

Analysis of the topology of the platform protein Gab1 and its interactions with the adaptor Grb2

Dissertation
zur Erlangung des
Doktorgrades der Naturwissenschaften (Dr. rer. nat.)

der

Naturwissenschaftlichen Fakultät I – Biowissenschaften –

der Martin-Luther-Universität
Halle-Wittenberg,



vorgelegt

von Katharina Mandel

geb. am 07.10.1985 in Bremen

Gutachter:

- 1: Prof. Hüttelmaier
- 2: Prof. Feller
- 3: Prof. Schaper

Tag der öffentlichen Verteidigung: 20.01.2020

Don't climb mountains so that
people can see you. Climb
mountains so that you can see the
world.

David McCullough Jr.

Table of contents

CHAPTER 1 INTRODUCTION

1.1	CELL SIGNALLING	1
1.2	ORGANISATION AND REGULATION OF MULTI-PROTEIN SIGNALLING COMPLEXES	2
1.3	SCAFFOLD PROTEINS AND RELATED PROTEINS IN SIGNALLING COMPLEXES	3
1.4	GRB2-ASSOCIATED BINDER (GAB) PROTEINS	5
1.4.1	THE GAB1 PROTEIN	5
1.4.2	THE GAB2 PROTEIN	9
1.4.3	THE GAB3 AND PUTATIVE GAB4 PROTEIN	9
1.5	THE GRB2 ADAPTOR PROTEIN	9
1.6	THE PHOSPHOINOSITIDE 3-KINASE (PI3K)	11
1.7	STRUCTURAL COMPOSITION OF GAB PARALOGUES	12
1.7.1	INTRINSICALLY DISORDERED PROTEINS (IDPS)	14
1.7.2	A GAB TOPOLOGY MODEL AND GAB1 IN SIGNAL COMPUTATION	15
1.8	GAB PROTEINS AND CANCER	17
1.9	AIM AND CONCEPT OF THIS STUDY	18

CHAPTER 2 MATERIAL AND METHODS

2.1	MATERIALS	19
2.1.1	EQUIPMENT	19
2.1.2	CONSUMABLES	20
2.1.3	REAGENTS AND CHEMICALS	21
2.1.4	PROTEASE AND PHOSPHATASE INHIBITORS	21
2.1.5	ANTIBODIES	21
2.1.6	BACTERIAL STRAINS AND EXPRESSION VECTORS	21
2.1.7	PEPTIDES	22
2.1.8	PROTEIN CRYSTALLOGRAPHY	23
2.1.9	MOLECULAR WEIGHT PROTEIN STANDARDS AND DNA KITS	23
2.1.10	SOLUTION AND BUFFERS	23
2.2	METHODS	28
2.2.1	BACTERIAL CULTURE AND EXPRESSION OF RECOMBINANT PROTEINS	28
2.2.2	PROTEIN SEPARATION AND DETECTION	33
2.2.3	ISOTHERMAL TITRATION CALORIMETRY (ITC)	34
2.2.4	X-RAY CRYSTALLOGRAPHY	36

CHAPTER 3 RESULTS AND DISCUSSION: PURIFICATION OF THE GAB2 PROTEIN

3.1	EXPRESSION AND PURIFICATION OF RECOMBINANT GAB2 PROTEIN	38
3.2	CONCLUSION AND DISCUSSION	41

CHAPTER 4 RESULTS AND DISCUSSION: PURIFICATION OF THE P35GAB1 – GRB2 COMPLEX

4.1	OVERVIEW	42
4.2	ESTABLISHING A PURIFICATION PROTOCOL FOR THE P35GAB1 – GRB2 COMPLEX	43
4.2.1	IMMOBILIZED METAL AFFINITY CHROMATOGRAPHY (IMAC) OF THE P35GAB1 – GRB2 COMPLEX	43
4.2.2	ANION EXCHANGE CHROMATOGRAPHY OF THE P35GAB1 – GRB2 COMPLEX	46
4.2.3	‘POLISHING’ STEP OF THE P35GAB1 – GRB2 COMPLEX PURIFICATION	48
4.3	TEST PURIFICATION OF THE P35GAB1 FRAGMENT ALONE	49
4.4	CONCLUSION AND DISCUSSION	50

CHAPTER 5 RESULTS AND DISCUSSION: ANALYSIS OF THE P35GAB1 – GRB2 COMPLEX

5.1	OVERVIEW	52
5.2	ANALYTICAL ULTRACENTRIFUGATION, ATOMIC FORCE MICROSCOPY AND ELECTRON MICROSCOPY OF THE P35GAB1 – GRB2 COMPLEX	52
5.3	IDENTIFYING INTERACTION SITES IN THE P35GAB1 – GRB2 COMPLEX BY CROSS-LINKING MASS SPECTROMETRY (XL-MS)	55
5.4	GRB2 SH3 DOMAIN-BINDING TANDEM MOTIF IN P35GAB1 (GAB1)	60
5.5	GAREM1 PROTEIN CONTAINS A SIMILAR TANDEM MOTIF AND INTERACTS WITH GRB2	62
5.6	ANALYSIS OF THE GAB1 – GRB2 INTERACTION BY X-RAY CRYSTALLOGRAPHY	64
5.6.1	OVERVIEW	64
5.6.2	GRB2 PURIFICATION	65
5.6.3	CRYSTAL SCREENS FOR GRB2 WITH P35GAB1 AND P35GAB1-DERIVED PEPTIDES	66
5.6.4	FINE SCREENING AND STRUCTURE DETERMINATION OF A GRB2 – GAB1 PEPTIDE (32 AA) CRYSTAL	68
5.6.5	OVERALL STRUCTURE OF THE GRB2 – GAB1 PEPTIDE COMPLEX	71
5.6.6	STRUCTURAL FEATURES OF THE GAB1 PEPTIDE – GRB2 SH3N INTERACTION	72
5.6.7	STRUCTURE OF THE GAB1 PEPTIDE – GRB2 SH3C DOMAIN COMPLEX	77
5.6.8	STRUCTURAL FEATURES OF THE GRB2 SH2 DOMAIN IN THE GRB2 – GAB1 CRYSTAL	80
5.6.9	GRB2 – GAB1 OLIGOMERISATION MODEL	82
5.7	CONCLUSION AND DISCUSSION	84

**CHAPTER 6 RESULTS AND DISCUSSION: ASSEMBLY OF A PROTEIN COMPLEX
COMPRISED OF P35GAB1, GRB2 AND THE PI3K REGULATORY SUBUNIT P85A**

6.1	OVERVIEW	86
6.2	PROTEIN PURIFICATION AND AFM ANALYSIS OF PI3K P85A	87
6.3	TYROSINE PHOSPHORYLATION OF P85 BINDING SITES IN P35GAB1	89
6.4	PRELIMINARY ANALYSIS OF THE ASSEMBLED P35GAB1 – GRB2 – P85 COMPLEX	90
6.5	CONCLUSION AND DISCUSSION	92

CHAPTER 7 FINAL CONCLUSIONS AND FUTURE STRATEGIES **93**

REFERENCES **96**

APPENDICES **I**

TABLE OF FIGURES AND TABLES **VI**

LIST OF ABBREVIATIONS **VIII**

EIDESSTATTLICHE ERKLÄRUNG **X**

LIST OF PUBLICATIONS **XI**

ACKNOWLEDGEMENT **XII**

CURRICULUM VITAE **XIII**

Chapter 1 Introduction

1.1 Cell signalling

Many bacteria can respond to signals secreted from other bacteria or sense nutrients in their environment which direct them towards food sources. This ability is based on cell signalling processes which provide the cell with the ability to rapidly and specifically respond to changes in its environment. During evolution from a unicellular to a multi-cellular organism a well-coordinated communication system emerged offering an information transfer over small and large distances. Examples are the endocrine system where hormones deliver information to cells in distant sites of the human body or the human central nervous system where electric signals are being transmitted between neighbouring cells. Signal transduction is provided by a variety of receptors and their corresponding ligands including peptides, lipids, nucleotides, proteins or chemical substances. Receptors are highly specialised sensors and are either located intracellularly or on the cell surface. The three main classes of receptors are G protein-coupled receptors (GPCR), ion channel receptors and enzyme-coupled receptors, for example receptor tyrosine kinases (RTKs) (Hille, 2001, Lemmon and Schlessinger, 2010, Audet and Bouvier, 2012). Intracellular receptors respond to ligands such as lipophilic or low molecular weight signalling molecules which can cross the plasma membrane. The three main stages of a simplified cell signalling pathway: signal reception, signal integration and response are illustrated in Figure 1.1.

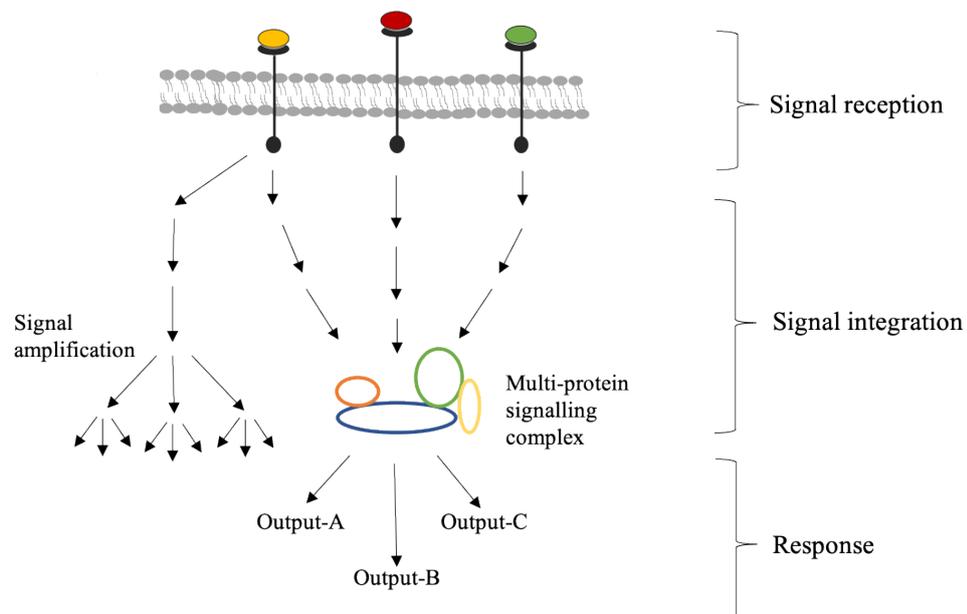


Figure 1.1 Schematic overview of a simplified cell signalling pathway

Extracellular signal molecules activate cell surface receptors and trigger a signalling cascade. Signalling proteins assemble to multi-protein signalling complexes and integrate the incoming signal into coordinated cellular processes including cell growth, survival and migration.

The signalling process is initiated by the binding of the cell surface receptor to a ligand (signal reception). Then, the receptor activation leads to the relocation of signalling proteins in close proximity to the receptor. Signals which are transmitted from various receptors at the same time, are being integrated by signalling proteins and/or signalling multi-protein complexes and distinct signalling pathways become activated (signal integration). In some cases, signals can become amplified by molecules known as second messengers, for example cyclic AMP. Signalling pathways regulate cellular processes such as cell survival, cell migration or proliferation (response).

Cells contain numerous signalling pathways which are arranged in signalling networks. These highly-coordinated signalling networks manage a precise signal integration and coordination of a multitude of information. Multi-protein signalling complexes in particular strongly contribute to the signal integration. They are described in the following subchapter 1.2.

1.2 Organisation and regulation of multi-protein signalling complexes

Multi-protein signalling complexes are assembled by protein-protein interactions (PPI). Proteins contain protein recognition modules, the protein domains to interact with other proteins. Protein domains often encompass a consensus amino acid sequence and characterised by a distinct fold (Yang and Bourne, 2009). The ligand binding site, which is responsible for specific ligand recognition, is located on the domain surface. Protein interactions can have dissociation constants in the nanomolar (permanent interaction) to high micromolar range (weak transient interaction) (Perkins et al., 2010). Frequently found domains in PPIs are the Src homology 3 (SH3) domains (ca. 300 in human genome (Karkkainen et al., 2006)) which commonly bind to pro-rich regions and pleckstrin homology (PH) domains enabling proteins to bind the plasma membrane. For example, the N-terminal PH domain of the large multisite docking protein Gab1 of the Gab family binds to phosphoinositides within the membrane and enables Gab1 proteins to interact with nearby membrane receptors (Maroun et al., 1999). Another common protein domain is the Src homology 2 (SH2) domain. It recognises protein surface epitopes with a central phosphorylated tyrosine residue (Yaffe, 2002, Kaneko et al., 2012). Within cell signalling processes, the posttranslational modification of protein binding sites by phosphorylation serves as an important regulatory mechanism. Identified protein domains have been deposited in various databases, e.g. the EMBL database SMART (smart.embl-heidelberg.de).

PPI are important in almost every signalling process and need to be generally controlled by mechanisms ensuring efficient and regulated signalling. One regulatory system in PPI is allostery (Hilser et al., 2012). It describes the modification of the activity or function of a protein, often enzymes, by binding of a regulatory ligand to another side ('allos stereos' = 'another object') outside of the primary protein functional site. The effector binding induces a conformational change of the functional site in the protein either inhibiting (negative allostery) or activating (positive allostery) it to bind to other interaction partners. Signalling processes can also be regulated by a simple on/off switch mechanism,

such as the phosphorylation of SH2 domain binding sites to initiate ligand binding. But, phosphorylation can also occur on multiple protein binding sites leading to a rheostat or ultrasensitive response (Tyson et al., 2003). The successive phosphorylation either additively increases (positive rheostat) or decreases (negative rheostat) the affinity of the protein to bind to another interaction partner. For example, in case of the important tumour suppressor protein p53, the cumulative phosphorylation of p53 increases the affinity to its interaction partner CBP/p300 (positive rheostat) (Lee et al., 2010). An ultrasensitive response can be based on the principle that a distinct number of phosphorylation sites (threshold) needs to be phosphorylated in order to trigger the downstream signalling cascade. Sequential multisite phosphorylation is required in yeast cell cycle regulation for instance. The cyclin-dependent kinase inhibitor Sic1 needs to be phosphorylated on any six (or more) of nine cyclin-dependent kinase sites to interact with Cdc4 leading to Sic1 ubiquitination and degradation (Borg et al., 2007).

PPIs are not only carefully regulated but also prominently affected by many intracellular factors. The cell is comprised of a densely-packed environment with approx. 40% of the total cell volume occupied by macromolecules (Ellis and Minton, 2003) whereby the macromolecule concentration in the cytoplasm of mammalian cells can reach 200-300 mg/ml (Luby-Phelps, 2000). This macromolecular crowding inhibits a free diffusion of proteins which means protein components of a pre-assembled signalling complex need to be temporally and spatially well-organized. Another factor influencing the formation of signalling complexes or in general protein-protein interactions is the compartmentalisation of the cell (Lewitzky et al., 2012). The cell contains various well-defined organelles with functional specialisation. Cellular organelles are classified as membrane-containing compartments such as the Golgi apparatus or ER and membrane-less compartments including nucleoli and nuclear speckles. Membrane-less compartments seem to be enormously dynamic assemblies of macromolecules which are formed by a still marginally understood phase separation mechanism (Mitrea and Kriwacki, 2016). The location of proteins which are part of a complex is therefore tightly regulated in order to co-localise them for complex assembly (Shiber et al., 2018).

Signalling complexes often contain proteins with scaffolding functions and adaptor proteins. The following subchapter 1.3 introduces the group of scaffold proteins and related proteins.

1.3 Scaffold proteins and related proteins in signalling complexes

Scaffold proteins and related proteins contain several modular interaction domains or motifs enabling them to tether interacting components in a signalling complex. Through their modular composition scaffold proteins can localize pathway specific enzymes in close contact to other pathway components or present enzymes close to their specific substrates. Therefore, they are contributing to the specificity of signal transduction events. Compared to many other signalling proteins which possess enzymatic activity, scaffold proteins lack any enzymatic functions.

Scaffold proteins can be separated into three main categories regarding their functional modes: adaptor proteins (A), scaffold/anchoring proteins (B) and docking proteins (C) (Buday and Tompa, 2010) (Figure 1.2). Adaptor proteins are small-sized proteins which usually have two domains involved in protein – protein interactions or use two regions located in two to three domains (Figure 1.2, A). The growth factor receptor-bound protein 2 (Grb2) or Crk/CrkL proteins are prominent examples for adaptor proteins (Maignan et al., 1995, Feller, 2001).

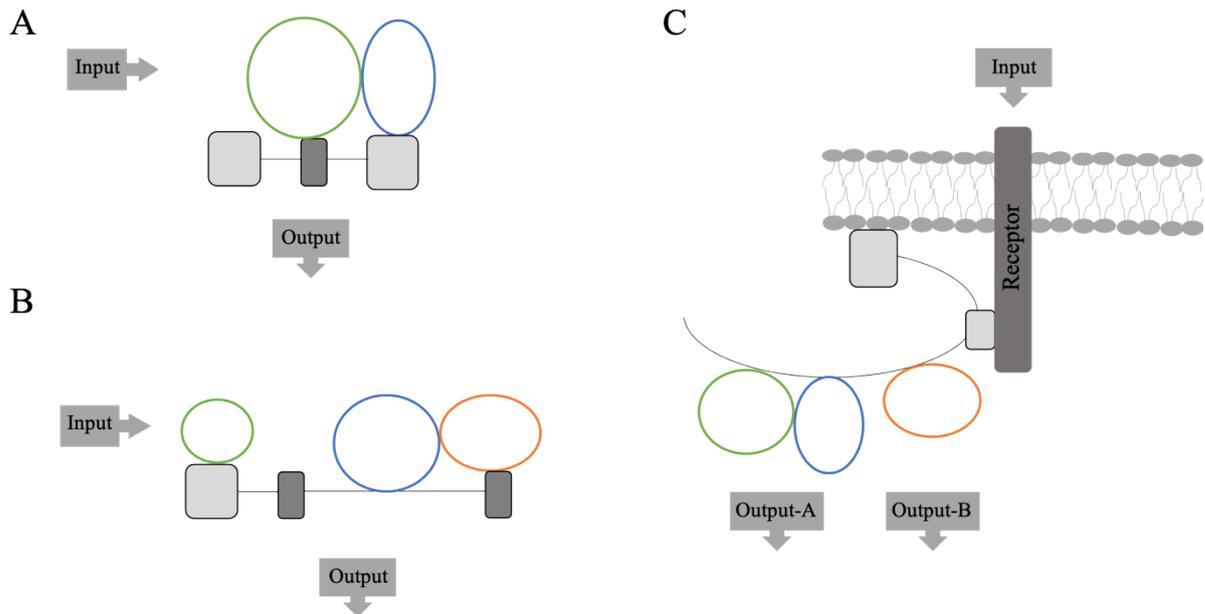


Figure 1.2 Various scaffold proteins or related proteins in cell signalling

There are three main categories of scaffold proteins or related proteins. (A) Adaptor proteins (grey), here shown with two interacting proteins. (B) Scaffold/anchoring proteins (grey) contain more than two interacting protein domains enabling them to interact with at least two signalling proteins. (C) Docking proteins (grey) are very similar composed as scaffold/anchor proteins but they can bind to the plasma membrane localizing them in close contact to membrane receptors. Figure was modified from Buday and Tompa, 2010.

Scaffold/anchoring proteins are composed of several protein domains or contain multiple protein binding motifs which can link at least two proteins in close proximity (Figure 1.2, B). It is believed that scaffold/anchoring proteins provide an active role in regulatory functions as fine-tuning or feedback mechanisms additionally to their linker functions (Shaw and Filbert, 2009). Docking proteins or large multisite docking proteins (LMD) are similar to scaffold/anchoring proteins in their ability to interact with several proteins but distinguishable in their ability to bind to the plasma membrane in order to be in direct contact with activated receptors (Figure 1.2, C).

The Grb2-associated binder (Gab) protein involved in many cellular signalling processes is a prominent example for LMD proteins (Wöhrle et al., 2009) and is described in more detail in the next subchapter 1.4.

1.4 Grb2-associated binder (Gab) proteins

The Grb2-associated binder (Gab) family of proteins belongs to the class of large multisite docking proteins (LMD) (Liu and Rohrschneider, 2002, Gu and Neel, 2003). Gab proteins play a significant role in many signalling cascades such as cytokine receptor, multichain immune recognition receptor or receptor tyrosine kinase pathways (Wöhrle et al., 2009). Gab proteins provide a docking platform for multiple proteins in order to form large, dynamically composed signalling complexes. Within these complexes, Gab proteins are believed to serve as a signal integration and processing platform coordinating signals to activate the corresponding signalling pathway (Simister and Feller, 2012). Well-characterised Gab protein orthologues such as Daughter of Sevenless (DOS) in the fruit fly *Drosophila melanogaster* and Suppressor of Clear (SOC) 1 in the nematode *Caenorhabditis elegans* are important key proteins in RTK signalling pathways (Herbst et al., 1996, Raabe et al., 1996, Schutzman et al., 2001).

Vertebrates have at least three Gab protein paralogues such as Gab1, Gab2 and Gab3 (Holgado-Madruga et al., 1996, Gu et al., 1998, Wolf et al., 2002, Abbeyquaye et al., 2003). A fourth member, the putative Gab4 protein, has only been identified on the DNA and transcript level so far. In contrast, invertebrates such as *Drosophila melanogaster* or *Caenorhabditis elegans* possess only one *gab* gene (Herbst et al., 1996, Schutzman et al., 2001). A *gab* gene sequence analysis revealed that genes from the *gab1*, *gab2* and *gab3* proteins diverged from a common ancestral *gab* gene with the emergence of early vertebrates which can be dated back about 500 million years (Dr. Davey, University College Dublin, personal communication). It appears that the *gab3* gene diverged first, slightly before the *gab1* and *gab2* genes. The putative *gab4* gene seems to have evolved from the *gab2* protein gene concurrent with the emergence of primates (about 70 million years ago). The following subchapters 1.4.1 - 1.4.3 provide a detailed description of each Gab family paralogue.

1.4.1 The Gab1 protein

The Gab1 protein (Gab1 isoform 1, aa 1 – aa 694; ca. 100 kDa; UniProt accession code: Q13480, aa sequence in appendix A.1) was the first member of the Gab family of proteins being identified (Holgado-Madruga et al., 1996). Gab1 was discovered in a human glial tumour as a target of tyrosine phosphorylation upon epidermal growth factor (EGF) or insulin stimulation. A human Gab1 isoform 2 (UniProt accession code: Q13480-2) has recently been identified (Yasui et al., 2016). Expression of the Gab1 isoform 1 protein is ubiquitously whereas the Gab1 isoform 2 protein is exclusively found in the heart. The Gab1 isoform 2 contains an extra in-frame coding exon resulting from alternative splicing. The additional exon consists of 30 amino acids (aa) and is located slightly downstream of the higher affinity Grb2 SH3 binding site (Harkiolaki et al., 2009). A biological function of the additional exon in Gab1 isoform 2 has not been examined yet. Also, there is a Gab1 variant that lacks most of the N-terminal PH domain implicated in the progression of tumours in Syrian hamster embryo cells (Kameda

et al., 2001). Recently, a study discovered that this Gab1 variant plays a role in LIF-mediated (leukaemia inhibitor factor) cell survival in mouse embryonic stem cells when cells are cultivated under limited nutrient availability (Sutherland et al., 2018).

The full length Gab1 isoform 1 protein (UniProt accession code: Q13480), referred to as Gab1 protein, was exclusively used in this thesis.

The Gab1 protein is implicated in signalling via RTKs such as the insulin or the c-Met receptor or non-RTKs, e.g. the B cell or EPO receptor (Liu and Rohrschneider, 2002). The c-Met pathway is highly important for processes such as embryonic development, wound healing and organ regeneration (Trusolino et al., 2010). Essential functions of Gab1 in embryonic development, related to its critical role in the c-Met pathway, were demonstrated by a *gab1* gene knockout (KO) mouse model exhibiting a strong phenotype (Itoh et al., 2000, Sachs et al., 2000). Mice impaired to express Gab1 died prenatal at day 14-18 following gestation. *Gab1*^{-/-} mice showed a reduced liver size, an impaired placental and heart development together with defects in muscles, skin of the limbs and in the diaphragm. The severe *Gab1*^{-/-} phenotype suggests a non-redundant Gab1 function in several Gab1-based signalling pathways (Itoh et al., 2000).

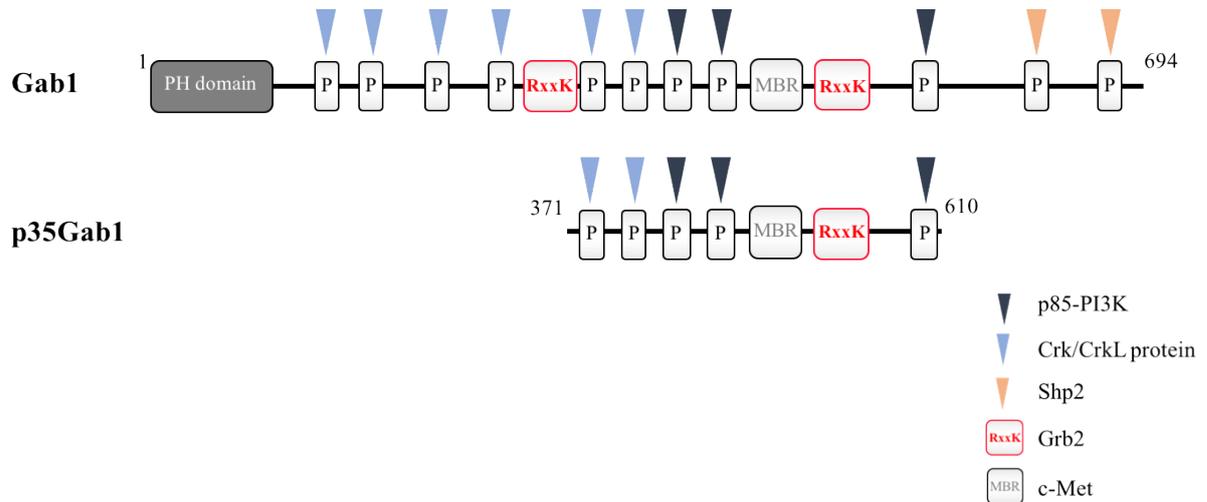
Not only full-length Gab1 occurs in cell, but also a physiological active Gab1 fragment of around 35 kDa (p35Gab1) (Le Goff et al., 2012) which is described in more detail in the following subchapter 1.4.1.1.

1.4.1.1 The p35Gab1, a shorter fragment of full-length Gab1

The p35Gab1 is a shorter but still physiologically active fragment of full-length Gab1 (Gab1, aa 371 – aa 610, aa sequence in appendix A.1). It was identified as a caspase cleavage product of full-length Gab1 playing an important role in the c-Met pathway when the cell is under moderate stress. Le Goff et al. could show that p35Gab1 maintains the survival pathway via c-Met under moderate stress conditions by preserving the PI3K interaction (Le Goff et al., 2012). The p35Gab1 fragment contains the c-Met receptor binding region (MBR), one of the Grb2 binding site, all three PI3K and two out of six Crk/CrkL binding sites (Figure 1.3, A). A sequence alignment of full-length Gab1 shows that the caspase cleavage sites for p35Gab1 are highly conserved throughout the evolution of tetrapods (Figure 1.3, B). Cleaving sites of caspases are usually characterized by a sequence of four amino acids with a terminal aspartate after which the caspase cleaves the protein.

It has not been examined yet whether the Gab1 isoform 2 is similarly cleaved in response to stress as shown for Gab1. If a Gab1 isoform 2 fragment is generated by caspases cleaving at the p35Gab1 cleavage sites as found in Gab1, it would contain the extra exon and would be slightly heavier than the p35Gab1 fragment from the Gab1 isoform 1.

A



B

	370	610
<i>H. sapiens</i>	SDTDS	NSLDG
<i>B. taurus</i>	SDTDS	NSLDG
<i>M. musculus</i>	SDTDS	NSLDG
<i>M. auratus</i>	SDTDS	NSLDG
<i>X. laevis</i>	SETDS	SSLDE
<i>D. rerio</i>	SETDT	MNADG
	* : ** :	. *

Figure 1.3 The p35Gab1 fragment (aa 371 – aa 610) of full-length Gab1 (aa 1 – aa 694)

Full-length Gab1 is cleaved by caspases to a functional fragment, the p35Gab1 (Caspase cleave sites: after Asp370 and after Asp610). (A) The full length Gab1 contains a PH domain, Grb2 SH3C domain binding sites (R-x-x-K), a c-Met binding region (MBR) and various binding sites for signalling proteins. The p35Gab1 fragment still contains the MBR, p85-PI3K, two out of six Crk/CrkL protein binding sites and one of the two Grb2 binding sites. The P in the box charts the phosphorylated tyrosine residue of the SH2 binding motif. (B) Gab1 sequence alignment of various species in respect of the p35Gab1 caspase cleavage sites in Gab1 (caspase cleaves after Asp370 and Asp610, indicated in bold letters) (Figure modified from Le Goff et al., 2012). UniProt accession code of Gab1 sequences: Q13480 (*Homo sapiens*); A6QLU3 (*Bos taurus*); Q9QYY0 (*Mus musculus*); Q99PF6 (*Mesocricetus auratus*); Q6AZI1 (*Xenopus laevis*) and B8A4S9 (*Danio rerio*). Single fully conserved residues are marked with an asterisk (*). A colon (:) indicates a conservation between groups of strongly similar properties and a period (.) denotes a conservation between groups of weakly similar properties.

As mentioned earlier, the full-length Gab1 protein as well as the p35Gab1 fragment play important roles in the c-Met pathway which is described in more detail in the following subchapter 1.4.1.2.

1.4.1.2 Gab1 and the c-Met signalling pathway

It has been shown that Gab1 plays a pivotal role in the c-Met signalling pathway (Maroun et al., 1999, Sachs et al., 2000). The c-Met pathway promotes main cellular processes such as morphogenesis, motility and cell survival, and plays an important role in embryonic development (Stoker et al., 1987,

Montesano et al., 1991, Schmidt et al., 1995, Uehara et al., 1995). The c-Met receptor becomes activated by binding to the matured form of HGF/SF (hepatocyte growth factor/scatter factor) (Stoker and Perryman, 1985, Bottaro et al., 1991) (Figure 1.4). Gab1 is recruited to the membrane and binds either directly to the cytoplasmic domain of the c-Met receptor via the c-Met binding region (MBR) or indirectly via Grb2 (Weidner et al., 1996, Bardelli et al., 1997). The Gab PH domain interacts with the cell membrane by binding to plasma membrane-located PIP₃. The group of Prof. Schaper suggested that the membrane recruitment of Gab1 is regulated by a specific phosphorylation event in Gab1 (Eulendorf and Schaper, 2009). They postulate that Gab1 is retained in the cytoplasm in unstimulated cells due to a Gab1 PH domain block by the proteins topological arrangement itself. Upon HGF stimulation, the PH domain block is released by a kinase-induced phosphorylation (e.g. Erk1/2) on residue Ser552 (murine Gab1) and Gab1 is translocated to the cytoplasmic membrane (Eulendorf and Schaper, 2009, Bongartz et al., 2017).

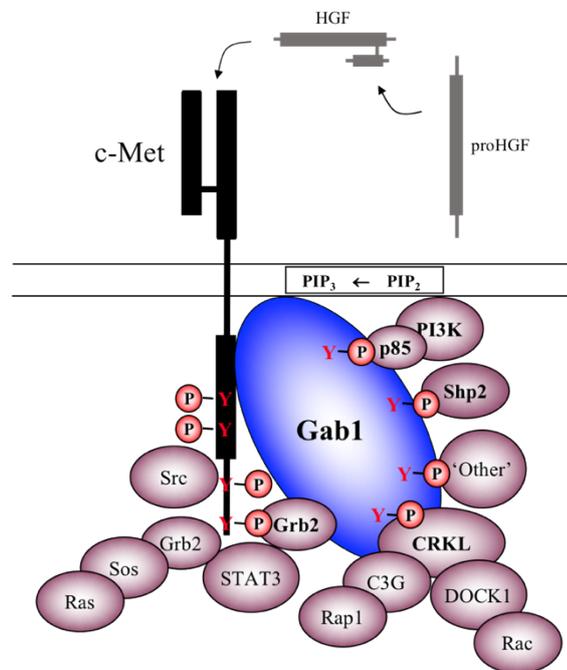


Figure 1.4 Simplified overview of the Gab1 protein in the c-Met pathway

Overview of c-Met signalling pathway. Upon c-Met receptor activation by HGF, Gab1 and Grb2 are recruited and activate downstream signalling molecules. In turn, cellular processes like proliferation or migration become activated. Tyrosyl phosphorylation sites (Y-P) in consensus binding motifs in Gab1 serve as a binding site for SH2 domain containing proteins. Direct interaction partner proteins of Gab1 are indicated in bold. Figure adapted from Furge et al. (Furge et al., 2000).

When coupled to the c-Met receptor, the Gab1 protein becomes phosphorylated by the c-Met tyrosine kinase domain, thereby “activating” SH2 binding sites in Gab1 for Crk or Crk-like (CrkL) adaptor proteins, phosphatase Shp2 (PTPN11) and the PI3K regulatory subunit p85 (Graziani et al., 1991, Garcia-Guzman et al., 1999, Cunnick et al., 2001, Gu and Neel, 2003). The assembled Gab-based signalling complex subsequently coordinates incoming signals and regulates processes such as cell

survival, proliferation and migration. Cell survival is mainly controlled by the PI3K/AKT pathway via the PI3K-AKT interaction (Fan et al., 2001, Xiao et al., 2001). Signalling via Ras superfamily GTPase Rap1 or the GTPase Rac mainly promotes cell migration and cell adhesion (Hordijk et al., 1997, Price et al., 2004).

1.4.2 The Gab2 protein

The Gab2 protein has similar functions as the Gab1 protein and acts downstream of various receptors, such as the B and T cell receptor, growth factor receptors and cytokine receptors (Gu et al., 1998, Nishida et al., 1999, Fan et al., 2001, Liu et al., 2001, Nishida et al., 2002, Yamasaki et al., 2003), but is more cell type-restricted in its expression. Gab2-deficient mice (Gab2^{-/-}) show a less severe phenotype than mice with a Gab1 KO (Gu et al., 2001, Nishida et al., 2002). They exhibit a normal development with no immediately obvious defects. However, detrimental effects in the hematopoietic lineage were observed on the cellular level upon closer inspection. The total number of mast cells was reduced and their responsiveness was impaired. This is most likely due to the essential role of Gab2 in the c-Kit signalling pathway which is highly relevant in mast cell development (Gu et al., 2001, Nishida et al., 2002, Gonzalez-Espinosa et al., 2003). Gab2^{-/-} mice also demonstrated an impaired osteoclastogenesis and phagocytosis as well as a reduced responsiveness of hematopoietic progenitor cells (Gu and Neel, 2003, Gu et al., 2003, Wada et al., 2005, Zhang et al., 2007).

1.4.3 The Gab3 and putative Gab4 protein

The Gab3 protein has been implicated in the macrophage differentiation process (Wolf et al., 2002). A Gab3 KO mouse model (Gab3^{-/-}) showed a normal development and no obvious phenotype compared to WT mice (Seiffert et al., 2003). Although Gab3 expression was restricted to hematopoietic murine tissue, especially lymphocytes and bone marrow-derived macrophages, the deficiency in the Gab3 expression caused no detectable defects in the murine immune system. The putative Gab4 protein was the last paralogue of the Gab family of proteins identified (Wöhrle et al., 2009). So far, a *gab4* gene has been only found in humans and chimpanzees and the characterization on the protein level still remains to be done. Gab proteins function as assembly platforms in signalling pathways and have many interaction partners such as the Grb2 adaptor protein or the phosphoinositide 3-kinases (PI3K) as demonstrated in the c-Met pathway (see subchapter 1.4.1.2). The Gab1 - Grb2 and Gab1 - PI3K interaction play a crucial role in the c-Met pathway (Schaeper et al., 2007) and are, therefore, discussed in more detail in the following subchapters 1.5 and 1.6.

1.5 The Grb2 adaptor protein

The growth factor receptor-bound protein 2 (Grb2) is an adaptor protein which plays an important role in cell signalling pathways (Brummer et al., 2010). The Grb2 protein links for example receptors with

downstream effector proteins. Grb2 is structurally composed of an N- and C-terminal Src-homology 3 (SH3) domain (SH3N; SH3C) which are separated by a single Src-homology 2 (SH2) domain (Maignan et al. 1995) (Figure 1.5). The Grb2 SH2 domain recognises binding epitopes with a pY-x-N-x consensus sequence (Ogura et al., 1999) whereas the Grb2 SH3 domain bind to pro-rich regions (Goudreau et al., 1994, Harkiolaki et al., 2009). Domains in Grb2 are connected by short linkers providing intramolecular protein flexibility which is necessary for target recognition (Yuzawa et al., 2001).

The Grb2 SH2 domain has a typical SH2 domain core structure which consists of a central anti-parallel β -sheet and two flanking α -helices (Rahuel et al., 1996, Thornton et al., 1996, Ogura et al., 1999, Nioche et al., 2002). The N- and C-terminal SH3 domains of Grb2 are, like many other SH3 domains, composed of five anti-parallel β -strands that form two perpendicular β -sheets. The RT-, the n-Src- and the distal loop connect the first three β -strands whereas the last two β -strands are linked by a 3_{10} helix (Musacchio et al., 1992, Yu et al., 1992).

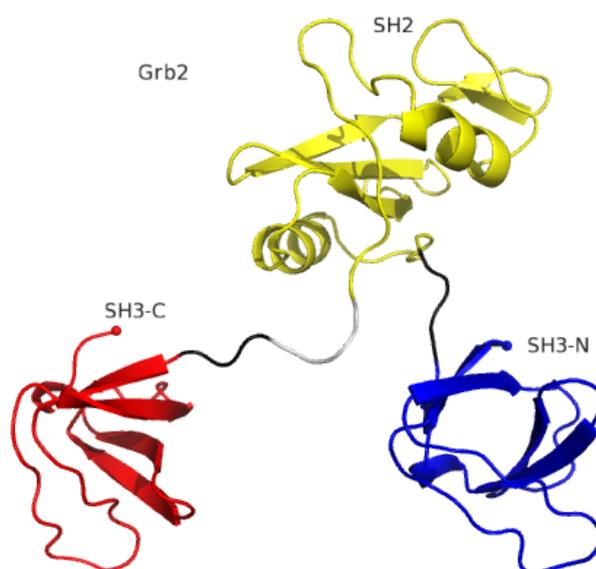


Figure 1.5 Grb2 ribbon model

Structural model of human Grb2 (C32S, C198A) (Grb2 – Gab1 peptide structure, see subchapter 5.6, Gab1 ligand is not shown). Grb2 is composed of one central SH2 domain (yellow) and two flanking SH3 domains (SH3C: red, SH3N: blue). Grb2 N- and C-terminus are presented as dots. C-terminal Grb2 residues 212 – 217 are not shown as they could not be built into the model. Grey coloured regions in Grb2 represent residues which could not be built into the model.

In the literature, functional Grb2 has been described as a Grb2 monomer as well as a Grb2 dimer. Full-length Grb2 (WT) was initially crystallized as a Grb2 dimer (Maignan et al., 1995). Later, NMR analysis of Grb2 suggested that Grb2 (C32S, C198A) exists in a monomeric form with flexible inter-domain linkers (Yuzawa et al., 2001, Yuzawa et al., 2003). Further studies showed that Grb2 occurs in a continuous Grb2 monomer - dimer equilibrium. A Grb2 dimer dissociation constant of approximately 0.7 - 5 μ M was determined (McDonald et al., 2008a, Ahmed et al., 2015). To define the Grb2 dimerisation process in more detail various studies have been performed. Grb2 residues Phe61, Phe182 and Arg207 which are located directly at the dimer-interface showed no critical role in Grb2 dimer

association (McDonald et al., 2008a). However, Grb2 dimerisation was abolished by either tyrosine phosphorylation of Tyr160_{Grb2} or by an SH2 domain binding phospho-tyrosine peptide (Ahmed et al., 2015). Until now, exact details of the underlying mechanisms of the Grb2 dimerisation as well as *in vivo* functions of Grb2 as a monomeric and/or dimeric protein, have not been fully elucidated yet.

Grb2 interacts with many different proteins, including the large multisite docking protein Gab1 (Holgado-Madruga et al., 1996). In signalling pathways, Grb2 couples Gab1 to a range of different receptors such as the c-Met receptor, the EGF receptor, the FGF receptor or the PDGF receptor (Lock et al., 2000, Ong et al., 2001, Liu and Rohrschneider, 2002, Gu and Neel, 2003, Kallin et al., 2004). Several studies have shown that the Grb2 – Gab1 interaction is constitutive (Holgado-Madruga et al., 1996, Schaeper et al., 2000).

A crystal structure of a single Grb2 SH3C domain complexed with a Gab2 peptide identified two Grb2 SH3C binding motifs in Gab2 with a core R-x-x-K motif. The two Grb2 interaction sites in Gab2 are distinct from each other and defined by their 10-fold difference in binding affinity ($K_{d \text{ Grb2 SH3C}}$: 3 μM vs. $K_{d \text{ Grb2 SH3C}}$: 30 μM). A sequence alignment of the Gab family of proteins in humans showed that the two Grb2 SH3C binding sites are conserved in Gab1, Gab2 and Gab3 (Harkiolaki et al., 2009).

1.6 The phosphoinositide 3-kinase (PI3K)

Phosphoinositide 3-kinases (PI 3-kinases or PI3Ks) belong to the family of intracellular lipid kinases that regulates major cellular processes such as cell growth, motility and metabolism (Engelman et al. 2006). PI3Ks are composed of a regulatory and a catalytic subunit and are subdivided into classes (I-III) based on their substrate preference and subunit composition (Vanhaesebroeck et al., 2012).

PI3Ks become activated by various receptors including RTKs or GPCRs and catalyse the formation of specific phosphoinositides. The heterodimer PI3K (Class IA), composed of a p110 catalytic and a p85 regulatory subunit, converts PIP₂ (phosphatidylinositol-4,5-bisphosphate) into the second messenger PIP₃ (phosphatidylinositol-3,4,5-triphosphate) (Backer, 2010, Vanhaesebroeck et al., 2012). PI3K signalling is generally terminated by the lipid phosphatase PTEN. Aberrations in PI3K signalling, for example due to mutations in p85 or p110, are found in several human cancer types (Vanhaesebroeck et al., 2012).

The Gab1/2 protein contains three SH2 domain consensus binding motifs (pY-x-x-M) for the PI3K subunit p85 (human Gab: Y447, Y472, Y589 and mouse Gab: Y448, Y473, Y590) (Holgado-Madruga et al., 1997, Ong et al., 2001, Gu and Neel, 2003, Liu and Rohrschneider, 2002, Simister and Feller, 2012). The interaction between Gab1 and the PI3K is important in developmental processes, in particular for eyelid closure and keratinocyte migration, as demonstrated by a mutant Gab1 protein in mice deficient in recruiting PI3K (Y448F/Y473F/Y590F) (Schaeper et al., 2007). A more indirect Gab/PI3K interaction is mediated by the Gab PH domain and the PI3K-generated product PIP₃ (Rameh et al., 1997, Rodrigues et al., 2000). After Gab recruitment to the plasma membrane, the N-terminal Gab

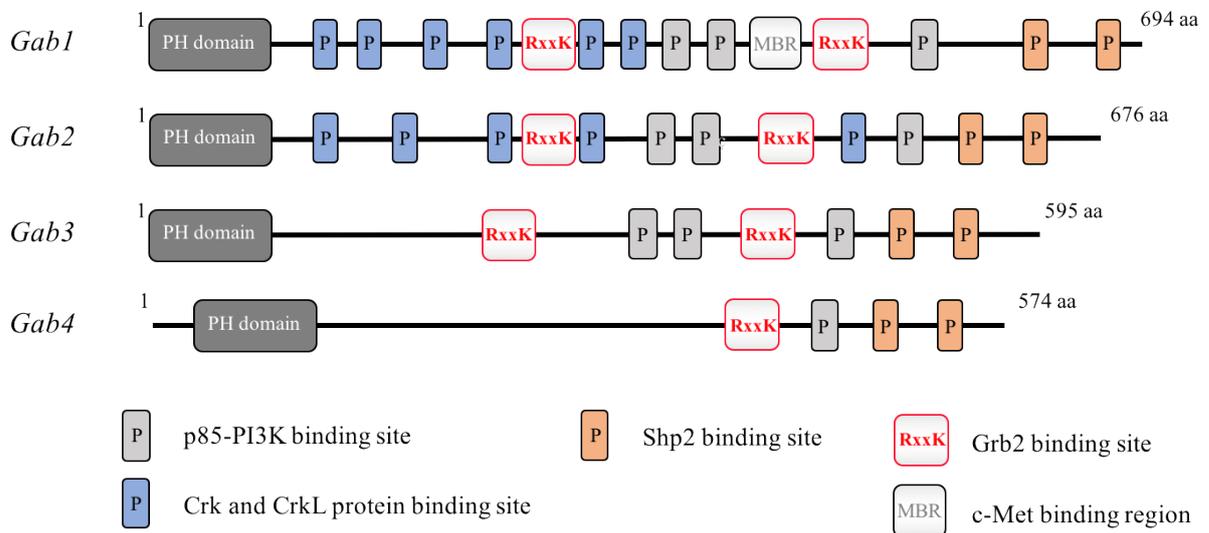
PH domain interacts with PIP₃ and tethers Gab in close proximity to activated receptors such as the c-Met receptor (Maroun et al., 1999, Rodrigues et al., 2000).

Gab proteins are involved in various pathways and interact with a series of different signalling proteins. Gabs structural composition, in particular their binding sites for interaction partners, is outlined in the following subchapter 1.7.

1.7 Structural composition of Gab paralogues

The determination of the protein structure is often very helpful to understand the actual protein function. Although Gab family members only share a 40-50% sequence identity, they possess a very similar topology marked by a well-folded N-terminal pleckstrin homology (PH) domain and a long protein ‘tail’ with multiple protein – protein interaction sites (Figure 1.6, A). The protein ‘tail’ of Gab contains several binding sites for signalling proteins including the phosphoinositide 3-kinase PI3K, the phosphatase Shp2 and adaptor proteins Grb2 and Crk/CrkL (Simister and Feller, 2012). Comparing the structural composition within the Gab family of proteins, differences are obvious with respect to the protein length, the number of protein interaction sites (p85-PI3K or Grb2) or the presence of specific binding motifs (Crk/CrkL) (Figure 1.6, A).

A



B

Human Gab1 (UniProt accession code: Q13480)

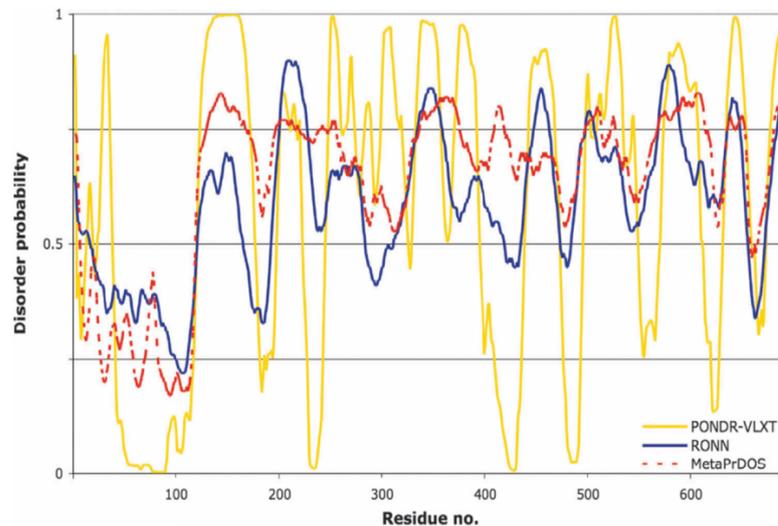


Figure 1.6 Structural composition of the Gab family of proteins (A) and predicted intrinsic disorder in the Gab1 protein (B)

(A) Schematic overview of Gab proteins structural composition. Gab proteins share common structural features such as an N-terminal (PH) domain, R-x-x-K Grb2 binding motifs and various binding sites for SH2 domain containing interaction partners, which need to be phosphorylated (indicated by an P) prior to binding. Only Gab1 contains a c-Met binding region (MBR). Figure legend is on the bottom (modified figure from Simister and Feller, 2012). (B) Intrinsic disorder prediction of the human Gab1 protein. The Gab1 protein contains intrinsically disordered regions locating after the well-folded PH domain. Gab1 (Q13480) was analysed by three different prediction programs as PONDRL-VLXT (yellow), RONN (blue) and MetaPrDOS (red dotted). Disorder probability (0-1, y-axis) is shown for each Gab1 amino acid residue (x-axis) (Figure from Simister and Feller, 2012).

The putative Gab4 protein lacks the lower affinity Grb2 binding site and contains only one of the three p85 binding sites. Interestingly, the proteins Gab3 and Gab4 lack all Crk/CrkL binding sites compared to Gab1 and Gab2 which strongly suggests a separate evolvement of functions between Gab family members. The Gab1 protein is also the only Gab protein within the Gab family of proteins that contains a c-Met binding region (MBR) enabling it to bind directly to the c-Met receptor (Weidner et al., 1996). Several disorder prediction programs (PONDRL-VLXT (yellow), RONN (blue) and MetaPrDOS (red)) indicate that the long Gab1 ‘tail’ is mostly intrinsically disordered whereas the PH domain is mostly structured (Figure 1.6, B). As the family of Gab proteins share a similar topology, the predicted ‘intrinsically disordered protein (IDP) character’ of the Gab1 protein is most likely applicable to other Gab proteins. So far, only little is known about the potential secondary structure elements within the Gab protein tails, although a few studies have shed first light on this. Binding sites in Gab1 and Gab2 for their interaction partners Grb2 and 14-3-3 have been structurally explored (Gab1 PDB codes: 4qsy, no publication.; Gab2 PDB codes: 2vwf, 2w0z, 5ewz, 5exa) (Harkiolaki et al., 2009, Bier et al., 2016). The PH domain of other proteins has been well studied (Scheffzek and Welti, 2012), but a structure of Gab PH is still missing.

One reason why the structure of Gab proteins have not been solved yet is the protein size (100 kDa) but most likely also the large degree of intrinsically disordered residues predicted to occur in the Gab protein. An detailed overview of intrinsically disordered proteins (IDPs) is given in the following subchapter 1.7.1.

