins and profilins were performed under native conditions and employed affinity chromatography (sepharose-glutathione) followed by protease cleavage and the final gel filtration. The general strategy for purification of formins and profilins is shown in Figure 4.3.3.

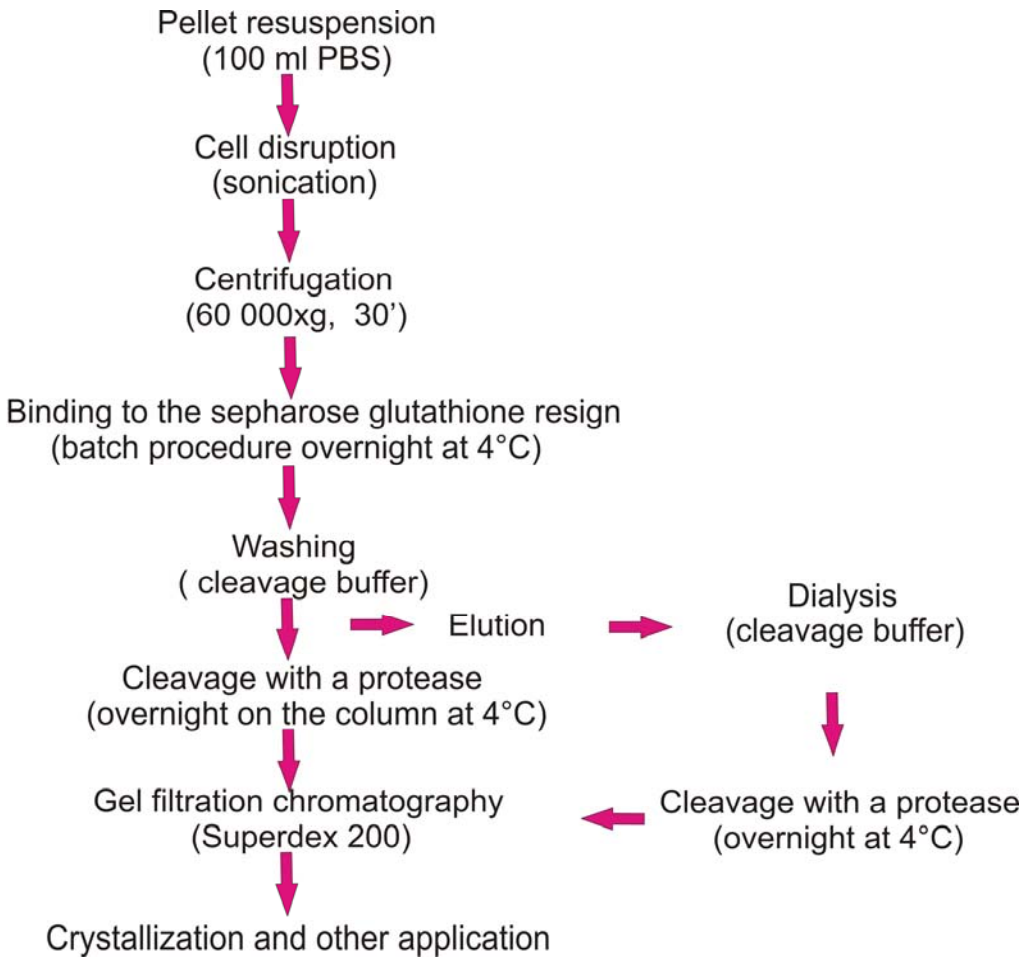


Figure 4.3.3. Flow chart of the purification scheme for formin and profilin constructs.

Theoretical isoelectric points, molecular masses, and extinction coefficients were estimated from the amino acid sequences of the studied proteins. Table 4.3.5 shows the basic physicochemical properties of the expressed constructs.

Table 4.3.5. Physicochemical properties of the studied formins and profilins

No.	Construct name	N.of aa	MW(Da)	E_{280} ($M^{-1}cm^{-1}$)	pI
1	DIAPH1(614-1134)	523	57780	21860	5.6
2	DAAM1(594-1030)	434	50161	22920	6.6
3	dDia2 (572-1087)	516	58049	33000	8.6
4	dDia2 (585-1053)	496	55935	33920	8.8
5	dDia2 (602-1053)	486	55014	33920	8.8
6	dDia2 (585-1004)	413	46788	33920	8.3
7	dDia2 (602-1004)	403	45867	33920	8.2
8	dDia2 (616-1053)	472	53791	33920	8.3
9	dDia2 (616-1004)	389	44643	33920	8.3
10	dDia2 (619-1053)	469	53548	33920	8.6
11	dDia2 (619-1004)	386	44401	33920	7.7
12	human Profilin1	140	15054	18575	8.4
13	human Profilin2	140	15046	21805	6.5
14	dicty Profilin1	126	13063	19940	6.1
15	dicty Profilin2	124	12729	18575	6.7

Affinity chromatography was the first step of the purification of formins and profilins. Presence of the GST-tag in all the expressed proteins enabled the use of sepharose-glutathione slurry under native conditions. The proteins obtained after this step of purification were usually pure, but in case of formins the samples were often a mixture of dimeric formins and aggregates. The formin dimer could be easily separated from aggregates by gel filtration. Removal of the

GST-tag depended on the size of the constructs. The GST-FH1FH2 and GST-profilins constructs were cleaved after elution from the sepharose-glutathione slurry, while the GST-FH2 constructs were cleaved on the sepharose–glutathione column. The final purification was achieved by gel filtration in the buffer proper for further applications of the protein.

4.3.2.1 Expression and purification of DIAPH1

The longest construct of the FH1FH2 domain (FH1FH2 (553-1200)) from human Diaphanous-related formin 1 were expressed in *E. coli* as inclusion bodies. Two other constructs (FH1FH2 (614-1134) and FH1FH2 (553-1134)) were produced in soluble fractions with very high yields but the purification was difficult because of strong aggregation and degradation.

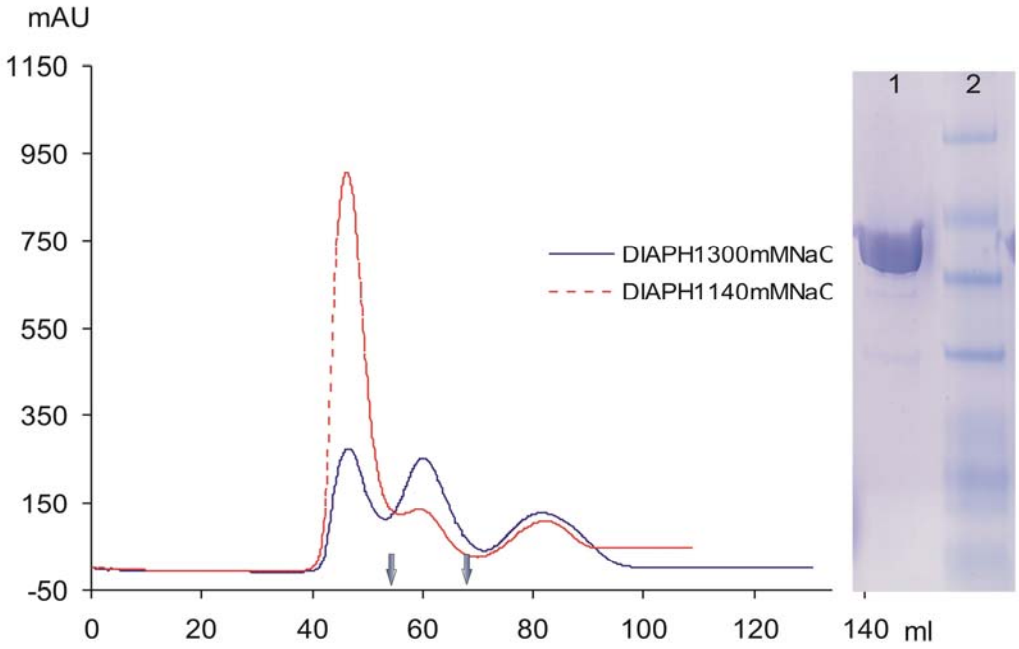


Figure 4.3.4. A typical chromatogram and the SDS-PAGE analysis of the DIAPH1FH1FH2 after gel filtration in buffer containing 300 mM (blue) and140 mM (red) NaCl. The fractions pooled are contained between the two arrows; SDS-PAGE (line 1, DIAPH1FH1FH2; line 2, marker).

In the PBS buffer the proteins were almost completely aggregated but increasing the concentration of NaCl up to 300 mM significantly decreased the aggregation and allowed the isolation of the dimeric formins. The gel filtration chromatogram and SDS-PAGE analysis are shown in Figure 4.3.4.

4.3.2.2 Expression and purification of DAAM1

Several constructs of FH1FH2 and FH2 of DAAM1 and DAAM2 were cloned but only one construct of DAAM1FH2 was used for further studies. The DAAM1FH2 domain was produced in *E.coli* in a soluble fraction with high efficiency. The pure FH2 dimer was obtained in flowthrough after cleavage of the GST-FH2 fusion protein bound to the sepharose-glutathione resign. The final gel filtration on Superdex 75 gave a pure and homogenous protein. The elution profile from gel filtration and SDS-PAGE analysis are shown in Figure 4.3.5.

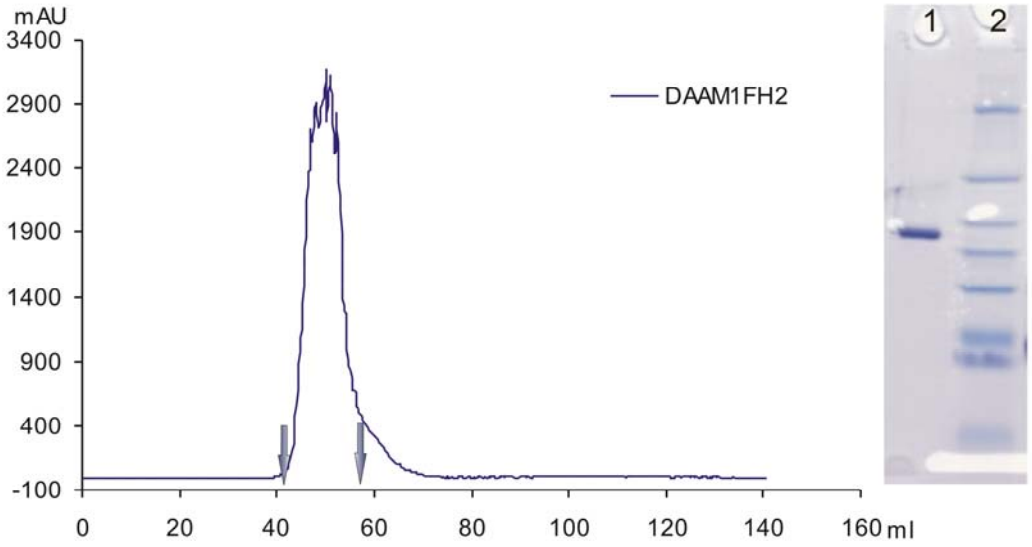


Figure 4.3.5. A typical chromatogram and the SDS-PAGE analysis of the DAAM1FH2 after gel filtration on Superdex 75 column. The fractions pooled are contained between the two arrows; SDS-PAGE (line 1, DAAM1FH2; line 2, marker).

4.3.2.3 Expression and purification of dDia2

Twenty constructs of dDia2 that differ in length were expressed in *E.coli* and purified. The dDia2(572-1087) construct containing three domains: FH1, FH2, and DAD, was used as a template for all other constructs and mutants. The yields of expression were very different depending on the construct. The expression level of the longest construct dDia2(572-1087) was low and the protein tended to aggregate and precipitate. The FH1FH2 constructs expressed well, however, the longer constructs with the C-terminal truncation at position 1053 were less soluble. The solubility of FH2 constructs rely on the length of the N-terminal region of the FH2 domain. The FH2(636-1004) was totally insoluble, most probably because of the lack of the lysine-rich N-terminal fragment (GKSNKPAKPIIKPSVKMRN). All other constructs of the FH1FH2(585-1053) and FH2(616-1053) that contained this fragment were soluble with the exception of mutants shown in Table 4.3.3. Mutations in five positions: K618G, K621A, K624N, and R688G, K689N, significantly decreased solubility, while the substitution of seven positively charged amino acids in positions K618G, K621A, K624N, K632A, R634G, and R688G, K689N made the protein insoluble. The second problem that occurred during purification was the degradation. Limited proteolysis and the N-terminal Edman sequencing of degradation products revealed that all wild-type constructs degraded between the R688 and K689 in the sequence: R▼KVVV. Mutations in these positions (R688G, K689N) made the proteins more stable and resistant for the protease cleavage. Four constructs of the FH1 domain, containing one and two poly-proline stretches, were expressed in order to investigate the interactions with profilins by means of NMR. The expression yield of FH1 constructs was high but only shorter constructs: FH1(602-670) and FH1(585-670) were soluble. All dDia2 constructs were expressed as GST-fusion proteins with a cleavage site for the PreScission protease. GST-FH1FH2 constructs were cleaved after elution from the resign, while all the FH2 constructs were cleaved on the column. A typical elution profile and the SDS-PAGE analysis of the isolated FH1FH2 are shown in Figure 4.3.6.

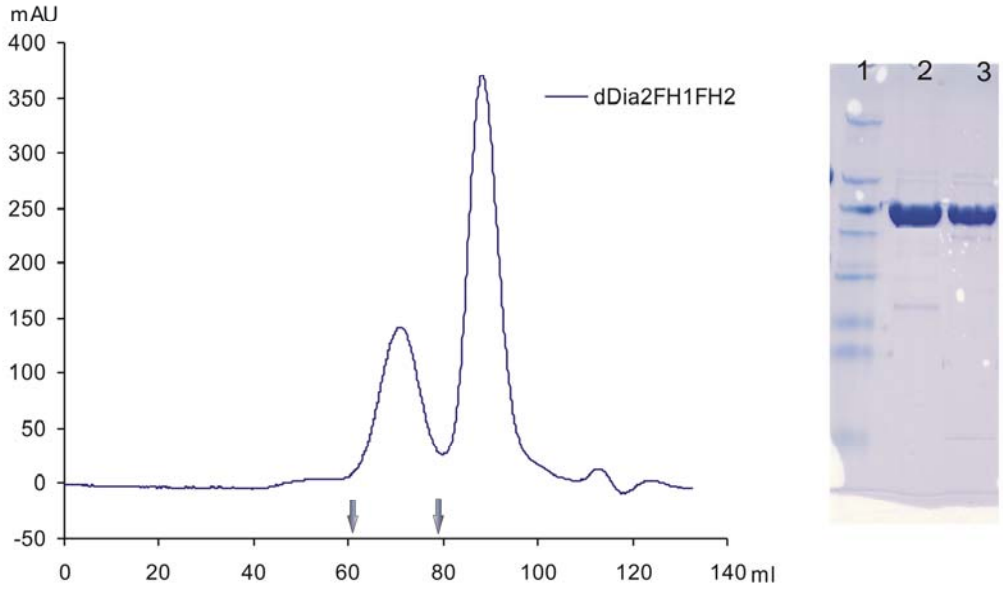


Figure 4.3.6. A typical chromatogram and the SDS-PAGE analysis of the dDia2FH1FH2 after gel filtration on Superdex 200 column. The fractions pooled are contained between the two arrows; SDS-PAGE (line 1, marker; line 2 and 3, dDia2FH1FH2).

