Identification and Characterization of GRP50: A Novel Golgi-Associated Ankyrin Repeat Protein

by

Michele Marie Maxwell

B.A. Wellesley College, 1990

Submitted to the Department of Biology in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

at the

Massachusetts Institute of Technology

February 2002

Signature of Author: _____
Department of Biology
February 4, 2002

Certified by: _____
David E. Housman
Ludwig Professor of Biology
Thesis Supervisor

Accepted by: _____
Alan D. Grossman
Professor of Biology
Chairperson, Biology Graduate Committee

Identification and Characterization of GRP50:
A Novel Golgi-Associated Ankyrin Repeat Protein

by

Michele Marie Maxwell

Submitted to the Department of Biology
on February 4, 2002 in Partial Fulfillment of the
Requirements for the Degree of Doctor of Philosophy in
Biology

ABSTRACT

This thesis describes the identification and characterization of a novel gene, *GRP50* (Golgi-associated ankyrin repeat protein of 50 kilodaltons), which encodes a highly conserved mammalian protein exhibiting a unique structural architecture. Homology searching tools reveal that no known orthologues of this protein have been identified in other species, indicating that it is not a member of a previously characterized family. GRP50 is abundantly expressed in fetal tissues and in adult brain, and is induced as part of the serum response in cultured primary fibroblasts.

Sequence analysis indicates that GRP50 encodes a medium-sized protein with multiple modular domains, including an ankyrin repeat region, a polyproline-rich region, and a leucine-rich region that may form α-helical coiled coils. Each of these domains has been shown in other proteins to function as a protein-protein interaction module. We have used immunolocalization techniques to show that GRP50 is peripherally associated with the cytoplasmic face of Golgi and vesicle membranes in cultured mammalian fibroblasts. Further, we have demonstrated that Golgi-association of GRP50 is sensitive to the effects of the fungal metabolite, Brefeldin A. In biochemical fractionation experiments, GRP50 was purified from cultured cells as part of a large macromolecular complex. The possible roles of GRP50 in biological events at the cytoplasmic face of Golgi and vesicle membranes are discussed.

Thesis Supervisor: David E. Housman
Ludwig Professor of Biology

ACKNOWLEDGEMENTS

I am indebted to the many individuals who have offered their assistance and support throughout my graduate school career. First, I would like to thank David for his advice, support and encouragement, and for providing an exceptional learning environment. I have learned a great deal from David, and I am truly grateful for the opportunities I have enjoyed in his laboratory

The members of my thesis committee—Jackie Lees, Frank Gertler, and Frank Solomon—have generously given their time, offered their insights, and shared their expertise. I would like to thank them all for their help. I would also like to thank Anne Young, who served as the external examiner on my thesis committee, for taking time to read my thesis and offer comments and suggestions.

I am truly indebted to all of the members of the Housman Lab, past and present. I have been exceedingly fortunate to know and work with all of them. The work described in this thesis began as a collaborative effort with Scott Lowe and Mila McCurrach, and would not have been possible without them. I would like to thank Julia Alberta, J. Michael Andresen, Jill Crittenden, Brigid Davis, Barbara Jordan, and Alex Kazantsev for sharing their expertise and advice, regarding both specific experiments and scientific life in general. Julian Borrow shared his knowledge of techniques, offered experimental and career advice, and drank late night coffee with me—I thank him and hope I've learned some of what he knows (and I wish I'd picked up his freezer organization scheme). Al Charest has been a terrific bay-mate and colleague, and I'm grateful to him for reagents, protocols, advice, encouragement, helpful discussions, and most of all, for Jacob. Dianne

Keen and Brenda Luciano have been my good friends and fellow graduate students, I thank them for many useful scientific ideas, great conversations, unfailing support, and a whole lotta deep-fried love.

During my time in the Housman Lab, I have been fortunate to work with three talented and hard-working undergraduate students. Vanessa Cheung, Nicole Lee, and Tina Pinto provided a great deal of help with experiments, and much more. I thank them for their hard work, and for teaching me (I hope) how to be a better teacher.

I would like to thank the members of the Lees lab, especially Raluca Verona, Jeff Trimarchi, Rebecca Landsberg, and Ulrike Ziebold, for their willingness to answer my many questions and their generosity in sharing reagents and protocols.

I am indebted to Jed Chatterton and Monty Krieger for discussions and advice about the analysis of Golgi proteins. Jed deserves extra thanks for his patient help with chromatographic fractionation experiments—he really went above and beyond the call.

On a personal note, I would like to thank all of my friends and family for their support throughout this long process. Special thanks to Tomonori Ishikawa, my dear friend of (gulp) fifteen years, for his support and encouragement, and for his incredibly generous gift of the super-cool hardware.

Extra-special thanks to my best friend, Matt, for his patience, love, and encouragement.. Matt has enriched every part of my life, and his unfailing support of the part that seems crazy to him has been an incredible gift.

TABLE OF CONTENTS

CHAPTER 1

Introduction

Introduction

The process of assigning biological functions to novel genes has long been at the heart of modern molecular biology. The tenet that clues to molecular function can be inferred from primary structure provides the underlying rationale for contemporary genomics and has significantly impacted the development of bioinformatics as an independent discipline. Traditionally, in the case of protein-coding genes, novel sequences are identified by genetic or genomic means, and their putative protein products are analyzed by a variety of methods designed to aid in functional characterization. These analyses are usually begun by electronic scanning of sequence databases in order to identify structurally related proteins or protein domains. A widespread computational application of this principle is currently underway, in the attempt to annotate the recently completed genomic sequences of a variety of organisms (Gaasterland and Oprea, 2001). Ultimately, of course, elucidation of the biological role of a newly identified protein requires a combination of computational analysis and experimental evidence.

In some instances, the path from primary structure to function is reasonably straightforward, as in the case when a newly identified protein has an obvious structural orthologue of known function in a different species. Numerous computational algorithms and databases have been created specifically to aid in the identification of structurally and functionally homologous sequences or sequence motifs (Altschul et al., 1997). Even when a pairwise sequence comparison search engine fails to detect highly related proteins, databases that rely on the generation of "signature" motifs from the sequences of functionally related proteins can often be used to distinguish these motifs in a novel sequence (Bateman et al., 1999; Henikoff et al., 1999). This type of analysis can prove

especially fruitful when a new protein contains conserved domains or motifs, such as those that mediate DNA binding or an enzymatic activity, that are clearly associated with a specific molecular function.

Even when the analysis of a new protein sequence reveals a highly conserved functional domain, the predictive value of such information can be limited. Identification of a conserved catalytic domain, for instance, is less valuable as an indicator of molecular function if no information is available regarding the endogenous substrates or regulators of the enzyme. In circumstances where a novel protein lacks homology to known functional domains, clues to biological activity can sometimes be inferred from the identification of conserved structural elements (Teichmann et al., 2001). These considerations are complicated, however, when a novel protein contains structural domains that are capable of mediating many distinct cellular processes.

This thesis describes the identification and characterization of a novel gene, *GRP50* (Golgi-associated Ankyrin-Repeat Protein of 50 kDa), which encodes a highly conserved mammalian protein exhibiting a unique structural architecture. No known orthologues of this gene have been identified in other species, indicating that it is not a member of a previously described family. The GRP50 protein contains several predicted structural domains, including an ankyrin repeat region, a proline-rich region, and a leucine-rich region exhibiting coiled coil-forming potential. Each of these types of domains has been shown in other proteins to function as protein-protein interaction modules. In addition, we have determined that GRP50 is associated with the cytoplasmic face of the Golgi apparatus in mammalian cells, where it may exist as part of a large multiprotein complex. In order to provide a framework for the analyses presented in the

chapters that follow, the sections below are devoted to an overview of the structural domains present in the GRP50 protein and their known functions in protein-protein interactions. The chapter concludes with a synopsis of some of the well-known multiprotein complexes that mediate membrane transport through and structural maintenance of the Golgi apparatus, and includes a discussion of the cytoplasmic factors that regulate these processes.

Structure and Function of Ankyrin Repeat Proteins

The ankyrin repeat is a 33 amino acid repeating motif that was first described in the sequences of the yeast cell-cycle regulators Swi6p and Cdc10p (Breeden and Nasmyth, 1987). In comparing the sequences of these two related proteins, the authors noted that similar repeating elements were also present in the sequences of the Notch protein in *Drosophila melanogaster* and of the lin-12 protein in *Caenorhabditis elegans*, two proteins involved in cell fate specification in differentiating tissues (Wharton et al., 1985; Yochem et al., 1988). Subsequently, cloning and analysis of the cDNA encoding human erythrocyte ankyrin revealed 22 copies of this repeat in the amino-terminal domain of the protein (Lux et al., 1990), and resulted in the designation of this element as the ankyrin (ANK) repeat.

Current analysis of the non-redundant protein databases indicates that ANK repeats are found in more than one thousand proteins from both prokaryotes and eukaryotes, with each protein containing between two and twenty (or more) copies of this motif (Bork, 1993; Michaely and Bennett, 1993). ANK repeats are present in proteins of diverse functions, including cytoskeletal proteins, transcription factors, cell cycle

inhibitors, developmental regulators, and toxins (Bork, 1993; Sedgwick and Smerdon, 1999). The common function of ANK repeat proteins is that they bind other proteins, but their binding partners are not restricted to a single class of target molecules. Further, although the interaction between an individual ANK repeat protein and its target protein is specific, the proteins that are known to bind ANK repeats are largely unrelated, both functionally and structurally. That a number of distinct biological roles have been ascribed to ANK repeat proteins is likely the result of the widely divergent set of proteins with which they interact.

This diversity of binding partners is illuminated by the structural organization of the ANK repeats themselves. Employing structural-prediction algorithms to analyze the repeats in erythrocyte ankyrin, Lux and colleagues (Lux et al., 1990) proposed that the principal structural elements of the motif consist of two amphipathic α-helices, separated by a loop, and followed by a β-hairpin. Subsequent structural prediction analyses, utilizing ANK repeat consensus sequences, have similarly proposed a β-hairpin-helix-loop-helix ($\beta_2\alpha_2$) structure (Figure 1A) in which the helices are aligned in an antiparallel fashion. In this structure, multiple repeats pack together via the hydrophobic side chains of the α-helices, such that β-hairpins of different repeats form an extended β-sheet (Bork, 1993; Michaely and Bennett, 1993; Sedgwick and Smerdon, 1999). The binding surface of a domain containing ANK repeats is therefore determined by the tertiary structure of multiple stacked repeats, and differences in the number of repeats can contribute to binding specificity (Andrade et al., 2001).

The proposed general structure of ANK repeat domains was first confirmed upon determination of the three-dimensional structure of the ANK repeat protein 53BP2 bound

to the p53 tumor suppressor protein (Gorina and Pavletich, 1996). Subsequently, the three-dimensional structures of additional ANK repeat proteins, either alone or bound to their known interacting partners, have been solved. These latter structures include those of the products of the *INK4* gene ($p16^{INK4a}$, $p18^{INK4c}$, and $p19^{INK4d}$) bound to the cyclin-dependent kinase CDK6 (Baumgartner et al., 1998; Brotherton et al., 1998; Jeffrey et al., 2000; Li et al., 1999; Russo et al., 1998), of the transcription factor GA-binding proteinβ (GABPβ) bound to GABPα and to DNA (Batchelor et al., 1998), and of the NF-κB inhibitor I-κBα, bound to NF-κB (Huxford et al., 1998; Jacobs and Harrison, 1998). These structures have revealed that the side-by-side stacking of the antiparallel α-helical pairs results in an extended β-sheet (composed of neighboring β-hairpins) that projects away from the helices, forming an L-shaped structure that has been compared to a cupped hand (Sedgwick and Smerdon, 1999). In the cupped hand analogy, the β-hairpins form the fingers and the α-helical bundle forms the palm. Figure 1B depicts the organization of an ANK repeat domain (viewed from two different angles), with a single ANK repeat highlighted in red.

The solved structures of ANK repeat proteins, in complexes with their known binding partners, provide further explanation for the observed heterogeneity of ANK repeat interacting proteins. Although the ANK repeat is defined as a consensus sequence, the most highly conserved residues among divergent ANK repeat proteins are those that confer the β-hairpin-helix-loop-helix secondary structure (Bork, 1993). This structure can tolerate considerable sequence variation, particularly in the residues that comprise the β-hairpin, which are not included in the ANK repeat consensus (Figure 1A and Bork, 1993; Michaely and Bennett, 1993). The above-mentioned structural studies have revealed that,

while binding can be mediated by any of the solvent-accessible residues of the stacked ANK repeat structure, it most often involves residues in the β-hairpins (Figure 1B, Sedgwick and Smerdon, 1999). Thus the specificity of a particular ANK repeat protein for its binding partner is conferred by the sequences that are least likely to be conserved among ANK repeat proteins in general.

These observations imply that binding specificity in ANK repeat proteins is determined both by the number of individual ANK repeats, which associate to form the tertiary structure of the domain, and by the sequences in and around the repeats themselves. That the number of ANK repeats present in a given protein can vary considerably is evident from scanning the protein sequence databases. The *Drosophila* plutonium protein contains two ankyrin repeats (Axton et al., 1994), the 53BP2 and INK4 proteins each contain four (Iwabuchi et al., 1994; Li et al., 1999), the I-κB protein contains six (Haskill et al., 1991), and ankyrin itself contains twenty-two copies (Lux et al., 1990). Most ANK repeat proteins contain four or more copies of the repeat; however, mutational analyses and biochemical techniques have recently been used to show that the carboxyl terminal two ANK repeats of $p16^{INK4a}$ can fold independently to assume a thermostable structure (Zhang and Peng, 2000). Based on this result, the authors propose that the two-ANK repeat fold represents the minimum structural unit for all ANK repeat proteins (Zhang and Peng, 2000).

As a consequence of the variation in the numbers and sequences of repeats found in ANK repeat proteins, these molecules have been found to interact with a widely divergent set of proteins and hence to perform a number of distinct biological roles. The ANK repeat protein GABPβ associates with GABPα, a transcription factor of the Ets-

domain family, to form a multisubunit complex that is capable of DNA binding and transcriptional activation (Batchelor et al., 1998; LaMarco and McKnight, 1989; Thompson et al., 1991). Similarly, the intracellular ANK repeat domain of the Notch receptor, which is released by proteolysis upon ligand binding, functions as a transcriptional co-activator when bound to the RBP-J (Su(H)) transcription factor (Beatus et al., 2001; Petcherski and Kimble, 2000; Tani et al., 2001; Wu et al., 2000). The biological role of the inhibitors of cyclin-dependent kinases differs from the above examples: $p16^{INK4a}$, $p15^{INK4b}$, $p18^{INK4c}$, and $p19^{INK4d}$ inhibit activity of CDK6 by stabilizing a catalytically inactive conformation of the enzyme (Brotherton et al., 1998; Jeffrey et al., 2000; Russo et al., 1998), thereby effecting an important cell cycle checkpoint. Lastly, the I-κB proteins inhibit the activity of the NF-κB transcription factor in two ways: I-κB binding both inhibits the ability of NF-κB to bind DNA and results in export of NF-κB from the nucleus and/or its sequestration in the cytoplasm (Li et al., 1998). The degree to which NF-κB activity is inhibited by a given I-κB depends upon the residues in the β-hairpins of the I-κB ANK repeats, which directly mediate binding to NF-κB (Huxford et al., 1998; Jacobs and Harrison, 1998; Simeonidis et al., 1999).

As is the case with many protein-protein interactions, binding of an ANK repeat protein to its target can be modulated rapidly in a variety of ways. Post-translational modifications (such as phosphorylation) of either binding partner can result in conformational changes that favor or impede binding. In addition, regulation of the steady state levels or subcellular locations of both proteins influences the likelihood of binding, as in the example of the Notch receptor described above. These examples

underscore the versatility of ANK repeat-mediated protein-protein interactions in effecting cellular responses to external stimuli.

The above examples illustrate the broad spectrum of biological functions attributed to ANK repeat proteins. The common feature of these domains is that almost all bind to other proteins, although one protein, the yeast Swi6, contains ANK repeats that are involved in intramolecular interactions (Foord et al., 1999). Thus, the biological role of an ANK repeat protein is determined by the function of its binding partner and by whether binding enhances or inhibits that function. The evolutionary conservation of the basic structure of the ANK repeat, and the large number of proteins that contain this element, emphasize its general importance as a protein-protein interaction domain. The exceptional versatility of a general structure that tolerates considerable sequence substitutions and varying repeat numbers contributes to its utility as a stable, yet adaptable, interaction domain.

Proline-Rich Sequences in Signaling and Scaffolding

Short proline-rich sequences also interact with a wide variety of cellular proteins, although the nature of these interactions is distinct from those of ANK repeat proteins in that proline-rich sequences bind specific classes of protein interaction domains. It is believed that proline-binding domains are among the most common types of protein interaction domains (Kay et al., 2000; Reiersen and Rees, 2001). Proline-rich motifs are the preferred ligands for modular interaction domains such as the Src homology 3 (SH3), WW, and EVH1 domains (Holt and Koffer, 2001; Kay et al., 2000; Mayer, 2001; Sudol, 1996; Zarrinpar and Lim, 2000).

The SH3 domain was among the first modular protein interaction domains to be identified (Mayer and Hanafusa, 1990; Stahl et al., 1988), and in subsequent years a substantial body of work has focussed on the numerous biological processes effected by these domains. SH3 domains are 50-70 amino acids in length and have been found in a large number of eukaryotic signal transduction and cytoskeletal proteins. These include the Src family of tyrosine kinases, the Ras effector Grb2, the regulatory subunits of phosphatidylinositol 3-kinase (PI3K) and phospholipase Cγ (PLCγ), and the actin binding protein α-spectrin. Mutations in the SH3 domains result in mislocalization of PLCγ and Grb2 (Bar-Sagi et al., 1993; Seidel-Dugan et al., 1992) and activate the transforming potential of Src and Abl (Franz et al., 1989; Seidel-Dugan et al., 1992). These results provided early evidence for a functional role for SH3 domains.

A screen of a cDNA expression library identified the first putative SH3-binding proteins (Cicchetti et al., 1992), and the binding sites in the target proteins were mapped to short (~10 amino acid) proline-rich sequences (Ren et al., 1993). More recently, the development of combinatorial peptide libraries has permitted determination of the ligand specificity of a number of SH3 domain proteins. These experiments have revealed that the preferred ligands for SH3 domains are proline-rich sequences containing the core conserved motif P–x–x–P, where x is any amino acid (Mayer and Eck, 1995; Ren et al., 1993). Moreover, individual SH3 domains exhibit preferences for distinct sequences around the core (Kay et al., 2000; Lim et al., 1994; Rickles et al., 1994; Sparks et al., 1996). Structural analyses of SH3 domains from a variety of proteins reveals a similar overall topology (Feng et al., 1995; Musacchio et al., 1994; Yu et al., 1994): the binding surface of an SH3 domain consists of three shallow grooves, two of which contact

hydrophobic-proline dipeptides in the P–x–x–P core (Kay et al., 2000; Mayer, 2001). One key determinant of specificity is given by the preference of most SH3 domains for a positively charged residue adjacent to the core motif; this residue is thought to occupy the third groove in the SH3 binding surface (Feng et al., 1995; Kay et al., 2000).

SH3 domains have been implicated in the control of a number of biological processes, including the regulation of enzymatic activity and the targeting of proteins to distinct subcellular locations. The classical example of the former function is the inhibition of Src kinase activity via the intramolecular interaction of its SH3 domain with a proline-rich linker region adjacent to its catalytic domain; this interaction stabilizes the catalytically inactive conformation of the kinase (Williams et al., 1997; Williams et al., 1998). The Src kinase is activated by a phosphorylation-induced conformational switch, which disrupts SH3 binding to the linker region. A prominent example of the role of SH3 domains in targeting proteins to particular locations is given by the recruitment by Grb2 (via its SH3 domain) of the Ras-GEF, Sos, to the plasma membrane in response to EGF-receptor activation (Lim et al., 1994; McCormick, 1993). Because Ras is localized at the plasma membrane, the Grb2-mediated translocation of Sos to the membrane couples Ras activation (by Sos) to the activation of receptor tyrosine kinases in response to extracellular signals.

The family of WW domain-containing proteins comprises a distinct class of proline-binding modules. The WW domain, named for two conserved tryptophan residues that are critical for function, is a small globular protein module of 38-40 amino acids which consists of a three-stranded antiparallel β-sheet (Bork and Sudol, 1994; Macias et al., 1996). Although WW domains are structurally unrelated to SH3 domains

and exhibit distinct substrate specificities, these domains bind proline-rich ligands in a similar manner (Sudol, 1996; Zarrinpar and Lim, 2000). WW domains bind proline-rich sequences containing the core consensus P–P–x–Y or P–P–L–P, usually flanked by additional proline residues (Bedford et al., 1997; Chen and Sudol, 1995; Linn et al., 1997; Zarrinpar and Lim, 2000).

Significant efforts have focussed on understanding the WW domain due to the fact that many signaling complexes involving this domain have been implicated in human diseases, including Liddle's syndrome of hypertension, muscular dystrophy, Alzheimer's disease, and Huntington's disease (Bork and Sudol, 1994; Chang et al., 1996; Faber et al., 1998; Rentschler et al., 1999; Schild et al., 1996). The function of WW domains is similar to that of other protein-interaction modules, in that binding of a WW domain to its polyproline ligand can alter the activity or subcellular localization of the ligand protein (Yagi et al., 1999). For example, the ubiquitin-ligase Nedd-4 has been shown to target the epithelial sodium channel for ubiquitin-dependent degradation, via interaction of the Nedd-4 WW domain with proline-rich sequences in the β subunit of the channel (Staub et al., 1996; Staub et al., 1997b). Mutations in and around the P–P–x–Y motif in the sodium channel, which have been detected in many patients with Liddle's syndrome, abolish binding and result in extended half-life of the channel and, ultimately, hypertension (Hansson et al., 1995; Staub et al., 1997a).

The biochemistry of proline provides some insights into the utility of proline-rich domains as ligands for protein interaction modules. The only naturally occurring amino acid with a cyclized side chain, proline restricts the conformation of polypeptide chains; proline-rich sequences tend to adopt a left-handed helical structure with three residues per

turn that has been termed the polyproline type II (PPII) helix (Reiersen and Rees, 2001). Although the PPII helical conformation is not limited to proline-rich sequences, the majority of PPII helices are proline-rich (Stapley and Creamer, 1999). In proline-rich PPII helices, the proline residues in a core P–x–x–P motif are on the same face of the helix and therefore are both accessible for interaction with other proteins. In addition, the interaction of PPII helices with both SH3 and WW domains is largely mediated by hydrophobic contacts, which result in relatively low-affinity binding (Kay et al., 2000; Zarrinpar and Lim, 2000). The relatively weak binding to polyproline domains suggests that the stability of these types of interactions may be modulated rapidly in response to cellular signals; for example, ligand binding may be stabilized through the additive affects of multiple interactions or by the assembly of a multiprotein complex. Alternatively, covalent modification of the ligand, such as by phosphorylation, may selectively stabilize or destabilize binding to its cognate domain.

The importance of proline-rich sequences as protein-interaction domains is emphasized by the frequency with which these sequences occur in a variety of organisms and by the large number of proteins that contain proline-binding domains. Comparative genomic analysis of the completed sequences of *C. elegans* and *D. melanogaster* indicates that proline-rich motifs are among the most common elements in these organisms (Rubin and Lewis, 2000). Recent insights into the structure of WW domains (Huang et al., 2000; Verdecia et al., 2000) revealed a striking similarity between the binding surfaces of the WW and SH3 modules. This observation is telling given that these two domains are otherwise structurally unrelated, and suggests that multiple mechanisms have evolved for interaction with proline-rich sequences.

In addition to the well-characterized SH3 and WW domains, several additional modules that bind proline-rich ligands have been described in recent years (Sudol 1998; Callebaut et al., 1998; Kay et al., 2000). These include the cytoskeletal regulatory protein profilin and the proteins of the EVH1 (Ena/VASP Homology) family, which regulate actin dynamics. EVH1 ligand sequences are present in components of focal adhesions, such as vinculin and zyxin, and are believed to influence cell motility (Laurent et al., 1999; Niebuhr et al., 1997). A common theme among the various proline-binding domains is their ability to recruit proteins containing proline-rich sequences to specific subcellular locations and/or into particular hetero-oligomeric complexes, often in response to a defined stimulus. It has therefore been suggested that the principal role of proline-binding modules in intermolecular interactions is as molecular scaffolds, which facilitate the regulated assembly of multiprotein complexes in situations requiring a rapid response, such as the initiation of signaling cascades, cytoskeletal rearrangement and synaptic vesicle endocytosis (Cohen et al., 1995; Kay et al., 2000; McPherson, 1999).

Multiprotein Complexes in Membrane Transport and Golgi Structure

Eukaryotic cells contain an elaborate network of membrane-bounded organelles, the maintenance of which requires the physical and functional separation of their resident components in the face of constant, bi-directional flow between compartments. The flow of membrane components and proteins from one organelle to another is mediated by transport vesicles, which are formed by pinching off of the membrane at one organelle and are consumed by fusing with the membrane at another. For instance, in the classical view of transport through the secretory pathway, membrane-bounded vesicles containing

newly synthesized cargo proteins bud from the endoplasmic reticulum (ER) and fuse with the *cis*-Golgi (Palade, 1975; Schekman and Orci, 1996). Similar yet biochemically distinct vesicles transport cargo through the *cis*-, *medial*-, and *trans*-Golgi and to the *trans*-Golgi network (TGN)-endosomal system. At the TGN, they are sorted to their correct destinations (e.g., plasma membrane, endosome, lysosomes, or secretory granules, Traub and Kornfeld, 1997). This process is mirrored in the constitutive and regulated endocytic pathways (Carpentier et al., 1992; Pearse, 1988), and in the retrograde transport of vesicles from the Golgi to the ER (Letourneur et al., 1994; Orci et al., 1997).

The structural maintenance of organelles and the processes of vesicle formation, transport, and fusion require the activities of specific hetero-oligomeric protein complexes that assemble on the cytoplasmic face of these membranes. An overview of these complexes is depicted in Figure 2. Formation of a vesicle from a donor membrane requires the prior assembly of a vesicle "coat", comprised of multiple subunit proteins, which deforms the membrane and facilitates budding of a nascent vesicle (Klumperman, 2000; Kreis et al., 1995; Lemmon and Traub, 2000; Schekman and Orci, 1996). A different set of proteins is required for docking and subsequent fusion of a vesicle with acceptor membranes: soluble proteins are recruited from the cytoplasm into a complex with integral membrane proteins on vesicles (v-SNAREs) and on the target membrane (t-SNAREs), thereby tethering the membranes to permit fusion (Clague, 1999; Hinshaw, 2000; Rothman, 1994; Sollner et al., 1993a). Specificity of vesicle formation and targeting is achieved through the use of heterotypic complexes at distinct subcellular sites, but the underlying principles that govern these processes in the ER, Golgi, and TGN-endosomal systems are similar. The sections below describe the structures of a few

of the membrane-associated complexes that mediate these events in the Golgi apparatus and discuss the regulatory mechanisms involved in complex assembly. The chapter concludes with a brief discussion of the role of these factors in the structural maintenance of the Golgi.

The coat complexes that mediate vesicle budding from Golgi membranes have been characterized extensively. Among the best understood of these are the clathrin/adaptor protein (AP) complexes, which are involved in receptor-mediated endocytosis at the plasma membrane and in the selective transport of proteins from the TGN to lysosomes or endosomes (Mostov et al., 2000; Robinson, 1994). Clathrin/AP complexes at different subcellular locations contain distinct AP complexes; to date, four AP complexes have been identified (Hirst et al., 1999; Robinson, 1994; Simpson et al., 1996; Simpson et al., 1997). Clathrin, which forms the vesicle scaffold, associates in the TGN with the AP1 complex, which consists of the soluble factors β'-adaptin, γ-adaptin, AP47, and AP19, to promote the formation of vesicles that transport cargo and membrane components to endosomes (Lemmon and Traub, 2000; Schekman and Orci, 1996).

In the endocytic pathway, a second adaptor protein complex, AP2, is recruited to the plasma membrane by interaction with the integral membrane protein synaptotagmin (von Poser et al., 2000; Zhang et al., 1994). The AP2 complex, which consists of distinct adaptin subunits (α-adaptin, β-adaptin, AP50, and AP17), associates with clathrin to form invaginations of the membrane known as coated pits (Robinson, 1994). Membrane fission requires the activity of yet another protein, the GTPase dynamin, which is recruited into this complex and forms a ring around the neck of the budding vesicle (Damke et al., 1995; Takei et al., 1995).

Although clathrin/AP1 coated vesicles are associated with the TGN, the majority of coated vesicles in the Golgi apparatus are decorated with the COPI (coat protein I) coat. COPI-coated vesicles are implicated both in anterograde intra-Golgi transport (of secretory cargo) and in retrograde transport (of ER-resident proteins) from the Golgi to the ER (Cosson and Letourneur, 1997; Kreis et al., 1995; Orci et al., 1997). The formation of COPI vesicles begins with the recruitment to the membrane of an ADP-ribosylation factor (ARF), a member of the family of Ras-related small GTPases (Serafini et al., 1991; Spang et al., 1998). The ARF, in its GTP-bound form, is required for the assembly of coatomer, a multiprotein complex composed of seven subunits (termed α, β, β', γ, δ, ε, and ζ), onto the membrane (Schekman and Orci, 1996; Waters et al., 1991). An essential role for COPI has been demonstrated in yeast, where deletion of any of the subunit components of this complex is lethal (Duden et al., 1994; Letourneur et al., 1994) and in mammals, where deletion of one of the coatomer subunits results in a global defect in secretion (Guo et al., 1994).

The assembly of the clathrin/AP1 complex (at the TGN) and of the coatomer complex (at the Golgi) is dependent upon the activity of ARFs. There are six ARF proteins in mammals (three in yeast), which are classified according to sequence similarity (Donaldson and Jackson, 2000; Roth, 1999); of these, the Class I ARFs (ARF1, ARF2, and ARF3) mediate the assembly of the clathrin/AP1 and COPI vesicle coats. The ARFs are themselves regulated by a family of guanine nucleotide exchange factors (ARF-GEFs), soluble cytoplasmic proteins which activate membrane-bound ARF by stimulating the exchange of GDP for GTP (Chardin and McCormick, 1999; Lippincott-Schwartz et al., 1998). In the case of COPI assembly, ARF-GTP then recruits the

coatomer components to the membrane to form the intact COPI coat; this assembly does not require GTP hydrolysis (Spang et al., 1998). At some point, either directly prior to or immediately following coatomer assembly, the molecular switch is terminated by the activity of the family of GTPase activating proteins (ARF-GAPs), which promote hydrolysis of ARF-GTP to ARF-GDP and subsequent dissociation of ARF-GDP from the membrane (Goldberg, 1999).

The protein complexes that mediate tethering and fusion of transport vesicles with acceptor membranes are distinct from those described above, although they exhibit some parallel features. The paradigm for fusion of a vesicle with an acceptor membrane was first elucidated in studies of synaptic vesicle release in neurons (Sollner et al., 1993a; Sollner et al., 1993b) and has since been proven broadly applicable to membrane fusion events. Membrane fusion requires recruitment of the soluble proteins NSF (N-ethyl maleimide-sensitive factor) and α-SNAP (α-soluble NSF attachment protein) into a complex with integral membrane proteins on vesicles (v-SNAREs) and on the target membrane (t-SNAREs) (Clague, 1999; Mochida, 2000; Waters and Pfeffer, 1999). A schematic representation of these events is depicted in Figure 3A. Vesicle fusion at a given subcellular site involves individual members of the v-SNARE and t-SNARE protein families, sometimes referred to as the "core components" of membrane fusion (Lin and Scheller, 2000).