1.7.1 Intrinsically disordered proteins (IDPs)

The classical structure – function paradigm for proteins, which describes that proteins need to adopt a tertiary structure to perform their functions, had been well accepted for a long time. The recognition of intrinsically disordered proteins (IDPs) has transformed the view on proteins from previously thought well-folded entities with relatively low flexibility to a more diverse structural spectrum with some highly dynamic macromolecules (Wright and Dyson, 1999). Although IDPs lack a well-defined tertiary structure, they are completely functional. IDPs or intrinsically disordered regions (IDRs) in proteins can be found in all three domains of life (archaea, bacteria, eukaryotes) and in viruses (Xue et al., 2012). An increase of the amount of intrinsic disorder is found in the eukaryotic compared to the prokaryote proteome (Xue et al., 2012). This might indicate a link between protein disorder and cellular complexity. More than one third of the eukaryote proteome contains proteins with intrinsically disordered regions which were defined by a length of 30 amino acids (Ward et al., 2004, Babu et al., 2011).

IDPs are characterised by a distinct amino acid composition creating a generally low hydrophobicity of the protein and a high protein net charge. They often contain elevated levels of charged and polar amino acid residues, but are depleted in hydrophobic and aromatic amino acid residues. Additionally, IDPs contain less asparagine and cysteine residues. Moreover, IDPs possess low sequence complexity such as pro-rich repetitive regions (Habchi et al., 2014). Another unique feature of IDPs is their flexibility and thereby their ability to adopt various functional conformations within cells. It is known that many IDPs partially fold upon binding to their interaction partners, which is referred to as ‘binding-induced folding’. These often transient interactions can include an IDP and a globular protein, or two IDPs (Dyson and Wright, 2002). For example, the intrinsically disordered phosphorylated kinase-inducible domain (pKID) of the cyclic-AMP responsive element binding protein (CREB) folds upon binding to the CREB-binding protein (Sugase et al., 2007). But, there are also IDRs in protein complexes which retain their conformational heterogeneity upon complex formation. These complexes are referred to as ‘fuzzy’ (Tompa and Fuxreiter, 2008). It has been shown that intracellular conditions, including macromolecular crowding, can also affect the topology of an IDP. An increase of protein compaction under crowded conditions was measured for IDPs such as the IDP carboxyamidated ribonuclease or the intrinsically disordered C-terminal domain of histone H1 (Qu and Bolen, 2002, Roque et al., 2007), whereas other IDPs remained unstructured (Flaugh and Lumb, 2001, Szasz et al., 2011).

IDPs are implicated in numerous biological processes, yet they are particularly found in cell regulation and cell signalling (Xie et al., 2007). It was shown that three out of four signalling proteins of the

mammalian proteome contain intrinsically disordered regions (IDR) (Theillet et al., 2014). Due to their importance in signalling pathways, IDPs are often associated with diseases such as neurodegenerative disorders, cancer, cardiovascular and metabolic diseases (Uversky et al., 2008, Vacic et al., 2012).

Hub proteins represent one example of IDPs involved in cellular signalling. Within signalling cascades, hub proteins are able to establish multivalent interactions with various interaction partners. The elevated levels of IDRs in hub proteins are believed to promote the binding versatility of hub proteins (Dunker et al., 2005, Haynes et al., 2006). The cell cycle regulator protein p53 is one prominent example for an intrinsically disordered hub protein (Bell et al., 2002). A spectroscopic analysis of the viral IDP (E1A) and its cellular host interaction partners (CBP and pRB) demonstrated that protein promiscuity and allostery, both known to be strongly associated with protein disorder, can contribute to the modulation of the hosts cellular signalling pathways (Ferreon et al., 2013, Hilser, 2013).

As IDPs such as Gab proteins play a major role in many important signalling pathways and are implicated in various diseases, their actual structural composition is of great interest. Structural methods like NMR or X-ray crystallography requires pure and soluble protein sample. However, the purification of IDPs, in particular proteins with large amounts of intrinsically disorder, is usually very challenging most likely due to their high inherent flexibility. Several studies have shown that the purification of IDPs is often complicated by protein aggregation (Lebendiker and Danieli, 2014). Also, the great flexibility of IDPs is a very major challenge for X-ray crystallography (Jensen et al., 2013, Konrat, 2014).

However, data from Prof. Schaper's group and preliminary work by our group proposed a topology model for the Gab1 protein which is described in more detail in the following subchapter 1.7.2

1.7.2 A Gab topology model and Gab1 in signal computation

A comparison of the topological composition of Gab proteins with other LMD proteins (e.g. DOK, FRS2 and IRS) showed that they share a similar structural composition of one or two well-folded N-terminal domains followed by a mostly disordered protein 'tail'. Additionally, the group of Prof. Schaper suggested an Gab1 interaction of the N-terminal PH domain and the protein 'tail', which was further supported by data from our group, albeit in the form of an intramolecular Gab1 interaction (Eulenfeld and Schaper, 2009, Simister et al., 2011). Surprisingly, a Gab1 peptide array probed with a recombinantly expressed Gab1 PH domain identified more than one intramolecular interactions between the Gab PH domain and distinct regions in the Gab 'tail' (Simister et al., 2011). Based upon those findings the N-terminal folding nucleation (NFN) hypothesis was proposed (Figure 1.7).

The NFN hypothesis describes a 'folding' mechanism of how LMD proteins adapt protein formations with increased compactness in order to avoid non-specific interactions with other proteins, protein aggregation or degradation (Simister et al., 2011). During ribosomal translation of Gab or other LMD proteins, the N-terminal chain instantly folds into a well-structured domain and serves as a nucleation point for the subsequent nascent protein chain. Protein 'tail' regions bind to the distant N-terminal

domain and let the protein adapt a loop-like structure (Figure 1.7). The more compact shape of the LMD protein could then protect against unspecific protein aggregation or proteolysis. Posttranslational modifications on the more compact shaped Gab or other LMD proteins could initiate interactions with other signalling proteins which were previously blocked by the initial protein shape.

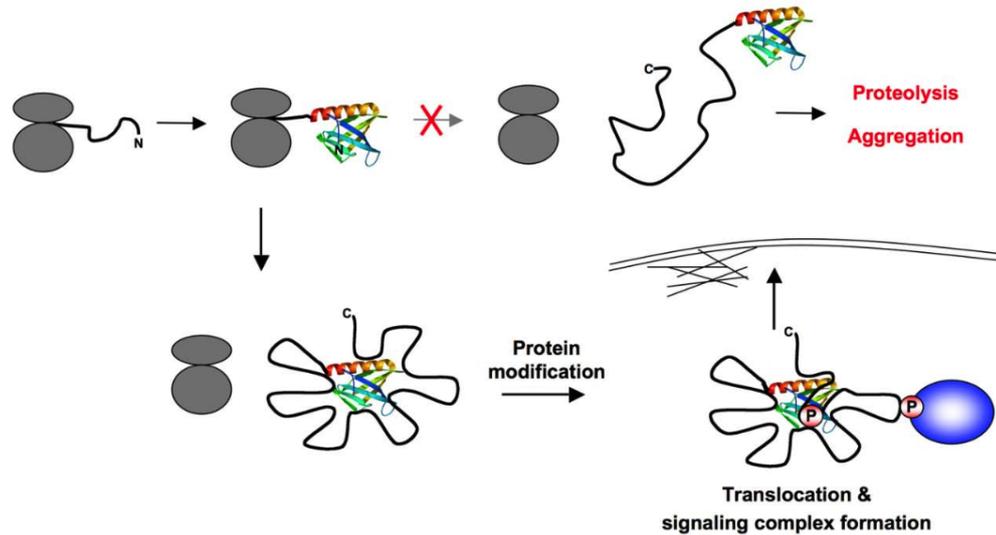


Figure 1.7 N-terminal folding nucleation (NFN) hypothesis for Gab and other LMD proteins

The NFN hypothesis describes a topological organisation mechanism by which LMD proteins adapt a more compact shape. Accordingly, the initial ribosomal translation of the N-terminal well-folded domain serves as a nucleation point for the nascent polypeptide chain and provides docking points for intramolecular chain contacts. The resulting overall more compact topology of the protein could prevent protein aggregation or proteolysis. Protein modifications such as phosphorylation enables interactions with other signalling partner proteins or protein translocation to the plasma membrane (Figure from Simister et al., 2011).

Interestingly, Crk/CrkL, p85(PI3K) and Shp2 protein binding sites in Gab1 are located in clustered regions. The proposed Gab1 loop-like structure separates these clustered binding sites in functional Gab1 loop regions such as a proliferation, a survival and a migration and invasion loop (Figure 1.8).

Within these loop regions, a defined group of signalling proteins could form stable sub-complexes to activate function related pathways. Phosphorylation of CRKL, p85(PI3K) and Shp2 interactions sites in Gab1 by e.g. c-Met is required prior to Gab signalling complex assembly (Wöhrle et al., 2009). The Grb2 SH3C – Gab1 interaction is constitutive and therefore independent of receptor activation (Holgado-Madruga et al., 1996). An inter-loop cross-talk could be facilitated by the inherent flexibility of the IDP Gab1. The protein architecture according to this model might explain for the first time how a spatial separation of different pathways can be achieved by Gab and also other LMD proteins.

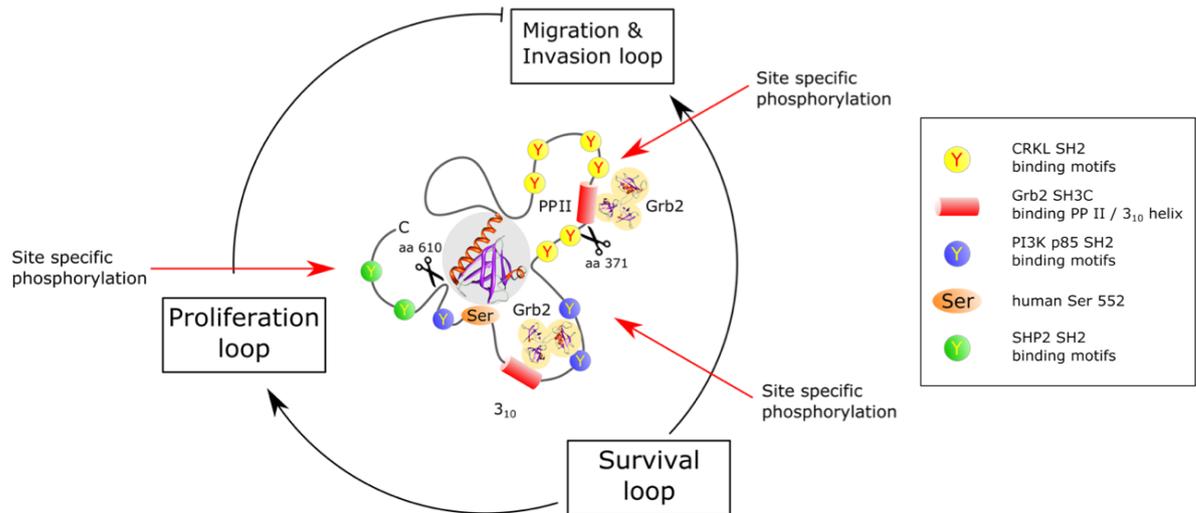


Figure 1.8 A topological model of Gab1 explaining how Gab can facilitate signal computation within cell signalling

The loop model of the Gab1 protein (NFN hypothesis) defines distinct loop areas which are specialised in specific biological processes as proliferation, survival and migration/invasion. Signalling proteins can assemble on spatially separated loops and form multi-protein complexes activating specific cellular pathways (Figure adapted from Lewitzky et al., 2012).

1.8 Gab proteins and cancer

Gab's pivotal role in signalling processes regulating migration, invasion or proliferation, is probably the main reason why Gab proteins are often associated with tumorigenesis. In particular, the function of Gab as a signal amplifier and / or its connection with proto-oncogenes strongly add to aberrant RTK signalling in tumours (Wöhrle et al., 2009). Gab2 overexpression was found in solid cancers such as breast cancer and melanoma (Daly et al., 2002, Bentires-Alj et al., 2006, Horst et al., 2009). A recent study has shown that human hepatocellular carcinoma (HCC) tissue contained overexpressed Gab2 in over 60% of the cases (Cheng et al., 2017). Also, the development of solid tumours growing in the extracranial space, called neuroblastomas, are associated with an increased amount of Gab2 protein (Zhang et al., 2017). Neuroblastomas are responsible for over 10% of paediatric cancer deaths.

Gab2 is also connected with neoplasms of the hematopoietic system, for example in chronic myeloid leukaemia (CML) (Aumann et al., 2011). The significance of Gab2 as a contributor in BCR/ABL transformation was shown by myeloid progenitor cells from Gab2^{-/-} mice, which exhibited a complete resistance to BCR/ABL (Sattler et al., 2002). The *bcr/abl* fusion gene encodes a constitutively active tyrosine kinase found in virtually all CML patients.

In contrast to the strong implication of Gab2 in tumorigenesis, roles of the Gab1 or Gab3 proteins have been less well defined. It has been shown that the Gab1 protein is a mediator of tumour progression and tumour metastasis due to its essential role in the c-Met signalling pathway. Many tumours show an aberrantly expressed or mutated c-Met receptor (Danilkovitch-Miagkova and Zbar, 2002). A colorectal cancer cell line (DLD-1) that overexpresses the c-Met receptor showed that the Met-Gab1 and not the

Met-Grb2 interaction, was mainly responsible for tumour growth and invasion motility (Seiden-Long et al., 2008). In conclusion, Gab proteins, especially Gab1 and Gab2, are often associated with carcinogenesis which can make them a great therapeutic target in cancer therapy.

1.9 Aim and concept of this study

The intrinsically disordered Gab family proteins are large multisite docking proteins and function as signal integration platforms for cellular signalling. Many studies have shown that Gab proteins play an important role in various signalling pathways and are also strongly implicated in tumourigenesis (Wöhrle et al., 2009). However, an ultrastructural characterization of full-length Gab proteins or of Gab proteins within signalling complexes has not been done, yet. An structure of the Gab protein could help us to get a better understanding of how the IDPs can act as a signal platform for different signalling pathways and, importantly, how signalling pathway crosstalk is mediated in cells. Therefore, a major aim of this thesis was the examination of structural features of Gab proteins and Gab-based complexes.

Methods to generate protein structure models with a high resolution such as NMR or X-ray crystallography require relatively large amounts of soluble and pure protein. Initially, the Gab2 protein and, subsequently, the p35Gab1 fragment were recombinantly expressed in *Escherichia coli* (*E. coli*). For the assembly of a Gab-based (signalling) complex, Gab interaction partners Grb2 and p85(PI3K) were recombinantly expressed in *E. coli* and purified. Different expression and purification strategies were employed to establish a purification protocol for soluble and pure protein.

Gab-based complexes were structurally examined by atomic force microscopy (AFM), electron microscopy (EM), cross-linking mass spectrometry (XL-MS), X-ray crystallography and isothermal titration calorimetry (ITC). The Gab2 – Grb2 interaction has already been structurally characterised, with respect to the individual Grb2 SH3C domain complexed with Gab2 peptides (Harkiolaki et al., 2009). In order to determine the contribution of the full-length Grb2 protein towards the Gab1 – Grb2 interaction, another aim of this thesis was to explore contact sites between full-length Grb2 and Gab1, or Gab1-derived peptides.

Chapter 2 Material and Methods

2.1 Materials

2.1.1 Equipment

Bacterial culture and lysis

Type B6120 incubator for bacterial plates

Heraeus, Hanau, D

Sonicator Sonopuls HD 2200

Bandelin, Berlin, D

Bacterial shaker Innova 4330, 43, 4200, 4400

New Brunswick Scientific, Edison,
NY, USA

Centrifuges

Cooled microfuge 5417R, 5415R

Eppendorf, Hamburg, D

Large volume low speed centrifuge J-6B

Beckman Coulter, Krefeld, D

Medium speed centrifuges Avanti J-25

Beckman Coulter

Chromatography

Äkta FPLC and Äkta pure 25

GE Healthcare, Freiburg, D

HiLoad 16/600 Superdex S75

GE Healthcare

HiLoad 16/600 Superdex S200

GE Healthcare

Mono Q HR 5/5

GE Healthcare

HiTrap columns (HiTrap IMAC FF,

GE Healthcare

HiTrap Q HP, HiTrap Benzamidine FF)

Isothermal Titration Calorimetry

MicroCal™ iTC₂₀₀

GE Healthcare

Miscellaneous

PCR cycler FlexCycler

Analytik Jena, Jena, D

Precision scale Scaltec SBC21

Scaltec, Göttingen, D

Scale Scaltec SP061

Scaltec

Scale PA64C

OHAUS, Greifensee, CH

pH meter inoLab

Mettler Toledo, Gießen, D

Autoclave ELV 5050

Tuttnauer, Breda, NL

Water purifying system MilliQ

Merck Millipore, Darmstadt, D

Water purifying system Siemens Ultra Clear™ UV UF

neoLab, Heidelberg, D

-80 °C freezer Panasonic MDF-974-PE

Ewald, Rodenberg, D

Material and Methods

-150 °C freezer Panasonic MDF-C2156VAN-PE

Ewald

Mixing

Clay Adams Nutator

BD, Heidelberg, D

Thermomixer Eppendorf Comfort

Eppendorf

Vortex-Genie 2

Scientific Industries, Inc., Bohemia,
NY, USA

Variomag Elektrorührer

Variomag, Daytona Beach, FL, USA

Photometry

BioPhotometer

Eppendorf

NanoVue™ Plus Spectrophotometer

GE Healthcare

Western Blotting

Protean II xi electrophoresis chamber

Bio-Rad, Munich, D

Mini-Protean 3

Bio-Rad

TransBlotSD Semi-Dry Transfer Cell

Bio-Rad

Gel documentation systems

Syngene G:BOX

Syngene, Cambridge, UK

Fusion FX7™

Vilber Lourmat, Eberhardzell, D

Glassware was purchased from Schott Duran.

2.1.2 Consumables

1.5 ml tubes Axygen

Corning, NY, USA

15 ml tubes with lid, sterile

Corning

50 ml tubes with lid, sterile

Corning

14 ml tubes

Greiner Bio-One, Frickenhausen, D

500 ml filter systems

Corning

0.2 µm syringe filter

Corning

30 ml centrifuge tubes

Sarstedt, Nümbrecht, D

Bacteriological Petri Dish, 100x15 mm

Corning

Push caps for 30 ml tubes

Sarstedt

Fisherbrand™ disposable cuvettes

Fisher Scientific, Schwerte, D

Dialysis membrane 3500 Da MWCO

Fisher Scientific

Spectra/Por dialysis membrane, 1000 Da MWCO

VWR, Darmstadt, D

Inoculation mini needles

VWR

Pipette tips

Corning

Material and Methods

Plastic pipettes	Corning
Poly-Prep Chromatography Columns	Bio-Rad
PVDF membrane	Fisher Scientific
Syringes	BD Plastipak, Heidelberg, D
Vivaspin Turbo 15, 5 000 Da MWCO	Sartorius, Göttingen, D

2.1.3 Reagents and Chemicals

Chemicals were purchased from Roth, Sigma Aldrich, Serva or Fisher Scientific unless stated otherwise. Thrombin and column materials (IMAC, GSH beads) were obtained from GE Healthcare. The gel-staining reagent InstantBlue was obtained from Expedeon.

2.1.4 Protease and phosphatase inhibitors

Antipain hydrochloride (1000x stock: 5 mg/ml in H₂O)

Aprotinin (100x stock: 1 mg/ml in H₂O)

Ethylenediaminetetraacetic acid (EDTA) (stock: 0.5 M in H₂O, adjusted to pH 7.5)

Leupeptin (1000x stock: 0.5 mg/ml in H₂O)

Pepstatin A (1000x stock: 0.7 mg/ml in methanol)

Phenylmethylsulfonylfluoride (PMSF) (500x: 100 mg/ml in DMSO)

Sodium molybdate (100x stock: 100 mM in H₂O)

Sodium orthovanadate (100x stock: 100 mM in H₂O)

2.1.5 Antibodies

Antibodies

Primary antibody	Host species	Dilution for WB	Source and/or reference
GST	mouse	1:500	2C8 1-4, E. Kremmer, GSF, Munich (Pietrek et al., 2010)
4G10 (pY antibody)	mouse	1:250	(Druker et al., 1989)
Secondary antibody	Host species	Dilution for WB	Source and/or reference
HRP-coupled anti-mouse IgG	donkey	1:10 000	715-036-151, Jackson ImmunoResearch™

Table 2.1 List of antibodies

2.1.6 Bacterial strains and expression vectors

Bacterial strains

E. coli expression strains BL21 (DE3), BL21-CodonPlus (DE3)-RIL and TKB1 were used for protein expression. The BL21 XL2 bacterial strain was used for DNA-plasmid preparation.

Host strain	Genotype	Source
BL21 (DE3)	<i>E. coli</i> BF ⁻ <i>dcm ompT hsdS</i> (<i>r_B</i> ⁻ <i>m_B</i> ⁻) <i>gal λ</i> (DE3)	NEB (C2527H)
BL21 XL2	<i>endA1 supE44 thi-1 hsdR17 recA1 gyrA96 relA1 lac</i> [F <i>proAB lacI^qZAM15 Tn10 (Tet^r) Amy Cam^r</i>]	Agilent Technologies (200150)
TKB1	<i>E. coli</i> B F ⁻ <i>dcm ompT hsdS</i> (<i>r_B</i> ⁻ <i>m_B</i> ⁻) <i>gal λ</i> (DE3) [pTK <i>Tet^r</i>]	Agilent Technologies (200134)
BL21-CodonPlus (DE3)-RIL	<i>E. coli</i> B F ⁻ <i>ompT hsdS</i> (<i>r_B</i> ⁻ <i>m_B</i> ⁻) <i>dcm</i> ⁺ <i>Tetr gal λ</i> (DE3) <i>endA Hte</i> [argU ileY leuW Cam ^r]	Agilent Technologies (230245)

Table 2.2 List of bacterial strains

Protein expression vectors

Insert	Insert origin	Tag	Vector	Source and/or reference
Gab2	mouse	GST	pGEX-6P-1	Insert Gab2 (Edmead et al., 2006), cloned into a pGEX-6P-1 vector.
p35Gab1	human	His	pET200-D-TOPO	(Le Goff et al., 2012)
p35Gab1 (C374A, C405A, C514A)	human	His	pET200-D-TOPO	p35Gab1 sequence mutations were made by Dr. Tobias Gruber (AG Prof. Feller, MLU).
Grb2	human	His	pET-28(+)	PhD thesis of Rebekah Bartelt (Prof. Houtman, University of Iowa).
Grb2 (C32S, C198A)	human	His	pET-28(+)	Grb2 sequence mutations were made by Dr. Marc Lewitzky (AG Prof. Feller, MLU)
Grb2	human	His	pMCSG7	DNASU plasmid collection
Grb2	human	-	pET-21d(+)	Insert Grb2 was cloned from the pMCSG7 vector into the pET-21d(+) vector (EMD Millipore).
p85α; cysteine-free (C146S, C167S, C498S, C656S, C659V)	human	GST	pGEX-6P-1	(LoPiccolo et al., 2015)

Table 2.3 List of protein expression vectors

All vector inserts were sequenced before use.

2.1.7 Peptides

Peptides

Peptide	Peptide sequence	Peptide-Source
Grb2 SH2 positive control	A(pY)VNVA (1)	Dr. Masch (MLU)
Grb2 SH2 negative control	AYVNVA (2)	Dr. Masch (MLU)
Grb2 SH3N positive control	PPPPLPPRRRR (3)	Dr. Masch (MLU)
Grb2 SH3N negative control	PPGALGPLLRR (4)	Dr. Masch (MLU)
15 aa Gab1, XL-MS Grb2 SH2 binder (C514S)	Gab1 514-SEPPPVDRNLKPDRK-528 (10)	Dr. Masch (MLU)

Peptide	Peptide sequence	Peptide-Source
15 aa Gab1, XL-MS Grb2 SH2 binder (C514S, R521A, K524A)	Gab1 514-SEPPPVDA ^{ANL} APDRK-528 (11)	Dr. Masch (MLU)
15 aa Gab1, XL-MS Grb2 SH3N binder (WT)	Gab1 497-FRSSPKTPRRRPVPV-511 (5)	Dr. Masch (MLU)
AA mutant 15 aa Gab1, XL-MS Grb2 SH3N binder (P501A, R506A)	Gab1 497-FRSS ^{AKT} PP ^{ARP} VVPV-511 (6)	Dr. Masch (MLU)
32 aa Gab1, SH3 tandem peptide (C514S)	Gab1 497-FRSSPKTPRRRPVPVADSEPPPVDRNLKPK-528 (7)	JPT
AA mutant 32 aa Gab1, SH3 tandem peptide (P501A, R506A, C514S)	Gab1 497-FRSS ^{AKT} PP ^{ARP} VVPVADSEPPPVDRNLKPK-528 (8)	JPT
45 aa Gab1, SH3 tandem peptide (C514S)	Gab1 497-FRSSPKTPRRRPVPVADSEPPPVDRNLKPK-541 (9)	Prof. Schutkowski (MLU); JPT
45 aa Gab1, SH3 tandem peptide (WT)	Gab1 497-FRSSPKTPRRRPVPVADCEPPPVDRNLKPK-541 (9)	Dr. Gruber (MLU)
32 aa Gareml, SH3 tandem peptide (WT)	Gareml 529-LLNAPPVPPRS ^{AKPL} STSPSIPPRTVKP-560 (12)	JPT

Table 2.4 List of peptides

Gab1 peptides were either synthesized by Dr. Antonia Masch (AG Prof. Schutkowski, MLU) or JPT Peptide Technologies (JPT). Mutated residues in the binding motif are coloured in red. Mutated residues outside the motif are coloured in grey.

2.1.8 Protein Crystallography

Sparse matrix screens

JBScreen JCSG++ (Jena BioScience), JBScreen classic 1-10 (Jena BioScience), JBScreen cryo 1-4 (Jena BioScience), Crystal screen 1+2 (Hampton Research), Morpheus (Molecular Dimensions), Low ionic strength crystallization kit + extension (Sigma-Aldrich)

2.1.9 Molecular weight protein standards and DNA Kits

Precision Plus Protein™ Dual Color Standard	Bio-Rad
Precision Plus Protein™ Kaleidoscope™ Standard	Bio-Rad
Blue Prestained Protein Standard	NEB
QIAprep Spin Miniprep Kit	Qiagen

2.1.10 Solution and Buffers

Bacterial culture and recombinant protein expression

Material and Methods

LB (Luria/Miller) medium (1L)

25 g LB granulate (6673, Carl Roth) was dissolved in 950 ml H₂O. LB medium was filled up to 1 litre with H₂O and autoclaved for 21 minutes at 121 °C.

LB agar plates

35.6 g LB (Lennox) agar EZMix™ powder (Sigma-Aldrich) is added to 1 litre H₂O. Medium is heated and stirred until the powder is completely dissolved. Afterwards, the medium is autoclaved for 21 minutes at 121 °C. Medium is cooled down to 45 °C and appropriate antibiotics are added. After thoroughly mixing, about 25-30 ml medium are poured into a petri dish (10 cm diameter) and left to solidify at RT. Plates are stored upside-down in sealed plastic bags for max. 4 weeks at 4 °C.

Terrific Broth (TB) medium (1 litre)

47.6 g granulated powder (BP9728, Fisher Bioreagents) was dissolved in 950 ml H₂O. 4 ml glycerol was added to the TB medium. Then, the TB medium was filled up to 1 litre with H₂O and autoclaved for 21 minutes at 121 °C before use.

TFB1

30 mM potassium acetate

10 mM CaCl₂

50 mM MnCl₂

100 mM RbCl

15% glycerol (v/v)

The pH needs to be adjusted to 5.8 with 1 M acetic acid. Filter sterilize (0.2 µm).

TFB1 can be stored at RT.

TFB2

100 mM MOPS (pH 6.5)

75 mM CaCl₂

10 mM RbCl

15% glycerol (v/v)

The pH needs to be adjusted to 6.5 with 1 M KOH. Filter sterilize (0.2 µm).

TFB2 can be stored at RT.

TPE lysis buffer

1% Triton X-100 (v/v)

100 mM EDTA pH 7.5

1x PBS

Material and Methods

Before use cool to 4 °C and add protease inhibitors if required.

GSH-bead wash buffer (GSH-WB)

50 mM Tris pH 7.5

100 mM EDTA pH 7.5

0.1% Tween 20 (v/v)

Tris-HCl (1 M, 1 litre)

121.1 g Tris base

Dissolve in H₂O and adjust with HCl to pH 6.8, 7.5 or 8.8.

Thrombin cleavage buffer

20 mM sodium phosphate pH 8.0

150 mM NaCl

PreScission protease 3C cleavage buffer

50 mM Tris-HCl pH 8.0

150 mM NaCl

1 mM EDTA

1 mM DTT

10x PBS (10 litres)

800 g NaCl

25 g KCl

143 g Na₂HPO₄

25 g KH₂PO₄

Dissolve in 10 litres H₂O.

4x Bradford protein assay reagent

For a 4x Bradford protein assay reagent, 250 mg Brilliant Blue G are dissolved in 120 ml ethanol. The solution is mixed very carefully with 250 ml concentrated phosphoric acid. Finally, H₂O is added to a final volume of 500 ml. The solution can be stored at RT and in a light-protected environment. For preparing a 1x solution, the concentrated solution is diluted 1:4 with H₂O, mixed for 15 min at RT and filtered through a pre-wetted filter. Solution is stored at RT in the dark.

Size exclusion chromatography (SEC) protein buffers

(All buffers used in chromatography were filter-sterilized (0.2 µm) and degassed before use)

Material and Methods

1. *Gab2*

50 mM Tris pH 7.5

2 mM β -mercaptoethanol

2. *p35Gab1*

20 mM sodium phosphate pH 7.5

250 mM NaCl

3. *p35Gab1 – Grb2 complex*

20 mM Tris or sodium phosphate pH 8 (dependent on the method used)

150 mM NaCl

+/- 0.1 mM sodium molybdate

4. *Grb2*

20 mM Tris pH 8

150 mM NaCl

5. *p85(PI3K)*

20 mM Tris pH 8

300 mM NaCl

SDS PAGE and Western Blotting

TBST

20 mM Tris 7.5

100 mM NaCl

0.1% (v/v) Tween 20

Semi-Dry blot-buffer (SDBB)

48 mM Tris base

38.6 mM Glycine

0.037% (w/v) SDS

Filter-sterilize (0.2 μ m) and store at RT.

10x SDS-PAGE running buffer

250 mM Tris base

1.9 M Glycine

1% (w/v) SDS

Dissolve in H₂O and store at RT. Leave out SDS for the 10x native gel running buffer.

Material and Methods

2-4x SDS protein gel sample buffer

70 mM Tris pH 6.8

40% (v/v) Glycerol

5% (v/v) β -mercaptoethanol

3% (w/v) SDS

0.05% (w/v) Bromophenol blue sodium salt

Aliquot and store at -20 °C. Mix carefully to dissolve SDS completely before use. For native gel sample buffer, use 70 mM Tris pH 8.8 and leave out SDS and β -mercaptoethanol.

SDS polyacrylamide gel

4% Stacking gel (10 ml)

Acrylamide/Bis-acrylamide solution	1.67 ml
------------------------------------	---------

1 M Tris-HCl pH 6.8	1.27 ml
---------------------	---------

10% SDS	0.1 ml
---------	--------

50% Glycerol	0.9 ml
--------------	--------

10% APS	0.1 ml
---------	--------

TEMED	10 μ l
-------	------------

12% Separating gel (10 ml)

Acrylamide/Bis-acrylamide solution	4 ml
------------------------------------	------

1 M Tris-HCl pH 8.8	3.75 ml
---------------------	---------

10% SDS	0.1 ml
---------	--------

10% APS	0.1 ml
---------	--------

TEMED	10 μ l
-------	------------

All SDS gel components, except of APS and TEMED, are mixed and filled up with H₂O to the final volume. Finally, APS and TEMED are added, the solution is mixed thoroughly and the gel is cast. To obtain other gel percentages, the amount of Acrylamide/Bis-acrylamide solution in the separating gel is varied.

Blocking buffer for Western blots

TBST with 5% low fat dry milk. For phospho-specific antibodies use 3% Ig-free bovine serum albumin (BSA).

Isothermal calorimetry titration

ITC buffer

25 mM HEPES-KOH pH 7.5

100 mM potassium acetate

5 mM magnesium acetate

Solution is filter-sterilized (0.2 μm) and degassed before use. When using phospho-tyrosine peptides, 100 μM sodium molybdate and sodium orthovanadate are added into the ITC buffer.

2.2 Methods

2.2.1 Bacterial culture and expression of recombinant proteins

2.2.1.1 *Bacterial expression strains*

The BL21 (DE3) bacterial expression strain was commonly used for recombinant protein expression (catalog number C2527H, NEB). The CodonPlus-RIL strain contains extra copies of the *argU*, *ileY*, and *leuW* tRNA genes (catalog number 230245, Agilent Technologies). BL21 (DE3) CodonPlus-RIL was used for the p85 α protein expression as described in LoPiccolo et al., 2015. The BL21 (DE3) TKB1 strain was employed to phosphorylate recombinantly expressed protein *in vivo* (catalog number 200134, Agilent Technologies). According to the TKB1 manual (Instruction manual, catalog number 200134, Agilent Technologies), the expression of the Tyr kinase domain (TK) needs to be activated by a special TK induction medium. However, the kinase was already active in TB medium and the TK induction medium step was skipped.

2.2.1.2 *Making chemically competent bacteria (Modified RbCl Method)*

The RbCl method was used to make competent bacteria. The protocol is an adaption from a protocol described by Hanahan (Hanahan, 1985). Bacteria were plated onto an LB plate supplemented with the appropriate antibiotic(s), if any, and incubated overnight at 37 °C. One colony from the LB plate was picked and incubated in 2.5 ml LB supplemented with appropriate antibiotic(s) for overnight at 37 °C, shaking at 225 rpm. On the next day, the culture was diluted 1:100 into a total volume of 125 ml. MgSO₄ was added to a final concentration of 20 mM. The bacteria culture was grown until an OD₆₀₀ of 0.4-0.6 was reached. Harvested cells (4000 x g, 4 °C for 15 min) were resuspended in 50 ml ice-cold TFB1 and pooled in one tube. Cells were incubated on ice for 5 min. After a second centrifugation step (4000 x g, 4 °C for 15 min), cells were resuspended in 5 ml ice-cold TFB2 using ice-cold pipettes. Finally, cells were incubated for 50 min on ice, then aliquoted (200 μl / 1.5 ml tube) with ice-cold tips and quick-frozen in liquid nitrogen to be stored at -80 °C.

2.2.1.3 *Heat-shock transformation of bacterial strains*

The heat-shock transformation was used to transform chemically competent bacteria. An aliquot of competent bacteria was thawed on ice and mixed with 1 μl (ca. 1 pg-100 ng) of the desired vector DNA dissolved in TE buffer or water. For the co-transformation with two DNA sources, as for p35Gab1-Grb2 expression, 1 μl of each plasmid were added to the competent bacteria. After the cells

were incubated for 20 min on ice, a heat shock for 45 sec at 42 °C was performed. The bacteria were incubated for 2 min on ice and subsequently, filled up with warm LB medium to 1 ml. After one hour incubation (37 °C, 225 rpm), cells were briefly centrifuged (950 x g, 1 min). Most of the LB medium was discarded and the pellet was resuspended in residual LB medium. Half or all cells, dependent on the colony numbers expected, were plated on an LB plate supplemented with the appropriate antibiotic(s) and incubated overnight at 37 °C.

2.2.1.4 Protein expression and preparing bacterial lysate supernatant (20K)

2-4 ml LB medium supplemented with the appropriate antibiotic(s) was inoculated with material from a bacteria glycerol stock or a single colony from an LB plate. The culture was incubated until bacteria growth was visible (37 °C, 220 rpm). Then, 0.5-1 ml bacteria culture was added to 400 ml of fresh LB medium supplemented with the appropriate antibiotic(s). The bacterial starter culture was incubated overnight at 37 °C (220 rpm). 550 ml (10x for a large expression) TB medium supplemented with the appropriate antibiotic(s) was mixed with 30 ml starter culture and bacteria were cultivated at 37 °C until an OD₆₀₀ of > 1.3 (max. 1.8) was reached. Protein expression was induced with 0.1 mM or 0.03 mM IPTG (Gold Biotechnology, Inc., St. Louis, Missouri, USA) and the rotation of the shaker was reduced to 130 rpm. Expression times and temperatures varied for each protein expression (see individual purification protocols). Bacteria were harvested by centrifugation (2500 x g for 30 min at 4 °C). Ice-cooled bacterial pellets were resuspended in an ice-cold lysis buffer supplemented with protease inhibitors (Aprotinin, Antipain, Leupeptin, Pepstatin A and PMSF). For buffer details see the specific protein purification protocol. The bacterial suspension was sonicated for 3x 1 min on ice to break cells and to shear the DNA. Intermittent cooling phases (ca. 2 min each) prevented sample overheating. A subsequent centrifugation cleared the bacterial lysate from cell debris (20,000 x g, 30 min, 4 °C; Avanti J-25, Beckman). Bacterial lysate (20K supernatant) was used immediately for further purification or frozen with LN₂ and stored at -80 °C.

2.2.1.5 Purification of recombinantly expressed GST-fusion proteins

2.2.1.5.1 Gab2 protein

GST-Gab2 protein (murine WT Gab2 in pGEX-6P-1; Gab2: ca. 75 kDa; GST: ca. 26 kDa) was expressed in *E. coli* BL21 or TKB1 for 5 h at 18 °C. Bacteria were lysed in TPE buffer (20 ml buffer per 1 litre bacterial pellet) supplemented with protein inhibitors. Bacterial lysate (20K supernatant) was generated. Details for protein expression and 20K preparation can be found in 2.2.1.4.

The 20K lysate was incubated with glutathione sepharose (GSH) beads on a nutator overnight at 4 °C. Protein-loaded beads were collected with TPE buffer and washed 3x with GSH wash buffer (3000 g, 4 °C, 10 min). Elution was done with 1x bead volume elution buffer (EB: 100 mM GSH, pH-adjusted to approx. pH 8 with 1 M Tris-HCl pH 8.8) by gentle mixing for overnight at 4 °C in a Poly-Prep column. The column flow-through was collected and a second elution step was performed with

fresh EB (0.5 x bead volume) for 2 h at 4 °C. A final elution with 1 ml EB replaced the buffer in the column void volume. Eluate was collected by gravity flow and dialysed at 4 °C against 3x 1 litre of 50 mM Tris pH 7.5 supplemented with fresh 2 mM β -mercaptoethanol. Dialysed eluate was analysed by size exclusion chromatography. Protein sample integrity and purity was analysed by SDS PAGE and Coomassie InstantBlue staining. For TKB1 expressed GST-Gab2, buffers were supplemented with 250 mM NaCl, 5% glycerol and 1 mM each of the phosphatase inhibitors sodium molybdate and/or vanadate.

2.2.1.5.2 *p85 α protein*

GST-p85 α (human cysteine-free p85 α in pGEX-6P-1 vector; p85 α : ca. 83 kDa; GST: ca. 25 kDa) was expressed in the CodonPlus-RIL bacteria strain. Expression and purification steps were performed according to the protocol described in LoPiccolo et al, 2015 with some modifications, as indicated below. Protein expression was induced with 0.1 mM IPTG and cells were incubated for overnight at 25 °C. Bacteria were lysed in PBS, pH 7.4 supplemented with 4 mM DTT, 2 mM EDTA, 2 mM PMSF and protease inhibitors Aprotinin, Leupeptin, Pepstatin A and Antipain. The 20 K supernatant was incubated with GSH beads for 3 h at 4 °C. PreScission protease cleavage was performed in a Poly-Prep column on a nutator at 4 °C ON. Optimal protease amounts for self-purified GST-3C had been determined by previously performed protease titration experiments. Cleaved p85 α protein was dialysed ON against 2x 2 litres Mono Q buffer (20 mM Tris pH 8, 20 mM NaCl) at 4 °C and loaded onto a Mono Q column (HiTrap Q HP). Protein was eluted with a 0-350 mM NaCl gradient over 40 column volumes (Binding buffer: 20 mM Tris pH 8; Elution buffer: 20 mM Tris pH 8, 1 M NaCl). Mono Q fractions containing p85 α were concentrated and analysed on a Superdex S200 column (HiLoad 16/600 Superdex S200, Buffer: 20 mM Tris pH 8, 300 mM NaCl). Fractions were analysed by SDS PAGE, and those with a 95% purity were pooled and stored at -80 °C.

2.2.1.6 *Purification of recombinantly expressed His-tagged proteins*

2.2.1.6.1 *p35Gab1 protein*

His-p35Gab1 (human WT p35Gab1 in pET200-D-TOPO; p35Gab1: ca. 35 kDa) protein was expressed in *E. coli* BL21 for 4 h at 18 °C. Bacterial pellet from 1 litre culture was lysed in 20 ml of 20 mM sodium phosphate pH 7.5, 500 mM NaCl, 5% glycerol supplemented with protease inhibitors. Bacterial lysate (20K supernatant) was generated. Details for protein expression and making bacterial lysate supernatant can be found in 2.2.1.4.

20K supernatant was incubated with cobalt-charged IMAC beads (20 μ l suspension IMAC beads/ 1 ml 20K) for 1 h at 4 °C on a roller. A second pull down was done with depleted 20K using the same bead to 20K ratio. Beads were washed with 12.5 mM imidazole wash buffer (5 min, 900 x g, 4 °C) and transferred into a Poly-Prep column for elution with elution buffer (EB) containing 500 mM imidazole (elution volume: 2x beads volume, 1 h, 4 °C, shaking). A second elution step was performed with 1x

bead volume EB (30 min, 4 °C, shaking). Eluate was collected by gravity flow. Residual protein in the column was washed out with 0.5 x beads volume EB. The pooled eluate was dialysed against 3x 1 litre of 20 mM sodium phosphate pH 7.5, 250 mM NaCl (S200 buffer) at 4 °C. Afterwards, eluate was concentrated (5000 MWCO Vivaspin concentrator), centrifuged (10 min, 4 °C, 18000 x g) and applied onto a S200 column (HiLoad 16/600 Superdex S200). Fractions were analysed by SDS PAGE, and those containing the complex were pooled and stored at -80 °C.

2.2.1.6.2 *p35Gab1 – Grb2 complex*

His-p35Gab1 (human WT or C374A, C405A, C514A mutant in pET200-D-TOPO vector; p35Gab1: ca. 35 kD) was co-expressed with Grb2 (human WT or C32S, C198A mutant in pET-21d(+) vector; or human His-Grb2 (WT) in pMCSG7 vector; Grb2: ca. 27 kDa) in *E. coli* TKB1 for overnight at 18 °C. Bacteria were lysed in lysis buffer (LB) of 20 mM Tris pH 8, 500 mM NaCl, 2.5 mM Imidazole, 1 mM PMSF, 0.1 mM sodium orthovanadate, 0.1 mM sodium molybdate, supplemented with protease inhibitors. A bacterial pellet from 1 litre ON culture was resuspended in 100 ml LB and bacterial lysate supernatant (20K) was generated. The ratio of lysis buffer to bacterial pellet was important to minimize the protein loss due to a too high protein concentration. Details for protein expression and 20K preparation can be found in 2.2.1.4.

Filtered 20K (0.2 µm filter pore size) was loaded onto a cobalt-charged HiTrap IMAC column (CV: 1 ml, 5 ml or 2x 5 ml). Binding buffer (BB_{IMAC}) was 20 mM Tris pH 8, 500 mM NaCl and elution buffer (EB_{IMAC}) was 20 mM Tris pH 8, 500 mM NaCl, 500 mM Imidazole. Filtration of the lysate was necessary to prevent column obstruction. The column was washed with 4% EB_{IMAC} until the UV signal was stable. The complex was then eluted with 50% EB_{IMAC}. Fraction samples were analysed by SDS-PAGE and fractions containing the complex were dialysed against 20 mM Tris pH 8 (3x 2 litres buffer, Spectra/Por, 1 kDa MWCO). Dialysate was loaded onto a HiTrap Q HP column (Binding buffer (BB_Q): 20 mM Tris pH 8; Elution buffer (EB_Q): 20 mM Tris pH 8, 1 M NaCl). The complex was washed with 0%, then 19% EB_Q and finally eluted with 23% EB_Q. Fractions containing the complex were concentrated (5000 MWCO Vivaspin Turbo 15) and loaded onto a S200 column (HiLoad 16/600 Superdex S200) equilibrated with 20 mM Tris pH 8 or sodium phosphate pH 8, 150 mM NaCl +/- 0.1 mM sodium molybdate. S200 complex eluate was concentrated to the desired concentration (5000 MWCO, Vivaspin Turbo 15) and used directly or frozen in LN₂ and stored at -80 °C. *E. coli* BL21 (DE3)-expressed complex was purified as described above but without sodium molybdate or vanadate.

2.2.1.6.3 *Grb2 protein*

His-Grb2 (human C32S, C198A mutant; pET-28(+) vector; Grb2: ca. 27 kDa) was purified with His-tag or the His-tag was cleaved off during purification. His-Grb2 expression was done in *E. coli* BL21 for ON at 25 °C. Bacteria were lysed in lysis buffer (LB) of 20 mM Tris or sodium phosphate pH 8, 500 mM NaCl, 5 mM imidazole, supplemented with protease inhibitors. A bacterial pellet from 1 litre

ON culture was resuspended in 100 ml LB and bacterial lysate supernatant (20K) was generated. All subsequent purification steps were performed at RT as concentrated Grb2 protein (ca. > 2 mg/ml) is better soluble at 23 °C (RT) than at 4 °C.

Details for protein expression and 20K preparation be found in 2.2.1.4. Filtered 20K (0.2 µm filter pore size) was loaded onto a cobalt-charged HiTrap IMAC column (5 ml or 2x 5 ml; Binding buffer (BB_{IMAC}): 20 mM Tris pH 8, 500 mM NaCl; Elution buffer (EB_{IMAC}): 20 mM Tris pH 8, 500 mM NaCl, 500 mM imidazole). The column was washed with 1% EB_{IMAC} until the UV signal was stable. His-Grb2 was either eluted immediately or the His-tag was cleaved on-column. His-tag cleavage was done with thrombin (250 Units thrombin/1 litre bacterial pellet) in 20 mM Tris pH 8, 150 mM NaCl for ON at RT on a nutator. His-tag cleaved Grb2 protein was eluted in 20 mM Tris pH 8, 150 mM NaCl. Thrombin was removed with a HiTrap Benzamidine column (1 ml). Fractions with Grb2/His-Grb2 were dialysed against 20 mM sodium phosphate pH 6.8 (2x 5 litres buffer, Spectra/Por, 1 kDa MWCO) and loaded onto a HiTrap Q column (HP; 5 ml; Binding buffer (BB_Q): 20 mM sodium phosphate pH 6.8; Elution buffer (EB_Q): 20 mM sodium phosphate pH 6.8, 1 M NaCl). The complex was washed with BB_Q and eluted with a 10% and 50% EB_Q step gradient. Clean Grb2/His-Grb2 fractions were collected, concentrated (5000 MWCO, Vivaspin Turbo 15) and loaded onto a S75 column (HiLoad 16/600 Superdex S75) equilibrated with 20 mM Tris pH 8, 150 mM NaCl. Grb2/His-Grb2 was finally concentrated to the desired concentration (5000 MWCO, Vivaspin Turbo 15) and used directly or frozen in LN₂ and stored at -80 °C.

2.2.1.7 Purification of recombinantly expressed untagged Grb2 protein

Human Grb2 (WT, Grb2 sequence from Grb2-pMCSG7 vector) was cloned in the pET-21d(x) vector containing no affinity-tag. Protein was expressed in *E. coli* BL21 for ON at 18 °C. Bacteria were lysed in ice-cold TPE lysis buffer supplemented with 1 mM Vanadate and protease inhibitors. Grb2 (untagged) was pulled down from the bacterial lysate supernatant with phospho-tyrosine-agarose (O-Phospho-L-tyrosine Agarose, Sigma) for 1 h at 4 °C on a nutator. Elution was done with free phospho-tyrosine (100 mM) for 1 h at 4 °C on a nutator. The pooled Grb2 eluate was dialysed against 3x 1 litre of 20 mM Tris pH, 220 mM NaCl (4 °C) and loaded onto a S75 column (HiLoad 16/600 Superdex S75). S75 fractions containing Grb2 were concentrated (5000 MWCO, Vivaspin Turbo 15). Protein aliquots were frozen in LN₂ and stored at -80 °C. As the Grb2 concentration purified from the untagged Grb2 construct was relatively low (< 1 mg/ml) compared to the purification with the His-tagged Grb2, purification steps for untagged Grb2 could be performed at 4 °C.

2.2.1.8 Protein concentration determination using Bradford and NanoVue Plus

The total protein concentration of an unknown sample was determined by Bradford assay. A protein standard curve was prepared using bovine serum albumin (BSA). 50 µl samples with BSA concentrations ranging from 2-12 µg/ml were mixed with 950 µl 1x Bradford solution. The sample

with the unknown concentration was diluted to an appropriate protein concentration (ca. 4-10 µg/ml) and also mixed with 950 µl 1x Bradford solution. After 10 min incubation, the $A_{595\text{ nm}}$ was determined by the photometer (BioPhotometer, Eppendorf). The unknown sample concentration was then calculated from the BSA standard curve.

The protein concentration of purified protein samples were measured on a NanoVue Plus Spectrophotometer (GE Healthcare) using extinction coefficients calculated in ProtParam (web.expasy.org/protparam).

2.2.2 Protein separation and detection

2.2.2.1 SDS polyacrylamide gel electrophoresis

Proteins were separated by size on a polyacrylamide gel according to the method of Laemmli (Laemmli, 1970). The protein was mixed with 2-4x SDS PAGE sample buffer and denatured at 100 °C for 5 min. The SDS PAGE sample buffer contains β-mercaptoethanol to reduce disulfide bonds in the protein and sodium dodecyl sulfate (SDS). SDS associates with hydrophobic residues of the protein and covers the protein with negative charges. The denatured protein sample is then loaded onto a discontinuous polyacrylamide gel which is composed of a stacking gel (pH 6.8) and a separating gel (pH 8.8). The separating gel was cast with a concentration appropriate to the molecular size of the proteins to be separated. The stacking gel (4%) ensures the simultaneous entry of the sample into the separating gel. The separating gel then separates the negatively charged proteins according to their size.

2.2.2.2 Native-PAGE

A native-PAGE separates a protein mixture in its native form according to the proteins migration ability which depends on its shape, overall charge and hydrodynamic size. Before the run, the gel chamber (Protean II xi electrophoresis chamber or Mini-Protean 3; Bio-Rad) was thoroughly cleaned to remove any residual SDS. Samples were resuspended in native gel sample buffer with a 1:1 or 1:2 ratio and loaded onto a non-denaturing polyacrylamid gel (stacking and separating gel with pH 8.8). The gel was run in native gel running buffer at 200 V for 2 to 4 h on ice at 4 °C. Stacking and separating gel were stained with Coomassie InstantBlue stain.