4.3.2.4 Expression and purification of profilins

Both human and Dictyostelium profilins 1 and 2 were expressed with a very high yield, even in minimal media. Proteins were produced in *E.coli* as GST-fusion proteins with the thrombin cleavage site. One of the constructs of Dictyostelium profilin 2 was expressed with the C-terminal His-tag. Purified profilins were soluble and stable in all used buffers. The final purification step, which allowed removing of the GST and thrombin, was achieved by gel filtration on Superdex 75 (Figure 4.3.7).

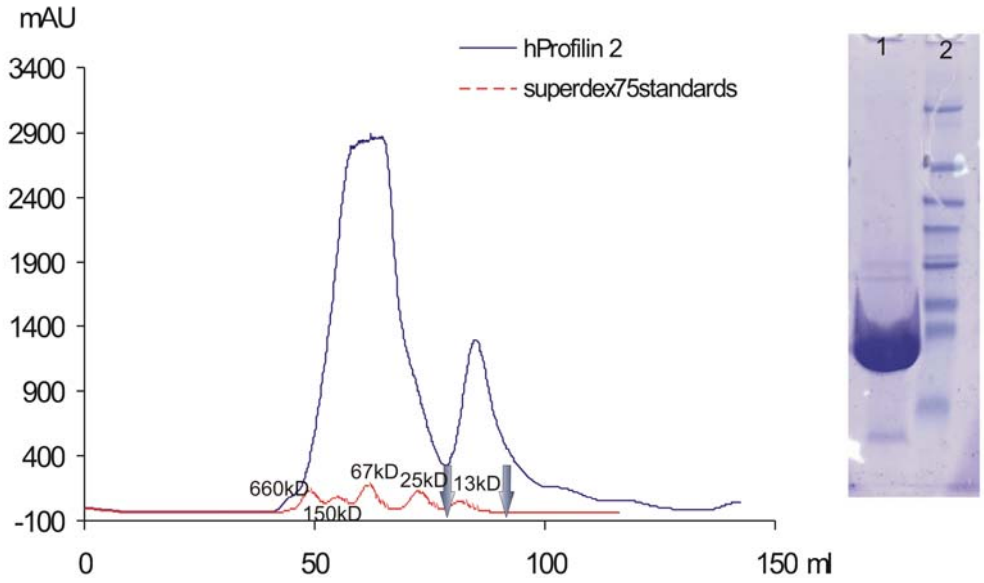


Figure 4.3.7. A typical chromatogram and the SDS-PAGE analysis of the human profilin 2 after gel filtration on Superdex 75 column. The fractions pooled are contained between the two arrows; SDS-PAGE (line 1, profilin 2; line 2, marker).

4.3.3 Functional and structural studies of formins and profilins

4.3.3.1 NMR analyses

All purified formins and profilins were examined by NMR in order to assess their structural integrity necessary for construct optimization and binding studies. The unique strength of NMR lies in its capability to estimate unstructured regions of the polypeptide chains in the partially folded proteins and to identify proteins that are heterogeneous because of aggregation or other conformational effects (Rehm et al., 2002). The signal dispersion and line shape of the resonances is dependent on the folding. The line widths of the signal in any NMR spectrum are correlated to the size of the molecule. The NMR signal of large biomolecules decays much faster than that of smaller ones which produces broader NMR signals. Figure 4.3.8 and Figure 4.3.9 show one-dimensional proton NMR spectra of several constructs of formins and profilins. All profilins

showed the signal dispersion downfield of 8.5 ppm and upfield of 1 ppm characteristic of structured proteins.

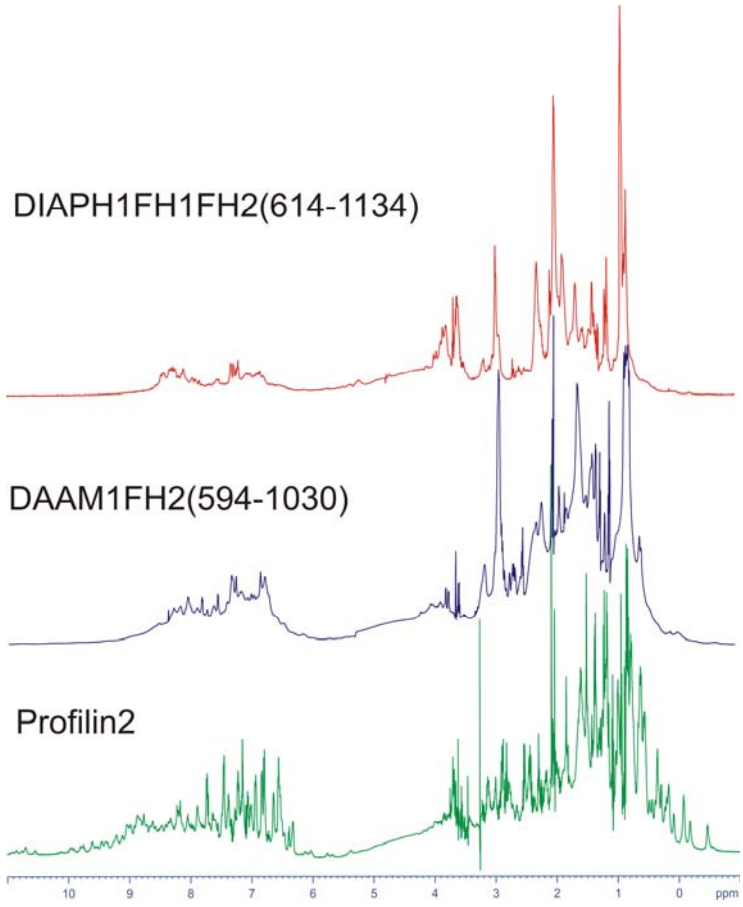


Figure 4.3.8. Characterization of the human formin and profilin structures by one-dimensional proton NMR spectra.

Formins yielded rather poor NMR spectra, most probably because of large size and dimerization. The quality of the one-dimensional spectrum was dependent on the length of a construct. The longest constructs of dDia2FH1FH2 gave large and broad signals at approximately 8.3 ppm and showed a small dispersion of the amide backbone chemical shift what suggested unfolding and/or aggregation. Spectra of dDia2FH1FH2(602-1004) and dDia2FH2(619-1004) were 'better' but indicated that these proteins were only partially folded and contained large unstructured fragments. Based on these spectra it could be concluded that not only the FH1 domain but also some parts of the FH2 domain

were unstructured. The spectrum of the monomeric FH2(690-1004) 'core' domain obtained after limited proteolysis revealed that this unstructured region was located in the N-terminal of the FH2 domain. Despite the not optimal features of NMR spectra of formins, further structural studies, including crystallization, were undertaken.

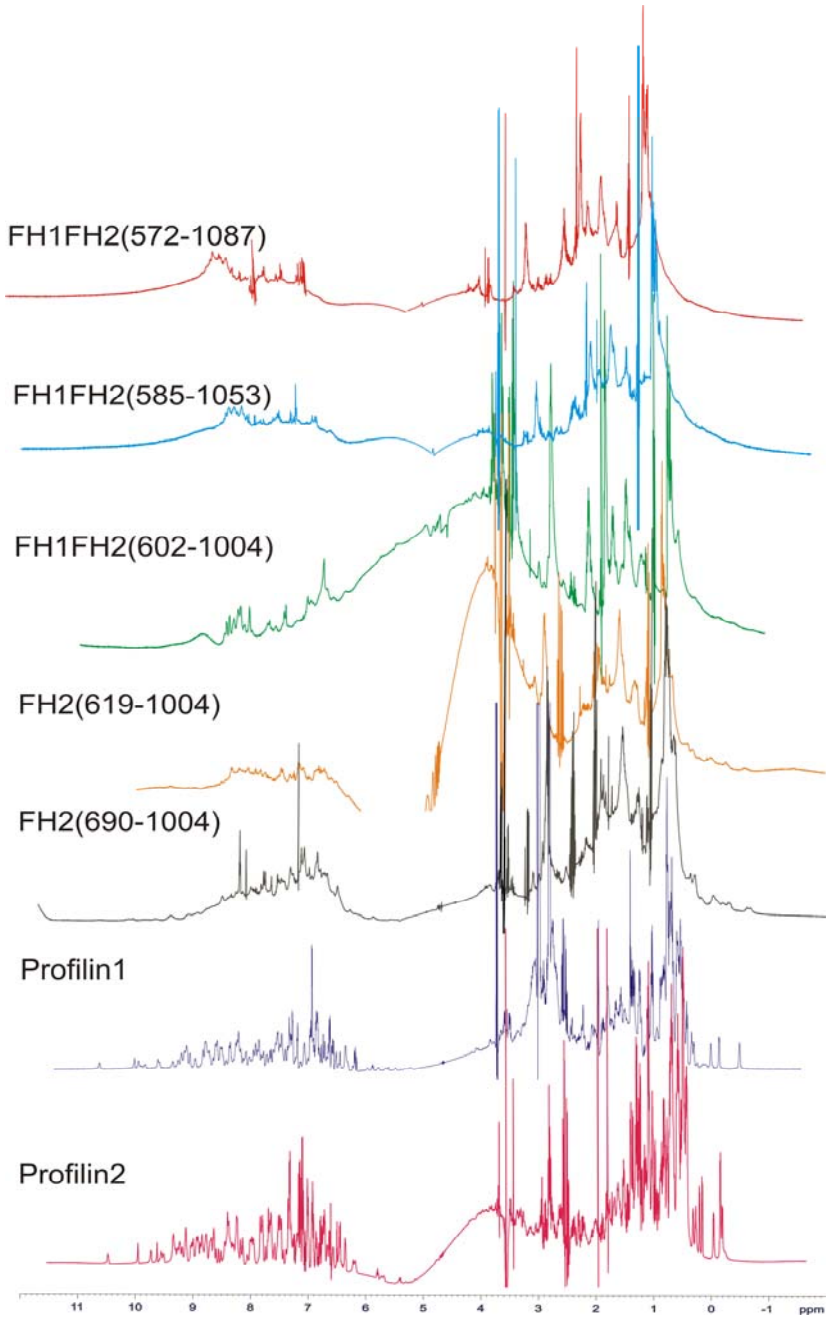


Figure 4.3.9. Characterization of Dictyostelium formins and profilins by one-dimensional proton NMR spectrum.

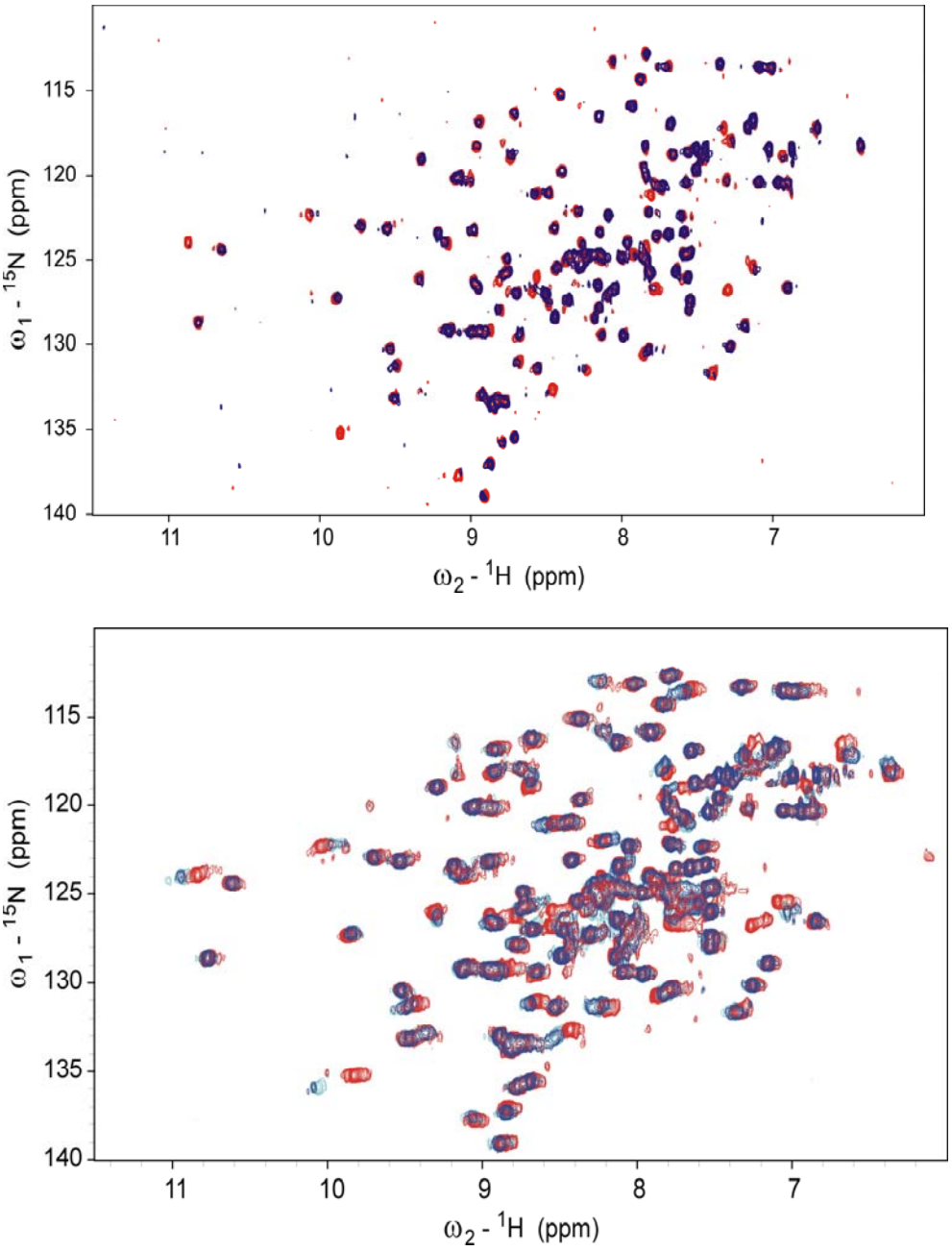


Figure 4.3.10. Titration of the ^{15}N-labeled human profilin 2 with the FH1 domain containing two different numbers of poly-proline stretches. Upper plot – titration with the construct containing nine possible binding sites (GGTAISPPPPLSGDATI PPPPPLPEGVGIPSPSSLPGGTAIPPPPPLPGSARIPPPPPPLPGSAGIPPPPPP LPGEAGMPPPPPPLPGGPGIPPPPPFPGPGIPPPPPGMGMPPPPPFG). Lower plot – titration with the construct containing one binding site (MSGGGGPP PPPPPPGG). Both plots show the ^{15}N-HSQC spectra of profilin 2 before (red), and after (blue) addition of the FH1 domain.