Initially, the identification of distinct isoforms of these proteins suggested that the specificity of membrane trafficking could be attributable to SNAREs (Waters and Pfeffer, 1999). Subsequently, however, the observation that some SNAREs can function promiscuously at multiple transport steps (Sevilla et al., 1997; Yang et al., 1999) cast

doubt on this notion, and led to the search for other proteins that might mediate the observed specificity of membrane targeting. A number of proteins and protein complexes have since been identified as tethering factors, because they interact with integral membrane proteins on vesicle and target membranes to physically link the separate compartments. Tethering components, or "accessory factors", that function in different parts of the transport pathway are structurally unrelated, suggesting that they could be responsible for the specificity observed in membrane targeting (Guo et al., 2000; Waters and Pfeffer, 1999).

In *S. cerevisiae*, fusion of ER-derived (COPII) vesicles with the cis-Golgi membranes requires at least two tethering factors, Uso1p and Sec35p (Nakajima et al., 1991; Wuestehube et al., 1996). Uso1p, the orthologue of mammalian p115, is a homodimer which self-associates via a long coiled coil region (Sapperstein et al., 1995; Waters et al., 1992) . Uso1p and Sec35p are both required for vesicle tethering, which precedes membrane fusion. In addition, Ypt1p (Rab1) association with the membrane is necessary for recruitment of Uso1p, suggesting that the Rab GTPase regulates Uso1p function (Cao et al., 1998). The mechanism by which Rab1 mediates tethering of COPII vesicles has been further elucidated in mammals, where it appears that Rab1 is activated by proteins on the *cis*-Golgi prior to p115 entry into the complex (Moyer et al., 2001).

Membrane tethering in intra-Golgi transport also requires p115 (Uso1p), though possibly in a different context (Waters et al., 1992). In this example, the amino terminus of p115 binds to Giantin, a large integral membrane protein in COPI vesicles, and the carboxyl terminus of p115 binds to GM130 (Figure 3B). GM130 is a peripheral Golgi protein which is in turn anchored to the Golgi through an interaction with GRASP65,

another peripheral Golgi protein which is myristoylated and anchored in the membrane via this lipid moiety (Barr et al., 1998; Sonnichsen et al., 1998). In this manner, p115 and GM130 tether COPI membranes to the Golgi apparatus and facilitate membrane fusion.

The p115-mediated tethering of COPI vesicles to Golgi membranes has important implications for the structural maintenance of the Golgi, which is a mitotically regulated process (Shorter and Warren, 1999). In animal cells, the Golgi apparatus appears as a ribbon-like structure adjacent to the nucleus and is composed of multiple stacked cisternae. At the onset of mitosis, however, this structure undergoes a dramatic fragmentation of its interphase morphology and is converted into a large number of small vesicles (Lucocq et al., 1989; Misteli and Warren, 1995); this process coincides with a global inhibition of membrane trafficking (Warren, 1993).

The observation that the majority of the small vesicles that are formed during mitotic fragmentation are COPI vesicles (Misteli and Warren, 1995) led to the suggestion that mitotic fragmentation occurs due to inhibition of vesicle fusion, while vesicle formation proceeds normally. This theory gained support when it was determined that GM130 is phosphorylated by the G_2/M specific kinase CDK1 and that p115 is unable to bind to phosphorylated GM130 (Lowe et al., 1998; Nakamura et al., 1997). Elucidation of the precise role of each of the tethering proteins in this process is complicated by their involvement in multiple reactions that mediate membrane recognition. Nevertheless, these examples emphasize the importance of the regulated assembly of peripheral membrane proteins for the maintenance of Golgi structure and the control of membrane trafficking.

The aim of this chapter is to provide a framework for the consideration of the experimental data presented in the following sections. Chapter 2, below, describes the cloning and initial characterization of *GRP50*, a highly conserved mammalian gene encoding a novel proline-rich, ankyrin repeat protein. The data presented in Chapter 3 demonstrate the Brefeldin A-sensitive association of this protein with the cytoplasmic face of Golgi and vesicle membranes in cultured cells, and suggest the possibility of a GRP50-containing multiprotein complex.

Figure 1

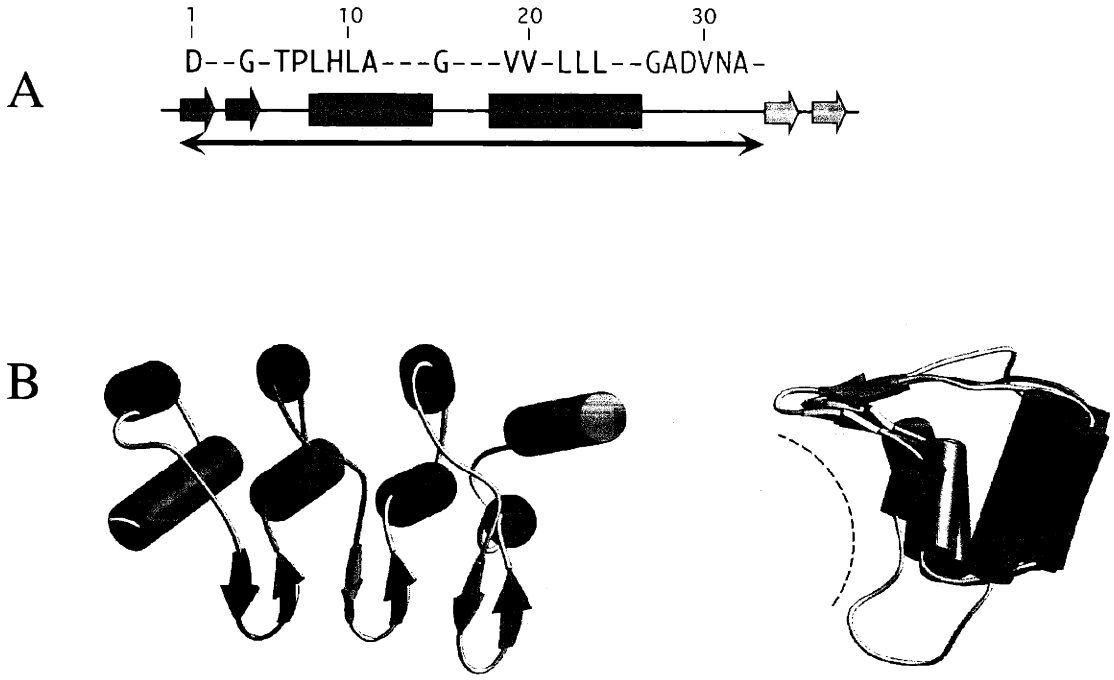


Figure 1. ***(A)*** The ANK-repeat consensus sequence aligned above the β2α2 structure. β-hairpins are depicted as arrows, and α-helical regions are depicted as rectangles. A single ANK repeat is denoted by the lower (black) arrow. ***(B)*** Structure of a four-ANK repeat domain, showing the arrangement of α-helices (cylinders) and β-hairpins (arrows). α-helices align in an antiparallel fashion to form a bundle, and β-hairpins project away from the helical bundle. A single ANK repeat is highlighted in red. Left panel: view from "top" of domain. Right panel: view from side, showing the "cupped hand" (see text). The dotted line in the right panel indicates the binding surface of the domain. Adapted from Sedgwick and Smerdon, 1999.

Figure 2

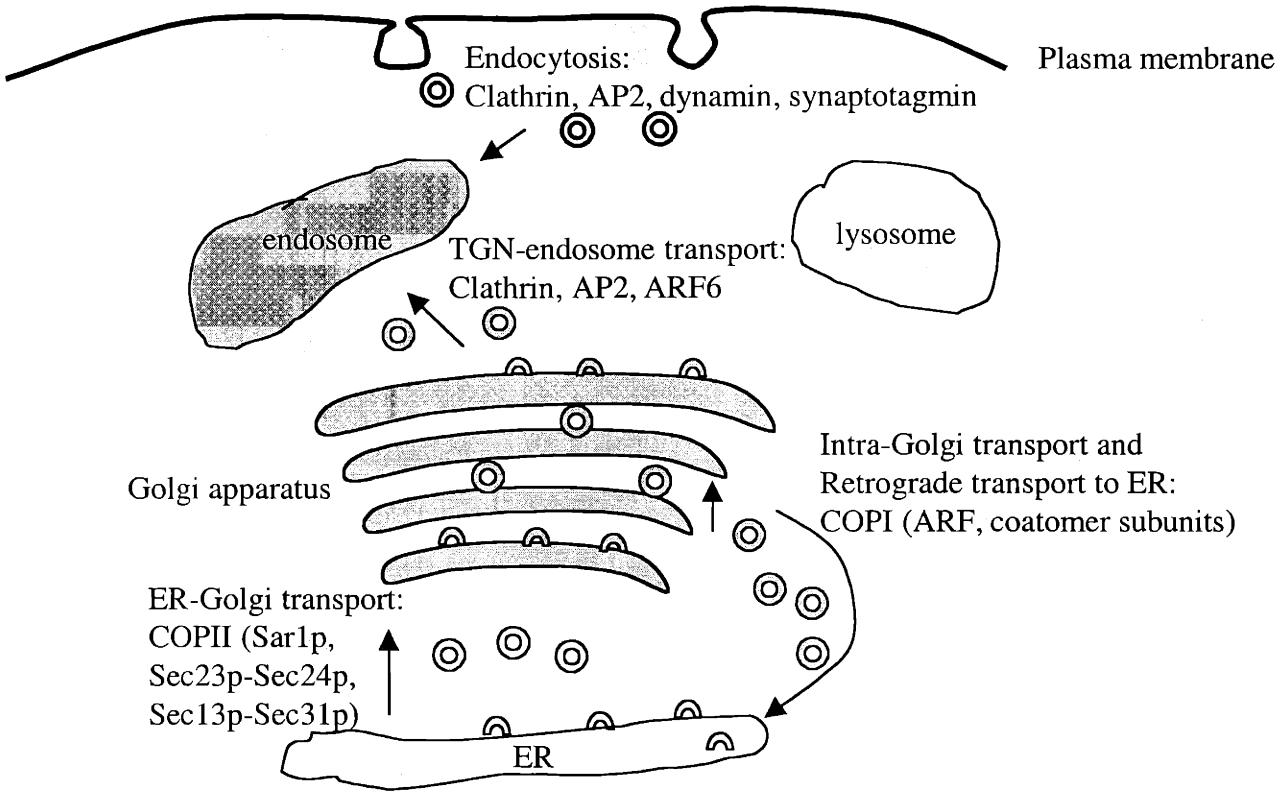


Figure 2. Schematic diagram of membrane transport in eukaryotic cells, indicating protein complexes involved in vesicle *formation* at various stages of the secretory or endocytic pathway.

Figure 3

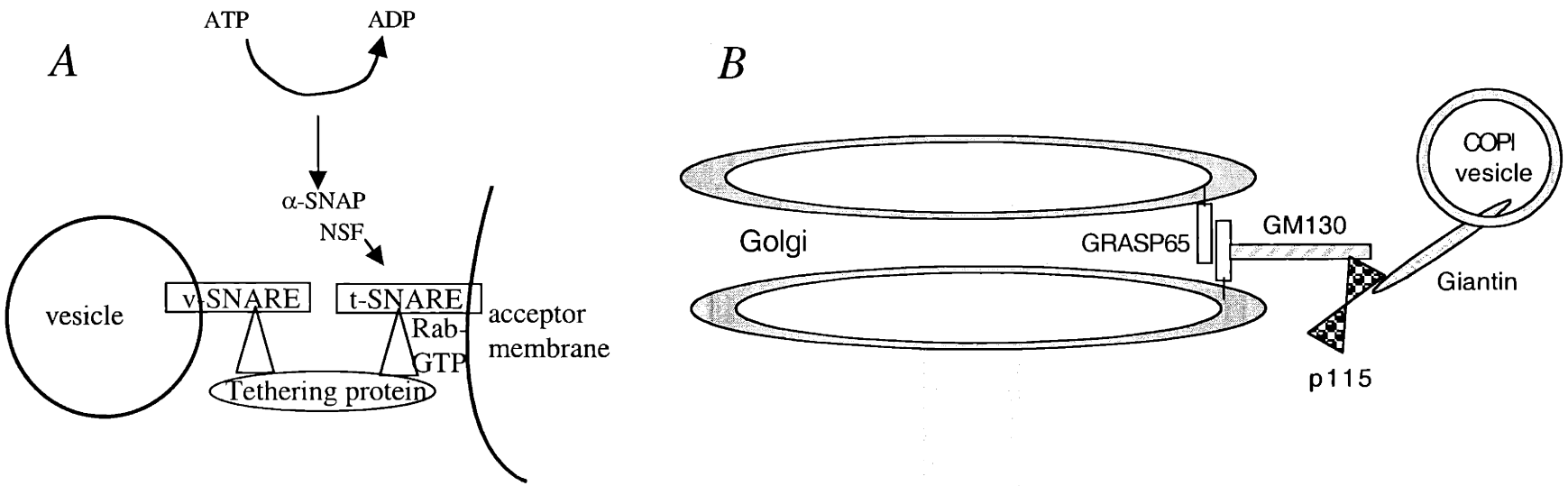


Figure 3. ***(A)*** The general scheme of protein complexes required for *fusion* of transport vesicles. α-SNAP and NSF are recruited to the membrane by interaction with a SNARE protein, resulting in activation of the Rab GTPase. Rab-GTP recruits a specific tethering protein, which provides a physical link between the membranes. ***(B)*** Example of a tethering complex. In ER-Golgi transport, p115 tethers COPI vesicles to the Golgi membranes via interactions with Giantin and GM130. Adapted from Guo et al., 2000.

REFERENCES

Altschul, S.F., T.L. Madden, A.A. Schaffer, J. Zhang, Z. Zhang, W. Miller, and D.J. Lipman. 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* 25:3389-3402.

Andrade, M.A., C. Perez-Iratxeta, and C.P. Ponting. 2001. Protein repeats: structures, functions, and evolution. *J Struct Biol.* 134:117-131.

Axton, J.M., F.L. Shamanski, L.M. Young, D.S. Henderson, J.B. Boyd, and T.L. Orr-Weaver. 1994. The inhibitor of DNA replication encoded by the Drosophila gene plutonium is a small, ankyrin repeat protein. *Embo J.* 13:462-470.

Bar-Sagi, D., D. Rotin, A. Batzer, V. Mandiyan, and J. Schlessinger. 1993. SH3 domains direct cellular localization of signaling molecules. *Cell.* 74:83-91.

Barr, F.A., N. Nakamura, and G. Warren. 1998. Mapping the interaction between GRASP65 and GM130, components of a protein complex involved in the stacking of Golgi cisternae. *Embo J.* 17:3258-3268.

Batchelor, A.H., D.E. Piper, F.C. de la Brousse, S.L. McKnight, and C. Wolberger. 1998. The structure of GABPalpha/beta: an ETS domain- ankyrin repeat heterodimer bound to DNA. *Science.* 279:1037-1041.

Bateman, A., E. Birney, R. Durbin, S.R. Eddy, R.D. Finn, and E.L. Sonnhammer. 1999. Pfam 3.1: 1313 multiple alignments and profile HMMs match the majority of proteins. *Nucleic Acids Res.* 27:260-262.

Baumgartner, R., C. Fernandez-Catalan, A. Winoto, R. Huber, R.A. Engh, and T.A. Holak. 1998. Structure of human cyclin-dependent kinase inhibitor p19INK4d: comparison to known ankyrin-repeat-containing structures and implications for the dysfunction of tumor suppressor p16INK4a. *Structure.* 6:1279-1290.

Beatus, P., J. Lundkvist, C. Oberg, K. Pedersen, and U. Lendahl. 2001. The origin of the ankyrin repeat region in Notch intracellular domains is critical for regulation of HES promoter activity. *Mech Dev.* 104:3-20.

Bedford, M.T., D.C. Chan, and P. Leder. 1997. FBP WW domains and the Abl SH3 domain bind to a specific class of proline-rich ligands. *Embo J.* 16:2376-2383.

Bork, P. 1993. Hundreds of ankyrin-like repeats in functionally diverse proteins: mobile modules that cross phyla horizontally? *Proteins.* 17:363-374.

Bork, P., and M. Sudol. 1994. The WW domain: a signalling site in dystrophin? *Trends Biochem Sci.* 19:531-533.

Breeden, L., and K. Nasmyth. 1987. Similarity between cell-cycle genes of budding yeast and fission yeast and the Notch gene of Drosophila. *Nature.* 329:651-654.

Brotherton, D.H., V. Dhanaraj, S. Wick, L. Brizuela, P.J. Domaille, E. Volyanik, X. Xu, E. Parisini, B.O. Smith, S.J. Archer, M. Serrano, S.L. Brenner, T.L. Blundell, and E.D. Laue. 1998. Crystal structure of the complex of the cyclin D-dependent kinase Cdk6 bound to the cell-cycle inhibitor p19INK4d. *Nature.* 395:244-250.

Callebaut, I., P. Cossart, and P. Dehoux. 1998. EVH1/WH1 domains of VASP and WASP proteins belong to a large family including Ran-binding domains of the RanBP1 family. *FEBS Lett.* 441:181-185.

Cao, X., N. Ballew, and C. Barlowe. 1998. Initial docking of ER-derived vesicles requires Uso1p and Ypt1p but is independent of SNARE proteins. *Embo J.* 17:2156-2165.

Carpentier, J.L., J.P. Paccaud, P. Gorden, W.J. Rutter, and L. Orci. 1992. Insulin-induced surface redistribution regulates internalization of the insulin receptor and requires its autophosphorylation. *Proc Natl Acad Sci U S A.* 89:162-166.

Chang, S.S., S. Grunder, A. Hanukoglu, A. Rosler, P.M. Mathew, I. Hanukoglu, L. Schild, Y. Lu, R.A. Shimkets, C. Nelson-Williams, B.C. Rossier, and R.P. Lifton. 1996. Mutations in subunits of the epithelial sodium channel cause salt wasting with hyperkalaemic acidosis, pseudohypoaldosteronism type 1. *Nat Genet.* 12:248-253.

Chardin, P., and F. McCormick. 1999. Brefeldin A: the advantage of being uncompetitive. *Cell.* 97:153-155.

Chen, H.I., and M. Sudol. 1995. The WW domain of Yes-associated protein binds a proline-rich ligand that differs from the consensus established for Src homology 3-binding modules. *Proc Natl Acad Sci U S A.* 92:7819-7823.

Cicchetti, P., B.J. Mayer, G. Thiel, and D. Baltimore. 1992. Identification of a protein that binds to the SH3 region of Abl and is similar to Bcr and GAP-rho. *Science.* 257:803-806.

Clague, M.J. 1999. Membrane transport: Take your fusion partners. *Curr Biol.* 9:R258-260.

Cohen, G.B., R. Ren, and D. Baltimore. 1995. Modular binding domains in signal transduction proteins. *Cell.* 80:237-248.

Cosson, P., and F. Letourneur. 1997. Coatomer (COPI)-coated vesicles: role in intracellular transport and protein sorting. *Curr Opin Cell Biol.* 9:484-487.

Damke, H., T. Baba, A.M. van der Bliek, and S.L. Schmid. 1995. Clathrin-independent pinocytosis is induced in cells overexpressing a temperature-sensitive mutant of dynamin. *J Cell Biol.* 131:69-80.

Donaldson, J.G., and C.L. Jackson. 2000. Regulators and effectors of the ARF GTPases. *Curr Opin Cell Biol.* 12:475-482.

Duden, R., M. Hosobuchi, S. Hamamoto, M. Winey, B. Byers, and R. Schekman. 1994. Yeast beta- and beta'-coat proteins (COP). Two coatomer subunits essential for endoplasmic reticulum-to-Golgi protein traffic. *J Biol Chem.* 269:24486-24495.

Faber, P.W., G.T. Barnes, J. Srinidhi, J. Chen, J.F. Gusella, and M.E. MacDonald. 1998. Huntingtin interacts with a family of WW domain proteins. *Hum Mol Genet.* 7:1463-1474.

Feng, S., C. Kasahara, R.J. Rickles, and S.L. Schreiber. 1995. Specific interactions outside the proline-rich core of two classes of Src homology 3 ligands. *Proc Natl Acad Sci U S A.* 92:12408-12415.

Foord, R., I.A. Taylor, S.G. Sedgwick, and S.J. Smerdon. 1999. X-ray structural analysis of the yeast cell cycle regulator Swi6 reveals variations of the ankyrin fold and has implications for Swi6 function. *Nat Struct Biol.* 6:157-165.

Franz, W.M., P. Berger, and J.Y. Wang. 1989. Deletion of an N-terminal regulatory domain of the c-abl tyrosine kinase activates its oncogenic potential. *Embo J.* 8:137-147.

Gaasterland, T., and M. Oprea. 2001. Whole-genome analysis: annotations and updates. *Curr Opin Struct Biol.* 11:377-381.

Goldberg, J. 1999. Structural and functional analysis of the ARF1-ARFGAP complex reveals a role for coatomer in GTP hydrolysis. *Cell.* 96:893-902.

Gorina, S., and N.P. Pavletich. 1996. Structure of the p53 tumor suppressor bound to the ankyrin and SH3 domains of 53BP2. *Science.* 274:1001-1005.

Guo, Q., E. Vasile, and M. Krieger. 1994. Disruptions in Golgi structure and membrane traffic in a conditional lethal mammalian cell mutant are corrected by epsilon-COP. *J Cell Biol.* 125:1213-1224.

Guo, W., M. Sacher, J. Barrowman, S. Ferro-Novick, and P. Novick. 2000. Protein complexes in transport vesicle targeting. *Trends Cell Biol.* 10:251-255.

Hansson, J.H., L. Schild, Y. Lu, T.A. Wilson, I. Gautschi, R. Shimkets, C. Nelson-Williams, B.C. Rossier, and R.P. Lifton. 1995. A de novo missense mutation of the beta subunit of the epithelial sodium channel causes hypertension and Liddle syndrome, identifying a proline-rich segment critical for regulation of channel activity. *Proc Natl Acad Sci U S A.* 92:11495-11499.

Haskill, S., A.A. Beg, S.M. Tompkins, J.S. Morris, A.D. Yurochko, A. Sampson-Johannes, K. Mondal, P. Ralph, and A.S. Baldwin, Jr. 1991. Characterization of an immediate-early gene induced in adherent monocytes that encodes I kappa B-like activity. *Cell.* 65:1281-1289.

Henikoff, S., J.G. Henikoff, and S. Pietrokovski. 1999. Blocks+: a non-redundant database of protein alignment blocks derived from multiple compilations. *Bioinformatics.* 15:471-479.

Hinshaw, J.E. 2000. Dynamin and its role in membrane fission. *Annu Rev Cell Dev Biol.* 16:483-519.

Hirst, J., N.A. Bright, B. Rous, and M.S. Robinson. 1999. Characterization of a fourth adaptor-related protein complex. *Mol Biol Cell.* 10:2787-2802.

Holt, M.R., and A. Koffer. 2001. Cell motility: proline-rich proteins promote protrusions. *Trends Cell Biol.* 11:38-46.

Huang, X., F. Poy, R. Zhang, A. Joachimiak, M. Sudol, and M.J. Eck. 2000. Structure of a WW domain containing fragment of dystrophin in complex with beta-dystroglycan. *Nat Struct Biol.* 7:634-638.

Huxford, T., D.B. Huang, S. Malek, and G. Ghosh. 1998. The crystal structure of the IkappaBalpha/NF-kappaB complex reveals mechanisms of NF-kappaB inactivation. *Cell.* 95:759-770.

Iwabuchi, K., P.L. Bartel, B. Li, R. Marraccino, and S. Fields. 1994. Two cellular proteins that bind to wild-type but not mutant p53. *Proc Natl Acad Sci U S A.* 91:6098-6102.

Jacobs, M.D., and S.C. Harrison. 1998. Structure of an IkappaBalpha/NF-kappaB complex. *Cell.* 95:749-758.

Jeffrey, P.D., L. Tong, and N.P. Pavletich. 2000. Structural basis of inhibition of CDK-cyclin complexes by INK4 inhibitors. *Genes Dev.* 14:3115-3125.

Kay, B.K., M.P. Williamson, and M. Sudol. 2000. The importance of being proline: the interaction of proline-rich motifs in signaling proteins with their cognate domains. *Faseb J.* 14:231-241.

Klumperman, J. 2000. Transport between ER and Golgi. *Curr Opin Cell Biol.* 12:445-449.

Kreis, T.E., M. Lowe, and R. Pepperkok. 1995. COPs regulating membrane traffic. *Annu Rev Cell Dev Biol.* 11:677-706.

LaMarco, K.L., and S.L. McKnight. 1989. Purification of a set of cellular polypeptides that bind to the purine- rich cis-regulatory element of herpes simplex virus immediate early genes. *Genes Dev.* 3:1372-1383.

Laurent, V., T.P. Loisel, B. Harbeck, A. Wehman, L. Grobe, B.M. Jockusch, J. Wehland, F.B. Gertler, and M.F. Carlier. 1999. Role of proteins of the Ena/VASP family in actin-based motility of Listeria monocytogenes. *J Cell Biol.* 144:1245-1258.

Lemmon, S.K., and L.M. Traub. 2000. Sorting in the endosomal system in yeast and animal cells. *Curr Opin Cell Biol.* 12:457-466.

Letourneur, F., E.C. Gaynor, S. Hennecke, C. Demolliere, R. Duden, S.D. Emr, H. Riezman, and P. Cosson. 1994. Coatomer is essential for retrieval of dilysine-tagged proteins to the endoplasmic reticulum. *Cell.* 79:1199-1207.

Li, J., I.J. Byeon, K. Ericson, M.J. Poi, P. O'Maille, T. Selby, and M.D. Tsai. 1999. Tumor suppressor INK4: determination of the solution structure of p18INK4C and demonstration of the functional significance of loops in p18INK4C and p16INK4A. *Biochemistry.* 38:2930-2940.

Li, T., L.O. Narhi, J. Wen, J.S. Philo, K. Sitney, J. Inoue, T. Yamamoto, and T. Arakawa. 1998. Interactions between NFkappaB and its inhibitor ikappaB: biophysical characterization of a NFkappaB/ikappaB-alpha complex. *J Protein Chem.* 17:757-763.

Lim, W.A., F.M. Richards, and R.O. Fox. 1994. Structural determinants of peptide-binding orientation and of sequence specificity in SH3 domains. *Nature.* 372:375-379.

Lin, R.C., and R.H. Scheller. 2000. Mechanisms of synaptic vesicle exocytosis. *Annu Rev Cell Dev Biol.* 16:19-49.

Linn, H., K.S. Ermekova, S. Rentschler, A.B. Sparks, B.K. Kay, and M. Sudol. 1997. Using molecular repertoires to identify high-affinity peptide ligands of the WW domain of human and mouse YAP. *Biol Chem.* 378:531-537.

Lippincott-Schwartz, J., N.B. Cole, and J.G. Donaldson. 1998. Building a secretory apparatus: role of ARF1/COPI in Golgi biogenesis and maintenance. *Histochem Cell Biol.* 109:449-462.

Lowe, M., C. Rabouille, N. Nakamura, R. Watson, M. Jackman, E. Jamsa, D. Rahman, D.J. Pappin, and G. Warren. 1998. Cdc2 kinase directly phosphorylates the cis-Golgi matrix protein GM130 and is required for Golgi fragmentation in mitosis. *Cell.* 94:783-793.

Lucocq, J.M., E.G. Berger, and G. Warren. 1989. Mitotic Golgi fragments in HeLa cells and their role in the reassembly pathway. *J Cell Biol.* 109:463-474.

Lux, S.E., K.M. John, and V. Bennett. 1990. Analysis of cDNA for human erythrocyte ankyrin indicates a repeated structure with homology to tissue-differentiation and cell-cycle control proteins. *Nature.* 344:36-42.

Macias, M.J., M. Hyvonen, E. Baraldi, J. Schultz, M. Sudol, M. Saraste, and H. Oschkinat. 1996. Structure of the WW domain of a kinase-associated protein complexed with a proline-rich peptide. *Nature.* 382:646-649.

Mayer, B.J. 2001. SH3 domains: complexity in moderation. *J Cell Sci.* 114:1253-1263.

Mayer, B.J., and M.J. Eck. 1995. SH3 domains. Minding your p's and q's. *Curr Biol.* 5:364-367.

Mayer, B.J., and H. Hanafusa. 1990. Mutagenic analysis of the v-crk oncogene: requirement for SH2 and SH3 domains and correlation between increased cellular phosphotyrosine and transformation. *J Virol.* 64:3581-3589.

McCormick, F. 1993. Signal transduction. How receptors turn Ras on. *Nature.* 363:15-16.

McPherson, P.S. 1999. Regulatory role of SH3 domain-mediated protein-protein interactions in synaptic vesicle endocytosis. *Cell Signal.* 11:229-238.

Michaely, P., and V. Bennett. 1993. The membrane-binding domain of ankyrin contains four independently folded subdomains, each comprised of six ankyrin repeats. *J Biol Chem.* 268:22703-22709.

Misteli, T., and G. Warren. 1995. Mitotic disassembly of the Golgi apparatus in vivo. *J Cell Sci.* 108:2715-2727.

Mochida, S. 2000. Protein-protein interactions in neurotransmitter release. *Neurosci Res.* 36:175-182.

Mostov, K.E., M. Verges, and Y. Altschuler. 2000. Membrane traffic in polarized epithelial cells. *Curr Opin Cell Biol.* 12:483-490.

Moyer, B.D., B.B. Allan, and W.E. Balch. 2001. Rab1 interaction with a GM130 effector complex regulates COPII vesicle cis--Golgi tethering. *Traffic.* 2:268-276.

Musacchio, A., M. Saraste, and M. Wilmanns. 1994. High-resolution crystal structures of tyrosine kinase SH3 domains complexed with proline-rich peptides. *Nat Struct Biol.* 1:546-551.

Nakajima, H., A. Hirata, Y. Ogawa, T. Yonehara, K. Yoda, and M. Yamasaki. 1991. A cytoskeleton-related gene, uso1, is required for intracellular protein transport in Saccharomyces cerevisiae. *J Cell Biol.* 113:245-260.

Nakamura, N., M. Lowe, T.P. Levine, C. Rabouille, and G. Warren. 1997. The vesicle docking protein p115 binds GM130, a cis-Golgi matrix protein, in a mitotically regulated manner. *Cell.* 89:445-455.

Niebuhr, K., F. Ebel, R. Frank, M. Reinhard, E. Domann, U.D. Carl, U. Walter, F.B. Gertler, J. Wehland, and T. Chakraborty. 1997. A novel proline-rich motif present in ActA of Listeria monocytogenes and cytoskeletal proteins is the ligand for the EVH1 domain, a protein module present in the Ena/VASP family. *Embo J.* 16:5433-5444.

Orci, L., M. Stamnes, M. Ravazzola, M. Amherdt, A. Perrelet, T.H. Sollner, and J.E. Rothman. 1997. Bidirectional transport by distinct populations of COPI-coated vesicles. *Cell.* 90:335-349.

Palade, G. 1975. Intracellular aspects of the process of protein synthesis. *Science.* 189:347-358.

Pearse, B.M. 1988. Receptors compete for adaptors found in plasma membrane coated pits. *Embo J.* 7:3331-3336.

Petcherski, A.G., and J. Kimble. 2000. Mastermind is a putative activator for Notch. *Curr Biol.* 10:R471-473.

Reiersen, H., and A.R. Rees. 2001. The hunchback and its neighbours: proline as an environmental modulator. *Trends Biochem Sci.* 26:679-684.