2.2.2.3 Detection of proteins immobilized on a cellulose membrane with antibodies (Western Blot)

The polyacrylamide gel was cut to an appropriate size and washed with semi dry blot buffer (SDBB). The proteins were then transferred from the gel onto a PVDF membrane (Fisher Scientific) via semi-dry blotting (TransBlotSD Semi-Dry Transfer Cell, Bio-Rad). Five layers of SDBB soaked chromatography paper (Whatman 3MM) were positioned on the Semi-Dry Transfer Cell. The PDVF

membrane was activated with methanol for 10 sec, washed with SDBB and placed on top of the Whatman paper. The gel was placed on top of the membrane and another set of five layers of SDBB soaked filter paper were put on top of the gel. Finally, air bubbles were removed. Proteins were blotted for 1 h at constant 20 V. To prepare the membrane for protein detection, non-specific binding sites on the membrane were saturated by incubation with a blocking buffer (skimmed milk or IgG-free BSA) appropriate for the antibody (1 h, RT on a nutator). The primary antibody incubation was done ON at 4 °C on a nutator. The membrane was washed with TBST (3x 15 min, nutator) before it was incubated with a secondary, HRP-coupled, antibody for 1 h at RT. The membrane was washed 3x 15 min with TBST. For the antibody detection, freshly mixed ECL detection solution (Thermo Scientific/Pierce) was left on top of the membrane for at least 1 min. The edges of the blot were briefly held against filter paper to remove excess ELC reagent. The light generated by the catalytic oxidation of the luminol in the ECL reagent by the antibody-coupled horseradish peroxidase was detected on a chemiluminescence imaging system (Fusion FX7™, Vilber Lourmat or G:BOX, Syngene).

2.2.3 Isothermal titration calorimetry (ITC)

Isothermal titration calorimetry (ITC) is an analytical method to explore binding processes, for example the kinetics of enzyme-catalysed reactions or the interaction between macromolecules, such as proteins, DNA or lipids (Doyle, 1997). The method is based on the accurate measurement of the heat associated with the binding process, as an endothermic reaction absorbs heat and an exothermic reaction releases heat.

ITC, in a common setup for protein-protein measurements, is able to measure dissociation constant (K_d) values between 10 nM and 0.1 mM (Ladbury, 2010). From a titration of the solution of one binding partner into the solution of the other one, parameters of the interaction as the binding affinity, the free energy of binding (ΔG), the enthalpy (ΔH) and the entropy (ΔS) of a binding reaction can be determined.

A representative result for an ITC measurement of a high-affinity binding event is shown in Figure 2.1. The top panel displays the actual measurements. The amount of energy used to compensate the temperature change induced by the interaction between the two molecules is shown in $\mu\text{cal/sec}$ in relation to the time of the measurement. Peaks are obtained by the consecutive injection of ligand solution to the cell which contains the macromolecule solution. The syringe is continuously stirring to mix the sample in the sample cell during ligand titration. After injection of a small volume of ligand, an equilibrium between ligand bound macromolecule and free macromolecule is established. The temperature difference between the sample cell and a reference cell, generated by the binding event, is constantly balanced. The energy used for this re-adjustment is calculated into heat generated or heat absorbed by the binding process.

A diffusion of very small volumes of ligand into the macromolecule solution during the setup of the measurement is unavoidable when the syringe is initially inserted into the sample cell solution. The

first injection contains typically less ligand solution than the succeeding injections to remove the mixture generated by this initial diffusion of solutions. The data point for this first injection is excluded from data analysis. Also, a diffusion of very small volumes of ligand into the macromolecule solution is unrestricted when the syringe is initially inserted into the sample cell solution. After several injections, the heat generated by the interaction is dropping quickly due to the reduced amount of free macromolecule. With an excess of ligand, all binding sites in the macromolecule are saturated and the heat generated approaches the heat generated by dilution alone. Therefore, control experiments using solely the buffer should be performed and subtracted from the original measurements to obtain accurate values.

During analysis of the raw ITC data, the peaks are integrated to yield the absolute amount of energy change for each injection. The binding isotherm (Figure 2.1, bottom panel) displays the energy change in kcal/mol injectant in relation to the molar ratio of the ligand to macromolecule. Fitting a curve to the data yields stoichiometry N , the binding enthalpy ΔH , and the equilibrium dissociation constant K_d , assuming the interaction occurred under equilibrium conditions.

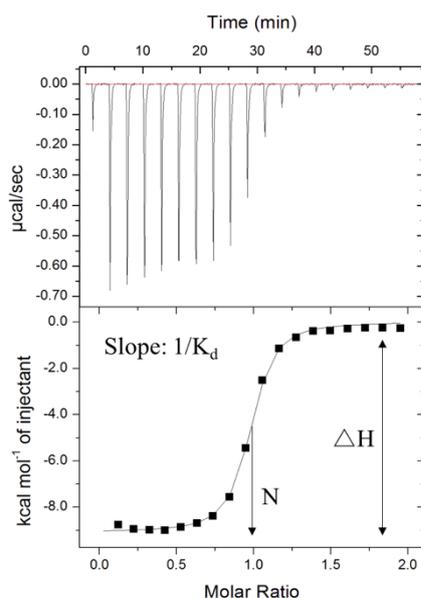


Figure 2.1 Representative ITC measurement

ITC analysis of the binding between GST-SH3C domain of Mona/Gads and a Gab2 peptide (Dr. Marc Lewitzky, AG Prof. Feller, MLU).

ITC measurements were done with a MicroCal™ iTC₂₀₀ calorimeter. 0.35 ml of the macromolecule solution (protein or protein domain), with a concentration between 0.02-0.05 mM in ITC buffer, was carefully transferred into the sample cell to prevent any air bubble formation. 0.1 ml of 0.2-0.5 mM of the small molecule (peptide) solution in ITC buffer was titrated from the syringe into the sample cell holding containing the macromolecule solution.

The titration was usually performed using 19 injections (2 μ l, first one 0.5 μ l) at an equilibrium temperature of 25 °C. The data was fitted using χ^2 minimization on a model assuming one set of sites

to determine the equilibration constant K_d . Data analysis was done with the MicroCal-enabled Origin™ software.

2.2.4 X-ray crystallography

X-ray crystallography is, besides nuclear magnetic resonance (NMR) spectroscopy and (Cryo)-Electron microscopy (EM), a powerful technique to study the three-dimensional structure of a protein. However, NMR and EM are methods with a size limitation, as NMR is used for proteins below 50 kDa and EM is usually applied on larger, well-ordered protein complexes.

In X-ray crystallography, the protein crystal is exposed to an X-ray beam and diffracted reflections are recorded. The diffraction pattern is then used to determine the 3D molecular structure of a protein by using mathematical tools, modelling and refinement programs. A major challenge in X-ray crystallography is the requirement that proteins need to crystallize. To induce crystal formation, a soluble protein is mixed with range of different precipitants such as polyethylene glycol or salts (McPherson, 1976, McPherson, 2001). Commercially available sparse matrix crystal screens make it possible to concurrently screen a large number of precipitants. The pre-screening and the optimization procedure are described in the following subchapter.

2.2.4.1 Initial crystallization-screening and optimization

Pre-screens were performed using the vapour diffusion sitting-drop method. This is a typical crystallization method based on the principles of diffusion. The drop usually contains a lower precipitant concentration than the reservoir. Hence, to obtain an equilibrium, water from the drop is moving to the reservoir in form of vapor. Due to the loss of water, the drop becomes gradually supersaturated until an equilibrium between the drop and reservoir concentration is achieved. Only very small amounts of protein and precipitant are used which allows testing of a large number of possible conditions. The protein was purified and concentrated to >10 mg/ml. Protein and peptide were mixed in a 1:1.5 molar ratio, trying to optimize binding according to the affinity determined prior to the crystal screen. 70 μ l reservoir solutions were filled into the reservoir chambers of MRC-2-well-sitting-drop-plates. Protein drops with 0.2 μ l protein and 0.2 μ l reservoir solution were mixed at 13 °C and placed on the plate by a Microsys SQ pipette robot (Cartesian). Plates were quickly sealed with foil and scanned by an automated imaging system (Desktop Minstrel UV-Imager, Rigaku). A list of all applied sparse matrix crystal screens can be found in 2.1.8. Selected crystals were pre-tested to examine if the crystal shape can diffract the X-ray beam and if the crystal consists of protein and not salt. The crystal was mixed with a cryoprotectant (10% (R,R)-(-)-2,3-Butandiol) to prevent ice-crystal formation at 100 K.

For follow-up fine screens, various precipitant concentrations and pH's were tested systematically. The vapour diffusion hanging drop method was used. This is a method similar to the vapour diffusion sitting-drop method except that the drop is hanging, not sitting, above the reservoir. In Easy Xtal 15-

or 24-well-hanging-drop-plates, 470 μ l reservoir solution were filled into the reservoir chambers. Those plates contain larger-sized wells compared to the pre-screens in order use more sample to increase the size of the crystal. Protein drops were placed underneath a cover slip and sealed properly. Two protein to reservoir solution ratios were tested, a 1:1 ratio (1 μ l protein : 1 μ l reservoir) which was also used for the pre-screen and a 2:1 ratio (2 μ l protein : 1 μ l reservoir). The plates were stored at 13 °C. Optimized crystallisation conditions of the Grb2 – Gab1 peptide complex are summarized in the following subchapter. Other crystallization screen setups are summarized in the appendix A.3.

2.2.4.2 *Crystallization of the Grb2 with a Gab1-derived peptide (32 aa peptide)*

Human Grb2 (His-tagged and His-tagged cleaved; C32S, C198A) was purified in 20 mM Tris pH 8, 150 mM NaCl as described in subchapter 2.2.1.6.3. 0.64 mM (17.7 mg/ml) His-Grb2 (C32S, C198A) or 1.19 mM (30.5 mg/ml) Grb2 (C32S, C198A) were mixed with the 32 amino acids-long Gab1 peptide (20 mM) in a molar ratio of 1:1.5 (protein : peptide). The Gab1 peptide (32 aa Gab1, SH3 tandem peptide, Gab1 497-FRSSPKTPRRPVPVADSEPPPVDRNLKPDRK-528, C514S) was purchased from JPT Peptide Technologies (Berlin, D).

Pre-screens were executed using the vapour diffusion sitting-drop method. The X-ray scattering properties of well-shaped crystals were tested. We found three interesting precipitant solutions for crystal formation (A) 0.05 M Tris-HCl pH 8.0, 28% PEG 3350 (Sigma Low Ionic Strength + Extension Kit), (B) 0.1 M KSCN, 30% PEGMME 2000 (Jena Bioscience JCSG ++ 1-4) and (C) 0.1 M HEPES pH 7.5, 20% 4000 PEG, 10% 2-propanol (Jena Bioscience Classic 1-4). However, only the fine screen of condition C revealed a reproducible crystal formation. A full data set from a crystal formed in 0.1 M Tris pH 8, 22% PEG 4000, 10% 2-Propanol was collected on the in-house Cu anode X-ray generator with a 2.41 Å resolution. Later, a second data set was taken on the synchrotron BESSY II (Helmholtz Centrum Berlin). Data recording, processing and structure determination was done by Dr. Constanze Breithaupt (Prof. Stubbs, MLU). The structural data discussed in chapter 5 is the current state of analysis from the in-house data. Structural models in figures were designed using PyMOL software (PyMOL Molecular Graphics System, Version 1.8.6.0 Schrödinger, LLC)

Chapter 3 Results and Discussion: Purification of the Gab2 protein

3.1 Expression and purification of recombinant Gab2 protein

We decided to initially concentrate on the purification of the Gab2 protein as a representative member of the Gab protein family since initial experiments had shown that a GST-Gab2 fusion protein (mouse Gab2, UniProt accession number: Q9Z1S8) was well expressed in *E. coli* (BL21). Protein expression was done for 5 h at 18 °C. IDP proteins, such as the Gab proteins, might be more susceptible for protein aggregation due to lack of a tertiary structure, and are therefore often favourably expressed at lower temperatures as hydrophobic interactions are temperature-dependent (Baldwin, 1986, Schellman, 1997) (see 2.2.1.5.1 for a detailed description of the GST-Gab2 purification protocol).

The expressed GST-Gab2 protein was purified by affinity chromatography and a subsequent size exclusion chromatography (SEC) step. GST-Gab2 from bacterial lysate was immobilized on GSH-sepharose beads (Figure 3.1, A, sample 'On beads'). The 100 kDa band most likely corresponded to the GST-Gab2 protein, as the GST-Gab2 construct is expected to have a size of 100 kDa (GST = approx. 26 kDa, Gab2 = approx. 75 kDa). As seen on the gel, the GST-Gab2 sample was contaminated with smaller-sized protein bands (Figure 3.1, A sample 'after washing'). These could be bacterial proteins or GST-Gab2 fragments resulting from an incomplete translation of the construct or due to cleavage by bacterial proteases which have not been inactivated by supplemented protease inhibitors. Gab2 was eluted with an excess of reduced glutathione and dialysed against 50 mM Tris pH 7.5 supplemented with freshly added β -mercaptoethanol (2 mM) each at 4 °C.

Gab2 separation on a gel filtration column (HiLoad 16/600 S75pg) indicated that most of the protein was highly aggregated. Gab2 eluted immediately after the void volume (VV) of the column (approx. 40 ml) (Figure 3.1, B and C). A second peak (approx. 57 ml) contains GST-tag only. The mass for the second peak, estimated from the elution volume, was approximately 47 kDa protein which would be consistent with a GST-dimer (ca. 50 kDa). The GST-tag can often be found to accumulate as a GST-dimer in the bacteria (Ji et al., 1992). It could originate from incomplete translation of the construct or due to an unwanted contamination of proteases which have not been inhibited with supplemented protease inhibitors. A sample taken from the pooled GST-Gab2 S75 peak and probed with an anti-GST antibody confirms that the approximately 100 kDa band is the GST-tagged Gab2 protein (Figure 3.1, D). Lower molecular weight bands could either result from GST-Gab2 fragments due to unspecific cleavage or incomplete expression or they could be non-specific background from non-related proteins.

A number of strategies can be employed to prevent the protein from aggregating. Varying the expression time from a short (5 h) to a longer expression (16 h) did not show any influence on the condition of Gab2 (data not shown). A reduction of the protein synthesis rate by decreasing the inducer (IPTG) concentration can sometimes increase the solubility of a protein. Furthermore, high-

salt buffers or solvent additives are known to stabilise the protein in its native form (Graslund et al., 2008). However, preliminary experiments with a lower IPTG concentration and various buffer additives such as glycerol, reducing agents and detergents or a high-salt buffer failed to prevent Gab2 from associating in multimeric forms (data not shown).

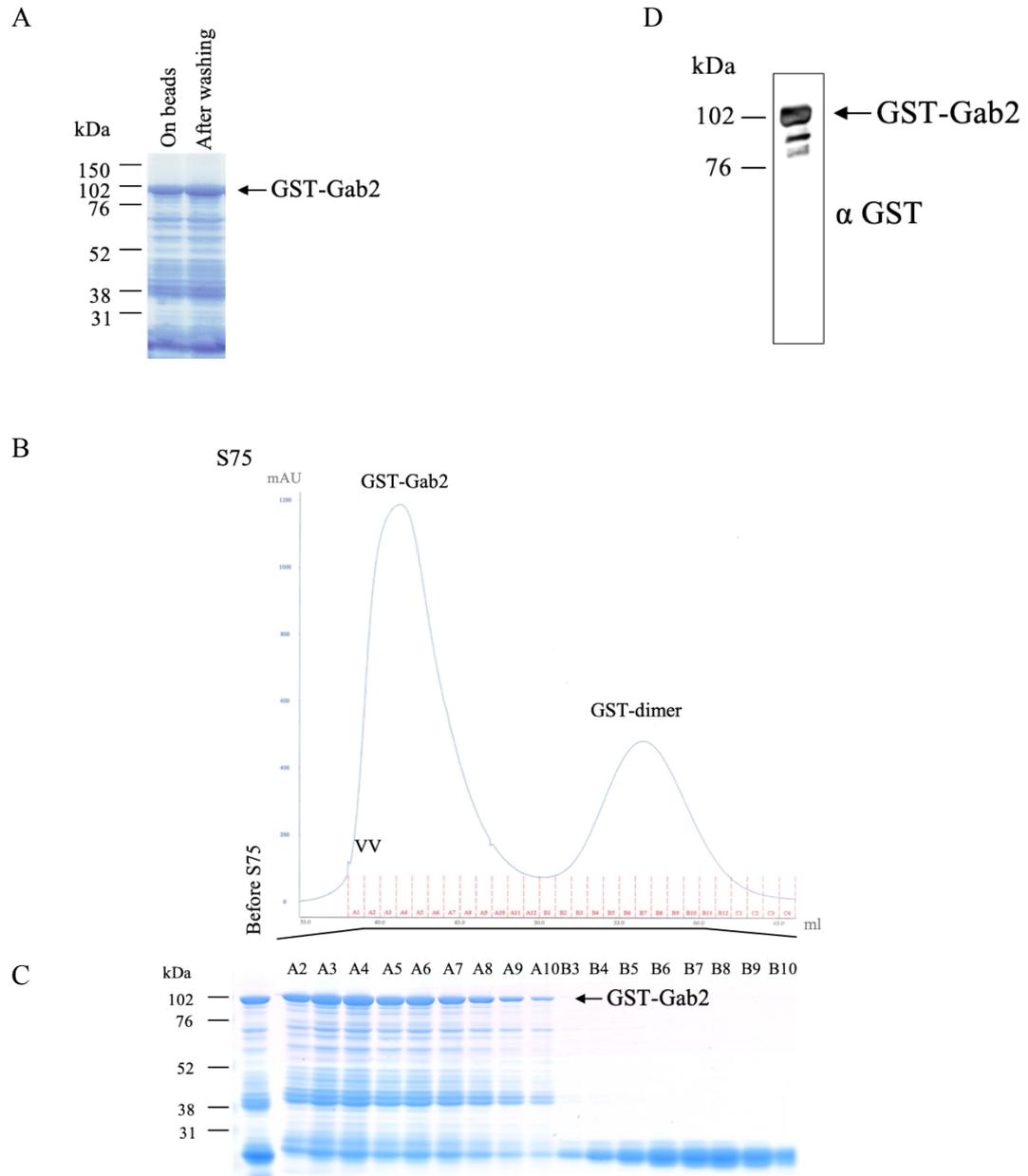


Figure 3.1 Purification of GST-Gab2

GST-Gab2 purification via affinity chromatography and size exclusion chromatography (SEC). (A) Affinity-purified GST-Gab2. (B) SEC analysis of the eluted GST-Gab1 protein sample by an Superdex S75 column (HiLoad™ 16/600 Superdex™ S75). The x-axis shows the elution volume of the sample in [ml]. The y-axis measures the UV absorbance in milli Absorbance Units [mAU]. (C) GST-Gab2 S75 fractions were resolved by SDS PAGE (9% separating gel) and stained with Coomassie InstantBlue. VV: Void volume of the column. (D) A sample from pooled GST-Gab2 S75 eluate was separated by SDS PAGE (7% separating gel) and immunoblotted with an anti-GST antibody. Protein marker bands are indicated on the left side.

Within the cell, the native Gab2 protein becomes tyrosine phosphorylated as a result of signalling processes in order to interact with SH2 domain containing proteins. Negatively charged phosphate groups (PO_4^{3-}) located on the protein surface can make a protein more hydrophilic and therefore could increase protein solubility when expressing it in a non-cellular environment. Furthermore, phosphorylation is required for the later on assembly of Gab-based complexes. The *E. coli* BL21 derivative strain TKB1 harbours the tyrosine kinase domain of the murine ephrin type-B receptor I and, unlike other *E. coli* strains that have only serine/threonine kinases, phosphorylates tyrosine residues of expressed proteins as well. GST-Gab2 expression in TKB1 led to Gab2 tyrosine phosphorylation. However, SEC analysis of phosphorylated Gab2 still displayed most of the Gab2 protein aggregated (data not shown). It is generally known that protein affinity tags can directly influence the solubility of a protein (Kimple et al., 2013, Lebendiker and Danieli, 2014). The relatively large glutathione S-transferase tag (GST, ca. 25 kDa) often enhances protein solubility, but it has also been shown to introduce protein aggregation owing to the GST dimerisation (Lebendiker and Danieli, 2014). Therefore, a removal of the GST tag could improve the protein solubility. However, preliminary GST test cleavages indicated an increase in Gab2 instability rather than an improvement of Gab2 solubility (data not shown). Other tags which are known to increase protein solubility could be tested for the Gab2 expression, as for example, the *E. coli* maltose binding protein (MBP, 40 kDa). Although, a detailed mechanism of how MBP improves protein solubility is still unknown, studies suggest that MBP has a chaperone-like ability that assists the folding of its fused to protein (Bach et al., 2001, Fox et al., 2001).

Protein purification goals are to obtain protein in its native form but also to gain high sample purity (>95%). Residual contaminations are usually removed by ion chromatography or SEC but, we observed a large number of impurities which were neither removable by extensive washing nor by SEC (see Figure 3.1, A). The co-purified proteins could be bacterial proteins which non-specifically stick to the Gab2 protein or the affinity material. One possibility to remove contaminants would be a denaturing step with for example high concentrations of guanidine hydrochloride (6 M) or Urea (8 M). While immobilization of the affinity tagged protein, denatured contaminants including bacterial proteases can be completely removed from the sample. In a subsequent step, the protein of interest is then refolded in its native state. As denatured GST is unable to bind to the affinity matrix, the affinity tag must be replaced with a more suitable tag, e.g. a polyhistidine tag. However, not every protein can be refolded to its native state. Some of the contaminants could also be GST-Gab2 fragments as a result of unspecific protease cleavage. IDPs might be more sensitive to protein degradation or fragmentation due to the full exposure of unstructured regions to the solvent. The Gab2 immunoblot with an anti-GST antibody indeed indicated a few shorter GST-Gab2 fragments (Figure 3.1, D). To remove incomplete recombinant protein, a second affinity tag could be introduced at the C-terminus. Protein purified using both tags consecutively should only be full length.

3.2 Conclusion and discussion

Attempts to purify recombinant GST-Gab2, expressed in *E. coli*, demonstrated that most of the protein formed soluble protein aggregates. Studies on IDPs have shown that the tendency to form protein aggregates is higher in IDPs which is most likely due to their lack of tertiary structural elements (Wright and Dyson, 1999, Lebendiker and Danieli, 2014, Oldfield and Dunker, 2014). The protein aggregation observed in the Gab2 purification process was addressed by modifying the protein expression and by expression in a different bacterial strain to introduce phosphorylation, but without success. We also started to examine a few common buffer additives by simply adding them into the buffers used in the purification process. This was time-consuming and not very efficient. As the list of buffer additives that potentially can increase the protein solubility is enormous, an efficient screening procedure could be used for further optimisation. Churion & Bondos published a detailed and easy-to-use protocol which can be applied to identify an optimal buffer composition for IDPs. The screen hierarchical tests various buffer additives, their concentration and potential combinations by using a filter-based aggregation assay (Churion and Bondos, 2012).

Protein aggregation can occur when the protein concentration exceeds the specific solubility limit of a protein. But there are also other reasons for protein aggregation such as the lack of native disulfide bond formations or the missing of a chaperone or essential partner protein (Lebendiker and Danieli, 2014). The multimeric self-association of the native protein conformation or of structurally altered protein states that can occur in a soluble and insoluble form (Lebendiker and Danieli, 2014). Insoluble aggregates precipitate out of solution and can be removed by centrifugation or filtration. Insoluble aggregates of a recombinant protein inside the bacterium often appear as so called inclusion bodies (Ventura and Villaverde, 2006). In contrast, soluble aggregates are completely dissolved in the buffer and are only detectable by methods such as ultracentrifugation, SEC, or dynamic light scattering (DLS). A final SEC is usually applied in a protein purification to separate the sample from possible occurring soluble aggregates.

The observed Gab2 instability during purification could be a more general issue for proteins from the Gab family as similar problems were encountered when expressing the Gab1 protein (Dr. Terumasa Sowa, unpublished data). One strategy to overcome challenges with difficult protein purifications is a divide-and-conquer approach. The protein of interest is either divided into functional fragments, domains, or existing functional fragments which are then used to make a step-by-step analysis (Gaudet, 2009, Edwin et al., 2014). As a functional Gab1 fragment (p35Gab1) already exists, and Gab2 and Gab1 are very similar on a structural level, we decided to focus next on the p35Gab1.

Chapter 4 Results and Discussion: Purification of the p35Gab1 – Grb2 complex

4.1 Overview

The p35Gab1 is a naturally occurring and functional fragment of the full-length Gab1 protein that performs anti-apoptotic functions when the cell is under moderate stress (Le Goff et al., 2012).

The p35Gab1 fragment was well expressed in *E. coli* but most of the protein was aggregated as shown by size exclusion chromatography (see subchapter 4.3). Gab1 belongs to the class of IDPs which are more prone to form protein aggregates than other proteins (Breydo and Uversky, 2011). One strategy to improve the expression of aggregation-prone proteins is the co-expression with a native interaction partner (Romier et al., 2006).

The Gab1 protein interacts with a range of different signalling proteins (Figure 4.1) but many of these, such as the p85-PI3K, Crk/CrkL and Shp2, are not constitutively bound to Gab1 and need tyrosine phosphorylation of consensus motifs (indicated as P) for binding. The interaction between Gab1 and the c-Met receptor via the c-Met binding region (MBR) is initiated by hepatocyte growth factor (HGF) stimulation (Schaeper et al., 2000).

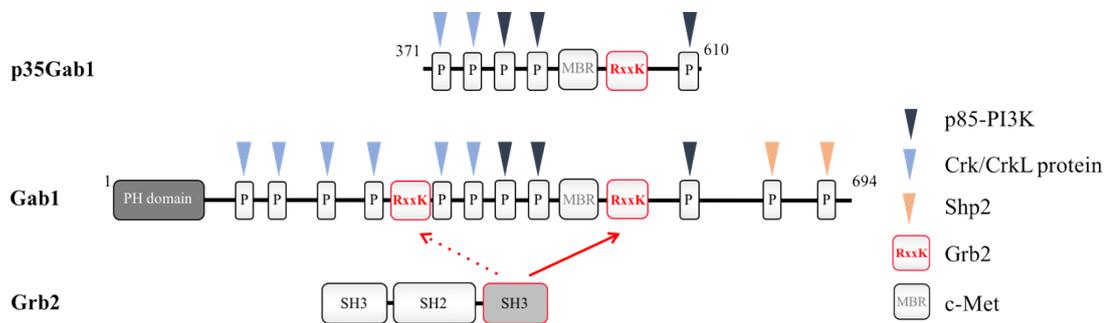


Figure 4.1 Schematic view of Gab1, p35Gab1 and Grb2

Stick representations of p35Gab1, Gab1 and Grb2. The Grb2 SH3C domain binds to two R-x-x-K motifs in Gab1 (Harkiolaki et al., 2009). The bold red arrow indicates the stronger Gab1 – Grb2 SH3C interaction site (K_d 3 μ M), the dotted red arrow indicates the second, weaker, Gab1 – Grb2 SH3C binding site (K_d 30 μ M). Boxes with the letter P indicate known tyrosine phosphorylation sites, that serve as binding sites for PI3K, Crk/CrkL and Shp2 proteins.

By contrast, the interaction between Gab1 and the Grb2 adaptor protein is independent of external signalling events. Co-immunoprecipitation assays indicated a constitutive binding between Grb2 and Gab1 (Holgado-Madruga et al., 1996). Grb2, composed of two flanking SH3 domains and one central SH2 domain, can interact via its C-terminal SH3 domain (SH3C) with two motifs (R-x-x-K) in Gab1, albeit not simultaneously (Maignan et al., 1995, Lock et al., 2000, Lewitzky et al., 2001, Harkiolaki et al., 2009). The two R-x-x-K Grb2 binding sites in Gab1 possess distinct binding affinities with a

stronger (K_d 3 μ M) and a weaker binding site (K_d 30 μ M). The p35Gab1 fragment only contains the higher affinity Grb2 binding site (Le Goff et al., 2012).

Co-expression of p35Gab1 with its interacting partner Grb2 greatly reduced the propensity for aggregation of p35Gab1. Protein expression was done in *E. coli* BL21 and TKB1. The TKB1 strain was used to allow tyrosine phosphorylation of the SH2 binding motifs in Gab1 as we were interested in a subsequently step-wise assembly of whole Gab signalling complexes.

The purification of p35Gab1 co-expressed with Grb2 as well as an initial test with p35Gab1 alone is described in the following subchapters 4.2 and 4.3.

4.2 Establishing a purification protocol for the p35Gab1 – Grb2 complex

The p35Gab1 – Grb2 was purified by an initial IMAC, an intermediate Mono Q step and a final SEC. An N-terminal polyhistidine-tagged p35Gab1 construct (human WT Gab1 aa 371 - aa 610; pET200-vector) for bacterial expression was kindly provided by Dr. Fafeur (Le Goff et al., 2012). We used wildtype (WT) and cysteine-free mutant His-p35Gab1. All p35Gab1 cysteines were mutated (C374A, C405A, C514A; Dr. Tobias Gruber, MLU) to prevent disulfide bond formation within p35Gab1 or with cysteines from unspecific *E. coli* proteins and thereby to inhibit potential protein misfolding or aggregation. Expressed mutated p35Gab1 (C374A, C405A, C514A) showed a similar SEC profile as WT p35Gab1.

For Grb2, we used WT and mutant Grb2 (C32S, C198A). Grb2 (C32S, C198A) had been used previously in Grb2 NMR studies (Yuzawa et al., 2003) and it is structurally very similar to WT Grb2 (PDB code: 1gri, Maignan et al, 1995). During purification of Grb2 alone, we found that homodimerisation of the WT Grb2 protein is reduced for the Grb2 (C32S, C198A, data not shown).

The purification protocol of the p35Gab1 – Grb2 complex is described in detail in the following subchapters (4.2.1 - 4.2.3).

4.2.1 Immobilized Metal Affinity Chromatography (IMAC) of the p35Gab1 – Grb2 complex

4.2.1.1 His-p35Gab1 co-expression with His-tagged Grb2 (*E. coli* BL21)

E. coli (BL21) was transformed with a His-p35Gab1 expression vector (WT p35Gab1; pET200) or co-transformed with a His-p35Gab1 (WT p35Gab1; pET200-D-TOPO) and a His-Grb2 expression vector (WT Grb2; pMCSG7). His-p35Gab1 and His-p35Gab1 – His-Grb2 were purified by immobilized metal affinity chromatography (IMAC) and IMAC fractions were analysed by SDS PAGE. Low imidazole concentrations (20% elution buffer (EB), 100 mM imidazole) eluted weakly bound protein

contaminants in the His-p35Gab1 and His-p35Gab1 – His-Grb2 sample (Figure 4.2). His-p35Gab1, when expressed alone, eluted with 50% elution buffer (250 mM imidazole) along with a series of unspecific proteins (Figure 4.2, A). In the His-p35Gab1 – His-Grb2 co-expression sample, His-p35Gab1 and His-Grb2 eluted in a single peak (40% EB; 200 mM imidazole; Figure 4.2, B). A comparison of His-p35Gab1 and His-p35Gab1 – His-Grb2 IMAC fractions that were separated by SDS-PAGE clearly shows a very substantial increase of sample purity for His-p35Gab1 – His-Grb2 co-expression compared to His-p35Gab1 alone.

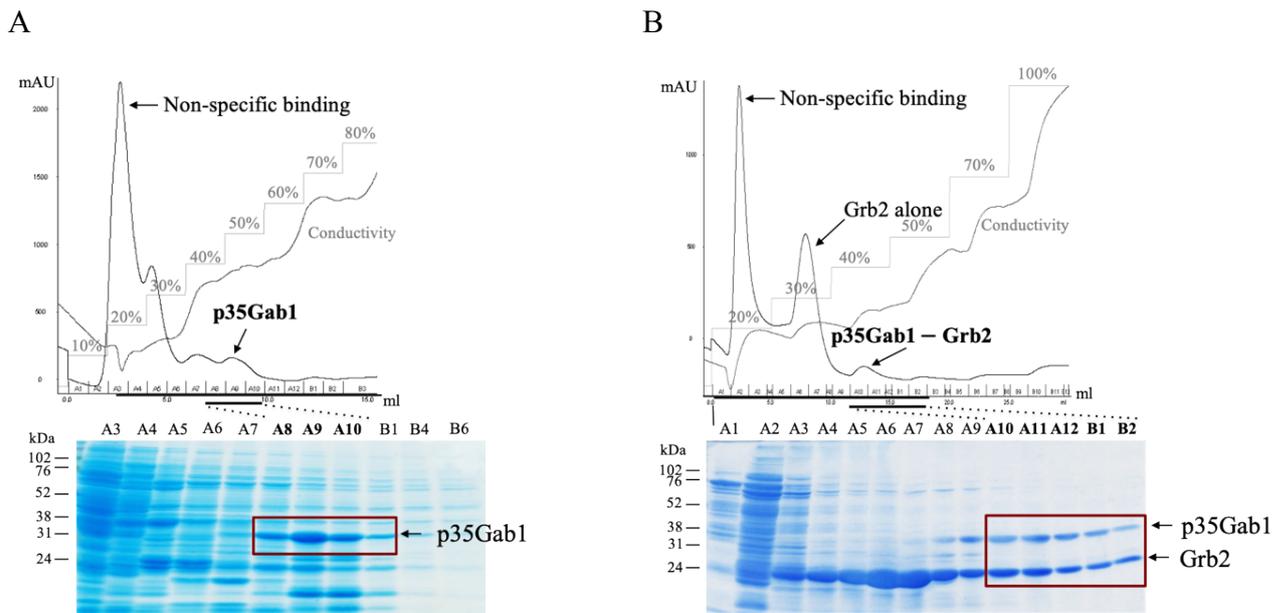


Figure 4.2 IMAC of His-p35Gab1 alone and His-p35Gab1 co-expressed with His-Grb2

Imidazole step gradient (0-0.5 M imidazole, percentages are indicated) elution of His-p35Gab1 (WT) (A) or His-p35Gab1 (WT) – His-Grb2 (WT) (B) from HiTrap IMAC column (cobalt charged). The conductivity increases during the run due to the increasing concentration of imidazole. UV absorbance is indicated on the y-axis in [mAU]. The run volume denoted in [ml] is shown on the x-axis. Selected fractions were analysed by SDS PAGE (12% resolving gel). Gels were stained with Coomassie InstantBlue and protein markers are indicated on the left. Bands containing the protein(s) of interest are marked with a red box.

The IMAC purified His-p35Gab1 – His-Grb2 sample contained a large excess of His-Grb2 (peak at 30% EB; 150 mM imidazole) due to its apparently high expression level and independent binding of the His-Grb2 to the affinity matrix via its own His-tag (Figure 4.2, B; fraction A6 and A7). Therefore, His-p35Gab1 was thereafter co-expressed with an untagged Grb2 construct (human WT and Grb2 C32S, C198A; pET-21d (+) vector) which is described in the following subchapter .

4.2.1.2 His-p35Gab1 co-expression with untagged Grb2 (*E. coli* TKB1 strain)

His-p35Gab1 (WT and p35Gab1 C374A, C405A, C514A) and untagged Grb2 (C32S, C198A), co-expressed in TKB1, were analysed on an IMAC HiTrap column. A linear gradient from 0 to 0.5 M

imidazole was chosen as an initial starting point to identify an imidazole concentration for washing and elution of His-p35Gab1 – Grb2 complex (Figure 4.3, A). The complex eluted in a rather broad peak, which could indicate an interaction of the protein with the column resin. The co-expression of His-p35Gab1 with untagged Grb2 significantly reduced the amount of Grb2 that was not co-eluting with His-p35Gab1 seen with the His-p35Gab1 – His-Grb2 complex (compare Figure 4.2 and Figure 4.3). Also, the His-p35Gab1 – Grb2 complex eluted at relatively low imidazole concentrations compared to the His-p35Gab1 – His-Grb2 complex which bound more strongly to the column, probably due to the presence of two His-tags per complex.

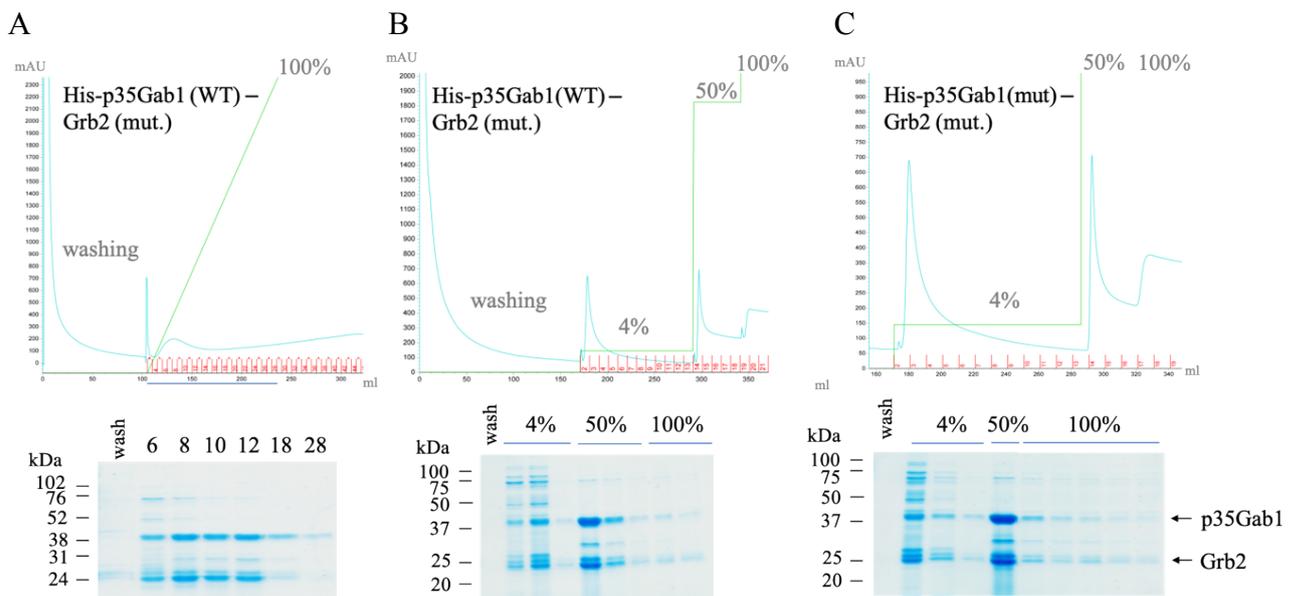


Figure 4.3 Testing of different IMAC gradients for His-p35Gab1 – Grb2 elution

Optimization of IMAC purification of His-p35Gab1 – Grb2 (protein expressed in TKB1). (A) Elution of His-p35Gab1 (WT) and Grb2 (untagged; C32S, C198A) from a HiTrap IMAC column with a linear imidazole gradient (40 column volumes). (B) Imidazole step gradient for elution of His-p35Gab1 (WT) co-expressed with Grb2 (untagged; C32S, C198A). (C) Imidazole step gradient for elution of cysteine-free mutant His-p35Gab1 (C374A, C405A, C514A) co-expressed with Grb2 (untagged; C32S, C198A). The HiTrap columns were charged with cobalt ions. The blue line indicates the UV absorbance in [mAU], shown at the y-axis, and the green line indicates the set percentage of elution buffer in [%]. The x-axis shows the elution volume of the sample in [ml]. IMAC wash samples and selected IMAC fractions were analysed by SDS PAGE (12% resolving gels). Gels were stained with Coomassie InstantBlue. The protein marker is indicated on the left.

In a second step, the His-p35Gab1 – Grb2 complex sample was eluted with a step gradient to better exploit the p35Gab1 – Grb2 complex elution behaviour that was observed for the linear gradient (Figure 4.3, B). With 4% EB (20 mM imidazole) non-specifically bound proteins were removed from the sample. A consecutive step to 50% EB (250 mM imidazole) was introduced to elute the entire bound complex in one step before washing the column with 100% EB.

A complex comprised of cysteine-free His-p35Gab1 (C374A, C405A, C514A) with Grb2 (C32S, C198A) showed a similar IMAC elution profile as wildtype His-p35Gab1 complexed with Grb2

(C32S, C198A) (Figure 4.3, compare B and C). The peak trailing observed in chromatograms (B) and (C) could be resulting from an interaction between column resin and sample. The increase of the UV absorbance in the 100% EB step is most likely a result of a high imidazole concentration (500 mM imidazole) as imidazole is known to absorb at 280 nm. Also, SDS PAGE-analysed fractions from the 100% EB step contained no protein. The protein complex was further purified by anion exchange chromatography on a Mono Q column as described in the following subchapter.

4.2.2 Anion exchange chromatography of the p35Gab1 – Grb2 complex

The p35Gab1 – Grb2 protein complex has an isoelectric point (pI) of 5.8 as estimated by isoelectric.ovh.org (Bjellqvist et al., 1993, Bjellqvist et al., 1994, Wilkins et al., 1999, Gasteiger et al., 2005). Therefore, it is expected that the complex acquires a net negative charge at a buffer pH above pH 5.8 and can be captured on a positively charged anion exchange column resin as in the Mono Q column. Before Mono Q chromatography, contaminating particles or protein precipitates which could possibly obstruct the Mono Q column were removed by centrifugation. The 30 kDa contaminant band in the His-p35Gab1 – Grb2 sample completely disappeared after centrifugation along with a slight loss of the complex (Figure 4.4, A, compare ‘before spin’ (-), ‘after spin’ (circular arrow)). The minor reduction of the total amount of complex could be partly due to aggregation of an insoluble complex fraction. A pellet sample with a comparable amount to the ‘before spin’ and ‘after spin’ samples showed a relatively small amount of complex. A flow through (FT) and a wash sample were free of protein which indicated a binding of all of the remaining complex onto the anion exchanger column material.

In an initial experiment, a small-step salt gradient (0-1 M NaCl) elution of the His-p35Gab1 – Grb2 sample was performed (Figure 4.4, A; sample: His-p35Gab1(WT) – Grb2 (untagged; C32S, C198A)). The EB concentration was initially manually increased after each peak. An elution of complexes at different salt concentrations was observed which suggests that the sample contains a mixture of complex species. One possibility could be a variation in the phosphorylation patterns as phosphorylation strongly affects the pI value.

A two-step Mono Q protocol with a 19% EB (190 mM NaCl) washing and a 23% EB (230 mM NaCl) elution step significantly reduced contaminant bands and yielded highly purified His-p35Gab1 – Grb2 complex sample (Figure 4.4, B; sample: His-p35Gab1 (C374A, C405A, C514A) – Grb2 (untagged; C32S, C198A); EB: elution buffer). Washing the column with 100% EB (1 M NaCl) regenerated the column material for the following run.

Results and Discussion: Purification of the p35Gab1 – Grb2 complex

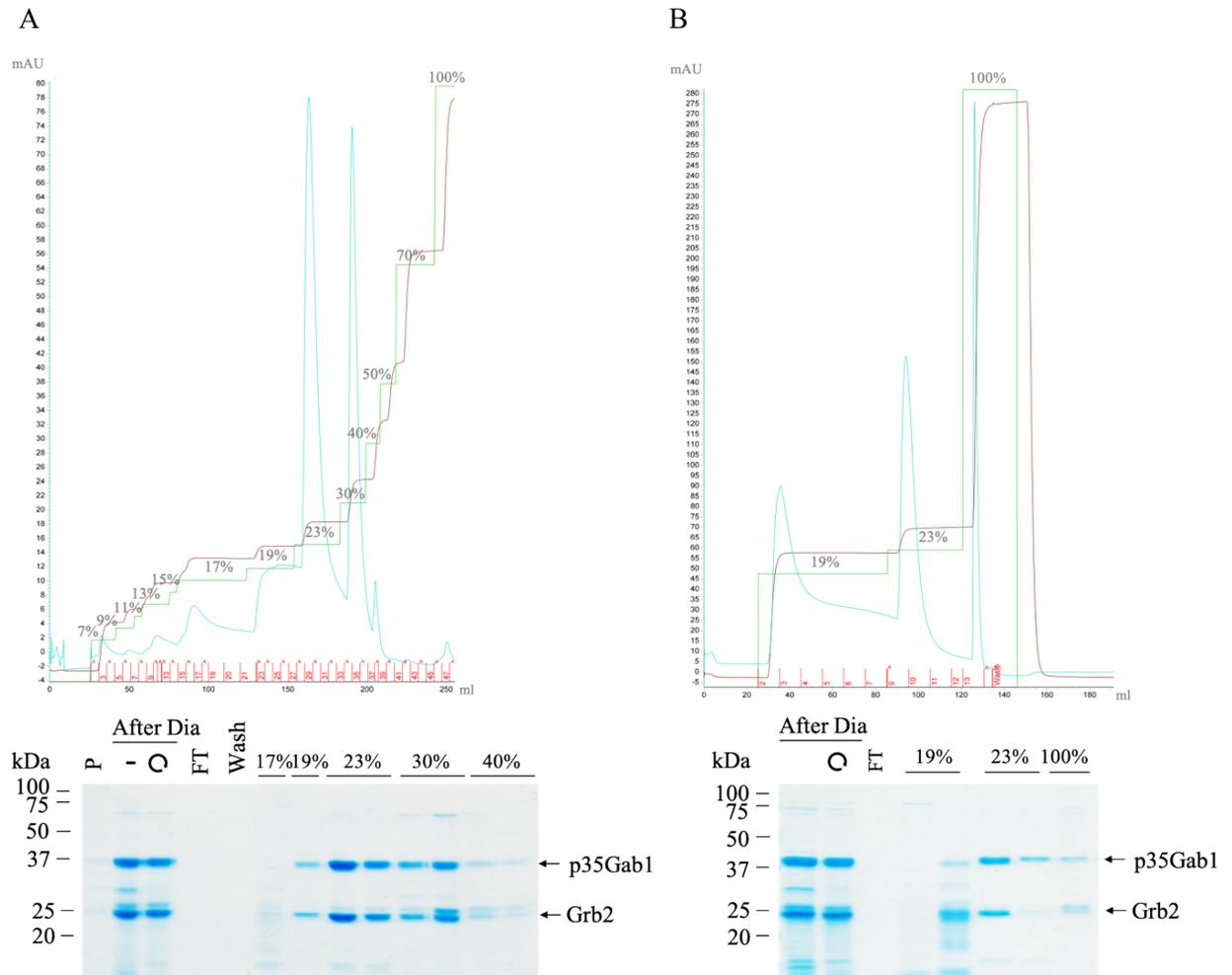


Figure 4.4 Mono Q purification of His-p35Gab1 – Grb2 complex

Grb2 (untagged; C32S, C198A) was co-expressed in *E. coli* TKB1 with wildtype His-p35Gab1 (A) and cysteine-free mutant His-p35Gab1 (C374A, C405A, C514A) (B). Samples were loaded onto a HiTrap Q column and eluted with a thirteen-step salt gradient (0-1M NaCl) (A) or a three-step salt gradient (0-1M NaCl) (B). The blue line indicates the UV absorbance in [mAU], shown at the y-axis, the green line set the EB percentage in [%] and the brown line the conductivity in [mS/cm]. The x-axis shows the elution volume of the sample in [ml]. Selected fractions (>5 mAU) were analysed on a SDS PAGE (12% resolving gel). Both gels were stained with InstantBlue. The protein markers are indicated on the left side of the gel.

While expressing the His-p35Gab1 – Grb2 complex in *E. coli* BL21, the complex precipitated after decreasing the high salt concentration to 150 mM NaCl (data not shown; sample: His-p35Gab1 (C374A, C405A, C514A)/ Grb2 (untagged; C32S, C198A)). Initial experiments showed that glycerol or sodium sulphate, along with a working temperature of around 23 °C, increased the solubility of the BL21-expressed complex but more data on this is definitely needed. This could indicate that the additional phosphate groups on the protein complex, obtained by the expression in TKB1, increase the complex stability.

4.2.3 ‘Polishing’ step of the p35Gab1 – Grb2 complex purification

The p35Gab1 – Grb2 complex purification protocol (IMAC, Mono Q) was followed by a size exclusion chromatography for removing remaining sample impurities, buffer exchange and for analysing the approximate molecular mass of the complex (Figure 4.5). The His-p35Gab1 – Grb2 complex eluted from a Superdex S200 column in a single peak at approx. 70 ml (Figure 4.5). A complex comprised of cysteine-free mutant His-p35Gab1 (C374A, C405A, C514A) with Grb2 eluted very similar to the complex with wildtype His-p35Gab1 and Grb2 (Figure 4.5, B).

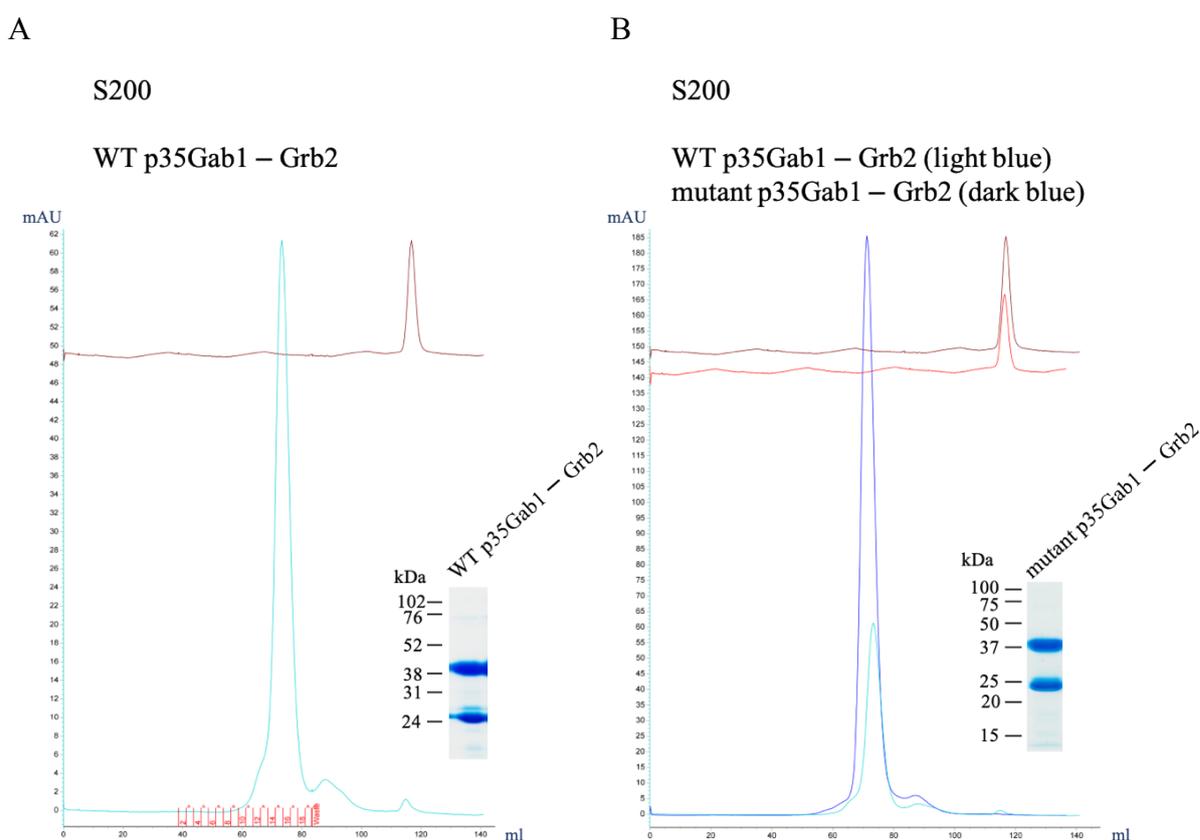


Figure 4.5 SEC analysis of Grb2 in complex with wildtype or cysteine-free mutant His-p35Gab1

Grb2 (untagged; C32S, C198A) was co-expressed in *E. coli* TKB1 with WT His-p35Gab1 (A) or cysteine-free mutant His-p35Gab1 (C374A, C405A, C514A) (B) (in B, UV chromatograms for cysteine-free mutant His-p35Gab1 (dark blue) and WT His-p35Gab1 (light blue, from A) are overlaid). Samples were separated on a SEC column (HiLoad™ 16/600 Superdex™ S200). Complex samples were resolved by SDS PAGE (12% resolving gel). SEC buffer: 20 mM Tris pH 8, 150 mM NaCl. The light and dark blue line chart the UV absorbance in [mAU], indicated at the y-axis. The x-axis shows the elution volume of the sample in [ml]. Protein gels were stained with Coomassie InstantBlue. Marker lines are indicated on the left.