Structural as well as binding properties of proteins can be investigated by two-dimensional heteronuclear single-quantum coherence experiment (HSQC). For the ^{15}N-HSQC spectrum, a ^{15}N-labeled protein is required. The HSQC shows one NMR peak for every proton bound directly to a nitrogen atom and thus exactly one signal per residue in the protein (Rehm et al., 2002). Monitoring changes in chemical shifts and line widths of amide resonances of a ^{15}N-labeled protein upon addition of a unlabeled protein allows to detect and characterize possible protein-protein binding. Because of a weak binding between formins and profilins NMR was employed to study these interactions. Figure 4.3.10. shows the titration of ^{15}N-labeled human profilin 2 with unlabeled FH1FH2 domains with various number of poly-proline stretches. The constructs of the FH1FH2 domain used for titrations contained nine (upper plot) and one (lower plot) potential binding sites. In both cases, the same peaks in the ^{15}N-profilin spectra were affected upon binding to the FH1 domain. Unfortunately, the residues involved in this interaction could not be identify because of the lack of the assignment of the profilin 2 spectrum. The longer construct of the FH1 domain saturated binding sites in profilin after one step of titration while the short construct in the same concentration saturated profilin after several steps of titration. This experiment proved that more than one profilin might interact with the FH1 domain at the same time. The NMR experiments confirmed the results of the previous studies (mobility shift assay) suggesting weak type of binding between formins and profilins.

4.3.3.2 Pyrene-actin assays

Pyrene-actin assays were carried out in order to check nucleation and polymerization activity of the purified constructs of formins and profilins. Actin filament assembly assays were performed in a buffer containing 10 mM imidazole, 2 mM $MgCl_2$, 0.2 mM $CaCl_2$, 1 mM Na-ATP, and 50 mM KCl, pH 7.2, as described (Eichinger and Schleicher, 1992). Actin polymerization was measured by fluorescence spectroscopy (Figure 4.3.11).

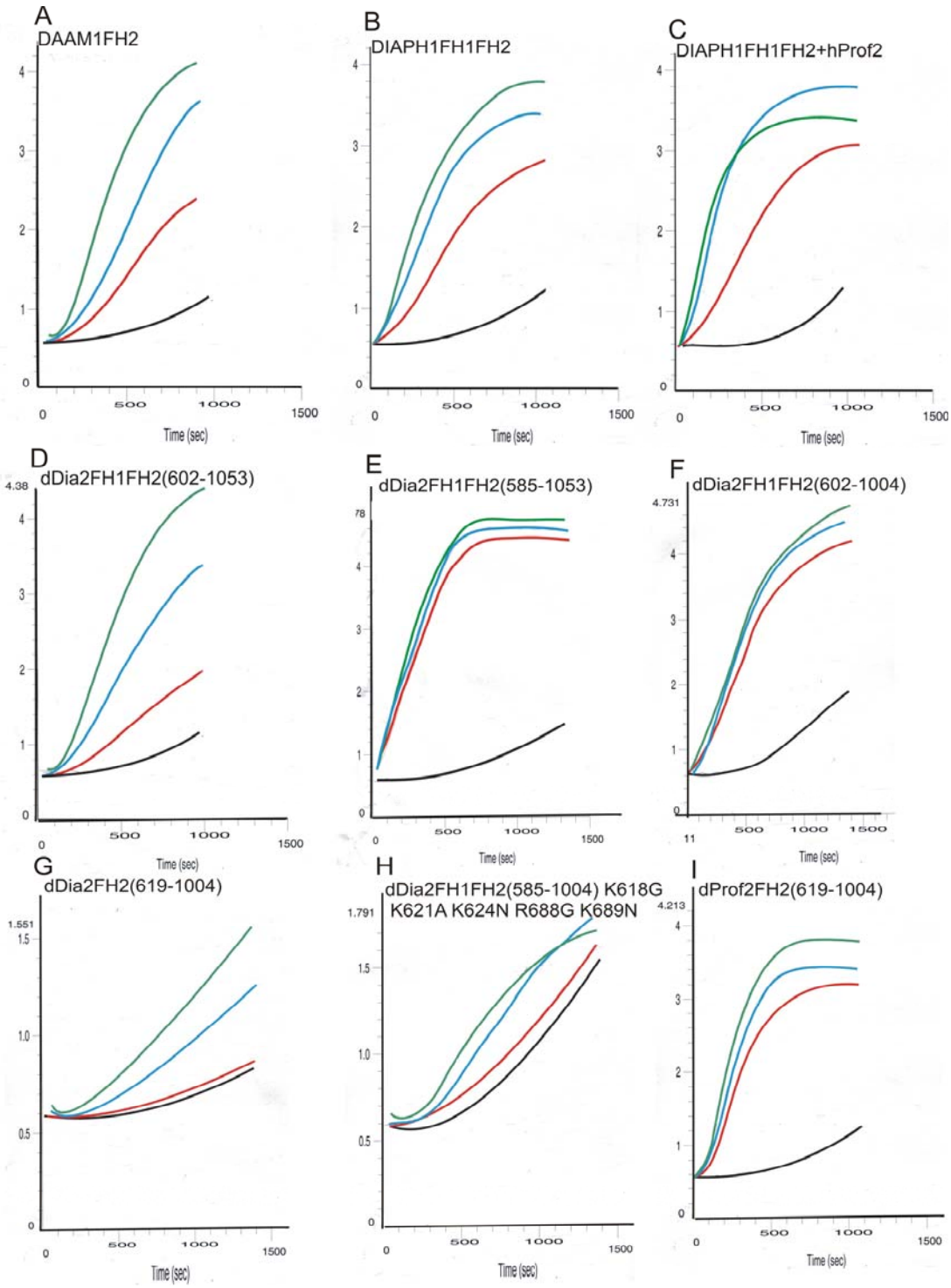


Figure 4.3.11. Interaction of formin constructs with actin. Actin (1.8 μM; 10% pyrene-labeled) was polymerized in the presence of 0 nM (black), 50 nM (red), 100 nM (blue) and 250 nM (green) formin. (C) – interaction of 50 nM DIAPH1FH1FH2 with actin in the presence of 0 nM (red), 50 nM (blue), 100 nM (green) profilin 2.

These experiments showed that the following constructs nucleate actin polymerization *in vitro*: DIAPH1FH1FH2(614-1134), DAAM1FH2(594-1030), and dDia2FH1FH2(585-1053), dDia2FH1FH2(602-1053), dDia2FH1FH2(585-1004), and dDia2FH1FH2(602-1004). Nanomolar concentrations of these constructs significantly increased the rate of spontaneous polymerization. Figure 4.3.11. C shows the influence of profilin 2 on the DIAPH1FH1FH2-induced polymerization. Addition of profilin 2 increased polymerization rate but there was no significant difference between 50 nM and 100 nM concentrations of profilin. Surprisingly, the dDia2FH2(619-1004) and all the dDiaFH1FH2 constructs containing mutations in positions K618G, K621A, K624N, R688G, K689N did not show any interactions with actin. As there was no obvious reason for the lack of the nucleation and polymerization activity for the dDia2FH2 construct, I designed and expressed the fusion protein containing Dictyostelium profilin 2 and 'non active' dDia2FH2 (619-1004) domain in order to check whether this FH2 construct was able to polymerize actin. The pyrene-actin assay showed that this construct was active. This suggested that the FH2 constructs might elongate actin filament but did not have nucleation activity.

4.3.4 Crystallization

The formin proteins used for crystallization trials were checked by mass spectrometry, N-terminal sequencing, and NMR to confirm their purities, homogeneities and folding. An extensive number of crystallization trials was carried out during this thesis in order to get the crystals of FH1FH2 or FH2 domains, formin-profilin complexes (FH1FH2-profilin or FH1-profilin), formin-actin complexes (FH2-actin) or formin-profilin-actin complexes (FH1FH2-profilin-actin). Four homologous formins were chosen for crystallization attempts. Among of the large number of formin constructs, only two gave nice but unfortunately poorly diffracting crystals. Crystals of DAAM1FH2(594-1030) grew in several conditions after a few days at room temperature and at 4°C (Table 4.3.6, Figure 4.3.12). Optimization of initial conditions did not improved the quality of the crystals.

Table 4.3.6. Crystallization conditions for dDia2FH2 and DAAM1FH2

No.	Protein	Crystallization conditions
1	DAAM1FH2 (594-1030)	**Index 19** 0.056 M Sodium phosphate monobasic monohydrate, 1.344 M Potassium phosphate dibasic pH 8.2 **Index 25** 3.5 M Sodium formate pH 7.0 **Index 46** 0.1 M Bis-Tris pH 6.5, 20% w/v Polyethylene glycolmonomethyl ether 5000 **Index 87** 0.2 M Malonate pH 7.0, 20 % w/v Polyethylene glycol 3350 **Index 88** 0.2 M tri-Ammonium citrate pH 7.0, 20 % w/v Polyethylene glycol 3350 **Index 94** 0.2 M Sodium citrate, 20 % w/v Polyethylene glycol 3350
2	dDia2FH2(619-1004) L644K Q647K R688G K689N	**Index 44** 0.1 M HEPES pH 7.5; 25% w/v Polyethylene glycol 3350

Crystallization of dDia2 was a very laborious task because more than 700 crystallization conditions were unsuccessfully tested for each construct. The main problems which became apparent during crystallization were: degradation and precipitation caused by low solubility of the proteins. The problem of degradation was resolved by introducing mutations in the positions R688G and K689N. The solubility of the FH2 construct was significantly improved by two substitutions: L644K and Q647K, which were designed after a sequence alignment to the previously crystallized yeast Bni1FH2 (Xu et al., 2004). The mutated dDia2FH2 (619-1004) L644K Q647K R688G K689N was used for further crystallization. Crystals of this mutant grew after two months at 4°C in 0.1 M HEPES pH 7.5; 25% w/v Polyethylene glycol 3350 (Figure 4.3.12) but unfortunately did not diffract below 6 Å resolution.

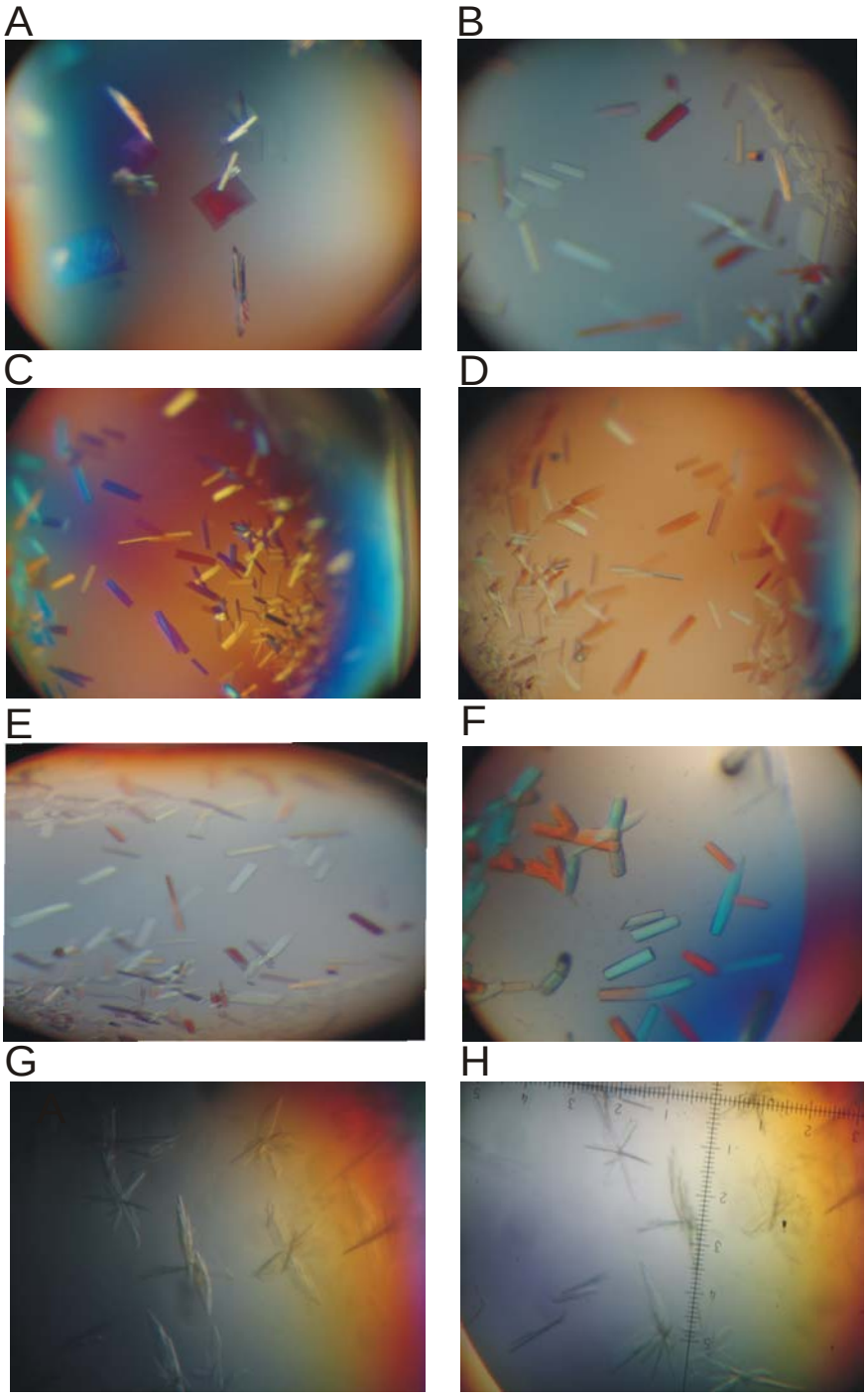


Figure 4.3.12. Crystals of formins grown under different conditions. (A-F) Crystals of the DAAM1FH2 (594-1030) grown in the INDEX 19 (A), INDEX 25 (B), INDEX 46 (C), INDEX 87 (D), INDEX 88 (E), INDEX 94 (F); Crystals of the dDia2 (619-1004) L644K Q647K R688G K689N grown in the INDEX 44.

Crystals were measured on a synchrotron source in DESY Hamburg. All other mutants designed according to the protocol for mutational surface engineering (Derewenda, 2004) did not work because of low solubility. Crystallization of the complexes of formin-profilin, formin-actin and formin-profilin-actin were unsuccessful most probably because of weak binding between these proteins. Despite of the fact that most of the initial crystallization trials have been unsuccessful so far, the project is still continued.