Ren, R., B.J. Mayer, P. Cicchetti, and D. Baltimore. 1993. Identification of a ten-amino acid proline-rich SH3 binding site. *Science.* 259:1157-1161.

Rentschler, S., H. Linn, K. Deininger, M.T. Bedford, X. Espanel, and M. Sudol. 1999. The WW domain of dystrophin requires EF-hands region to interact with beta-dystroglycan. *Biol Chem.* 380:431-442.

Rickles, R.J., M.C. Botfield, Z. Weng, J.A. Taylor, O.M. Green, J.S. Brugge, and M.J. Zoller. 1994. Identification of Src, Fyn, Lyn, PI3K and Abl SH3 domain ligands using phage display libraries. *Embo J.* 13:5598-5604.

Robinson, M.S. 1994. The role of clathrin, adaptors and dynamin in endocytosis. *Curr Opin Cell Biol.* 6:538-544.

Roth, M.G. 1999. Snapshots of ARF1: implications for mechanisms of activation and inactivation. *Cell.* 97:149-152.

Rothman, J.E. 1994. Mechanisms of intracellular protein transport. *Nature*. 372:55-63.

Rubin, G.M., and E.B. Lewis. 2000. A brief history of Drosophila's contributions to genome research. *Science*. 287:2216-2218.

Russo, A.A., L. Tong, J.O. Lee, P.D. Jeffrey, and N.P. Pavletich. 1998. Structural basis for inhibition of the cyclin-dependent kinase Cdk6 by the tumour suppressor p16INK4a. *Nature*. 395:237-243.

Sapperstein, S.K., D.M. Walter, A.R. Grosvenor, J.E. Heuser, and M.G. Waters. 1995. p115 is a general vesicular transport factor related to the yeast endoplasmic reticulum to Golgi transport factor Uso1p. *Proc Natl Acad Sci U S A*. 92:522-526.

Schekman, R., and L. Orci. 1996. Coat proteins and vesicle budding. *Science*. 271:1526-1533.

Schild, L., Y. Lu, I. Gautschi, E. Schneeberger, R.P. Lifton, and B.C. Rossier. 1996. Identification of a PY motif in the epithelial Na channel subunits as a target sequence for mutations causing channel activation found in Liddle syndrome. *Embo J.* 15:2381-2387. abs.html.

Sedgwick, S.G., and S.J. Smerdon. 1999. The ankyrin repeat: a diversity of interactions on a common structural framework. *Trends Biochem Sci.* 24:311-316.

Seidel-Dugan, C., B.E. Meyer, S.M. Thomas, and J.S. Brugge. 1992. Effects of SH2 and SH3 deletions on the functional activities of wild- type and transforming variants of c-Src. *Mol Cell Biol.* 12:1835-1845.

Serafini, T., G. Stenbeck, A. Brecht, F. Lottspeich, L. Orci, J.E. Rothman, and F.T. Wieland. 1991. A coat subunit of Golgi-derived non-clathrin-coated vesicles with homology to the clathrin-coated vesicle coat protein beta-adaptin. *Nature*. 349:215-220.

Sevilla, L., E. Tomas, P. Munoz, A. Guma, Y. Fischer, J. Thomas, B. Ruiz-Montasell, X. Testar, M. Palacin, J. Blasi, and A. Zorzano. 1997. Characterization of two distinct intracellular GLUT4 membrane populations in muscle fiber. Differential protein composition and sensitivity to insulin. *Endocrinology*. 138:3006-3015.

Shorter, J., and G. Warren. 1999. A role for the vesicle tethering protein, p115, in the post-mitotic stacking of reassembling Golgi cisternae in a cell-free system. *J Cell Biol.* 146:57-70.

Simeonidis, S., D. Stauber, G. Chen, W.A. Hendrickson, and D. Thanos. 1999. Mechanisms by which IkappaB proteins control NF-kappaB activity. *Proc Natl Acad Sci U S A*. 96:49-54.

Simpson, F., N.A. Bright, M.A. West, L.S. Newman, R.B. Darnell, and M.S. Robinson. 1996. A novel adaptor-related protein complex. *J Cell Biol.* 133:749-760.

Simpson, F., A.A. Peden, L. Christopoulou, and M.S. Robinson. 1997. Characterization of the adaptor-related protein complex, AP-3. *J Cell Biol.* 137:835-845.

Sollner, T., M.K. Bennett, S.W. Whiteheart, R.H. Scheller, and J.E. Rothman. 1993a. A protein assembly-disassembly pathway in vitro that may correspond to sequential steps of synaptic vesicle docking, activation, and fusion. *Cell.* 75:409-418.

Sollner, T., S.W. Whiteheart, M. Brunner, H. Erdjument-Bromage, S. Geromanos, P. Tempst, and J.E. Rothman. 1993b. SNAP receptors implicated in vesicle targeting and fusion. *Nature.* 362:318-324.

Sonnichsen, B., M. Lowe, T. Levine, E. Jamsa, B. Dirac-Svejstrup, and G. Warren. 1998. A role for giantin in docking COPI vesicles to Golgi membranes. *J Cell Biol.* 140:1013-1021.

Spang, A., K. Matsuoka, S. Hamamoto, R. Schekman, and L. Orci. 1998. Coatomer, Arf1p, and nucleotide are required to bud coat protein complex I-coated vesicles from large synthetic liposomes. *Proc Natl Acad Sci U S A.* 95:11199-11204.

Sparks, A.B., J.E. Rider, N.G. Hoffman, D.M. Fowlkes, L.A. Quillam, and B.K. Kay. 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.* 93:1540-1544.

Stahl, M.L., C.R. Ferenz, K.L. Kelleher, R.W. Kriz, and J.L. Knopf. 1988. Sequence similarity of phospholipase C with the non-catalytic region of src. *Nature.* 332:269-272.

Stapley, B.J., and T.P. Creamer. 1999. A survey of left-handed polyproline II helices. *Protein Sci.* 8:587-595.

Staub, O., S. Dho, P. Henry, J. Correa, T. Ishikawa, J. McGlade, and D. Rotin. 1996. WW domains of Nedd4 bind to the proline-rich PY motifs in the epithelial Na+ channel deleted in Liddle's syndrome. *Embo J.* 15:2371-2380.

Staub, O., I. Gautschi, T. Ishikawa, K. Breitschopf, A. Ciechanover, L. Schild, and D. Rotin. 1997a. Regulation of stability and function of the epithelial Na+ channel (ENaC) by ubiquitination. *Embo J.* 16:6325-6336.

Staub, O., H. Yeger, P.J. Plant, H. Kim, S.A. Ernst, and D. Rotin. 1997b. Immunolocalization of the ubiquitin-protein ligase Nedd4 in tissues expressing the epithelial Na+ channel (ENaC). *Am J Physiol.* 272:C1871-1880.

Sudol, M. 1996. Structure and function of the WW domain. *Prog Biophys Mol Biol.* 65:113-132.

Sudol, M. 1998. From Src Homology domains to other signaling modules: proposal of the 'protein recognition code'. *Oncogene* 17:1469-74.

Takei, K., P.S. McPherson, S.L. Schmid, and P. De Camilli. 1995. Tubular membrane invaginations coated by dynamin rings are induced by GTP-gamma S in nerve terminals. *Nature*. 374:186-190.

Tani, S., H. Kurooka, T. Aoki, N. Hashimoto, and T. Honjo. 2001. The N- and C-terminal regions of RBP-J interact with the ankyrin repeats of Notch1 RAMIC to activate transcription. *Nucleic Acids Res.* 29:1373-1380.

Teichmann, S.A., A.G. Murzin, and C. Chothia. 2001. Determination of protein function, evolution and interactions by structural genomics. *Curr Opin Struct Biol.* 11:354-363.

Thompson, C.C., T.A. Brown, and S.L. McKnight. 1991. Convergence of Ets- and notch-related structural motifs in a heteromeric DNA binding complex. *Science*. 253:762-768.

Traub, L.M., and S. Kornfeld. 1997. The trans-Golgi network: a late secretory sorting station. *Curr Opin Cell Biol.* 9:527-533.

Verdecia, M.A., M.E. Bowman, K.P. Lu, T. Hunter, and J.P. Noel. 2000. Structural basis for phosphoserine-proline recognition by group IV WW domains. *Nat Struct Biol.* 7:639-643.

von Poser, C., J.Z. Zhang, C. Mineo, W. Ding, Y. Ying, T.C. Sudhof, and R.G. Anderson. 2000. Synaptotagmin regulation of coated pit assembly. *J Biol Chem.* 275:30916-30924.

Warren, G. 1993. Membrane partitioning during cell division. *Annu Rev Biochem.* 62:323-348.

Waters, M.G., D.O. Clary, and J.E. Rothman. 1992. A novel 115-kD peripheral membrane protein is required for intercisternal transport in the Golgi stack. *J Cell Biol.* 118:1015-1026.

Waters, M.G., and S.R. Pfeffer. 1999. Membrane tethering in intracellular transport. *Curr Opin Cell Biol.* 11:453-459.

Waters, M.G., T. Serafini, and J.E. Rothman. 1991. 'Coatomer': a cytosolic protein complex containing subunits of non- clathrin-coated Golgi transport vesicles. *Nature.* 349:248-251.

Wharton, K.A., B. Yedvobnick, V.G. Finnerty, and S. Artavanis-Tsakonas. 1985. opa: a novel family of transcribed repeats shared by the Notch locus and other developmentally regulated loci in D. melanogaster. *Cell.* 40:55-62.

Williams, J.C., A. Weijland, S. Gonfloni, A. Thompson, S.A. Courtneidge, G. Superti-Furga, and R.K. Wierenga. 1997. The 2.35 A crystal structure of the inactivated form of chicken Src: a dynamic molecule with multiple regulatory interactions. *J Mol Biol.* 274:757-775.

Williams, J.C., R.K. Wierenga, and M. Saraste. 1998. Insights into Src kinase functions: structural comparisons. *Trends Biochem Sci.* 23:179-184.

Wu, Y., H. Luo, N. Kanaan, and J. Wu. 2000. The proteasome controls the expression of a proliferation-associated nuclear antigen Ki-67. *J Cell Biochem.* 76:596-604.

Wuestehube, L.J., R. Duden, A. Eun, S. Hamamoto, P. Korn, R. Ram, and R. Schekman. 1996. New mutants of Saccharomyces cerevisiae affected in the transport of proteins from the endoplasmic reticulum to the Golgi complex. *Genetics.* 142:393-406.

Yagi, R., L.F. Chen, K. Shigesada, Y. Murakami, and Y. Ito. 1999. A WW domain-containing yes-associated protein (YAP) is a novel transcriptional co-activator. *Embo J.* 18:2551-2562.

Yang, B., L. Gonzalez, Jr., R. Prekeris, M. Steegmaier, R.J. Advani, and R.H. Scheller. 1999. SNARE interactions are not selective. Implications for membrane fusion specificity. *J Biol Chem.* 274:5649-5653.

Yochem, J., K. Weston, and I. Greenwald. 1988. The Caenorhabditis elegans lin-12 gene encodes a transmembrane protein with overall similarity to Drosophila Notch. *Nature.* 335:547-550.

Yu, H., J.K. Chen, S. Feng, D.C. Dalgarno, A.W. Brauer, and S.L. Schreiber. 1994. Structural basis for the binding of proline-rich peptides to SH3 domains. *Cell.* 76:933-945.

Zarrinpar, A., and W.A. Lim. 2000. Converging on proline: the mechanism of WW domain peptide recognition. *Nat Struct Biol.* 7:611-613.

Zhang, B., and Z. Peng. 2000. A minimum folding unit in the ankyrin repeat protein p16(INK4). *J Mol Biol.* 299:1121-1132.

Zhang, J.Z., B.A. Davletov, T.C. Sudhof, and R.G. Anderson. 1994. Synaptotagmin I is a high affinity receptor for clathrin AP-2: implications for membrane recycling. *Cell.* 78:751-760.

CHAPTER 2

Identification and Characterization of A Novel Ankyrin Repeat Protein

INTRODUCTION

Many biological processes are mediated by specific protein-protein interactions, often occurring over relatively short regions of modular protein domains. Temporal and spatial regulation of transient protein-protein interactions permits exquisite control over cellular responses to external stimuli. Over the past decade, the rapid expansion of protein sequence and structure databases has revealed an increasingly diverse multitude of conserved sequence elements that have been shown to mediate such interactions. These include short sequence motifs of a few amino acids, longer conserved protein domains, and repeat elements consisting of multiple tandem copies of a particular sequence motif. In some cases, the identification of a conserved protein-interaction domain in a novel sequence can provide clues to the function of the new protein, particularly when the domain is known to bind only one or a few types of ligands. Elucidation of biological function is more complicated, however, when a novel protein contains an interaction domain that is formally capable of binding many different types of molecules.

One such protein binding domain is the ankyrin repeat, a 33 amino acid motif that was first described in the sequences of several cell-cycle regulators in yeast and in tissue differentiation factors in *Drosophila* and *C. elegans* (Breeden and Nasmyth, 1987), and later found in the cytoskeletal protein ankyrin (Lux et al., 1990). To date, ankyrin repeats have been detected in more than one thousand proteins from both prokaryotes and eukaryotes, with each protein containing as few as two to as many as twenty copies (Bork, 1993; Michaely and Bennett, 1993). Each repeat adopts a β-hairpin-helix-loop-

helix ($\beta_2\alpha_2$) structure; structural studies have revealed that this conformation can tolerate considerable sequence variation in the residues that directly mediate protein binding (Bork, 1993; Sedgwick and Smerdon, 1999). Perhaps as a result of this tolerance for sequence variation, ankyrin repeat proteins have been shown to interact with a widely divergent set of proteins and hence to perform a number of distinct biological roles. Examples of ankyrin repeat proteins with known biological functions include the Swi6 and GABPα-GAPBβ transcription factors (Batchelor et al., 1998; Foord et al., 1999; Thompson et al., 1991), the cyclin-dependent kinase inhibitors encoded by the *INK4* locus (Baumgartner et al., 1998; Li et al., 1999; Luh et al., 1997; Russo et al., 1998), and the inhibitor of the NF-κB transcription factor, I-κBα (Huxford et al., 1998; Jacobs and Harrison, 1998).

Among short motifs that mediate protein-protein interactions, proline-rich sequences are perhaps the most common. Short proline-rich sequences mediate interactions with a wide variety of proteins, and are the preferred ligands for several modular protein interaction motifs, including Src homology 3 (SH3), WW, and EVH1 domains (Holt and Koffer, 2001; Kay et al., 2000; Mayer, 2001; Sudol, 1996). SH3 domains have critical functions in a large number of eukaryotic signal transduction and cytoskeletal proteins, including the Src family of tyrosine kinases, the Ras effector Grb2, the regulatory subunits of phosphatidylinositol 3-kinase (PI3K) and phospholipase Cγ (PLCγ), and the actin binding protein α-spectrin (Bar-Sagi et al., 1993; Franz et al., 1989; McCormick, 1993; Seidel-Dugan et al., 1992). WW domains are also found in a number of signaling and cytoskeletal proteins, where they act to target their proline-rich ligands to distinct subcellular destinations (Bork and Sudol, 1994; Staub et al., 1996; Staub et al.,

1997; Sudol et al., 2001). EVH1 (Ena/VASP Homology 1) domains are found in proteins of the Ena/VASP superfamily, which are believed to modulate cellular motility and axon outgrowth by regulating the dynamics of the actin cytoskeleton (Bear et al., 2001; Callebaut et al., 1998; Goldberg et al., 2000; Korey and Van Vactor, 2000; Machesky, 1997).

The biochemistry of proline (with its cyclized side chain) restricts the conformation of proline-rich sequences. These sequences tend to adopt a helical structure with three residues per turn that has been termed the polyproline-2 (PPII) helix (Kay et al., 2000; Reiersen and Rees, 2001). The binding of the PPII helix to SH3 domains is largely mediated by hydrophobic interactions, which result in relatively low affinity binding that can be rapidly modulated (Kay et al., 2000). The above examples illustrate the importance of proline-rich sequences as intermolecular ligands, and illustrate the utility of these sequences in situations requiring the rapid response of cellular processes to external signals.

The types of signals mediated by transient protein-protein interactions can exert broad physiological consequences, as is illustrated by the examples given above. We have identified a novel gene, encoding a protein that contains ankyrin repeats and a polyproline-rich domain, in the course of screening a murine cDNA library to identify evolutionarily conserved genes that promote apoptosis. Apoptosis, or programmed cell death, is an important physiological process in many multicellular organisms. Apoptosis is a key component of the processes of tissue remodeling and neural patterning during development, and mediates tissue homeostasis in adult organisms. In addition, apoptosis can be induced in many cell types by exposure to environmental mutagens that cause

genotoxic stress (Raff, 1992; Vaux and Weissman, 1993). The first pro-apoptotic genes were identified in the nematode *C. elegans* and have since been shown to have orthologues in mammalian systems (Hengartner and Horvitz, 1994; Miura et al., 1993).

In the experiments described below, we utilized low-stringency hybridization techniques to screen for murine homologues of the *D. melanogaster* cell death gene, *reaper* (White et al., 1994; White et al., 1996) One of the cDNA clones obtained in this screen represents a novel mammalian gene, which we have called *GRP50. GRP50.*is ubiquitously expressed and does not appear to mediate cell death, but is induced in response to serum stimulation of cell cycle-arrested fibroblasts in culture.

Stimulation of growth-arrested fibroblasts with serum causes cell cycle re-entry and induces expression of many genes involved in cell growth, proliferation, and differentiation (Winkles, 1998). Although this system has been used extensively in the study of genes that regulate cell cycle progression, only a fraction of serum-induced genes directly mediate this process. Many of the genes that are induced in response to serum are involved in diverse cellular processes, such as signal transduction, coagulation, cytoskeletal reorganization, tissue remodeling, and wound healing (Iyer et al., 1999). The expression pattern and structural architecture of the GRP50 protein suggest that it may play a role in the biology of the serum response, through interaction with other cellular proteins.

RESULTS

Cloning of GRP50 cDNA

The GRP50 cDNA was originally isolated in a low-stringency hybridization screen for murine homologues of the *D. melanogaster* cell death gene, *reaper* (see Materials and Methods). Because our goal was to identify genes whose expression is increased in apoptotic cells, we constructed an oligo(dT)-primed cDNA library using RNA isolated from cells induced to undergo apoptosis by exposure to γ-irradiation. The cells used in these experiments were mouse embryonic fibroblasts (MEFs) which had been transformed by stable coexpression of the adenovirus early region 1A gene (*E1A*) and an activated *ras* allele (T24 H-*ras*). These cells, referred to here as E1A/Ras MEFs, readily undergo apoptosis in response to low levels of γ-irradiation and other genotoxic agents (Lowe et al., 1993). In contrast, treatment of normal MEFs with low levels of γ-irradiation causes arrest in the G_1 phase of the cell cycle (Kastan et al., 1992).

Positive clones from the initial cDNA library screening were analyzed by hybridization to poly(A)+ RNA isolated from normal cells (MEFs) and from apoptotic cells (γ-irradiated, E1A/ras MEFs). Preliminary analysis of cDNA clones obtained in this screen revealed that one clone detects messenger RNAs (mRNAs) that appeared to be enriched in oncogenically transformed cells undergoing apoptosis. Figure 1A (left panels) shows northern hybridization of this clone, a 2.2 kb cDNA originally designated ER18 and later named GRP50, to poly(A)+ RNA from MEFs (upper panel, lane 1) and from γ-irradiated, E1A/ras MEFs (lane 2). The ER18 probe detects two mRNA species: one with an apparent molecular mass of ~1.7 kb and another of ~3.5 kb; the 3.5 kb mRNA is more

abundant in γ-irradiated, E1A/ras MEFs than in normal MEFs (compare Figure 1A, lanes 1 and 2). In the lower panel, the blot was reprobed with another cDNA clone from this library (ELP-1, see legend) as a loading control. The ELP-1 probing, together with ethidium bromide staining of the RNA gel (not shown), indicated that the RNA in lane 2 was underloaded, and suggested that the GRP50 transcript was induced two- to four-fold in γ-irradiated, E1A/ras MEFs. The induction of GRP50 in E1A/ras MEFs appeared to be due, at least in part, to the effects of γ-irradiation in these cells (Figure 1A, right panel). Analysis of poly(A)+ RNA from unirradiated (lane 1) and from γ-irradiated (lanes 2 and 3) E1A/ras MEFs indicated that the RNA is induced by 60 minutes following irradiation.

Detection of a 3.5 kb transcript on the northern blots in Figure 1A suggested that the 2.2 kb clone represented a truncated GRP50 cDNA. In addition, sequencing of the 2.2 kb clone revealed a single open reading frame lacking an in-frame stop codon upstream of the first AUG. Because this clone was isolated from an oligo(dT)-primed cDNA library and contained a poly(A) tract, it seemed likely that the 2.2 kb clone lacked the 5-prime end of the longer GRP50 transcript. In order to isolate additional GRP50 cDNAs, the E1A/ras MEF library was rescreened, at high stringency, using the 2.2 kb cDNA as a probe. A schematic representation of the additional cDNA clones obtained is shown in Figure 1B. Surprisingly, restriction mapping and sequencing revealed that two of the new cDNA clones share the same 5-prime end as the original 2.2 kb clone but contain different 3-prime ends. Although each of the cDNAs isolated contains a poly(A) tail, all but one differ in the length of their 3-prime untranslated regions (UTRs). A 3.5 kb cDNA was isolated whose sequence, over the first 2.2 kb, is identical to that of the original 2.2 kb ER18 clone; this larger cDNA contain an additional ~1.2 kb of sequence downstream.

One clone of 1.8 kb (1B, last line), contains only 3-prime UTR sequences, and is identical to the last 1.8 kb of the 3.5 kb cDNA.

Isolation of 1.6 kb and 3.5 kb cDNAs with identical 5-prime and different 3-prime sequence suggested that these clones represent the ~1.7 and ~3.5 kb GRP50 transcripts observed on the northern blot in Figure 1A. To determine whether these transcripts differ in their 3-prime UTRs, two restriction fragments (depicted in red lines, 1B) were used as probes for subsequent northern analysis. One probe corresponds to the 5-prime end of the GRP50 cDNA, and contains a portion of the open reading frame (ORF), and the second probe was derived from the 3-prime-most end of the 2.2 kb ER18 clone. A northern blot containing poly(A)+ RNA was hybridized to each of these probe in succession (Figure 1C). The 5-prime probe detects both the 1.7 and the 3.5 kb transcripts (1C, left panel), but the 3-prime probe detects only the longer transcript (right panel). These results imply that the sole difference between the two observed GRP50 mRNAs is the length of their 3-prime UTRs, and that our 1.6 and 3.5 kb clones represent full-length cDNAs derived from each of these transcripts.

GRP50 mRNA expression profile

Although preliminary experiments suggested that GRP50 mRNA levels are increased in γ-irradiated, E1A/ras MEFs as compared to normal MEFs, further analyses failed to confirm these results. GRP50 mRNA expression is not generally induced in response to γ-irradiation in MEFs (Figure 2A) or in E1A/ras MEFs (Figure 2B), indicating that GRP50 is not a radiation-response gene. In addition, comparison of GRP50 mRNA levels in unstimulated MEFs and E1A/ras MEFs (Figure 2C)

demonstrated that GRP50 was not induced in response to expression of the activating oncogenes, *E1A* and *ras*. Different batches of cells were used in the experiments shown in Figure 2 and in Figure 1A, and the possibility that the cells acquired distinct mutations in culture cannot be excluded; however, these inconsistencies suggest that the induction of GRP50 mRNA observed in our initial experiments is not a general feature of GRP50 regulation.

In the course of investigating the expression profile of GRP50 mRNA in response to radiation and other stimuli, we analyzed the regulation of this gene as a function of cell cycle progression. Figure 3 (top panels) shows the expression of GRP50 in MEFs that have been arrested in the G_0 phase of the cell cycle by serum deprivation, and then stimulated to re-enter the cell cycle (as a synchronous population) by addition of serum. The left and right panels show RNA isolated from two independently derived MEF populations; cells were harvested and RNA isolated at the indicated time points following serum stimulation. A probe derived from the *ARPP PO* gene was used as a loading control for these blots (Figure 3, bottom panels) because expression of this gene does not vary over the course of the cell cycle (Hurford et al., 1997).

The time course of these experiments (8-24 hours post serum stimulation) corresponds to the G_1 and S phases of the cell cycle; ^{3}H-thymidine incorporation assays demonstrate that these cells are in S phase between ~16 and ~20 hours following serum addition (Humbert et al., 2000). Surprisingly, GRP50 is expressed at very low levels in quiescent cells, and is induced at least 3-5 fold in serum-stimulated cells; this induction is evident by 8 hours following serum stimulation and remains constant over the next 16 hours. These results indicate that GRP50 is markedly induced in MEFs as they exit G_0

and re-enter the cell cycle, and identify *GRP50* as a serum-response gene. The temporal pattern and level of induction of GRP50 mRNA in response to serum mirrors that of other known serum-response genes, including those involved in cytoskeletal reorganization, tissue remodeling, and wound healing (Iyer et al., 1999, and see Discussion, below).

The global expression pattern of GRP50 mRNA was investigated using northern blots containing RNA derived from human tissue. The tissue distribution of GRP50 mRNA was examined in adult (Figure 4A) and fetal (4B) tissues and in adult brain structures (4C). Although low levels of GRP50 mRNA were detected in all tissues examined, GRP50 is most strongly expressed in adult brain, heart, and skeletal muscle and is weakly expressed in lung, liver, and kidney (Figure 4A). A similar pattern of GRP50 expression was detected in northern blots of RNA isolated from mouse tissues (not shown). In contrast to the low expression levels in some adult tissues, GRP50 is expressed at high levels in all fetal tissues examined, including fetal lung and kidney (Figure 4B). Because GRP50 is expressed at high levels in adult brain, we wished to determine whether this expression is restricted to particular substructures. Figure 4C shows that GRP50 is expressed throughout the adult brain, and is only slightly more abundant in the occipital pole and in the frontal lobe than in other brain structures.

Structural features of the GRP50 protein

The sequence of the longest (3.5 kb) GRP50 cDNA is depicted in Figure 5. As mentioned above, none of the cDNA clones we isolated exhibited sequence differences in their presumptive protein coding regions. A search for putative open reading frames

(ORFs) along the cDNA sequence revealed several interesting features. First, the GRP50 transcript contains a single large ORF containing several in-frame methionine (AUG) codons and no in-frame upstream stop codons. However, this putative ORF is preceded by a short (39 nucleotide, underlined in 5-prime noncoding sequence in Figure 5) ORF, which lies in a different reading frame from that proposed for the GRP50 protein and which terminates prior to the proposed start for GRP50 translation. Second, although neither of the first two AUG codons conforms precisely to the consensus sequence for translation initiation, the second is flanked by the most conserved element of this consensus, a purine at position –3 (Kozak, 1987).

These observations are important because they describe the two most common exceptions to the general rule that translation of eukaryotic RNA is initiated at the first AUG codon: initiation at a downstream AUG can occur when the 5-prime proximal AUG is followed closely by a terminator codon (Kozak, 1999) or when it lacks the consensus sequence for translation initiation (Kozak, 1995). The facts that none of the cDNA clones obtained in our screen contain extended 5-prime sequences, and that the largest of these cDNAs is the same size as the larger GRP50 mRNA, suggest that our cDNA clones contain the intact 5-prime end of the GRP50 transcript. Taken together, these data imply that the second AUG codon may function as the translation initiator. This presumptive initiator defines an ORF of 1092 bp, conceptual translation of which yields a polypeptide of 363 amino acids with an estimated molecular mass of 41 kDa (Figure 5).

The structural features of this polypeptide include two ankyrin-like repeats (Figure 5, boxed) located near the NH_2-terminus, a proline-rich central domain (underlined), and a leucine-rich region at the COOH-terminus. The proline-rich central

domain (25% proline over a 100 amino acid stretch) is also rich in serine residues (15%), and includes several sites containing the SH3 domain binding consensus P–x–x–P (Mayer, 2001). Although the carboxyl-terminal region is leucine-rich, these residues do not exhibit the heptad periodicity characteristic of a classical leucine zipper (Landschulz et al., 1988). The Simple Modular Architecture Research Tool (SMART) (Schultz et al., 2000) identified the ankyrin repeats and the proline-rich low-complexity repeat as the only motifs corresponding to known protein domains. However, examination of the GRP50 protein sequence using the coiled coil prediction program COILS (Lupas, 1996; Lupas et al., 1991) identified sequences in the leucine-rich region that could mediate formation of a coiled coil structure (indicated with asterisks in Figure 5). Coiled coils are protein oligomerization domains in which two to five amphipathic α-helices associate to form a supercoil (Burkhard et al., 2001). Additional analyses of the GRP50 protein sequence using a suite of software tools (see Materials and Methods) did not reveal other conserved domains or motifs, apart from consensus sites for phosphorylation and N-linked glycosylation (not shown).

GRP50 is highly conserved among mammals

A search of the GenBank sequence databases has identified GRP50-homologous expressed sequence tags (ESTs) of human, rat, rabbit, and swine origins (Accession numbers AL543701, BG665112, C82830, and BI343865, respectively). Each of these ESTs shares significant similarity to murine GRP50 at both the nucleotide (~90%) and amino acid (~95%) levels, suggesting that they represent GRP50 orthologues in these species. Recently, a 2 kb human cDNA (GenBank Accession number BC012978) that is

homologous to the first 2 kb of murine GRP50 cDNA was identified. The human and mouse cDNAs share ~90% identity in the nucleotide sequences of their presumptive ORFs and ~80% identity in portions of their 5-prime and 3-prime UTR sequences. The translated amino acid sequences of mouse and human GRP50 share 92% overall identity (339/368 residues) and 95% overall similarity (350/368 identities plus conservative substitutions). Alignment of the mouse and human GRP50 protein sequences is shown in Figure 6A. It is interesting to note that the 5-prime UTR sequences of the mouse and human cDNAs exhibit an unusual degree of similarity, and that the human GRP50 cDNA contains a 39 nucleotide upstream ORF like that found in the mouse (Figure 6B). Conservation of 5-prime UTR sequences, including this short upstream ORF, suggests that these features may play a role in regulating translation of the GRP50 mRNA.

Protein sequence databases were scanned with the GRP50 protein sequence using the BLASTP algorithm (Altschul et al., 1997), and the results of these searches are summarized in Figure 7A. Although the ankyrin repeat region of GRP50 is homologous to ankyrin repeats in a variety of mammalian proteins (Figure 7A), no mammalian proteins with overall structural architecture similar to GRP50 were identified. A BLASTP search using the GRP50 sequence minus the ankyrin repeats did not return any homologous sequences (not shown). These results imply that GRP50 is a novel ankyrin repeat containing protein with a unique structural composition. Within the ankyrin repeat region, GRP50 is most similar to the ankyrin repeats of a predicted protein in *C. elegans*(45/97 positives), of mouse, human and rat I-κBα (49/93), and of *C. elegans* unc-44 (44/67). Of those proteins with homology to the ankyrin repeat region of GRP50, only the predicted protein from *C. elegans* (hypothetical protein R31.2, GenBank Accession

number T24241) exhibited significant similarity to GRP50 that extends beyond this domain (Figure 7B). Over the first ~160 amino acids of GRP50 and *C. elegans* R31.2, these proteins share 31% identity (51/164 residues) and 45% similarity (74/164, probability score 1e-12). The observations that the ankyrin repeats in GRP50 are more similar to those of *C. elegans* R31.2 than to those of other mammalian proteins, and that the similarity between these two proteins extends beyond the repeat region, suggest that R31.2 may be the orthologue of GRP50 in *C. elegans*. A search of the complete genomic sequences of *S. cerevisiae* and *D. melanogaster* failed to detect GRP50-homologous proteins in these organisms.