According to a S200 calibration curve, the elution volume of the complex would correspond to a protein-complex with an approximate molecular mass of around 99 kDa. The SEC calibration calculation is based on elution profiles of globular proteins which may differ from the elution behaviour of complexes involving intrinsically disordered proteins. However, it could still suggest that

p35Gab1 is bound to two Grb2 molecules, for example a Grb2 dimer (1x35 kDa (p35Gab1) + 2x 27 kDa (Grb2) = ca. 89 kDa). This complex ratio was also observed in the Grb2 – Gab1 peptide crystal structure (see subchapter 5.6).

4.3 Test purification of the p35Gab1 fragment alone

After an initial immobilized metal affinity chromatography purification step, His-p35Gab1 was analysed by size exclusion chromatography (Figure 4.6). A detailed description of the His-p35Gab1 expression and purification protocol can be found in 2.2.1.6.1.

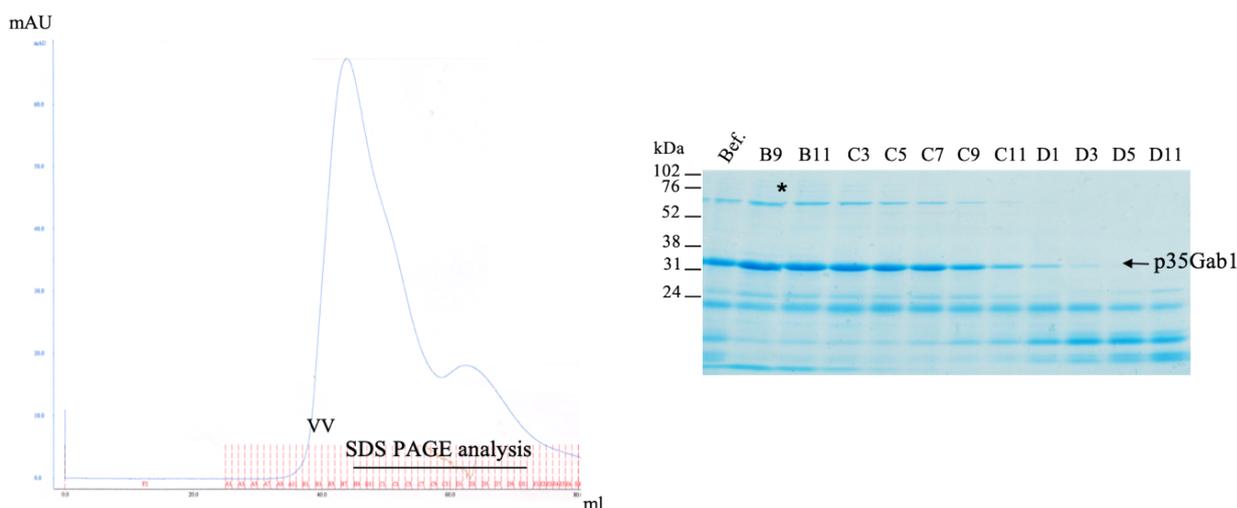


Figure 4.6 SEC analysis of His-p35Gab1

His-p35Gab1 (WT) was separated on a SEC column (HiLoad™ 16/600 Superdex™ S75). The blue line charts the UV absorbance in [mAU], shown at the y-axis. The x-axis shows the elution volume of the sample in [ml]. Selected S75 fractions were resolved by SDS PAGE (12% resolving gel) and stained with Coomassie InstantBlue. Marker lanes are indicated on the left. VV: void volume of S75 column. Black line in chromatogram indicates SEC samples analysed by SDS. Bef.: Before, Sample before S75.

The p35Gab1 SEC chromatogram and SDS PAGE-analysed SEC fractions indicated that most of His-p35Gab1 protein was aggregated. His-p35Gab1 eluted directly after the column void volume (VV; approx. 40 ml), which suggests that it was associated with high molecular weight protein aggregates. The second peak contained mainly low molecular weight proteins (<35 kDa), which could be bacterial proteins or possible p35Gab1 cleavage products. Attempts to increase the protein solubility of p35Gab1 by adding glycerol (5%) or by increasing the NaCl concentration (500 mM) could not reduce the propensity for aggregation (data not shown).

His-p35Gab1 co-eluted with a number of other proteins. One contaminant protein (indicated by an asterisk *, Figure 4.6) was identified as the bacterial chaperone DnaK (70 kDa), a Hsp70 homolog of *E. coli* by mass spectrometry (Prof. Kessler, Target Discovery Institute, University of Oxford) (Hestekamp and Bukau, 1998). DnaK is a common contaminant in recombinant protein expression in

E. coli (Birringer et al., 2006, Agou et al., 2002, Fleischer et al., 2007). One method to remove unwanted DnaK is by washing with magnesium (Mg^{2+}) and adenosine triphosphate (ATP) (Rial and Ceccarelli, 2002). This was not tested due to the lack of time.

4.4 Conclusion and discussion

Full-length p35Gab1 protein was well expressed in *E. coli* BL21 but it aggregated during purification demonstrated by SEC analysis. While testing various strategies to prevent p35Gab1 from aggregation, we obtained the most promising result with a co-expression approach. Co-expressing p35Gab1 with its binding partner, the Grb2 adaptor, substantially improved the protein stability of p35Gab1 and decreased the binding of p35Gab1 to non-specifically associated proteins.

Another co-expression strategy, used to improve the purification of a low-soluble protein, is the co-expression with a bacterial chaperone system such as the DnaK or / and the GroEL/GroES strategy (Goloubinoff et al., 1989, Georgiou and Valax, 1996, Fink, 1999, Schlieker et al., 2002, Sorensen and Mortensen, 2005). However, others have reported that this approach may only lead to an increase in the amount of soluble aggregates but not of properly folded proteins (Haacke et al., 2009) and for an IDP like p35Gab1 this does not seem promising.

The co-expression of His-p35Gab1 with Grb2 (untagged or His-Grb2) resulted in stable p35Gab1 – Grb2 complex formation. With WT or cysteine-free mutant His-p35Gab1 showing no difference. The use of untagged Grb2 compared to His-Grb2 in the p35Gab1 – Grb2 complex enormously reduced the amount of extra Grb2 which was seen in the His-p35Gab1 – His-Grb2 complex. The complex was purified by a three-step protocol consisting of an initial IMAC step, followed by a Mono Q purification and a final SEC step. We observed differences in the complex stability of BL21- and TKB1-expressed complexes. p35Gab1 – Grb2 expressed in BL21 was less stable and only reached low maximal protein concentrations ($< 1\text{ mg/ml}$), whereas complex expressed in TKB1 could be concentrated to 10 mg/ml or even 30 mg/ml . This could be due to the addition of phosphate groups on the surface of the complex that increases the hydrophilicity of the protein. The gradual elution of the p35Gab1 – Grb2 complex (expressed in TKB1) from the Mono Q column suggests a possible variation in the tyrosine phosphorylation patterns. A variation in the phosphorylation pattern is not unlikely, as the expressed complex contains 10 tyrosine residues in total (5xY residues on the surface of Grb2; 5xY residues in p35Gab1) which have been found phosphorylated (www.phosphosite.org, UniProt accession codes: human aa 371 – aa 610 of Gab1 protein, Q13480; human Grb2 protein, P62993). A method to detect and quantify p35Gab1 – Grb2 complex species varying in their degree of phosphorylation would be for example phosphopeptide mass spectrometry (McLachlin and Chait, 2001) or chromatofocussing to separate complex species with distinct pI values (Pepaj et al., 2006).

An alternative technique to potentially fine tune phosphorylation on the p35Gab1 – Grb2 complex could be an *in vitro* phosphorylation of the non-phosphorylated complex with purified tyrosine kinases / kinase domains.

Although, p35Gab1 was purified with its binding protein Grb2, the goal of single purified p35Gab1 or even full length Gab1 or Gab2 for structural analysis still remains. As optimisation approaches of p35Gab1 or full length Gab2 expression in *E. coli* showed no improvement in the protein solubility, other commonly used protein expression systems could be tested such as yeast, insect or mammalian cells (Bill, 2014). Another approach would be to dissociate p35Gab1 from Grb2 after initial co-purification. This would have the advantage that during purification p35Gab1 is stabilised by Grb2. An excess of a Gab1 peptide, that competes for the same binding site as p35Gab1 on Grb2, could lead to a separation of Grb2 and p35Gab1. A following cleaning step by phospho-tyrosine could be tested to remove unbound Grb2.

In conclusion, the optimized purification protocol of p35Gab1 – Grb2 (see 2.2.1.6.2) yielded sufficient amounts of pure complex for further analysis, as described in the following Chapter 5.

Chapter 5 Results and Discussion: Analysis of the p35Gab1 – Grb2 complex

5.1 Overview

With the p35Gab1 – Grb2 complex purification protocol, protein was purified to >95% purity (approximation from the SDS gel band). The complex was assessed by various biophysical methods such as analytical ultracentrifugation (AUC), atomic force microscopy (AFM) and electron microscopy (EM). By using cross-linking mass spectrometry (XL-MS), we identified a new Grb2 SH3N – Gab1 binding site. This new interaction site in Gab1 is located in close proximity to the already known R-x-x-K Grb2 SH3C binding motif (Harkiolaki et al., 2009). Crystal screens for Grb2 with a Gab1 peptide containing the binding sites for Grb2 SH3C and SH3N yielded a Grb2 – Gab1 peptide crystal structure and revealed new structural insights of the Grb2 – Gab1 interaction.

5.2 Analytical ultracentrifugation, atomic force microscopy and electron microscopy of the p35Gab1 – Grb2 complex

In collaboration with Dr. Ingrid Tessmer (Rudolf Virchow Center, University of Würzburg) we performed analytical ultracentrifugation (AUC) of the p35Gab1 – Grb2 sample. Sedimentation velocity AUC determines the shape and macromolecular size of a protein in solution (Lebowitz et al., 2002). Samples of Grb2 (WT) and His-p35Gab1 (WT) – Grb2 (WT) complex were subjected to centrifugation in an analytical ultracentrifuge (Beckman Coulter). Sedimentation coefficient distributions for both samples are displayed in Figure 5.1. The peak at small sedimentation coefficients probably resulted from a buffer mismatch between sample and reference buffer solution. The Grb2 sample showed a single peak corresponding to approximately 27 kDa which is in line with the predicted molecular weight for recombinant Grb2 of 27 kDa. The distinct peak in the His-p35Gab1 – Grb2 sample which corresponds to a molecular weight of approximately 61 kDa most likely represents a 1:1 heterodimer complex comprised of the 35 kDa p35Gab1 and the 27 kDa Grb2 proteins. A strong increase of $c(s)$ at high S values indicated the presence of aggregates in the His-p35Gab1 – Grb2 sample (Figure 5.1; right graph was cut off at a sedimentation coefficient of 10). The calculated friction coefficient (f/f_0) of the His-p35Gab1 – Grb2 complex was 1.8, which indicates an elongated shape of the protein complex. In contrast, Grb2 has a f/f_0 value of 1.37, which is very close to other friction coefficients of globular proteins (f/f_0 : 1.2 – 1.3) (Smith, 1988, Dam and Schuck, 2004). Altogether, the AUC results demonstrated a homogenous soluble complex suitable for further structural analysis.

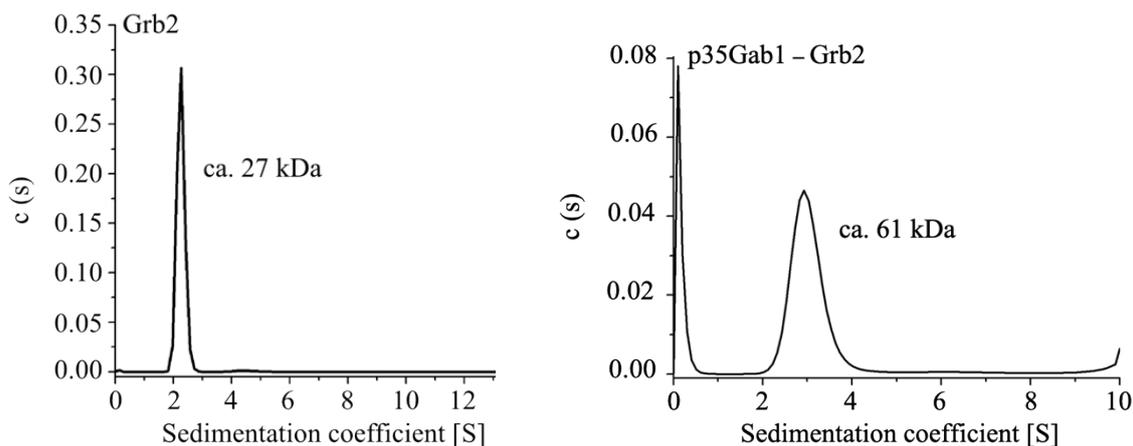


Figure 5.1 Analytical ultracentrifugation of Grb2 and His-p35Gab1 – Grb2 complex

The Grb2 (WT) sample showed a single peak corresponding to a MW of ca. 27 kDa, RMSD: < 0.0049, friction coefficient $f/f_0=1.37$. The peak for the His-p35Gab1(wt) – Grb2 (untagged, wt) sample suggested a 1:1 heterodimer complex of ca. 61 kDa. RMSD: < 0.0046, friction coefficient $f/f_0=1.8$. Proteins were expressed in BL21. Run speed was 40000 rpm at 20 °C (Beckman Optima XL-I analytical ultracentrifuge, An-50 Ti rotor, Beckman Coulter). Data was collected in continuous mode at 280 nm wavelength and analysed by SEDFIT.

The p35Gab1 – Grb2 complex was subsequently analysed by atomic force microscopy (AFM) and electron microscopy (EM) to obtain first structural complex features. AFM analyses the surface characteristics of a protein by scanning the sample with a tip attached to a micro-cantilever and was also done in collaboration with Dr. Ingrid Tessmer. Two p35Gab1 – Grb2 samples, His-p35Gab1 (WT) – His-Grb2 (WT) and His-p35Gab1 (WT) – Grb2 (WT), were analysed by AFM (Figure 5.2).

The His-p35Gab1 – His-Grb2 sample with the His-tag on both proteins of the complex, demonstrated a high heterogeneity in the particle size (Figure 5.2, image on the left). This may be due to an excess of Grb2 in the sample (see Figure 4.2, B). Also, unbound His-tagged Grb2 is co-purified with the p35Gab1 – Grb2 complex and can dimerise in a concentration-dependent manner (Ahmed et al., 2015). This results in a mixture of different species such as the p35Gab1 – Grb2 complex, monomeric Grb2 and possibly dimeric Grb2 (grey arrows point to particles with different size).

AFM analysis of the His-p35Gab1 – Grb2 sample demonstrated a largely homogenous sample (Figure 5.2, image in the middle). This agrees with the SEC profile of the His-p35Gab1 – Grb2 sample which exhibits a single peak (see Figure 4.5). Complex particles were mostly of round shape and stable for at least one hour at RT (Figure 5.2, image on the right).

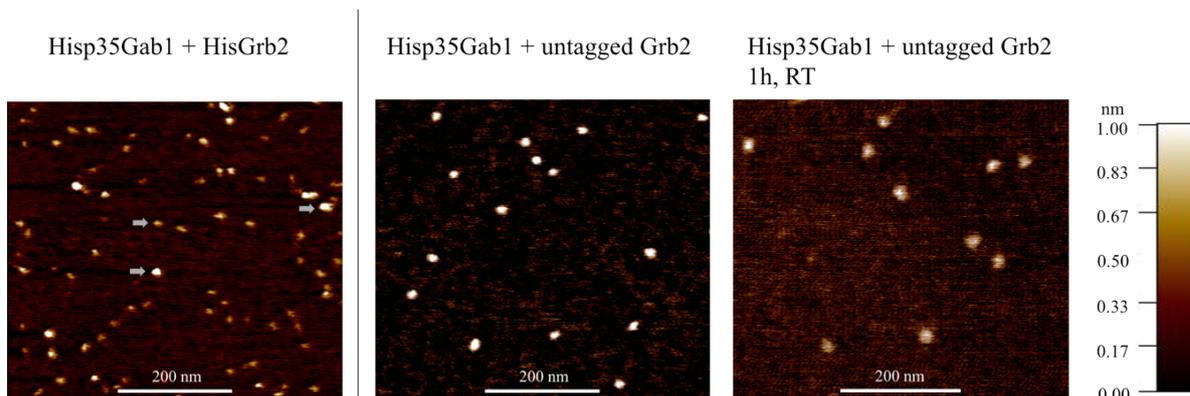


Figure 5.2 AFM analysis of the His-p35Gab1 – His-Grb2 and the His-p35Gab1 – Grb2 complex

His-p35Gab1 (WT) – His-Grb2 (WT) and His-p35Gab1 (WT) – Grb2 (WT), were expressed in *E. coli* (BL21) and purified. Samples were immobilized on a mica surface, dried and analysed by AFM (Molecular Force Probe - 3D AFM, Asylum Research). The AFM scanning was done in tapping mode for high resolution with minimum sample damage and images were taken immediately (left and middle picture) or after waiting for 1 hour at RT (right picture). Vertical scale (colour) is indicated on the right. Grey arrows indicate particles of different sizes in the His-p35Gab1 (WT)/ His-Grb2 (WT) sample.

The His-p35Gab1 – His-Grb2 complex was also very preliminary analysed by negative stain transmission electron microscopy (TEM), in cooperation with Dr. Thorsten Mielke (MPI for Molecular Genetics, Berlin). Complex samples (His-p35Gab1 (WT) – His-Grb2 (WT)) were stained with tungsten phosphate. The more commonly used uranyl acetate staining was incompatible with the phosphate buffer utilised for complex purification.

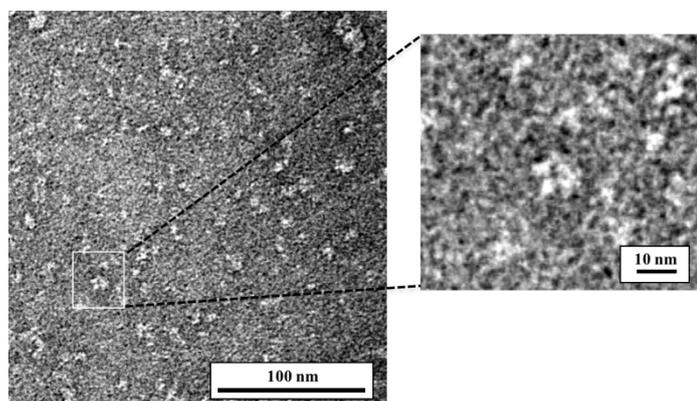


Figure 5.3 Negative TEM stain of the p35Gab1 – Grb2 complex

A His-p35Gab1 (WT) – His-Grb2 (WT) sample (approx. 35 µg/ml), obtained from expression in *E. coli* (BL21), was stained with 4% tungsten phosphate and analysed with a Philips CM100 TEM (100 kV). Scale bar is indicated.

The picture of the TEM scan in Figure 5.3 clearly shows discrete, defined and asymmetrical particles. Zooming in onto one of the structures, it appears to be composed of three distinct domains (Grb2 consists of two SH3 and one SH2 domain) and some extra protein material which would be consistent with Grb2 bound to p35Gab1 (particle indicated by an arrow). The p35Gab1 – Grb2 complex with a

molecular size of only about 61 kDa is at the lower limit of this imaging technique. EM is usually employed to the analysis of much larger protein complexes, for example the characterization of the 26S proteasome, a multi-protein complex with a molecular weight of multiple megadaltons (Walz et al., 1998). Improvements to the resolution could be made by using a more powerful 300 kV TEM along with a newly developed CCD camera system (Sirohi et al., 2016). However, this initial EM experiment already demonstrated that even a relatively small-sized complex can possibly be visualized and yield meaningful results after method optimisation.

To identify p35Gab1 – Grb2 interaction sites, the complex was analysed by chemical cross-linking coupled with mass spectrometry (XL-MS). As described in chapter 1, two Grb2 – Gab2 interaction sites between the C-terminal SH3 domain of Grb2 and two R-x-x-K motifs in the Gab2 protein have been characterized already (Harkiolaki et al., 2009). The two R-x-x-K consensus motifs are highly conserved throughout the Gab family of proteins but the p35Gab1 fragment of Gab1 contains only one of these interaction sites (see Figure 4.1). XL-MS analysis was done in collaboration with Prof. Andrea Sinz (MLU, Halle Saale) and results are discussed in the following subchapter 5.3.

5.3 Identifying interaction sites in the p35Gab1 – Grb2 complex by cross-linking mass spectrometry (XL-MS)

XL-MS examines native or native-like state protein-protein interactions in solution. Closely located protein regions in the protein-protein interface are cross-linked with reagents (crosslinkers) and cross-linked peptides are subsequently analysed by mass spectrometry (Rappsilber, 2011, Leitner et al., 2016). A His-p35Gab1 (WT) – Grb2 (WT) complex sample was mixed with a BS2G (bis[sulfosuccinimidyl]glutarate) crosslinker. BS2G contains two functional groups which form amide bonds and therefore cross-link two closely spaced primary amines. BS2G reacts with lysine (K) residues, with the amino group at the protein N-terminus and also, but to a lesser extent, with the hydroxyl groups of tyrosine, threonine, and serine residues. After cross-linking, the complex sample was in-gel digested with trypsin and endoproteinase Asp-N and analysed by tandem MS. Table 5.1 depicts all identified intermolecular p35Gab1 – Grb2 complex cross-links. In the $_0\text{MEAI AKY}_7$ Grb2 peptide it was not possible to define the cross-linked amino acid, as it could be either crosslinked via the N-terminus or via the lysine 6 residue (K6_{Grb2}). Intramolecular p35Gab1 and Grb2 cross-links are listed in appendix A.2.

p35Gab1 – Grb2 interaction sites identified by XL-MS are visualized on a Grb2 ribbon model for a better understanding of their locations (Figure 5.4, cross-linked Grb2 lysine residues (K) are highlighted in light green; A refers to the chain ID of one Grb2 protomer; PDB code: 1gri). The Grb2 model shows the protein surface (light grey) combined with secondary structures elements of Grb2 (light brown).

p35Gab1			Grb2		
Sequence	Region	Amino acid	Sequence	Domain	Amino acid
499[SSPKTPPR] ₅₀₆	pro-rich	K502	8[DFKATA] ₁₃	SH3N	K10
499[SSPK] ₅₀₂	pro-rich	K502	8[DFKATA] ₁₃	SH3N	K10
499[SSPKTPPR] ₅₀₆	pro-rich	K502	0[MEAI AKY] ₇	SH3N	N-term
499[SSPKTPPR] ₅₀₆	pro-rich	K502	0[MEAI AKY] ₇	SH3N	N-term
499[SSPKTPPR] ₅₀₆	pro-rich	S499	23[DILKVLNEEB] ₃₂	SH3N	K26
522[NLKPDR] ₅₂₇	R-x-x-K Grb2 SH3C	K524	101[FGNDVQHFVKVLR] ₁₁₂	SH2	K109
392[DLNKL RK] ₃₉₈	Not defined	K395	68[AKAEEMLSK] ₇₆	SH2	K69

Table 5.1 Identified intermolecular cross-linked peptides in the p35Gab1 – Grb2 complex

The cross-linked His-p35Gab1 (WT) – Grb2 (WT) complex was analysed by nano-HPLC/nano-ESI-MS/MS (UltiMate RSLC nano-HPLC system coupled to an Orbitrap Fusion Tribrid mass spectrometer, Thermo Fisher Scientific). His-p35Gab1 (WT) – Grb2 (WT), expressed in BL21, was mixed with a BS2G isotope d₀/d₄ mix (100-fold molar excess) for subsequent identification of cross-linked peptides in MS analysis. Data were evaluated with StavroX software.

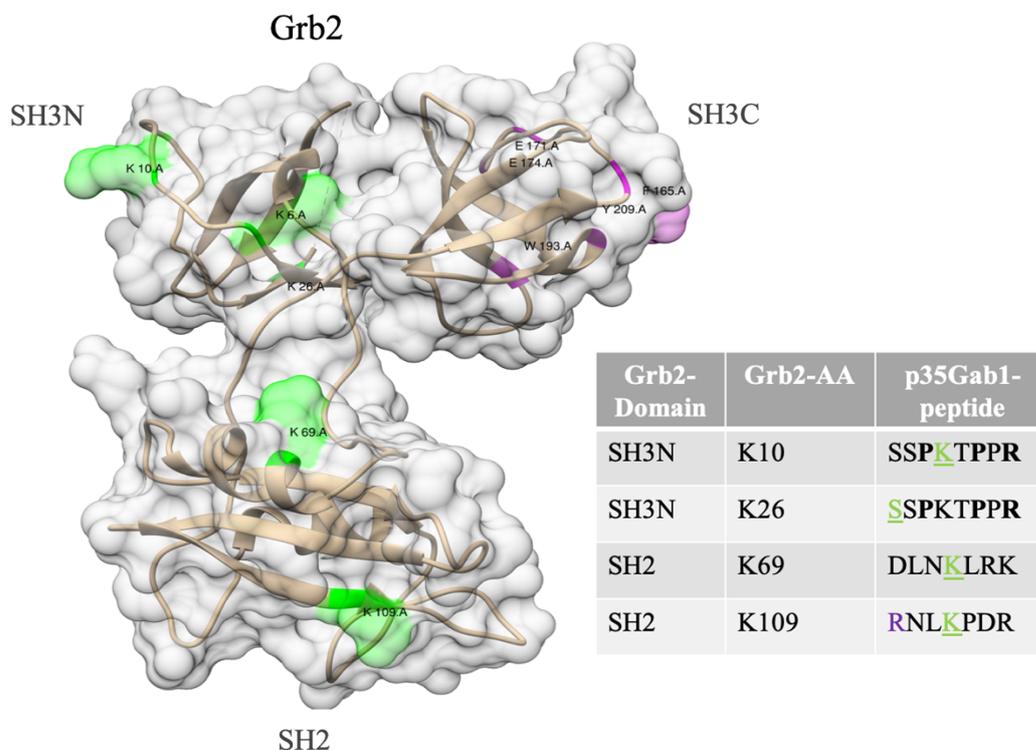


Figure 5.4 XL-MS-identified interaction sites in the p35Gab1 – Grb2 complex are displayed on a Grb2 model

The Grb2 protein (PDB code: 1gri) is shown as a ribbon model combined with a surface depiction. Grb2 lysine residues (K) which have been identified as cross-linked with p35Gab1 are highlighted in light green. The table (right side) summarizes identified p35Gab1 peptides and the corresponding cross-linked Grb2 domain. The cross-linked p35Gab1 lysine residue is underlined and coloured in green. Amino acids highlighted in purple are interacting residues from a Grb2 SH3C – Gab2 peptide complex crystal structure (PDB code: 2vuf). A potential SH3N binding motif in p35Gab1 is indicated by bold residues. The letter ‘A’ refers to the chain ID of one of the two Grb2 protomers in the crystal asymmetric unit.

XL-MS results of the p35Gab1 – Grb2 complex suggest that a p35Gab1 peptide (₄₉₉SSPKTPPR₅₀₆) interacts with Grb2 residues on the Grb2 SH3N surface (K10_{Grb2}; Grb2 N-terminus). This p35Gab1

peptide ${}_{499}\text{SSPKTPPRR}_{506}$ contains a typical **P-x-x-P-x-R** motif which is known to bind to SH3 domains (Feng et al. 1994; Goudreau et al. 1994; Terasawa et al. 1994; Wu et al. 1995). K10_{Grb2} and the N-terminal Grb2 residue (in Figure 5.4 highlighted as K6_{Grb2}) are in close proximity of the canonical Grb2 SH3N ligand binding site as shown for the Grb2 SH3N – Sos peptide complex in 1994 (Goudreau et al., 1994, Terasawa et al., 1994). Lysine K26_{Grb2} is located on the opposite site of the SH3N domain, which makes it unlikely that this residue is involved in the same Gab1 – Grb2 SH3N binding surface.

Isothermal titration calorimetry (ITC) was conducted to measure binding affinities of Grb2 with Gab1 peptides containing the potential Grb2 binding sites identified by cross-linking MS. ITC measurements are summarized in Table 5.2. Non-specific interactions between peptide and the GST affinity tag of the purified protein domain could be excluded (Table 5.2, no. 1 – 3). The binding integrity of purified proteins or protein domains was confirmed with domain-specific control peptide ligands (Table 5.2, no. 4 – 7, peptides 1 – 4). ITC for the Gab1 peptide containing the potential Grb2 SH3N binding motif (FRSSPKTPPRRPVPV) with Grb2 confirmed binding to the full-length Grb2 protein ($K_d 36.9 \pm 1.6 \mu\text{M}$) and the Grb2 SH3N domain ($K_d 50.5 \pm 3.8 \mu\text{M}$) in solution (Table 5.2, no. 13 – 14). SH3 domains typically bind proline-rich ligands with modest affinities in the micromolar range (1 to 100 μM) (Li, 2005). Replacement of the first proline and the arginine of the **P-x-x-P-x-R** motif with alanine (FRSSAKTPPARPVPV) abolished binding and therefore confirmed the critical position of these residues (Table 5.2, no. 15).

For two cross-links, that were found between p35Gab1 and the Grb2 SH2 domain ($\text{K69}_{\text{Grb2}} - \text{K395}_{\text{Gab1}}$ and $\text{K109}_{\text{Grb2}} - \text{K524}_{\text{Gab1}}$), a binding could not be confirmed by ITC (Table 5.2, no. 8; p35Gab1 peptide SEPPPVDRLNK $_{524}$ PDRK, C514S and Grb2 SH2). SH2 domains usually bind phosphorylated tyrosine ligands (pY) (McNemar et al., 1997, Ogura et al., 1999). Grb2 $\text{K109}_{\text{Grb2}}$ is closely located to the Grb2 SH2 pY ligand binding site whereas Grb2 K69_{Grb2} is located 12 Å away from the Grb2 SH2 binding site (Ogura et al., 1999, Nioche et al., 2002).

In fact, lysine K524 in Gab1 is part of the R-x-x-K Grb2 SH3C binding motif which was first described in the Gab2 protein but which is also found in the Gab1 protein as it is highly conserved throughout the Gab family proteins (Harkiolaki et al., 2009). Indeed, a binding of the SH2 cross-linked p35Gab1 peptide (15 aa, SEPPPVDRLNK $_{524}$ PDRK, C514S, R-x-x-K motif) to the Grb2 SH3C domain ($K_d 5.5 \mu\text{M}$) and to full length Grb2 ($K_d 3.9 \mu\text{M}$) was confirmed (Table 5.2, no. 9 – 10). This affinity is similar to the one measured for a Gab2 peptide (Grb2 SH3C – Gab2b $K_d 3.7 \mu\text{M}$; (Harkiolaki et al., 2009). Alanine mutation of the arginine and lysine residues in the Gab1 **R-x-x-K** motif (SEPPPVDANLAPDRK) abolished binding to the Grb2 SH3C domain and the full length Grb2 protein (Table 5.2, no. 11 – 12). A binding of the second identified SH2-cross-linked Gab1 peptide ($\text{K395}_{\text{Gab1}}$; ${}_{392}\text{DLNKLRRK}_{398}$) to the Grb2 SH2 domain has not been confirmed, yet. However, binding is unlikely as the peptide or close amino acids contain no phosphorylated tyrosine residues similar to

Results and Discussion: Analysis of the p35Gab1 – Grb2 complex

the other SH2-cross-linked Gab1 peptide (K524_{Gab1}; 522¹NLKPDR₅₂₇). Representative ITC results are shown in appendix A.4.

No.	Peptide	Peptide sequence	Protein/ Domain	K _d [μM]	N
1	15 aa Gab1, XL-MS Grb2 SH3N binder (WT)	Gab1 497-FRSSPKTPRRRPVPV-511 (5)	GST	-	-
2	15 aa Gab1, XL-MS Grb2 SH2 binder (C514S)	Gab1 514-SEPPPVDRLNLPDRK-528 (10)	GST	-	-
3	Grb2 SH2 positive control peptide	A(pY)VNVA (1)	GST	-	-
4	Grb2 SH2 positive control peptide	A(pY)VNVA (1)	GST-Grb2 SH2	0.9	0.7
5	Grb2 SH2 negative control peptide	AYVNVA (2)	GST-Grb2 SH2	-	-
6	Grb2 SH3N positive control peptide	PPPLPPRRRR (3)	GST-Grb2 SH3N	5.3	0.9
7	Grb2 SH3N negative control peptide	PPGALGPLLR (4)	GST-Grb2 SH3N	-	-
8	15 aa Gab1, XL-MS Grb2 SH2 binder (C514S)	Gab1 514- SEPPPVDRLNLPDRK -528 (10)	GST-Grb2 SH2	-	-
9	15 aa Gab1, XL-MS Grb2 SH2 binder (C514S)	Gab1 514- SEPPPVDRLNLPDRK -528 (10)	GST-Grb2 SH3C	5.5	1.0
10	15 aa Gab1, XL-MS Grb2 SH2 binder (C514S)	Gab1 514- SEPPPVDRLNLPDRK -528 (10)	His-Grb2 full C32S C198A	3.9	1.2
11	AA mutant 15 aa Gab1, XL-MS Grb2 SH2 binder (C514S, R521A, K524A)	Gab1 514- SEPPPVDANLAPDRK -528 (11)	GST-Grb2 full	-	-
12	AA mutant 15 aa Gab1, XL-MS Grb2 SH2 binder (C514S, R521A, K524A)	Gab1 514- SEPPPVDANLAPDRK -528 (11)	GST-Grb2 SH3C	-	-
13	15 aa Gab1, XL-MS Grb2 SH3N binder (WT)	Gab1 497- FRSSPKTPRRRPVPV -511 (5)	GST-Grb2 SH3N	50.5 ± 3.8	1.0 *
14	15 aa Gab1, XL-MS Grb2 SH3N binder (WT)	Gab1 497- FRSSPKTPRRRPVPV -511 (5)	GST-Grb2 full	36.9 ± 1.6	1.0 *
15	AA mutant 15 aa Gab1, XL-MS Grb2 SH3N binder (P501A, R506A)	Gab1 497- FRSSAKTPPARVPV -511 (6)	His-Grb2 full C32S C198A	-	-
16	32 aa Gab1, SH3 tandem peptide (C514S)	Gab1 497- FRSSPKTPRRRPVPVADSEPPPVDRLNLPDRK -528 (7)	Grb2 full C32S C198A	1.1 ± 0.3	1.1 ± 0.2
17	32 aa Gab1, SH3 tandem peptide (C514S)	Gab1 497- FRSSPKTPRRRPVPVADSEPPPVDRLNLPDRK -528 (7)	His-Grb2 full C32S C198A	1.0	0.8

No.	Peptide	Peptide sequence	Protein/ Domain	K _d [μM]	N
18	32 aa Gab1, SH3 tandem peptide (C514S)	Gab1 497- FRSSPKTPRRPVPVADSEPPPVDRLNKPDRK -528 (7)	GST-Grb2 SH3N	-	-
19	AA mutant 32 aa Gab1, SH3 tandem peptide (P501A, R506A, C514S)	Gab1 497- FRSS AK TPP AR PVPVADSEPPPVDRLNKPDRK -528 (8)	His-Grb2 full C32S C198A	1.0	0.8
20	AA mutant 32 aa Gab1, SH3 tandem peptide (P501A, R506A, C514S)	Gab1 497- FRSS AK TPP AR PVPVADSEPPPVDRLNKPDRK -528 (8)	Grb2 full C32S C198A	1.1	0.9
21	AA mutant 32 aa Gab1, SH3 tandem peptide (P501A, R506A, C514S)	Gab1 497- FRSS AK TPP AR PVPVADSEPPPVDRLNKPDRK -528 (8)	Grb2 full C32S C198A	1.0	0.8
22	45 aa Gab1, SH3 tandem peptide (C514S)	Gab1 497- FRSSPKTPRRPVPVADSEPPPVDRLNKPDRKVKPAPLEIKPLPE -541 (9)	Grb2 full C32S C198A	0.6	1.3
23	45 aa Gab1, SH3 tandem peptide (C514S)	Gab1 497- FRSSPKTPRRPVPVADSEPPPVDRLNKPDRKVKPAPLEIKPLPE -541 (9)	Grb2 full C32S C198A	0.9	0.9
24	45 aa Gab1, SH3 tandem peptide (WT) **	Gab1 497- FRSSPKTPRRPVPVADCEPPPVDRLNKPDRKVKPAPLEIKPLPE -541 (9)	Grb2 SH3C	1.2	1.0
25	32 aa Garem1, SH3 tandem peptide (WT)	Garem1 529- LLNAPPVPPRS AK PLSTSPSIPPRTVKPARQQ -560 (12)	His-Grb2 full C32S C198A	0.8	0.9
26	32 aa Garem1, SH3 tandem peptide (WT)	Garem1 529- LLNAPPVPPRS AK PLSTSPSIPPRTVKPARQQ -560 (12)	Grb2 full C32S C198A	0.7	0.9

Table 5.2 ITC measurements of Grb2 interactions with Gab1- and Garem1-derived peptides

Peptides are acetylated on the N-terminus and amidated on the C-terminus except the sumo-tag purified peptide. Standard deviations are presented for three or more independent measurements. Mutations are either shown in bold red (motif mutations) or grey letters. N values marked with an asterisk (*) indicate that N is set manually to 1 due to very low protein/protein domain concentrations. Sumo purification tag purified peptides are marked with an double asterisks (**). A list of all peptides can be found in 2.1.7. XL-MS: cross-linking mass spectrometry. ITC analysis was done by Dr. Marc Lewitzky and Dr. Tobias Gruber (AG Prof. Feller, MLU) and representative ITC results are shown in appendix A.4.

The XL-MS approach did not detect the well-established interaction of the p35Gab1 R-x-x-K motif with the Grb2 SH3C domain. According to the Grb2 SH3C – Gab2b crystal structure (PDB code: 2vwf), the Gab2b peptide R-x-x-**K** lysine residue is not in close proximity (< 7.7 Å, BS2G crosslinker spacer arm length) to any lysine residue Grb2 of the Grb2 SH3-domain. The Grb2 – Gab1 peptide crystal structure (see subchapter 5.6) shows that the same is true for the Gab1 R-x-x-K motif. This could be one reason why BS2G is unable to cross-link the R-x-x-K motif in Gab1 to the Grb2 SH3C domain.

Grb2 or Grb2-like proteins are conserved as well (compare Grb2-SH3N alignment in subchapter Figure 5.17 and Grb2-SH3C alignment in Harkiolaki et al., 2009).

A	Grb2 SH3N P-x-x-P-x-R	Grb2 SH3C R-x-x-K-P
Gab1 human	MQVPPPAHMGFRSS PKT PPRR PVPVADCEPPPVD	RNLK PKDRK
Gab2 human	-----SLGYPSTTLPVHRGPSRGSEIQPPPV	RNLK PKDRK
Gab3 human	-----LEPPP	RDLK PQRK
Gab4 human	-----LLGSPLTELSMHQDLSQGHEVQLPPV	RSLK PNQK
B		
H. sapiens	MQVPPPAHMGFRSS PKT PPRR PVPVADCEPPPVD	RNLK PKDRK
M. musculus	MQVPPPAHMGFRSS PKT PPRR PVPVADCEPPPVD	RNLK PKDRK
G. gallus	KQVPPPSHMGFRSS PKT PPRR SMPAEKCEPPPVD	RNLK PKDRK
L. chalumnae	MQVPPPAHMGFRSS PKT PPRR LVPAADCQPPPVD	RNLK PKDRK
T. rubripes	KQVPPPAHMGFRSS PKT PPRR PI-LSDCQPPPVD	RNLK PKDRK
D. rerio	KQVPPPAHMGFRSS PKT PPRR PIPISECQPPPID	RNLK PKDRK
C. milii	MQVPPPAHMGFRV S PKT PPRR PVLAVDCQPPPVD	RNLK PKDRK
B. floridae	IPSPPLVEVTPGPAPAPSATPEKKPADLPPPI	RDLK PGSR
N. vectensis	-PGFTPTPKADTAPAVDRSVKPPPTVDRLTKPPV	RTLK PSMD

Figure 5.6 Grb2 SH3 domain-binding tandem motif in Gab proteins and Gab-like proteins

(A) The newly found Grb2 tandem motif in Gab1 is not conserved in the Gab family of proteins. Sequence alignment of the putative tandem motif in the Gab family of proteins. References for human Gab1 family proteins sequences were retrieved from UniProt (Accession codes: human Gab1 protein: Q13480; human Gab2 protein: Q9UQC2; human Gab3 protein: Q8WWW8; human Gab4 protein: Q2WGN8). (B) The Grb2 SH3 domain-binding tandem motif in Gab1 or Gab1-like proteins is highly conserved throughout metazoan evolution. Sequence alignment of Gab1 or Gab1-like proteins harbouring the interaction motif P-x-x-P-x-R-x(n)-R-x-x-K-P. Accession codes for Gab1 or Gab1-like proteins sequences were retrieved from UniProt (*Homo sapiens* (Human) Gab1 protein: Q13480; *Mus musculus* (Mouse) Gab1 protein: Q9QYY0; *Gallus gallus* (Chicken) Gab1 protein: F1NHL8; *Latimeria chalumnae* (West Indian Ocean Coelacanth) Gab1 protein: H3AJG3; *Takifugu rubripes* (Japanese pufferfish) Gab1 protein: H2UFJ2; *Danio rerio* (Zebrafish) Gab1 protein: F1Q7S5; *Callorhynchus milii* (Elephant fish) Gab1 protein: V9KEN5; *Branchiostoma floridae* (Florida lancelet) Gab1 protein (uncharacterised): C3XWJ7; *Nematostella vectensis* (starlet sea anemone) predicted Gab1 protein: A7RVJ5). Residues in the binding motif (framed by boxes) which have been found to be critical for Grb2 binding are indicated in bold (Goudreau et al., 1994, Terasawa et al., 1994, Harkiolaki et al., 2009). Initial sequence alignments were made with Clustal Omega. Sequences near the tandem peptide were manually aligned.

Tandem motifs have been implicated in other protein- protein interactions. A tandem SH3 domain motif was described in a host-pathogen interaction (Aitio et al., 2010). The enterohemorrhagic *E. coli* protein EspF_U requires a tandem P-x-x-P SH3 domain motif to interact with the SH3 domain of the human insulin receptor tyrosine kinase substrate (IRTKS) in order to hijack the mammalian actin assembly machinery. Additional to the *E. coli* protein other cellular proteins also contain a IRTKS SH3 domain tandem motif which seems to be critical for its interaction. Furthermore, there are short

linear tandem motifs such as ITAMS (immunoreceptor tyrosine-based activation motifs) which are important in SH2 domain mediated interactions (Underhill and Goodridge, 2007).

In conclusion, the conservation of the tandem Grb2 interaction module in Gab1 for over 500 million years is suggestive of its important role throughout evolution. A alignment search of the tandem motif was done to examine if there are other tandem motifs in the human proteome that are similar to the one mediating the Gab1 – Grb2 interaction. The search results are discussed in the following subchapter.

5.5 Garem1 protein contains a similar tandem motif and interacts with Grb2

A search for motifs similar to the Grb2 tandem motif in the human proteome was performed with ScanProsite (de Castro et al., 2006). This was done by using the input sequence P-x-x-P-x-R(x)₁₁₋₁₇-R-x-x-K-P specified with a spacer length between 11 and 17 amino acids whereby x representing any amino acid (Dr. Marc Lewitzky, AG Prof. Feller, MLU). The search identified a number of candidate proteins and a shortened list is shown in Figure 5.7. Some of the identified proteins, are known interacting partners of Grb2 (proteins coloured in blue), but so far, only Garem1 has been found to interact with Grb2 *in vivo* (Tashiro et al., 2009).

GABI**	GFRSS PKT PPR ---RPVPVADCEPPPVD RNLK PDRKVK
GAREM1**	RLLN A PPV PPR ----SAKPLSTSPSIP PR TV K PARQQT
KMT2B*	SEPGG PPA PRR RTPRENELPLPEPEEQ S R PRK PTLQPV
SYNJ2*	TPQ A PL L PRR ---PPPRVPAIKKPTLR R T GK PLSPEE
SPKAP	EFLMT P NV PCR --SLKRKKESQGS TAVR KH K PRLSE
VIP1	QASDN P F S PPR -----TLHSPPLQLQ QR SE K PPWYSS
NOL6	YTEVF P PT P VR PA FSFYETLRERSLL P RLD K PCPAYV
CD11B	HWSRS P PR P PRERFELGDGRKPGEAR PARA Q K PAQLKE

Figure 5.7 Proteins with a Grb2 tandem binding motif

Shortened list of results from the alignment search with ScanProsite (Input: P-x-x-P-x-R(x)₁₁₋₁₇-R-x-x-K-P, Homo sapiens). Proteins that interact with Grb2 *in vitro* (*) or *in vivo* (**) are coloured in blue. Residues from the motif are marked in bold.

Garem1 (Grb2-associated and regulator of Erk/MAPK) is an adaptor protein that is involved in the epidermal growth factor (EGF) receptor-signalling pathway. It contains an N-terminal CABIT (cysteine-containing, all in Themis) domain, a proline rich region (aa 511 – aa 550) and a C-terminal SAM-PNT domain. Based on co-immunoprecipitation studies, the Grb2 interaction site could be located to a pro-rich region in Garem1 (aa 498 – aa 550) (Tashiro et al., 2009). We confirmed by ITC a 1:1 stoichiometric binding of Grb2 with a Garem1 peptide (32 aa, ₅₂₉LLN**A**PPV**PPR****S**AKPLSTSPSIP**PR**TV**K**PARQ**Q**₅₆₀) containing the double motif (K_d 0.7 – 0.8 μM). To analyse the interaction between Grb2 and the newly found Grb2 tandem motif in Garem1 on a

structural level, crystal screens with Grb2 (His-Grb2/ Grb2 C32S, C198A) in complex with the 32 aa Garem1 peptide were set up. However, despite extensive screening efforts, we only obtained very small Grb2 – Garem1 peptide complex crystals, which weakly diffracted, if at all. Crystal screen optimisation is definitely needed.

The putative Grb2 tandem motif (aa 529 – aa 560) in the Garem1 protein overlaps with the Grb2 interacting region identified by Tashiro et al. and is located in an intrinsically disordered region of Garem1 (PONDR-VLXT) (Figure 5.8).

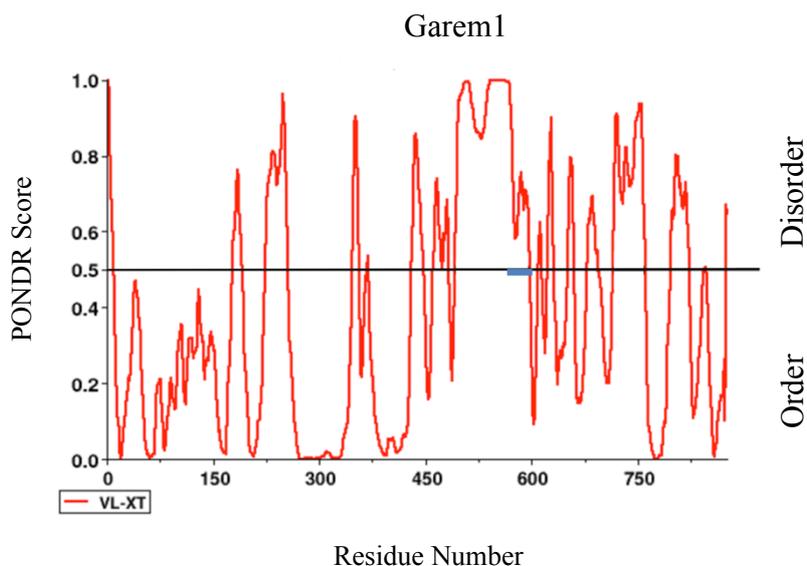


Figure 5.8 The putative Grb2 tandem motif in Garem1 is found in a disordered region

Human Garem1 (Uniprot code: Q9H706) was analysed with the disorder predictor PONDR-VLXT. Region with the Grb2 tandem motif in Garem1 is marked with a blue rectangle. Values above 0.5 PONDR score display disordered regions and below 0.5 PONDR score ordered regions.

A Garem1 protein sequence alignment revealed a similar conservation of the tandem motif in Garem1 as in Gab1 (Figure 5.9). Comparing the two alignments with each other shows that the Grb2 SH3 domain-binding tandem motif in Gab1 has been conserved for a longer time period than in the Garem1 protein. When comparing the appearance of the single domain motifs, in Garem1 the Grb2 SH3N binding site evolved earlier than the Grb2 SH3C binding site. In contrast to the Gab1 protein where the Grb2 SH3C binding site appeared before the Grb2 SH3N binding motif.

The Grb2 SH3 domain-binding tandem motif appears to be the result of convergent evolution as it appeared in two unrelated proteins, Gab1 and Garem1. Convergent evolution with respect to protein evolution is a mechanism by which similar protein features or motifs independently evolve in unrelated species or different protein families. It has been shown that, in particular, SLiMs (short, linear motifs) have a higher tendency to evolve in a convergent manner (Davey et al., 2012). A potential convergent evolution of the Grb2 SH3 domain-binding tandem motif in Garem1 and Gab1 would strongly imply a functional role tandem motif in cellular processes.