4.4 Discussion

The Dictyostelium discoideum diaphanous-related formin 2 (dDia2) shows high sequence homology with three human formins (DAAM1, DAAM2, DIAPH1), as well as with the yeast formin Bni1p (Figure 4.4.1). The formin homology 2 (FH2) domain, which is responsible for the formin-mediated actin assembly, is one of the most conserved domains among the formin family. The high sequence homology suggests that the 3D structure of the FH2 domain would also be conserved, however, the structure of only one dimeric formin is known so far. Studies carried out in this thesis suggest that despite of the sequence and structural similarities of various formins, their biochemical properties might be significantly different. Pyrene-actin assays showed that the dDia2FH2(619-1004) domain does not have any nucleation activity while DAAM1FH2(594-1030) meaningfully increases the polymerization rate. The other constructs of dDia2 with the longer N-terminus, which contained a part or full FH1 domain (dDia2FH1FH2(602-1004), dDia2FH1FH2(585-1004), respectively), are active. The fusion protein, comprising the Dictyostelium profilin 2 and dDia2FH2(619-1004), also increases polymerization rate (Figure 4.3.11). Taken together, these data suggest that dDia2FH2(619-1004) elongates actin filaments but alone cannot initiate the nucleation event. Mutational studies on dDia2FH1FH2(585-1004) show that mutated residues (K618G, K621A, K624N, R688G, K689N) may be involved in the formin-actin interaction.

```
dDia2  ----------------------------------------NSTEPILGSPP-------
DAAM1  ------------------------------------------------------------
DAAM2  ----------------------------------------------SSPPPPGGP--
DIAPH1 MASLSAAAITVPPSVPSRAPVPPAPPLPGDSGTIIPPPPAPGDSTTPPPPPPPPPPPPL
Bni1   -------------------------------------SSVLSSQPPPPPPPPPPVPAK

dDia2  ------------------------------------------------------PPPPPP
DAAM1  ----------------------------SSVPGSL------------LPPPPPPP
DAAM2  ------------------------LTL----SSSMTTND-------------LPPPPPP
DIAPH1 PGGTAISPPPPLSGDATIPPPPPLPEGVGIPSPSSLPGGTAIPPPPPLPGSARIPPPPPP
Bni1   LFGESLEKEKK----------SEDDTV----KQETTGDSPA------------PPPPPP
                                                          ******

dDia2  MSGGG---------------GPPPPPPPP--G---GKSNK-PAKP---------------
DAAM1  LPGGM-LPPPPPPLPP---GGPPPPPGPPP-L---GAIMPPPGAPM-------------
DAAM2  LPFACCPPPPPPPLPP---GGPPTPPGAPPCL---GMGLPLPQDPYPSS----------
DIAPH1 LPGSAGIPPPPPPLPGEA-GMPPPPPPLP--G---GPGIP-PPPPFPGGPGIPPPPPGMG
Bni1   P-----PPPPPMALFGKPKGETPPPPPLPSVLSSSTDGVIPPAPPMM------------
                          .*.**  *           *  *

dDia2  -------------------------IIKPSVKMRNFNWITI-P--ALKVQGTFWDKLDET
DAAM1  -----------------GLALKKKSIPQPTNALKSFNWSKL-P--ENKLEGTVWTEIDDT
DAAM2  -----------------DVPLRKKRVPQPSHPLKSFNWVKL-N--EERVPGTVWNEIDDM
DIAPH1 MPPPPPFGFGVPAAPVLPFGLTPKKLYKPEVQLRRPNWSKLVA--EDLSQDCFWTKVKED
Bni1   PASQIKSAVTSPLLPQSPSLF--EKYPRPHKKLKQLHWEKL-DCTDNSIWGTGKAEKFAD
                                  :*   ::  :* .:         .    :

dDia2  SFIQSLDKVELESLFSAKAPT----------VK-----VESKQLTRKVV--VTVIDMKKA
DAAM1  KVFKILDLEDLERTFSAYQRQQDFFVNSNSKQKEA-DAIDDTLSSKLKVKELSVIDGRRA
DAAM2  QVFRILDLEDFEKMFSAYQRH----------QKEL-GSTEDIYLASRKVKELSVIDGRRA
DIAPH1 RFENNELFAKLTLTFSAQTKT----------KKDQEGGEEKKSVQKKKVKELKVLDSKTA
Bni1   DLYEKGVLADLEKAFAAREIKSLASKRKEDL---QKITFLSRDISQQFGINLHMYSSLSV
        . .     .:   *:*                        .         : : .   .

dDia2  NNCAIMLQHFKIPNEQLKKM--QIMLDEKH-FSQENAIYLLQFAPTKEDIEAIKEYQ---
DAAM1  QNCNILLSRLKLSNDEIKRA--ILTMDEQEDLPKDMLEQLLKFVPEKSDIDLLEEHK---
DAAM2  QNCIILLSKLKLSNEEIRQA--ILKMDEQEDLAKDMLEQLLKFIPEKSDIDLLEEHK---
DIAPH1 QNLSIFLGSFRMPYQEIKNV--ILEVNEAV-LTESMIQNLIKQMPEPEQLKMLSELK---
Bni1   ADLVKKILNCDRDFLQTPSVVEFLSKSEIIEVSVNLARNYAPYSTDWEGVRNLEDAKPPE
        :    :        :      :  .*   .. .         .  . :  :.: :

dDia2  GDQMQLGAAEQYMLTV--MDIPKLDSRLKAFIFKQKFEGLVEDLVPDIKAIKAASLELKK
DAAM1  HELDRMAKADRFLFEM--SRINHYQQRLQSLYFKKKFAERVAEVKPKVEAIRSGSEEVFR
DAAM2  HEIERMARADRFLYEM--SRIDHYQQRLQALFFKKKFQERLAEAKPKVEAILLASRELVR
DIAPH1 DEYDDLAESEQFGVVM--GTVPRLRPRLNAILFKLQFSEQVENIKPEIVSVTAACEELRK
Bni1   KDPNDLQRADQIYLQLMVNLESYWGSRMRALTVVTSYEREYNELLAKLRKVDKAVSALQE
        :   :  :::    :          *:.:: .  .:     :  ..:  :  .   : .

dDia2  SKRLSDILKFILAIGNYVNGSTTRGGAFGFKVLETLPKMRDARSN-DNKLSLLHFLAKTL
DAAM1  SGALKQLLEVVLAFGNYMNKGQRG-NAYGFKI-SSLNKIADTKSSIDKNITLLHYLITIV
DAAM2  SKRLRQMLEVILAIGNFMNKGQRG-GAYGFRV-ASLNKIADTKSSIDRNISLLHYLIMIL
DIAPH1 SESFSNLLEITLLVGNYMNAGSRNAGAFGFNI-SFLCKLRDTKST-DQKMTLLHFLAELC
Bni1   SDNLRNVFNVILAVGNFMNDTSKQ--AQGFKL-STLQRLTFIKDT-TNSMTFLNYVEKIV
       *  : ::::. * .**::*       * **.:   * ::   :..  ..:::*:::
```

```
dDia2  QDRIPEIWNIGAELPHIEHASEVSLNNIISDSSEIKRSIDLIERDFVP---MINDPLFAH
DAAM1  ENKYPSVLNLNEELRDIPQAAKVNMTELDKEISTLRSGLKAVETELEY---QKSQPPQPG
DAAM2  EKHFPDILNMPSELQHLPEAAKVNLAELEKEVGNLRRGLRAVEVELEY---QRRQVREPS
DIAPH1 ENDYPDVLKFPDELAHVEKASRVSAENLQKNLDQMKKQISDVERDVQ----NFPAATDEK
Bni1   RLNYPSFNDFLSELEPVLDVVKVSIEQLVNDCKDFSQSIVNVERSVEIGNLSDSSKFHPL
       .   *.. .:  **  : .. .*.  :: .:   :   :  :* ..

dDia2  DKHWIHKITEFQKIAKVQYQRIEKEIDEMNKAFEEITSYFGEPKS-T-QPDVFFSTINNF
DAAM1  DK-FVSVVSQFITVASFSFSDVEDLLAEAKDLFTKAVKHFGEEAG-KIQPDEFFGIFDQF
DAAM2  DK-FVPVMSDFITVSSFSFSELEDQLNEARDKFAKALMHFGEHDS-KMQPDEFFGIFDTF
DIAPH1 DK-FVEKMTSFVKDAQEQYNKLRMMHSNMETLYKELGEYFLFDPK-KLSVEEFFMDLHNF
Bni1   DK-VLIKTLPVLPEARKKGDLLEDEVKLTIMEFESLMHTYGEDSGDKFAKISFFKKFADF
       **  :     .   :  . . :.         : .    :      .     **  :  *


dDia2  LEDLEKAYGEYQAMIRKAELENSK-------MEDP--------------------EK--
DAAM1  LQAVSEAKQENENMRKKKEEEERRARMEAQLKEQRERER--KMRKAK-----ENSEES--
DAAM2  LQAFSEARQDLEAMRRRKEEEERRARMEAMLKEQRERER--WQRQRKVLAAGSSLEEG--
DIAPH1 RNMFLQAVKENQK-RRETEEKMRRAKLAKEKAEKERLEK--QQKREQLIDMNAEGDET--
Bni1   INEYKKAQAQNLAA---EEEERLYIKHKKIVEEQQKRAQEKEKQKENSNSPSSEGNEEDE
        :   :*  :        * :           *.                      *

dDia2  -----GGLQDLSSQIRSGQLFKDRR-----------------------------------
DAAM1  -----GEFDDLVSALRSGEVFDKDLSKLKRNRKR-------------------------
DAAM2  -----GEFDDLVSALRSGEVFDKDLCKLKRSRKR-------------------------
DIAPH1 -----GVMDSLLEALQSGAAFRRKRGPRQANRKA-----------------------GC
Bni1   AEDRRAVMDKLLEQLKNA--GPAKSDPSSARKRALVRKKYLSEKDNAPQLLNDLDTEEGS
            . ::.* . ::..

dDia2  VGDSVIAQMQNVDSLRKN----------LKSTSTTTPNTPPTIKIELPSQSILKPSGQLK
DAAM1  --ITNQMTDSSRERPITK----------LNF---------------------------
DAAM2  --SGSQALEVTRERAINR----------LNY---------------------------
DIAPH1 AVTSLLASELTKDDAMAA----------VPAKVSKNSETFPTI-LEEAKELVGRAS----
Bni1   ILYSPEAMDPTADTVIHAESPTPLATRGVMNTSEDLPSPSKTSALEDQEEISDRARMLLK
                  . :              :

dDia2  K---------------------------------------
DAAM1  ----------------------------------------
DAAM2  ----------------------------------------
DIAPH1 ----------------------------------------
Bni1   ELRGSDTPVKQNSILDEHLEKLRARKERSIGEASTGNRLSFK
```

Figure 4.4.1. Sequence alignment of dDia2, DAAM1, DAAM2, DIAPH1, and Bni1. Domains are coloured blue (FH1), red (FH2) and green (DAD). dDia2 has two regions, which might be identified as DAD domains. Conserved residues responsible for dimerization are indicated in dark yellow. Residues of Bni1 involved in interactions with actin are coloured grey.

This can explain the lack of the nucleation and polymerization activities of these constructs (Figure 4.3.11). The FH1 domain and the N-terminal fragment of the FH2 domain, which correspond to the lasso, linker and knob domains of Bni1, are necessary for the full activity of the dDia2 formin (Figure 4.4.2). In contrast to dDia2FH2, DAAM1FH2(594-1030) does not need the FH1 domain for nucleation and polymerization activities. Comparison of amino acid sequences of dDia2FH2(619-1004) and DAAM1FH2(594-1030) in the region containing the lasso, linker and knob subdomains, shows high identity but DAAM1FH2 is 16 amino acids longer (26 up to first poly-proline stretch in the FH1 domain), what might significantly affect the activity, especially that these additional residues are located in the interface which interacts with actin molecules (Figure 4.4.2).

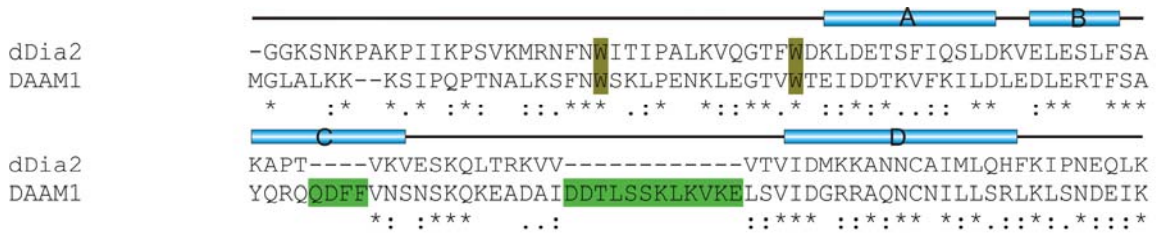


Figure 4.4.2. Sequence alignment of N-terminal subdomains of dDia2FH2 and DAAM1FH2 and the secondary structure of the Bni1FH2. Secondary structure elements of Bni1FH2 are shown above the sequences, with cylinders representing helices and thin lines indicating unstructured regions (helices A and B belong to the lasso, helix C forms the linker, helix D belongs to the knob subdomain). Conserved tryptophans are coloured brown; differences in the primary sequence are shown in green.

The crystal structure of Bni1FH2 complexed with TMR-actin showed that each FH2 domain has two actin-binding sites: the knob and lasso/post. Binding of the knob is anchored by insertion of the D helix of the FH2 domain into the cleft between actin subdomains 1 and 3. The second FH2-actin interface involves a composite surface formed by elements from both the lasso and post region of the FH2 domain. Interactions between FH2 and actin are mediated by highly polar, basic surfaces in the lasso/post binding site (Xu et al., 2004; Otomo et al.,

2005). High homology of formins suggests that the molecular bases for the FH2-actin interaction are similar for all members of the formin family. Investigations of dDia2FH1FH2 show that the FH1 domain and the basic region between the last poly-proline stretch and the first conserved tryptophan in the FH2 domain might be involved in actin nucleation.

Stretches of poly-prolines in the FH1 domain are believed to recruit profilin-actin from a large cytoplasmatic reservoir for addition to the barbed end. According to the *in vitro* NMR studies, the profilin-formin binding is weak but the pyrene-actin assays of DIAPH1FH1FH2 and dDia2FH1FH2 show that profilin significantly increases polymerization rates (Figure 4.3.11 C). The profilin binding sites are likely to differ in affinity because both the nonproline composition and the contiguity of prolines vary (Figure 4.4.1). Proline hexamers and octamers bind profilin much more weakly than decamers, whereas substitution of proline residues by glycine or alanine destabilizes binding (Petrella et al., 1996; Perelroizen et al., 1994). The ^{15}N-HSQC experiment (Chapter 4.3.3) shows that the FH1 domain with multiple poly-proline stretches is able to bind several profilins simultaneously, and this binding influences the actin polymerization rate (Vavylonis et al., 2006; Kovar et al., 2006; Kovar and Pollard, 2004). The fusion protein of the ‘non active’ dDia2FH2(619-1004) and profilin 2 provides evidence that profilin plays an important role during actin nucleation.

The crystal structures of Bni1pFH2 and Bni1pFH2 in complex with TMR-actin (Xu et al., 2004; Otomo et al., 2005) which have been resolved recently shed light on the molecular mechanism of actin assembly, but still a lot of questions have to be elucidated. The first question is as follows: do all formins share the same mechanism of actin nucleation and polymerization? So far, the mechanism for actin assembly seems to be very conserved, but it has been recently reported that the *Arabidopsis* Formin 1 is a nonprocessive formin that moves from the barbed end to the side of an actin filament after the nucleation event (Michelot et al., 2006). The other important question is whether the full-length formins as compared to isolated domains retain the same molecular characteristics, especially in a cellular environment. The structural arrangement

of the ‘open’ and ‘closed’ states of the FH2 domain on the barbed end of an actin filament is also not clearly understood.