A BLAST search of the assembled human genome sequence with the GRP50 cDNA identified a locus (LOC91369) on the long arm of chromosome 17 that corresponds to the human *GRP50* gene. The *GRP50* gene lies within contig AC005921, which maps to chromosomal band 17q21.33 (Figure 8A, boxed). This region of chromosome 17, which is syntenic with mouse chromosome 11, is densely populated with genes (8B and C), and *GRP50* is separated from neighboring genes by less than 10 kb. A schematic representation of the *GRP50* locus is shown in Figure 8C; the gene contains 5 exons and spans approximately 15-20 kb of genomic sequence. As indicated in Figure 8C, transcription of *GRP50* proceeds in the opposite direction to that of the neighboring genes *ABCC3* (upstream) and *LUC7A* (downstream). *ABBC3* encodes the multidrug resistance protein-3 (MRP-3), a member of a family of ATP-binding cassette (ABC) proteins expressed in the liver which mediate transport of bile acids, phospholipids, and organic anions across the canalicular membrane (Borst et al., 2000; Uchiumi et al., 1998). *LUC7A* encodes a recently described protein called CROP, an

arginine- and serine-rich protein identified in a differential display-based screen for genes that are overexpressed in cisplatin-resistant cultured renal cancer cells (Nishii et al., 2000). The transcriptional orientations and close physical proximity of these three genes indicate that the 3-prime noncoding regions of *ABCC3* and *GRP50* are likely to overlap, while the 5-prime noncoding regions of *GRP50* and *LUC7A* may share common transcriptional regulatory elements.

DISCUSSION

This chapter describes the cloning and initial characterization of a novel gene, *GRP50*, which encodes a highly conserved mammalian protein with a unique structural architecture. Sequence analysis indicates that *GRP50* encodes a medium-sized protein with multiple modular domains, including an ankyrin repeat region, a polyproline-rich region, and a leucine-rich region that may form α-helical coiled coils. The GRP50 protein is highly conserved among mammals, but homology with non-mammalian proteins is in most cases limited to the ankyrin repeat region. However, one hypothetical protein (R31.2) in the nematode *C. elegans* contains sequences that are homologous to GRP50 outside of the ankyrin repeats themselves. The *C. elegans* R31.2 protein also displays an overall architecture that is partly similar to GRP50, in that it is composed of two NH_2-terminal ankyrin repeats, followed by low complexity repeat regions (some of which are proline rich).

Because of the structural features of ankyrin repeats discussed above, orthologous ankyrin repeat proteins from divergent species tend to exhibit sequence conservation outside of the repeat consensus (Bork, 1993; Sedgwick and Smerdon, 1999), while in unrelated proteins the homology is limited to the consensus sequence itself. In addition, binding specificity is conferred by the number of repeats that associate to form a tertiary structure as well as by the sequences in and around the repeats. Thus the observation that GRP50 and the R31.2 protein contain the same number of ankyrin repeats, and exhibit sequence similarity outside of the repeats, might be taken as a more convincing indication of functional homology than that implied by the high degree of similarity between

GRP50 and I-κBα, which is limited to the ankyrin repeats proper. Although the absence of a known biological function for either GRP50 or the R31.2 protein precludes a convincing demonstration of orthology, the structural similarities between these two proteins parallel those of other distantly related orthologues (Zapata et al., 2000; Zhang and Emmons, 2000).

Although we have not detected putative GRP50 orthologues in yeast or in *Drosophila*, GRP50 is highly conserved among mammals. The predicted amino acid sequences of ESTs from human, rat, rabbit, and swine each exhibit greater than 95% similarity to murine GRP50. More striking, however, is the high degree of nucleotide sequence similarity observed in the untranslated regions of the human and mouse GRP50 mRNAs. Like the mouse *GRP50*, the human gene appears to produce two alternatively spliced transcripts (although the smaller transcript is less abundant in human tissues than in mouse), and no differences in the presumptive open reading frame have been detected in human ESTs. In addition, analysis of the human *GRP50* genomic region using programs designed to detect protein-coding sequences indicates that a single polypeptide is produced from this locus. These observations suggest that alternative splicing of the human *GRP50* gene yields mRNA species that differ only in the lengths of their 3-prime UTRs, as is apparently the case with murine *GRP50*.

A comparison of the 5-prime and 3-prime UTRs of murine and human GRP50 reveals that these sequences share greater than 80% similarity, and indicates that they may play a role in the post-transcriptional regulation of GRP50 expression by influencing the stability or translational accessibility of the GRP50 mRNA. In particular, the presence of a small open reading frame upstream of the putative translation initiator codons in both

the mouse and the human GRP50 transcripts suggests an evolutionary conservation of intricate mechanisms for modulating GRP50 expression levels. Conservation of regulatory sequences implies a cellular requirement that GRP50 protein levels be tightly controlled.

The profile of GRP50 mRNA expression supports the presumption that GRP50 protein levels are tightly regulated. The tissue distribution of GRP50 mRNA indicates that, while this gene is ubiquitously expressed in adult tissues, expression levels appear to be regulated in a tissue- or cell-type specific manner: GRP50 is more highly expressed in brain, heart, and skeletal muscle than in liver or lung. In addition, GRP50 is expressed at high levels in embryonic tissues, including those that show low expression levels in adults. Although our data do not permit a direct comparison between GRP50 mRNA levels in fetal and adult tissues, *in silico* analysis of GRP50 representation in the EST databases suggests that this gene is more highly expressed in embryonic tissues and in tumors than in normal adult tissues. Thus GRP50 may be a developmentally regulated gene that exhibits distinct patterns of expression in fetal and adult tissues.

Although originally of interest because of its apparent enrichment in cells undergoing apoptosis, GRP50 expression is not induced specifically in response to an apoptotic stimulus or more generally in response to genotoxic stress. Multiple analyses of GRP50 mRNA levels in γ-irradiated MEFs and E1A/ras MEFs did not reveal a consistent pattern of GRP50 induction under these conditions. Further, the enrichment of GRP50 mRNA in γ-irradiated, E1A/ras MEFs observed in our initial experiments was not attributable to the expression of transforming oncogenes in these cells, since untransformed cells exhibit similar levels of the transcripts.

In contrast, GRP50 expression is markedly induced in the response of MEFs to serum stimulation. The ability of primary fibroblasts to proliferate in culture is dependent upon exogenously added serum or purified growth factors; when deprived of serum, these cells enter a quiescent, or nonproliferative, state termed G_0, which is characterized by low metabolic activity (Winkles, 1998). Quiescent MEFs are released from this state upon re-exposure to serum and subsequently re-enter the cell cycle as a synchronous population. The response of primary fibroblasts to serum starvation and release has therefore provided a model system for the study of genes that mediate growth and cell cycle progression (Iyer et al., 1999; Winkles, 1998).

Our results indicate that GRP50 mRNA is induced several fold in response to serum stimulation; the increase in mRNA levels is evident by as few as 8 hours following exposure to serum and remains constant over the next 16 hours. Because the induction of GRP50 mRNA was observed in the earliest time points included in these experiments, these results do not indicate the precise timing of GRP50 induction in the course of the serum response. In addition, because samples from unsynchronized cells were not included in these experiments, we cannot determine whether the differences observed result primarily from a marked repression of GRP50 expression in quiescent cells. Nevertheless, these results clearly demonstrate that GRP50 is induced in cells as they exit from G_0, and permit some insight into the temporal pattern of GRP50 expression in the response of fibroblasts to serum stimulation.

Although serum-stimulation experiments are frequently employed in the study of the cellular response to mitogenic signals, the biology of the serum response is not limited to cell cycle progression. Recent work by Iyer and colleagues (Iyer et al., 1999)

describes the analysis of the transcriptome in serum-stimulated fibroblasts using a microarray-based approach. In this study, the expression profiles of close to 10,000 genes were examined in serum-starved and serum-stimulated human primary fibroblasts. In addition to cell cycle regulators and immediate early transcription factors, genes involved in coagulation, inflammation, angiogenesis, tissue remodeling, cytoskeletal reorganization, and wound healing were among those that exhibited significant induction in response to serum.

We searched the data set from these experiments (genome-www.stanford.edu/serum) for genes whose expression profiles resembled that of GRP50 (i.e., genes that were induced greater than two-fold at each time point from 8 to 24 hours). This search revealed that fewer than fifty genes in the 10K dataset exhibited a serum induction profile like that of GRP50 (GRP50 was not included in the dataset). Of those genes that did show a similar profile of RNA induction in response to serum, there were several that are implicated in tissue remodeling and wound healing (Iyer et al., 1999). Iyer and colleagues have suggested that the diverse set of serum-inducible genes represents the broad transcriptional response to a specific physiological stimulus, serum, that is normally encountered in the context of a wound, and that the view of the serum stimulus as a mitogenic signal is an oversimplification of a broader biological process.

The experiments described in this chapter detail the cloning of GRP50 and its identification as a novel member of the ankyrin repeat family of proteins. The predicted structure of the GRP50 protein indicates the presence of multiple protein binding domains, and implies that the function of GRP50 is mediated through intermolecular interactions with another protein or class of proteins. In the absence of a structural

orthologue of known function, however, it is not possible to predict what types of proteins might interact with GRP50, and these analyses await further investigation.

MATERIALS AND METHODS

Cell culture

E1A/ras MEFs were generated by Scott Lowe, and have been described previously (Lowe et al., 1993). MEFs and E1A/ras MEFs were cultured in Dulbecco's modified Eagle's medium (DMEM) supplemented with 10% heat-inactivated fetal calf serum (FCS), 50 U/ml penicillin, 50 μg/ml streptomycin, and 2mM L-glutamine and incubated at 37°C in a humidified chamber containing 5% CO_2. For irradiation experiments, cells were plated in 15 cm dishes and exposed to 7Gy of γ-radiation in a GAMMACELL 40 with a ^{137}Cs source. For irradiation time-course experiments, cells were harvested at the indicated time points following γ-irradiation, and pellets were flash-frozen and stored at –80°C prior to RNA isolation.

Library construction and screening

The E1A/ras cDNA library was constructed using the SuperScript™ Choice System for cDNA Synthesis (Gibco BRL) according to the manufacturer's instructions. RNA from E1A/*ras* MEFs, harvested one hour following γ-irradiation, was subjected to two rounds of poly(A) selection and used as the template for cDNA synthesis. First strand synthesis was primed with a NotI-oligo(dT) primer/adapter, maximizing the likelihood that it would initiate from the poly(A) tail of messenger RNAs. Double stranded cDNA was size-selected and directionally cloned into the bacteriophage vector λgt22a. The library was packaged using Gigapack™ Gold packaging extract (Stratagene) according to the manufacturer's instructions and propagated in *E. coli* strain Y1090

according to standard methods (Sambrook et al., 1989). The total size of the unamplified cDNA library was ~7 x 10^6 clones, of which10^6 clones were plated and screened.

For low-stringency screening to identify *reaper*-homologous cDNAs, a PCR product containing the 200 bp reaper open reading frame was labeled with α^{32}P-dCTP by de novo synthesis using a 3-prime specific primer. Library filters were hybridized to the resulting probe at > 2 x 10^6 cpm/ml in Church's buffer (0.5 M sodium phosphate, pH 7.0, 7% SDS, 1% bovine serum albumin, 1mM EDTA, and 100 μg/ml salmon sperm DNA) at 55°. Filters were washed 4 x 20 minutes in 2 x SSC, 0.1% SDS at 50°C. Between probings, library filters were stripped in 0.4M NaOH for 30 minutes at 45°C, followed by 0.1 x SSC, 0.1% SDS, 0.2M Tris, pH 7.5 for 15 minutes at 45°C. For high-stringency screening to identify additional GRP50 clones, the original 2.2 kb ER18 cDNA was labeled with α^{32}P-dCTP by random priming and hybridized to the stripped library filters in Church's buffer at 65°C. The filters were washed 2 x 20 minutes in 2 x SSC, 0.1% SDS, followed by 2 x 20 minutes in 0.1 x SSC, 0.1% SDS at 55°C. DNA was prepared from plaque-purified positive phage clones using standard methods (Sambrook et al., 1989), and cDNAs were excised from phage vectors and subcloned into pBluescript II (Stratagene) for sequencing and subsequent analysis.

RNA isolation and northern analysis

Human Multiple Tissue Northerns, containing 2μg of poly(A)+ RNA per lane, were purchased from Clontech and were hybridized and washed according to the manufacturer's instructions. Serum starvation and release northerns were a kind gift of Raluca Verona, and are described elsewhere (Humbert et al., 2000). RNA from cultured

MEFs and E1A/ras MEFs was prepared as follows: total cellular RNA was isolated using either the Ultraspec™ RNA isolation system (Biotecx, Houston, TX.) or RNeasy Midi Kit (Qiagen) according to the manufacturers' instructions. For poly(A) selection, a minimum of 500 μg total RNA was bound in batch to oligo(dT)-cellulose (New England Biolabs) in 0.4 M NaCl, 10 mM Tris, pH 7.5, 1 mM EDTA. The slurry was loaded onto a disposable column (BioRad), washed extensively with 0.2 M NaCl, 10 mM Tris, pH 7.5, 1 mM EDTA, and poly(A)+ RNA was eluted in DEPC-treated H_2O. For northern blotting, 2 to 4 μg of poly(A)+ RNA was electrophoresed in MOPS buffer (20 mM Na-MOPS, 5 mM Na-acetate, 1 mM EDTA, pH 7.0) through 2.2 M formaldehyde gels. Gel-fractionated RNA was then transferred to nylon membranes (Hybond N) and hybridized with the indicated probes in Church's buffer (0.5M sodium phosphate, pH 7.0, 7% SDS, 1% bovine serum albumin, 1 mM EDTA, and 100 μg/ml salmon sperm DNA) at 65°C. Northern blots were washed 2 x 20 minutes in 2 x SSC, 0.1% SDS, followed by 2 x 20 minutes in 0.1 x SSC, 0.1% SDS at 55°C.

Sequencing and sequence analysis

Each of the GRP50 cDNAs depicted in Figure 1C was sequenced in its entirety by generating a series of overlapping subclones. Sequencing of both strands was performed on an automated ABI sequencer using vector-derived primers or, occasionally, GRP50-specific primers. DNA and protein similarity searching was performed using the BLAST server at the NCBI (ncbi.nlm.nih.gov/blast). The BLAST search of the assembled human genome sequence was carried out using the Ensembl Genome Server at <ensembl.org>. Protein sequences were analyzed using the software DNASTAR and the suites of tools

available at the following web servers: Baylor College of Medicine (searchlauncher.bcm.tmc.edu), ExPASy (expasy.ch/tools/), and EMBL (embl–heidelberg.de/Services/index.html). Multiple sequence alignments (DNA and protein) were generated using the web-based ClustalW 1.8 program, located at searchlauncher.bcm.tmc.edu/multi-align/multi-align.html. Aligned sequences were shaded using the BOXSHADE server at ch.embnet.org/software/BOX form.html.

ACKNOWLEDGEMENTS

I thank Scott Lowe for the gift of the E1A/Ras cell line. I am grateful to Raluca Verona for the use of her serum starvation and release northerns and for the AR PPO control probe.

Figure 1

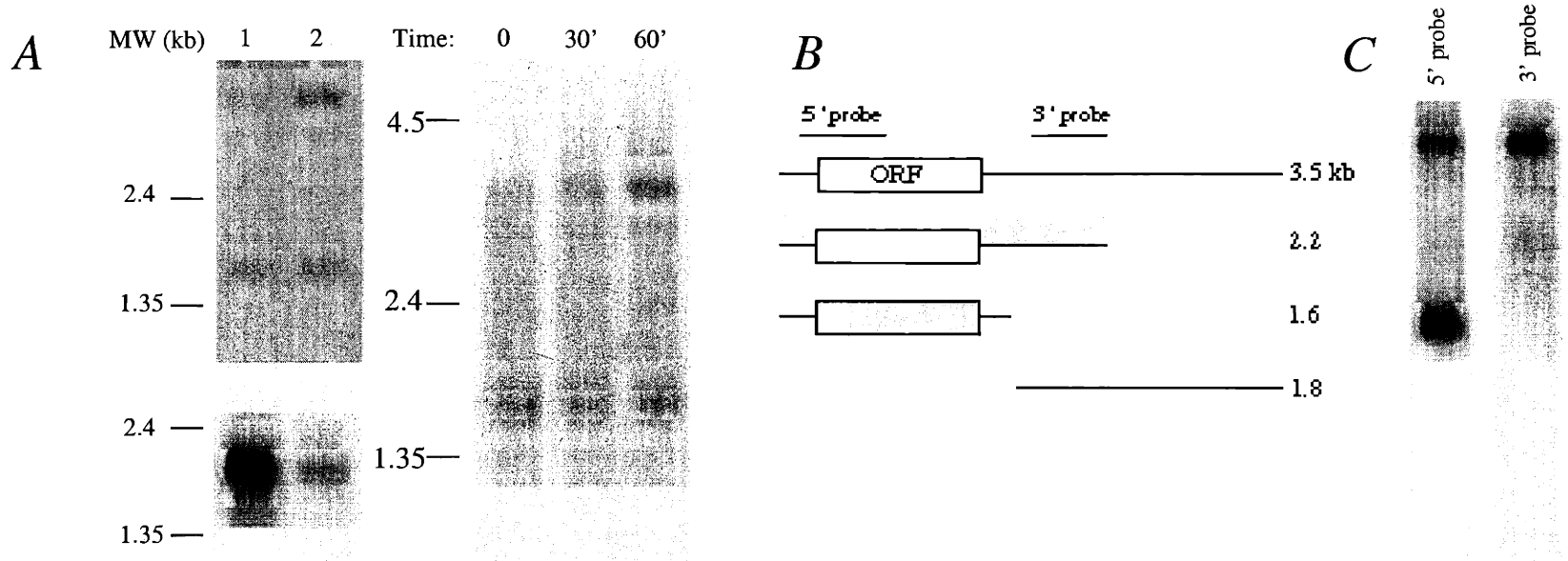


Figure 1. ***(A)*** Left, upper panel: expression of GRP50 in MEFs (lane 1) and in γ-irradiated, E1A/ras MEFs (lane 2). 4 μg of polyA-selected RNA was loaded per lane. RNA in lane 2 was isolated from E1A/ras MEFs one hour following exposure to 7 Gy of γ-radiation. Right: expression of GRP50 in E1A/ras MEFs at 0, 30, and 60 minutes following γ-radiation. The blots were probed with the 2.2 kb cDNA depicted in (B). The lower panel shows the above blot reprobed with ELP-1 (Erd2-like protein, an ER to Golgi transporter) cDNA as a loading control. ***(B)*** GRP50 cDNA clones isolated from E1a/ras library. The open reading frame is marked with a box, and the probes used in (C) are indicated. Aligned portions of cDNA clones share identical sequence. ***(C)*** Detection of GRP50 transcripts in polyA+ RNA from E1A/ras MEFs,using 5' and 3' probes indicated in (B). The panels show successive probings of the same blot.

Figure 2

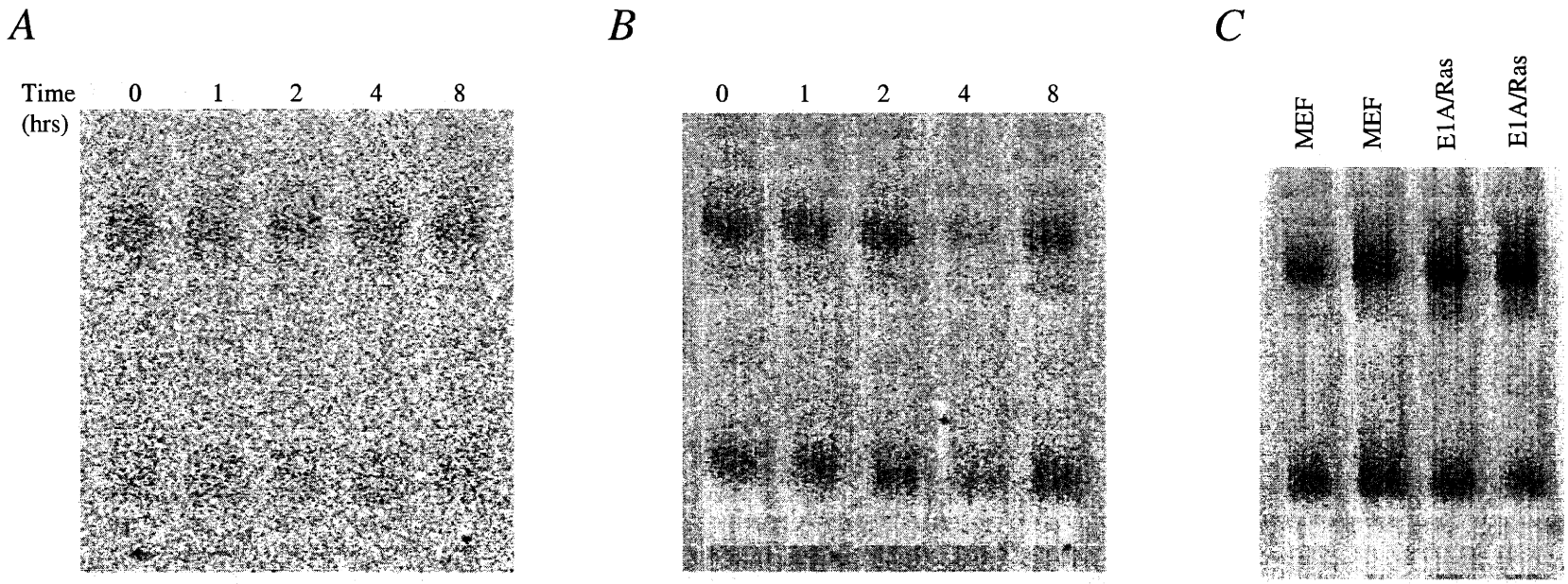


Figure 2. ***(A), (B)*** Expression of GRP50 mRNA is not induced in response to γ-irradiation in MEFs (A) or in E1A/Ras MEFs (B). Cells were harvested at the indicated time points (hours) following exposure to 7Gy of γ-radiation. ***(C)*** Expression of GRP50 mRNA in unirradiated MEFs and E1A/Ras Mefs. All blots were hybridized with *AR PPO* as a loading control (not shown).

Figure 3

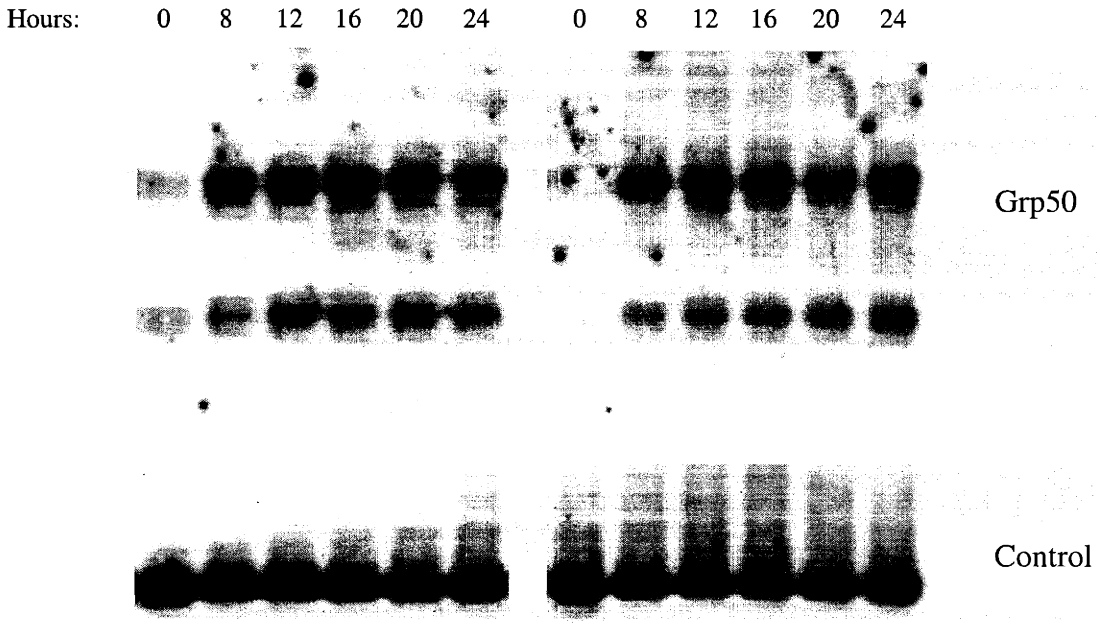


Figure 3. Expression of GRP50 mRNA is serum-induced. MEFs were arrested in G_0 via serum deprivation and harvested at indicated time points (hours) after serum addition. Upper panel: GRP50 expression; lower panel; ARPP PO loading control (see text). Each blot contains RNA isolated from independently derived MEFs.

Figure 4

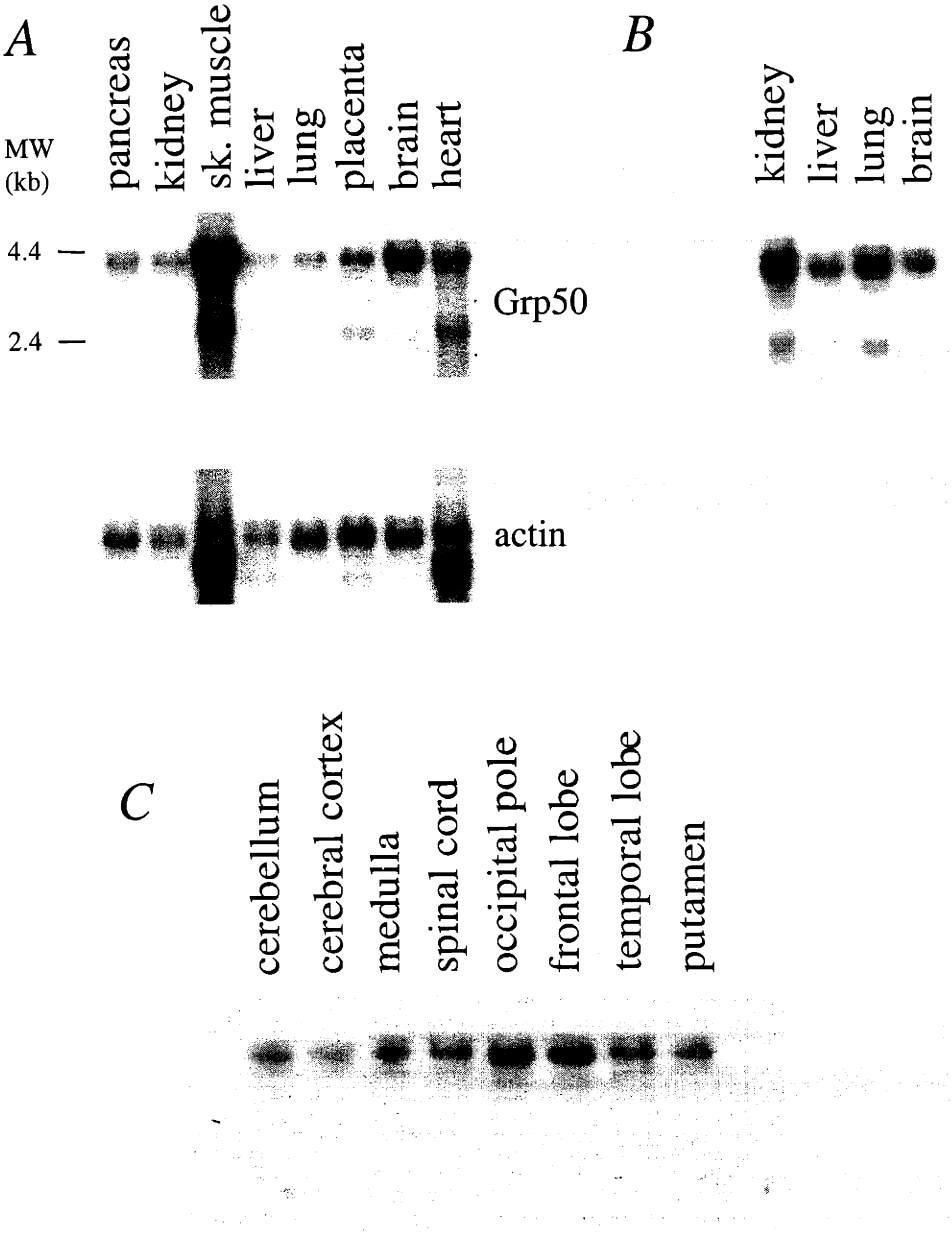


Figure 4. Expression of human GRP50 in adult ***(A)*** and ***(B)*** fetal tissues, and **(*C*)** in adult human brain structures. Northern Blots were purchased from Clontech. In the lower panel in (A), the blot was reprobed with actin to control for loading. The GRP50 probe used is an 800 bp PvuII-ScaI fragment of the open reading frame.