	Grb2 SH3N P-x-x-P-x-R		Grb2 SH3C R-x-x-K		Grb2 SH3N P-x-x-P-x-R		Grb2 SH3C R-x-x-K	
	Garem1				Gab1			
H. sapiens	KSEAVREECRLIN APPVPPRS AKPLSTSPSIE PR TVK PAR QQ				MQVPPPAHMGFRSS PKT PPRR PFVADCEPPP VD RNL K PD RK			
M. musculus	KSEAVREECRLIN APPVPPRS AKPLSTSPSIE PR TVK PV RPQ				MQVPPPAHMGFRSS PKT PPRR PFVADCEPPP VD RNL K PD RK			
G. gallus	KSEAVKEECRLIN APPVPPRS SKPSSTSPSIE PR TVK PAR QQ				KQVPPPSHMGFRSS PKT PPRR SMPAEKCEPPP VD RNL K PD RK			
L. chalumnae	KSEAVKEECRLIN APPVPPRS SKPSSTSPSIE PR AT KQ TRQQ				MQVPPPAHMGFRSS PKT PPRR LVAADCQPPP VD RNL K PD RK			
T. rubripes	KSEAVREECRLIN APP IP PPRS SKQMPILVPILSKSLQQDTRCP				KQVPPPAHMGFRSS PKT PPRR PI-LSDCQPPP VD RNL K PD RK			
D. rerio	KSEAVKEECRLIN APP IP PPRS SKQAGSSSATVPYPSAKPRQK				KQVPPPAHMGFRSS PKT PPRR PIPISECQPPP VD RNL K PD RK			
C. milii	KSEAVKEECRLIN TP PPV PPR NTKPSLSPSIPPRQIRQQARS				MQVPPPAHMGFRV S PKT PPRR PFVLAVDCQPPP VD RNL K PD RK			
B. floridae	LGELKKAQESKKT PP VP PPR PKAPTVTSDQGTSDDYDTDLTGG				IPSPPLVEVTPGPAPAPSATPEKPPADLPPP IR RD L K PG SR			

Figure 5.9 Strong evolutionary conservation of the putative Grb2 SH3 domain-binding tandem peptide in Garem1 and Gab1 throughout metazoans

Sequence alignment of putative Grb2-binding tandem motif (framed by black boxes) in Garem1 or Garem1-like proteins and Gab1 or Gab-like proteins. Protein sequences were retrieved from UniProt. (Accession codes: *Homo sapiens* (Human) Garem1 protein: Q9H706; *Mus musculus* (Mouse) Garem1 protein: Q3UFT3; *Gallus gallus* (Chicken) Garem1 protein: F1NTK0; *Latimeria chalumnae* (West Indian Ocean Coelacanth) Garem1 protein: H2ZWY3; *Takifugu rubripes* (Japanese pufferfish) Garem1 protein: H2SB82; *Danio rerio* (Zebrafish) Garem1 protein: Q7ZVU1; *Branchiostoma floridae* (Florida Lancelet) Garem1 protein: C3ZD55. *Callorhynchus milii* (Elephant fish) predicted Garem1 protein (XP_007886156.1) is from the NCBI. See Figure 5.6 for Gab1 alignment accession codes. Sequences near the tandem peptide were manually aligned. Residues in the binding motif important for binding in Garem1 and Gab1 are marked in bold.

In order to analyse the binding features of the Grb2 tandem motif in Gab1 in complex with Grb2, we performed crystallization screens, which are described in the following subchapter 5.6.

5.6 Analysis of the Gab1 – Grb2 interaction by X-ray crystallography

5.6.1 Overview

We investigated the Gab1 and Grb2 protein interaction by X-ray crystallography to obtain a better understanding of the functionality of the p35Gab1 – Grb2 protein complex. The structural characterization of the intrinsically disordered protein, full-length Gab1 (aa 1 – aa 694) is only in very early stages. So far, a C-terminal Gab1 phosphopeptide (aa 621 – aa 633) has been crystallized with the N-terminal SH2 domain of SHP2 (2.1 Å, PDB code: 4qsy, 2014, unpublished data). There are also crystal structures available of the Grb2 SH3C and of the 14-3-3 protein in complex with a Gab2 protein-derived peptide (PDB codes: 2vwf; 2w0z; 5ewz; 5exa, respectively) (Harkiolaki et al., 2009, Bier et al., 2016). There are no structures for Gab3 and Gab4 available, so far. The adaptor protein Grb2 (aa 1 – aa 217) in contrast has been extensively studied over the last decades. Many crystal structures are available for the C-terminal SH3 and the SH2 domain (Nishida et al., 2001, Nioche et al., 2002, Harkiolaki et al., 2009, Higo et al., 2013). The isolated N-terminal Grb2 SH3 domain has mainly been studied by NMR (Goudreau et al., 1994, Terasawa et al., 1994, Wittekind et al., 1994,

Guruprasad et al., 1995). The only full length Grb2 (WT, ligand-free) crystal structure was solved to a resolution of 3.1 Å (PDB code: 1gri, Maignan et al., 1995).

Therefore, a crystal structure of Grb2 in complex with the p35Gab1 protein could elucidate new structural characteristics and shed light into the binding features of the newly identified Grb2-binding tandem motif in Gab1. We conducted crystal screens with the p35Gab1 – Grb2 complex and with Grb2 in complex with p35Gab1-peptides. Finally, we obtained a structure from full length Grb2 in complex with a p35Gab1 protein-derived peptide. Grb2 purification, the process of crystal optimization and structure determination are described in the following subchapters.

5.6.2 Grb2 purification

We successfully expressed a mutant His-Grb2 (human Grb2 C32S, C198A; pET28(+); 27 kDa) in *E. coli* BL21 and purified His-Grb2 or Grb2 without a His-tag. The Grb2 C32S and C198A mutations were inserted to favour monomeric Grb2, as wildtype Grb2 can form homodimers (McDonald et al., 2008b, Ahmed et al., 2015). A Grb2 (C32S, C198A) mutant had been used previously for analysing monomeric Grb2 by NMR spectroscopy (Yuzawa et al., 2001).

His-tagged Grb2 (C32S, C198A) was initially immobilized on a HiTrap IMAC column and eluted with an imidazole gradient. For the purification of untagged Grb2, the His-tag was removed by on-column cleavage with thrombin. This has the advantage that only the portion of Grb2 where the His-tag has been cleaved off is carried on to the next purification step. After a Mono Q (HP HiTrap Q) purification step, Grb2 (C32S, C198A) or His-Grb2 (C32S, C198A) was finally run on a SEC S75 column (Figure 5.10, S75 of Grb2 C32S, C198A; without His-tag). Grb2 eluted at 65 ml, which corresponds to a globular protein with a molecular mass of 33 kDa (according to a S75 calibration curve). The predicted molecular weight for recombinant Grb2 is 27 kDa. The minor discrepancy between calculated and SEC-estimated molecular mass of Grb2 can be due to a more extended Grb2 structure in solution. Grb2 dimerisation was observed with higher Grb2 (C32S, C198A) sample concentrations (Figure 5.10, chromatogram on the right). The S75 profile shows two peaks corresponding to dimeric Grb2 and monomeric Grb2. SDS-PAGE-analysis of Grb2 purification samples demonstrated a relatively pure sample of Grb2 (27 kDa band) with minor contamination by a 20 kDa-sized protein (Figure 5.10, final Grb2 sample). Grb2 purifications were conducted at room temperature (RT) as we observed an decrease in the protein yield when the purification was performed at 4 °C rather than RT (data not shown). A detailed Grb2 expression and purification protocol can be found in 2.2.1.6.3.

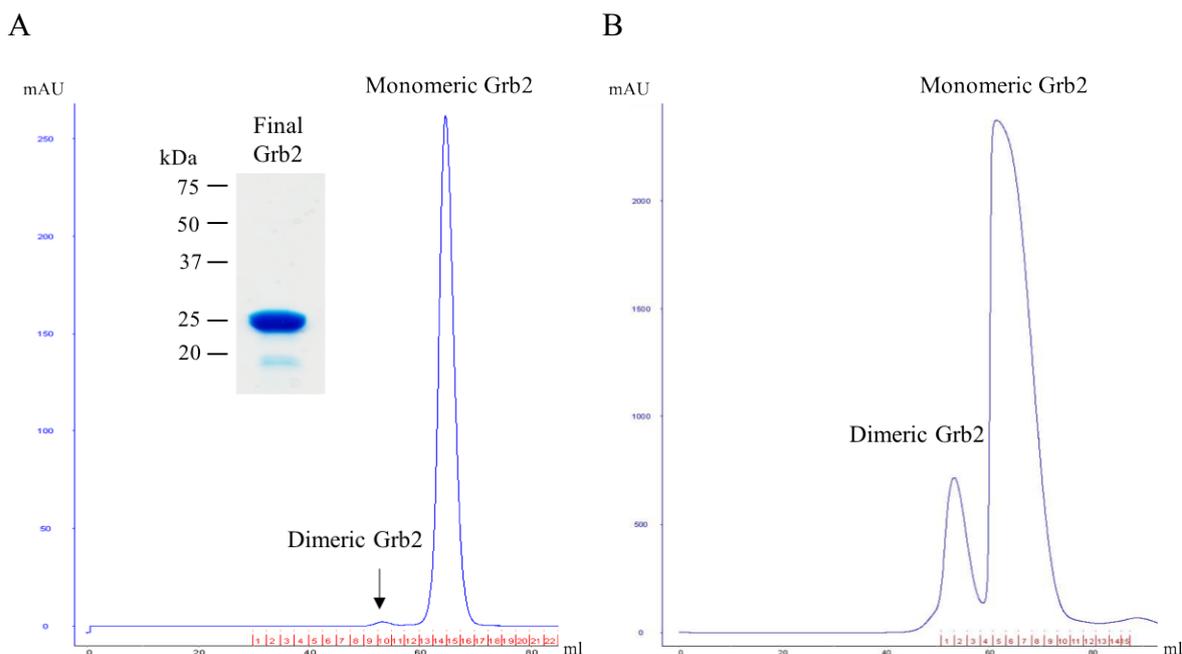


Figure 5.10 SEC of two differently concentrated Grb2 (C32S, C198A) samples on a S75 column

S75 (HiLoad™ 16/600 Superdex™ S75) loaded with a Grb2 (C32, C198A) sample (A) and a 4x higher concentrated and 24x larger amount loaded Grb2 (C32S, C198A) sample (B) (A: 4 mg/ml, total amount on S75: ca 4 mg; B: ca 16 mg/ml; total amount on S75: 88 mg). The blue line indicates the UV absorption in [mAU], shown at the y-axis. The x-axis shows the elution volume of the sample in [ml]. Fractions collected during the run are marked by numbers (red). A final Grb2 sample after SEC was separated on a SDS PAGE (12% separating gel) (A). Protein marker bands are indicated on the left side.

5.6.3 Crystal screens for Grb2 with p35Gab1 and p35Gab1-derived peptides

The Grb2 – p35Gab1 complex (Grb2 C32S, C198A; p35Gab1 C374A, C405A, C514A) was expressed in TKB1, purified (purification protocol described in 2.2.1.6.2) and used to set up sparse matrix crystallization screens. Another set of crystal screens was set up with untagged or His-tagged full-length Grb2 (C32S, C198A) and two, in length different, Gab1 peptides (32 aa and 45 aa) (see Table 2.4 Peptides). Both Gab1 peptides cover the potential Grb2 tandem binding site, but the 45 aa Gab1 peptide is C-terminally extended by 13 aa. The 45 aa Gab1 peptide was used to examine the potential influence of a proline-rich stretch downstream of the tandem motif on Grb2 binding. Cysteine 514 in the p35Gab1 peptide was mutated to serine in order to prevent artificial disulfide bond formations. The C514S mutation is located in the linker region between the two Grb2 binding motifs, thus a potential effect of the mutation on peptide binding is expected to be negligible. ITC measurements were performed to determine the binding affinity between Gab1 peptides and full length Grb2 which is described in the following subchapter.

5.6.3.1 ITC of Grb2 and Gab1 peptides with the SH3 domain tandem motif

ITC measurements confirmed the binding of the 32 aa Gab1 peptide (FRSSPKTPRRRPVPVADSEPPPVDRLNLPDRK, C415S) to full length Grb2 (C32S, C198A) and to His-tagged Grb2 (C32S, C198A) (K_d Grb2 $1.1 \pm 0.3 \mu\text{M}$; K_d His-Grb2 1.0; Table 5.2, no. 16 – 17). The increased affinity measured between Grb2 and the Gab1 peptide with the Grb2 tandem motif (32 aa), compared to Grb2 and Gab1 peptides with single motifs (15 aa; Gab1 SH3N peptide: FRSSPKTPRRRPVPV, K_d full length Grb2 $36.9 \pm 1.6 \mu\text{M}$; Gab1 SH3C peptide: SEPPPVDRLNLPDRK, C415S, K_d full length Grb2 $3.9 \mu\text{M}$), is most likely due to the two SH3 domain-binding motifs in the 32 aa Gab1 peptide (FRSSPKTPRRRPVPVADSEPPPVDRLNLPDRK, C415S). Interestingly, we could not detect any binding between the 32 aa Gab1 tandem peptide and the individual Grb2 SH3N domain by ITC (Table 5.2, no. 18). The low affinity Grb2 SH3N – Gab1 peptide interaction (15 aa Gab1 SH3N peptide; K_d Grb2 SH3N $50.5 \pm 3.8 \mu\text{M}$, Table 5.2, no. 13) could have been destabilised by the residues belonging to the additional Grb2 SH3C binding site, which is not taking part in the interaction. Or the 32 aa Gab1 tandem peptide requires the higher affinity Grb2 SH3C – Gab1 peptide interaction (K_d SH3C $5.5 \mu\text{M}$, Table 5.2, no. 9) to stabilise its interaction with the Grb2 SH3N domain.

The extended 45 aa Gab1 peptide, which contains additional prolines located C-terminal of the double motif (FRSSPKTPRRRPVPVADCEPPPVDRLNLPDRKVKPAPLEIKPLPE, WT), bound to the single Grb2 SH3C domain with a K_d of $1.2 \mu\text{M}$ (Table 5.2, no. 24). The longer Gab1 peptide only slightly increased the binding affinity to Grb2 compared to the 32 aa Gab1 peptide (32 aa Gab1 peptide: K_d full length Grb2 $1.1 \pm 0.3 \mu\text{M}$; 45 aa Gab1 peptide: K_d full length Grb2 $0.6\text{-}0.9 \mu\text{M}$; Table 5.2, no. 16; 22 – 23). Therefore, it seems that the proline-rich amino acid sequence, adjacent to the Grb2 binding double motif, does not notably impact the Gab1 peptide-Grb2 binding.

To examine the contribution of the Grb2 SH3N binding to the overall binding of the 32 aa Gab1 peptide, an alanine-mutated SH3N motif Gab1 peptide was synthesized (FRSSAKTTPARPVPVADSEPPPVDRLNLPDRK). Interestingly, the mutant Gab1 peptide bound with the similar binding affinity to Grb2 as the wildtype Gab1 peptide (mutant Gab1 peptide K_d full length Grb2: $1.0 - 1.1 \mu\text{M}$; Gab1 peptide K_d full length Grb2: $1.1 \pm 0.3 \mu\text{M}$; Table 5.2, no. 20-21, 16). This would indicate that the Grb2 SH3N binding site only marginally contributes to this interaction.

5.6.3.2 Grb2 – Gab1 crystal screens

Analysis of the Grb2 – Gab1 peptide (32 aa), Grb2 – Gab1 peptide (45 aa) and p35Gab1 – Grb2 complex crystal setups showed rather different results (Figure 5.11). The Grb2 – Gab1 peptide (32 aa) screen demonstrated crystal with various shapes like needles and rods. Some crystals scattered X-rays to a resolution of 4 \AA . This was the highest resolution which we obtained for crystals from the initial screens. Crystals from the Grb2 – Gab1 peptide (45 aa) screen yielded a maximum resolution of 11 \AA .

Crystal set up	Initial Screen result
His-/ Grb2 + 32 aa Gab1 peptide	Crystals in approx. 20 conditions with best conditions (A – C, Grb2 + Gab1 peptide) in: <ul style="list-style-type: none"> • A: 0.05 M Tris/HCl pH 8.0, 28% (w/v) PEG 3350 • B: 0.1 M KSCN, 30% (w/v) PEGMME 2000 • C: 0.1 M HEPES pH 7.5, 20% (w/v) PEG 4000, 10% 2-propanol
Grb2 + 45 aa Gab1 peptide	Weakly diffracted irregular crystals (~ 11 Å resolution) in <ul style="list-style-type: none"> • 0.05 M Imidazole pH 7.0, 20% or 28% (w/v) PEG 3350 • 0.05 M HEPES pH 7.5, 20% (w/v) PEG 3350
Grb2 + p35Gab1 (TKB1)	Microcrystals in <ul style="list-style-type: none"> • 0.1 M Tris/HCl pH 8.5, 0.2 M tri-sodium citrate, 30% (v/v) PEG 400

Figure 5.11 Overview of initial crystallization screens

His-Grb2 (C32S, C198A) or Grb2 (C32S, C198A) without a His-tag was expressed in BL21, the p35Gab1 – Grb2 complex (His-p35Gab1 C374A, C405A, C514A; Grb2 C32S, C198A) was expressed in TKB1. Gab1 peptides were synthesized as indicated in Table 2.4.

It was difficult to crystallize the p35Gab1 – Grb2 complex, most likely due to the intrinsically disordered nature of the p35Gab1 protein. Only one condition in the p35Gab1 – Grb2 screen showed microcrystals after 147 days (0.1 M Tris/HCl pH 8.5, 0.2 M tri-sodium citrate, 30% (v/v) PEG 400). Removing some flexible parts of p35Gab1 could increase the possibility of crystal formation. Limited proteolysis with low concentrations of protease would be one method to establish more stable ‘subdomains’ of p35Gab1. Usually, exposed regions such as loops or other flexible regions rather than folded or bound regions are first attack points of proteases. Possibly, stable fragments could then be identified by SDS PAGE and mass spectrometry.

Due to limited time, we only focused on the most encouraging results obtained with Grb2 and the 32 aa Gab1 peptide (condition A-C). The Grb2 – Gab1 peptide (32 aa) complex fine screening and a full data set for a Grb2 – Gab1 peptide (32 aa) crystal is described in the following subchapter.

5.6.4 Fine screening and structure determination of a Grb2 – Gab1 peptide (32 aa) crystal

Grb2 – Gab1 peptide (32 aa) crystal screen conditions A, B and C (see Figure 5.11) from the initial screen were used as starting points for further fine screenings testing different buffer pH values and various precipitant concentrations. The fine screen based on condition A yielded crystals but with low

diffraction power, whereas the condition B-based fine screen failed to produce crystals. Crystallization in condition B of the initial screen was possibly initiated by unknown crystallization seeds specific to the well. However, the screen based on condition C (0.1 M HEPES pH 7.5, 20% (w/v) PEG 4000, 10% (v/v) 2-propanol) reproducibly generated crystals with various morphologies (Figure 5.12).

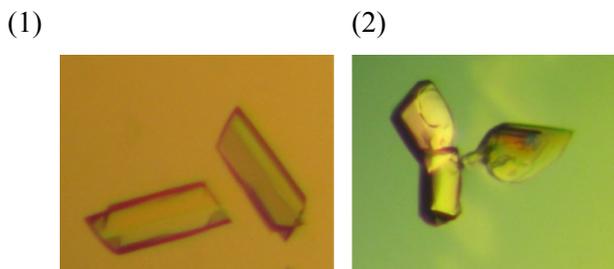


Figure 5.12 Grb2 – Gab1 peptide (32 aa) crystals (fine screen of condition C)

Different shapes of Grb2 – Gab1 peptide (32 aa) crystals. (1): 0.1 M Tris pH 8, 22% PEG 4000, 10% 2-Propanol (2): 0.1 M HEPES pH 7.5 10% 2-Propanol, 20% PEG 4000.

One Grb2 – Gab1 peptide crystal from the condition C fine screen (approx. crystal size: 0.12 x 0.06 x 0.03 mm) diffracted to $< 3 \text{ \AA}$ resolution in a diffraction pattern pre-test. The crystal was obtained after 14 days in the precipitant solution of 0.1 M Tris pH 8, 22% PEG 4000, 10% 2-Propanol at 13 °C. A full X-ray diffraction data set was collected on the in-house Cu anode (Prof. Milton Stubbs, MLU) and was processed to a final resolution of 2.41 Å (Table 5.3).

Beamline	In house Cu anode	BESSY II
Radiation wavelength (Å)	$\lambda = 1.5418$	$\lambda = 0.918$
Unit cell parameters (Å)	53.6 39.6 61.9	53.6 39.6 61.9
(°)	90 114.5 90	90 114.5 90
Space group	P2 ₁	P2 ₁
Resolution (Å)	2.41	2.2
Completeness (%)	98.9 (96.9)	96.2 (87.3)
R _{meas} (%)	10.8 (56.7)	6.6 (90.3)
I / $\sigma(I)$	10.8 (2.7)	17.5 (2.1)
CC _{1/2} (%)	99.5 (81.9)	99.9 (78.0)
Residues built / Total residues	234 / 255	n.d.
Number of protein atoms	2258	n.d.
Number of water atoms	90	n.d.
Resolution range (Å)	32.4 – 2.41	n.d.
Molecules / asymmetric unit	1	n.d.
R _{work} (%)	20.1	n.d.
R _{free} (%)	23.4	n.d.
R.m.s.d. bond angle (Å)	0.003	n.d.
R.m.s.d. bond lengths (°)	0.565	n.d.

Table 5.3 Data collection for Grb2 – Gab1 peptide (32 aa) crystal and structure building

Grb2 (C32S, C198A) – Gab1 peptide (32 aa) crystal structure refinement of in-house data set (values are not final). Values in the parentheses indicate the highest resolution shell: 2.41 – 2.56 Å. No model has been built for the BESSY II, yet (n.d.). The Grb2 – Gab1 peptide structure model contains a total number of residues of 255 (217 residues of Grb2, 32 residues of the Gab1-derived peptide and six residues from the Grb2 vector backbone).

An additional data set for the same Grb2 – Gab1 32 aa peptide crystal was generated at the BESSY II synchrotron (Helmholtz Centrum Berlin) at a 2.2 Å resolution. The Grb2 – Gab1 peptide crystal belongs to the monoclinic space group $P2_1$ ($a = 53.6$ Å; $b = 39.6$ Å; $c = 61.9$ Å; $\alpha = 90^\circ$; $\beta = 114.5^\circ$; $\gamma = 90^\circ$) with one molecule per asymmetric unit. Molecular replacement was used for the determination of initial phases. For this, data for individual Grb2 domains from various crystal structures was used (SH2 domain of hsGrb2/CD28pep, 1.4 Å, PDB code: 3wa4; SH3C of hsGrb2/Gab2bpep, 1.6 Å, PDB code: 2vwf and SH3N of hsGrb2, 3.1 Å, PDB code: 1gri) (Maignan et al., 1995, Harkiolaki et al., 2009, Higo et al., 2013). The structure was refined to R-values of $R = 20.1\%$ and $R_{\text{free}} = 23.4\%$. Crystal harvesting, data collection and structure determination was done by Dr. Constanze Breithaupt (Stubbs group, MLU).

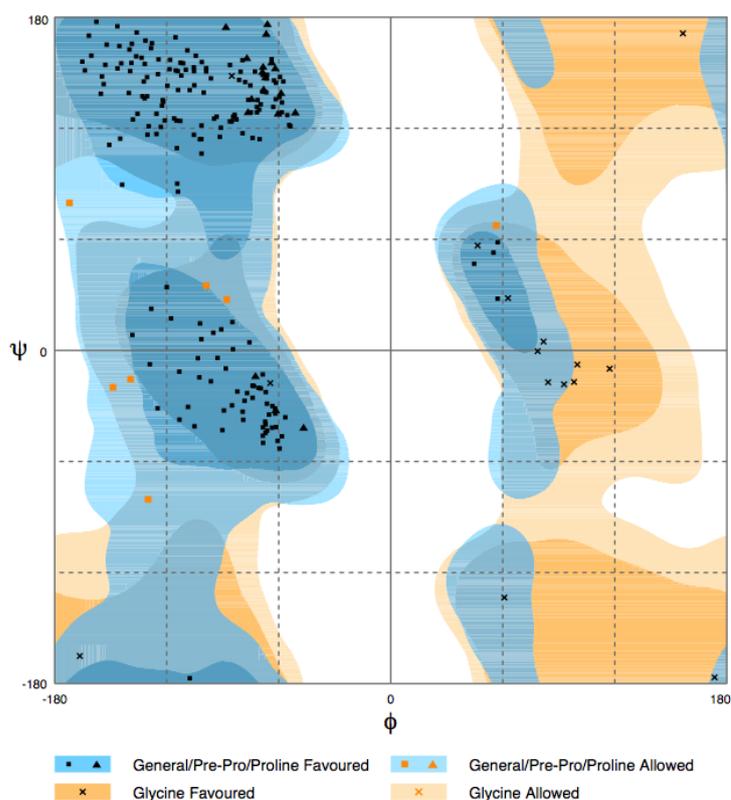


Figure 5.13 Ramachandran plot of the Grb2 – Gab1 peptide complex crystal structure

The Ramachandran plot depicts the dihedral angles Φ (phi, x-axis) and ψ (psi, y-axis) for an amino acid residue in the model. Shaded areas indicate allowed combinations of angles.

The quality of a structural model can be evaluated by a Ramachandran plot where backbone dihedral angles Φ (phi) and ψ (psi) of residues are plotted against each other. A RAMPAGE analysis of the Ramachandran plot for the modelled Grb2 – Gab1 peptide structure (Figure 5.13) showed that 96.9% residues are positioned in regions which are favoured for peptide bonds (saturated coloured regions), 3.1% are lying in allowed regions (light coloured) and none of the residues are found in the outlier regions (white coloured) (Lovell et al., 2003).

The synchrotron data set (BESSY) is currently under analysis. However, large differences between the two data sets are not expected. The structure which is discussed in the thesis represents the current state of data analysis calculated from the in-house Cu anode data set.

5.6.5 Overall structure of the Grb2 – Gab1 peptide complex

All Grb2 domains (SH3N: residues 1 – 57; SH2: residues 60 – 152; SH3C: residues 160 – 217) and the Gab1 peptide (FRSSPKTPRRPVPVADSEPPPVDRLNLPDRK; C514S) were visible in the crystal structure. Further 90 water molecules could be modelled. Grb2 SH3C residues 213 – 217 and Grb2 residues 153 – 156 from the highly flexible linker region between the SH2 and SH3C domain could not be modelled. Also, five of the six residues, originating from the Grb2 vector backbone (GSHMA), are not modelled. The accurate location of Gab1 residues 1 – 3, 16 – 18, and 32 could not be defined, most likely due to the high flexibility of those residues as they are not taking part in the Grb2 interaction. The overall stoichiometry in the Grb2 – Gab1 peptide crystal structure is 1:1. This agrees with the results from ITC measurements (Table 5.2; no. 16; $N = 1.1 \pm 0.2$). In the crystal, the Gab1 peptide region harbouring the SH3N binding site interacts with the Grb2 SH3N domain of one Grb2 molecule, whereas the SH3C binding region from the same peptide is bound to the SH3C domain of a neighbouring Grb2 molecule (see Figure 5.14). The Grb2 SH2 domain and the Grb2 SH3 domains show a typical SH2 and SH3 domain core structure (Maignan et al., 1995).

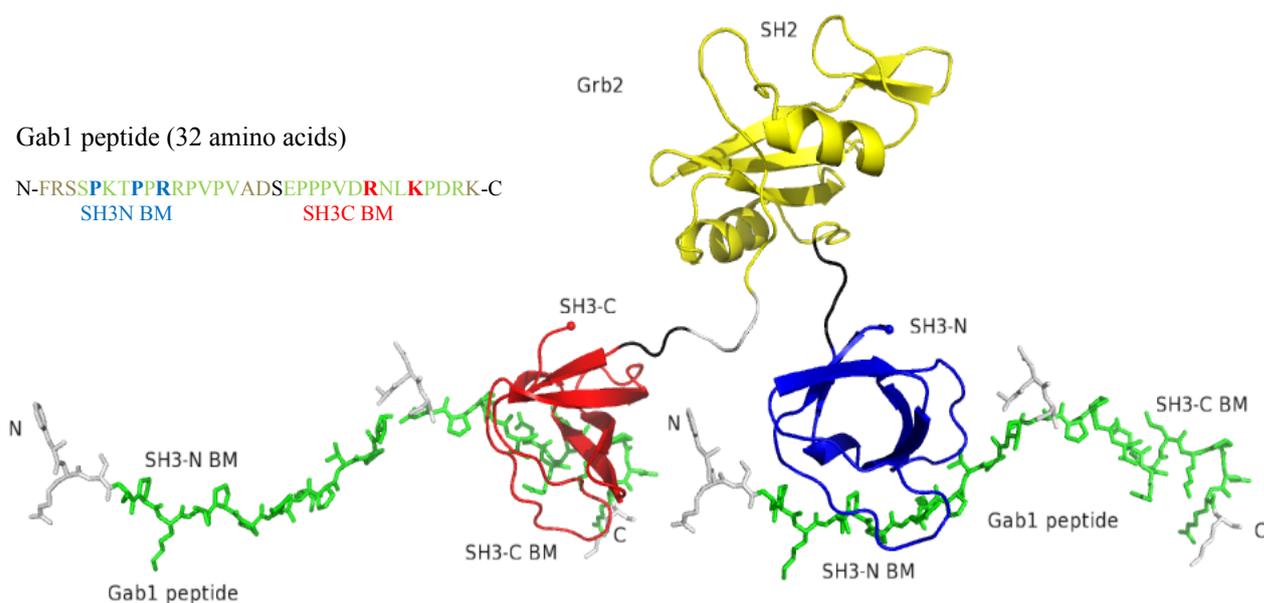


Figure 5.14 Structure of Grb2 complexed with two Gab1-derived peptides harbouring the Grb2 tandem motif

Ribbon model of Grb2 (human; C32S, C198A) in complex with two Gab1 peptides (32 aa, green). The amino acid sequence of the Gab1 peptide with indicated Grb2 SH3 domain binding motifs (binding motif: BM) is shown on the top. Grb2 is composed of one central SH2 domain (yellow) and two flanking SH3 domains (SH3C or SH3-C: red, SH3N or SH3-N: blue). Grb2 linker regions are coloured

in black. Grey coloured regions in the Grb2 model and the Gab1 peptide represent residues which could not be built into the model (Grb2 residues 153 – 156; Gab1 peptide residue 1 – 3, 16 – 18 and 32). Grb2 C-terminal residues 213 – 217 are not shown in the Grb2 model as they could not be built. N: N-terminus, C: C-terminus. Grb2 N- and C-terminus are presented as dots.

Gab1 peptide interaction modes with the Grb2 SH3N and the Grb2 SH3C are individually discussed in the following subchapters.

5.6.6 Structural features of the Gab1 peptide – Grb2 SH3N interaction

The Gab1 peptide interacts with the Grb2 SH3N domain via the P-x-x-P-x-R motif (Figure 5.15). This type of motif binding has been already shown for the N-terminal Grb2 SH3 domain complexed with a Sos-derived peptide by using NMR (Goudreau et al., 1994, Terasawa et al., 1994, Wittekind et al., 1994). Proline residues in the Gab1 peptide interact with the Grb2 SH3N by forming hydrophobic contacts with the Grb2 surface. The Gab1 peptide ligand adopts an PolyProline-II (PPII-helix) when interacting with the Grb2 SH3N domain. Studies have shown that proline-rich SH3 domain ligands often adapt a PPII helix conformation (Goudreau et al., 1994, Musacchio et al., 1994, Yu et al., 1994). (Gab1 peptide Lys6 - Arg10). The PPII-helices are characterized by an extended left-handed helix, defined by a rise per residue of ~ 3.1 Å in contrast to ~ 1.5 Å in α -helices (reviewed in (Adzhubei et al., 2013)).

P-x-x-P-x-R motif mediated SH3 domain-ligand interactions are characterised by ligand binding to three adjacent SH3 domain surface subsites S1, S2 and S3 (Feng et al., 1994, Goudreau et al., 1994, Lim et al., 1994, Musacchio et al., 1994, Terasawa et al., 1994, Wittekind et al., 1994, Wu et al., 1995, Wittekind et al., 1997). In more detail, the pyrrolidine ring of Pro5 stacks into a groove on Grb2 formed by Tyr7_{Grb2} and Tyr52_{Grb2}, defined as subsite S1 (Figure 5.15, B). Pro8 interacts with a groove on Grb2 formed by Pro49_{Grb2}, Trp36_{Grb2} and the aliphatic portion of Asn35_{Grb2}. This groove is described as subsite S2.

H-bonds between the peptide carboxyl oxygen of Pro5, Lys6, Pro8 and Pro9 with the Tyr52_{Grb2}, Asn51_{Grb2}, Trp36_{Grb2} and Asn35_{Grb2} respectively are stabilizing the interaction (Figure 5.16). In addition, the positively charged side chain of arginine 10 (Arg10) (P-x-x-P-x-R) sticks into a negatively-charged Grb2 groove (subsite S3) (Figure 5.15). It forms hydrogen bonds with its Nⁿ² to the O^{δ2} of Asp15_{Grb2}, and with its Nⁿ¹ to the O^{δ1} of Asp15_{Grb2} and the O^{ε2} of Glu16_{Grb2}. The interaction is strengthened by hydrophobic contacts between the aliphatic portion of Arg10 and Trp36_{Grb2} with a distance between 3 to 3.7 Å. The binding is further sustained by H-bond formation between the backbone of Arg11 with Asp33_{Grb2} and Gln34_{Grb2} (Figure 5.16).

A

Gab1 peptide (aa 1 – aa 32)

FRS **SPKTPRRPVP**VADSEPPPVDRLNKPDRK

B

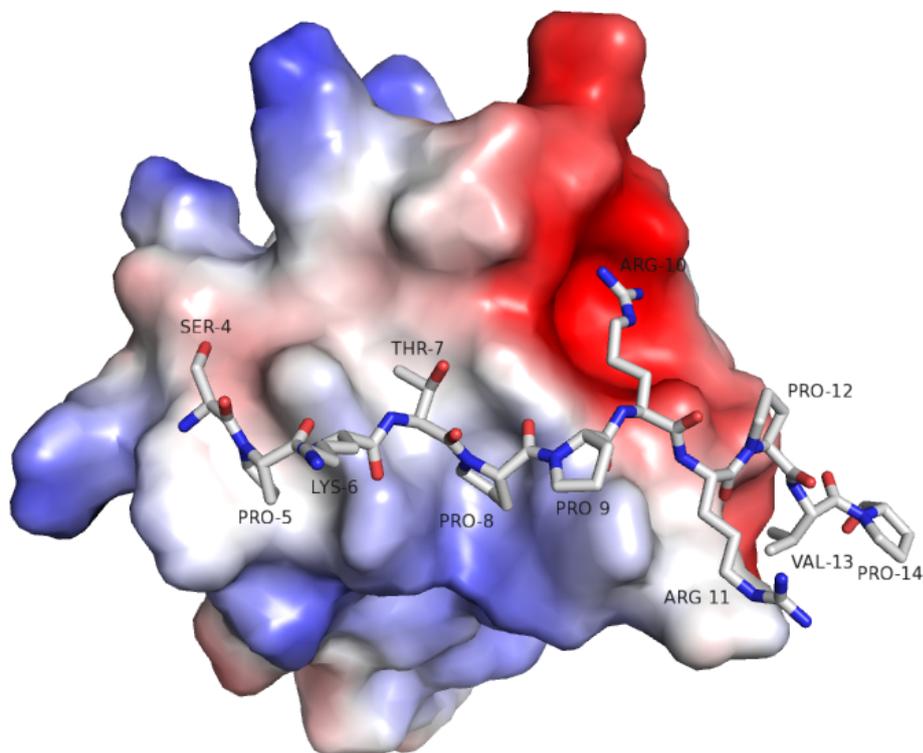


Figure 5.15 Grb2-SH3N interaction with a Gab1-derived peptide

The Gab1 peptide binds the Grb2 SH3N domain via a P-x-x-P-x-R motif. (A) Depiction of the full length Gab1 peptide amino acid sequence (aa 1 – aa 32, grey coloured). Residues interacting with the Grb2 SH3 domain are coloured in black and are underlined. P-x-x-P-x-R motif residues in the Gab1 peptide are indicated in bold. (B) The electrostatic potential was mapped onto the surface of the N-terminal SH3 Grb2 domain with the Gab1 peptide (aa 4 – aa 14, stick representation). The Grb2 SH3N – Gab1 peptide interaction area is 690.6 Å² in size (PDBePISA). Negative potential is drawn in red and positive potential in blue. Electrostatic potential surface calculation was done with PyMOL. The numbering of the Gab1 residues results from the original position in the peptide. Colour code for atoms in Gab1 peptide representation: oxygen: red, nitrogen: blue, carbon: grey.

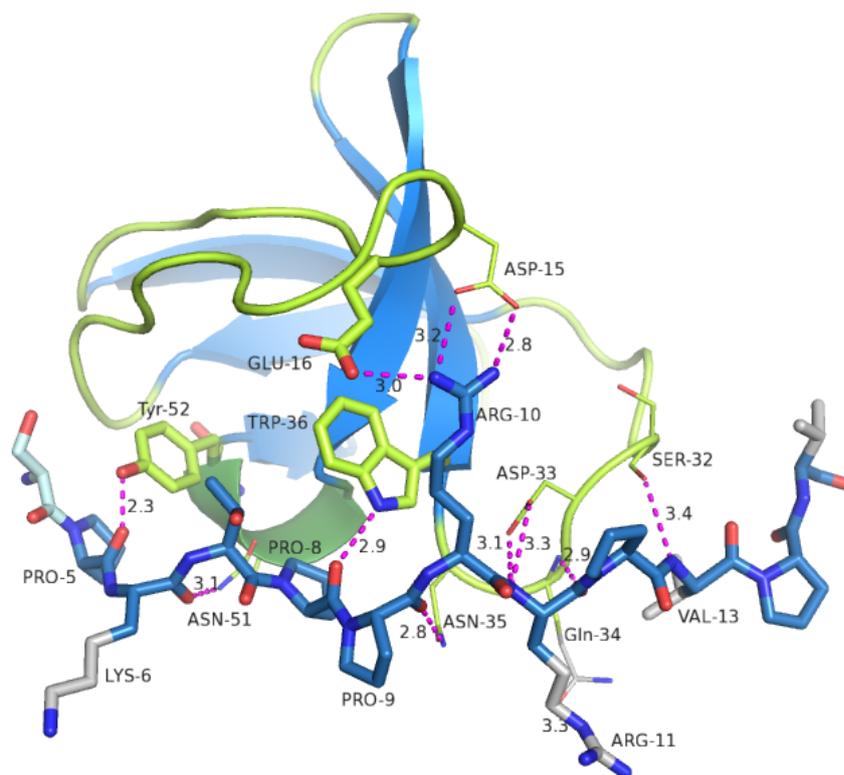


Figure 5.16 Detailed representation of the Grb2 SH3N – Gab1 peptide interaction site

A close-up view of the Grb2 SH3N domain ligand-binding site bound with a Gab1 peptide (aa 4 – aa 14 from the full length 32 aa Gab1 peptide) with indicated hydrogen bonds. Hydrogen bonds are shown as magenta dashed lines and measured distances are indicated in [Å]. Grb2 SH3N is represented in a cartoon mode and residues forming H-bonds with Gab1 residues are shown in stick representation. The Gab1 peptide is blue coloured and displayed in a stick representation. Labels from Grb2 or Gab1 peptide residues are only shown for residues taking part in an intermolecular H-bond formation. Residues or elements of residues which have not been built in the crystal structure are coloured in grey. Colour code for shown residue atoms: oxygen: red, nitrogen: blue, carbon: metallic blue (Gab1 peptide) and green (Grb2 SH3N).

As mentioned before, we performed the Grb2 – Gab1 peptide crystal screens with a double mutant Grb2 (C32S, C198A) to prevent the formation of disulfide bonds. Previous NMR studies had shown that the C32S mutation, which is located in the Grb2 SH3N domain, had no effect on the overall Grb2 SH3N domain structure (Goudreau et al., 1994). In the crystal, Ser32_{Grb2} forms an intramolecular H-bond and its carboxyl-group forms a weak polar contact with the main chain NH-group of Val13_{Gab1} (Figure 5.16). As Ser32_{Grb2} contributes only peripherally to the Gab1 peptide binding, the binding is probably not greatly affected by the mutation. However, ITC measurements between the Gab1 peptide and mutant/ WT Grb2 SH3N domain should be done to show if the mutation has an effect on the Gab1 peptide – Grb2 SH3N interaction.

Many of the Grb2 SH3N residues which have been found to interact with the Gab1 peptide in the Grb2 – Gab1 peptide structure have been previously identified to be involved in Sos peptide binding or as critical residues for other Grb2 SH3 domain-mediated interactions (Terasawa et al., 1994). A sequence alignment of the SH3N domains of Grb2 or Grb2-like proteins shows that most of the Grb2

SH3N domain residues contributing in the Gab1 – Grb2 SH3N interaction are conserved throughout metazoan evolution which underlines their importance in binding proline-rich ligands (Figure 5.17). The importance of the Gab1 P-x-x-P-x-R motif in the Grb2 SH3N – Gab1 interaction, as seen in the crystal structure, was confirmed by ITC. Binding between Grb2 and a Gab1 peptide that was mutated to alanine at positions P-x-x-P-x-R (FRSSAKTPPARPVPV) was completely abolished (Table 5.2, no. 15).

	10	20	30	40	50
H. sapiens Grb2	MEAIAKYDFKATAD	DELSFKRGDILKVLN	EECDQN	WYKAELNGKDGFI	PKNYIEM
D. rerio Grb2	MEAIAKYDFKATAD	DELSFKRGEILKVLN	EECDQN	WYKAELNGKDGFI	PKNYIEM
D. melanogaster Drk	MEAIAKHDFSATAD	DELSFRKTQILKILN	MEDDSN	WYRAELDGKEGLI	PSNYIEM
C. elegans Sem5	MEAVAEHDFQAGSP	DELSFKRGNTLKV	LNKDED	PHWYKAELDGNEGFI	PSNYIRM
N. vectensis Grb2-like	MEARAKHEFKATQ	EDDLGFAKGSILN	VLDMDQ	DKN	WYKAEQDGREGWI

Figure 5.17 Grb2 SH3N residues contributing in the P-x-x-P-x-R Gab1 peptide interaction are highly conserved

Sequence alignment of the SH3N domains of Grb2 or Grb2-like proteins in selected metazoan species. Amino acids highlighted in red are involved in the Grb2 SH3N – Gab1 peptide ligand binding, identified in the Grb2 - Gab1 peptide structure, and if conserved in Grb2 or Grb2-like proteins of other species marked as bold. Protein sequences were retrieved from UniProt. Accession codes: *Homo sapiens* (human) Grb2 protein: P62993; *Danio rerio* (zebrafish) Grb2 protein: Q6PC73; *Drosophila melanogaster* (fruit fly) Drk (Esev2B) protein: Q6YKA8; *Caenorhabditis elegans* (roundworm) Sem-5 protein: P29355; *Nematostella vectensis* (starlet sea anemone) predicted Grb2-like protein: A7SE08. Sequence from the *N. vectensis* genome are from predicted Grb2-like proteins. All sequences were manually aligned.

There are two opposite ligand orientations possible in a P-x-x-P-mediated ligand binding (Feng et al., 1994, Lim et al., 1994). The orientation of the ligand is determined by the relative position of a positively charged residue (K/R) interacting with a conserved acidic amino acid on the SH3 peptide-binding surface (in the Grb2 – Gab1 peptide structure: Arg10 of the Gab1 peptide). Class I ligand binding is mediated by an (K/R)-x-x-P-x-x-P motif whereas a class II peptide binds via an x-P-x-x-P-x-(K/R) motif. Figure 5.18 clearly illustrates that the Gab1 peptide (green) binds in a class II ligand orientation to the Grb2 SH3N, similar to a FynSH3 – APP12 class II peptide (magenta) (Camara-Artigas et al., 2016). A class I consensus peptide (PI3K SH3 domain complexed with a PD1R peptide, yellow) binds in the opposite direction (Batra-Safferling et al., 2010). The N-terminal Grb2 SH3 domain seems to preferentially bind peptides with class II ligand orientation and not class I orientation for, so far, unknown reasons (Sparks et al., 1996).

Superimposing Gab1 peptide-bound Grb2 SH3N onto free Grb2 SH3N (PDB code: 1gri) illustrates that the binding of the Gab1 peptide does not significantly change the overall Grb2 SH3N domain structure (Figure 5.19) which had been previously shown for the Sos – Grb2 SH3N complex (Goudreau et al., 1994).

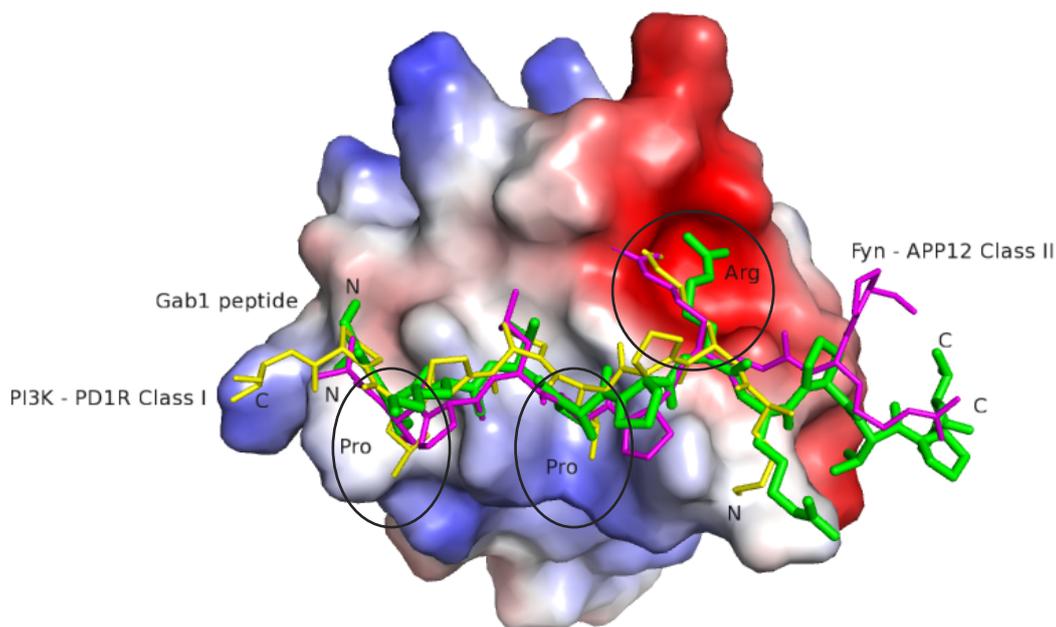


Figure 5.18 Gab1 peptide binds to the Grb2 SH3N in a class II orientation

Electrostatic potential mapped onto the surface of the Grb2 SH3N domain in complex with Gab1 peptide ligand (green), class II peptide (purple, PDB code: 4znx) and class I peptide (yellow, PDB code: 3lr5) ligands. All peptides are shown in a stick representation. Black circled areas indicate the location of critical interaction cavities for the P-x-x-P-x-R motif. Electrostatic potential surface calculation was done with PyMOL. Negative potential is drawn in red and positive potential in blue. N: N-terminus, C: C-terminus



Figure 5.19 The overall structure of the Grb2 SH3N domain remains merely unchanged upon Gab1 peptide binding

Gab1 peptide-complexed Grb2 SH3N compared to free Grb2 SH3N. Comparison of Gab1 ligand-bound Grb2 SH3N (coloured in blue) to free Grb2 SH3N (PDB code: 1gri, coloured in purple). Both SH3 domains are presented as $C\alpha$ -traces.

5.6.7 Structure of the Gab1 peptide – Grb2 SH3C domain complex

The interaction between the Grb2 SH3C domain and the SH3C binding motif in Gab1 (R-x-x-K) is stabilised by salt bridges, hydrogen bonds and hydrophobic interactions (Figure 5.20). The Gab1 peptides adapts a 3_{10} helix conformation by binding to the Grb2 SH3 C domain similar to the Gab2 peptide in the Grb2 SH3C – Gab2b crystal structure (Harkiolaki et al., 2009). The 3_{10} helix is more compact, having 3.0 amino acids per turn compared to the more common α helix, having 3.6 amino acids per turn (Toniolo and Benedetti, 1991). Gab1 residues which are important in the Grb2 SH3C binding are numbered by their relative position regarding to the arginine residue in the R-x-x-K binding motif (see 32 aa Gab1 peptide amino acid sequence, Figure 5.20, A).

A

Gab1 peptide (aa 1 – aa 32)

FRSSPKTPRRRPVPVADS E P P P V D R N L K P D R K
 -5 -4 -3 -2 -1 0 +1 +2 +3

B

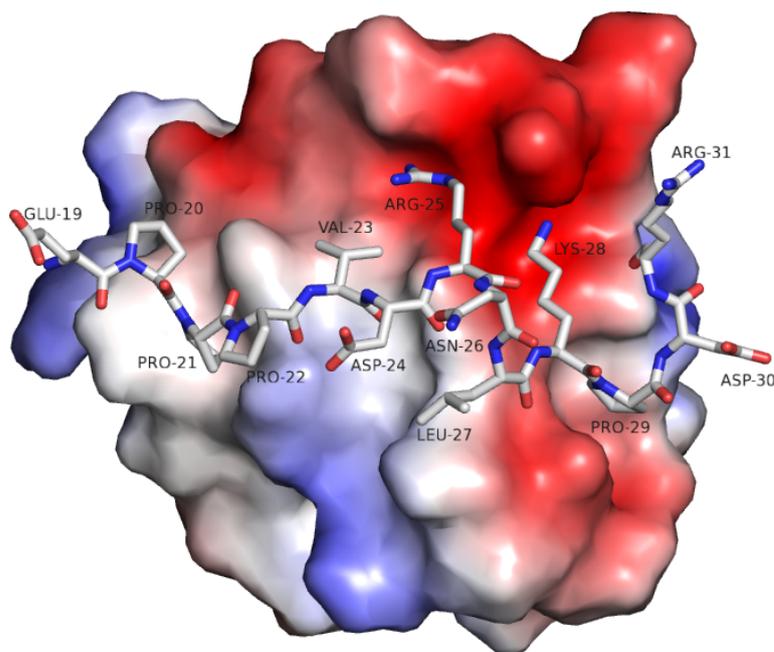


Figure 5.20 Complex between the Grb2 SH3C domain and a Gab1-derived peptide

Grb2 SH3C domain in complex with a Gab1-derived peptide. The domain-ligand interaction is mediated via a R-x-x-K motif (Arg-25; Lys-28). (A) Depiction of the 32 aa full length amino acid sequence of the Gab1 peptide (aa 1 – aa 32, coloured in grey). Gab1 peptide residues (aa 19 – aa 31) which are seen in complex with the Grb2-SH3C in figure part (B) are coloured in black and are underlined. Highly critical residues, as arginine 25 (Arg25) and lysine 28 (Lys28), are indicated in bold letters. Gab1 peptide residues are numbered by their relative position regarding to the arginine residue from the R-x-x-K motif. (B) Surface of the Grb2 SH3C domain in complex with the Gab1 peptide (aa 19 – 31 aa, grey) in stick representation. Negative potential is drawn in red and positive potential in blue. Electrostatic potential surface calculation was done with PyMOL. Negative potential is drawn in red and positive potential in blue. In the Grb2 SH3C – Gab1 complex, Gab1 peptide residues are numbered with the original number as in the full length 32 aa Gab1 peptide sequence. Colour code for atoms in Gab1 peptide representation: oxygen: red, nitrogen: blue, carbon: grey.

atoms in the backbone of Leu₊₂ (Leu27) and Asp₋₁ (Asp24), and Lys₊₃ (Lys28) and Arg₀ (Arg25). These H-bonds stabilise the 3₁₀ helix conformation in this region of the Gab1 peptide (Figure 5.21).