5 Summary

The focus of this thesis was on the biophysical and biochemical characterization of two groups of proteins: insulin-like growth factor binding proteins (IGFBPs) and formins.

Insulin-like growth factor binding proteins (IGFBPs) control bioavailability, activity and distribution of IGF-1 and -2 through high-affinity IGFBP-IGF complexes. IGF binding sites are found on the conserved amino- and carboxyl-terminal domains of IGFBPs. Their relative contributions to IGFBP-IGF complexation have been difficult to analyze, in part, because of the lack of appropriate three-dimensional structures. To analyze the effects of N- and C-terminal domain interactions, we determined several X-ray structures, first, of a ternary complex of N- and C-terminal domain fragments of IGFBP-4 and IGF-1 and second, of a "hybrid" ternary complex using the C-terminal domain fragment of IGFBP-1 instead of IGFBP-4. We also solved the binary complex of the N-terminal domains of IGFBP-4 and IGF-1, again to analyze C- and N-terminal domain interactions by comparison with the ternary complexes. These crystal structures provide the molecular basis for the IGFBPs regulation of IGF signaling and support research into the design of IGFBP variants as therapeutic IGF inhibitors for diseases of IGF disregulation. Key features of the IGFBP-4/IGF-1 complexes include: 1) a disulphide bond ladder in the N-terminal domain and the first 5 N-terminal “thumb” residues, which bind to IGF-1 and partially mask the IGF residues responsible for the type 1 IGF receptor binding; 2) a high affinity IGF-1 interaction site formed by residues 39-82 in a globular fold; 3) although CBP-4 does not bind individually to either IGF-1 or NBP-4, in the ternary complex CBP-4 contacts both, stabilizes the complex, and contributes to blocking of the IGF-1R binding region of IGF-1; 4) the central domain, which is unstructured and flexible even when IGFBPs are bound to IGFs, covers the IGF not yet covered by the N- and C-terminal domains and additionally blocks the access of IGF-1R to IGF through steric hindrance.

Formin proteins are nucleating regulators of eukaryotic actin filament assembly and elongation. Formins are modular proteins, containing a series of domains and functional motifs. The formin homology 2 domain (FH2) associates

processively with actin-filament barbed ends and modifies their rate of growth. The FH1 domain influences the function of the FH2 domain through binding to the actin monomer-binding protein, profilin. Two recently resolved crystal structures of Bni1pFH2 and Bni1pFH2 in complex with TMR-actin provide new insights into the molecular details of formin-mediated actin assembly but the mechanism of actin nucleation remains still not clear. Structural analysis of other members of the formin family is important for understanding their precise molecular mechanisms. The studies presented in this thesis were carried out on three human formins (DIAPH1, DAAM1, DAAM2), one Dictyostelium discoideum formin (dDia2) and four profilins (profilin 1 and profilin 2 from both species). Various biochemical and biophysical techniques, including NMR spectroscopy, X-ray crystallography, as well as actin nucleation/polymerization assays, were employed to investigate formins, profilins, actins and their interactions. The main findings of these studies are: 1) the N-terminal fragment of the FH2 domain is flexible, mostly unstructured, and contains a conserved Trp responsible for dimerization; 2) the nucleation activity is depended on the length of the N-terminal fragment of the FH2 domain; 3) dDia2FH2 does not have any nucleation activity; 4) the FH1 domain binds profilins with low affinity; 5) one FH1 domain, with multiple poly-proline stretches, binds simultaneously several profilin molecules; 6) profilins increase the rate of the FH1FH2-mediated actin polymerization.

6 Zusammenfassung

Der Fokus dieser Doktorarbeit war die biochemische und biophysikalische Charakterisierung zweier Gruppen von Proteinen: Insulin like growth factor binding Proteine (IGFBPs) und Formins.

Die Insulin-like growth factor binding proteine (IGFBPs) steuern die Bioverfügbarkeit, Aktivität, und Verteilung von IGF-1 und -2 durch die hochaffienen IGFBP-IGF Komplexe. IGF Bindingsstellen befinden sich sowohl im amino- als auch im carboxyterminalen Fragment der IGFBPs, die auch die konservierten Domänen der IGFBPs bilden. Der jeweilige Beitrag der Domänen zur IGFBP-IGF Komplexbildung war schwer zu analysieren, da die notwendigen dreidimensionale Strukturen fehlten. Um die Effekte der N- und C-terminalen Domäne auf die Interaktion genauer zu untersuchen haben wir verscheiden Röntgenkristall-Strukturen gelöst. Die erste Struktur war der ternäre Komplex aus den N- und C-terminalen Fragmenten von IGFBP-4 und IGF-1, die zweite Struktur war eine hybrid Komplex bei dem das C-terminale Fragment von IGFBP-4 gegen IGFBP-1 ausgetauscht wurde. Wir haben auch den binären Komplex zwischen dem N-terminalen Fragment von IGFBP-4 und IGF-1 bestimmt, um die N- und C-terminale Interaktion durch vergleich mit der ternären Struktur analysierenzu können. Die beschriebene Kristallstruktur gibt die molekulare Grundlage für die IGFBP Regulation des IGF Signalwegs. Dies unterstützt die Forschung in der Suche/Design nach IGFBP Varianten, die als therapeutische IGF Inhibitoren bei Krankheiten mit einer Disregulation von IGF. Schlüssel-Merkmale des IGFBP/IGF Komplex beihalten: 1) Eine Disulfidleiter in der N-terminalen Domäne und die ersten 5 N-terminalen „Daumen“ Aminosäurereste, die an IGF binden und teilweise die IGF Reste, die für die Bindung an den Typ 1 IGF Rezeptor wichtig sind, maskieren. 2) Eine hochaffienen IGF-1 Interaktionsbindungsstelle, die von den Resten 39-83 in einer globulären Faltung gebildet wird. 3)Obwohl CBP-4 alleine nicht an IGF-1 oder NBP-4 bindet, so bindet es im ternären Komplex an beide Proteine , stabilisiert den Komplex und trägt zur Blockierung der IGF-1R Bindungsregion auf IGF-1 bei. 4) Die zentrale Domäne von IGFBP, die auch im gebundenen Zustand an IGF, ungefaltet und flexibel ist, decket den Teil von IGF ab, der von

der N- und C-terminalen Domäne noch nicht erfasst wurde und blockiert zusätzlich den Zugang von IGF-1R an IGF durch eine sterische Hinderung.

Formin Proteine sind Regulatoren der Nukleation der eukaryotischen Aktin Filament Zusammenlagerung und Verlängerung. Formins sind modulare Proteine, die eine Serien von Domänen und Motiven enthalten. Die Formin homology 2 (FH2) Domäne assoziiert mit den „stacheligen“ Actin-Filamentenden und modifiziert die Wachstrumsrate. Die FH1 Domäne beeinflusst die Funktion der FH2 Domäne in dem sie an das Actin-Monomer-binding Protein Profilin bindet. Zwei kürzlich gelöste Kristallstrukturen von Bni1pFH2 und Bni1pFH2 im Komplex mit TMR-Actin gaben neue Einblick in die molekularen Details der Formin gesteuerten Actin Gruppierung. Der Mechanismus der Actin Keimbildung bleibt weiterhin unklar. Strukturuntersuchungen der andern Mitglieder der Formin Familie ist sehr wichtig für das Verständnis des exakten molekularen Mechanismus. Die Forschungsergebnisse die in dieser Arbeit vorgestellt werden umfassen die Untersuchungen an drei humanen Formins: DIAPH1, DAAM1, DAAM2, einen Formin aus Dictyostelium discoideum (dDia2) und vier Profilins (profilin 1 und profilin 2 beider Spezien). Verscheidene biochemische und biophysikalische Techniken, darunter NMR Spektroskopie, Röntgenkristallographie sowie Aktin Nucleation/Polymerisierungs Experimente wurden durchgeführt um Formins, Profilins, Actins und ihre Interaktionen zu untersuchen. Die Ergebnisse dieser Untersuchung sind: 1) Das N-terminal Fragment der FH2 Domäne ist flexibel und hauptsächlich unstrukturiert und enthält ein konservierter Trp das für die Dimerisierung notwendig ist. 2) Die Nucleations Aktivität ist abhängig von der Länge des N-terminalen Fragments der FH2 Domäne. 3) dDia2FH2 hat keine Nucletions Aktivität. 4) Die FH1 Domäne bindet Profilins mit niedriger Affinität. 5) Die FH1 Domäne mit einem multiplen Polyprolin Abschnitten bindet mehrere Profiline gleichzeitig. 6) Profilin steigert die Rate der FH1 FH2-gesteuerten Polymerisation.

7 Appendix

7.1 Abbreviations and symbols

- 1D one-dimensional
- 2D two-dimensional
- Å Ångstrøm (10^{-10} m)
- aa amino acid
- ALS acid labile subunit
- APS ammonium peroxodisulfate
- bp base pair
- BSA bovine serum albumin
- cDNA complimentary DNA
- Da Dalton (g mol-1)
- DAAM1 dishevelled-associated activator of morphogenesis 1
- DAAM2 dishevelled-associated activator of morphogenesis 2
- DAD diaphanous-autoregulatory domain
- dDia2 Dictyostelium discoideum Diaphanous-related formin 2
- DIAPH1 human Diaphanous-related formin 1
- DMSO dimethylsulfoxide
- DNA deoxyribonucleic acid
- EDTA ethylenediamine tetraacetic acid
- FH1 formin homology 1 domain
- FH2 formin homology 2 domain
- g gravity (9.81 m s^{-2})
- GH growth hormone
- GSH reduced glutathione
- GSSG oxidized glutathione
- GST glutathione S-transferase
- HSQC heteronuclear single quantum coherence
- Hz Hertz
- IGF insulin-like growth factor
- IGFBP IGF binding protein

- IGF-1R IGF receptor type I
- IGF-2R IGF receptor type II
- IMAC immobilized metal affinity chromatography
- IPTG isopropyl-β-thiogalactopyranoside
- IR insulin receptor
- IRS insulin receptor substrate(s)
- ITC isothermal titration calorimetry
- K_D dissotation constant
- LB Luria-Broth medium
- MAP mitogen-activated protein kinase
- MM minimal medium
- MW molecular weight
- NiNTA nickel-nitrilotriacetic acid
- NLS nuclear localization signal
- NMR nuclear magnetic resonance
- OD optical density
- P3K phosphatidylinositol 3-kinase
- PAGE polyacrylamide gel electrophoresis
- PBS phosphate-buffered saline
- ppm parts per million
- RMSD root mean square deviation
- SDS sodium dodecyl sulfate
- TB terrific broth
- TEMED N,N,N’,N’-tetramethylethylendiamine

7.2 Full-length IGFBP-1 and IGFBP- 4 sequences

IGFBP-1

```
        10         20         30         40         50
MSEVPVARVW LVLLLLTVQV GVTAGAPWQC APCSAEKLAL CPPVSASCSE VTRSAGCGCC 60
PMCALPLGAA CGVATARCAR GLSCRALPGE QQPLHALTRG QGACVQESDA SAPHAAEAGS 120
PESPESTEIT EEELLDNFHL MAPSEEDHSI LWDAISTYDG SKALHVTNIK KWKEPCRIEL 180
YRVVESLAKA QETSGEEISK FYLPNCNKNG FYHSRQCETS MDGEAGLCWC VYPWNGKRIP 240
GSPEIRGDPN CQIYFNVQN
```

IGFBP- 4

```
       10         20         30         40         50
MLPLCLVAAL LLAAGPGPSL GDEAIHCPPC SEEKLARCRP PVGCEELVRE PGCGCCATCA 60
LGLGMPCGVY TPRCGSGLRC YPPRGVEKPL HTLMHGQGVC MELAEIEAIQ ESLQPSDKDE 120
GDHPNNSFSP CSAHDRRCLQ KHFAKIRDRS TSGGKMKVNG APREDARPVP QGSCQSELHR 180
ALERLAASQS RTHEDLYIIP IPNCDRNGNF HPKQCHPALD GQRGKCWCVD RKTGVKLPGG 240
LEPKGELDCH QLADSFRE
```

7.3 Full-length formins sequences

dDia2

```
        10         20         30         40         50
MSFDLESNSS GGSTIGRNSS IRLSSGLAPS ESTVSLNEII DLDREFELLL DKLAIEDPIK 60
RKQMQSLPDI SKRTLLEQNK ADIYRTVKHK GPIESFADVK SVISSINTKH VPIDIIKTLR 120
IHLNTADRDW IQSFLDNDGV QPILNILKRL ERNKNRKRKE HSILQWECTR CIAALMKIKI 180
GMEYIASFPQ TTNLMVLCLD TPLIKAKTLV LELLAAIAVT DRGHGAVLTS MIYHKEVKKE 240
ITRYFNLVQS LKIEKNAEYL TTCMSFINCI ISSPSDLPSR IEIRKAFLNL KILKYIENLR 300
ADYNEDKNLL TQLDVFEEEL STDEQLNSQQ GTQIGIEDLF SQISSRVTGT PSQQELITLM 360
THFQRMSSSN LGLGVWTLYN ALANQLEDEL KIHPDLDVTL VSLLFPEVKK SSSGLFGFGS 420
KSKSPSSSPA LSSMAKTELK KDNEEKQKTI EHLLKQLNKF SGGQNTERWM IEREEKNKLI 480
AQLMAQTKNG GGGGGGRVGG DSSLSNDEAL KRENQLLRME IENIKNNPSV LLNSGNSING 540
DVPNLFISSP GSTLSPSPSG EPPIPSTDFG ITSSSIHTST DKLTNSTEPI LGSPPPPPPP 600
PMSGGGGPPP PPPPPGGKSN KPAKPIIKPS VKMRNFNWIT IPALKVQGTF WDKLDETSFI 660
QSLDKVELES LFSAKAPTVK VESKQLTRKV VVTVIDMKKA NNCAIMLQHF KIPNEQLKKM 720
QIMLDEKHFS QENAIYLLQF APTKEDIEAI KEYQGDQMQL GAAEQYMLTV MDIPKLDSRL 780
KAFIFKQKFE GLVEDLVPDI KAIKAASLEL KKSKRLSDIL KFILAIGNYV NGSTTRGGAF 840
GFKVLETLPK MRDARSNDNK LSLLHFLAKT LQDRIPEIWN IGAELPHIEH ASEVSLNNII 900
SDSSEIKRSI DLIERDFVPM INDPLFAHDK HWIHKITEFQ KIAKVQYQRI EKEIDEMNKA 960
FEEITSYFGE PKSTQPDVFF STINNFLEDL EKAYGEYQAM IRKAELENSK MEDPEKGGLQ 1020
DLSSQIRSGQ LFKDRRVGDS VIAQMQNVDS LRKNLKSTST TTPNTPPTIK IELPSQSILK 1080
PSGQLKK
```