Figure 5

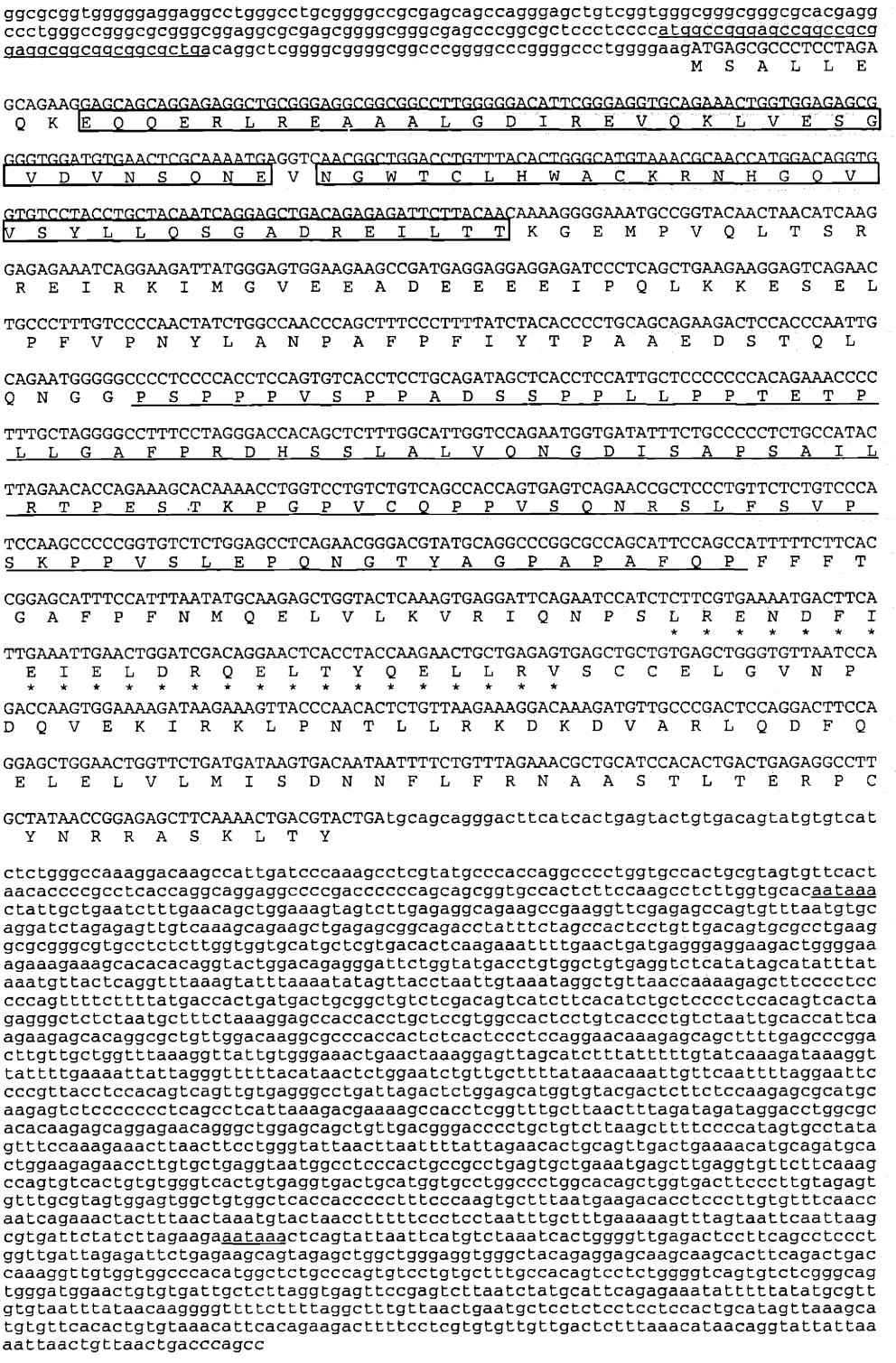
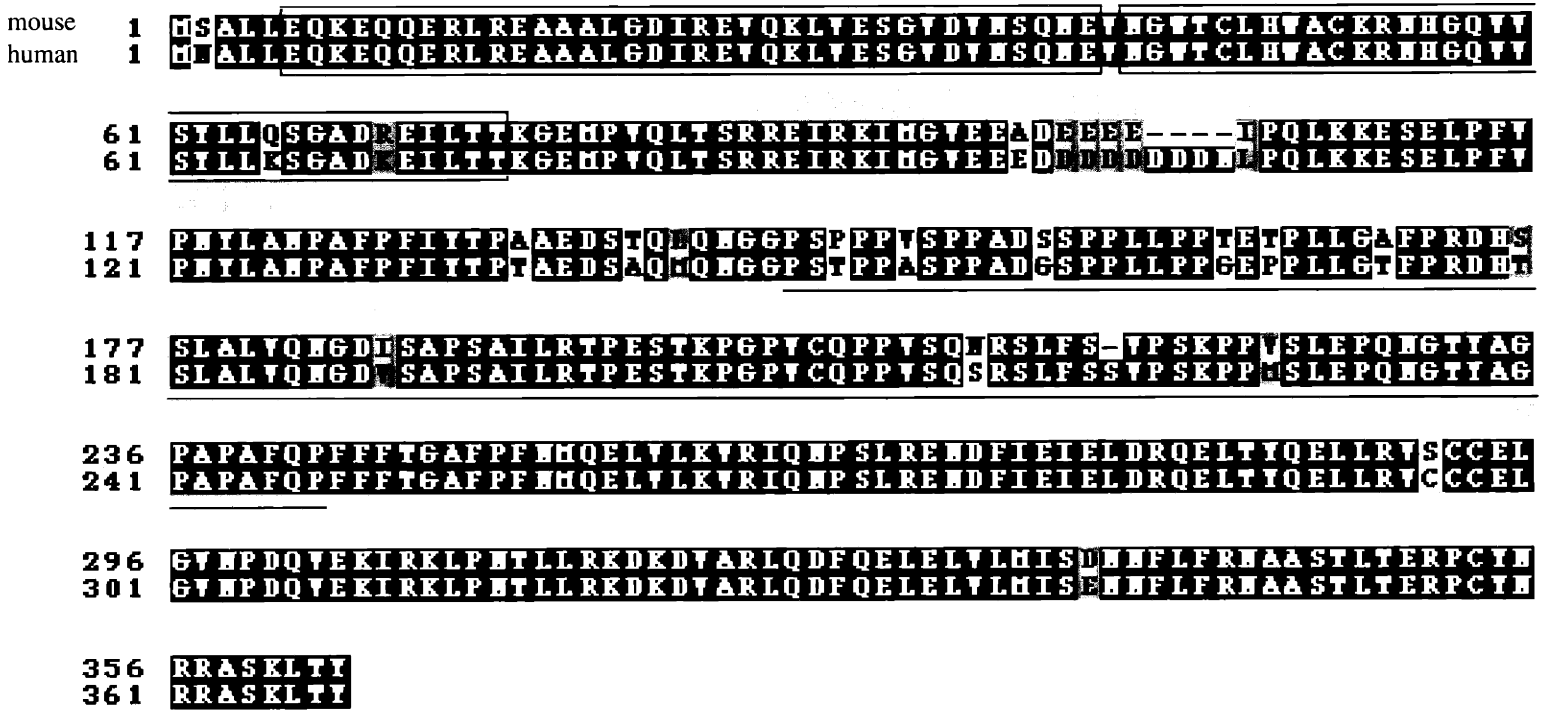


Figure 5. Sequence of the 3.5 kb murine GRP50 cDNA, and predicted translation. In the untranslated regions, the short upstream ORF (5') and consensus polyadenylation signals (3') are underlined. In the putative protein sequence, the two ankyrin repeats are boxed, the proline-rich domain is underlined, and the putative coiled coil forming region is indicated with asterisks.

Figure 6

A



B

Figure 6. *(A)* Alignment of the predicted amino acid sequences of mouse (top) and human (bottom) GRP50 proteins. Identical residues (339/368, 92%) are shaded black; conservative substitutions are shaded blue (conservative substitutions plus identities: 350/368, 95%). The two N-terminal ankyrin repeats are boxed, and the proline rich domain is underlined. ***(B)*** Alignment of a short (~100 bp) region of the mouse (top) and human (botttom) 5' UTRs (lower case), including the first 27 bp of the presumptive ORFs (upper case). The 39 bp upstream ORFs present in both sequences are boxed.

Figure 7

A

BLASTP results

Sequences producing significant alignments:		Score (bits)	E Value
C. elegans R31.2	gi|7506777|pir||T24241 hypothetical protein R31.2 - Caenorh...	56	4e-07
Human IκBα	gi|6688192|emb|CAB65113.1| (AJ249290) IkappaBalpha [Homo sa...	50	3e-05
Rat IκBα	gi|14548069|sp|Q63746|IKBA_RAT NF-KAPPAB INHIBITOR ALPHA (I...	49	9e-05
Mouse IκBα	gi|14548087|sp|Q9Z1E3|IKBA_MOUSE NF-KAPPAB INHIBITOR ALPHA ...	49	1e-04
Ankyrin	gi|1841966|gb|AAB47551.1| (U65916) ankyrin [Rattus norvegicus]	48	1e-04
RNAse L	gi|10863929|ref|NP_066956.1| ribonuclease L (2',5'-oligoiso...	48	2e-04
unc-44	gi|7494531|pir||T15347 ankyrin-related unc-44 - Caenorhabdi...	47	2e-04

B

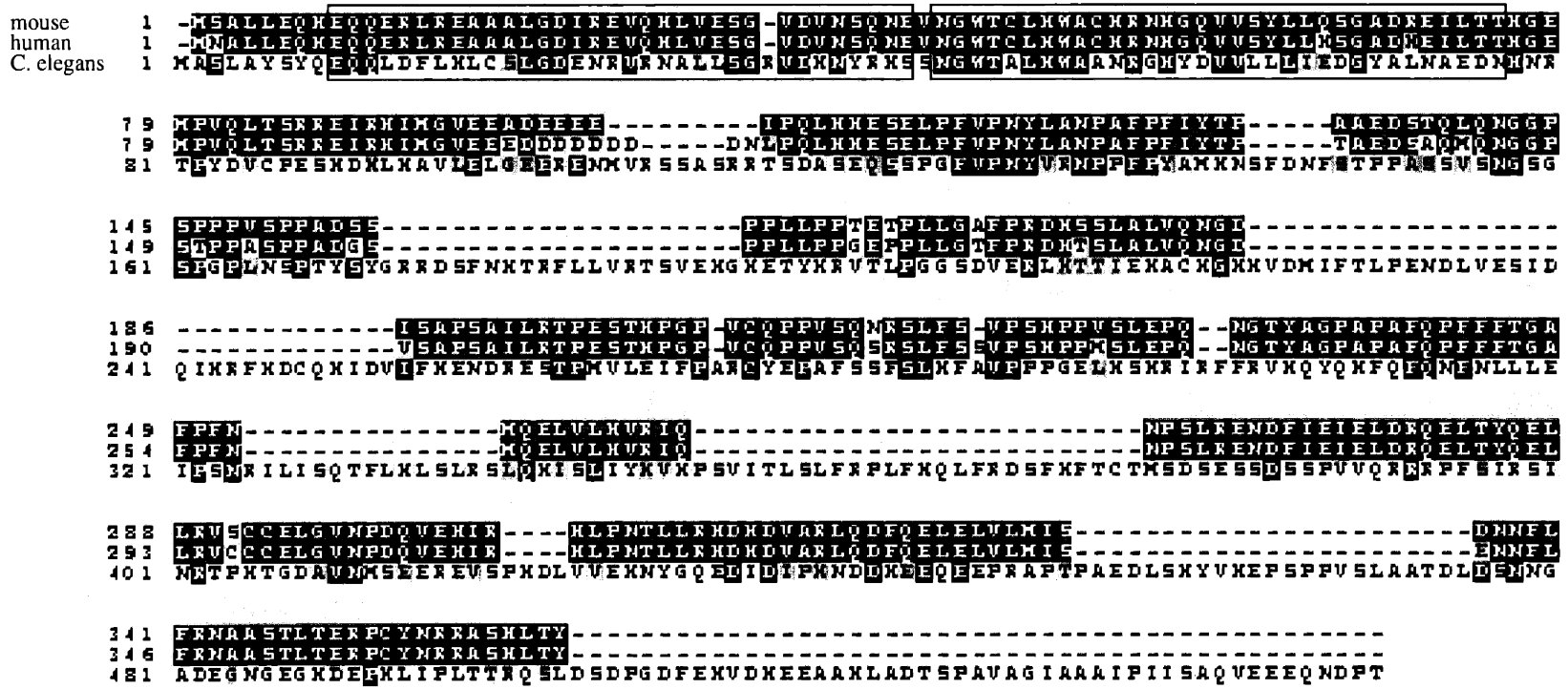


Figure 7. *(A)*. Results (edited for clarity) of BLASTP search of known proteins for similarity to GRP50, showing the highest scoring sequences. *(B)*. Alignment of the predicted amino acid sequences of murine (top) and human (middle) GRP50 with putative C. elegans protein R31.2 (bottom). The C-terminal domain of R31.2 protein is omitted. Identical residues are shaded black, conservative substitutions are shaded blue, and the two amino-terminal ankyrin repeats are boxed.

Figure 8

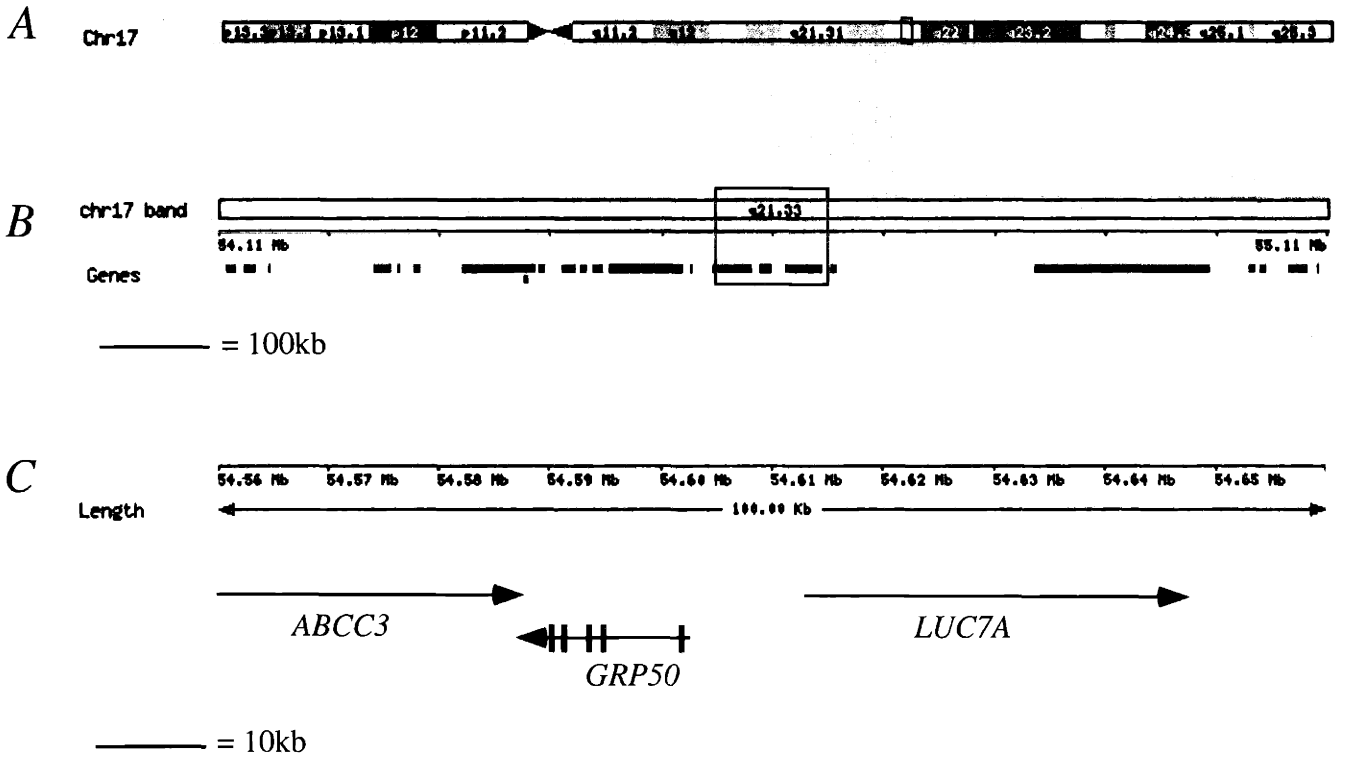


Figure 8. Chromosomal location and predicted structure of the human *GRP50* locus, adapted from the output file at www.ensembl.org/Homo_sapiens/contigview?chr=chr17 ***(A)*** The human *GRP50* locus lies at 17q21.33. ***(B)*** Expanded view of the 1 Mb boxed region in (A). The relative positions of genes are indicated by black lines. ***(C)*** Expanded view of the 100 kb boxed region in (B), showing transcriptional orientation of *GRP50* and neighboring genes. The relative positions of the *GRP50* exons are indicated by narrow rectangles.

REFERENCES

Altschul, S.F., T.L. Madden, A.A. Schaffer, J. Zhang, Z. Zhang, W. Miller, and D.J. Lipman. 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* 25:3389-3402.

Bar-Sagi, D., D. Rotin, A. Batzer, V. Mandiyan, and J. Schlessinger. 1993. SH3 domains direct cellular localization of signaling molecules. *Cell.* 74:83-91.

Batchelor, A.H., D.E. Piper, F.C. de la Brousse, S.L. McKnight, and C. Wolberger. 1998. The structure of GABPalpha/beta: an ETS domain- ankyrin repeat heterodimer bound to DNA. *Science.* 279:1037-1041.

Baumgartner, R., C. Fernandez-Catalan, A. Winoto, R. Huber, R.A. Engh, and T.A. Holak. 1998. Structure of human cyclin-dependent kinase inhibitor p19INK4d: comparison to known ankyrin-repeat-containing structures and implications for the dysfunction of tumor suppressor p16INK4a. *Structure.* 6:1279-1290.

Bear, J.E., M. Krause, and F.B. Gertler. 2001. Regulating cellular actin assembly. *Curr Opin Cell Biol.* 13:158-166.

Bork, P. 1993. Hundreds of ankyrin-like repeats in functionally diverse proteins: mobile modules that cross phyla horizontally? *Proteins.* 17:363-374.

Bork, P., and M. Sudol. 1994. The WW domain: a signalling site in dystrophin? *Trends Biochem Sci.* 19:531-533.

Borst, P., R. Evers, M. Kool, and J. Wijnholds. 2000. A family of drug transporters: the multidrug resistance-associated proteins. *J Natl Cancer Inst.* 92:1295-1302.

Breeden, L., and K. Nasmyth. 1987. Similarity between cell-cycle genes of budding yeast and fission yeast and the Notch gene of Drosophila. *Nature.* 329:651-654.

Burkhard, P., J. Stetefeld, and S.V. Strelkov. 2001. Coiled coils: a highly versatile protein folding motif. *Trends Cell Biol.* 11:82-88.

Callebaut, I., P. Cossart, and P. Dehoux. 1998. EVH1/WH1 domains of VASP and WASP proteins belong to a large family including Ran-binding domains of the RanBP1 family. *FEBS Lett.* 441:181-185.

Foord, R., I.A. Taylor, S.G. Sedgwick, and S.J. Smerdon. 1999. X-ray structural analysis of the yeast cell cycle regulator Swi6 reveals variations of the ankyrin fold and has implications for Swi6 function. *Nat Struct Biol.* 6:157-165.

Franz, W.M., P. Berger, and J.Y. Wang. 1989. Deletion of an N-terminal regulatory domain of the c-abl tyrosine kinase activates its oncogenic potential. *Embo J.* 8:137-147.

Goldberg, D.J., M.S. Foley, D. Tang, and P.W. Grabham. 2000. Recruitment of the Arp2/3 complex and mena for the stimulation of actin polymerization in growth cones by nerve growth factor. *J Neurosci Res.* 60:458-467.

Hengartner, M.O., and H.R. Horvitz. 1994. C. elegans cell survival gene ced-9 encodes a functional homolog of the mammalian proto-oncogene bcl-2. *Cell.* 76:665-676.

Holt, M.R., and A. Koffer. 2001. Cell motility: proline-rich proteins promote protrusions. *Trends Cell Biol.* 11:38-46.

Humbert, P.O., R. Verona, J.M. Trimarchi, C. Rogers, S. Dandapani, and J.A. Lees. 2000. E2f3 is critical for normal cellular proliferation. *Genes Dev.* 14:690-703.

Hurford, R.K., Jr., D. Cobrinik, M.H. Lee, and N. Dyson. 1997. pRB and p107/p130 are required for the regulated expression of different sets of E2F responsive genes. *Genes Dev.* 11:1447-1463.

Huxford, T., D.B. Huang, S. Malek, and G. Ghosh. 1998. The crystal structure of the IkappaBalpha/NF-kappaB complex reveals mechanisms of NF-kappaB inactivation. *Cell.* 95:759-770.

Iyer, V.R., M.B. Eisen, D.T. Ross, G. Schuler, T. Moore, J.C. Lee, J.M. Trent, L.M. Staudt, J. Hudson, Jr., M.S. Boguski, D. Lashkari, D. Shalon, D. Botstein, and P.O. Brown. 1999. The transcriptional program in the response of human fibroblasts to serum. *Science.* 283:83-87.

Jacobs, M.D., and S.C. Harrison. 1998. Structure of an IkappaBalpha/NF-kappaB complex. *Cell.* 95:749-758.

Kastan, M.B., Q. Zhan, W.S. el-Deiry, F. Carrier, T. Jacks, W.V. Walsh, B.S. Plunkett, B. Vogelstein, and A.J. Fornace, Jr. 1992. A mammalian cell cycle checkpoint pathway utilizing p53 and GADD45 is defective in ataxia-telangiectasia. *Cell.* 71:587-597.

Kay, B.K., M.P. Williamson, and M. Sudol. 2000. The importance of being proline: the interaction of proline-rich motifs in signaling proteins with their cognate domains. *Faseb J.* 14:231-241.

Korey, C.A., and D. Van Vactor. 2000. From the growth cone surface to the cytoskeleton: one journey, many paths. *J Neurobiol.* 44:184-193.

Kozak, M. 1987. An analysis of 5'-noncoding sequences from 699 vertebrate messenger RNAs. *Nucleic Acids Res.* 15:8125-8148.

Kozak, M. 1995. Adherence to the first-AUG rule when a second AUG codon follows closely upon the first. *Proc Natl Acad Sci U S A.* 92:2662-2666.

Kozak, M. 1999. Initiation of translation in prokaryotes and eukaryotes. *Gene.* 234:187-208.

Landschulz, W.H., P.F. Johnson, and S.L. McKnight. 1988. The leucine zipper: a hypothetical structure common to a new class of DNA binding proteins. *Science.* 240:1759-1764.

Li, J., I.J. Byeon, K. Ericson, M.J. Poi, P. O'Maille, T. Selby, and M.D. Tsai. 1999. Tumor suppressor INK4: determination of the solution structure of p18INK4C and demonstration of the functional significance of loops in p18INK4C and p16INK4A. *Biochemistry.* 38:2930-2940.

Lowe, S.W., H.E. Ruley, T. Jacks, and D.E. Housman. 1993. p53-dependent apoptosis modulates the cytotoxicity of anticancer agents. *Cell.* 74:957-967.

Luh, F.Y., S.J. Archer, P.J. Domaille, B.O. Smith, D. Owen, D.H. Brotherton, A.R. Raine, X. Xu, L. Brizuela, S.L. Brenner, and E.D. Laue. 1997. Structure of the cyclin-dependent kinase inhibitor p19Ink4d. *Nature.* 389:999-1003.

Lupas, A. 1996. Prediction and analysis of coiled-coil structures. *Methods Enzymol.* 266:513-525.

Lupas, A., M. Van Dyke, and J. Stock. 1991. Predicting coiled coils from protein sequences. *Science.* 252:1162-1164.

Lux, S.E., K.M. John, and V. Bennett. 1990. Analysis of cDNA for human erythrocyte ankyrin indicates a repeated structure with homology to tissue-differentiation and cell-cycle control proteins. *Nature.* 344:36-42.

Machesky, L.M. 1997. Cell motility: complex dynamics at the leading edge. *Curr Biol.* 7:R164-167.

Mayer, B.J. 2001. SH3 domains: complexity in moderation. *J Cell Sci.* 114:1253-1263.
McCormick, F. 1993. Signal transduction. How receptors turn Ras on. *Nature.* 363:15-16.
Michaely, P., and V. Bennett. 1993. The membrane-binding domain of ankyrin contains four independently folded subdomains, each comprised of six ankyrin repeats. *J Biol Chem.* 268:22703-22709.

Miura, M., H. Zhu, R. Rotello, E.A. Hartwieg, and J. Yuan. 1993. Induction of apoptosis in fibroblasts by IL-1 beta-converting enzyme, a mammalian homolog of the C. elegans cell death gene ced-3. *Cell.* 75:653-660.

Nishii, Y., M. Morishima, Y. Kakehi, K. Umehara, N. Kioka, Y. Terano, T. Amachi, and K. Ueda. 2000. CROP/Luc7A, a novel serine/arginine-rich nuclear protein, isolated from cisplatin-resistant cell line. *FEBS Lett.* 465:153-156.

Raff, M.C. 1992. Social controls on cell survival and cell death. *Nature.* 356:397-400.

Reiersen, H., and A.R. Rees. 2001. The hunchback and its neighbours: proline as an environmental modulator. *Trends Biochem Sci.* 26:679-684.

Russo, A.A., L. Tong, J.O. Lee, P.D. Jeffrey, and N.P. Pavletich. 1998. Structural basis for inhibition of the cyclin-dependent kinase Cdk6 by the tumour suppressor p16INK4a. *Nature.* 395:237-243.

Sambrook, J., E.F. Fritsch, and T. Maniatis. 1989. Molecular cloning : a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.

Schultz, J., R.R. Copley, T. Doerks, C.P. Ponting, and P. Bork. 2000. SMART: a web-based tool for the study of genetically mobile domains. *Nucleic Acids Res.* 28:231-234.

Sedgwick, S.G., and S.J. Smerdon. 1999. The ankyrin repeat: a diversity of interactions on a common structural framework. *Trends Biochem Sci.* 24:311-316.

Seidel-Dugan, C., B.E. Meyer, S.M. Thomas, and J.S. Brugge. 1992. Effects of SH2 and SH3 deletions on the functional activities of wild- type and transforming variants of c-Src. *Mol Cell Biol.* 12:1835-1845.

Staub, O., S. Dho, P. Henry, J. Correa, T. Ishikawa, J. McGlade, and D. Rotin. 1996. WW domains of Nedd4 bind to the proline-rich PY motifs in the epithelial Na+ channel deleted in Liddle's syndrome. *Embo J.* 15:2371-2380.

Staub, O., H. Yeger, P.J. Plant, H. Kim, S.A. Ernst, and D. Rotin. 1997. Immunolocalization of the ubiquitin-protein ligase Nedd4 in tissues expressing the epithelial Na+ channel (ENaC). *Am J Physiol.* 272:C1871-1880.

Sudol, M. 1996. Structure and function of the WW domain. *Prog Biophys Mol Biol.* 65:113-132.

Sudol, M., K. Sliwa, and T. Russo. 2001. Functions of WW domains in the nucleus. *FEBS Lett.* 490:190-195.

Thompson, C.C., T.A. Brown, and S.L. McKnight. 1991. Convergence of Ets- and notch-related structural motifs in a heteromeric DNA binding complex. *Science.* 253:762-768.

Uchiumi, T., E. Hinoshita, S. Haga, T. Nakamura, T. Tanaka, S. Toh, M. Furukawa, T. Kawabe, M. Wada, K. Kagotani, K. Okumura, K. Kohno, S. Akiyama, and M. Kuwano. 1998. Isolation of a novel human canalicular multispecific organic anion transporter,

cMOAT2/MRP3, and its expression in cisplatin-resistant cancer cells with decreased ATP-dependent drug transport. *Biochem Biophys Res Commun.* 252:103-110.

Vaux, D.L., and I.L. Weissman. 1993. Neither macromolecular synthesis nor myc is required for cell death via the mechanism that can be controlled by Bcl-2. *Mol Cell Biol.* 13:7000-7005.

White, K., M.E. Grether, J.M. Abrams, L. Young, K. Farrell, and H. Steller. 1994. Genetic control of programmed cell death in Drosophila. *Science.* 264:677-683.

White, K., E. Tahaoglu, and H. Steller. 1996. Cell killing by the Drosophila gene reaper. *Science.* 271:805-807.

Winkles, J.A. 1998. Serum- and polypeptide growth factor-inducible gene expression in mouse fibroblasts. *Prog Nucleic Acid Res Mol Biol.* 58:41-78.

Zapata, J.M., S. Matsuzawa, A. Godzik, E. Leo, S.A. Wasserman, and J.C. Reed. 2000. The Drosophila tumor necrosis factor receptor-associated factor-1 (DTRAF1) interacts with Pelle and regulates NFkappaB activity. *J Biol Chem.* 275:12102-12107.

Zhang, H., and S.W. Emmons. 2000. A C. elegans mediator protein confers regulatory selectivity on lineage- specific expression of a transcription factor gene. *Genes Dev.* 14:2161-2172.

CHAPTER 3

GRP50 Is A Brefeldin A-Sensitive Peripheral Golgi Protein

INTRODUCTION

The Golgi apparatus is a complex organelle, which not only mediates the modification of newly synthesized proteins and lipids, but also serves as the major sorting point in the secretory pathway (Farquhar and Palade, 1998). In mammalian cells, it appears in microscopic studies as a ribbon-like structure adjacent to the nucleus and is composed of multiple stacked membrane-bounded compartments, termed cisternae. In the classical view of transport through the secretory pathway, membrane-bounded vesicles containing newly synthesized cargo proteins bud from the endoplasmic reticulum (ER) and fuse with the *cis*-Golgi (Schekman and Orci, 1996). Similar yet biochemically distinct vesicles transport cargo through the *cis*-, *medial*-, and *trans*-Golgi and to the *trans*-Golgi network (TGN), where they are sorted to their correct destinations (e.g., cell surface, endosome, lysosomes, or secretory granules, (Glick, 2000; Traub and Kornfeld, 1997).

A number of cellular proteins are specifically targeted to the Golgi apparatus, and these fall into two predominant classes. The first of these includes the multipass transmembrane proteins that transport molecules across the bilayer as well as the enzymes that mediate post-translational processing of secretory cargo as it passes through the lumen of the Golgi. For example, the glycosyltransferases, which are believed to comprise the bulk of Golgi membrane proteins, are type II integral membrane proteins whose catalytic domains face the Golgi lumen (Shaper et al., 1988; Taatjes et al., 1992). These proteins reside exclusively in the membranes of the Golgi apparatus.

The second category of Golgi-associated proteins includes both integral and peripheral membrane proteins involved in maintaining the structural organization of the Golgi and in regulating membrane transport events such as vesicle budding, transport, and fusion (Gleeson, 1998; Munro, 1998). The peripheral membrane proteins of this latter category are associated with the cytoplasmic surface of Golgi and vesicle membranes, to which they are transiently recruited from soluble cytoplasmic pools. At the membrane, these proteins associate with the cytoplasmic domains of integral membrane proteins and/or with other cytosolic proteins to form multiprotein complexes of diverse functions (Guo et al., 2000; Robinson et al., 1997; Robinson, 1994; Simpson et al., 1997).

Several of the multiprotein complexes that preferentially associate with Golgi and vesicle membranes have been characterized extensively. Among the best understood of these are the clathrin/adaptor protein (AP) complexes, which are involved in receptor-mediated endocytosis and in the selective transport of proteins from the TGN to lysosomes or endosomes (Mostov et al., 2000; Pearse, 1988; Robinson, 1994). In addition, over the past decade the components of the non-clathrin coat complexes that define coat protein I (COPI) and coat protein II (COPII) vesicles have been elucidated. Coatomer is a hetero-oligomeric complex that associates with coat protein I (COPI)-coated vesicles, which are implicated in both anterograde intra-Golgi transport and retrograde transport from the Golgi to the ER (Cosson and Letourneur, 1997; Kreis et al., 1995; Orci et al., 1997). A distinct set of proteins assembles to form the complex that coats COPII vesicles, which transport cargo from the ER to the Golgi (Bannykh and Balch, 1998; Barlowe et al., 1994).

Although some of the individual proteins that comprise the above-mentioned types of complexes are share limited sequence homology, they are structurally and functionally distinct. The common features among these complexes are that they are required for vesicle budding at donor membranes, that their components are recruited to the membrane from cytosolic pools, and that this assembly requires hydrolyzable GTP (Barlowe and Schekman, 1993; Donaldson et al., 1992; Melancon et al., 1987; Robinson and Kreis, 1992; Simpson et al., 1996). Coat complex formation begins when a small GTPase is recruited to the donor membrane by its cognate guanine nucleotide exchange factor (GEF). Binding of the GTPase, in its GTP-bound form, permits assembly of additional coat proteins onto the membrane. Activation of the GTPase then results in hydrolysis of GTP, release of the GTPase (in its GDP-bound form) from the membrane, and termination of the cycle of coat complex assembly (Guo et al., 2000; Schekman and Orci, 1996; Spang et al., 1998; Springer et al., 1999).

In addition to the molecules mentioned above, which are implicated in the formation of vesicles from donor membranes, other Golgi and vesicle associated integral and peripheral membrane protein complexes are required for fusion of vesicles with acceptor membranes and for maintenance of the structure of the Golgi apparatus itself. The importance of these complexes was first elucidated in neurons, where the soluble proteins NSF and SNAP are recruited from the cytoplasm into a complex with integral membrane proteins on vesicles (v-SNAREs) and on the plasma membrane (t-SNAREs); the association of these proteins is required for membrane fusion (Mochida, 2000; Sollner et al., 1993). In the Golgi apparatus, fusion of COPI vesicles with Golgi membranes is mediated by the peripheral membrane proteins p115 and GM130, which tether the

membranes through interactions with Giantin and GRASP65 (Barr et al., 1998; Nakamura et al., 1997; Sonnichsen et al., 1998). Additional studies have implicated p115 in the reassembly of the Golgi apparatus following mitosis (Lowe et al., 1998; Shorter and Warren, 1999) and in the general maintenance of Golgi structure in interphase cells (Puthenveedu and Linstedt, 2001).