SH3 domain ligand interactions mediated by an R-x-x-K motif have been shown in other SH3 domain-ligand complex structures (Harkiolaki et al., 2003, Kaneko et al., 2003, Lewitzky et al., 2004, Harkiolaki et al., 2009). The critical positions of the arginine and lysine residues in the Gab1 R-x-x-K motif were determined by ITC. Binding between full-length Grb2 and Gab1 peptide with an R-x-x-K motif mutated to alanine (15 aa; SEPPPVDANLAPDRK; C514S, R521A, K524A; Table 5.2, no. 11) was completely abolished.

The 32 aa Gab1 tandem peptide in the crystal structure contains the sequence equivalent of the R-x-x-K binding motif of Gab2 (Gab2b) which has been characterised in the Grb2 SH3C – Gab2b crystal structure (Harkiolaki et al., 2009). The 32 aa Gab1 peptide is more than twice as long as the 14 aa Gab2b peptide and contains more residues N-terminal of the R-x-x-K motif. It also harbours the newly found Grb2 SH3N binding motif (Figure 5.22).

```

human Gab1      490-  PPPAHMGFRSSPKTPPRRVPVADCEPPPVDRNLKPDRKVKPAPLEIKPLPEW
human Gab2      494-  T-----TLPVHRGPSRGSEIQPPPVNRNLKPDRKAKPTPLDLRNNTVI
                                * *      : :      : * * * * : * * * * * * * * . * * : * * : :
    
```

Figure 5.22 Sequences of the human 32 aa Gab1 and the 14 aa Gab2b peptides that were crystalized with the Grb2 SH3C domain

The 32 aa Gab1 peptide (C514S) crystallized with full length Grb2 is highlighted in green. The Gab2b peptide crystallized with the Grb2 SH3C is highlighted in orange (PDB code: 2vwf, (Harkiolaki et al., 2009). Alignment was made with Clustal Omega with * (asterisk) for fully conserved residues and : (colon) for conserved residues with similar properties. Protein sequences were retrieved from UniProt (Accession codes of human Gab1: Q13480 and human Gab2: Q9UQC2).

An interface area of 582.3 Å² was calculated for the Gab1 peptide complexed with the Grb2 SH3C domain (PDBePISA server). In comparison, the interaction surface of the Gab2b peptide – Grb2 SH3C complex is only slightly smaller, 537.7 Å².

A comparison of the Gab1 and the Gab2b peptide interactions with the Grb2 SH3C domain was done by superimposition of both structures in PyMOL (Figure 5.23, A, PDB code for Grb2 SH3C – Gab2b: 2vwf). The Gab1 – Grb2 SH3C domain interaction is very similar to the Gab2 – Grb2 SH3C domain interaction, especially in regards to the residues from the R-x-x-K motif significantly contributing to the ligand - domain interaction (PDB code: 2vwf, Harkiolaki et al., 2009). This can be explained by the high conservation of the R-x-x-K motif in Gab1 and Gab2 (Harkiolaki et al., 2009). Non-conserved residues in the Gab1 R-x-x-K motif were found to be not involved in the Gab1 – Grb2 binding. The Gab1 and the Gab2b peptide adopt an extended conformation allowing multiple contacts with the SH3C binding site. The Gab1 peptide has, similar to the Gab2b peptide, a 3₁₀ helix conformation. A difference between the Gab1 – and Gab2 – Grb2 SH3C structures can be found in the orientation of an arginine residue outside of the R-x-x-K motif (Arg31 in Gab1 and Arg14 in Gab2b).

In the Grb2 SH3C – Gab2b structure, Arg14 forms two strong H-bonds (2.5 Å and 2.9 Å) with Glu171_{Grb2}. In contrast, Arg31 of the Gab1 peptide is not involved in the Grb2 SH3C – Gab1 peptide binding at all and therefore flexible and not visible in the electron density map. Figure 5.23 shows the Gab1 Arg31 side chain in the calculated most favoured position.

To examine changes in the overall Grb2 SH3C domain structure upon Gab1 peptide binding, the SH3C of the Grb2 – Gab1 crystal structure was superimposed in PyMOL onto a SH3C domain of the unbound dimeric Grb2 without peptide ligand (Figure 5.23, B). The Gab1 peptide binding only minimally changes the overall structure of the Grb2 SH3C domain, which is not unexpected given to the compactness and rigidity of SH3 domains.

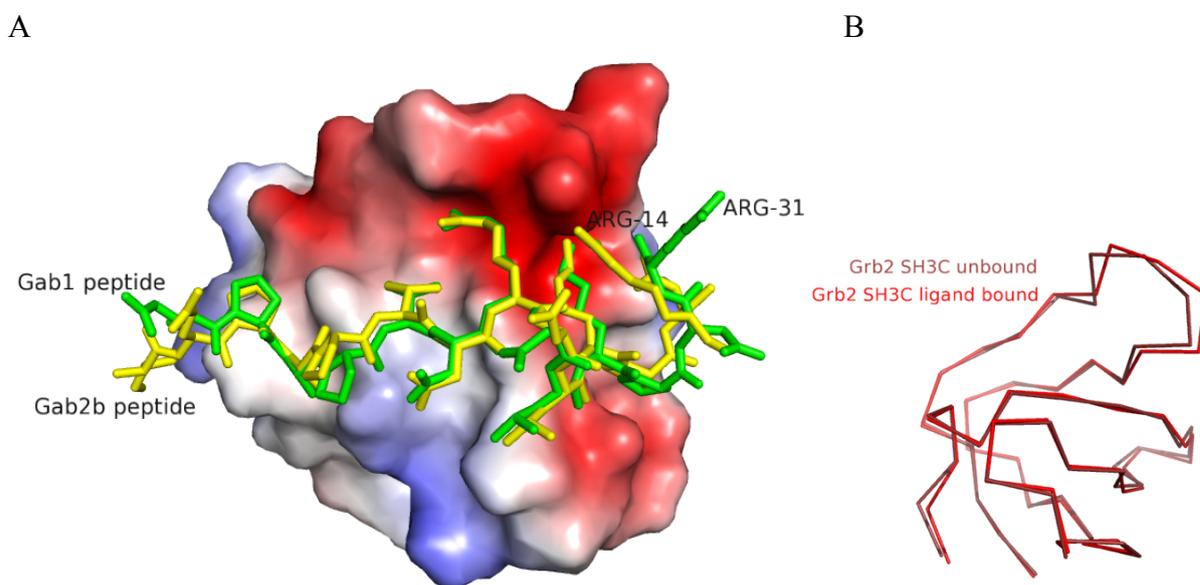


Figure 5.23 Grb2 SH3C domain complexed with Gab family peptides and superposition of Gab1 peptide-bound and unbound Grb2 SH3C domains

(A) Surface of the C-terminal Grb2 SH3 domain bound to the 32 aa Gab1 peptide (green) and a Gab2b peptide (yellow) (PDB code: 2vwf). Electrostatic potential surface calculation was done with PyMOL. Negative potential is drawn in red and positive potential in blue. (B) Superposition of Gab1 peptide-bound and unbound Grb2 SH3C domains. C α -traces of Grb2 SH3C Gab1 bound (red) and Grb2 SH3C unbound (purple, PDB code: 1gri, Maignan et al., 1995).

5.6.8 Structural features of the Grb2 SH2 domain in the Grb2 – Gab1 crystal

The Grb2 SH2 domain from Grb2 – Gab1 peptide crystal (coloured in yellow) was compared to a ligand-free Grb2 SH2 domain (coloured in light brown) and to an isolated Grb2 SH2 domain in complex with a phosphorylated tyrosine-ligand (coloured in green) (Figure 5.24). Most of the more rigid structural features of the Grb2 SH2 domain, β -sheets and α -helices, are very similar in all of the SH2 domains. Differences were detected in the flexible loop regions connecting β -sheets and α -helices of the SH2 domain. The orientation of the Grb2 SH2 BG- and BC-loops in the Grb2 – Gab1 peptide

crystal (yellow) was found to be more similar to the pY ligand-bound Grb2 SH2 domain (green) than to the unbound Grb2 SH2 domain (brown) although no specific SH2 domain ligand was added to the crystallization solution (Grb2 loops named according to Eck et al., 1993). A crystallographic Grb2 neighbour molecule interacts with the SH2 binding site and mimics the binding of a pY-ligand (residues are coloured in blue, Figure 5.24). A Grb2 aspartate side chain binds into the same SH2 binding pocket which is usually occupied by the phosphate of a pY-ligand.

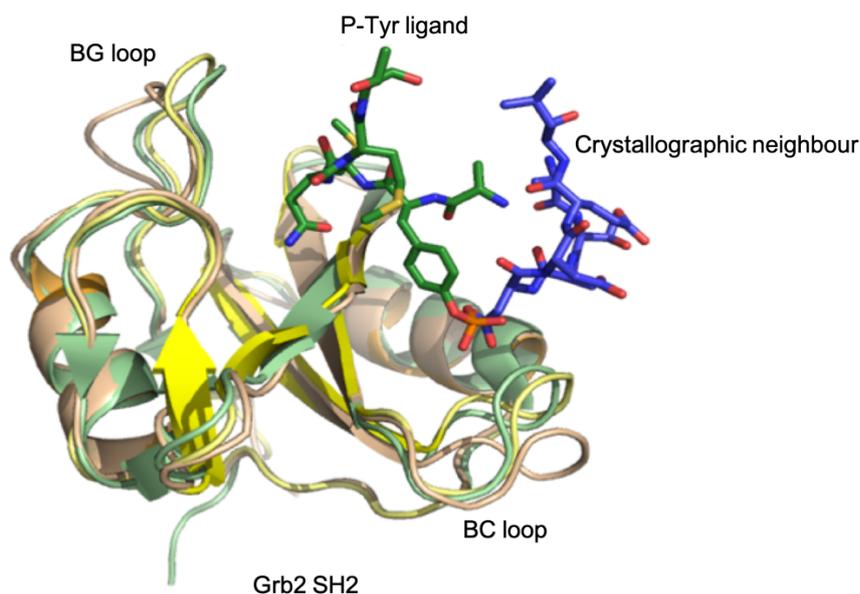


Figure 5.24 Superposition of Grb2 SH2 domains

The Grb2 SH2 domain of the Grb2 – Gab1 peptide complex (coloured in yellow) is superimposed on an SH2 domain from unbound Grb2 (dimeric Grb2 PDB code: 1gri, coloured in light brown) and a P-Tyr ligand-bound SH2 domain (PDB code: 3wa4, coloured in green, P-Tyr ligand peptide is shown in a stick representation). A crystallographic neighbour peptide is shown in stick representation coloured in blue.

Moreover, the crystal packing shows a domain - domain interface between the SH2 domain of one Grb2 molecule and the Grb2 SH3C domain from a neighbouring Grb2 molecule which is illustrated in Figure 5.25 (SH2 domain of Grb2 molecule 1 is coloured in yellow, SH3C domain of Grb2 molecule 2 is coloured in red). The originally introduced C198A mutation in the Grb2 SH3C domain is directly located in the SH2 – SH3C interface area but it does not influence the SH3 domain structure (Figure 5.25, B). However, binding studies between the Grb2 SH2 and Grb2 SH3C WT or a C198A mutant should be conducted, to show if the mutation impacts onto the formation of the interface in solution.

Interestingly, the same SH2 – SH3C interface was found in the dimeric Grb2 structure (Maignan et al., 1995). In the Grb2 dimer, the SH2 domain of one Grb2 protomer interacts with the SH3 domain of the other Grb2 protomer and thereby forming a large part of the Grb2 dimerisation interface (Figure in appendix A.6). Superimposing the SH2 and SH3C domains of the Grb2 – Gab1 peptide complex onto

the SH2 and SH3C domain of the dimeric Grb2 demonstrates the high similarity of the relative domain orientations (Figure 5.25, Gab1-bound Grb2 SH2 domain: light colour, unbound Grb2: dark colour; PDB code: 1gri).

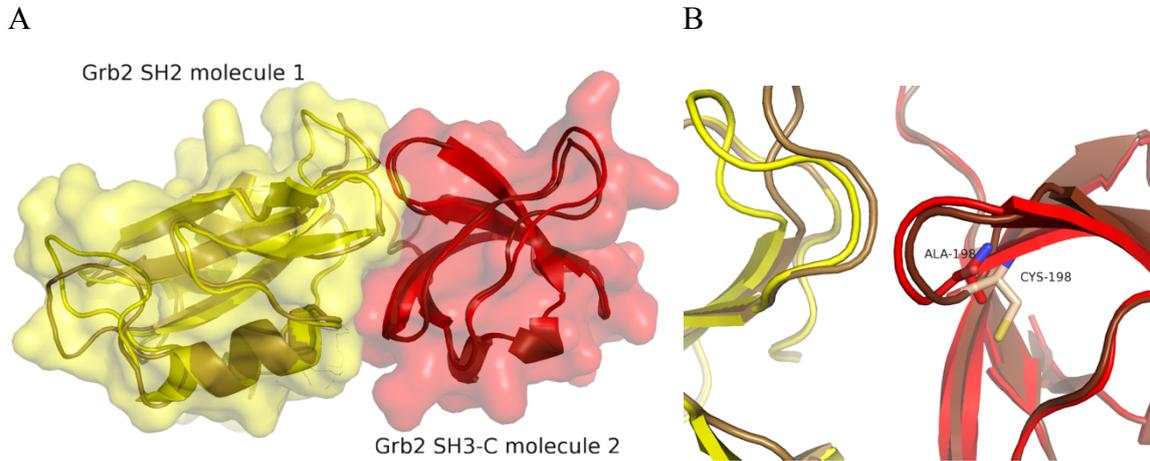


Figure 5.25 Interface between Grb2 SH2 domain and neighbouring Grb2 SH3C domain

The crystal package of the Grb2 – Gab1 peptide complex depicts an interface between the Grb2 SH2 domain (coloured in yellow; Grb2 molecule 1) and a neighbouring Grb2 SH3C domain from another Grb2 molecule (coloured in red; Grb2 molecule 2). (A) An overlay of the Grb2 domains from the Gab1-bound Grb2 (light colour) with the Grb2 domains from the (unbound) dimeric Grb2 (dark colour; PDB code: 1gri) shows that a very similar interface is present in the dimeric Grb2. The surface model is represented from the Gab1-bound Grb2 structure. (B) The Grb2 SH3C located C198A mutation is located directly in this SH2 – SH3 interface but it does not change the general structure of the SH3C domain.

Altogether, the emerging Grb2 SH2 – SH3C interface in two independent crystal packings could suggest that it plays a physiological role rather than being an artefact due to crystal packing. In the following subchapter, a theory is postulated of how the domain – domain interface can play a role inside the cell.

5.6.9 Grb2 – Gab1 oligomerisation model

The interaction between the Grb2 SH2 and Grb2 SH3C domains of two individual Grb2 protomers could provide the basis for a Grb2 – Gab1 protein cluster formation which extends into two dimensions (Figure 5.26). Individual Grb2 molecules would be linked intermolecularly by the SH2 – SH3 domain interface leading to an oligomerization in one axis. A second extension of the Grb2 – Gab1 cluster along the other axis could be mediated by the Gab1 peptide which links the SH3C domain from one Grb2 protomer with the SH3N domain from another Grb2 protomer.

Similar protein cluster formations have been shown to facilitate important functions in various cell processes such as in receptor signalling or autophagosome assembly (Huang et al., 2016, Yamamoto et al., 2016). The Grb2 protein together with the Sos protein oligomerize the LAT protein involved in T-

cell receptor activation to enhance T-cell receptor signalling power (Huang et al., 2016). In yeast, the intrinsically disordered protein Atg13 provides an linkage between pre-assembled molecules to form the autophagosome membrane (Yamamoto et al., 2016). In these examples, the protein oligomerization is activated by tyrosine phosphorylation. The potential Grb2 – Gab1 cluster consists of only SH3 mediated interactions which are stimulus-independent. Grb2 could be a regulator itself, as highly concentrated Grb2 has a tendency to form oligomers (Ahmed et al., 2015).

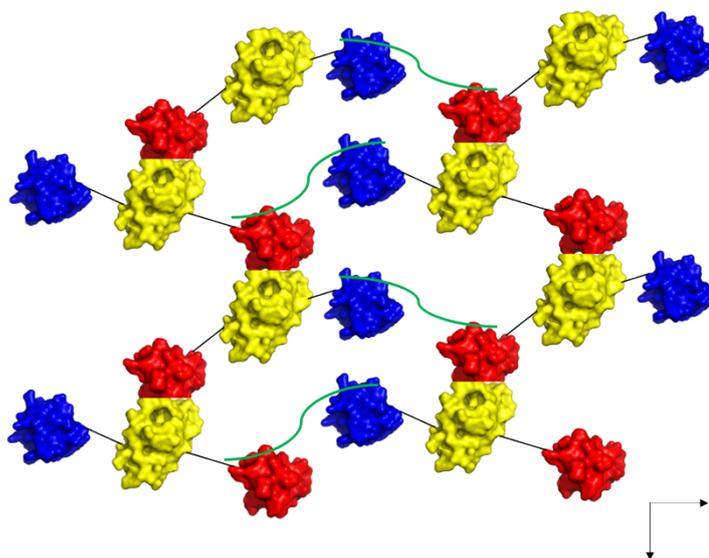


Figure 5.26 Schematic depiction of potential 2D Grb2 – Gab1 cluster

Colour coding: Grb2 SH3N: blue, Grb2 SH2: yellow, Grb2 SH3C: red, Gab1-derived peptide: green; linker regions: dark grey. Grb2 domains are represented in surface mode. The linker region and the Gab1 peptide are shown as lines.

The protein cluster is more likely to be established close to the membrane than in the cytoplasm, as Grb2 and Gab1 are highly involved in signalling processes which take place in close proximity to the receptor, such as the c-Met receptor. Then, upon stimulation, accumulated Grb2 and Gab1 clusters could contribute to an efficient signalling transmission. As the Grb2 – Gab1 cluster is mediated by SH3 domain interactions, which have usually weak binding affinities, and an SH3 – SH2 domain interface with an unknown binding affinity, it could be more of a transient Grb2 – Gab1 cluster than a robust oligomeric protein complex. Possibly, the Gab1 PH domain interaction with membrane-located PIP₃ could also play a role in stabilizing a cluster.

Studies have shown that the epidermal growth factor receptor (EGFR) dimerise and form clusters of receptors on the membrane upon stimulation (Chung et al., 2010, Gao et al., 2015). EGFR is clustered to enable efficient signal transduction but could also raise the question if EGFR pathway signalling proteins as Grb2 and Gab1 also need to be available in a relatively high concentration. A Grb2 – Gab1 cluster could be potentially useful to generate a local enrichment of ready-to-use Grb2 and Gab1 signalling proteins.

Important for further characterization of a potential Grb2 – Gab1 cluster is the analysis of a possible influence of the introduced C198A mutation (Grb2 SH3C domain) as it is directly located in the Grb2 SH2 – SH3 interface (Figure 5.25).

5.7 Conclusion and discussion

We were able to purify sufficient amounts of homogenous p35Gab1 – Grb2 complex to analyse the Grb2 – Gab1 interaction by various methods. Low resolution images of single p35Gab1 – Grb2 complex particles were detected by initial AFM and EM analysis. Sample resolution in both methods was relatively poor due to the small size of the p35Gab1 – Grb2 complex. Another alternative method to try would be Cryo-electron microscopy which has recently made major progress in the analysis of small proteins (< 100 kDa) (Wu et al., 2012, Fernandez-Leiro and Scheres, 2016, Merk et al., 2016).

We identified a new interaction site between Gab1 and the N-terminal Grb2 SH3 domain by combining various methods such as XL-MS, ITC and X-ray crystallography. The Gab1 – Grb2 SH3N interaction is mediated by a highly conserved P-x-x-P-x-R motif in the Gab1 protein. In fact, a Gab1 – Grb2 SH3N interaction had been previously proposed but precipitation experiments failed to confirm this interaction (Lewitzky et al., 2001). One possibility for that could be that the low binding affinity of the Gab1 – Grb2 SH3N interaction ($K_{d \text{ Grb2 SH3N}} 50.5 \pm 3.8 \mu\text{M}$) prevented an effective pull down of Gab1 – Grb2 SH3N complexes from cell lysates.

The Grb2 SH3N interaction site in Gab1 is located only 14 aa upstream of the R-x-x-K Grb2 SH3C binding site. Both binding sites were therefore denoted as a Grb2 SH3 domain-binding tandem motif. From crystal screens with full length Grb2 (C32S, C18A) and a Gab1 peptide containing the Grb2 tandem motif, we obtained a crystal structure at a 2.41 Å resolution. This is a significant improvement in structural resolution compared to the only previously reported Grb2 full-length crystal structure from 1995 with a resolution of 3.1 Å (Maignan et al., 1995). In the crystal the Gab1 tandem peptide binds to both Grb2 SH3 domains but from different Grb2 molecules. Due to the observed peptide binding mode, the Gab1 peptide could neither bind to the SH3N and SH3C domain of one Grb2 molecule in the dimer nor to the SH3N domain of one dimer-protomer and the SH3C domain of another protomer, simultaneously. If the Gab1 peptide might be able to bind simultaneously to both SH3 domains in a single Grb2 molecule could be examined by *in silico* modelling.

The structural features of the Gab1 – Grb2 SH3C structure closely resembled the published crystal structure of the Grb2 SH3C domain in complex with a Gab2b peptide harbouring a similar R-x-x-K motif (PDB code: 2wvf, Harkiolaki et al., 2009). A comparison of the Grb2 – Gab1 and Grb2 – Gab2 crystal structures demonstrated that the R-x-x-K motif-related interactions are highly conserved between Gab1 and Gab2. In the Gab1 – Grb2 SH3N interaction, the Gab1 peptide with the P-x-x-P-x-R motif bound to the Grb2 SH3N domain with a class II orientation similar to previously

published class II-oriented P-x-x-P-x-R ligands of SH3 domains (Terasawa et al., 1994, Wittekind et al., 1994, Sparks et al., 1996, Wittekind et al., 1997).

The overall Grb2 – Gab1 peptide crystal packing showed a cluster of Grb2 proteins connected by Gab1 peptides. Interestingly, the Grb2 – Gab1 cluster was so far only observed by X-ray crystallography and not by other methods (AFM, EM, AUC or SEC). One reason for this could be that it only forms when Grb2 is mixed with a Gab1 peptide and not p35Gab1; another reason could be that cluster formation could be concentration-dependent, and X-ray crystallography uses a highly concentrated Grb2 – Gab1 peptide sample compared to the other methods. However, the obtained X-ray crystallography result reflects only one single possible interaction mode and the potential risk of crystallization artefacts cannot be excluded. Therefore, a potential Grb2 – Gab1 cluster formation definitely needs further experimental confirmation. An approach could be to use cells with fluorescently labelled Grb2 and/or Gab1 proteins and analyse stimulated and unstimulated cells by microscopy.

A motif search with the Grb2 SH3 domain-binding tandem motif sequence in humans identified another protein, the adaptor protein Garem1, which contains a similar tandem binding motif. Sequence analysis demonstrated that the Grb2-binding tandem motif in Garem1 is similarly conserved as in the Gab1 protein. The emergence of the motif in two unrelated proteins could be a hint that the Grb2 tandem motif emerged twice as a result of convergent evolution, which could suggest that the Grb2 SH3 domain-binding tandem motif plays an important role in cellular function.

Chapter 6 Results and Discussion: Assembly of a protein complex comprised of p35Gab1, Grb2 and the PI3K regulatory subunit p85 α

6.1 Overview

Cell signalling protein complexes can be structurally characterised by combining methods such as NMR, EM and X-ray crystallography (Francis et al., 2004, Stauber et al., 2006, Qiao et al., 2013). This chapter describes the initial attempts of assembling a minimal signalling complex comprised of three signalling proteins, the Gab1 fragment p35Gab1 (Le Goff et al., 2012), Grb2 and the PI3K (Phosphoinositide 3-kinase) regulatory subunit p85, to study the structural features of this complex and its dynamics. The Gab1 (p35Gab1) – Grb2 – p85(PI3K) complex plays, for example, a crucial role in the HGF/SF c-Met receptor signalling pathway (see Figure 1.4). The modular composition of the p35Gab1, Grb2 and p85 α protein is displayed in Figure 6.1. p85 α is an isoform of the p85 subunit of Class IA PI3Ks (Backer, 2010). PI3Ks are classified into three classes (I-III), by which the class IA PI3Ks are coupled to receptor tyrosine kinases such as the c-Met receptor and composed of a p110 catalytic and a p85 regulatory subunit (Vanhaesebroeck et al., 2012).

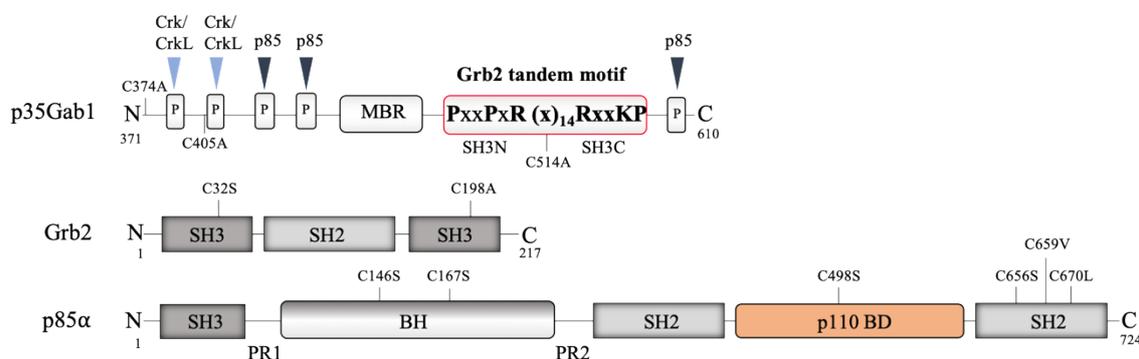


Figure 6.1 Schematic overview of p35Gab1, Grb2 and the PI3K regulatory subunit p85 α (cysteine-free mutant)

p35Gab1 (aa 371 – aa 610), the functional fragment of full-length Gab1, contains SH2 domain binding sites (P) for Crk/CrkL proteins and p85 α , a c-Met binding region (MBR) and a Grb2 tandem binding motif. Grb2 (aa 1 – aa 217) with its central SH2 domain flanked by an N- and C-terminal SH3 domain. p85 α (aa 1 – aa 724) contains a SH3 domain, two SH2 domains, two proline-rich regions (PR1 and PR2), a Breakpoint Cluster Region homology domain (BH) and the p110 binding domain (p110 BD) for PI3K catalytic subunit p110 dimerisation. Generated mutations, as described in the main text, are indicated for each protein.

The p35Gab1 fragment still contains all three p85 pY-x-x-M consensus motif binding sites (Y447, Y472, Y589), the c-Met binding region and our newly described Grb2 tandem binding motif (Rocchi et al., 1998, Lehr et al., 2000, Le Goff et al., 2012) (see Figure 1.3).

The p85 α protein (aa 1 – aa 724) is composed of an SH3 domain, two SH2 domains, two proline-rich regions (PR1 and PR2), a Breakpoint Cluster Region homology domain (BH) and the p110 binding domain (p110 BD) for PI3K catalytic subunit p110 dimerisation (Figure 6.1). For the p35Gab1, Grb2 and p85 complex assembly studies, a cysteine-free mutant p85 α protein was recombinantly expressed in *E. coli* (BL21, CodonPlus-RIL), purified and analysed by AFM. p85 α (cysteine-free mutant) was added to the previously purified tyrosine-phosphorylated His-p35Gab1 – Grb2 complex from *E. coli* TKB1 cells. We used mutant Grb2 (C32S, C198A) and mutant p35Gab1 (C374A, C405A, C514A) vector to obtain a complex with high purity (previously shown in Chapter 4). Successful complex formation was subsequently assessed by methods that maintain the native protein structure such as native polyacrylamide gel electrophoresis (native-PAGE) and SEC.

6.2 Protein purification and AFM analysis of PI3K p85 α

For the p85 α protein purification, we used a cysteine-free human p85 α construct (C146S, C167S, C498S, C656S, C659V) kindly provided by Dr. Backer (Albert Einstein College, New York). The cysteine-free mutant p85 α protein was previously characterised by LoPiccolo et al. and was more soluble than the wildtype p85 α (LoPiccolo et al., 2015). The group confirmed similar functionality in cysteine-free and wildtype p85 α protein. The cysteine-free mutant p85 α subunit of PI3K is referred to as p85 α (mut.) in the following text.

p85 α (mut.) expression and purification was performed according to a modified version of the protocol described in LoPiccolo et al, 2015 (see 2.2.1.5.2). GST-tagged p85 α (mut.) (ca. 100 kDa; p85 α : ca. 84 kDa, GST: ca. 27 kDa) from bacterial lysate was immobilized on GSH beads (Figure 6.2, A; sample: ‘On beads’). Optimal GSH beads amount had been initially determined by a titration experiment. After extensive washing, the immobilized GST-tag of p85 α was cleaved with PreScission protease 3C (ON at 4 °C). The optimal 3C amount of a self-made 3C protease was determined by titration experiments. Using an GST-tagged 3C protease (GST-3C) enabled us to easily separate cleaved p85 α protein (ca. 84 kDa, sample of fractions (Fr) 1-3) from the protease (ca. 40 kDa) and residual GST-p85 α protein (ca. 100 kDa) which were both immobilized on GSH beads. A GSH bead sample after cleavage indicated a virtually complete cleavage of GST-p85 α as the sample only detectably contained the GST-tag (25 kDa) and residual amounts of cleaved p85 α (ca. 84 kDa) (see sample ‘On beads after elution’). The obtained p85 α protein was dialysed against Mono Q buffer (20 mM Tris pH 8, 20 mM NaCl) and centrifuged after dialysis to remove dirt particles and precipitates (Figure 6.2, B). Protein precipitation was not observed (compare ‘after dialysis’ samples: ‘-’ = before centrifugation and ‘cyclic arrow’ = after centrifugation).

Cleaved p85 α (mut.) was loaded on a Mono Q column and eluted with a linear salt gradient which separated the p85 α protein from smaller contaminants. The Mono Q flow-through (FT) and the wash sample did not contain any p85 α protein which suggests a strong interaction between p85 α and the Mono Q material. The sharp UV peak at the end of the Mono Q elution profile is not caused by

Results and Discussion: Assembly of a protein complex comprised of p35Gab1, Grb2 and the PI3K regulatory subunit p85 α

proteins as seen on the SDS gel (fraction 40). SEC of a concentrated p85 α (mut.) sample suggests a p85 dimer formation as it eluted with ca. 63 ml (Monomer: approx. 84 kDa, expected: 74 ml elution volume; Dimer: approx. 168 kDa, 66 ml elution volume) (Figure 6.2, C). A dynamic p85 α monomer - dimer equilibrium has been described with both, monomer and dimer, performing discrete functions in PI3K signalling (LoPiccolo et al., 2015). As described in the paper, a high protein concentration, a low NaCl concentration and a low temperature promote the formation of dimeric p85 α in solution.

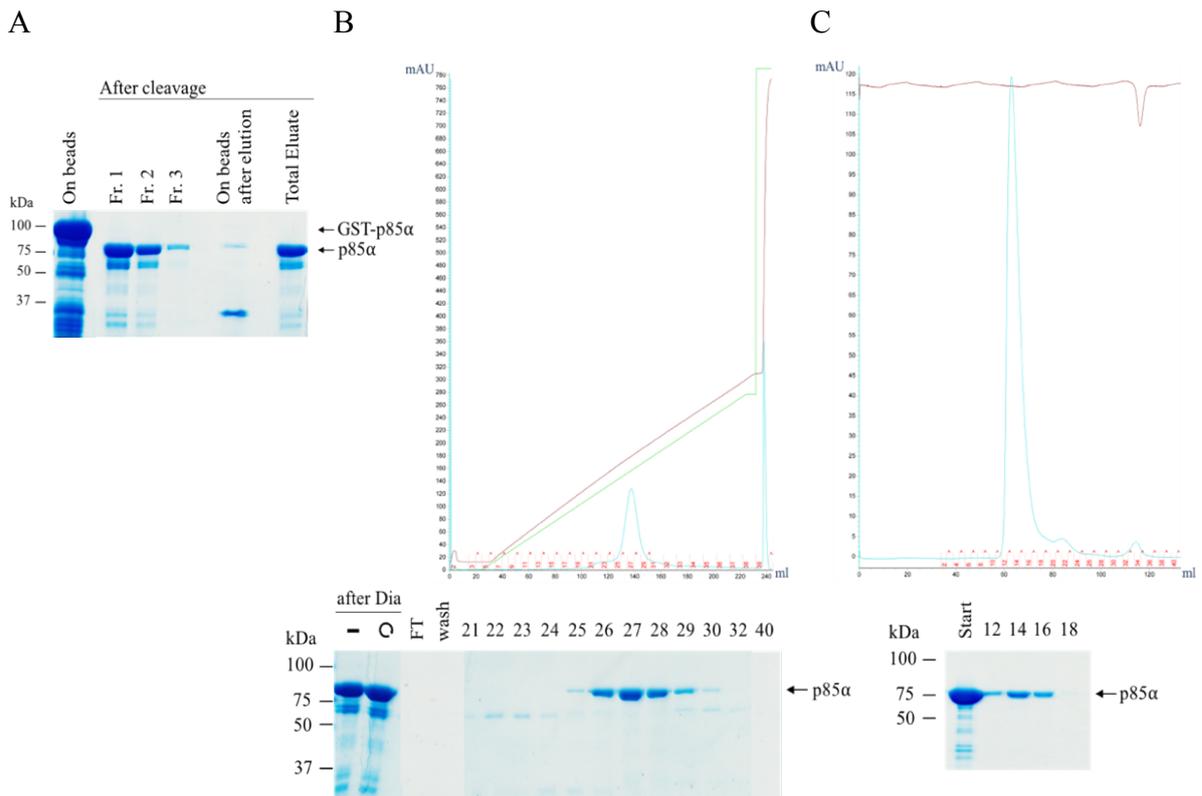


Figure 6.2 Purification of the PI3K regulatory subunit p85 α (mut.)

Three-step purification of p85 α (mut.). Samples taken during the purification were analysed by SDS PAGE (7.5% resolving gel). (A) GSH-immobilized GST-p85 α was cleaved with 3C protease. (B) Mono Q (HiTrap Q HP) analysis of dialysed p85 α eluate sample. A linear gradient from 0 – 350 mM NaCl for 40 column volumes (indicated by the green line) was chosen. (C) SEC of p85 α on a S200 column (HiLoadTM 16/600 SuperdexTM S200). In the histogram, the blue line charts the UV absorbance in [mAU], shown at the y-axis, and the measured conductivity is shown as the brown line. The x-axis shows the elution volume of the sample in [ml]. Collected Mono Q and SEC fractions are indicated with red numbers. Fr: fraction, FT: flow through.

For Atomic Force Microscopy (AFM) analysis, purified p85 α (mut.) was immobilized on a mica surface, dried and analysed by a Molecular Force Probe - 3D AFM. The analysis was performed in AFM tapping mode which offers high resolution with minimum sample damage. AFM was done in collaboration with Dr. Ingrid Tessmer (Rudolf Virchow Center, University of Würzburg). As seen in Figure 6.3 (A), the AFM scan of p85 α (mut.) demonstrates distinct single particles. A zoom into one image shows heterogeneously distributed particles and potentially p85 α domain boundaries. The fact that p85 α can adopt a range of distinct conformations could contribute to the heterologous sample

Results and Discussion: Assembly of a protein complex comprised of p35Gab1, Grb2 and the PI3K regulatory subunit p85 α

appearance in the AFM image (LoPiccolo et al., 2015). The ability to change between conformational states expands the capacity of the PI3K regulatory subunit p85 to interact with a range of different partner proteins.

An AFM volumetric analysis estimated the molecular weight distribution in the p85 α (mut.) sample against an intern AFM calibration using proteins with known molecular masses. The p85 α (mut.) sample contained a 90 kDa population, which most likely corresponds to a monomeric p85 α (Figure 6.3, B). Interestingly, no p85 α dimers were measured with AFM (ca. 180 kDa). The largest population in the sample belongs to smaller sized protein populations (60 kDa, 40 kDa, and smaller proteins). These have been also seen in the SEC elution profile but as a very small population (see Figure 6.2, C). One explanation could be that these are cleavage products generated by a contaminating protease that were co-purified with the target protein.

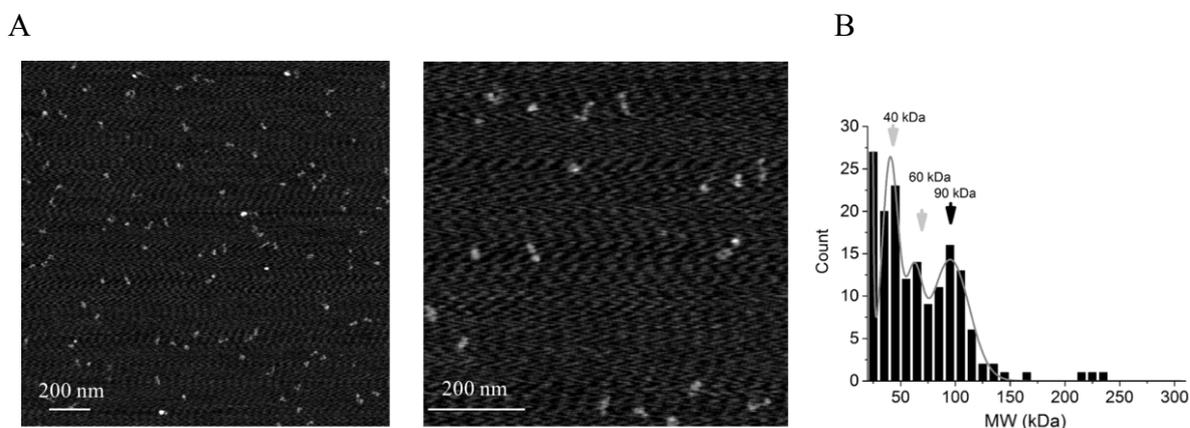


Figure 6.3 Atomic Force Microscopy (AFM) of PI3K regulatory subunit p85 α

AFM analysis of p85 α . (A) AFM image of p85 α (mut.) (left) and a zoomed image (right) show individual particles with various shapes. (B) AFM volumetric analysis of the p85 α (mut.) sample displays a mixture of protein species with distinct molecular weights of 40 kDa, 60 kDa and 90 kDa. The graph illustrates their various distribution in the sample. Volumetric analysis estimated the molecular weight [kDa] according to a protein standard.

6.3 Tyrosine phosphorylation of p85 binding sites in p35Gab1

The p35Gab1 – Grb2 complex was expressed in TKB1 to phosphorylate p85-binding sites in p35Gab1 and to enable binding of the p85 SH2 domain. Immunoblotting the p35Gab1 – Grb2 complex, expressed either in BL21 or TKB1, with an anti-phosphotyrosine antibody (4G10) demonstrated phosphorylation of p35Gab1 and, more moderately, Grb2 protein in TKB1 (Figure 6.4, A). Tyrosine phosphorylation was not observed for BL21-expressed proteins. Grb2, which contains several tyrosine residues in its sequence, was also phosphorylated. A broader detected band (like here in Grb2) could be the result of different phosphorylation patterns on the same protein. A p35Gab1 mass spectrometry analysis in collaboration with Prof. Benedikt Kessler (Target Discovery Institute, University of

Results and Discussion: Assembly of a protein complex comprised of p35Gab1, Grb2 and the PI3K regulatory subunit p85 α

Oxford) confirmed phosphorylation on all p85 binding sites (pY447, pY472 and pY589) and on one of the two Crk/ CrkL proteins binding sites (Y406) (Figure 6.4, B).

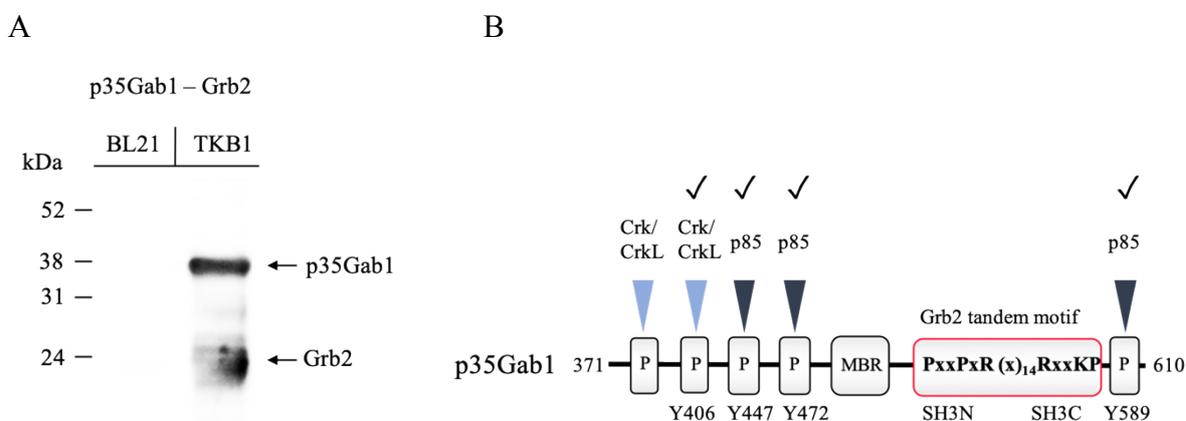


Figure 6.4 Phosphorylation of p85 binding sites in p35Gab1

p85 binding sites in the p35Gab1 protein are phosphorylated by complex expression in the bacterial strain TKB1. (A) BL21- or TKB1-expressed p35Gab1 – Grb2 complex was immunoblotted with an anti-phosphotyrosine antibody (4G10). (B) Schematic overview of p35Gab1 and phosphorylated p85 binding sites in p35Gab1 (black tick: detected by mass spectrometry analysis). Specific tyrosine residues (Y) are indicated at the bottom. MBR: c-Met binding region, P: tyrosine residues which could be phosphorylated and serve as binding sites for SH2 domain proteins.

6.4 Preliminary analysis of the assembled p35Gab1 – Grb2 – p85 complex

A tyrosine-phosphorylated p35Gab1 – Grb2 sample and a p85 α (mut.) sample (approx. 1 mg/ml, 20 mM Tris, 300 mM NaCl) were mixed at a 1:1 molar ratio. The SEC elution profile of the p85 α suggests that it forms homodimers in the absence of other proteins (see Figure 6.2, C). Phosphatase inhibitors sodium orthovanadate and sodium molybdate were added to the mixed samples to prevent dephosphorylation of p85 binding sites in p35Gab1. Complex formation was assessed after a short (1 h) and a long (ON) incubation at RT by native-PAGE (Figure 6.5, A). One sample was incubated ON at 4 °C to examine complex formation at a low temperature. Protein integrity of each sample was verified by SDS PAGE stained with colloidal Coomassie InstantBlue.

As seen on the native-PAGE, the mixed p35Gab1 – Grb2 – p85 α sample forms largely a single protein band with a strong shift in comparison to the individual control samples. This most likely corresponds to a p35Gab1 – Grb2 – p85 α complex. The complex quickly assembles and remains stable ON (compare samples 1 h, RT; ON, RT and ON, 4 °C).

The SEC profile for a p35Gab1 – Grb2 – p85 α sample confirmed the formation of a trimeric complex (Figure 6.5, B). A comparison of the elution profile for a p35Gab1 – Grb2 – p85 α mixture (red) with the elution profiles for p85 α (brown) or for the p35Gab1 – Grb2 complex (blue) clearly showed that in the mixed p35Gab1 – Grb2 – p85 α sample (red) a large part of the protein eluted in a single peak and

Results and Discussion: Assembly of a protein complex comprised of p35Gab1, Grb2 and the PI3K regulatory subunit p85 α

at an earlier time than for the single p35Gab1 – Grb2 complex (blue) control sample. This strongly indicates the formation of a larger protein complex in the mixed sample. The molecular size of each protein sample was estimated according to the elution volume of each sample and a protein-size standard curve specific for the column.

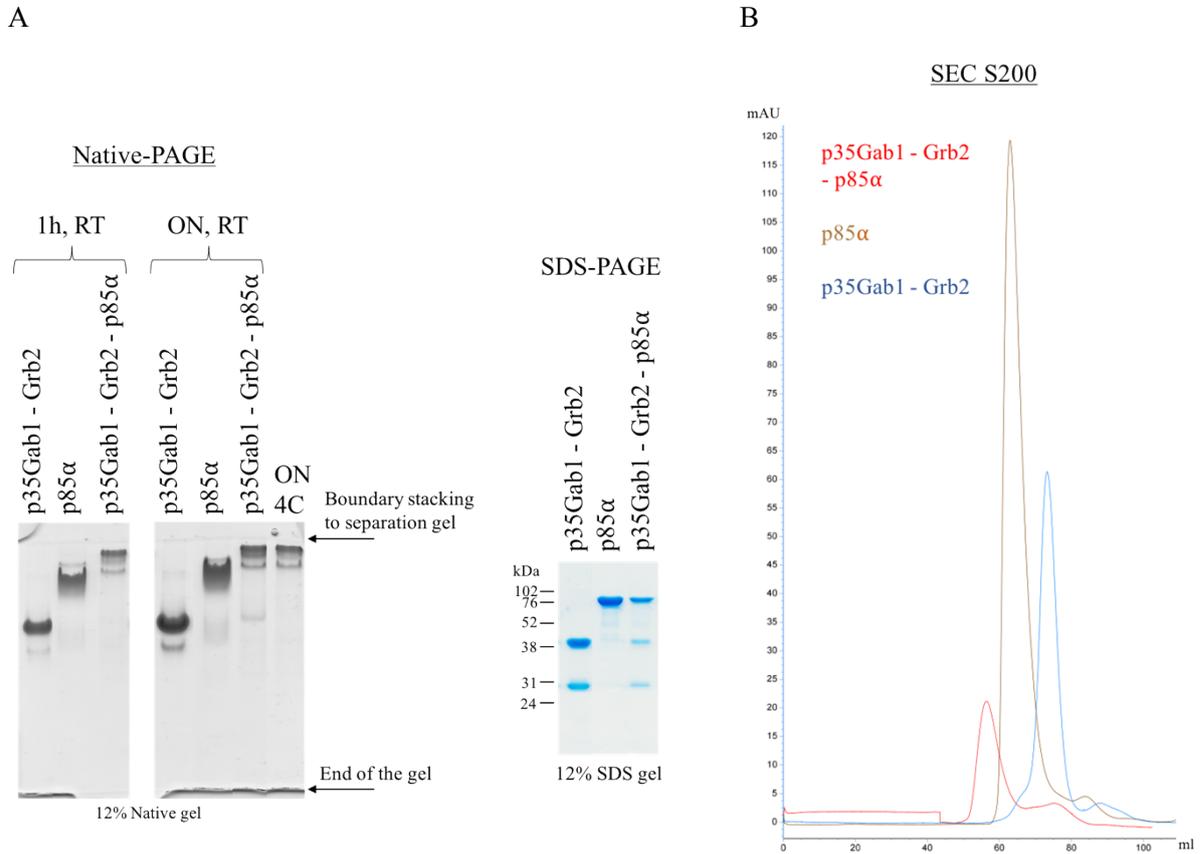


Figure 6.5 p35Gab1 – Grb2 – p85 α complex formation

(A) Native-PAGE (12% resolving gel) loaded with p35Gab1 – Grb2 – p85 α , p85 α and p35Gab1 – Grb2 complex (left). Purified p35Gab1 – Grb2 complex (cysteine-free His-p35Gab1 C374A, C405A, C514A; Grb2 C32S, C198A; approx. 1 mg/ml complex; 20 mM Tris pH 8, 150 mM NaCl) expressed in TKB1 were mixed at a 1:1 ratio with p85 α (p85 α mut.; approx. 1 mg/ml; 20 mM Tris, 300 mM NaCl) and incubated for 1 h or ON. Boundary of stacking and separating gel and bottom edge of native-PAGE gel are indicated by arrows. Control samples (1 h) resolved on a 12% SDS gel (right). Marker lane is indicated on the left. (B) Mixed p35Gab1 – Grb2 – p85 α sample (cysteine-free p35Gab1 C374A, C405A, C514A; p85 α (mut.); Grb2 C32S, C198A) (1 h, RT) was analysed on a Superdex S200 column (HiLoad 16/600 Superdex) in 20 mM Tris pH 8, 150 mM NaCl, 0.1 mM Molybdate, 0.1 mM Vanadate. Comparison of SEC S200 elution profiles of p35Gab1 – Grb2 – p85 α (red), p85 α alone (brown) and p35Gab1 – Grb2 complex (blue). The x-axis shows the elution volume of the sample in [ml]. The y-axis measures the UV absorbance in [mAU].

The p85 α (mut.) sample eluted presumably as a p85 α -dimer at an elution volume that corresponds to a size of approx. 213 kDa. Similar apparent molecular masses for p85 α were estimated by SAXS (LoPiccolo et al., 2015). The mass, estimated for the elution volume, for the p35Gab1 – Grb2 complex was 89 kDa protein which would be consistent with a complex comprised of one p35Gab1 molecule

Results and Discussion: Assembly of a protein complex comprised of p35Gab1, Grb2 and the PI3K regulatory subunit p85 α

(35 kDa) and two Grb2 molecules (2x 27 kDa). The calculated molecular mass for the formed complex is approx. 325 kDa which would fit a p35Gab1 : Grb2 : homo-p85 α -dimer complex at a 1:2:1 ratio. However, based upon a single technique, it is not possible to conclude with confidence any kind of stoichiometry, in particular, since p35Gab1 and p85 encompass intrinsically disordered regions of substantial size, which may lead to deviations from properties expected for globular proteins (predicted disorder: 34% for human p85A; UniProt accession code: P27986; 61% for human p35Gab1; Gab1 (Q13480) aa 371 – aa 610; determined by PONDR-VLXT). Data such as from AUC, NMR or X-ray crystallography would be useful to further investigate the ratio of the protein components in the complex. An initial attempt at AFM analysis was also performed of the trimeric complex but failed to obtain useful data, possibly due to inadequate sample storage, sample preparation, or other unknown reasons.