DAAM1

```
        10         20         30         40         50
MAPRKRGGRG ISFIFCCFRN NDHPEITYRL RNDSNFALQT MEPALPMPPV EELDVMFSEL 60
VDELDLTDKH REAMFALPAE KKWQIYCSKK KDQEENKGAT SWPEFYIDQL NSMAARKSLL 120
ALEKEEEEER SKTIESLKTA LRTKPMRFVT RFIDLDGLSC ILNFLKTMDY ETSESRIHTS 180
LIGCIKALMN NSQGRAHVLA HSESINVIAQ SLSTENIKTK VAVLEILGAV CLVPGGHKKV 240
LQAMLHYQKY ASERTRFQTL INDLDKSTGR YRDEVSLKTA IMSFINAVLS QGAGVESLDF 300
RLHLRYEFLM LGIQPVIDKL REHENSTLDR HLDFFEMLRN EDELEFAKRF ELVHIDTKSA 360
TQMFELTRKR LTHSEAYPHF MSILHHCLQM PYKRSGNTVQ YWLLLDRIIQ QIVIQNDKGQ 420
DPDSTPLENF NIKNVVRMLV NENEVKQWKE QAEKMRKEHN ELQQKLEKKE RECDAKTQEK 480
EEMMQTLNKM KEKLEKETTE HKQVKQQVAD LTAQLHELSR RAVCASIPGG PSPGAPGGPF 540

PSSVPGSLLP PPPPPPLPGG MLPPPPPPLP PGGPPPPPGP PPLGAIMPPP GAPMGLALKK 600
KSIPQPTNAL KSFNWSKLPE NKLEGTVWTE IDDTKVFKIL DLEDLERTFS AYQRQQDFFV 660
NSNSKQKEAD AIDDTLSSKL KVKELSVIDG RRAQNCNILL SRLKLSNDEI KRAILTMDEQ 720
EDLPKDMLEQ LLKFVPEKSD IDLLEEHKHE LDRMAKADRF LFEMSRINHY QQRLQSLYFK 780
KKFAERVAEV KPKVEAIRSG SEEVFRSGAL KQLLEVVLAF GNYMNKGQRG NAYGFKISSL 840
NKIADTKSSI DKNITLLHYL ITIVENKYPS VLNLNEELRD IPQAAKVNMT ELDKEISTLR 900
SGLKAVETEL EYQKSQPPQP GDKFVSVVSQ FITVASFSFS DVEDLLAEAK DLFTKAVKHF 960
GEEAGKIQPD EFFGIFDQFL QAVSEAKQEN ENMRKKKEEE ERRARMEAQL KEQRERERKM 1020
RKAKENSEES GEFDDLVSAL RSGEVFDKDL SKLKRNRKRI TNQMTDSSRE RPITKLNF
```

DAAM2

```
        10         20         30         40         50
MAPRKRSHHG LGFLCCFGGS DIPEINLRDN HPLQFMEFSS PIPNAEELNI RFAELVDELD 60
LTDKNREAMF ALPPEKKWQI YCSKKKEQED PNKLATSWPD YYIDRINSMA AMQSLYAFDE 120
EETEMRNQVV EDLKTALRTQ PMRFVTRFIE LEGLTCLLNF LRSMDHATCE SRIHTSLIGC 180
IKALMNNSQG RAHVLAQPEA ISTIAQSLRT ENSKTKVAVL EILGAVCLVP GGHKKVLQAM 240
LHYQVYAAER TRFQTLLNEL DRSLGRYRDE VNLKTAIMSF INAVLNAGAG EDNLEFRLHL 300
RYEFLMLGIQ PVIDKLRQHE NAILDKHLDF FEMVRNEDDL ELARRFDMVH IDTKSASQMF 360
ELIHKKLKYT EAYPCLLSVL HHCLQMPYKR NGGYFQQWQL LDRILQQIVL QDERGVDPDL 420
APLENFNVKN IVNMLINENE VKQWRDQAEK FRKEHMELVS RLERKERECE TKTLEKEEMM 480
RTLNKMKDKL ARESQELRQA RGQVAELVAQ LSELSTGPVS SPPPPGGPLT LSSSMTTNDL 540
PPPPPPLPFA CCPPPPPPPL PPGGPPTPPG APPCLGMGLP LPQDPYPSSD VPLRKKRVPQ 600
PSHPLKSFNW VKLNEERVPG TVWNEIDDMQ VFRILDLEDF EKMFSAYQRH QKELGSTEDI 660
YLASRKVKEL SVIDGRRAQN CIILLSKLKL SNEEIRQAIL KMDEQEDLAK DMLEQLLKFI 720
PEKSDIDLLE EHKHEIERMA RADRFLYEMS RIDHYQQRLQ ALFFKKKFQE RLAEAKPKVE 780
AILLASRELV RSKRLRQMLE VILAIGNFMN KGQRGGAYGF RVASLNKIAD TKSSIDRNIS 840
LLHYLIMILE KHFPDILNMP SELQHLPEAA KVNLAELEKE VGNLRRGLRA VEVELEYQRR 900
QVREPSDKFV PVMSDFITVS SFSFSELEDQ LNEARDKFAK ALMHFGEHDS KMQPDEFFGI 960
FDTFLQAFSE ARQDLEAMRR RKEEEERRAR MEAMLKEQRE RERWQRQRKV LAAGSSLEEG 1020
GEFDDLVSAL RSGEVFDKDL CKLKRSRKRS GSQALEVTRE RAINRLNY
```

DIAPH1

```
         10         20         30         40         50
MEPPGGSLGP GRETRDKKKG RSPDELPSAG GDGGKSKKFL ERFTSMRIKK EKEKPNSAHR 60
NSSASYGDDP TAQSLQDVSD EQVLVLFEQM LLDMNLNEEK QQPLREKDII IKREMVSQYL 120
YTSKAGMSQK ESSKSAMMYI QELRSGLRDM PLLSCLESLR VSLNNNPVSW VQTFGAEGLA 180
SLLDILKRLH DEKEETAGSY DSRNKHEIIR CLKAFMNNKF GIKTMLETEE GILLLVRAMD 240
PAVPNMMIDA AKLLSALCIL PQPEDMNERV LEAMTERAEM DEVERFQPLL DGLKSGTTIA 300
LKVGCLQLIN ALITPAEELD FRVHIRSELM RLGLHQVLQD LREIENEDMR VQLNVFDEQG 360
```

```
EEDSYDLKGR LDDIRMEMDD FNEVFQILLN TVKDSKAEPH FLSILQHLLL VRNDYEARPQ 420
YYKLIEECIS QIVLHKNGAD PDFKCRHLQI EIEGLIDQMI DKTKVEKSEA KAAELEKKLD 480
SELTARHELQ VEMKKMESDF EQKLQDLQGE KDALHSEKQQ IATEKQDLEA EVSQLTGEVA 540
KLTKELEDAK KEMASLSAAA ITVPPSVPSR APVPPAPPLP GDSGTIIPPP PAPGDSTTPP 600
PPPPPPPPPP PLPGGTAISP PPPLSGDATI PPPPPLPEGV GIPSPSSLPG GTAIPPPPPL 660
PGSARIPPPP PPLPGSAGIP PPPPPLPGEA GMPPPPPPLP GGPGIPPPPP FPGGPGIPPP 720
PPGMGMPPPP PFGFGVPAAP VLPFGLTPKK LYKPEVQLRR PNWSKLVAED LSQDCFWTKV 780
KEDRFENNEL FAKLTLTFSA QTKTKKDQEG GEEKKSVQKK KVKELKVLDS KTAQNLSIFL 840
GSFRMPYQEI KNVILEVNEA VLTESMIQNL IKQMPEPEQL KMLSELKDEY DDLAESEQFG 900
VVMGTVPRLR PRLNAILFKL QFSEQVENIK PEIVSVTAAC EELRKSESFS NLLEITLLVG 960
NYMNAGSRNA GAFGFNISFL CKLRDTKSTD QKMTLLHFLA ELCENDYPDV LKFPDELAHV 1020
EKASRVSAEN LQKNLDQMKK QISDVERDVQ NFPAATDEKD KFVEKMTSFV KDAQEQYNKL 1080
RMMHSNMETL YKELGEYFLF DPKKLSVEEF FMDLHNFRNM FLQAVKENQK RRKTEEKMRR 1140
AKLAKEKAEK ERLEKQQKRE QLIDMNAEGD ETGVMDSLLE ALQSGAAFRR KRGPRQANRK 1200
AGCAVTSLLA SELTKDDAMA AVPAKVSKNS ETFPTILEEA KELVGRAS
```

8 References

Alberts, A. S. (2001). Identification of a carboxyl-terminal diaphanous-related formin homology protein autoregulatory domain. *J. Biol. Chem.* 276, 2824-2830.

Allan, G. J., Tonner, E., Szymanowska, M., Shand, J. H., Kelly, S. M., Phillips, K., Clegg, R. A., Gow, I. F., Beattie, J. & Flint, D. J. (2006) Cumulative mutagenesis of the basic residues in the 201-218 region of insulin-like growth factor-binding protein 5 results in progressive loss of both IGF-1 binding and inhibition of IGF-1 biological actions. *Endocrinology* 147**,** 338-349.

Andress, D.L. (1995) Heparin modulates the binding of insulin-like growth factor (IGF) binding protein-5 to a membrane protein in osteoblastic cells. J. Biol. Chem. *47*, 28289–28296.

Babajko, S., and Binoux, M. (1996) Modulation by retinoic acid of insulin-like growth factor (IGF) and IGF binding protein expression in human SK-N-SH neuroblastoma cells. E. J. Endocrinology, *134*, 474–480.

Bach, L.A., (1999) The insulin-like growth factor system: basic and clinical aspects. Aust. N.Z. J. Med. *29*, 355-361.

Bach, L.A., and Rechler, M.M. (1995) Insulin-like growth factor binding proteins. Diabetes Rev. *3*, 38-61.

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.

Ballard, F.J., Wallace, J.C., Francis, G.L., Read, L.C., and Tomas, F.M. (1996) Des (1-3)-IGF-I: a truncated form of insulin-like growth factor-I. Int. J. Biochem. Cell. Biol. *28*, 1085-7.

Baxter, R.C. (1988) Characterisation of the acid-labile subunit of the growth hormone-dependent insulin-like growth factor binding protein complex. J. Clin. Endocrinol. Metab. *67*, 265-72.

Baxter, R.C. (1991) Insulin-like growth factor (IGF) binding proteins: the role of serum IGFBPs in regulating IGF availability. Acta Paediatr. Scand. *372*, 107-14.

Baxter, R.C., Bayne, M.L., and Cascieri, M. A. (1992) Structural determinants for binary and ternary complex formation between insulin-like growth factor-I (IGF-I) and IGF binding protein-3. J. Biol. Chem. *267*, 60-65.

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

Beck, K.D., Knusel, B., and Hefti, F. (1993) The nature of the trophic action of brain-derived neurotrophic factor des(1-3)-insulin-like growth factor-1, and basic fibroblast growth factor on mesencephalic dopaminergic neurons developing in culture*.* Neurosci. *52*, 855-66.

Boisclair, Y.R., Rhoads, R.P., Ueki, I. Wang, J. and. Ooi, G.T. (2001) Regulation of the acid-labile subunit of the 150-kDa IGF-binding protein complex and its role in the circulating IGF system. J. Anim. Sci. *79*, E41-E47.

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.

Branden, C. and Tooze, J. (1999) Introduction to Protein Structure. Garland Publ., New York.

Braulke, T. (1999) Type-2 IGF receptor: a multi-ligand binding protein. Hormone and Metabolism Research, *31,* 242–246.

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

Burtrum, D., Zhu, Z., Lu, D., Anderson, D.M., Prewett, M., Pereira, D.S., Bassi, R., Abdullah, R., Hooper, A.T., Koo, H., Jimenez, X., Johnson, D., Apblett, R., Kussie, P., Bohlen, P., Witte, L., Hicklin, D.J., and Ludwig, D.L. (2003) A fully human monoclonal antibody to the insulin-like growth factor I receptor blocks ligand-dependent signaling and inhibits human tumor growth *in vivo*. Cancer Research, *63*, 8912–8921.

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. (2001a) Localization of the IGF binding domain and evaluation of the role of cysteine residues in IGF binding in IGF binding protein-4. Journal of Endocrinology, *169*, 135–143.

Byun, D., Mohan, S., Yoo, M., Sexton, C., Baylink, D.J., and Qin, X. (2001b) Pregnancy-associated plasma protein-A accounts for the insulin-like growth factor (IGF)-binding protein-4 (IGFBP-4) proteolytic activity in human pregnancy serum and enhances the mitogenic activity of IGF by degrading IGFBP-4 in vitro. Journal of Clinical Endocrinology and Metabolism, *86*, 847–854.

Campbell, P.G. et al. (1998) Plasminogen binds the heparin-binding domain of insulin-like growth factor-binding protein 3. Am. J. Physiol. 275: E321-E331.

Carlsson–Skwirut, C., Jornvall, H., Holmgren, A., Andersson, C., Bergman, T., Lundquist, B., and Sara, V.R. (1986) Isolation and characterisation of variant IGF-I as well as IGF-2 from adult human brain. FEBS Lett. *201*, 46-50.

Carlson-Skwirut, C., Lake, M., Hartmanis, M., Hall, K., and Sara, V.R. (1989) A comparison of the biological activity of the recombinant intact, and truncated insulin-like growth factor 1 (IGF-1). Biochim. Biophys. Acta. *1011*, 192-7.

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

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

CCP4, Collaborative Computational Project, Number 4. (1994) The CCP4 Suite: Programs for protein crystallography. Acta Cryst., *D50*, 760-763.

Chan, J.M., Stampfer, M.J., Giovannucci, E., Gann, P.H., Ma, J., Wilkinson, P., Henneken, C.H., and Pollak, M. (1998) Plasma insulin-like growth factor-I and prostate cancer risk: A prospective study. Science, *279*, 563–566.

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*, 283-96.

Chen, C., Zhu, Y.F., Liu, X.J., Lu, Z.X., Xie, Q., and Ling, N. (2001) Discovery of a series of nonpeptide small molecules that inhibit the binding of insulin-like growth factor (IGF) to IGF-binding proteins. J Med Chem. *44*,:4001-10.

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.

Clemmons, D.R., Cascieri, M.A., Camacho-Hübner, C., Mc-Cusker, R.H., and Bayne, M.L. (1990) Discrete alterations of the insulin-like growth factor I molecule which alter its affinity for insulin-like growth factor-binding proteins result in changes in bioactivity. J. Biol. Chem. *265*, 12210–12216.

Clemmons, D.R. (1999) IGF binding proteins and extracellular matrix. In Rosenfeld, R.G., and Roberts, C.T. (eds), The IGF system. Molecular biology, physiology, and clinical applications. Humana Press, Totowa, pp. 273-279.

Clemmons, D.R. (2001) Use of mutagenesis to probe IGF-binding protein structure/function relationships. Endocr. Rev. *22*, 800-817.

Cohen, P., Lamson, G., Okajima, T., and Rosenfeld, R.G. (1993) Transfection of the human insulin-like growth factor binding protein-3 gene into Balb/c fibroblasts inhibits cellular growth. Mol. Endocrinol. *7*, 380-386.