Golgi- and vesicle-associated proteins are often categorized according to their relative resistance or sensisitivity to Brefeldin A (BFA), a small hydrophobic compound produced by toxic fungi, which exerts pleiotropic effects on the Golgi apparatus (Chardin and McCormick, 1999; Dinter and Berger, 1998; Roth, 1999). BFA has been shown to promote dissociation of coatomer and other multiprotein complex components from Golgi membranes (Donaldson et al., 1990; Podos et al., 1994) and ultimately leads to the redistribution of Golgi membranes (and their lumenal cargo) to the ER (Dascher and Balch, 1994; Fujiwara et al., 1988; Lippincott-Schwartz et al., 1989; Sciaky et al., 1997). The effects of BFA are due to inhibition of ADP-ribosylation factors (ARFs) (Donaldson and Jackson, 2000; Donaldson et al., 1990) a family of Ras-related small GTPases whose activity is required for coat protein assembly onto budding vesicles (Schekman and Orci, 1996; Springer et al., 1999).

The previous section described the cloning and characterization of GRP50, which encodes a novel mammalian protein containing multiple protein-interaction domains. The results below show that GRP50 is peripherally associated with the cytoplasmic face of the Golgi apparatus. Further, Golgi-association of GRP50 is sensitive to the effects of Brefeldin A and may require the activity of one or more AFF-GEFs. Localization of GRP50 to the Golgi appears to require sequences in the carboxyl-terminal domain of

GRP50 which are predicted to form a coiled coil. Additionally, these results show that GRP50 co-purifies from cultured cells as part of a large macromolecular protein complex, and suggest that it may be associated with other proteins *in vivo* via intermolecular interactions.

RESULTS

GRP50 localizes to the Golgi apparatus in cultured fibroblasts

In order to facilitate analysis of the cellular function of the GRP50 protein, we generated a series of GRP50 constructs in mammalian expression vectors. In the first of these, the entire coding sequence of GRP50 was ligated in-frame to a pCDNA3-derived vector coding for six NH_2-terminal myc-epitope tags (vector pJF-Tag, see Materials and Methods). A second construct, in which a single myc-epitope tag (11 amino acids) was added to the NH_2-terminus of GRP50 using the polymerase chain reaction (PCR), was generated in pCDNA3. These constructs, schematic representations of which are shown in Figure 1A, permit detection of the resultant fusion proteins using commercially available anti-myc tag antibodies.

The myc_6-GRP fusion protein has a calculated molecular mass of 51 kilodaltons (kDa), and an apparent molecular mass, as measured by migration in SDS-polyacrylamide gel electrophoresis (SDS-PAGE), of ~62 kDa (Figure 1B). The increased apparent molecular mass is not due to post-translational modification, since migration of myc_6-GRP present in transfected cos7 cell lysates is indistinguishable from that of in vitro translated myc_6-GRP (Figure 1B). The myc_1-GRP fusion protein, which has a predicted molecular mass of 42 kDa, also exhibits an increased apparent molecular mass (~51 kDa) in SDS-PAGE (Figure 1B, right panel). It is likely that the decreased mobility of myc-GRP proteins in SDS-PAGE is due to the primary structure of the GRP50 protein (which is highly proline-rich). Like myc_1-GRP, endogenous GRP50 (calculated at 41 kDa) is therefore expected to migrate as an apparent ~50 kDa protein in SDS-PAGE.

Analysis of the GRP50 protein sequence using the PSORT algorithm yielded a 60% probability that GRP50 is a nuclear protein, based on the overall content of basic residues (Reinhardt and Hubbard, 1998). To determine the subcellular distribution of myc_6-GRP, cells transfected with the myc_6-GRP construct were analyzed using indirect immunofluorescence microscopy. Immunofluorescent labeling of myc_6-GRP transfected Rat-1 fibroblasts revealed a distinct perinuclear staining pattern in interphase cells and a diffuse overall staining in mitotic cells (Figure 1C, upper panels). Although the short (approximately 12 kDa) myc_6 peptide expressed from the empty vector exhibits weak nuclear staining (1C, lower panels), the myc_6-GRP fusion protein is excluded from the nucleus. The staining pattern observed in myc_6-GRP transfected cells was reproducible in all cell lines examined (including NIH/3T3, 293, and differentiated PC12 cells, data not shown), and is characteristic of the Golgi apparatus in cultured mammalian cells (Warren, 1993).

To determine whether myc_6-GRP is localized to Golgi membranes, pools of stably transfected Rat-1 fibroblasts were analyzed by double-immunofluorescent labeling using anti-myc epitope antibodies, together with antibodies to organelle marker proteins (Figure 2). In transfected cells, myc_6-GRP labeling (Figure 2A, left panels) exhibited considerable overlap with that of the Golgi- and vesicle-associated protein β-COP (2A, middle panels). β-COP is a subunit of coatomer, a multiprotein complex associated with coat protein I (COPI)-coated vesicles, which transport cargo within and from the Golgi apparatus (Duden et al., 1994; Orci et al., 1997). Myc_6-GRP staining was observed in the Golgi and in small punctate structures throughout the cytoplasm but was not associated

with the endoplasmic reticulum (ER); figure 2B shows that the staining pattern of myc_6-GRP does not overlap with that of the ER-resident protein BiP.

Although the observed co-localization of myc_6-GRP with the Golgi marker protein β-COP was striking, this staining pattern could be an artifact of ectopic overexpression of the tagged protein. To permit analysis of the endogenous GRP50 protein, rabbit polyclonal antiserum was raised against the carboxyl-terminal 126 amino acids of murine GRP50, expressed as a hexahistidine fusion protein in *E. coli* (see Materials and Methods). Whole serum from immunized animals failed to detect either endogenous or transfected GRP50 (over background bands) on western blots of total cell lysates (Figure 3A, lanes 1 and 2 and data not shown). However, whole serum was successfully used to immunoprecipitate myc_6-GRP from total cell lysates of transfected cells (not shown).

In order to reduce background in these experiments due to nonspecific serum immunoglobulins, antibodies specific for GRP50 were isolated from immune serum by affinity purification on an immobilized antigen column (see Materials and Methods). The affinity-purified anti-GRP antibodies weakly detect a single band of ~62 kDa on immunoblots of total cell lysate from myc_6-GRP transfected cos7 cells (Figure 3A, lane 6); this band is recognized by immune serum but not by preimmune serum (3A, lanes 1-4), and is the same size as the myc_6-GRP protein (3A, lane 8) detected by the anti-myc epitope antibody. Although the purified antibody failed to detect the endogenous GRP50 protein at 50 kDa (likely present at much lower levels than myc_6-GRP in cos7 cell extracts), the absence of contaminating bands in Figure 3, lane 6 suggested that the anti-GRP antibodies were relatively pure.

The affinity-purified anti-GRP antibodies retained the ability to immunoprecipitate the myc_6-GRP fusion protein from total cell lysates (Figure 3B). Purified anti-GRP (left lanes) or anti-myc epitope (polyclonal, right lanes) antibodies were incubated with total cell lysates from vector-transfected or myc_6-GRP transfected EcR-293 cells (see Materials and Methods) and precipitated using protein-A agarose. Figure 3B shows detection of myc_6-GRP in both anti-GRP and anti-myc immunoprecipitates analyzed by western blotting with the anti-myc epitope monoclonal antibody, 9e10. Myc_6-GRP was not detected in immunoprecipitates obtained using preimmune rabbit serum (not shown).

In a related experiment, anti-GRP or anti-myc antibodies were used to immunoprecipitate proteins from total cell lysates of ^{35}S-methionine labeled cells, and the resultant proteins were analyzed by SDS-PAGE (Figure 3C). Purified anti-GRP immunoprecipitates a major protein of ~62 kDa from myc_6-GRP transfected EcR-293 cells, as does the anti-myc antibody; this protein is likely the myc_6-GRP fusion (Figure 3C, lanes 2 and 4, upper arrow). In addition, both anti-GRP and anti-myc immunoprecipitates from myc_6-GRP transfected cells contain a second major protein of ~50 kDa (3C, lower arrow). Because this protein is present in anti-myc immunoprecipitates from myc_6-GRP transfected but not from vector-transfected cells (compare 3C, lanes 3 and 4), it might represent a protein that co-purifies with myc_6-GRP using this method. The 50 kDa protein is also present in anti-GRP immunoprecipitates from vector-transfected cells (3C, lane 1), and it migrates at the size expected for endogenous GRP50. These results suggest that the 50 kDa protein may be the

endogenous GRP50 protein, and that endogenous GRP50 and myc_6-GRP could be associated with each other, directly or indirectly, in myc_6-GRP EcR-293 cells.

In order to determine whether the subcellular distribution of endogenous GRP50 resembles that of exogenous myc_6-GRP, we used affinity purified anti-GRP in indirect immunofluorescence microscopy. The staining patterns detected by anti-GRP (Figure 4A) and by anti-β-COP (Figure 4B) in NIH/3T3 cells are remarkably similar; endogenous GRP50 exhibits the perinuclear localization characteristic of the Golgi apparatus in interphase cells. The staining is diffuse in the mitotic cells shown (arrows in A correspond to mitotic nuclei in B), when Golgi cisternae fragment prior to cell division (Lucocq et al., 1989; Lucocq and Warren, 1987). These results imply that endogenous GRP50, like exogenous myc_6-GRP, is a Golgi-associated protein.

GRP50 is a Brefeldin A-sensitive peripheral Golgi protein

The predicted GRP50 amino acid sequence (Figure 5) does not contain any potential membrane-spanning domains, and it does not appear to possess a signal sequence for translocation into the ER. Although the protein contains consensus sequences for N-linked glycosylation (indicated with "*" in Figure 5), the fact that in vitro translated and exogenously expressed myc_6-GRP exhibit identical electrophoretic mobilities (Figure 1B) implies that the protein is not normally glycosylated. GRP50 does, however, contain sequences in its leucine-rich region that are likely to form an α-helical coiled coil structure, according the coiled coil prediction program COILS (Lupas, 1996; Lupas et al., 1991). In recent years, the coiled coil has emerged as a key structural element that mediates the association of soluble cytoplasmic proteins with Golgi and

vesicle membranes (Clague, 1999; Kjer-Nielsen et al., 1999). These observations suggest that GRP50 is a soluble protein that does not enter the secretory pathway, and that the protein most likely resides at the cytoplasmic, rather than lumenal, face of the Golgi apparatus.

To test this hypothesis, we investigated the subcellular distribution of GRP50 in response to treatment of cells with the fungal metabolite, Brefeldin A (BFA), which promotes dissociation of coatomer and other multiprotein complex components from Golgi membranes (Dinter and Berger, 1998; Donaldson et al., 1990; Podos et al., 1994). As discussed above, BFA acts by inhibiting the ARF family of small GTPases, which are required for recruitment, onto membranes, of coat proteins from soluble cytosolic pools (Donaldson and Jackson, 2000; Donaldson et al., 1990; Schekman and Orci, 1996). BFA therefore permits the classification of Golgi-associated proteins based on their relative sensitivities to this drug: following a short exposure to BFA, peripherally associated proteins (such as the coatomer subunit β-COP) rapidly redistribute to the cytoplasm, while integral membrane proteins (such as the enzyme mannosidase II) remain associated with the Golgi. After a longer exposure to BFA, integral membrane proteins and lumenal cargo of the Golgi redistribute to the ER (Dascher and Balch, 1994; Sciaky et al., 1997).

The redistribution of GRP50 following BFA treatment more closely resembles that of β-COP (Figure 6, A and B) than that of mannosidase II (Figure 6C). All three proteins exhibit Golgi localization in untreated cells (Figure 6, left panels). After exposure to BFA for two minutes, both GRP50 and β-COP showed a reduced staining of the perinuclear region and an increased overall diffuse staining (6A and B, middle panels), and by five minutes (right panels) the perinuclear staining was almost completely

abolished. Although GRP50 staining was less diffusely cytoplasmic than β-COP after five minutes, it was clearly no longer confined to the Golgi apparatus. In contrast, mannosidase II staining (Figure 6C) was unchanged after two and five minutes of BFA treatment, indicating that the overall morphology of the Golgi apparatus was not disrupted in these experiments. These results suggest that GRP50 is peripherally associated with Golgi membranes, and that this association is dependent on the activity of a BFA-sensitive ARF.

Golgi-localization is mediated by the C-terminal domain of GRP50

The observation that GRP50 association with the Golgi apparatus can be abolished by BFA implies that this association is a reversible event, likely mediated by interaction of a structural domain of GRP50 with an unknown component of these membranes. To investigate whether a particular structural domain of GRP50 is required for Golgi localization, additional constructs were generated in the myc_6-epitope vector described above (see Materials and Methods) which permit expression of truncated versions of myc_6-GRP. Analysis of the subcellular distribution of these truncated myc_6-GRP fusion proteins in transiently transfected Rat-1 fibroblasts is shown in Figure 7. Deletion of the carboxyl terminal third of GRP50 (Figure 7, B, C, and E) abolished Golgi localization, while all fusion proteins containing this domain were targeted to Golgi-like structures to a greater (D and G) or lesser degree (F). Overexpression of truncated myc_6-GRP fusion proteins did not grossly affect the overall morphology of the Golgi apparatus, as determined by β-COP staining (not shown). Given these results, it appears that the

leucine-rich carboxyl terminal domain of GRP50, which contains potential coiled coil forming sequences, is necessary for Golgi localization of the protein.

Myc_6-GRP is part of a large macromolecular complex

Many peripheral Golgi and vesicle proteins are present in cells as part of large hetero-oligomeric protein complexes. In addition, preliminary fractionation and immunoprecipitation experiments (e.g., Figure 3C) suggested that GRP50, which contains multiple putative protein interaction domains, might be associated with other proteins in the cell. In order to explore whether GRP50 is normally part of a multiprotein complex, the elution profile of myc_6-GRP in gel chromatographic fractionation was investigated. Because our affinity purified anti-GRP antibodies perform poorly in immunoblotting assays (Figure 3A), size fractionation experiments were conducted using extracts from stably transfected myc_6-GRP Rat-1 cells (see Materials and Methods).

In this procedure, concentrated cytoplasmic extracts were prepared in detergent-free buffer and subsequently size-fractionated using FPLC on a Superose 6 column (AP Biotech). Forty-eight fractions were collected and analyzed by immunoblotting with anti-myc-epitope antibody and antibodies to β-COP and ldlCp, two peripheral membrane proteins known to elute as components of high-molecular weight oligomers using this method (Chatterton et al., 1999; Waters et al., 1991). The bulk of cellular proteins, as measured by spectrophotometric absorption at 280 nm, were eluted from the column in fractions 35-42 (data not shown). In contrast, the elution profiles of myc_6-GRP, ldlCp, and β-COP (Figure 8, A and B) indicate that these proteins purify as components of high molecular weight complexes.

Under these conditions, myc_6-GRP was found exclusively in the high molecular weight fractions numbered 20-25, with an elution peak in fraction 23 (Figure 8A). As previously reported, ldlCp, a member of the ~950 kDa "ldlCp complex" (Chatterton et al., 1999) eluted in fractions 18-21 (Figure 8B, left panel), and β-COP, a component of the ~700 kDa coatomer complex (Waters et al., 1991), eluted in fractions 24-28 with a peak in fraction 25 (8B, right panel). The size of the myc_6-GRP complex is estimated to be ~800 kDa, based on comparison of the elution peaks of these proteins. Surprisingly, no myc_6-GRP was detected in later fractions (Figure 8C, top and data not shown) where a ~60 kDa monomer would be expected to elute. β-tubulin, which has a monomeric mass of 55 kDa, was detected in fractions 32-35 (Figure 8C, bottom). These results suggest that, in the soluble fraction of cells, myc_6-GRP is present mostly or exclusively as a component of a large protein complex (or complexes).

Two proteins of ~50 kDa co-immunoprecipitate with myc_6-GRP

Because the results of the fractionation experiments described above suggest that myc_6-GRP may exist in cells as part of a soluble multiprotein complex, immunoprecipitation experiments were used to probe the composition of the putative myc_6-GRP-assocciated structures. As described above (Figure 3C), earlier results suggested that additional protein species co-immunoprecipitated with both endogenous GRP50 and myc_6-GRP. Since at least one of these proteins migrates at ~50 kDa (the expected size of endogenous GRP50), immunoprecipitates were analyzed from myc_6-GRP transfected cells, to permit distinction between the immunoprecipitated GRP protein and any associated proteins.

Robust expression of full length and truncated myc_6-GRP fusion proteins (depicted in Figure 9A) was achieved by generating stable transfectants in EcR-293 cells, which allow expression of the proteins in an ecdysone-inducible manner (No et al., 1996, see Materials and Methods). The myc_6-GRP fusion proteins were immunoprecipitated from total cell lysates of ^{35}S-methionine labeled cells and analyzed by SDS-PAGE (Figure 9B). The bands corresponding to myc_6-GRP fusion proteins are indicated with black arrows in Figure 9B (compare with Figure 9C, left panel). At least 2 proteins were consistently observed in anti-myc immunoprecipitates from myc_6-GRP transfected cells that were not present in immunoprecipitates from vector-transfected cells (Figure 9B, lanes 1 and 2). These proteins are similar in size, both migrating at ~48-50 kDa, and are present in immunoprecipitates from cells transfected with full length myc_6-GRP (construct #1) and with construct #4, which lacks the animo-terminal ankyrin repeat region. The co-immunoprecipitated proteins are not present, however, in immunoprecipitates from cells transfected with construct #3, which lacks the carboxyl-terminal third of GRP50 (containing the putative coiled coil). Additional protein species were observed in immunoprecipitates from cells transfected with construct #3, though these species were not present in immunoprecipitates from cells expressing the full length myc_6-GRP.

Endogenous GRP50 co-immunoprecipitates with myc_6-GRP

The additional proteins observed in immunoprecipitates from myc_6-GRP expressing cells do not appear to be proteolytic fragments of the fusion proteins. These proteins are not detected by the anti-myc antibody on immunoblots (Figure 9C, left

panel), and proteolytic fragments lacking the myc epitope tags should not be immunoprecipitated by the anti-myc-tag antibody. As described above, a protein of ~50 kDa was previously detected in anti-GRP immunoprecipitates from vector transfected cells, and we speculated that this 50 kDa protein could be the endogenous GRP50 species.

To assess whether one of the ~50 kDa proteins detected in the current experiments was the endogenous GRP50, the affinity purified anti-GRP antibody was used to probe immuoblots containing the anti-myc immunoprecipitates (Figure 9C, right panel). Of the five myc_6-GRP fusion proteins expressed in these cells, the anti-GRP antibody detects only the full length fusion (construct #1) and the fusion containing the intact carboxyl terminus (construct #4); this result was expected because only the carboxyl terminal region of GRP50 was used as the immunogen in the generation of these antibodies. However, the anti-GRP antibodies also detect one of the ~50 kDa proteins present in anti-myc immunoprecipitates from cells expressing these fusion proteins (indicated by red arrow in Figure 9C, right panel). These results suggest that this ~50 kDa protein is in fact the endogenous GRP50, and that endogenous GRP50 is capable of associating (directly or via an intermediate protein) with myc_6-GRP in lysates from transfected cells. In addition, this association appears to require sequences in the carboxyl terminus of GRP50, since it is observed only when the intact C-terminal domain is present in the fusion protein. To date, the identity of the other proteins present in these immunoprecipitates has not been determined.

DISCUSSION

The results described above identify GRP50 as a novel peripheral Golgi associated protein. The localization of GRP50 to the Golgi apparatus was initially observed in cells in which myc-tagged GRP50 was ectopically overexpressed, and was confirmed by immunostaining of endogenous GRP50. The overall structural architecture of GRP50 suggests that the protein is a soluble cytoplasmic molecule that is peripherally associated with Golgi and vesicle membranes, perhaps via interactions with resident integral or peripheral membrane proteins. The experiments demonstrating that localization of GRP50 to the Golgi apparatus is Brefeldin A sensitive bolster this claim, as BFA has been shown to cause dissociation of many peripherally-associated proteins from the membranes of the Golgi (Donaldson et al., 1990; Podos et al., 1994).

Although the results showing immunolocalization of GRP50 to the Golgi apparatus in cultured fibroblasts are compelling, they almost certainly represent an incomplete picture of the cellular role of this protein. First, the resolution of the light microscope does not permit clear distinction between the various subcompartments of the Golgi and the TGN-endosomal system, and therefore we cannot discern from these experiments whether GRP50 is restricted to a particular region of this system. In these experiments, both endogenous GRP50 and ectopically expressed myc_6-GRP were detected in punctate structures scattered throughout the cytoplasm, suggesting that the function of the protein is not restricted to the Golgi apparatus proper. Moreover, in immunofluorescent labeling assays, the concentration of some protein molecules in close physical proximity in a defined structure can result in increased local fluorescence, which

in our experiments may have obscured the signal of diffusely cytosolic molecules. Finally, we have investigated the localization of GRP50 in a limited number of cell types in culture, and our experiments do not offer any indication of the subcellular distribution of this protein in other types of cells. While these results therefore permit us to conclude that GRP50 is associated with the Golgi apparatus, they do not necessarily imply that the protein is found only in this location.

In fact, several lines of evidence suggest that the association of GRP50 with Golgi membranes is a transient event and that GRP50 is recruited to these membranes from a cytosolic pool. First, the GRP50 sequence suggests that the protein is not fatty-acylated and targeted to membranes by this mechanism. In addition, the myc_6-GRP fusion protein was detected (on immunoblots) in the soluble fraction of cells disrupted by mechanical force in the absence of detergents. Although it is possible that the proteins found in the soluble fraction represent Golgi-bound myc_6-GRP that was disrupted in the process of homogenization, a more likely interpretation is that GRP50 is present in both membrane-associated and soluble forms in cells. Lastly, the fact that GRP50 association with the Golgi apparatus is sensitive to the effects of BFA implies that this association is a regulated, rather than constitutive, event, which occurs under specific conditions conducive to binding.

In recent years, the mechanism by which BFA promotes dissociation of some peripheral membrane proteins from the Golgi apparatus has been characterized in some detail. This disruption is due to the effects of BFA on ADP-ribosylation factors (ARFs), a family of small G-proteins which, in their GTP-bound form, are required for the assembly of vesicle coat proteins (Schekman and Orci, 1996). BFA inhibits the guanine

nucleotide exchange factors (GEFs) that stimulate the exchange of GDP for GTP on ARFs (Peyroche et al., 1999). BFA acts as a noncompetitive inhibitor of ARF, by binding to the normally transient complex of the GEF and ARF-GDP, resulting in stabilization of an inactive complex and in reduction of the amount of free GEF available to bind (and activate) other ARFs (Chardin and McCormick, 1999). Thus the initial effect of BFA treatment is the disruption of complexes whose association depends of the activities of ARFs and of BFA-sensitive ARF-GEFs. The kinetics of GRP50 redistribution to the cytoplasm in response to BFA treatment suggests that association of GRP50 with the Golgi apparatus is an ARF (or ARF-GEF)-dependent process.

We used myc_6-GRP fusion proteins to identify the region of GRP50 that mediates localization to the Golgi apparatus. These experiments indicated that the carboxyl-terminal region of GRP50, which contains a potential coiled coil domain, is required for targeting of the myc_6-GRP fusion proteins to the Golgi apparatus. The α-helical coiled coil mediates oligomerization in many proteins, and has recently been implicated in the targeting of several peripheral membrane proteins to the Golgi (Kjer-Nielsen et al., 1999). In our experiments, myc_6-GRP fusions lacking the C-terminal region of GRP50 were no longer targeted to the Golgi, but were concentrated near the nucleus. The exception to this pattern was observed in cells expressing a myc_6-GRP fusion that contained only the central, proline-rich domain; this protein localized exclusively to the nucleus. In transiently transfected cells, none of the myc_6-GRP fusion proteins were found to be diffusely cytosolic.

Although these experiments do not implicate specific residues in binding to the Golgi, they do suggest that residues in the C-terminal domain are required for Golgi

localization. The observation that a myc_6-GRP fusion containing only this C-terminal region of GRP50 is at least capable of Golgi localization argues that this domain is sufficient for the appropriate targeting of GRP50. We cannot conclude whether the mislocalization of myc_6-GRP fusion proteins that lack this domain mimics an alternative distribution pattern for endogenous GRP50, since these effects could easily be due to overexpression, to misfolding of the truncated proteins, or both.

Gel chromatographic fractionation of cytosol prepared from myc_6-GRP transfected cells resulted in the elution of myc_6-GRP in fractions corresponding to high-molecular weight complexes. Based upon the elution profiles of ldlCp and β-COP, we estimate the size of the myc_6-GRP complex to be ~800 kDa. To our surprise, we did not detect myc_6-GRP in any of the later fractions, where a ~62 kDa monomer would be expected to elute. One interpretation of this result is that myc_6-GRP is normally present in cells as part of a large protein complex and that very little protein exists as a monomer. Even if this is the case, however, we cannot conclude that the same is true for endogenous GRP50, since overexpression of the myc_6-GRP fusion protein could easily have shifted the equilibrium in favor of complex formation.

Additionally, it is possible that the extra domain provided by the myc_6-epitope tag results in a fusion protein with very different binding properties than those of the endogenous protein. These same caveats must be applied to the estimation of the size of a putative GRP50 complex, because the myc_6–tag might reasonably be expected to alter the chromatographic elution profile of the protein. Still, these results suggest that GRP50 could exist in cells as a member of a large macromolecular complex; the structural features of this protein, coupled with its subcellular localization, contribute to the

plausibility of this interpretation. Ultimately, however, a meaningful interpretation of these results requires that these experiments be repeated in untransfected cells and that the elution profile of endogenous GRP50 be assessed.

The above caveats apply equally to the analyses of our immunoprecipitation experiments. The results of these experiments suggest that myc_6-GRP associates, in cells or in total cell lysates, with at least two protein species of ~50 kDa. That one of these ~50 kDa proteins is detected by affinity-purified anti-GRP antibody implies both myc_6-GRP and endogenous GRP50 are present in a complex that co-immunoprecipitates with myc_6-GRP. In addition, the observation that this protein species is present only when the myc_6-GRP fusion protein contains the intact carboxyl-terminal domain of GRP50 argues that the interaction of myc_6-GRP with the complex requires this domain. Confirmation of these results would be aided, however, by a demonstration that this interaction could be abolished by a single amino acid substitution.

The fact that the carboxyl-terminal domain contains sequences with the potential to form an α-helical coiled coil suggests the possibility that GRP50 could be present in cells as an oligomer, since this structure is among the principal oligomerization domains found in proteins (Burkhard et al., 2001). However, our results do not directly address this possibility, since the observed co-immunoprecipitation of GRP50 with. myc_6-GRP could easily be due to the presence of one or more additional protein species that interact directly with the GRP proteins. Additional experiments will be required to address this issue.

The data described above demonstrate that GRP50 is associated with the Golgi apparatus in a Brefeldin A-sensitive manner, and suggest that this association is almost

certainly mediated by intermolecular interactions with other proteins. The structural features of the GRP50 protein argue strongly for a function conferred by protein-protein interactions. Our results, although qualified by the difficulties inherent in ectopic overexpression studies, suggest the existence of a GRP50 protein complex, the nature of which merits further investigation.

MATERIALS AND METHODS

Expression Constructs

The myc_6-GRP fusion constructs were constructed by subcloning various restriction fragments of the GRP50 open reading frame into a pcDNA3 (Invitrogen) derived vector containing sequences coding for six NH_2-terminal myc epitope tags followed by a multiple cloning site (pJF-Tag, a gift of Alain Charest). The single-myc-tagged GRP50 construct was generated via the polymerase chain reaction using a proofreading thermostable polymerase (PfuTurbo, Stratagene) and the following primers: 5'–acggatccaccatggaacaaaaacttatttctgaagaagatctgatgagcgccctcctagagc-3' (where the 5' sequence encodes the 11 amino acid myc epitope tag, and the 3' sequence corresponds to the first 20 nucleotides of the GRP50 open reading frame), and 5'–agctcgagtcagtacgtcagttttgaagctctcc-3' (corresponding to the last 25 nucleotides of the GRP50 open reading frame). The PCR product was digested with BamHI and XhoI and ligated into the mammalian expression vector pcDNA3 (Invitrogen); this construct was sequenced fully to ensure that no mutations were generated during the PCR amplification step. In order to permit stable, inducible expression of myc-GRP fusion proteins, the entire coding sequence of each of the myc_6-GRP fusion constructs (including the myc_6-tags) was excised from pJF-Tag and subcloned into the expression vector pIND (Invitrogen). In pIND constructs, expression of the fusion protein is driven by an ecdysone-inducible promoter and can be selectively modulated by addition of an ecdysone analog to the culture medium (see below).

Cell Culture

Rat-1 fibroblasts were maintained under standard conditions in Dulbecco's Modified Eagle's Medium (DMEM), supplemented with 5% heat-inactivated Fetal Calf Serum (FCS), 50 U/ml penicillin, 50 μg/ml streptomycin, and 2mM L-glutamine. NIH/3T3 cells were cultured in DMEM plus 10% calf serum (CS) and the above concentrations of penicillin, streptomycin, and L-glutamine. Myc-GRP-transfected Rat-1 and NIH/3T3 cells were maintained in the above media plus 600 μg/ml G418 (Geneticin, GibcoBRL).

EcR-293 cells were purchased from Invitrogen as part of the Ecdysone-Inducible Mammalian Expression System (No et al., 1996) and maintained, according to the manufacturer's instructions, in DMEM plus 10% FCS, 2mM L-glutamine, and 400 μg/ml Zeocin™ (to maintain selection on the pVgRXR plasmid). EcR-293 cells transfected with myc-GRP constructs (or control plasmids) were maintained in this medium plus 400 μg/ml G418. In EcR-293 transfectants, expression of myc-GRP or control fusion proteins was induced by incubation in 1-5 μM ponasterone A (Invitrogen, diluted from 1 mM stock) in selective medium for 16-24 hours prior to harvesting or fixation.

All transfections were performed in serum-free medium (Optimem, GibcoBRL) using LipofectAMINE reagent (GibcoBRL), according to the manufacturer's recommendations. Generally, cells plated in 6-well dishes were transfected using 1 μg of column-purified (Qiagen) plasmid DNA and 3-6 μl LipofectAMINE reagent (depending upon the cell line used); components were scaled up for transfections in larger culture dishes. In transient expression experiments, transfected cells were incubated 36-48 hours prior to harvesting or fixation. Cell lines containing stably integrated expression

constructs were generated as above, except that plasmids were linearized prior to introduction into cells, and transfectants were diluted into selective medium, as described above, after 72 hours.

When indicated, cellular proteins were labeled with ^{35}S-methionine prior to harvesting for immunoprecipitation. For these experiments, cells in 100mm dishes were washed twice with PBS and incubated for one hour in methionine-free DMEM (GibcoBRL), followed by incubation for 5 hours in methionine-free DMEM plus 300-500mCi ^{35}S-methionine, supplemented with 1% dialyzed FCS. Myc-GRP-transfected EcR-293 cells were labeled as follows: Myc-GRP expression was inducted by addition of 1 μM ponasterone A to the culture medium. After 10 hours, the culture medium was removed, cells were washed twice with PBS and then incubated for 6 hours in methionine-free DMEM plus 300-500mCi ^{35}S-methionine, supplemented with 1% dialyzed FCS and 5 μM ponasterone A. This procedure was designed to ensure that transcription of myc-GRP (induced by ponasterone) was occurring at a high rate prior to addition of ^{35}S-methionine.