6.5 Conclusion and discussion

Our long-term goal is the assembly of Gab-based signalling complexes *in vitro* to better understand Gab functions in signalling pathway crosstalks. Here, we demonstrate that the requirements for a stable complex formation between the p35Gab1 – Grb2 complex and the PI3K regulatory subunit p85 α were fulfilled. The p85 binding sites in p35Gab1 were phosphorylated and p85 α (mut.) could be purified to high purity in milligram quantities. By native-PAGE and SEC, we were able to demonstrate that p35Gab1 – Grb2 and p85 α form a defined complex in solution. However, protein precipitation during or after complex assembly strongly reduced the protein amount of complex for SEC analysis. Due to complex formation, parts of the protein that were previously hidden by the protein could be exposed to the solvent and therefore change the requirements of a buffer to fulfil stabilizing features. Therefore, a strategy to enhance the solubility of the complex could be to modify the current buffer (20 mM Tris pH 8, 150 mM NaCl, 0.1 mM Molybdate, 0.1 mM Vanadate).

The size of the complex was estimated to be around 325 kDa which could represent a p35Gab1 : Grb2 : homo-p85 α -dimer complex at a 1:2:1 ratio. Interestingly, this would seem to agree with data from our Grb2 – Gab1 peptide crystal structure (see subchapter 5.6), which also indicates a p35Gab1 – Grb2 interaction involving two Grb2 molecules and one p35Gab1 peptide molecule. Once the currently observed solubility problems are resolved, the p35Gab1 – Grb2 – p85 α complex generated here, might be useful for structure analyses by NMR, EM or even X-ray crystallography.

Chapter 7 Final conclusions and future strategies

The large multisite docking protein Gab1 is an important protein in cell signalling. Gab1 binds to PIP₃-containing plasma membranes with its N-terminal PH domain and acts as an assembly platform for other signalling proteins. Gab1-based signalling complexes coordinate multiple signals and activate appropriate pathways, in normal development but also in various cancers. The Gab1 PH domain is most probably well-folded whereas the Gab1 protein ‘tail’, which contains most of the interaction sites, is predicted to be mostly disordered. Until now, there are no structural data available for the PH-domain or full-length Gab1, or the other Gab family members, Gab2 and Gab3. Preliminary work by our group and others has suggested that the Gab1 protein adapts a more compact structural topology to fulfil its function in cellular signalling processes (Eulendorf and Schaper, 2009, Simister et al., 2011). Major aims of this thesis were structural analysis of the Gab protein and Gab-based signalling complexes. Therefore, we attempted to purify the full-length Gab2 protein, and subsequently the functional p35Gab1 fragment, as well as the Gab1 interaction partners Grb2 and PI3K(p85).

During full-length Gab2 purification we observed strong protein aggregation. Purification of the Gab1 fragment p35Gab1 showed similar problems with protein solubility and stability. A substantial improvement to the p35Gab1 protein stability was accomplished by a co-expression of p35Gab1 with the adaptor protein Grb2 (Chapter 3 and Chapter 4). This finding suggests a new chaperone-like function of the Grb2 protein, additionally to its well-known adaptor functions. Intriguingly, a study by Kortum et al. revealed chaperone-like features of Grb2 (Kortum et al., 2013). The group demonstrated a Grb2/Sos1-dependent LAT oligomerisation in T cell receptor (TCR) signalling. The proposed chaperone-like feature of Grb2 is remotely reminiscent of the newly identified SecB-substrate interaction. The chaperone SecB interacts in a ‘wrapping’ mode with its substrate proteins to maintain them in an unfolded state (Huang et al., 2016). However, in contrast to the SecB-substrate interaction, Grb2 potentially preserves the functional topology of the IDP p35Gab1, rather than maintaining its unfolded state. A potential chaperone-like function of Grb2 could be a further explanation for the strong conservation of the Grb2 interaction motifs in Gab proteins (Gab1-3) in evolution (Harkiolaki et al., 2009). If the purification of full-length Gab or other Gab proteins could be improved by a co-expression of Grb2 still remains elusive.

We established a p35Gab1 – Grb2 complex purification protocol which yielded a relatively pure protein complex (Chapter 4). AUC and SEC analysis of the p35Gab1 – Grb2 complex confirmed a complex sample that was free of any major aggregates. First images of the p35Gab1 – Grb2 complex were obtained by AFM and EM but with a relatively low resolution (Chapter 5). In future, sample resolution could be either increased by optimizing the methodology, using more powerful electron

microscopes or by Cryo-EM. Another possibility would be to employ high resolution structural methods like X-crystallography or NMR, but obviously this is likely to require substantial advances of current methodologies.

XL-MS analysis of the p35Gab1 – Grb2 complex identified a new likely binding site between the N-terminal Grb2 SH3 domain and the p35Gab1, or Gab1 protein. The new Gab1 – Grb2SH3N interaction site in the Gab1 protein is located only a few amino acids upstream of the Gab – Grb2SH3C binding site previously characterised in a Grb2SH3C – Gab2 peptide crystal (PDB code: 2vwf, Harkiolaki et al., 2009). Both Grb2 binding sites (P-x-x-P-x-R-x_(n)-R-x-x-K) in Gab1 could function as a Grb2 SH3 domain-binding tandem motif. A sequence alignment of human Gab proteins showed that the tandem motif is specific to Gab1. However, it is well conserved throughout ca. 500 million years of evolution. The interaction between full-length Grb2 and a Gab-derived peptide containing the Grb2 binding tandem motif was examined by X-ray crystallography. The X-ray structure of the Grb2 – Gab1 peptide complex was the first full-length Grb2 structure complexed with a relatively long 32 aa Gab1 peptide. The obtained 2.41 Å resolution for the full-length Grb2 – Gab1 peptide complex was a significant improvement compared to the previously reported 3.1 Å full-length Grb2 crystal structure (unbound Grb2, 1gri; Maignan et al., 1995). The Grb2 – Gab1 peptide crystal structure demonstrated structural details of the P-x-x-P-x-R motif mediated Gab1 – Grb2SH3N interaction and the R-x-x-K motif mediated Gab1 – Grb2SH3C interaction which showed a high structural similarity to the published Gab2 – Grb2SH3C domain interaction (PDB code: 2vwf, Harkiolaki et al., 2009).

Surprisingly, the overall Grb2 – Gab1 peptide crystal package showed a Grb2 – Gab1 oligomerisation which was not seen in previous experiments. The Grb2 – Gab1 peptide complex might only oligomerise under specific conditions as for example in a Grb2 concentration-dependent manner. If existent *in vivo*, the Grb2 – Gab1 network may function to provide pre-formed Grb2 – Gab protein complexes, or individual proteins, that are easily and fast accessible for signal transduction events. However, further studies are required to confirm a Grb2 – Gab1 oligomerization model. DLS or AUC are methods which can be employed to analyse if a Gab1 peptide with the Grb2 tandem motif possibly drives Grb2 oligomerisation *in vitro*. Purified individual Grb2 SH3C and SH2 domains could be used in ITC or NMR to examine a potential Grb2 SH2 – SH3C domain binding in solution.

A similar Grb2 SH3 domain-binding tandem motif was found in other human proteins including the Grb2-interacting adaptor protein Garem1. Sequence alignments showed that the Grb2 binding tandem motif is highly conserved in Garem1 similarly as in the Gab1 protein. The independent appearance of the motif in different protein families suggests that the Grb2 binding tandem motif convergently evolved, which supports the idea of a functional role of the Grb2 binding tandem motif in cellular processes. Further functional studies of the Grb2 SH3 domain-binding tandem motif are, however,

required. One approach could be to compare cell lines, bearing either Gab1 with a single mutated Grb2 binding site (e.g. Gab1₄₉₇FRSSAKTPPARPVPVADCEPPPVDRLKPKDRK₅₂₈; ₄₉₇FRSSPKTPRRPVPVADCEPPPVDANLAPDRK₅₂₈), mutations in both Grb2 binding sites or even mutate all three Grb2 binding motifs, the Grb2 tandem binding motif and the lower affinity Grb2 binding site (Harkiolaki et al., 2009). Cellular phenotypes could be then analysed, for example, the impact on the c-Met pathway, especially the ability to recruit Grb2. Nowadays, with the CRISPR/Cas9 methodology, targeted mutations in cells or genetically engineered mouse models can be generated in a much shorter time period and with a simplified protocol compared to the conventional gene targeting.

A p35Gab1 – Grb2 – p85 complex was successfully rebuilt from purified recombinant proteins (Chapter 6). The assembled complex is a starting point for a detailed p35Gab1 – Grb2 – p85 complex analysis but definitely needs further optimization especially in terms of protein stability. Initial p35Gab1 – Grb2 – p85 SEC suggests a 1:2:1 ratio (p35Gab1 : Grb2 : homo-p85 α -dimer). However, other complex analysis by methods such as AUC and DLS are required to better define complex size and composition. For example, cross-linking mass spectrometry or X-ray crystallography could be applied to unravel structural features of the complex and to examine inter- and intramolecular complex contact sites. Due to the increased molecular size of the p35Gab1 – Grb2 – p85 complex, compared to the heterodimer p35Gab1 – Grb2 complex, the p35Gab1 – Grb2 – p85 complex might also be more suitable for imaging techniques such as EM or AFM.

In the future, even larger complexes could be assembled on the p35Gab1 or full-length Gab1 protein by adding CrkL, or the c-Met receptor cytoplasmic domain. Additional partner proteins might contribute to the overall complex stability.

References

- ABBEYQUAYE, T., RIESGO-ESCOVAR, J., RAABE, T. & THACKERAY, J. R. 2003. Evolution of Gab family adaptor proteins. *Gene*, 311, 43-50.
- ADZHUBEI, A. A., STERNBERG, M. J. & MAKAROV, A. A. 2013. Polyproline-II helix in proteins: structure and function. *J Mol Biol*, 425, 2100-32.
- AGOU, F., YE, F., GOFFINONT, S., COURTOIS, G., YAMAOKA, S., ISRAEL, A. & VERON, M. 2002. NEMO trimerizes through its coiled-coil C-terminal domain. *J Biol Chem*, 277, 17464-75.
- AHMED, Z., TIMSAH, Z., SUEN, K. M., COOK, N. P., LEE, G. R. T., LIN, C. C., GAGEA, M., MARTI, A. A. & LADBURY, J. E. 2015. Grb2 monomer-dimer equilibrium determines normal versus oncogenic function. *Nat Commun*, 6, 7354.
- AITIO, O., HELLMAN, M., KAZLAUSKAS, A., VINGADASSALOM, D. F., LEONG, J. M., SAKSELA, K. & PERMI, P. 2010. Recognition of tandem PxxP motifs as a unique Src homology 3-binding mode triggers pathogen-driven actin assembly. *Proc Natl Acad Sci U S A*, 107, 21743-8.
- AUDET, M. & BOUVIER, M. 2012. Restructuring G-protein- coupled receptor activation. *Cell*, 151, 14-23.
- AUMANN, K., LASSMANN, S., SCHOPFLIN, A., MAY, A. M., WOHRLE, F. U., ZEISER, R., WALLER, C. F., HAUSCHKE, D., WERNER, M. & BRUMMER, T. 2011. The immunohistochemical staining pattern of Gab2 correlates with distinct stages of chronic myeloid leukemia. *Hum Pathol*, 42, 719-26.
- BABU, M. M., VAN DER LEE, R., DE GROOT, N. S. & GSPONER, J. 2011. Intrinsically disordered proteins: regulation and disease. *Curr Opin Struct Biol*, 21, 432-40.
- BACH, H., MAZOR, Y., SHAKY, S., SHOHAM-LEV, A., BERDICHEVSKY, Y., GUTNICK, D. L. & BENHAR, I. 2001. Escherichia coli maltose-binding protein as a molecular chaperone for recombinant intracellular cytoplasmic single-chain antibodies. *J Mol Biol*, 312, 79-93.
- BACKER, J. M. 2010. The regulation of class IA PI 3-kinases by inter-subunit interactions. *Curr Top Microbiol Immunol*, 346, 87-114.
- BALDWIN, R. L. 1986. Temperature dependence of the hydrophobic interaction in protein folding. *Proc Natl Acad Sci U S A*, 83, 8069-72.
- BARDELLI, A., LONGATI, P., GRAMAGLIA, D., STELLA, M. C. & COMOGLIO, P. M. 1997. Gab1 coupling to the HGF/Met receptor multifunctional docking site requires binding of Grb2 and correlates with the transforming potential. *Oncogene*, 15, 3103-11.
- BATRA-SAFFERLING, R., GRANZIN, J., MODDER, S., HOFFMANN, S. & WILLBOLD, D. 2010. Structural studies of the phosphatidylinositol 3-kinase (PI3K) SH3 domain in complex with a peptide ligand: role of the anchor residue in ligand binding. *Biol Chem*, 391, 33-42.
- BELL, S., KLEIN, C., MULLER, L., HANSEN, S. & BUCHNER, J. 2002. p53 contains large unstructured regions in its native state. *J Mol Biol*, 322, 917-27.
- BENTIREN-ALJ, M., GIL, S. G., CHAN, R., WANG, Z. C., WANG, Y., IMANAKA, N., HARRIS, L. N., RICHARDSON, A., NEEL, B. G. & GU, H. 2006. A role for the scaffolding adapter GAB2 in breast cancer. *Nat Med*, 12, 114-21.
- BIER, D., BARTEL, M., SIES, K., HALBACH, S., HIGUCHI, Y., HARANOSONO, Y., BRUMMER, T., KATO, N. & OTTMANN, C. 2016. Small-Molecule Stabilization of the 14-3-3/Gab2 Protein-Protein Interaction (PPI) Interface. *ChemMedChem*, 11, 911-8.
- BILL, R. M. 2014. Playing catch-up with Escherichia coli: using yeast to increase success rates in recombinant protein production experiments. *Front Microbiol*, 5, 85.

- BIRNINGER, M. S., PEROZZO, R., KUT, E., STILLHART, C., SURBER, W., SCAPOZZA, L. & FOLKERS, G. 2006. High-level expression and purification of human thymidine kinase 1: quaternary structure, stability, and kinetics. *Protein Expr Purif*, 47, 506-15.
- BJELLQVIST, B., BASSE, B., OLSEN, E. & CELIS, J. E. 1994. Reference points for comparisons of two-dimensional maps of proteins from different human cell types defined in a pH scale where isoelectric points correlate with polypeptide compositions. *Electrophoresis*, 15, 529-39.
- BJELLQVIST, B., HUGHES, G. J., PASQUALI, C., PAQUET, N., RAVIER, F., SANCHEZ, J. C., FRUTIGER, S. & HOCHSTRASSER, D. 1993. The focusing positions of polypeptides in immobilized pH gradients can be predicted from their amino acid sequences. *Electrophoresis*, 14, 1023-31.
- BONGARTZ, H., HESSENKEMPER, W., MULLER, C., FENSKY, M., FRITSCH, J., MANDEL, K., BEHRMANN, I., HAAN, C., FISCHER, T., FELLER, S. M. & SCHAPER, F. 2017. The multi-site docking protein Gab1 is constitutively phosphorylated independent from its recruitment to the plasma membrane in Jak2-V617F-positive cells and mediates proliferation of human erythroleukaemia cells. *Cell Signal*, 35, 37-47.
- BORG, M., MITTAG, T., PAWSON, T., TYERS, M., FORMAN-KAY, J. D. & CHAN, H. S. 2007. Polyelectrostatic interactions of disordered ligands suggest a physical basis for ultrasensitivity. *Proc Natl Acad Sci U S A*, 104, 9650-5.
- BOTTARO, D. P., RUBIN, J. S., FALETTI, D. L., CHAN, A. M., KMIECIK, T. E., VANDE WOUDE, G. F. & AARONSON, S. A. 1991. Identification of the hepatocyte growth factor receptor as the c-met proto-oncogene product. *Science*, 251, 802-4.
- BREYDO, L. & UVERSKY, V. N. 2011. Role of metal ions in aggregation of intrinsically disordered proteins in neurodegenerative diseases. *Metallomics*, 3, 1163-80.
- BRUMMER, T., SCHMITZ-PEIFFER, C. & DALY, R. J. 2010. Docking proteins. *FEBS J*, 277, 4356-69.
- BUDAY, L. & TOMPA, P. 2010. Functional classification of scaffold proteins and related molecules. *FEBS J*, 277, 4348-55.
- CAMARA-ARTIGAS, A., ORTIZ-SALMERON, E., ANDUJAR-SANCHEZ, M., BACARIZO, J. & MARTIN-GARCIA, J. M. 2016. The role of water molecules in the binding of class I and II peptides to the SH3 domain of the Fyn tyrosine kinase. *Acta Crystallogr F Struct Biol Commun*, 72, 707-12.
- CHENG, J., ZHONG, Y., CHEN, S., SUN, Y., HUANG, L., KANG, Y., CHEN, B., CHEN, G., WANG, F., TIAN, Y., LIU, W., FENG, G. S. & LU, Z. 2017. Gab2 mediates hepatocellular carcinogenesis by integrating multiple signaling pathways. *FASEB J*, 31, 5530-5542.
- CHUNG, I., AKITA, R., VANDLEN, R., TOOMRE, D., SCHLESSINGER, J. & MELLMAN, I. 2010. Spatial control of EGF receptor activation by reversible dimerization on living cells. *Nature*, 464, 783-7.
- CHURION, K. A. & BONDOS, S. E. 2012. Identifying solubility-promoting buffers for intrinsically disordered proteins prior to purification. *Methods Mol Biol*, 896, 415-27.
- CUNNICK, J. M., MEI, L., DOUPNIK, C. A. & WU, J. 2001. Phosphotyrosines 627 and 659 of Gab1 constitute a bisphosphoryl tyrosine-based activation motif (BTAM) conferring binding and activation of SHP2. *J Biol Chem*, 276, 24380-7.
- DALY, R. J., GU, H., PARMAR, J., MALANEY, S., LYONS, R. J., KAIROUZ, R., HEAD, D. R., HENSHALL, S. M., NEEL, B. G. & SUTHERLAND, R. L. 2002. The docking protein Gab2 is overexpressed and estrogen regulated in human breast cancer. *Oncogene*, 21, 5175-81.

- DAM, J. & SCHUCK, P. 2004. Calculating sedimentation coefficient distributions by direct modeling of sedimentation velocity concentration profiles. *Methods Enzymol*, 384, 185-212.
- DANILKOVITCH-MIAGKOVA, A. & ZBAR, B. 2002. Dysregulation of Met receptor tyrosine kinase activity in invasive tumors. *J Clin Invest*, 109, 863-7.
- DAVEY, N. E., VAN ROEY, K., WEATHERITT, R. J., TOEDT, G., UYAR, B., ALTENBERG, B., BUDD, A., DIELLA, F., DINKEL, H. & GIBSON, T. J. 2012. Attributes of short linear motifs. *Mol Biosyst*, 8, 268-81.
- DE CASTRO, E., SIGRIST, C. J., GATTIKER, A., BULLIARD, V., LANGENDIJK-GENEVAUX, P. S., GASTEIGER, E., BAIROCH, A. & HULO, N. 2006. ScanProsite: detection of PROSITE signature matches and ProRule-associated functional and structural residues in proteins. *Nucleic Acids Res*, 34, W362-5.
- DOYLE, M. L. 1997. Characterization of binding interactions by isothermal titration calorimetry. *Curr Opin Biotechnol*, 8, 31-5.
- DRUKER, B. J., MAMON, H. J. & ROBERTS, T. M. 1989. Oncogenes, growth factors, and signal transduction. *N Engl J Med*, 321, 1383-91.
- DUNKER, A. K., CORTESE, M. S., ROMERO, P., IAKOUCHEVA, L. M. & UVERSKY, V. N. 2005. Flexible nets. The roles of intrinsic disorder in protein interaction networks. *FEBS J*, 272, 5129-48.
- DYSON, H. J. & WRIGHT, P. E. 2002. Coupling of folding and binding for unstructured proteins. *Curr Opin Struct Biol*, 12, 54-60.
- ECK, M. J., SHOELSON, S. E. & HARRISON, S. C. 1993. Recognition of a high-affinity phosphotyrosyl peptide by the Src homology-2 domain of p56lck. *Nature*, 362, 87-91.
- EDMEAD, C. E., FOX, B. C., STACE, C., KTISTAKIS, N. & WELHAM, M. J. 2006. The pleckstrin homology domain of Gab-2 is required for optimal interleukin-3 signalsome-mediated responses. *Cell Signal*, 18, 1147-55.
- EDWIN, A., GRUNDSTROM, C., WAI, S. N., OHMAN, A., STIER, G. & SAUER-ERIKSSON, A. E. 2014. Domain isolation, expression, purification and proteolytic activity of the metalloprotease PrtV from *Vibrio cholerae*. *Protein Expr Purif*, 96, 39-47.
- ELLIS, R. J. & MINTON, A. P. 2003. Cell biology: join the crowd. *Nature*, 425, 27-8.
- EULENFELD, R. & SCHAPER, F. 2009. A new mechanism for the regulation of Gab1 recruitment to the plasma membrane. *Journal of Cell Science*, 122, 574-574.
- FAN, S., MA, Y. X., GAO, M., YUAN, R. Q., MENG, Q., GOLDBERG, I. D. & ROSEN, E. M. 2001. The multisubstrate adapter Gab1 regulates hepatocyte growth factor (scatter factor)-c-Met signaling for cell survival and DNA repair. *Mol Cell Biol*, 21, 4968-84.
- FELLER, S. M. 2001. Crk family adaptors-signalling complex formation and biological roles. *Oncogene*, 20, 6348-71.
- FENG, S., CHEN, J. K., YU, H., SIMON, J. A. & SCHREIBER, S. L. 1994. Two binding orientations for peptides to the Src SH3 domain: development of a general model for SH3-ligand interactions. *Science*, 266, 1241-7.
- FERNANDEZ-LEIRO, R. & SCHERES, S. H. 2016. Unravelling biological macromolecules with cryo-electron microscopy. *Nature*, 537, 339-46.
- FERREON, A. C., FERREON, J. C., WRIGHT, P. E. & DENIZ, A. A. 2013. Modulation of allostery by protein intrinsic disorder. *Nature*, 498, 390-4.
- FINK, A. L. 1999. Chaperone-mediated protein folding. *Physiol Rev*, 79, 425-49.
- FLAUGH, S. L. & LUMB, K. J. 2001. Effects of macromolecular crowding on the intrinsically disordered proteins c-Fos and p27(Kip1). *Biomacromolecules*, 2, 538-40.

- FLEISCHER, R., HEERMANN, R., JUNG, K. & HUNKE, S. 2007. Purification, reconstitution, and characterization of the CpxRAP envelope stress system of *Escherichia coli*. *J Biol Chem*, 282, 8583-93.
- FOX, J. D., KAPUST, R. B. & WAUGH, D. S. 2001. Single amino acid substitutions on the surface of *Escherichia coli* maltose-binding protein can have a profound impact on the solubility of fusion proteins. *Protein Sci*, 10, 622-30.
- FRANCIS, N. R., WOLANIN, P. M., STOCK, J. B., DEROSIER, D. J. & THOMAS, D. R. 2004. Three-dimensional structure and organization of a receptor/signaling complex. *Proc Natl Acad Sci U S A*, 101, 17480-5.
- FURGE, K. A., ZHANG, Y. W. & VANDE WOUDE, G. F. 2000. Met receptor tyrosine kinase: enhanced signaling through adapter proteins. *Oncogene*, 19, 5582-9.
- GAO, J., WANG, Y., CAI, M., PAN, Y., XU, H., JIANG, J., JI, H. & WANG, H. 2015. Mechanistic insights into EGFR membrane clustering revealed by super-resolution imaging. *Nanoscale*, 7, 2511-9.
- GARCIA-GUZMAN, M., DOLFI, F., ZEH, K. & VUORI, K. 1999. Met-induced JNK activation is mediated by the adapter protein Crk and correlates with the Gab1 - Crk signaling complex formation. *Oncogene*, 18, 7775-86.
- GASTEIGER, E., HOOGLAND, C., GATTIKER, A., DUVAUD, S., WILKINS, M. R., APPEL, R. D. & A., B. 2005. Protein Identification and Analysis Tools on the ExPASy Server. In: WALKER, J. M. (ed.) *The Proteomics Protocols Handbook*. Humana Press.
- GAUDET, R. 2009. Divide and conquer: high resolution structural information on TRP channel fragments. *J Gen Physiol*, 133, 231-7.
- GEORGIU, G. & VALAX, P. 1996. Expression of correctly folded proteins in *Escherichia coli*. *Curr Opin Biotechnol*, 7, 190-7.
- GOLOUBINOFF, P., GATENBY, A. A. & LORIMER, G. H. 1989. GroE heat-shock proteins promote assembly of foreign prokaryotic ribulose biphosphate carboxylase oligomers in *Escherichia coli*. *Nature*, 337, 44-7.
- GONZALEZ-ESPINOSA, C., ODOM, S., OLIVERA, A., HOBSON, J. P., MARTINEZ, M. E., OLIVEIRA-DOS-SANTOS, A., BARRA, L., SPIEGEL, S., PENNINGER, J. M. & RIVERA, J. 2003. Preferential signaling and induction of allergy-promoting lymphokines upon weak stimulation of the high affinity IgE receptor on mast cells. *J Exp Med*, 197, 1453-65.
- GOUDREAU, N., CORNILLE, F., DUCHESNE, M., PARKER, F., TOCQUE, B., GARBAY, C. & ROQUES, B. P. 1994. NMR structure of the N-terminal SH3 domain of GRB2 and its complex with a proline-rich peptide from Sos. *Nat Struct Biol*, 1, 898-907.
- GRASLUND, S., NORDLUND, P., WEIGELT, J., HALLBERG, B. M., BRAY, J., GILEADI, O., KNAPP, S., OPPERMANN, U., ARROWSMITH, C., HUI, R., MING, J., DHE-PAGANON, S., PARK, H. W., SAVCHENKO, A., YEE, A., EDWARDS, A., VINCENTELLI, R., CAMBILLAU, C., KIM, R., KIM, S. H., RAO, Z., SHI, Y., TERWILLIGER, T. C., KIM, C. Y., HUNG, L. W., WALDO, G. S., PELEG, Y., ALBECK, S., UNGER, T., DYM, O., PRILUSKY, J., SUSSMAN, J. L., STEVENS, R. C., LESLEY, S. A., WILSON, I. A., JOACHIMIAK, A., COLLART, F., DEMENTIEVA, I., DONNELLY, M. I., ESCHENFELDT, W. H., KIM, Y., STOLS, L., WU, R., ZHOU, M., BURLEY, S. K., EMTAGE, J. S., SAUDER, J. M., THOMPSON, D., BAIN, K., LUZ, J., GHEYI, T., ZHANG, F., ATWELL, S., ALMO, S. C., BONANNO, J. B., FISER, A., SWAMINATHAN, S., STUDIER, F. W., CHANCE, M. R., SALI, A., ACTON, T. B., XIAO, R., ZHAO, L., MA, L. C., HUNT, J. F., TONG, L., CUNNINGHAM, K., INOUE, M., ANDERSON, S., JANJUA, H., SHASTRY, R., HO, C. K., WANG, D., WANG, H.,

- JIANG, M., MONTELIONE, G. T., STUART, D. I., OWENS, R. J., DAENKE, S., SCHUTZ, A., HEINEMANN, U., YOKOYAMA, S., BUSSOW, K. & GUNSALUS, K. C. 2008. Protein production and purification. *Nat Methods*, 5, 135-46.
- GRAZIANI, A., GRAMAGLIA, D., CANTLEY, L. C. & COMOGLIO, P. M. 1991. The tyrosine-phosphorylated hepatocyte growth factor/scatter factor receptor associates with phosphatidylinositol 3-kinase. *J Biol Chem*, 266, 22087-90.
- GU, H., BOTELHO, R. J., YU, M., GRINSTEIN, S. & NEEL, B. G. 2003. Critical role for scaffolding adapter Gab2 in Fc gamma R-mediated phagocytosis. *J Cell Biol*, 161, 1151-61.
- GU, H. & NEEL, B. G. 2003. The 'Gab' in signal transduction. *Trends in Cell Biology*, 13, 122-130.
- GU, H., PRATT, J. C., BURAKOFF, S. J. & NEEL, B. G. 1998. Cloning of p97/Gab2, the major SHP2-binding protein in hematopoietic cells, reveals a novel pathway for cytokine-induced gene activation. *Mol Cell*, 2, 729-40.
- GU, H., SAITO, K., KLAMAN, L. D., SHEN, J., FLEMING, T., WANG, Y., PRATT, J. C., LIN, G., LIM, B., KINET, J. P. & NEEL, B. G. 2001. Essential role for Gab2 in the allergic response. *Nature*, 412, 186-90.
- GURUPRASAD, L., DHANARAJ, V., TIMM, D., BLUNDELL, T. L., GOUT, I. & WATERFIELD, M. D. 1995. The crystal structure of the N-terminal SH3 domain of Grb2. *J Mol Biol*, 248, 856-66.
- HAACKE, A., FENDRICH, G., RAMAGE, P. & GEISER, M. 2009. Chaperone over-expression in Escherichia coli: apparent increased yields of soluble recombinant protein kinases are due mainly to soluble aggregates. *Protein Expr Purif*, 64, 185-93.
- HABCHI, J., TOMPA, P., LONGHI, S. & UVERSKY, V. N. 2014. Introducing protein intrinsic disorder. *Chem Rev*, 114, 6561-88.
- HANAHAN, D. 1985. Techniques for transformation of E. coli. In: GLOVER, D. M. (ed.) *DNA cloning: a practical approach*.: Oxford, United Kingdom: IRL Press.
- HARKIOLAKI, M., LEWITZKY, M., GILBERT, R. J., JONES, E. Y., BOURETTE, R. P., MOUCHIROUD, G., SONDERMANN, H., MOAREFI, I. & FELLER, S. M. 2003. Structural basis for SH3 domain-mediated high-affinity binding between Mona/Gads and SLP-76. *EMBO J*, 22, 2571-82.
- HARKIOLAKI, M., TSIRKA, T., LEWITZKY, M., SIMISTER, P. C., JOSHI, D., BIRD, L. E., JONES, E. Y., O'REILLY, N. & FELLER, S. M. 2009. Distinct binding modes of two epitopes in Gab2 that interact with the SH3C domain of Grb2. *Structure*, 17, 809-22.
- HAYNES, C., OLDFIELD, C. J., JI, F., KLITGORD, N., CUSICK, M. E., RADIVOJAC, P., UVERSKY, V. N., VIDAL, M. & IAKOUCHEVA, L. M. 2006. Intrinsic disorder is a common feature of hub proteins from four eukaryotic interactomes. *PLoS Comput Biol*, 2, e100.
- HERBST, R., CARROLL, P. M., ALLARD, J. D., SCHILLING, J., RAABE, T. & SIMON, M. A. 1996. Daughter of sevenless is a substrate of the phosphotyrosine phosphatase Corkscrew and functions during sevenless signaling. *Cell*, 85, 899-909.
- HESTERKAMP, T. & BUKAU, B. 1998. Role of the DnaK and HscA homologs of Hsp70 chaperones in protein folding in E.coli. *EMBO J*, 17, 4818-28.
- HIGO, K., IKURA, T., ODA, M., MORII, H., TAKAHASHI, J., ABE, R. & ITO, N. 2013. High resolution crystal structure of the Grb2 SH2 domain with a phosphopeptide derived from CD28. *PLoS One*, 8, e74482.
- HILLE, B. 2001. *Ion channels of excitable membranes*, Sinauer Associates Inc.
- HILSER, V. J. 2013. Structural biology: Signalling from disordered proteins. *Nature*, 498, 308-10.

- HILSER, V. J., WRABL, J. O. & MOTLAGH, H. N. 2012. Structural and energetic basis of allostery. *Annu Rev Biophys*, 41, 585-609.
- HOLGADO-MADRUGA, M., EMLET, D. R., MOSCATELLO, D. K., GODWIN, A. K. & WONG, A. J. 1996. A Grb2-associated docking protein in EGF- and insulin-receptor signalling. *Nature*, 379, 560-4.
- HOLGADO-MADRUGA, M., MOSCATELLO, D. K., EMLET, D. R., DIETERICH, R. & WONG, A. J. 1997. Grb2-associated binder-1 mediates phosphatidylinositol 3-kinase activation and the promotion of cell survival by nerve growth factor. *Proc Natl Acad Sci U S A*, 94, 12419-24.
- HORDIJK, P. L., TEN KLOOSTER, J. P., VAN DER KAMMEN, R. A., MICHIELS, F., OOMEN, L. C. & COLLARD, J. G. 1997. Inhibition of invasion of epithelial cells by Tiam1-Rac signaling. *Science*, 278, 1464-6.
- HORST, B., GRUVBERGER-SAAL, S. K., HOPKINS, B. D., BORDONE, L., YANG, Y., CHERNOFF, K. A., UZOMA, I., SCHWIPPER, V., LIEBAU, J., NOWAK, N. J., BRUNNER, G., OWENS, D., RIMM, D. L., PARSONS, R. & CELEBI, J. T. 2009. Gab2-mediated signaling promotes melanoma metastasis. *Am J Pathol*, 174, 1524-33.
- HUANG, W. Y., YAN, Q., LIN, W. C., CHUNG, J. K., HANSEN, S. D., CHRISTENSEN, S. M., TU, H. L., KURIYAN, J. & GROVES, J. T. 2016. Phosphotyrosine-mediated LAT assembly on membranes drives kinetic bifurcation in recruitment dynamics of the Ras activator SOS. *Proc Natl Acad Sci U S A*, 113, 8218-23.
- ITOH, M., YOSHIDA, Y., NISHIDA, K., NARIMATSU, M., HIBI, M. & HIRANO, T. 2000. Role of Gab1 in heart, placenta, and skin development and growth factor- and cytokine-induced extracellular signal-regulated kinase mitogen-activated protein kinase activation. *Mol Cell Biol*, 20, 3695-704.
- JENSEN, M. R., RUIGROK, R. W. & BLACKLEDGE, M. 2013. Describing intrinsically disordered proteins at atomic resolution by NMR. *Curr Opin Struct Biol*, 23, 426-35.
- JI, X., ZHANG, P., ARMSTRONG, R. N. & GILLILAND, G. L. 1992. The three-dimensional structure of a glutathione S-transferase from the mu gene class. Structural analysis of the binary complex of isoenzyme 3-3 and glutathione at 2.2-Å resolution. *Biochemistry*, 31, 10169-84.
- KALLIN, A., DEMOULIN, J. B., NISHIDA, K., HIRANO, T., RONNSTRAND, L. & HELDIN, C. H. 2004. Gab1 contributes to cytoskeletal reorganization and chemotaxis in response to platelet-derived growth factor. *J Biol Chem*, 279, 17897-904.
- KAMEDA, H., RISINGER, J. I., HAN, B. B., BAEK, S. J., BARRETT, J. C., ABE, T., TAKEUCHI, T., GLASGOW, W. C. & ELING, T. E. 2001. Expression of Gab1 lacking the pleckstrin homology domain is associated with neoplastic progression. *Mol Cell Biol*, 21, 6895-905.
- KANEKO, T., JOSHI, R., FELLER, S. M. & LI, S. S. 2012. Phosphotyrosine recognition domains: the typical, the atypical and the versatile. *Cell Commun Signal*, 10, 32.
- KANEKO, T., KUMASAKA, T., GANBE, T., SATO, T., MIYAZAWA, K., KITAMURA, N. & TANAKA, N. 2003. Structural insight into modest binding of a non-PXXP ligand to the signal transducing adaptor molecule-2 Src homology 3 domain. *J Biol Chem*, 278, 48162-8.
- KARKKAINEN, S., HIIPAKKA, M., WANG, J. H., KLEINO, I., VAHA-JAAKKOLA, M., RENKEMA, G. H., LISS, M., WAGNER, R. & SAKSELA, K. 2006. Identification of preferred protein interactions by phage-display of the human Src homology-3 proteome. *EMBO Rep*, 7, 186-91.
- KIMPLE, M. E., BRILL, A. L. & PASKER, R. L. 2013. Overview of affinity tags for protein purification. *Curr Protoc Protein Sci*, 73, Unit 9 9.

- KONRAT, R. 2014. NMR contributions to structural dynamics studies of intrinsically disordered proteins. *J Magn Reson*, 241, 74-85.
- KORTUM, R. L., BALAGOPALAN, L., ALEXANDER, C. P., GARCIA, J., PINSKI, J. M., MERRILL, R. K., NGUYEN, P. H., LI, W., AGARWAL, I., AKPAN, I. O., SOMMERS, C. L. & SAMELSON, L. E. 2013. The ability of Sos1 to oligomerize the adaptor protein LAT is separable from its guanine nucleotide exchange activity in vivo. *Sci Signal*, 6, ra99.
- LADBURY, J. E. 2010. Calorimetry as a tool for understanding biomolecular interactions and an aid to drug design. *Biochem Soc Trans*, 38, 888-93.
- LAEMMLI, U. K. 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature*, 227, 680-5.
- LE GOFF, A., JI, Z., LECLERCQ, B., BOURETTE, R. P., MOUGEL, A., GUERARDEL, C., DE LAUNOIT, Y., VICOONE, J., GOORMACHTIGH, G. & FAFEUR, V. 2012. Anti-apoptotic role of caspase-cleaved GAB1 adaptor protein in hepatocyte growth factor/scatter factor-MET receptor protein signaling. *J Biol Chem*, 287, 35382-96.
- LEBENDIKER, M. & DANIELI, T. 2014. Production of prone-to-aggregate proteins. *FEBS Lett*, 588, 236-46.
- LEBOWITZ, J., LEWIS, M. S. & SCHUCK, P. 2002. Modern analytical ultracentrifugation in protein science: a tutorial review. *Protein Sci*, 11, 2067-79.
- LEE, C. W., FERREON, J. C., FERREON, A. C., ARAI, M. & WRIGHT, P. E. 2010. Graded enhancement of p53 binding to CREB-binding protein (CBP) by multisite phosphorylation. *Proc Natl Acad Sci U S A*, 107, 19290-5.
- LEHR, S., KOTZKA, J., HERKNER, A., SIKMANN, A., MEYER, H. E., KRONE, W. & MULLER-WIELAND, D. 2000. Identification of major tyrosine phosphorylation sites in the human insulin receptor substrate Gab-1 by insulin receptor kinase in vitro. *Biochemistry*, 39, 10898-907.
- LEITNER, A., FAINI, M., STENGEL, F. & AEBERSOLD, R. 2016. Crosslinking and Mass Spectrometry: An Integrated Technology to Understand the Structure and Function of Molecular Machines. *Trends Biochem Sci*, 41, 20-32.
- LEMMON, M. A. & SCHLESSINGER, J. 2010. Cell signaling by receptor tyrosine kinases. *Cell*, 141, 1117-34.
- LEWITZKY, M., HARKIOLAKI, M., DOMART, M. C., JONES, E. Y. & FELLER, S. M. 2004. Mona/Gads SH3C binding to hematopoietic progenitor kinase 1 (HPK1) combines an atypical SH3 binding motif, R/KXXXK, with a classical PXXP motif embedded in a polyproline type II (PPII) helix. *J Biol Chem*, 279, 28724-32.
- LEWITZKY, M., KARDINAL, C., GEHRING, N. H., SCHMIDT, E. K., KONKOL, B., EULITZ, M., BIRCHMEIER, W., SCHAEFER, U. & FELLER, S. M. 2001. The C-terminal SH3 domain of the adapter protein Grb2 binds with high affinity to sequences in Gab1 and SLP-76 which lack the SH3-typical P-x-x-P core motif. *Oncogene*, 20, 1052-62.
- LEWITZKY, M., SIMISTER, P. C. & FELLER, S. M. 2012. Beyond 'furballs' and 'dumpling soups' - towards a molecular architecture of signaling complexes and networks. *FEBS Lett*, 586, 2740-50.
- LI, S. S. 2005. Specificity and versatility of SH3 and other proline-recognition domains: structural basis and implications for cellular signal transduction. *Biochem J*, 390, 641-53.
- LIM, W. A., RICHARDS, F. M. & FOX, R. O. 1994. Structural determinants of peptide-binding orientation and of sequence specificity in SH3 domains. *Nature*, 372, 375-9.

- LIU, Y., JENKINS, B., SHIN, J. L. & ROHRSCHEIDER, L. R. 2001. Scaffolding protein Gab2 mediates differentiation signaling downstream of Fms receptor tyrosine kinase. *Mol Cell Biol*, 21, 3047-56.
- LIU, Y. & ROHRSCHEIDER, L. R. 2002. The gift of Gab. *FEBS Lett*, 515, 1-7.
- LOCK, L. S., ROYAL, I., NAUJOKAS, M. A. & PARK, M. 2000. Identification of an atypical Grb2 carboxyl-terminal SH3 domain binding site in Gab docking proteins reveals Grb2-dependent and -independent recruitment of Gab1 to receptor tyrosine kinases. *J Biol Chem*, 275, 31536-45.
- LOPICCOLO, J., KIM, S. J., SHI, Y., WU, B., WU, H., CHAIT, B. T., SINGER, R. H., SALI, A., BRENOWITZ, M., BRESNICK, A. R. & BACKER, J. M. 2015. Assembly and Molecular Architecture of the Phosphoinositide 3-Kinase p85alpha Homodimer. *J Biol Chem*, 290, 30390-405.
- LOVELL, S. C., DAVIS, I. W., ARENDALL, W. B., 3RD, DE BAKKER, P. I., WORD, J. M., PRISANT, M. G., RICHARDSON, J. S. & RICHARDSON, D. C. 2003. Structure validation by C α geometry: phi,psi and Cbeta deviation. *Proteins*, 50, 437-50.
- LUBY-PHELPS, K. 2000. Cytoarchitecture and physical properties of cytoplasm: volume, viscosity, diffusion, intracellular surface area. *Int Rev Cytol*, 192, 189-221.
- MAIGNAN, S., GUILLOTEAU, J. P., FROMAGE, N., ARNOUX, B., BECQUART, J. & DUCRUIX, A. 1995. Crystal structure of the mammalian Grb2 adaptor. *Science*, 268, 291-3.
- MAROUN, C. R., HOLGADO-MADRUGA, M., ROYAL, I., NAUJOKAS, M. A., FOURNIER, T. M., WONG, A. J. & PARK, M. 1999. The Gab1 PH domain is required for localization of Gab1 at sites of cell-cell contact and epithelial morphogenesis downstream from the met receptor tyrosine kinase. *Mol Cell Biol*, 19, 1784-99.
- MCDONALD, C. B., SELDEEN, K. L., DEEGAN, B. J. & FAROOQ, A. 2008a. Structural basis of the differential binding of the SH3 domains of Grb2 adaptor to the guanine nucleotide exchange factor Sos1. *Arch Biochem Biophys*, 479, 52-62.
- MCDONALD, C. B., SELDEEN, K. L., DEEGAN, B. J., LEWIS, M. S. & FAROOQ, A. 2008b. Grb2 adaptor undergoes conformational change upon dimerization. *Arch Biochem Biophys*, 475, 25-35.
- MCLACHLIN, D. T. & CHAIT, B. T. 2001. Analysis of phosphorylated proteins and peptides by mass spectrometry. *Curr Opin Chem Biol*, 5, 591-602.
- MCNEMAR, C., SNOW, M. E., WINDSOR, W. T., PRONGAY, A., MUI, P., ZHANG, R., DURKIN, J., LE, H. V. & WEBER, P. C. 1997. Thermodynamic and structural analysis of phosphotyrosine polypeptide binding to Grb2-SH2. *Biochemistry*, 36, 10006-14.
- MCPHERSON, A. 2001. A comparison of salts for the crystallization of macromolecules. *Protein Sci*, 10, 418-22.
- MCPHERSON, A., JR. 1976. Crystallization of proteins from polyethylene glycol. *J Biol Chem*, 251, 6300-3.
- MERK, A., BARTESAGHI, A., BANERJEE, S., FALCONIERI, V., RAO, P., DAVIS, M. I., PRAGANI, R., BOXER, M. B., EARL, L. A., MILNE, J. L. S. & SUBRAMANIAM, S. 2016. Breaking Cryo-EM Resolution Barriers to Facilitate Drug Discovery. *Cell*, 165, 1698-1707.
- MITREA, D. M. & KRIWACKI, R. W. 2016. Phase separation in biology; functional organization of a higher order. *Cell Commun Signal*, 14, 1.
- MONTESANO, R., MATSUMOTO, K., NAKAMURA, T. & ORCI, L. 1991. Identification of a fibroblast-derived epithelial morphogen as hepatocyte growth factor. *Cell*, 67, 901-8.

- MUSACCHIO, A., NOBLE, M., PAUPTIT, R., WIERENGA, R. & SARASTE, M. 1992. Crystal structure of a Src-homology 3 (SH3) domain. *Nature*, 359, 851-5.
- MUSACCHIO, A., SARASTE, M. & WILMANN, M. 1994. High-resolution crystal structures of tyrosine kinase SH3 domains complexed with proline-rich peptides. *Nat Struct Biol*, 1, 546-51.
- NIOCHE, P., LIU, W. Q., BROUTIN, I., CHARBONNIER, F., LATREILLE, M. T., VIDAL, M., ROQUES, B., GARBAY, C. & DUCRUIX, A. 2002. Crystal structures of the SH2 domain of Grb2: highlight on the binding of a new high-affinity inhibitor. *J Mol Biol*, 315, 1167-77.
- NISHIDA, K., WANG, L., MORII, E., PARK, S. J., NARIMATSU, M., ITOH, S., YAMASAKI, S., FUJISHIMA, M., ISHIHARA, K., HIBI, M., KITAMURA, Y. & HIRANO, T. 2002. Requirement of Gab2 for mast cell development and KitL/c-Kit signaling. *Blood*, 99, 1866-9.
- NISHIDA, K., YOSHIDA, Y., ITOH, M., FUKADA, T., OHTANI, T., SHIROGANE, T., ATSUMI, T., TAKAHASHI-TEZUKA, M., ISHIHARA, K., HIBI, M. & HIRANO, T. 1999. Gab-family adapter proteins act downstream of cytokine and growth factor receptors and T- and B-cell antigen receptors. *Blood*, 93, 1809-16.
- NISHIDA, M., NAGATA, K., HACHIMORI, Y., HORIUCHI, M., OGURA, K., MANDIYAN, V., SCHLESSINGER, J. & INAGAKI, F. 2001. Novel recognition mode between Vav and Grb2 SH3 domains. *EMBO J*, 20, 2995-3007.
- OGURA, K., TSUCHIYA, S., TERASAWA, H., YUZAWA, S., HATANAKA, H., MANDIYAN, V., SCHLESSINGER, J. & INAGAKI, F. 1999. Solution structure of the SH2 domain of Grb2 complexed with the Shc-derived phosphotyrosine-containing peptide. *J Mol Biol*, 289, 439-45.
- OLDFIELD, C. J. & DUNKER, A. K. 2014. Intrinsically disordered proteins and intrinsically disordered protein regions. *Annu Rev Biochem*, 83, 553-84.
- ONG, S. H., HADARI, Y. R., GOTOH, N., GUY, G. R., SCHLESSINGER, J. & LAX, I. 2001. Stimulation of phosphatidylinositol 3-kinase by fibroblast growth factor receptors is mediated by coordinated recruitment of multiple docking proteins. *Proc Natl Acad Sci U S A*, 98, 6074-9.
- PEPAJ, M., WILSON, S. R., NOVOTNA, K., LUNDANES, E. & GREIBROKK, T. 2006. Two-dimensional capillary liquid chromatography: pH gradient ion exchange and reversed phase chromatography for rapid separation of proteins. *J Chromatogr A*, 1120, 132-41.
- PERKINS, J. R., DIBOUN, I., DESSAILLY, B. H., LEES, J. G. & ORENGO, C. 2010. Transient protein-protein interactions: structural, functional, and network properties. *Structure*, 18, 1233-43.
- PIETREK, M., BRINKMANN, M. M., GLOWACKA, I., ENLUND, A., HAVEMEIER, A., DITTRICH-BREIHOLZ, O., KRACHT, M., LEWITZKY, M., SAKSELA, K., FELLER, S. M. & SCHULZ, T. F. 2010. Role of the Kaposi's sarcoma-associated herpesvirus K15 SH3 binding site in inflammatory signaling and B-cell activation. *J Virol*, 84, 8231-40.
- PRICE, L. S., HAJDO-MILASINOVIC, A., ZHAO, J., ZWARTKRUIS, F. J., COLLARD, J. G. & BOS, J. L. 2004. Rap1 regulates E-cadherin-mediated cell-cell adhesion. *J Biol Chem*, 279, 35127-32.
- QIAO, Q., YANG, C., ZHENG, C., FONTAN, L., DAVID, L., YU, X., BRACKEN, C., ROSEN, M., MELNICK, A., EGELMAN, E. H. & WU, H. 2013. Structural architecture of the CARMA1/Bcl10/MALT1 signalosome: nucleation-induced filamentous assembly. *Mol Cell*, 51, 766-79.
- QU, Y. & BOLEN, D. W. 2002. Efficacy of macromolecular crowding in forcing proteins to fold. *Biophys Chem*, 101-102, 155-65.