Cohen, P., Peehl, D.M., Graves, H.C., and Rosenfeld, R.G. (1994) Biological effects of prostate specific antigen as an insulin-like growth factor binding protein-3 protease. J. Endocrinol. *142*, 407-15.

Cohen, P., Clemmons, D.R., and Rosenfeld, R.G. (2000) Does the GH-IGF axis play a role in cancer pathogenesis? Growth Hor. IGF Res. *10*, 297-305.

Conover, C.A., Oxvig, C., Overgaard, M.T., Christiansen, M., and Giudice, L.C. (1999) Evidence that the insulin-like growth factor binding protein-4 protease in human ovarian follicular fluid is pregnancy associated plasma protein-A. Journal of Clinical Endocrinology and Metabolism, *84*, 4742–4745.

Cooke, R.M., Harvey, T.S., and Campbell, I.D. (1991) Solution structure of human insulin-like growth factor 1: a nuclear magnetic resonance and restrained molecular dynamics study. Biochemistry, *30*, 5484-5491.

De Meyts, P., Wallach, B., Christoffersen, C.T., Urso, B., Gronskov, K., Latus, L.J., Yakushiji, F., Ilondo, M.M., and Shymko, R.M. (1994) The insulin-like growth factor-I receptor: Structure, ligand-binding mechanism and signal transduction. Horm. Res. *42*, 152-69.

Derewenda, Z. (2004) Rational protein crystallization by mutational surface engineering. Structure, 12, 529-535.

Dong, Y., Pruyne, D., Bretscher, A. (2003). Formin- dependent actin assembly is regulated by distinct modes of Rho signaling in yeast. *J. Cell Biol.* 161, 1081-1092.

Drenth, J. (1994) Principles of protein X-ray crystallography. Springer-Verlag, New York

Duan, C. (1997) The insulin-like growth factor system and its biological actions in fish. Am. Zool. 37: 489-501.

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*, 6386-96.

Evangelista, M., Pruyne, D., Amberg, D. C., Boone, C., Bretscher, A. (2002). Formins direct Arp2/3-independent actin filament assembly to polarize the growth in yeast. *Nat. Cell Biol.* 4, 260-269.

Evangelista, M., Blundell, K., Longtine, M. S., Chow, C. J., Adames, N., Pringle, J. R., Peter, M., Boone, C. (1997). Bni1p, a yeast forming linking cdc42p and the actin cytoskeleton during polarized morphogenesis. *Science* 276, 118-112.

Faix, J., and Rottner, K., (2006) The making of filopodia. Curr. Opin. Cell Biol.18, 18-25

Fanayan, S., Firth, S.M., Butt, A.J., and Baxter, R.C. (2000) Growth inhibition by insulin-like growth factor binding protein-3 in T47D breast cancer cells requires transforming growth factor-β (TGF-β) and type II TGF-β receptor. J. Biol. Chem. *275*, 39146-39151.

Feierbach, B., Chang, F. (2001). Roles of the fission yeast formin for3p in cell polarity, actin cable formation and symmetric cell division. *Curr. Biol.* 11, 1656-1665.

Fernandez-Tornero, C., Lozano, R. M., Rivas, G., Jimenez, M. A., Ständker, L., Diaz-Gonzalez, D., Forssmann, W.G., Cuevas, P., Romero, A. & Giménez-Gallego, G. (2005) Synthesis of the blood circulating C-terminal fragment of insulin-like growth factor binding protein –4 in its native conformation. Crystallization, heparin and IGF binding, and osteogenic activity. *J. Biol. Chem.* 280, 18899–18907. 13.

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, *14*, 2147-2150.

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., 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.

La Fortelle, E. de and Bricogne, G. (1997) Maximum-likelihood heavy-atom parameter refinement in MIR and MAD methods. Meth. Enzym. *276*, 472-494.

Fowlkes, J., and Freemark, M. (1992) Evidence for a novel insulin-like growth factor (IGF)-dependent protease regulating IGF-binding protein-4 in dermal fibroblasts. Endocrinology, *131*, 2071-2076.

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. Biochemical J. *251*, 95-103.

Frost, R.A., and Tseng L. (1991) Insulin-like growth factor binding protein-1 is phosphorylated by cultured human endometrial stromal cells and multiple protein kinases in vitro. J. Biol. Chem. *266*, 18082-18088.

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

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.

Gallicchio, M.A., Kneen, M., Hall, C., Scott, A.M., and Bach, L.A. (2001) Overexpression of insulin-like growth factor binding protein-6 inhibits rhabdomyosarcoma growth in vivo. Int. J. Cancer. *94*, 645-651.

Garrett, T.P., McKern, N.M., Lou, M., Frenkel, M.J., Bentley, J.D., Lovrecz, G.O., Elleman, T.C., Cosgrove, L.J., and Ward, C.W. (1998) Crystal structure of the first three domains of the type-1 insulin-like growth factor receptor. Nature, *394*, 395-9.

Gasman, S., and Zerial, M. (2003) RhoD regulates endosome dynamics through Diaphanous-related Formin and Src tyrosine kinase. Nat. Cell. Biol.5: 195-204

Giacovazzo, C., Monaco, H.L., Viterbo, D., Scordari, F., Gilli, G., Zanotti, M. and Catti, M. (1992) Fundamentals of crystallography. Oxford University Press, Oxford.

Gill, S.C., and von Hippel, P.H. (1989) Calculation of protein extinction coefficients from amino acid sequence data. Anal. Biochem. 182, 319-26.

Girnita, A., Girnita, L., del Prete, F., Bartolazzi, A., Larsson, O., and Axelson, M. (2004) Cyclolignans as inhibitors of the insulin-like growth factor-1 receptor and malignant cell growth. Cancer Research, 64, 236–42.

Giudice, L.C., Conover, C.A., Bale, L., Faessen, G.H., Ilg, K., Sun, I., Imani, B., Suen, L.-F., Irwin, J.C., Christiansen, M., Overgaard, M.T., and Oxvig, C. (2002) Identification and Regulation of the IGFBP-4 Protease and Its Physiological Inhibitor in Human Trophoblasts and Endometrial Stroma: Evidence for Paracrine Regulation of IGF-II Bioavailability in the Placental Bed during Human Implantation. The Journal of Clinical Endocrinology & Metabolism, *87*, 2359-66.

Glotzer,M., (2005) The molecular requirements for cytokinesis. Science. 307: 1735-1739

Grellier, P., Degalle, B., Babajko, S. (1998) Expression of insulin-like growth factor-binding protein 6 complementary DNA alters neuroblastoma cell growth. Cancer Res. *58*, 1670-1676.

Grosshans, J., et al., (2005) RhoGEF2 and the forming Dia control the formation of the furrow canal by directed actin assembly during Drosophila cellularization. Development 132: 1009-1020.

Grzmil, M., Hemmerlein, B., Thelen, P., Schweyer, S., and Burfeind, P. (2004) Blockade of the type I IGF receptor expression in human prostate cancer cells inhibits proliferation and invasion, up-regulates IGF binding protein-3, and suppresses MMP-2 expression. J. Pathol. *202*, 50-9.

Gundersen, G.G., and Wen, Y. (2004) Cortical control of microtubule stability and polarization. Curr. Opin. Cell Biol. 16: 106-112.

Habas, R. and He, X. (2001) Wnt/Frizzeld activation of Rho regulates vertebrate gastrulation and requires a novel formin homology protein Daam1. Cell. 107: 843-857.

Halloran, B.P., and Spencer, E.M. (1988) Dietary phosphorus and 1,25-dihydroxyvitamin D metabolism: influence of insulin-like growth factor I. Endocrinology, *123*, 1225-9.

Harris, E. S., Li, F., Higgs, H. N. (2004). The mouse formin, FRLα, slows actin filament barbed end elongation, competes with capping protein, accelerates polymerization from monomers, and severs filaments. *J. Biol. Chem.* 279, 20076-20087

Hashimoto, R., Ono, M., Fujiwara, H., Higashihashi, N., Yoshida, M., Enjohkimura, 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*, 27936-42.

Hassan, A.B. (2003) Keys to the hidden treasures of the mannose 6-phosphate/insulin-like growth factor 2 receptor. American Journal of Pathology, *162*, 3–6.

Headey, S.J., Keizer, D.W., Yao, S., Wallace, J.C., Bach, L.A., and Norton, R.S. (2004a) 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., Keizer, D.W., Yao, S., Brasier, G., Kantharidid P., Bach, L.A., and Norton, R.S. (2004b) C-terminal domain of insulin-like growth factor binding protein-6: structure and interaction with insulin-like growth factor-II Mol. Endocrinol. *18*, 2740-50.

Headey, S.J., Leeding, K.S., Norton, R.S., and Bach, L.A. (2004c) Contributions of the N- and C-terminal domains of IGF binding protein-6 to IGF binding. J. Mol. Endocr. *33*, 377-86.

Higashida, C., Miyoshi, T., Fujita, A., Oceguera-Yanez, F., Monypenny, J., Andou, Y., Narumiya, S., Watanabe, N. (2004). Actin polymerization-driven molecular movementof mDia1 in living cells. *Science* 303, 2007-2010.

Higgs, H. N., Peterson, K. J. (2005). Phylogenetic analysis of the forming homology 2 domain. *Mol. Biol. Cell* 16, 1-13.
Higgs, H. N. (2005). Formin proteins: a domain-based approach. *Trends Biochem. Sci.* 30, 342-353.

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.

Hodgson, D.R., May, F.E.B., and Westley, B.R. (1996) Involvement of phenylalanine 23 in the binding of IGF-1 to the insulin and type I IGF receptor. Regul. Pept. *66,* 191-196.

Huang, S.S., Leal, S.M., Chen, C-L., Liu, I-H., and Huang, J.S. (2004) Identification of insulin receptor substrate proteins as key molecules for the TβR-V/LRP-1-mediated growth inhibitory signaling cascade in epithelial and myeloid cells. FASEB, 18, 1230-54.

Humbel, R.E. (1990) Insulin-like growth factor-I, and II. Eur. J. Biochem. *190*, 445-62.

Imai, Y., Moralez, A., Andag, U., Clarke, J.B., Busby, W.H., 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*, 18188-94.

Imamura, H., Tanaka, K., Hihara, T., Umikawa, M., Kamei, T., Takahashi, K., Sasaki, T., Takai, Y. (1997). Bni1p and Bnr1p: downstream targets of the Rho family small G-proteins which interact with profilin and regulate actin cytoskeleton in *Saccharomyces cerevisiae. EMBO J.* 16**,** 2745-2755.

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*, 7481-5.

Jones, J.L., Gockerman, A., Busby, W.H., Wrigh, G. and Clemmons, D.R. (1993) Insulin-like growth factor binding protein 1 stimulates cell migration and binds to the a5b integrin receptor by means of its Arg-Gly-Asp sequence. Proc. Natl. Acad. Sci. USA, *90*, 10553-7.

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

Jung, G.W., Spencer, E.M., and Lue, T.F. (1998) Growth hormone enhances regeneration of nitric oxide synthase-containing penile nerves after cavernous nerve neurotomy in rats. Journal of Urology, *160*, 1899-904

Kabsch, W. (1993) Automatic processing of rotation diffraction data from crystals of initially unknown symmetry and cell constants. J. Appl. Cryst. *26*, 795-800.

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*, 6558-72.

Kelley, K.M., Oh, Y., Gargosky, S.E., Gucev, Z., Matsumoto, T., Hwa, V., Ng, L., Simpson, D.M., and Rosenfeld, R.G. (1996) Insulin-like growth factor-binding proteins (IGFBPs) and their regulatory dynamics. Int. J. Biochem. and Cell Biol. *28*, 619–637.

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

Khandwala, H.M., McCutcheon, I.E., Flyvbjerg, A., and Friend, K.E. (2000) The effects of insulin-like growth factors on tumorigenesis and neoplastic growth. Endocr. Rev. *21*, 215-44.

Knudtson, K.L., Boes, M., Sandra, A., Dake, B.L., Booth, B.A., and Bar, R.S. (2001) Distribution of chimeric IGF binding protein (IGFBP)-3 and IGFBP-4 in the rat heart: importance of C-terminal basic region. Endocrinology, *142*, 3749–3755.

Kobielak, A., Passoli, H. A., Fuchs, E. (2004). Mammalian formin-1 participates In adherens junctions and polymerization of linear actin cables. *Nat. Cell Biol.* 6, 21-30.

Kovar, D. R., Pollard, T. D. (2004). Insertional assembly of actin filament barbed ends in association with formins produces piconewton forces. *Proc. Natl. Acad. Sci. USA* 101, 14725-14730.

Kovar, D. R., Wu, J. Q., Pollard, T. D. (2005). Profilin-mediated competition between capping proteins and forming Cdc12p during cytokinesis in fission yeast. *Mol. Biol. Cell* 16, 2313-2324.

Kovar, D. R., Kuhn, J. R., Tichy, A. L., Pollard, T. D. (2003). The fission yeast cytokinesis formin Cdc12p is a barbed end actin filament capping protein gated by profilin. *J. Cell Biol.* 161, 875-887.

Lalou, C., Lassarre, C., and Binoux, M. (1996). A proteolytic fragment of insulin-like growth factor (IGF) binding protein-3 that fails to bind IGFs inhibits the mitogenic effects of IGF-I and insulin. Endocrinology, *137*, 3206-12.

Lamzin, V.S., and Wilson, K.S. (1993) Automated refinement of protein models. Acta Cryst. *D49*, 129-47.

Landale, E.C., Strong, D.D., Mohan, S. and Baylink, D.J. (1995) Sequence comparison and predicted structure for the four exon-encoded regions of human insulin-like growth factor binding protein 4. Growth Factors, *12*, 245–50.

Lassarre, C., and Binoux, M. (1994) Insulin-like growth factor binding protein-3 is functionally altered in pregnancy plasma. Endocrinology, *134*, 1254–62.

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-β receptor in mink lung epithelial cells. J. Biol. Chem*. 274*, 6711–7.

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. J. Biol. Chem. *279*, 469–76.

LeRoith, D. and Helman, L. (2004) The new kid on the block(ade) of the IGF-1 receptor. Cancer Cell, *5*, 201-2.

LeRoith, D., Werner, H., Faria, T.N., Kato, H., Adamo, M., and Roberts, C.T. (1993) Insulin-like growth factor receptor: implications for nervous function. Ann. NY Acad. Sci. *692*, 22-32.

LeRoith, D., Werner, H., Beitner-Johnson, D., and Roberts, C.T. (1995) Molecular and cellular aspects of the insulin-like growth factor I receptor. Endocr. Rev. *16*, 143-63..

Li, W., Fawcett, J., Widmer, H.R., Fielder, P.J., Rabkin, R., and Keller, G.A. (1997) Nuclear transport of insulin-like growt factor-I and insulin-like growth factor binding protein-3 in opossum kidney cells. Endocrinology, *138*, 1763-6.

Li, F., Higgs, H. N. (2003). The mouse forming mDia1 is a potent actin nucleation factor regulated by autoinhibition. *Curr. Biol.* 13, 1335-1340.