For immunofluorescence microscopy, cells were cultured on 22 x 22 mm glass coverslips (No. 1 thickness) in 6-well culture dishes, and treated or transfected as described above. For Brefeldin A (BFA) experiments, cells plated on coverslips were incubated in culture medium containing 5 μg/ml BFA (diluted from 5 mg/ml stock) for the indicated times (0, 2, or 5 minutes) and immediately fixed (without intermediate washing steps) and permeabilized as described below.

Immunofluorescence Microscopy

For indirect immunofluorescence microscopy, cells plated on glass coverslips were washed several times with phosphate-buffered saline (PBS), fixed in 2% formalin (2% formaldehyde in PBS) for 15 minutes at room temperature, washed with PBS, and permeabilized in 0.1% Triton-X100 for 15 minutes at room temperature. Prior to addition of primary antibodies, fixed samples were blocked in 10% goat serum (Sigma), 0.2% Tween-20 in PBS for a minimum of 30 minutes at room temperature. Primary antibodies were diluted in blocking buffer and incubated with cells for 4 hours to overnight at 4°C. Samples were washed extensively with PBS and incubated 1-2 hours with one of the following secondary antibodies, diluted in blocking buffer: Cy3-conjugated anti-mouse IgG (diluted 1:1000), Cy3-conjugated anti-rabbit IgG (1:1000), Texas Red-conjugated anti-mouse IgG (1:250), or FITC-conjugated anti-rabbit IgG (1:150). All secondary antibodies were obtained from AP Biotech, and the appropriate dilution range for each was determined empirically. Texas Red anti-mouse and FITC anti-rabbit IgG's were used together for all double labeling experiments, because the emission spectra of these two fluorochromes exhibit the least amount of overlap.

Following incubation with fluorochrome-linked secondary antibodies, cells were washed extensively with PBS and nuclei were counterstained with 4,6-diamidino-2-phenylindole (DAPI, added to penultimate PBS wash). Processed coverslips were mounted onto slides using Vectashield mounting medium (Vector Labs) and viewed using a Zeiss Axioplan2 microscope fitted with filters for fluorescence microscopy. Digital images were captured using a Quantix CCD camera (Photometrics Ltd.) driven by the IPLab Spectrum software package. Images were analyzed, false-colored, and merged

using IPLab Spectrum and subsequently exported to Adobe Photoshop for additional manipulations.

Antibodies

The following antibodies were obtained from the indicated commercial sources and were diluted (for immunostaining, immunoblotting, or immunoprecipitation) according to the manufacturers' recommendations: mouse monoclonal anti-myc tag (9E10 and 9E10-agarose conjugate) and rabbit polyclonal anti-myc tag (A-14) (Santa Cruz Biotechnology); mouse monoclonal anti-BiP (StressGen); rabbit polyclonal anti-β-COP (Affinity Bioreagents); mouse monoclonal anti-mannosidase II (BABCO); and mouse monoclonal anti-β-tubulin (Sigma). Affinity purified rabbit polyclonal antibodies to ldlCp were a kind gift of Jed Chatterton and were used as described previously (Chatterton et al., 1999). Fluorochrome-labeled secondary antibodies are described above (see Immunofluorescence Microscopy). Horseradish peroxide-labeled secondary antibodies were used in immunoblotting experiments as follows: HRP-conjugated anti-mouse IgG, FAB-specific (Sigma) was diluted 1:4000, and HRP-conjugated anti-rabbit IgG (AP Biotech) was diluted 1:10,000 in immunoblot blocking buffer (see below).

Antibodies to GRP50 were generated as follows: a part of the murine GRP50 cDNA encoding the carboxyl-terminal 126 amino acids of the protein was cloned into pQE (Qiagen) and expressed in *E. coli* strain XL1-Blue (Stratagene) as a hexahistidine fusion protein (his_6-C126). The his_6-C126 fusion protein was purified from bacterial inclusion bodies by adsorption to Ni-NTA-agarose (Qiagen) under denaturing conditions, according to the manufacturer's recommendations. Purified his_6-C126 was refolded by

gradual exchange from 8 M urea into physiological buffer by dialysis (at 4°C) against PBS/Urea. Beginning with 4 M Urea/PBS, the Urea concentration in the dialysis buffer was halved every two hours over a period of 12 hours, at which point samples were dialyzed for 2 hours against pure PBS. During the refolding procedure, the his_6-C126 protein formed insoluble aggregates (in 2 M Urea/PBS), which were subsequently used to generate custom rabbit polyclonal antisera at Covance, Inc. A fine suspension of purified his_6-C126 aggregates was used to immunize two female New Zealand white rabbits, and bleeds were performed and serum collected according to standard procedures (Harlow and Lane, 1988).

Affinity Purification of Anti-GRP Antibodies

We tested the rabbit antisera (obtained from Covance) for their ability to recognize recombinant GRP50 in immunoblotting and immunoprecipitation experiments; these experiments indicated a low titer of anti-GRP antibodies in whole sera. In addition, non-specific antibodies in the immune sera resulted in very high background in these experiments. For these reasons, we chose to affinity-purify anti-GRP antibodies from our rabbit antisera. Polyclonal antibodies specific for GRP50 were subsequently isolated and concentrated as follows: the entire GRP50 open reading frame was cloned into the pQE vector (Qiagen), expressed in *E. coli* strain BL21 as a hexahistidine fusion protein (his_6-GRP), and purified as described above. Unlike the his_6-C126 protein, however, the his_6-GRP fusion protein was successfully refolded into a soluble form during this procedure. The soluble his_6-GRP fusion protein was used to generate an antigen affinity column, as described below.

The soluble his_6-GRP protein was coupled to cyanogen bromide (CNBr)-activated Sepharose 4B (AP Biotech) according to the manufacturer's instructions. Briefly, 0.3 g (dry weight) CNBr-activated separose beads were resuspended in 1 mM HCl and subsequently washed for 15 minutes with 100 ml 1 mM HCl. Two milligrams of soluble his_6-GRP protein were dissolved in 10 ml coupling buffer (0.1 M $NaHCO_3$, 0.5 M NaCl, pH 8.3) at a concentration of 0.2 mg/ml and added to the washed beads (1 ml bed volume). The protein/sepharose mixture was incubated overnight at 4°C with constant rotation, and the antigen-coupled beads were collected by gravity sedimentation. Coupling efficiency was >90%, and was determined by measuring the protein concentration in the post-coupling flow through using the BCA protein assay kit (Pierce). The coupled beads were washed with 10 ml coupling buffer (to remove unbound antigen), and the coupling reaction was stopped by overnight incubation in 10 ml 1 M ethanolamine, pH 8. Antigen-coupled beads were loaded onto an EconoColumn (BioRad) and washed extensively with 10 column volumes (CV) of each of the following buffers: 10 mM Tris (pH 7.5), 0.5 M NaCl, followed by 100 mM glycine (pH 2.5), 0.5 M NaCl, followed by 10 mM Tris (pH 8.8), 0.5 M NaCl (until the flow-through was pH 8.8). The column was then washed with 10 CV each of 100mM triethylamine (pH 11), 0.5 M NaCl, followed by 10 mM Tris (pH 7.5), 0.5 M NaCl, followed by PBS. The washed antigen column was stored at 4°C in PBS plus 0.01% thimerosal (Sigma) until ready to use.

Immunoglobulins were collected from 50 ml of whole rabbit antiserum by saturated ammonium sulfate precipitation (Harlow and Lane, 1988), dissolved in 10 ml PBS, and dialyzed >2 hours against PBS (to remove excess salt). Concentrated antibodies were then added to the antigen-coupled sepharose beads, and the mixture was incubated

overnight at 4°C with constant rotation. The column was washed with 20 CV 10 mM Tris (pH 7.5), 0.5 M NaCl followed by 20 CV 10 mM Tris (pH 7.5), 0.1 M NaCl. Antibodies bound by acid-sensitive interactions were eluted in 10 CV 100 mM glycine (pH 2.5), 0.1 M NaCl and neutralized with 1 CV 1 M Tris (pH 8). The column was then washed with 10 CV 10 mM Tris (pH 8.8), 0.1 M NaCl, and base-sensitive antibodies were eluted with 10 CV 100 mM triethylamine (pH 11), 0.1 M NaCl and neutralized with 1 CV 1 M Tris (pH 8). Eluted antibodies were concentrated by saturated ammonium sulfate precipitation and redissolved in PBS.

Protein Preparation, Immunoprecipitations, and Immunoblotting

Total cell protein extracts were prepared by lysing cultured mammalian cells in either radio-immunoprecipitation assay buffer (RIPA, 50 mM Tris, pH 8, 150 mM NaCl, 1% (w/v) Triton X-100, 1% (w/v) NaDeoxycholate, 0.1% (w/v) SDS) or NP40 lysis buffer (50 mM Tris, pH 8, 150 mM NaCl, 1% (w/v) NonidetP-40) (Harlow and Lane, 1988), supplemented with 2 mM EGTA, 0.2 mM Na_3VO_4, 50 mM NaF, 1 mM PMSF, and a cocktail of protease inhibitors (Complete™, Roche). Lysates were prepared by adding lysis buffer to culture dishes (on ice) and scraping adherent cells into buffer. Lysates were incubated, with rotation, for 15 minutes at 4°C and clarified by centrifugation. Proteins were quantified using the BCA protein assay kit (Pierce).

For immunoblotting, 10 to 50 μg of total cell lysate was fractionated by SDS-PAGE according to standard methods (Harlow and Lane, 1988), and transferred to PVDF membrane (ImmobilonP, Millipore) using a Mini Trans-Blot™ cell (BioRad). Blots were blocked in PBS, 0.1% Tween-20 (PBST) plus 5% (w/v) nonfat dry milk. Primary

antibodies were added in blocking buffer and incubated 2 hours to overnight at 4°C. Blots were washed in PBST, incubated with HRP-conjugated anti-mouse or anti-rabbit IgG (diluted in PBST plus 3% (w/v) nonfat dry milk) for 30-60 minutes, and washed extensively in PBST. Bound antibodies were visualized using enhanced chemiluminescence (ECL, AP Biotech) according to the manufacturer's instructions.

For immunoprecipitations, total cell lysates were quantified and equal amounts of lysate were precleared by incubation for 1 hour with 50 µl normal rabbit serum and 50 µl (packed volume) heat-killed, fixed Staph A Cowan (SAC, Zymed Laboratories) followed by centrifugation for 10 minutes at 10,000 x g. Precleared lysates were incubated with 1-2 µg purified antibody for 3 hours to overnight at 4°C. Antibody-antigen complexes were isolated by addition of 50 µl of a 50% (v/v) slurry of protein-G agarose (for the anti-myc monoclonal antibody 9e10) or protein-A agarose (all other antibodies) followed by incubation for 1-2 hours at 4°C. (Protein-A and Protein-G agarose were purchased from Roche). Beads were collected by centrifugation and washed a minimum of four times in lysis buffer. Beads were boiled in SDS-PAGE loading buffer, and immunoprecipitated proteins were analyzed by SDS-PAGE, followed by autoradiography (for ^{35}S-labeled samples) or immunoblotting (for unlabeled protein samples).

Gel Chromatographic Fractionation

Cytosol was prepared, from 40 subconfluent 150 mm dishes of Rat-1 fibroblasts stably expressing myc_6-GRP, using a previously described method (Balch et al., 1984; Chatterton et al., 1999). Cells were harvested by trypsinization and washed first with medium, followed by 140 mM potassium acetate/10 mM triethylamine-acetic acid (pH

7.2) supplemented with 20 μg/ml type II trypsin inhibitor, and then with 125 mM potassium acetate/25 mM Hepes-KOH (pH 7.2) supplemented with 20 μg/ml type II trypsin inhibitor. Cells were homogenized using a ball bearing homogenizer (H&Y Enterprise, Redwood City CA) in 125 mM potassium acetate/25 mM Hepes-KOH (pH 7.2) supplemented with 0.1 units aprotinin, 5 μg/ml leupeptin, and 2 μg/ml type II trypsin inhibitor. The homogenate was clarified by centrifugation for one hour at 200,000 x g. Three hundred microliters of cytosol were size-fractionated by gel chromatography at a flow rate of 0.25 ml/min on a 24 ml Superose 6 column (AP Biotech). Forty-eight 0.5ml fractions were collected and 30μl aliquots were analyzed by SDS-PAGE followed by immunoblotting.

ACKNOWLEDGEMENTS

I thank Alain Charest for his kind gift of the pFJ-Tag expression vector and for many helpful discussions, Julian Borrow for advice regarding immunofluorescence experiments, and Jeff Trimarchi for advice regarding immunoprecipitation experiments. I am grateful to Jed Chatterton for assistance with gel chromatographic fractionation experiments and for his gift of the ldlCp antibody.

Figure 1

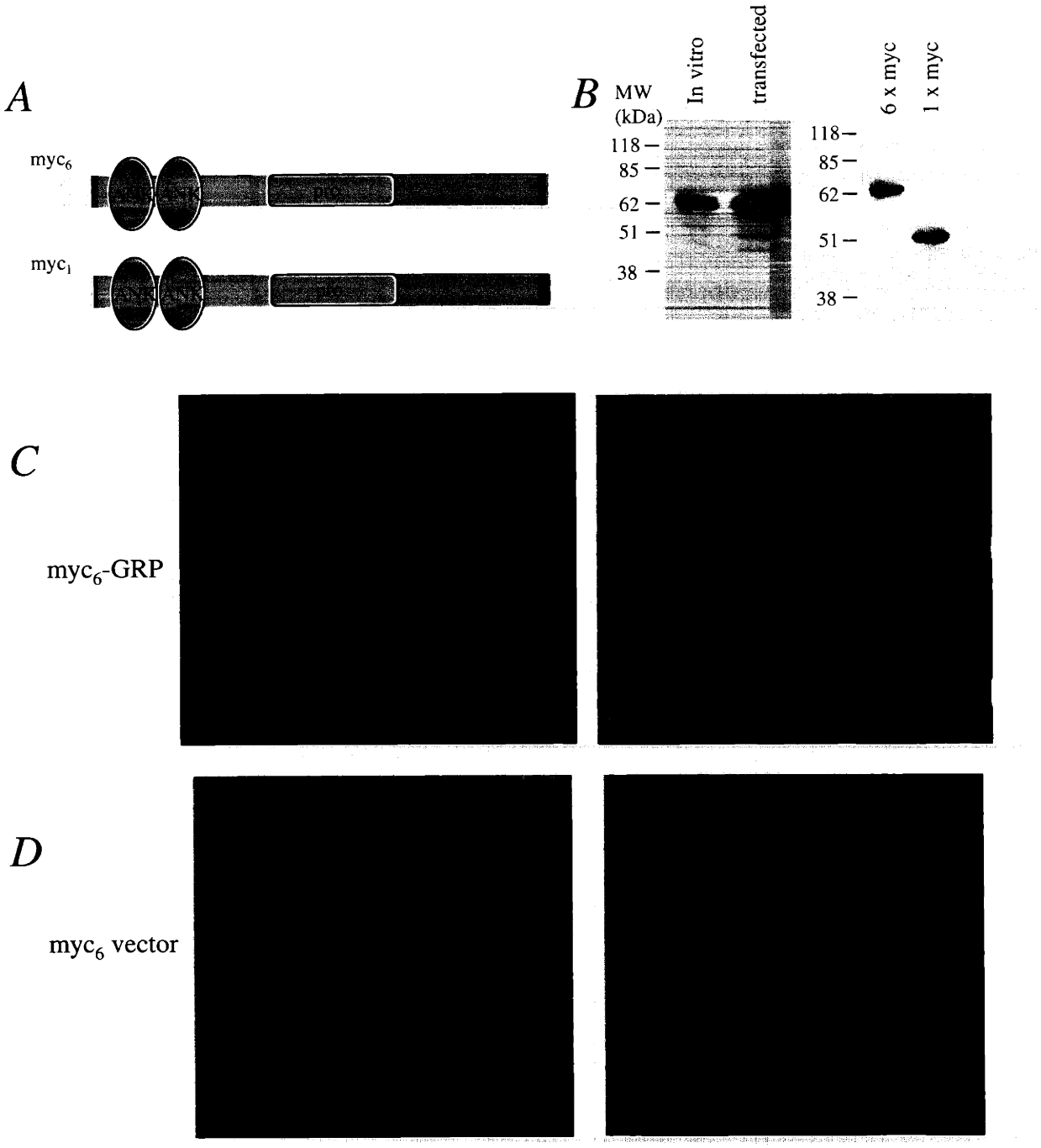


Figure 1. ***(A)*** Schematic of the myc_6-GRP expression constructs used in (B) and (C). ***(B)*** Left panel: immunoblot showing migration of in vitro translated (left) and transfected (right) myc_6-GRP. The right lane contains total cell lysate from myc_6-GRP transfected cos7 cells. Right panel: immunoblot of total cell lysates from myc_6-GRP (left) and myc_1-GRP transfected cos7 cells. ***(C), (D)*** Immunofluorescent labeling of rat1 cells transfected with (C) myc_6 -GRP and (D) empty myc_6 vector. Myc-tagged proteins in (B), (C) and (D) were detected with the anti-myc monoclonal antibody 9e10. In (C) and (D), antibody staining was visualized using cy3-linked anti-mouse IgG; nuclei are counterstained with DAPI (right panels).

Figure 2

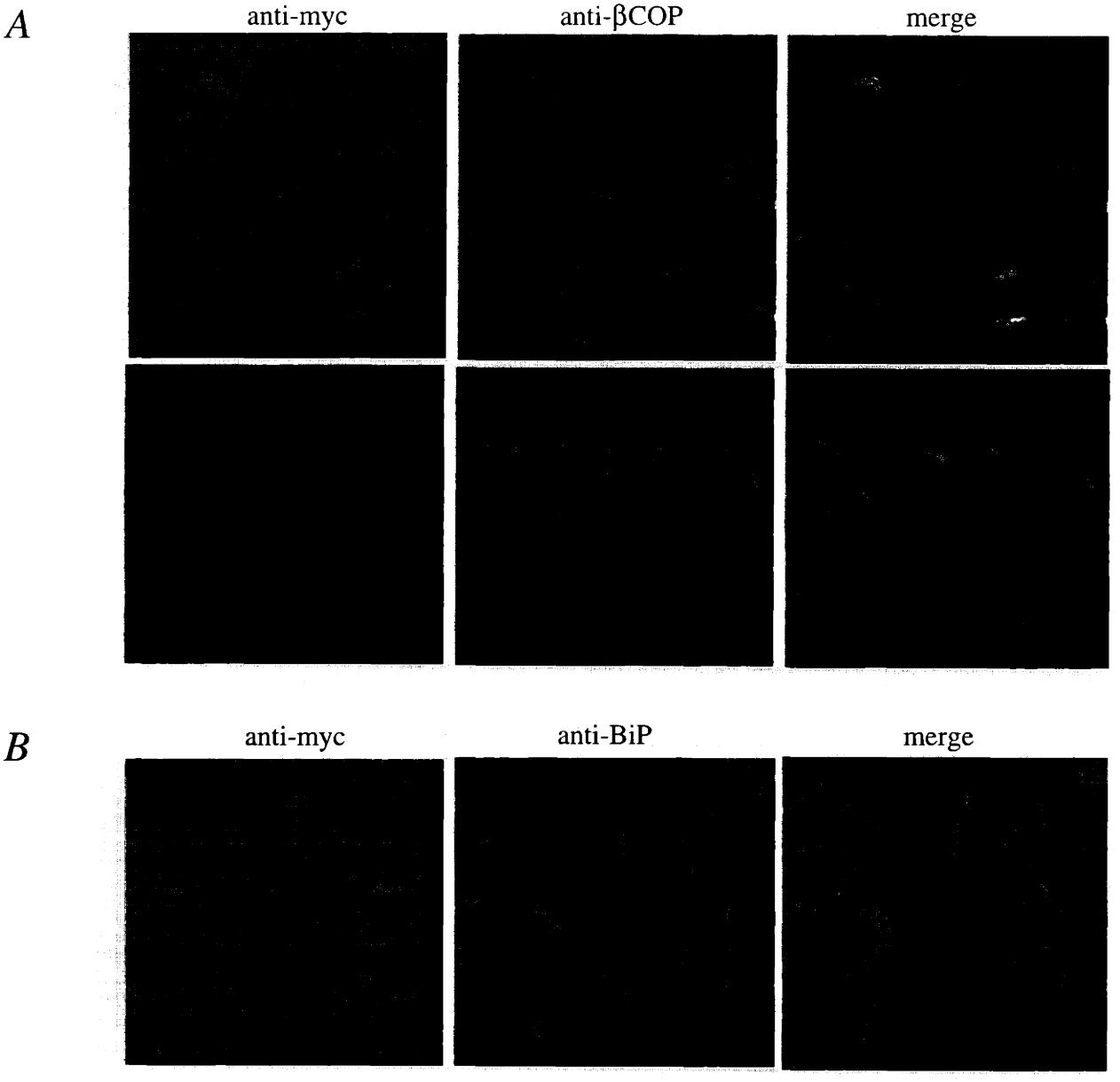


Figure 2. Immunofluorescent labeling in pools of stably transfected, myc-GRP expressing Rat-1 fibroblasts. ***(A).*** Myc-GRP is localized to the Golgi apparatus. The upper and lower panels show two different pools of myc-GRP transfected Rat-1 cells, double-labeled with the anti-myc monoclonal antibody 9e10 (left) and anti-βCOP polyclonal antibody (middle), followed by Texas Red- and FITC-conjugated secondary antibodies. The digital images were merged (right) to indicated overlapping staining. ***(B).*** Myc-GRP transfected Rat-1 cells double-labeled with anti-myc polyclonal antibody and anti-BiP monoclonal antibody show no overlap between myc-GRP staining and the endoplasmic reticulum.

Figure 3

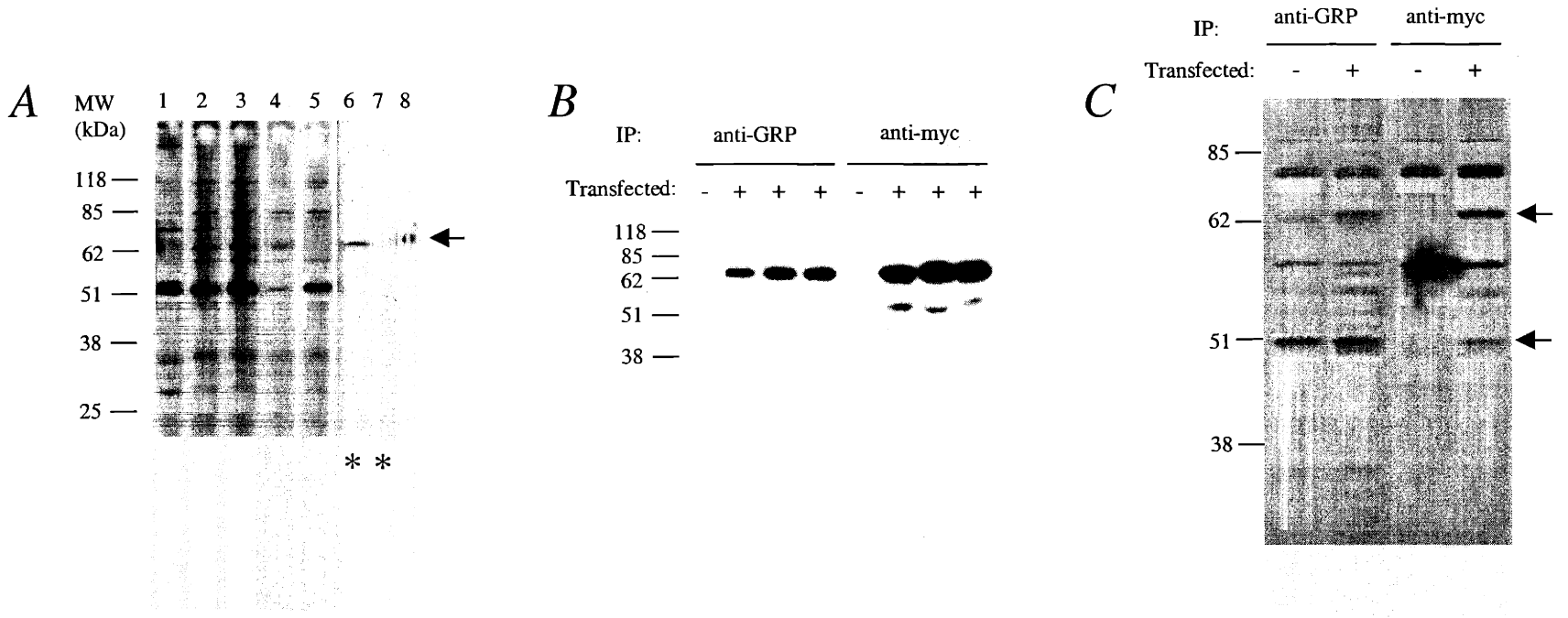


Figure 3. *(A)* Immunoblot of total cell lysate from myc_6-GRP transfected cos7 cells. Primary antibodies used: (lane1) preimmune rabbit serum; (2), GRP-immunized rabbit serum; (3), 30% and (4), 50% ammonium sulfate fractions of GRP-immunized rabbit serum; (5), flow-through from binding of antiserum to antigen affinity column; (6), affinity purified anti-GRP polylclonal antibodies; (7), no primary antibody; (8), anti-myc epitope polyclonal antibody. Primary antibodies were detected with horseradish peroxidase (HRP)-conjugated anti-rabbit IgG. Lanes 6 and 7 (*) shown from longer exposure. ***(B)*** Immunoblot containing anti-GRP and anti-myc immunoprecipitates from vector transfected (-) and myc_6-GRP transfected (+) EcR-293 cells, probed with anti-myc monoclonal antibody 9e10 followed by HRP-conjugated anti-mouse IgG. *(C)* Autoradiograph of SDS-PAGE fractionated anti-GRP and anti-myc immunoprecipitates, from ^{35}S-methionine labeled vector transfected (-) and myc_6-GRP transfected (+) EcR-293 cells. Arrows indicate the myc_6-GRP (62 kDa) and putative endogenous GRP (50 kDa) proteins.

Figure 4

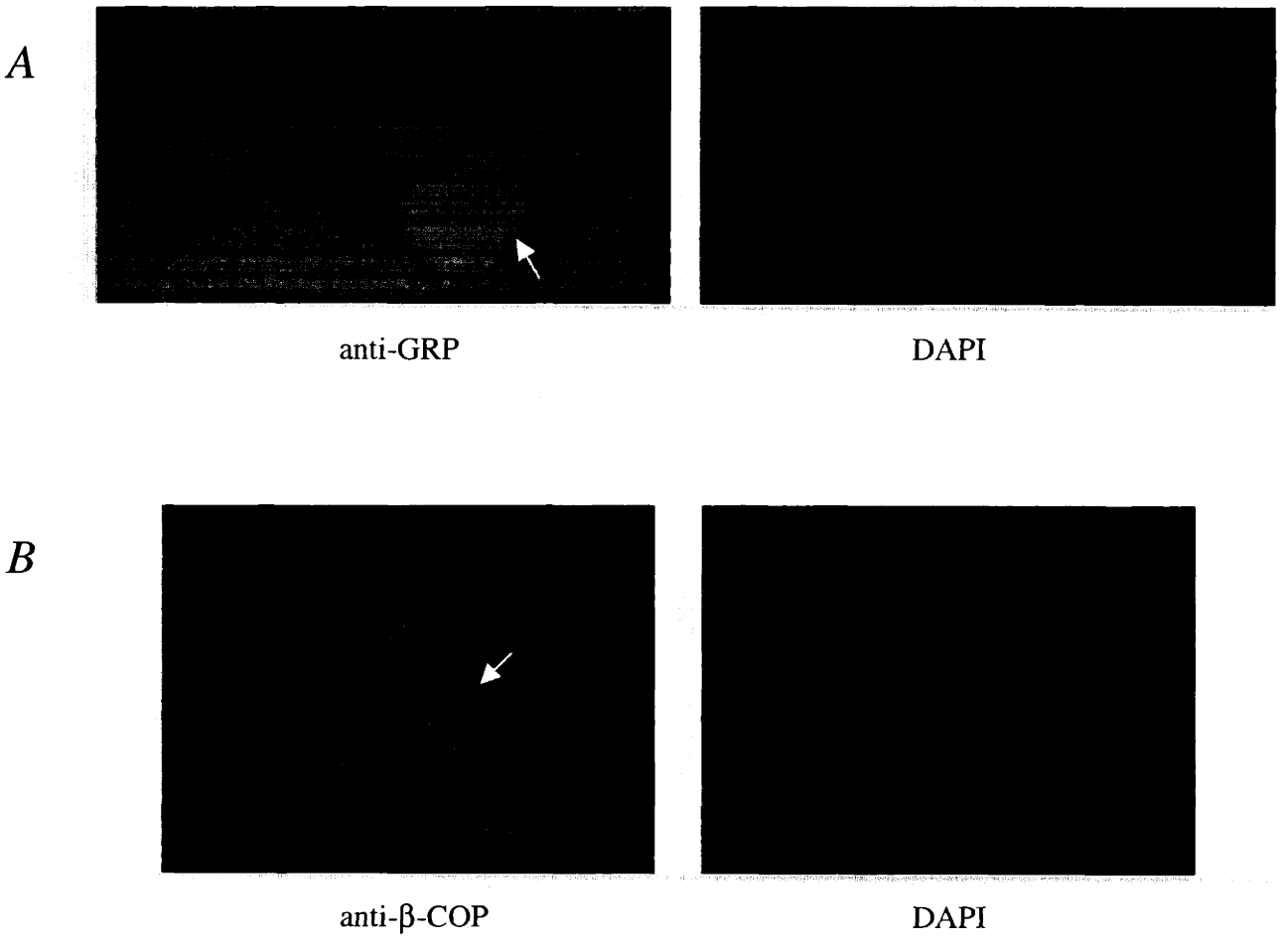


Figure 4. Immunofluorescent labeling of GRP ***(A)*** and b-COP ***(B)*** in untransfected NIH/3T3 cells. Cells were labeled with affinity purified rabbit anti-GRP or rabbit anti-β-COP (Affinity Bioreagents, Inc.), followed by FITC-conjugated anti-rabbit IgG (left panels). Arrows indicate mitotic cells, as shown by DAPI counterstaining of nuclei (right panels).

Figure 5

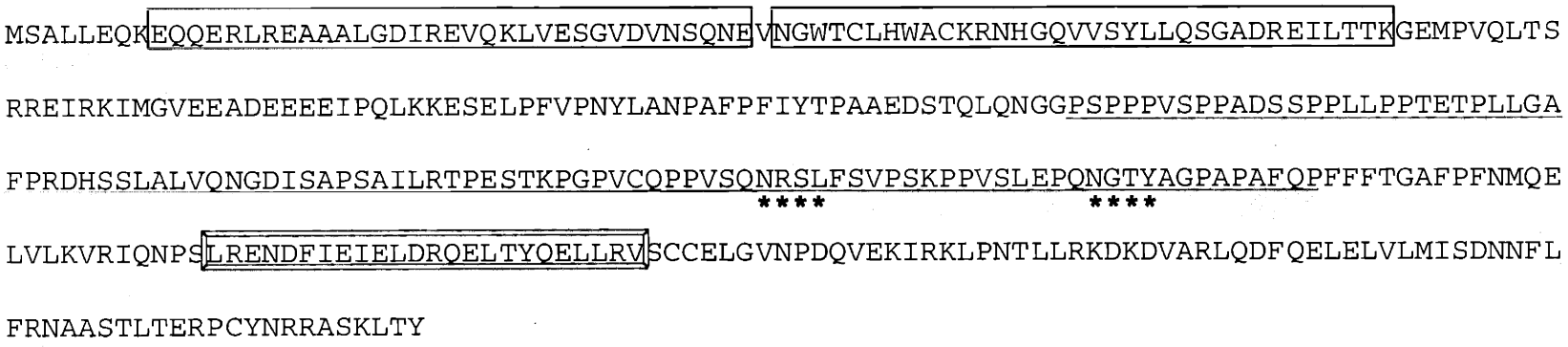


Figure 5. Amino acid sequence of murine GRP indicating the following motifs: the two NH_2-terminal ankyrin repeats (boxed), the proline-rich domain (underlined), the putative coiled coil-forming residues (double-boxed), and the consensus sites for N-linked glycosylation (*). Coiled coil prediction was obtained using the COILS program(see text) with a 21 residue scanning window.