- RAABE, T., RIESGO-ESCOVAR, J., LIU, X., BAUSENWEIN, B. S., DEAK, P., MAROY, P. & HAFEN, E. 1996. DOS, a novel pleckstrin homology domain-containing protein required for signal transduction between sevenless and Ras1 in *Drosophila*. *Cell*, 85, 911-20.
- RAHUEL, J., GAY, B., ERDMANN, D., STRAUSS, A., GARCIA-ECHEVERRIA, C., FURET, P., CARAVATTI, G., FRETZ, H., SCHOEPFER, J. & GRUTTER, M. G. 1996. Structural basis for specificity of Grb2-SH2 revealed by a novel ligand binding mode. *Nat Struct Biol*, 3, 586-9.
- RAMEH, L. E., ARVIDSSON, A., CARRAWAY, K. L., 3RD, COUVILLON, A. D., RATHBUN, G., CROMPTON, A., VANRENTERGHEM, B., CZECH, M. P., RAVICHANDRAN, K. S., BURAKOFF, S. J., WANG, D. S., CHEN, C. S. & CANTLEY, L. C. 1997. A comparative analysis of the phosphoinositide binding specificity of pleckstrin homology domains. *J Biol Chem*, 272, 22059-66.
- RAPPSILBER, J. 2011. The beginning of a beautiful friendship: cross-linking/mass spectrometry and modelling of proteins and multi-protein complexes. *J Struct Biol*, 173, 530-40.
- RIAL, D. V. & CECCARELLI, E. A. 2002. Removal of DnaK contamination during fusion protein purifications. *Protein Expr Purif*, 25, 503-7.
- ROCCHI, S., TARTARE-DECKERT, S., MURDACA, J., HOLGADO-MADRUGA, M., WONG, A. J. & VAN OBBERGHEN, E. 1998. Determination of Gab1 (Grb2-associated binder-1) interaction with insulin receptor-signaling molecules. *Mol Endocrinol*, 12, 914-23.
- RODRIGUES, G. A., FALASCA, M., ZHANG, Z., ONG, S. H. & SCHLESSINGER, J. 2000. A novel positive feedback loop mediated by the docking protein Gab1 and phosphatidylinositol 3-kinase in epidermal growth factor receptor signaling. *Mol Cell Biol*, 20, 1448-59.
- ROMIER, C., BEN JELLOUL, M., ALBECK, S., BUCHWALD, G., BUSSO, D., CELIE, P. H., CHRISTODOULOU, E., DE MARCO, V., VAN GERWEN, S., KNIPSCHER, P., LEBBINK, J. H., NOTENBOOM, V., POTERSZMAN, A., ROCHEL, N., COHEN, S. X., UNGER, T., SUSSMAN, J. L., MORAS, D., SIXMA, T. K. & PERRAKIS, A. 2006. Co-expression of protein complexes in prokaryotic and eukaryotic hosts: experimental procedures, database tracking and case studies. *Acta Crystallogr D Biol Crystallogr*, 62, 1232-42.
- ROQUE, A., PONTE, I. & SUAUI, P. 2007. Macromolecular crowding induces a molten globule state in the C-terminal domain of histone H1. *Biophys J*, 93, 2170-7.
- SACHS, M., BROHMANN, H., ZECHNER, D., MULLER, T., HULSKEN, J., WALTHER, I., SCHAEFER, U., BIRCHMEIER, C. & BIRCHMEIER, W. 2000. Essential role of Gab1 for signaling by the c-Met receptor in vivo. *J Cell Biol*, 150, 1375-84.
- SATTLER, M., MOHI, M. G., PRIDE, Y. B., QUINNAN, L. R., MALOUF, N. A., PODAR, K., GESBERT, F., IWASAKI, H., LI, S., VAN ETEN, R. A., GU, H., GRIFFIN, J. D. & NEEL, B. G. 2002. Critical role for Gab2 in transformation by BCR/ABL. *Cancer Cell*, 1, 479-92.
- SCHAEFER, U., GEHRING, N. H., FUCHS, K. P., SACHS, M., KEMPKES, B. & BIRCHMEIER, W. 2000. Coupling of Gab1 to c-Met, Grb2, and Shp2 mediates biological responses. *J Cell Biol*, 149, 1419-32.
- SCHAEFER, U., VOGEL, R., CHMIELOWIEC, J., HUELSKEN, J., ROSARIO, M. & BIRCHMEIER, W. 2007. Distinct requirements for Gab1 in Met and EGF receptor signaling in vivo. *Proc Natl Acad Sci U S A*, 104, 15376-81.
- SCHEFFZEK, K. & WELTI, S. 2012. Pleckstrin homology (PH) like domains - versatile modules in protein-protein interaction platforms. *FEBS Lett*, 586, 2662-73.

- SCHELLMAN, J. A. 1997. Temperature, stability, and the hydrophobic interaction. *Biophys J*, 73, 2960-4.
- SCHLIEKER, C., BUKAU, B. & MOGK, A. 2002. Prevention and reversion of protein aggregation by molecular chaperones in the E. coli cytosol: implications for their applicability in biotechnology. *J Biotechnol*, 96, 13-21.
- SCHMIDT, C., BLADT, F., GOEDECKE, S., BRINKMANN, V., ZSCHIESCHE, W., SHARPE, M., GHERARDI, E. & BIRCHMEIER, C. 1995. Scatter factor/hepatocyte growth factor is essential for liver development. *Nature*, 373, 699-702.
- SCHUTZMAN, J. L., BORLAND, C. Z., NEWMAN, J. C., ROBINSON, M. K., KOKEL, M. & STERN, M. J. 2001. The Caenorhabditis elegans EGL-15 signaling pathway implicates a DOS-like multisubstrate adaptor protein in fibroblast growth factor signal transduction. *Mol Cell Biol*, 21, 8104-16.
- SEIDEN-LONG, I., NAVAB, R., SHIH, W., LI, M., CHOW, J., ZHU, C. Q., RADULOVICH, N., SAUCIER, C. & TSAO, M. S. 2008. Gab1 but not Grb2 mediates tumor progression in Met overexpressing colorectal cancer cells. *Carcinogenesis*, 29, 647-55.
- SEIFFERT, M., CUSTODIO, J. M., WOLF, I., HARKEY, M., LIU, Y., BLATTMAN, J. N., GREENBERG, P. D. & ROHRSCHEIDER, L. R. 2003. Gab3-deficient mice exhibit normal development and hematopoiesis and are immunocompetent. *Mol Cell Biol*, 23, 2415-24.
- SHAW, A. S. & FILBERT, E. L. 2009. Scaffold proteins and immune-cell signalling. *Nat Rev Immunol*, 9, 47-56.
- SHIBER, A., DORING, K., FRIEDRICH, U., KLANN, K., MERKER, D., ZEDAN, M., TIPPMANN, F., KRAMER, G. & BUKAU, B. 2018. Cotranslational assembly of protein complexes in eukaryotes revealed by ribosome profiling. *Nature*, 561, 268-272.
- SIMISTER, P. C. & FELLER, S. M. 2012. Order and disorder in large multi-site docking proteins of the Gab family--implications for signalling complex formation and inhibitor design strategies. *Mol Biosyst*, 8, 33-46.
- SIMISTER, P. C., SCHAPER, F., O'REILLY, N., MCGOWAN, S. & FELLER, S. M. 2011. Self-organization and regulation of intrinsically disordered proteins with folded N-termini. *PLoS Biol*, 9, e1000591.
- SIROHI, D., CHEN, Z., SUN, L., KLOSE, T., PIERSON, T. C., ROSSMANN, M. G. & KUHN, R. J. 2016. The 3.8 Å resolution cryo-EM structure of Zika virus. *Science*, 352, 467-70.
- SMITH, C. A. 1988. Estimation of sedimentation coefficients and frictional ratios of globular proteins. *Biochemical Education*, 16, 104-106.
- SORENSEN, H. P. & MORTENSEN, K. K. 2005. Soluble expression of recombinant proteins in the cytoplasm of Escherichia coli. *Microb Cell Fact*, 4, 1.
- SPARKS, A. B., RIDER, J. E., HOFFMAN, N. G., FOWLKES, D. M., QUILLAM, L. A. & KAY, B. K. 1996. Distinct ligand preferences of Src homology 3 domains from Src, Yes, Abl, Cortactin, p53bp2, PLCgamma, Crk, and Grb2. *Proc Natl Acad Sci U S A*, 93, 1540-4.
- STAUBER, D. J., DEBLER, E. W., HORTON, P. A., SMITH, K. A. & WILSON, I. A. 2006. Crystal structure of the IL-2 signaling complex: paradigm for a heterotrimeric cytokine receptor. *Proc Natl Acad Sci U S A*, 103, 2788-93.
- STOKER, M., GHERARDI, E., PERRYMAN, M. & GRAY, J. 1987. Scatter factor is a fibroblast-derived modulator of epithelial cell mobility. *Nature*, 327, 239-42.
- STOKER, M. & PERRYMAN, M. 1985. An epithelial scatter factor released by embryo fibroblasts. *J Cell Sci*, 77, 209-23.

- SUGASE, K., DYSON, H. J. & WRIGHT, P. E. 2007. Mechanism of coupled folding and binding of an intrinsically disordered protein. *Nature*, 447, 1021-5.
- SUTHERLAND, L., RUHE, M., GATTEGNO-HO, D., MANN, K., GREAVES, J., KOSCIELNIAK, M., MEEK, S., LU, Z., WATERFALL, M., TAYLOR, R., TSAKIRIDIS, A., BROWN, H., MACIVER, S. K., JOSHI, A., CLINTON, M., CHAMBERLAIN, L. H., SMITH, A. & BURDON, T. 2018. LIF-dependent survival of embryonic stem cells is regulated by a novel palmitoylated Gab1 signalling protein. *J Cell Sci*, 131.
- SZASZ, C. S., ALEXA, A., TOTH, K., RAKACS, M., LANGOWSKI, J. & TOMPA, P. 2011. Protein disorder prevails under crowded conditions. *Biochemistry*, 50, 5834-44.
- TASHIRO, K., TSUNEMATSU, T., OKUBO, H., OHTA, T., SANO, E., YAMAUCHI, E., TANIGUCHI, H. & KONISHI, H. 2009. GAREM, a novel adaptor protein for growth factor receptor-bound protein 2, contributes to cellular transformation through the activation of extracellular signal-regulated kinase signaling. *J Biol Chem*, 284, 20206-14.
- TERASAWA, H., KOHDA, D., HATANAKA, H., TSUCHIYA, S., OGURA, K., NAGATA, K., ISHII, S., MANDIYAN, V., ULLRICH, A., SCHLESSINGER, J. & ET AL. 1994. Structure of the N-terminal SH3 domain of GRB2 complexed with a peptide from the guanine nucleotide releasing factor Sos. *Nat Struct Biol*, 1, 891-7.
- THEILLET, F. X., BINOLFI, A., FREMBGEN-KESNER, T., HINGORANI, K., SARKAR, M., KYNE, C., LI, C., CROWLEY, P. B., GIERASCH, L., PIELAK, G. J., ELCOCK, A. H., GERSHENSON, A. & SELENKO, P. 2014. Physicochemical properties of cells and their effects on intrinsically disordered proteins (IDPs). *Chem Rev*, 114, 6661-714.
- THORNTON, K. H., MUELLER, W. T., MCCONNELL, P., ZHU, G., SALTIEL, A. R. & THANABAL, V. 1996. Nuclear magnetic resonance solution structure of the growth factor receptor-bound protein 2 Src homology 2 domain. *Biochemistry*, 35, 11852-64.
- TOMPA, P. & FUXREITER, M. 2008. Fuzzy complexes: polymorphism and structural disorder in protein-protein interactions. *Trends Biochem Sci*, 33, 2-8.
- TONIOLO, C. & BENEDETTI, E. 1991. The polypeptide 310-helix. *Trends Biochem Sci*, 16, 350-3.
- TRUSOLINO, L., BERTOTTI, A. & COMOGLIO, P. M. 2010. MET signalling: principles and functions in development, organ regeneration and cancer. *Nat Rev Mol Cell Biol*, 11, 834-48.
- TYSON, J. J., CHEN, K. C. & NOVAK, B. 2003. Sniffers, buzzers, toggles and blinkers: dynamics of regulatory and signaling pathways in the cell. *Curr Opin Cell Biol*, 15, 221-31.
- UEHARA, Y., MINOWA, O., MORI, C., SHIOTA, K., KUNO, J., NODA, T. & KITAMURA, N. 1995. Placental defect and embryonic lethality in mice lacking hepatocyte growth factor/scatter factor. *Nature*, 373, 702-5.
- UNDERHILL, D. M. & GOODRIDGE, H. S. 2007. The many faces of ITAMs. *Trends Immunol*, 28, 66-73.
- UVERSKY, V. N., OLDFIELD, C. J. & DUNKER, A. K. 2008. Intrinsically disordered proteins in human diseases: introducing the D2 concept. *Annu Rev Biophys*, 37, 215-46.
- VACIC, V., MARKWICK, P. R., OLDFIELD, C. J., ZHAO, X., HAYNES, C., UVERSKY, V. N. & IAKOUCHEVA, L. M. 2012. Disease-associated mutations disrupt functionally important regions of intrinsic protein disorder. *PLoS Comput Biol*, 8, e1002709.

- VANHAESEBROECK, B., STEPHENS, L. & HAWKINS, P. 2012. PI3K signalling: the path to discovery and understanding. *Nat Rev Mol Cell Biol*, 13, 195-203.
- VENTURA, S. & VILLAVARDE, A. 2006. Protein quality in bacterial inclusion bodies. *Trends Biotechnol*, 24, 179-85.
- WADA, T., NAKASHIMA, T., OLIVEIRA-DOS-SANTOS, A. J., GASSER, J., HARA, H., SCHETT, G. & PENNINGER, J. M. 2005. The molecular scaffold Gab2 is a crucial component of RANK signaling and osteoclastogenesis. *Nat Med*, 11, 394-9.
- WALZ, J., ERDMANN, A., KANIA, M., TYPKE, D., KOSTER, A. J. & BAUMEISTER, W. 1998. 26S proteasome structure revealed by three-dimensional electron microscopy. *J Struct Biol*, 121, 19-29.
- WARD, J. J., SODHI, J. S., MCGUFFIN, L. J., BUXTON, B. F. & JONES, D. T. 2004. Prediction and functional analysis of native disorder in proteins from the three kingdoms of life. *J Mol Biol*, 337, 635-45.
- WEIDNER, K. M., DI CESARE, S., SACHS, M., BRINKMANN, V., BEHRENS, J. & BIRCHMEIER, W. 1996. Interaction between Gab1 and the c-Met receptor tyrosine kinase is responsible for epithelial morphogenesis. *Nature*, 384, 173-6.
- WILKINS, M. R., GASTEIGER, E., BAIROCH, A., SANCHEZ, J. C., WILLIAMS, K. L., APPEL, R. D. & HOCHSTRASSER, D. F. 1999. Protein identification and analysis tools in the ExPASy server. *Methods Mol Biol*, 112, 531-52.
- WITTEKIND, M., MAPELLI, C., FARMER, B. T., 2ND, SUEN, K. L., GOLDFARB, V., TSAO, J., LAVOIE, T., BARBACID, M., MEYERS, C. A. & MUELLER, L. 1994. Orientation of peptide fragments from Sos proteins bound to the N-terminal SH3 domain of Grb2 determined by NMR spectroscopy. *Biochemistry*, 33, 13531-9.
- WITTEKIND, M., MAPELLI, C., LEE, V., GOLDFARB, V., FRIEDRICH, M. S., MEYERS, C. A. & MUELLER, L. 1997. Solution structure of the Grb2 N-terminal SH3 domain complexed with a ten-residue peptide derived from SOS: direct refinement against NOEs, J-couplings and 1H and 13C chemical shifts. *J Mol Biol*, 267, 933-52.
- WÖHRLE, F. U., DALY, R. J. & BRUMMER, T. 2009. Function, regulation and pathological roles of the Gab/DOS docking proteins. *Cell Commun Signal*, 7, 22.
- WOLF, I., JENKINS, B. J., LIU, Y., SEIFFERT, M., CUSTODIO, J. M., YOUNG, P. & ROHRSCHEIDER, L. R. 2002. Gab3, a new DOS/Gab family member, facilitates macrophage differentiation. *Mol Cell Biol*, 22, 231-44.
- WRIGHT, P. E. & DYSON, H. J. 1999. Intrinsically unstructured proteins: re-assessing the protein structure-function paradigm. *J Mol Biol*, 293, 321-31.
- WU, S., AVILA-SAKAR, A., KIM, J., BOOTH, D. S., GREENBERG, C. H., ROSSI, A., LIAO, M., LI, X., ALIAN, A., GRINER, S. L., JUGE, N., YU, Y., MERGEL, C. M., CHAPARRO-RIGGERS, J., STROP, P., TAMPE, R., EDWARDS, R. H., STROUD, R. M., CRAIK, C. S. & CHENG, Y. 2012. Fabs enable single particle cryoEM studies of small proteins. *Structure*, 20, 582-92.
- WU, X., KNUDSEN, B., FELLER, S. M., ZHENG, J., SALI, A., COWBURN, D., HANAFUSA, H. & KURIYAN, J. 1995. Structural basis for the specific interaction of lysine-containing proline-rich peptides with the N-terminal SH3 domain of c-Crk. *Structure*, 3, 215-26.
- XIAO, G. H., JEFFERS, M., BELLACOSA, A., MITSUUCHI, Y., VANDE WOUDE, G. F. & TESTA, J. R. 2001. Anti-apoptotic signaling by hepatocyte growth factor/Met via the phosphatidylinositol 3-kinase/Akt and mitogen-activated protein kinase pathways. *Proc Natl Acad Sci U S A*, 98, 247-52.
- XIE, H., VUCETIC, S., IAKOUCHEVA, L. M., OLDFIELD, C. J., DUNKER, A. K., UVERSKY, V. N. & OBRADOVIC, Z. 2007. Functional anthology of intrinsic

- disorder. 1. Biological processes and functions of proteins with long disordered regions. *J Proteome Res*, 6, 1882-98.
- XUE, B., DUNKER, A. K. & UVERSKY, V. N. 2012. Orderly order in protein intrinsic disorder distribution: disorder in 3500 proteomes from viruses and the three domains of life. *J Biomol Struct Dyn*, 30, 137-49.
- YAFFE, M. B. 2002. Phosphotyrosine-binding domains in signal transduction. *Nat Rev Mol Cell Biol*, 3, 177-86.
- YAMAMOTO, H., FUJIOKA, Y., SUZUKI, S. W., NOSHIRO, D., SUZUKI, H., KONDO-KAKUTA, C., KIMURA, Y., HIRANO, H., ANDO, T., NODA, N. N. & OHSUMI, Y. 2016. The Intrinsically Disordered Protein Atg13 Mediates Supramolecular Assembly of Autophagy Initiation Complexes. *Dev Cell*, 38, 86-99.
- YAMASAKI, S., NISHIDA, K., SAKUMA, M., BERRY, D., MCGLADE, C. J., HIRANO, T. & SAITO, T. 2003. Gads/Grb2-mediated association with LAT is critical for the inhibitory function of Gab2 in T cells. *Mol Cell Biol*, 23, 2515-29.
- YANG, S. & BOURNE, P. E. 2009. The evolutionary history of protein domains viewed by species phylogeny. *PLoS One*, 4, e8378.
- YASUI, T., MASAKI, T., ARITA, Y., ISHIBASHI, T., INAGAKI, T., OKAZAWA, M., OKA, T., SHIOYAMA, W., YAMAUCHI-TAKIHARA, K., KOMURO, I., SAKATA, Y. & NAKAOKA, Y. 2016. Molecular Characterization of Striated Muscle-Specific Gab1 Isoform as a Critical Signal Transducer for Neuregulin-1/ErbB Signaling in Cardiomyocytes. *PLoS One*, 11, e0166710.
- YU, H., CHEN, J. K., FENG, S., DALGARNO, D. C., BRAUER, A. W. & SCHREIBER, S. L. 1994. Structural basis for the binding of proline-rich peptides to SH3 domains. *Cell*, 76, 933-45.
- YU, H., ROSEN, M. K., SHIN, T. B., SEIDEL-DUGAN, C., BRUGGE, J. S. & SCHREIBER, S. L. 1992. Solution structure of the SH3 domain of Src and identification of its ligand-binding site. *Science*, 258, 1665-8.
- YUZAWA, S., OGURA, K., HATANAKA, H., MIURA, K. & INAGAKI, F. 2003. Backbone assignments of Grb2 complexed with ligand peptides for SH3 and SH2 domains. *J Biomol NMR*, 27, 185-6.
- YUZAWA, S., YOKOCHI, M., HATANAKA, H., OGURA, K., KATAOKA, M., MIURA, K., MANDIYAN, V., SCHLESSINGER, J. & INAGAKI, F. 2001. Solution structure of Grb2 reveals extensive flexibility necessary for target recognition. *J Mol Biol*, 306, 527-37.
- ZHANG, X., DONG, Z., ZHANG, C., UNG, C. Y., HE, S., TAO, T., OLIVEIRA, A. M., MEVES, A., JI, B., LOOK, A. T., LI, H., NEEL, B. G. & ZHU, S. 2017. Critical Role for GAB2 in Neuroblastoma Pathogenesis through the Promotion of SHP2/MYCN Cooperation. *Cell Rep*, 18, 2932-2942.
- ZHANG, Y., DIAZ-FLORES, E., LI, G., WANG, Z., KANG, Z., HAVIERNIKOVA, E., ROWE, S., QU, C. K., TSE, W., SHANNON, K. M. & BUNTING, K. D. 2007. Abnormal hematopoiesis in Gab2 mutant mice. *Blood*, 110, 116-24.

Appendices

A.1 Gab1 amino acid sequence (UniProt accession code: Q13480)

¹MSGGEVVC¹⁰WLRKSPPEK²⁰LKRYAWKRR³⁰FVLRSGRLT⁴⁰DPDV⁴⁴LEYYKN⁵⁰DHAK
⁶⁰KPIRI⁶⁰DLNLCQQVDA⁷⁰GLTFNKKEFE⁸⁰NSYIFDINTI⁹⁰DRIFYLVADS¹⁰⁰EEEMNKWVRC¹¹⁰ICD
¹²⁰ICGFNPT¹²⁰EEDPVKPPGS¹³⁰SLQAPADLPL¹⁴⁰AINAPPSTQ¹⁵⁰ADSSSATLPP¹⁶⁰PYQLINVPPH¹⁷⁰
¹⁸⁰LETLGIQEDP¹⁸⁰QDYLLINCQ¹⁹⁰SKKPEPTRTH²⁰⁰ADSAKSTSSE²¹⁰TDCNDNVPSH²²⁰KNPASSQ
²³⁰SKH²³⁰GMNGFFQQM²⁴⁰TYDSSPPSRAP²⁵⁰SASVDSSLYN²⁶⁰LPRSYSHDVL²⁷⁰PKVSPSSTE²⁸⁰DG
²⁹⁰ELYVFNT²⁹⁰SGTSSVETQM³⁰⁰RHVSISYDIP³¹⁰PTPGNTYQIP³²⁰RTFPEGTLGQ³³⁰TSKLDTIPDI³⁴⁰
³⁵⁰PPRPPKPHP³⁵⁰AHDRSPVETC³⁶⁰SIPRTASDTD³⁷⁰**SSYCIPTAGM³⁸⁰SPRSNTIST³⁹⁰VDLNKL**
⁴⁰⁰**KDA⁴⁰⁰SSQDCYDIPR⁴¹⁰AFPSDRSSSL⁴²⁰EGFHNHFVKV⁴³⁰NVLTVGSVSS⁴⁴⁰EELDENYVPM⁴⁵**
⁴⁶⁰**NPNSPPRQHS⁴⁶⁰SSFTEPIQEA⁴⁷⁰NYVPMTPGTF⁴⁸⁰DFSSFGMQVP⁴⁹⁰PPAHMGFRSS⁵⁰⁰PKTP**
⁵¹⁰**RRRVPV⁵¹⁰VADCEPPPVD⁵²⁰RNLKPDRKVK⁵³⁰PAPLEIKPLP⁵⁴⁰EWEELQAPVR⁵⁵⁰SPITRSFA**
⁵⁶⁰**RD⁵⁶⁰SSRFPMSRP⁵⁷⁰DSVHSTTSSS⁵⁸⁰DSHDSEENYV⁵⁹⁰PMNPNLSSD⁶⁰⁰PNLFGSNSLD⁶¹⁰GG**
⁶²⁰SSPMIKPK⁶²⁰GDKQVEYLDL⁶³⁰DLDSGKSTPP⁶⁴⁰RKQKSSGSGS⁶⁵⁰SVADERVDYV⁶⁶⁰VVDQQKT
⁶⁷⁰LAL⁶⁷⁰KSTREAWTDG⁶⁸⁰RQSTESSETPA⁶⁹⁰KSVK⁶⁹⁴

Figure A.1 Amino acid sequence of full-length human WT Gab1 (aa 1 – aa 694) (UniProt accession code: Q13480). Bold and grey coloured amino acids: p35Gab1 (371 aa – 610 aa)

A.2 Table of intramolecular crosslinks found in p35Gab1 (His-p35Gab1 WT) and Grb2 (WT).

Intramolecular crosslinks in Hisp35Gab1

Peptide 1			Peptide 2		
Sequence	Domain	Amino acid	Sequence	Domain	Amino acid
499[SSPKTPPR]506	pro-rich	K (502)	392[DLNKLK]397		K (395)
499[SSPK]502	pro-rich	K (502)	551[SPITR]555		S (551)
499[SSPKTPPR]506	pro-rich	K (502)	392[DLNKLK]398		K (395)
499[SSPKTPPR]506	pro-rich	K (502)	398[KDASSQ]403		K (398)
499[SSPKTPPR]506	pro-rich	K (502)	392[DLNKLK]397		K (395)
499[SSPKTPPR]506	pro-rich	K (502)	396[LRK]398		K (398)
499[SSPKTPPR]506	pro-rich	K (502)	392[DLNKLK]398		K (395)
499[SSPKTPPR]506	pro-rich	K (502)	27[DDDDK]31	Enterokinase Cleavage Site	K (31)

Intramolecular crosslinks in Grb2

Peptide 1			Peptide 2		
Sequence	Domain	Amino acid	Sequence	Domain	Amino acid
14[DDELSFKR]21	SH3 N	K (20)	39[AELNGK]44	SH3 N	K (44)
14[DDELSFKR]21	SH3 N	K (20)	8[DFKATA]13	SH3 N	K (10)

Table A.2.1 Intramolecular crosslinks found in p35Gab1 and Grb2

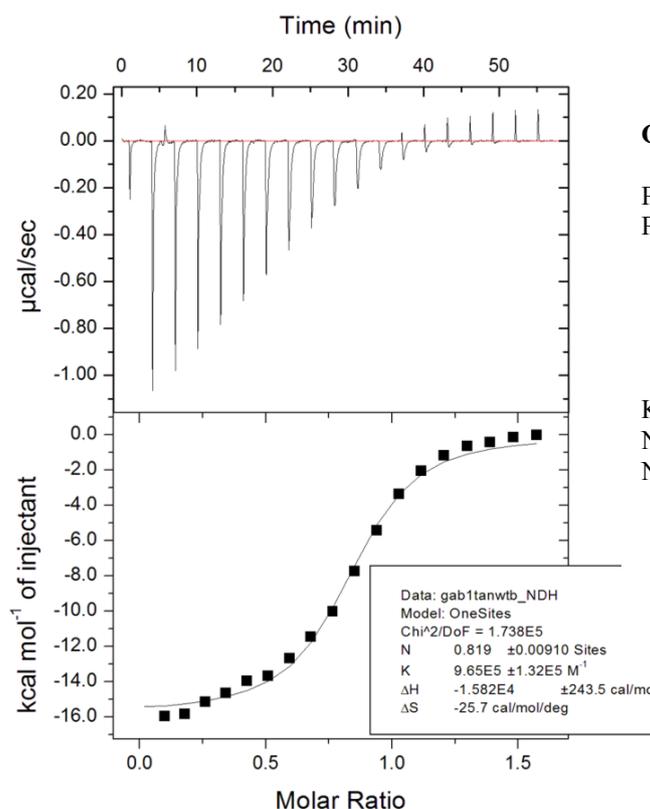
The numbering in p35Gab1 and Grb2 represent the position in full-length Gab1 and full-length Grb2, respectively.

A.3 Table of p35Gab1-Grb2 and Grb2-Gab1 peptides crystal screens

Crystal screen	Sparse Matrix Crystal Screen
His-Grb2/ Grb2 (C32S, C198A) + 32 aa p35Gab1 peptide	JBScreen JCSG ++ 1-4 (Jena Bioscience) JBScreen classic 1-10 (Jena Bioscience) JBScreen cryo 1-4 (Jena Bioscience) Crystal screen™ HR2-110, HR-112 (Hampton Research)
Grb2 (C32S, C198A) + 45 aa p35Gab1 peptide	Morpheus™ MD1-46 / MD1-47 (Molecular Dimensions) Low ionic strength crystallization kit + extension (Sigma-Aldrich)
His-Grb2/ Grb2 (C32S, C198A) + 32 aa Garem1 peptide	
Full-length p35Gab1 (C374A, C405A, C514A) + Grb2 (C32S, C198A)	JBScreen JCSG ++ 1-4 (Jena Bioscience) JBScreen classic 1-10 (Jena Bioscience) JBScreen cryo 1- 4 (Jena Bioscience) Crystal screen™ HR2-110, HR-112 (Hampton Research) Low ionic strength crystallization kit + extension (Sigma-Aldrich)

Table A.3.1 Crystal screens used in the thesis

A.4 Representative ITC results

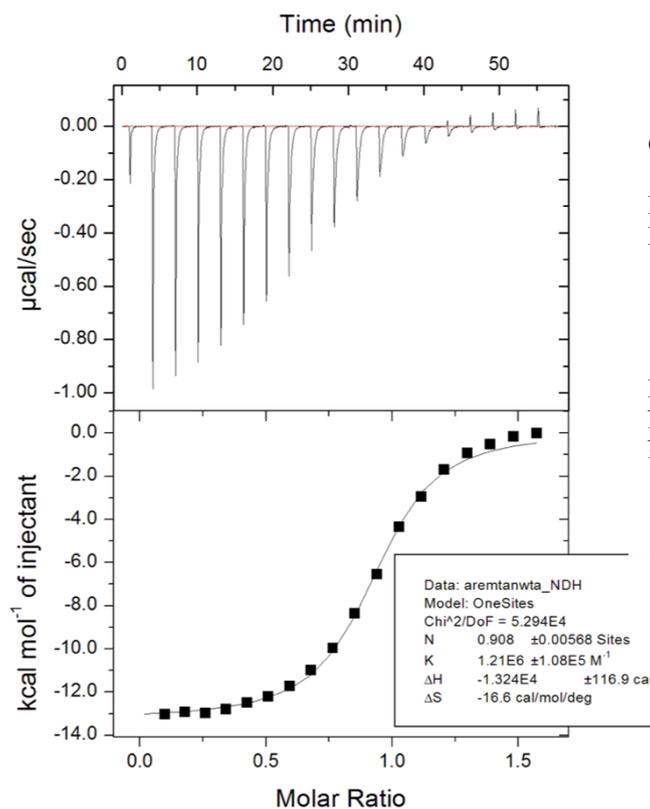


Grb2 + Gab1 peptide (tandem motif)

Protein/domain: His-Grb2 full, C32S C198A
Peptide: 32 aa Gab1,
SH3 tandem peptide
(C514S)

FRSSPKTPRRRPVPVAD
SEPPPVDRLNLPDRK

K_d : 1 μ M
N: 0.8
No.: 17

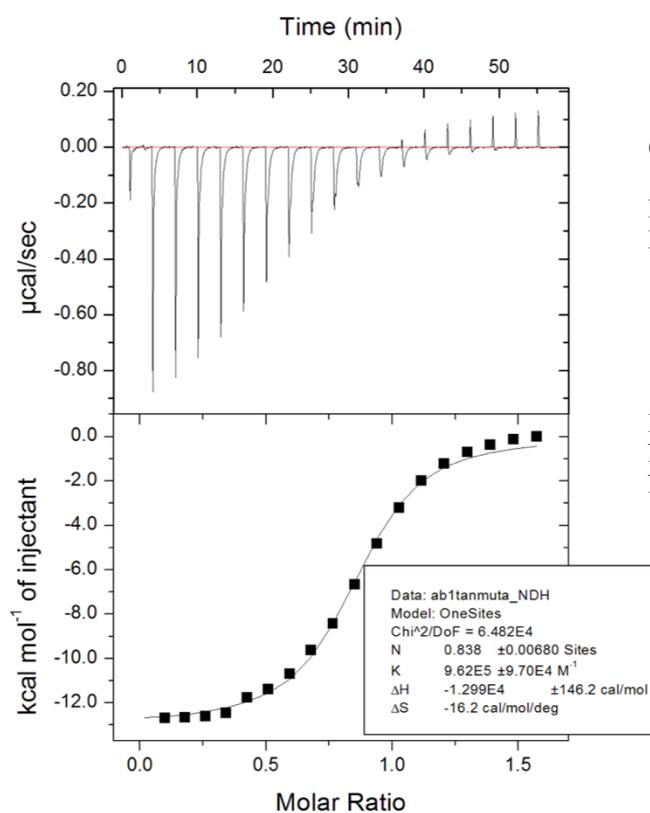


Grb2 + Garem1 peptide (tandem motif)

Protein/domain: His-Grb2 full, C32S C198A
 Peptide: 32 aa Garem1
 SH3 tandem peptide (WT)

LLNAPPVPPRS AKPLSTS
 PSIPRRTVKPARQQ

K_d 0.8 μ M
 N: 0.9
 No.: 25



Grb2 + Gab1 peptide (mutated tandem motif)

Protein/domain: His-Grb2 full, C32S C198A
 Peptide: AA mutant 32 aa Gab1
 SH3 tandem peptide (P501A,
 R506A, C514S)

FRSSAKTPPARPVPVADSEP
 PPVDRNLKPRDK

K_d 1.0 μ M
 N: 0.8
 No.: 19

Figure A.4.1 ITC results

A.5 Table of hydrogen bonds in Gab1 peptide – Grb2 SH3N and Gab1 peptide – Grb2 SH3C

Amino acid Gab1 peptide	Position in full-length Gab1	Atoms involved in Gab1 residue	Atoms involved in Grb2 residue	Distance [Å]
Pro5	501	O in peptide backbone	Tyr52 OH	2.3
Lys6	502	O in peptide backbone	Asn51 ND2	3.1
Pro8	504	O in peptide backbone	Trp36 NE1	2.9
Pro9	505	O in peptide backbone	Asn35 ND2	2.8
Arg10	506	NH2 in residue NH1 in residue NH1 in residue	Asp15 OD2 Asp15 OD1 Glu16 OE2	2.8 3.2 3.0
Arg11	507	O in peptide backbone N in peptide backbone	Gln34 N Asp33 OD1 Asp33 OD2	2.9 3.3 3.1
Val13	509	N in peptide backbone	Ser32 O in peptide backbone	3.4

Table A.5.1 H-bonds between the Gab1 peptide and the Grb2 SH3N domain

Amino acid Gab1 peptide	Position in full-length Gab1	Atoms involved in Gab1 residue	Atoms involved in Grb2 residue	Distance [Å]
Pro21	517	O in peptide backbone	Tyr209 OH	2.8
Pro22	518	O in peptide backbone	Asn208 ND2	2.8
Asp24	520	O in peptide backbone	Trp193 NE1	2.8
Arg25	521	NE atom NH1 atom	Glu174 OE2 OE1	2.8 3.0
Lys28	524	NZ atom	Glu171 OE1 Glu174 OE2	3.0 2.6

Table A.5.2 H-bonds between the Gab1 peptide and the Grb2 SH3C domain

Amino acid 1	Position in full-length Gab1	Atom involved	Amino acid 2	Position in full-length Gab1	Atoms involved	Distance [Å]
Asp24	520	O (peptide backbone)	Leu27	523	N-H (peptide backbone)	3.2
Arg25	521	O (peptide backbone)	Lys28	524	N-H (peptide backbone)	3.2

Table A.5.3 Identified intramolecular H-bonds in the 32 aa Gab1 peptide (bound to Grb2 SH3C domain)

A.6 SH2 – SH3C domain interface in Grb2-dimer (pdb code: 1gri)

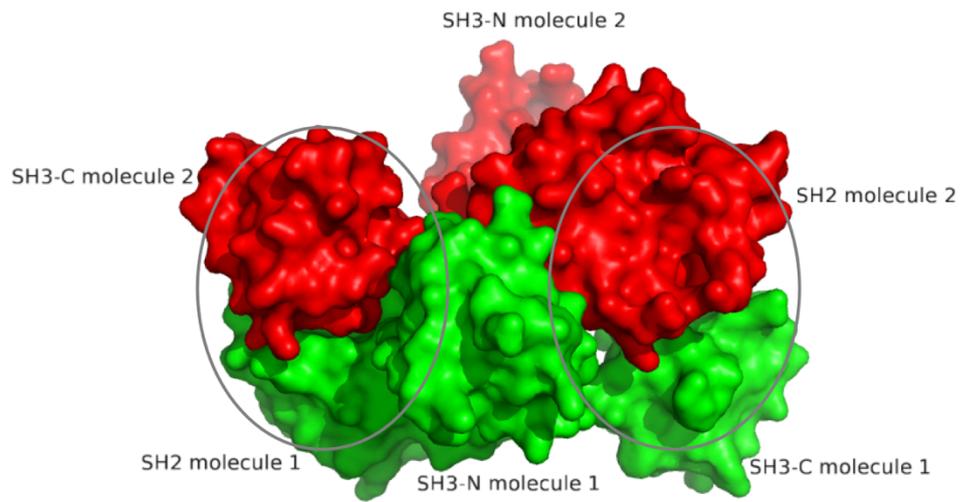


Figure A.6.1 SH2 – SH3C domain interface in Grb2-dimer

Unbound dimeric Grb2 structure (pdb code 1gri) with SH2 – SH3C interface between each Grb2 molecule (Grb2 molecule 1 coloured in red, Grb2 molecule 2 coloured in green). Red circles highlight the contact area between the SH2 domain and the C-terminal Grb2 SH3 domain.

Table of figures and tables

Figure 1.1 Schematic overview of a simplified cell signalling pathway.....	1
Figure 1.2 Various scaffold proteins or related proteins in cell signalling.....	4
Figure 1.3 The p35Gab1 fragment (aa 371 – aa 610) of full-length Gab1 (aa 1 – aa 694).....	7
Figure 1.4 Simplified overview of the Gab1 protein in the c-Met pathway.....	8
Figure 1.5 Grb2 ribbon model.....	10
Figure 1.6 Structural composition of the Gab family of proteins (A) and predicted intrinsic disorder in the Gab1 protein (B).....	13
Figure 1.7 N-terminal folding nucleation (NFN) hypothesis for Gab and other LMD proteins.....	16
Figure 1.8 A topological model of Gab1 explaining how Gab can facilitate signal computation within cell signalling.....	17
Figure 2.1 Representative ITC measurement.....	35
Figure 3.1 Purification of GST-Gab2.....	39
Figure 4.1 Schematic view of Gab1, p35Gab1 and Grb2.....	42
Figure 4.2 IMAC of His-p35Gab1 alone and His-p35Gab1 co-expressed with His-Grb2.....	44
Figure 4.3 Testing of different IMAC gradients for His-p35Gab1 – Grb2 elution.....	45
Figure 4.4 Mono Q purification of His-p35Gab1 – Grb2 complex.....	47
Figure 4.5 SEC analysis of Grb2 in complex with wildtype or cysteine-free mutant His-p35Gab1.....	48
Figure 4.6 SEC analysis of His-p35Gab1.....	49
Figure 5.1 Analytical ultracentrifugation of Grb2 and His-p35Gab1 – Grb2 complex.....	53
Figure 5.2 AFM analysis of the His-p35Gab1 – His-Grb2 and the His-p35Gab1 – Grb2 complex.....	54
Figure 5.3 Negative TEM stain of the p35Gab1 – Grb2 complex.....	54
Figure 5.4 XL-MS-identified interaction sites in the p35Gab1 – Grb2 complex are displayed on a Grb2 model.....	56
Figure 5.5 Putative Grb2 SH3 domain-binding tandem motif in p35Gab1 (Gab1 aa 371 – aa 610).....	60
Figure 5.6 Grb2 SH3 domain-binding tandem motif in Gab proteins and Gab-like proteins.....	61
Figure 5.7 Proteins with a Grb2 tandem binding motif.....	62
Figure 5.8 The putative Grb2 tandem motif in Garem1 is found in a disordered region.....	63
Figure 5.9 Strong evolutionary conservation of the putative Grb2 SH3 domain-binding tandem peptide in Garem1 and Gab1 throughout metazoans.....	64
Figure 5.10 SEC of two differently concentrated Grb2 (C32S, C198A) samples on a S75 column.....	66
Figure 5.11 Overview of initial crystallization screens.....	68
Figure 5.12 Grb2 – Gab1 peptide (32 aa) crystals (fine screen of condition C).....	69
Figure 5.13 Ramachandran plot of the Grb2 – Gab1 peptide complex crystal structure.....	70
Figure 5.14 Structure of Grb2 complexed with two Gab1-derived peptides harbouring the Grb2 tandem motif.....	71
Figure 5.15 Grb2-SH3N interaction with a Gab1-derived peptide.....	73
Figure 5.16 Detailed representation of the Grb2 SH3N – Gab1 peptide interaction site.....	74
Figure 5.17 Grb2 SH3N residues contributing in the P-x-x-P-x-R Gab1 peptide interaction are highly conserved.....	75
Figure 5.18 Gab1 peptide binds to the Grb2 SH3N in a class II orientation.....	76
Figure 5.19 The overall structure of the Grb2 SH3N domain remains merely unchanged upon Gab1 peptide binding.....	76
Figure 5.20 Complex between the Grb2 SH3C domain and a Gab1-derived peptide.....	77
Figure 5.21 Formed H-bonds between the Grb2 SH3C domain and the Gab1 peptide.....	78
Figure 5.22 Sequences of the human 32 aa Gab1 and the 14 aa Gab2b peptides that were crystallized with the Grb2 SH3C domain.....	79
Figure 5.23 Grb2 SH3C domain complexed with Gab family peptides and superposition of Gab1 peptide-bound and unbound Grb2 SH3C domains.....	80
Figure 5.24 Superposition of Grb2 SH2 domains.....	81
Figure 5.25 Interface between Grb2 SH2 domain and neighbouring Grb2 SH3C domain.....	82
Figure 5.26 Schematic depiction of potential 2D Grb2 – Gab1 cluster.....	83
Figure 6.1 Schematic overview of p35Gab1, Grb2 and the PI3K regulatory subunit p85 α (cysteine-free mutant).....	86
Figure 6.2 Purification of the PI3K regulatory subunit p85 α (mut.).....	88
Figure 6.3 Atomic Force Microscopy (AFM) of PI3K regulatory subunit p85 α	89
Figure 6.4 Phosphorylation of p85 binding sites in p35Gab1.....	90

Table of figures and tables

<i>Figure 6.5 p35Gab1 – Grb2 – p85α complex formation</i>	91
<i>Table 2.1 List of antibodies</i>	21
<i>Table 2.2 List of bacterial strains</i>	22
<i>Table 2.3 List of protein expression vectors</i>	22
<i>Table 2.4 List of peptides</i>	23
<i>Table 5.1 Identified intermolecular cross-linked peptides in the p35Gab1 – Grb2 complex</i>	56
<i>Table 5.2 ITC measurements of Grb2 interactions with Gab1- and Garem1-derived peptides</i>	59
<i>Table 5.3 Data collection for Grb2 – Gab1 peptide (32 aa) crystal and structure building</i>	69

List of abbreviations

3C protease	Precision protease
Å	Angstrom
A	Ala, Alanin
aa	Amino Acid
AFM	Atomic force microscopy
APS	Ammonium persulfate
AUC	Analytical Ultracentrifugation
BB	Binding buffer
BS	Binding site
BS2G	bis[sulfosuccinimidyl]glutarate
BSA	Bovine serum albumin
Da	Dalton
DLS	Dynamic light scattering
DMSO	Dimethyl sulfoxide
DTT	Dithiothreitol
<i>E. coli</i>	<i>Escherichia coli</i>
EB	Elution buffer
EDTA	Ethylene diamine tetraacetic acid
EGFR	Epidermal growth factor receptor
EM	Electron microscopy
Gab	Grb2-associated binder (Gab) protein
Grb2	Growth factor receptor-bound protein 2
GSH	Gluthathione sepharose
GST	Glutathione S-transferase tag
HGF	Hepatocyte growth factor
His-tag	Polyhistidine tag
HRP	Horseradish peroxidase
IDP	Intrinsically disordered protein
IDR	Intrinsically disordered region
IgG	Immunoglobulin G
IMAC	Immobilized Metal Affinity Chromatography
IPTG	Isopropyl- β -D-thiogalactopyranoside
ITC	Isothermal titration calorimetry
K	Lysine
K_d	Dissociation constant
KO	Knock out

List of abbreviations

LB	Luria-Bertani medium
LMD protein	Large multisite docking protein
mAU	Milli Absorbance units
MBR	c-Met binding region
mut.	Mutant
MW	Molecular weight
MWCO	Molecular weight cut off
N	Stoichiometry
NFN hypothesis	N-terminal folding nucleation hypothesis
NMR	Nuclear magnetic resonance
OD ₆₀₀	Optical density at a wavelength of 600 nm
ON	Overnight
PDB	Protein Data Bank
PEG	Polyethylene glycol
PH domain	Pleckstrin homology domain
PMSF	Phenylmethanesulfonyl fluoride
R	Arginine
rpm	Rotations per minute
RT	Room temperature
RTK	Receptor tyrosine kinase
SDS	Sodium dodecyl sulfate
PAGE	Polyacrylamide gel electrophoresis
SEC	Size exclusion chromatography
SH2 domain	Src homology 2 domain
SH3 domain	Src homology 3 domain
TB	Terrific broth medium
TBST	Tris Buffered Saline with Tween® 20
TCR	T-cell receptor
TEM	Transmission electron microscope
TEMED	Tetramethyl ethylene diamine
Tris	Tris-(hydroxymethyl)-aminomethan
UV	Ultraviolet
v/v	Volume per volume
w/v	Weight per volume
WT	Wildtype
XL-MS	Cross-linking mass spectrometry
Y	Tyrosine

Eidesstattliche Erklärung

Hiermit erkläre ich, diese Dissertation selbstständig und ohne fremde Hilfe verfasst zu haben. Es wurden nur die angegebenen Hilfsmittel und Quellen verwendet, inhaltlich oder wörtlich entnommene Stellen sind als solche markiert. Weiter erkläre ich, dass ich die vorliegende Dissertation an keiner anderen wissenschaftlichen Einrichtung zur Bewerbung eines Doktorgrades eingereicht habe. Ich selbst habe mich noch nicht um einen Doktorgrad beworben.

Bremen, den 19.04.2019

Katharina Mandel

List of publications

- Bongartz H., Hessenkemper W., Müller C., Fensky M., Fritsch J., **Mandel K.**, Behrmann I., Haan C., Fischer T., Feller M.F., Schaper F. (2017) The multi-site docking protein Gab1 is constitutively phosphorylated independent from its recruitment to the plasma membrane in Jak2-V617F- positive cells and mediates proliferation of human erythroleukaemia cells. *Cellular Signal*. 35: 37-47
- **Mandel K.**, Yang Y., Schambach A., Glage S., Otte A., Hass R. (2014) Mesenchymal stem cells directly interact with breast cancer cells and promote tumor cell growth in vitro and in vivo. *Stem Cells Dev*. 22(23): 3114-27.
- **Mandel K.**, Seidl D., Rades D., Lehnert H., Gieseler F., Hass R., Ungefroren H. (2013) Characterization of spontaneous and TGF- β -induced cell motility of primary human normal and neoplastic mammary cells in vitro novel real-time technology. *PLoS One* 8(2): e56591.
- Külzer S., Petersen W., Baser A., **Mandel K.**, Przyborski J.M. (2013) Use of self-assembling GFP to determine protein topology and compartmentalisation in the *Plasmodium falciparum*-infected erythrocyte. *Mol Biochem Parasitol*. 187(2): 87-90.
- Bucan V., **Mandel K.**, Bertram C., Lazaridis A., Reimers K., Park-Simon TW, Vogt P.M., Hass R. (2012) LEF-1 regulates proliferation and MMP-7 transcription in breast cancer cells. *Genes Cells* 17(7): 559-67.
- Otte A., **Mandel K.**, Reinstrom G., Hass R. (2011) Abolished adherence alters signaling pathways in phorbol ester-induced human U937 cells. *Cell Commun Signal* 9: 20
- Spork S., Hiss J.A., **Mandel K.**, Sommer M., Kooij T.W. A., Chu T., Schneider G., Maier U.G., Przyborski J.M. (2009) An unusual ERAD-like complex is targeted to the apicoplast of *Plasmodium falciparum*. *Eukaryot Cell*. 8(8): 1134-45

Acknowledgement

At this point, I am eternally thankful for all those who contributed to my PhD thesis.

Foremost, I would like to thank my supervisor Professor Stephan Feller for the opportunity to work on such fascinating and challenging project. I am grateful for his guidance, his kind support and advice, as well as his enthusiasm for all kind of science related discussions.

I am deeply grateful to Dr. Marc Lewitzky for teaching me the experimental methods, for his continuous support and his endless and dedicated time. His advices, comments and patience have been a great help during my PhD.

I would like to express my gratitude to Professor Milton Stubbs for the fruitful collaboration and the use of the X-ray crystallography facility. I would like to thank Dr. Christoph Parthier and Dr. Constanze Breithaupt for introducing me into the methodology of X-ray crystallography. In particular, I am grateful to Dr. Constanze Breithaupt for determining the Grb2 – Gab1 peptide crystal structure, her insightful discussions, her help to answer all my questions and her seemingly endless time.

I would like to thank Prof. Andrea Sinz and Nadja Mallock for the cross-linking mass spectrometry analysis of the complex and express a special thanks to all my other collaborators who contributed to this work.

Furthermore, I am thankful to previous and current lab members, Dr. Jenny Rouka, Dr. Philip Simister, Professor Toshi Menju, Dr. Martin Frejno, Mandy Bühnemann, Dr. Ilka Böhme, Heinrich Klose, Dr. Tobias Gruber and Dr. Terumasa Sowa, for the nice work atmosphere and their continual support. Especially, I would like to thank Dr. Tobias Gruber for his insightful comments and suggestions on my thesis and Dr. Jenny Rouka for her constant support and motivation.

Finally, I would like to thank my family, my friends Christine, Svenja and Jenny, as well as my friends spread over the world for their continuous support, their encouragement and their trust in me.

Curriculum Vitae

Persönliche Daten

Name: Katharina Mandel
Geburtsdatum: 07.10.1985
Geburtsort: Bremen

Ausbildung

Seit 03/2019 Wissenschaftliche Mitarbeiterin, Qualitätskontrolle BioNTech
IMFS

07/2018 – 10/2018 Junior Consultant, Unternehmensberatung, Thescon GmbH

01/2015 – 12/2016 Wissenschaftliche Mitarbeiterin, Martin-Luther Universität
Halle-Wittenberg, Institut für molekulare Medizin, Sektion
Tumorbiologie
*Thema: Untersuchung von Signalkomplexen aufbauend auf
Proteinen der Gab Familie (AG von Herrn Prof. Feller)*

10/2012 – 12/2014 Scientific Officer, Universität Oxford (GB), Fachbereich
Onkologie/Tumorbiologie

04/2012 – 09/2012 Wissenschaftliche Mitarbeiterin, Medizinische Hochschule
Hannover, Klinik für Frauenheilkunde und Geburtshilfe

10/2009 – 10/2011 Medizinische Hochschule Hannover, M. Sc. in Biomedizin

10/2006 – 06/2009 Philipps-Universität Marburg, B. Sc. in Biologie

08/1998 – 06/2005 Gymnasium Vegesack, Bremen, Abitur