Li, F., Higgs, H. N. (2005). Dissecting requirements for auto-inhibition of actin nucleation by the formin mDia1. *J. Biol. Chem.* 280, 6986-6992.

Lin, T.M., Galbert, S.P., Kiefer, D., Spellacy, W.N., and Gall, S. (1974) Characterization of four human pregnancy-associated plasma proteins. American Journal of Obstetrics and Gynecology, *118*, 223–36.

Liu, X.J., Xie, Q., Zhu, Y.F., Chen, C., Ling, N. (2001) Identification of a nonpeptide ligand that releases bioactive insulin-like growth factor-I from its binding protein complex. J. Biol. Chem. *31*, 32419-22.

Loddick, S.A., Liu, X.J., Lu, Z.X., Liu, C.L., Behan, D.P., Chalmers, D.C., Foster, A.C., Vale, W.W., Ling, N., and Desouza, E.B. (1998) Displacement of insulin-like growth factors from their binding proteins as a potential treatment for stroke. Proc. Nat. Acad. Sci. USA, *95*, 1894-8.

Maile, L.A., Gill, Z.P., Perks, C.M., and Holly, J.M. (1999) The role of cell surface attachment and proteolysis in the insulin-like growth factor (IGF)-independent effects of IGF-binding protein-3 on apoptosis in breast epithelial cells. Endocrinology, *140*, 4045-5.

Martin, J.L, and Baxter, R.C. (1999) IGF binding proteins as modulators of IGF action.
Makrides, S.C. (1996) Strategies for achieving high-level expression of genes in Escherichia coli. Microbiol. Rev., *60*, 512-538.

Maures, T.J., and Duan, C. (2002) Structure, developmental expression, and physiological regulation of zebrafish IGF-binding protein-1. Endocrinology 143: 1858-1871.

McCusker, R.H., and Clemmons, D.R. (1997) Use of lanthanum to accurately quantify insulin-like growth factor binding to proteins on cell surfaces. J. Cell. Biochem. *66*, 256–267.

McKern, N., et al., (2006) Structure of the insulin receptor ectodomain reveals a folded-over conformation. Nature; 443

McRee, D.E. (1999) XtalView/Xfit - A versatile program for manipulating atomic coordinates and electron density. J. Struc. Biol. *125*, 156-65.

Mejillano, M. R., Kojima, S., Applewhite, D. A., Gertler, F. B., Svitkina, T. M., Borisy, G. G. (2004). Lamelliopodial versus filopodial mode of the actin nanomachinery: pivotal role of the filament barbed end. *Cell* 118, 363-373.

Michelot, A., Guerin, C., Huang, S., Ingouff, M., Richard, S., Rodiuc, N., Staiger, C. J., Blanchoin, L. (2005). The formin homology-1 domain modulates the actin nucleation and bundling activity of *Arabidopsis* FORMIN1. *Plant Cell* 17, 2296-2313.

Miralles, F., et al., (2003) Actin dynamics control SRF activity by regulation of its coactivator MAL. Cell 113: 329-342.

Mishra, S., and Murphy, L. J. (2003) Phosphorylation of insulin-like growth factor (IGF) binding protein-3 by breast cancer cell membranes enhances IGF-I binding. Endocrinology, *144*, 4042–50.

Miyagi, et al., (2002) Delphilin: a novel PDZ domain and forming homology domain-containing protein that synapticaly colocalizes and interacts with glutamate receptor delta 2 subunit. J. Neurosci. 22: 803-814

Mosley, J. B., Sagot, I., Manning, A. L., Xu, Y., Eck, M. J., Pellman, D., Goode,. B. L. (2004).A conserved mechanism for Bni1- and mDia1- induced actin assembly and dual regulation of Bni1 by Bud6 and profilin. *Mol. Biol. Cell* 15, 896-907

Moseley, J. B., Goode, B. L. (2005). Differential activities and regulation of *Saccharomyces cerevisiae* formin proteins Bni1 and Bnr1 by Bud6. *J. Biol. Chem.* 280, 28023-28033.

Nakano, K., Imai, J., Arai, R., Toh, E. A., Matusi, Y., Mabuchi, I. (2002). The small GTPase Rho3 and the diaphanous/formin For3 function in polarized cell growthin fission yeast. *J. Cell Sci.* 115, 4629-4639.

Nam, T.J., Busby, W.H., and Clemmons, D.R. (1994) Human fibroblasts secrete a serine protease that cleaves insulin-like growth factor binding-5. Endocrinology, *135*, 1385–1391.

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

Nef, S. et al. (2003) Testis determination requires insulin receptor family function in mice. Nature 426: 291-295.

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.

Nezami, A.G., Poy, F., Eck, M.J. (2006) Structure of the autoinhibitory switch in formin mDia1. Structure 14: 257-263.

Novinec, M., Kordis, D., Turk, V., & Lenarcic, B. (2006) Diversity and evolution of the thyroglobulin type-1 domain superfamily. *Mol. Biol. Evol.* 23**,** 744-755.

Otomo, T., Tomchick, D. R., Otomo, C., Panchal, S. C., Machius, M., Rosen, M. K. (2005). Structural basis of actin filament nucleation and processive capping by a forming homology-2 domain. *Nature* 433, 488-494.

Otomo, T., Otomo, C., Tomchick, D. R., Machius, M., Rosen, M. K. (2005). Structural basis of Rho GTPase-mediated activation of the forming mDia1. *Mol. Cell* 18, 273-281.

Payet, L.D., Wang, X.H., Baxter, R.C., and Firth, 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*, 2797-806.

Perdue, J.F., Bach L.A., Hashimoto, R., Sakano, K., Fujita-Yamaguchi, Y., Fujiwara, H., and Rechler, M.M. (1994) in: The insulin-like growth factors and their regulatory proteins, Baxter, R.C., Gluckman, P.D., and Rosenfeld, R.G. (eds), Elsevier Science Publishing Co., New York, pp. 67-76.

Perks, C.M., Newcomb, P.V., Norman, M.R., and Holly, J.M. (1999) Effect of insulin-like growth factor binding protein-1 on integrin signaling and the induction of apoptosis in human breast cancer cells. J. Mol. Endocrinol. *22*, 141–50.

Perrakis, A., Harkiolaki, M., Wilson, K.S., and Lamzin, V.S. (2001) ARP/wARP and molecular replacement. Acta Cryst. *D57*, 1445-50.

Pellegrin, S., Mellor, H. (2005). The Rho family GTPase Rif induces filopodia through mDia2. *Curr. Biol.* 15, 129-133.

Petersen, J., Nielsen, O., Egel, R., Hagan, I. M. (1998). FH3 domain found in formins, targets the fission yeast formin Fus1 to the projection tip during conjugation. *J. Cell Biol.* 141, 1217-1228.

Perelroizen, I., Marchand, J.B., Blanchoin, L., Didry, D., and Carlier, M.F. (1994) Interaction of profiling with G-actin and poly(L-proline). Biochemistry, 33, 8472-8478.

Petrella, E.C., Machesky, L.M., Kaiser, D.A., and Pollard, T.D. (1996) Structural requirements and thermodynamics of the interaction of proline peptides with profiling. Biochemistry, 35, 16535 – 16543.

Pollak, M.N., Schernhammer, E.S., and Hankinson, S.E. (2004) Insulin-like growth factors and neoplasia. Nature Rev. Cancer, *4*, 505-18.

Pollard, T. D. (2004). Formins coming into focus. *Dev. Cell* 6, 312-314.

Pring, M., Evangelista, M., Boone, C., Zigmond, S. H. (2003). Mechanism of formin-induced nucleation of actin filaments. *Biochemistry* 42, 486-496

Pruyne, D., Evangelista, M., Yang, C., Bi, E., Zigmond, S., Bretscher, A., Boone, C. (2002). Role of formins in actin assembly: Nucleation and barbed-end association. *Science* 297, 612-615

Quinlan, M. E., Heuser, J. E., Kerkhoff, E., Mullins, R. D. (2005). Drosophila Spire in an actin nucleation factor. *Nature* 433, 382-388.

Qin, X., Byun, D., Strong, D.D., Baylink, D.J. and Mohan, S. (1999) Studies on the role of human insulin-like growth factor-II (IGF-II)-dependent IGF binding protein (hIGFBP)-4 protease in human osteoblasts using protease-resistant IGFBP-4 analogs. Journal of Bone and Mineral Research, *14*, 2079–2088.

Qin, X.Z., 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*, 23509-16.

Rajah, R., Valentinis, B., and Cohen, P. (1997) Insulin-like growth factor (IGF)-binding protein-3 induces apoptosis and mediates the effects of transforming growth factor-b1 on programmed cell death through a p53- and IGF-independent mechanism. J. Biol. Chem. *272*, 12181–8.

Rehm, T., Huber, R., and Holak, T.A. (2002) Application of NMR in structural proteomics: screening for proteins amenable to structural analysis. Structure, *10*, 1613-18.

Romero, S., Le Clainche, C., Didry, D., Egile, C., Pantaloni, D., Carlier, M. F. (2004). Formin is a processive motor that requires profilin to accelerate actin assembly and associated ATP hydrolysis. *Cell* 119, 419-429.

Rose, R., Weyand, M., Lammers, M., Ishizaki, T., Ahmadian, M. R., Wittinghofer, A. (2005). Structural and mechanistic insights into the interaction between Rho and mammalian Dia. *Nature* 435, 513-518.

Rosenberg, A.H., Goldman, E., Dunn, J.J., Studier, F.W. and Zubay, G. (1993) Effects of consecutive AGG codons on translation in Escherichia coli, demonstrated with a versatile codon test system. J. Bacteriol. *175*, 716–22.

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*, 4858–67.

Sagot, I., Rodal, A. A., Moseley, J., Goode, B. L., Pellman, D. (2002). An actin nucleation mechanism mediated by Bni1 and profilin. *Nat. Cell Biol*. 4, 626-631.

Sagot, I., Klee, S. K., Pellman, D. (2002). Yeast formins regulate cell polarity by the assembly of actin cables. *Nat. Cell Biol.* 4, 42-50.

Sakano, K., Enjoh, T., Numata, F., Fujiwara, H., Marumoto, Y., Higashihashi, N., Sato, Y., Perdue, J.F., and Fujita-Yamaguchi Y. (1991) The design, expression, and characterization of human insulin-like growth factor II (IGF-II) mutants specific for either the IGF-II/cation- independent mannose 6-phosphate receptor or IGF-I receptor J. Biol. Chem. 266, *31*, 20626-35.

Sala, A., Capaldi, S., Campagnoli, M., Faggion, B., Labo, S., Perduca, M., Romano, A., Carrizo, M. E., Valli, M., Visai, L., Minchiotti, L., Galliano, M. & Monaco, M. L. (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.

Sara, V.R., Carlsson-Skwirut, C., Anderson, C., Hall, E., Sjorgen, B., Holmgren, A, and Jornvall, H. (1986) Characterisation of somatomedins from human foetal brain: identification of a variant form of insulin-like growth factor I. Proc. Natl. Acad. Sci. USA, *83*, 4904-7.

Sato, H., Nishimura, S., Ohkubo, T., Kyoguku, Y., Koyama, S. and Kobayashi, Y. (1993) Three-dimensional structure of human insulin-like growth factor I (TGF-I) determined by ^{1}H-NMR and distance geometry, Int. J. Pept. Protein Res. *41*, 433–40.

Schagger, H., von Jagow, G. (1987) Tricine-sodium dodecyl sulfate-polyacrylamide gel electrophoresis for the separation of proteins in the range from 1 to 100 kDa. Anal Biochem. *166*, 368-79.

Schedlich, L.J., Young, T., Firth S., and Baxter, R. (1998) Insulin-like growth factor-binding protein (IGFBP9-3 and IGFBP-5 share a common nuclear trasport pathway in T47D human breast carcinoma cells. J. Biol. Chem. *273*, 18347-52.

Schedlich, L.J., O'Han, M.K., Leong, G.M., and Baxter, R.C. (2004) Insulin-like growth factor binding protein-3 prevents retinoid receptor heterodimerization: implications for retinoic acid-sensitivity in human breast cancer cells. Bioch. Bioph. Res. Comm. *314*, 83-8.

Schirenbeck, A., Bretschneider, T., Arasada, R., Schleicher, M., Faix, J. (2005). The Diaphanous-related formin dDia2 is required for the formation and maintenance of filopodia *Nat. Cell Biol.* 7, 619-625.
.
Schirenbeck, A., Arasada, R., Bretschneider, T., Stradal, T. E., Schleicher, M., Faix, J. (2006). The bundling activity of vasodilator-stimulated phosphoprotein is required for filopodium formation. *Proc. Natl. Acad. Sci. USA* 103, 7694-7699.

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*, 17859-66.

Shemesh, T., Otomo, T., Rosen, M. K., Bershadsky, A. D., Kozlov, M. M. (2005). A novel mechanism of actin filament processive capping by formin: solution of the rotation paradox. *J. Cell Biol.* 170, 889-893.

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*, 477–87.

Siwanowicz, I., Popowicz, G.M., Wisniewska, M., Huber, H., Kuenkele, K.P.Lang, K., and Holak, T. A. (2005) Structural basis mfor the regulation of insulin-like growth factors by IGF binding proteins. Structure. 13: 155-167.

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*, 454-7.

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,* 11022-29.

Vavylonis, D., Kovar, D.R., O'Shaughnessy, B., Pollard, T.D., (2006). Model of formin-associated actin filament elongation. Molecular Cell, 21: 455-466.

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*, 1677-85.

Watanabe, N., Kato, T., Fujita, A., Ishizaki, T., Narumiya, S. (1999). Cooperation between mDia1 and ROCK in Rho-iduced actin reorganization. *Nat. Cell Biol.* **1**, 136-143.

Wallar, B. J., Alberts, A. S. (2003). The formins: active scaffolds that remodel the cytoskeleton. *Trends Cell Biol.* 13, 435-446.

White, M.F. and Kahn, C.R. (1989) Cascade of autophosphorylation in the beta-subunit of the insulin receptor. J. Cell. Biochem. *39*, 429-41.

Wu, Y., Cui, K., Miyoshi, K., Hennighausen, L., Green, J.E., Setser, J., LeRoith, D., and Yakar, S. (2003) Reduced circulating insulin-like growth factor I levels delay the onset of chemically and genetically induced mammary tumors. Cancer Res. *63*, 4384–8.

Wüthrich, K. (1986) NMR of Proteins and Nucleic Acids, Wiley-Interscience, New York.

Yao, S.G., 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*, 11187-95.

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*, 3638-44.

Zigmond, S. H., Evangelista, M., Boone, C., Yang, C., Dar, A. C., Sicheri, F., Forkey, J., Pring, M. (2003). Formin leaky cap allows elongation in the presence of tight capping proteins. *Curr. Biol.* 13, 1820-1823.

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*, 4269–77.

Xu, Y., Moseley, J. B., Sagot, I., Poy, F., Pellman, D., Goode, B. L., Eck, M. J. (2004). Crystal structures of a formin homology-2 domain reveal a tethered dimer architecture. *Cell* **116**, 711-723.