Figure 6

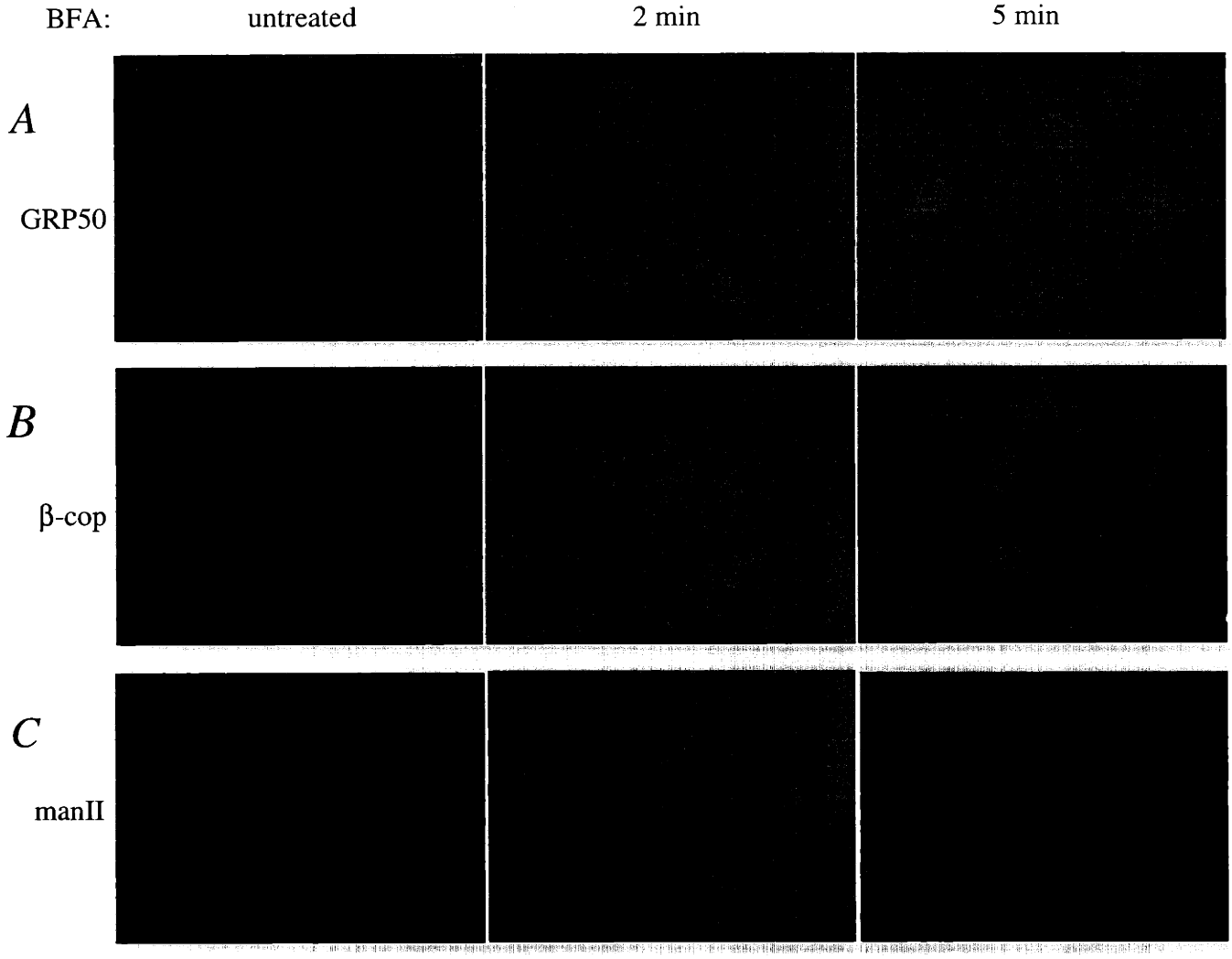


Figure 6. Immunolocalization of GRP50 (A), β-COP (B), and mannosidase II (C) in untreated (left panels) and Brefeldin A treated (middle, right panels) rodent fibroblast cells. Cells were treated with 5 μg/ml BFA for the indicated times and fixed immediaely. Primary antibodies (affinity purified anti-GRP polyclonal, anti-β-COP polyclonal, and anti manII monoclonal antibodies, see Materials and Methods) were detected with cy3-linked anti-rabbit (A and B) or anti-mouse (C) IgG.

Figure 7

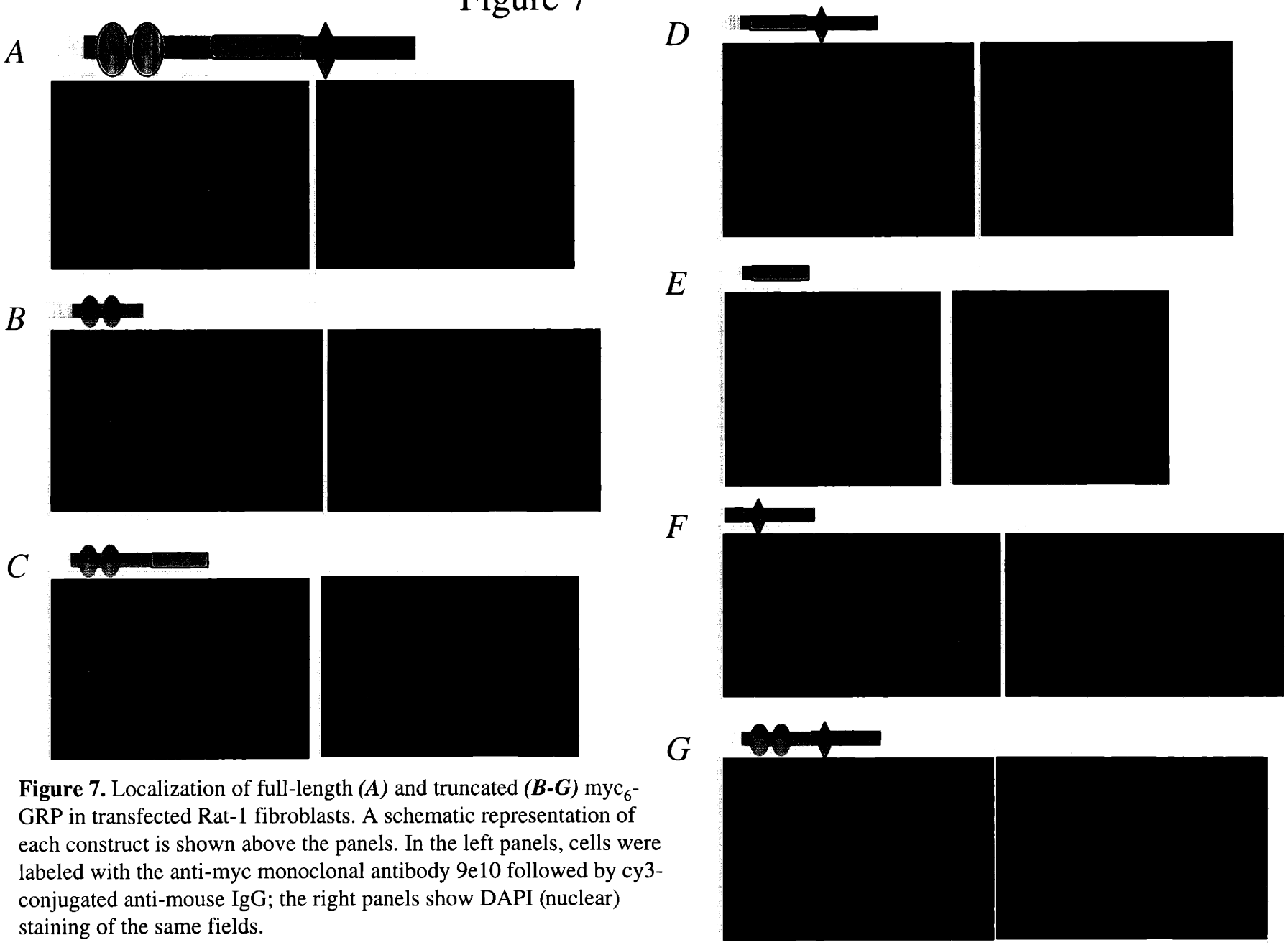


Figure 7. Localization of full-length ***(A)*** and truncated ***(B-G)*** myc_6-GRP in transfected Rat-1 fibroblasts. A schematic representation of each construct is shown above the panels. In the left panels, cells were labeled with the anti-myc monoclonal antibody 9e10 followed by cy3-conjugated anti-mouse IgG; the right panels show DAPI (nuclear) staining of the same fields.

Figure 8

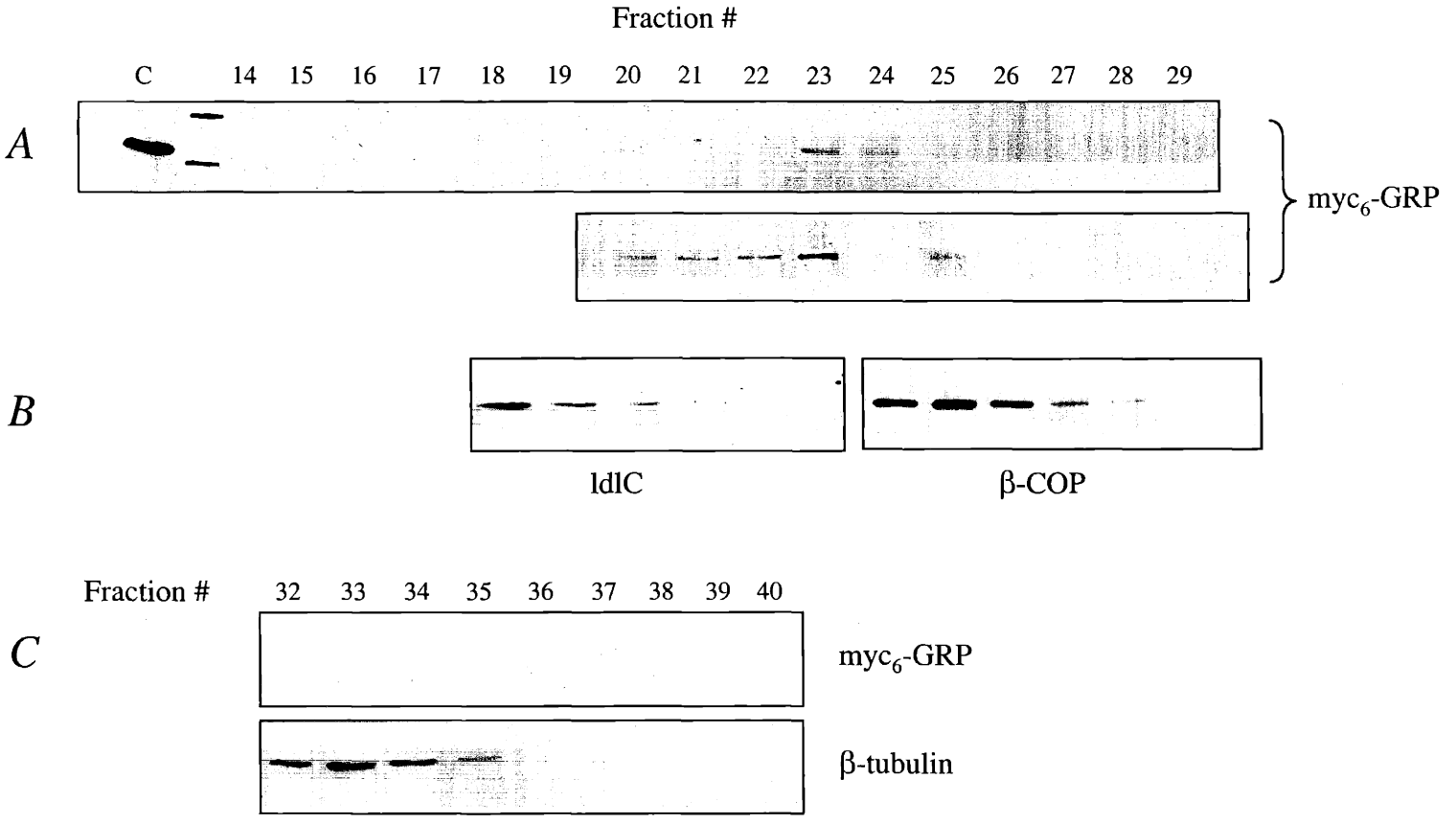


Figure 9. Immunoblot analysis of gel chromatographic fractions of cytosol from myc_6-GRP stably transfected Rat-1 cells. ***(A).*** Detection of myc_6-GRP, with the anti-myc monoclonal antibody 9e10, in unfractionated cytosol (C) and in fractions 20-25. The panels show two separate immunoblotting experiments using the samefractions. **(B).** Detection of ldlC (complex size ~950 kDa, left panel) in fractions 18-21, and of β-cop (complex size ~700 kDa, right panel) in fractions 24-28. ***(C).*** Top panel: myc_6-GRP is not detected in fractions 32-40, where smaller molecular weight species are expected to elute. Bottom panel: Detection of β-tubulin (apparent monomer mass: 55 kDa) on the above blot, in fractions 32-35.

Figure 9

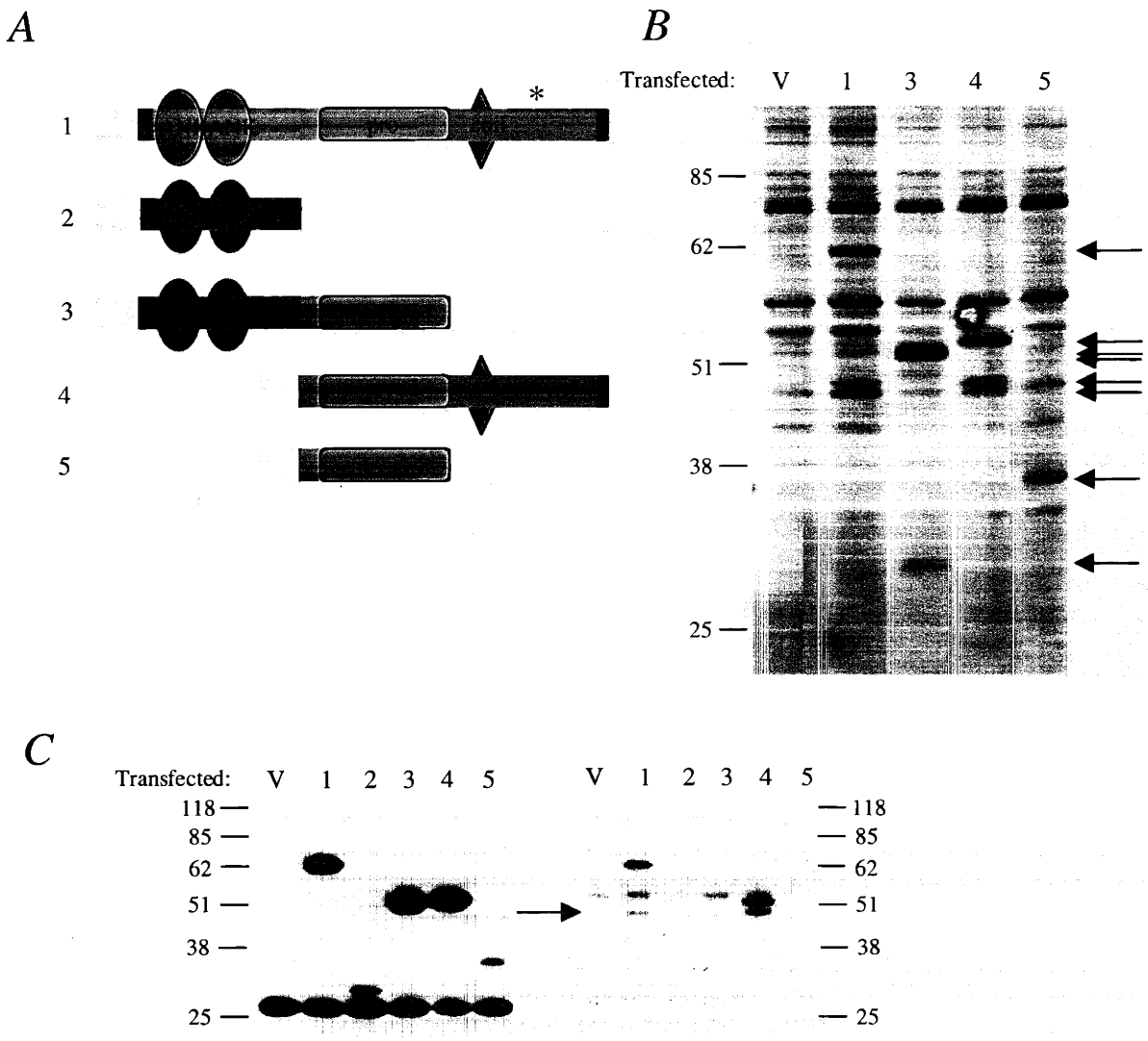


Figure 9. ***(A)*** myc_6-GRP constructs used in (B) and (C). The domain used to generate anti-GRP antibodies used in (C) is denoted with an asterisk. ***(B)*** Anti-myc immunoprecipitates from ^{35}S-methionine labeled EcR-293 cells, transfected with empty vector (V) or myc_6-GRP constructs shown in (A). Black arrows indicate the myc_6-GRP proteins; red arrows highlight putative co-immunoprecipitated proteins. ***(C)*** Western blots of anti-myc immunoprecipitates from EcR-293 cells, transfected with myc_6-GRP constructs shown in (A). Left panel, probed with 9e10; right panel probed with anti-GRP purified polyclonal antibodies. The red arrow highlights the band presumed to be endogenous GRP. Immunoprecipitations were performed using a 9e10-agarose conjugate (see Materials and Methods).

REFERENCES

Balch, W.E., W.G. Dunphy, W.A. Braell, and J.E. Rothman. 1984. Reconstitution of the transport of protein between successive compartments of the Golgi measured by the coupled incorporation of N- acetylglucosamine. *Cell.* 39:405-416.

Bannykh, S.I., and W.E. Balch. 1998. Selective transport of cargo between the endoplasmic reticulum and Golgi compartments. *Histochem Cell Biol.* 109:463-475.

Barlowe, C., L. Orci, T. Yeung, M. Hosobuchi, S. Hamamoto, N. Salama, M.F. Rexach, M. Ravazzola, M. Amherdt, and R. Schekman. 1994. COPII: a membrane coat formed by Sec proteins that drive vesicle budding from the endoplasmic reticulum. *Cell.* 77:895-907.

Barlowe, C., and R. Schekman. 1993. SEC12 encodes a guanine-nucleotide-exchange factor essential for transport vesicle budding from the ER. *Nature.* 365:347-349.

Barr, F.A., N. Nakamura, and G. Warren. 1998. Mapping the interaction between GRASP65 and GM130, components of a protein complex involved in the stacking of Golgi cisternae. *Embo J.* 17:3258-3268.

Burkhard, P., J. Stetefeld, and S.V. Strelkov. 2001. Coiled coils: a highly versatile protein folding motif. *Trends Cell Biol.* 11:82-88.

Chardin, P., and F. McCormick. 1999. Brefeldin A: the advantage of being uncompetitive. *Cell.* 97:153-155.

Chatterton, J.E., D. Hirsch, J.J. Schwartz, P.E. Bickel, R.D. Rosenberg, H.F. Lodish, and M. Krieger. 1999. Expression cloning of LDLB, a gene essential for normal Golgi function and assembly of the ldlCp complex. *Proc Natl Acad Sci U S A.* 96:915-920.

Clague, M.J. 1999. Membrane transport: Take your fusion partners. *Curr Biol.* 9:R258-260.

Cosson, P., and F. Letourneur. 1997. Coatomer (COPI)-coated vesicles: role in intracellular transport and protein sorting. *Curr Opin Cell Biol.* 9:484-487.

Dascher, C., and W.E. Balch. 1994. Dominant inhibitory mutants of ARF1 block endoplasmic reticulum to Golgi transport and trigger disassembly of the Golgi apparatus. *J Biol Chem.* 269:1437-1448.

Dinter, A., and E.G. Berger. 1998. Golgi-disturbing agents. *Histochem Cell Biol.* 109:571-590.

Donaldson, J.G., D. Finazzi, and R.D. Klausner. 1992. Brefeldin A inhibits Golgi membrane-catalysed exchange of guanine nucleotide onto ARF protein. *Nature*. 360:350-352.

Donaldson, J.G., and C.L. Jackson. 2000. Regulators and effectors of the ARF GTPases. *Curr Opin Cell Biol*. 12:475-482.

Donaldson, J.G., J. Lippincott-Schwartz, G.S. Bloom, T.E. Kreis, and R.D. Klausner. 1990. Dissociation of a 110-kD peripheral membrane protein from the Golgi apparatus is an early event in brefeldin A action. *J Cell Biol*. 111:2295-2306.

Duden, R., M. Hosobuchi, S. Hamamoto, M. Winey, B. Byers, and R. Schekman. 1994. Yeast beta- and beta'-coat proteins (COP). Two coatomer subunits essential for endoplasmic reticulum-to-Golgi protein traffic. *J Biol Chem*. 269:24486-24495.

Farquhar, M.G., and G.E. Palade. 1998. The Golgi apparatus: 100 years of progress and controversy. *Trends Cell Biol*. 8:2-10.

Fujiwara, T., K. Oda, S. Yokota, A. Takatsuki, and Y. Ikehara. 1988. Brefeldin A causes disassembly of the Golgi complex and accumulation of secretory proteins in the endoplasmic reticulum. *J Biol Chem*. 263:18545-18552.

Gleeson, P.A. 1998. Targeting of proteins to the Golgi apparatus. *Histochem Cell Biol*. 109:517-532.

Glick, B.S. 2000. Organization of the Golgi apparatus. *Curr Opin Cell Biol*. 12:450-456.
Guo, W., M. Sacher, J. Barrowman, S. Ferro-Novick, and P. Novick. 2000. Protein complexes in transport vesicle targeting. *Trends Cell Biol*. 10:251-255.

Harlow, E., and D. Lane. 1988. Antibodies : a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY. xiii, 726 pp.

Kjer-Nielsen, L., R.D. Teasdale, C. van Vliet, and P.A. Gleeson. 1999. A novel Golgi-localisation domain shared by a class of coiled-coil peripheral membrane proteins. *Curr Biol*. 9:385-388.

Kreis, T.E., M. Lowe, and R. Pepperkok. 1995. COPs regulating membrane traffic. *Annu Rev Cell Dev Biol*. 11:677-706.

Lippincott-Schwartz, J., L.C. Yuan, J.S. Bonifacino, and R.D. Klausner. 1989. Rapid redistribution of Golgi proteins into the ER in cells treated with brefeldin A: evidence for membrane cycling from Golgi to ER. *Cell*. 56:801-813.

Lowe, M., C. Rabouille, N. Nakamura, R. Watson, M. Jackman, E. Jamsa, D. Rahman, D.J. Pappin, and G. Warren. 1998. Cdc2 kinase directly phosphorylates the cis-Golgi

matrix protein GM130 and is required for Golgi fragmentation in mitosis. *Cell.* 94:783-793.

Lucocq, J.M., E.G. Berger, and G. Warren. 1989. Mitotic Golgi fragments in HeLa cells and their role in the reassembly pathway. *J Cell Biol.* 109:463-474.

Lucocq, J.M., and G. Warren. 1987. Fragmentation and partitioning of the Golgi apparatus during mitosis in HeLa cells. *Embo J.* 6:3239-3246.

Lupas, A. 1996. Prediction and analysis of coiled-coil structures. *Methods Enzymol.* 266:513-525.

Lupas, A., M. Van Dyke, and J. Stock. 1991. Predicting coiled coils from protein sequences. *Science.* 252:1162-1164.

Melancon, P., B.S. Glick, V. Malhotra, P.J. Weidman, T. Serafini, M.L. Gleason, L. Orci, and J.E. Rothman. 1987. Involvement of GTP-binding "G" proteins in transport through the Golgi stack. *Cell.* 51:1053-1062.

Mochida, S. 2000. Protein-protein interactions in neurotransmitter release. *Neurosci Res.* 36:175-182.

Mostov, K.E., M. Verges, and Y. Altschuler. 2000. Membrane traffic in polarized epithelial cells. *Curr Opin Cell Biol.* 12:483-490.

Munro, S. 1998. Localization of proteins to the Golgi apparatus. *Trends Cell Biol.* 8:11-15.

Nakamura, N., M. Lowe, T.P. Levine, C. Rabouille, and G. Warren. 1997. The vesicle docking protein p115 binds GM130, a cis-Golgi matrix protein, in a mitotically regulated manner. *Cell.* 89:445-455.

No, D., T.P. Yao, and R.M. Evans. 1996. Ecdysone-inducible gene expression in mammalian cells and transgenic mice. *Proc Natl Acad Sci U S A.* 93:3346-3351.

Orci, L., M. Stamnes, M. Ravazzola, M. Amherdt, A. Perrelet, T.H. Sollner, and J.E. Rothman. 1997. Bidirectional transport by distinct populations of COPI-coated vesicles. *Cell.* 90:335-349.

Pearse, B.M. 1988. Receptors compete for adaptors found in plasma membrane coated pits. *Embo J.* 7:3331-3336.

Peyroche, A., B. Antonny, S. Robineau, J. Acker, J. Cherfils, and C.L. Jackson. 1999. Brefeldin A acts to stabilize an abortive ARF-GDP-Sec7 domain protein complex: involvement of specific residues of the Sec7 domain. *Mol Cell.* 3:275-285.

Podos, S.D., P. Reddy, J. Ashkenas, and M. Krieger. 1994. LDLC encodes a brefeldin A-sensitive, peripheral Golgi protein required for normal Golgi function. *J Cell Biol.* 127:679-691.

Puthenveedu, M.A., and A.D. Linstedt. 2001. Evidence that Golgi structure depends on a p115 activity that is independent of the vesicle tether components giantin and GM130. *J Cell Biol.* 155:227-238.

Reinhardt, A., and T. Hubbard. 1998. Using neural networks for prediction of the subcellular location of proteins. *Nucleic Acids Res.* 26:2230-2236.

Robinson, L.J., F. Aniento, and J. Gruenberg. 1997. NSF is required for transport from early to late endosomes. *J Cell Sci.* 110:2079-2087.

Robinson, M.S. 1994. The role of clathrin, adaptors and dynamin in endocytosis. *Curr Opin Cell Biol.* 6:538-544.

Robinson, M.S., and T.E. Kreis. 1992. Recruitment of coat proteins onto Golgi membranes in intact and permeabilized cells: effects of brefeldin A and G protein activators. *Cell.* 69:129-138.

Roth, M.G. 1999. Snapshots of ARF1: implications for mechanisms of activation and inactivation. *Cell.* 97:149-152.

Schekman, R., and L. Orci. 1996. Coat proteins and vesicle budding. *Science.* 271:1526-1533.

Sciaky, N., J. Presley, C. Smith, K.J. Zaal, N. Cole, J.E. Moreira, M. Terasaki, E. Siggia, and J. Lippincott-Schwartz. 1997. Golgi tubule traffic and the effects of brefeldin A visualized in living cells. *J Cell Biol.* 139:1137-1155.

Shaper, J.H., G.F. Hollis, and N.L. Shaper. 1988. Evidence for two forms of murine beta-1,4-galactosyltransferase based on cloning studies. *Biochimie.* 70:1683-1688.

Shorter, J., and G. Warren. 1999. A role for the vesicle tethering protein, p115, in the post-mitotic stacking of reassembling Golgi cisternae in a cell-free system. *J Cell Biol.* 146:57-70.

Simpson, F., N.A. Bright, M.A. West, L.S. Newman, R.B. Darnell, and M.S. Robinson. 1996. A novel adaptor-related protein complex. *J Cell Biol.* 133:749-760.

Simpson, F., A.A. Peden, L. Christopoulou, and M.S. Robinson. 1997. Characterization of the adaptor-related protein complex, AP-3. *J Cell Biol.* 137:835-845.

Sollner, T., M.K. Bennett, S.W. Whiteheart, R.H. Scheller, and J.E. Rothman. 1993. A protein assembly-disassembly pathway in vitro that may correspond to sequential steps of synaptic vesicle docking, activation, and fusion. *Cell.* 75:409-418.

Sonnichsen, B., M. Lowe, T. Levine, E. Jamsa, B. Dirac-Svejstrup, and G. Warren. 1998. A role for giantin in docking COPI vesicles to Golgi membranes. *J Cell Biol.* 140:1013-1021.

Spang, A., K. Matsuoka, S. Hamamoto, R. Schekman, and L. Orci. 1998. Coatomer, Arf1p, and nucleotide are required to bud coat protein complex I-coated vesicles from large synthetic liposomes. *Proc Natl Acad Sci U S A.* 95:11199-11204.

Springer, S., A. Spang, and R. Schekman. 1999. A primer on vesicle budding. *Cell.* 97:145-148.

Taatjes, D.J., J. Roth, N.L. Shaper, and J.H. Shaper. 1992. Immunocytochemical localization of beta 1,4 galactosyltransferase in epithelial cells from bovine tissues using monoclonal antibodies. *Glycobiology.* 2:579-589.

Traub, L.M., and S. Kornfeld. 1997. The trans-Golgi network: a late secretory sorting station. *Curr Opin Cell Biol.* 9:527-533.

Warren, G. 1993. Membrane partitioning during cell division. *Annu Rev Biochem.* 62:323-348.

Waters, M.G., T. Serafini, and J.E. Rothman. 1991. 'Coatomer': a cytosolic protein complex containing subunits of non- clathrin-coated Golgi transport vesicles. *Nature.* 349:248-251.

CHAPTER 4

Conclusions

CONCLUSIONS

This thesis describes the identification of a highly conserved mammalian gene, *GRP50*, which encodes a novel Golgi-associated ankyrin repeat protein. In an effort to glean clues to the molecular and biological functions of the GRP50 protein, we have utilized a variety of approaches. This section summarizes the results of these analyses, suggests avenues for future investigation, and offers an anecdotal perspective on the contemporary approaches to the problem of assigning function to novel sequences.

Computerized sequence analysis and structure prediction tools have permitted the identification of GRP50 as a novel member of the ankryin repeat superfamily of proteins (Sedgwick and Smerdon, 1999). In addition, we have identified a putative orthologue of GRP50 in a hypothetical protein sequence (R31.2) in the nematode, *C. elegans*. The N-terminal domains of the mammalian and nematode proteins share considerable sequence similarity, both within and adjacent to their ankyrin repeat domains. The central and carboxyl-terminal regions of these proteins are only slightly similar, however, and the *C. elegans* protein contains a carboxyl-terminal transmembrane domain.

In view of our results demonstrating localization, via sequences in the carboxyl terminus, of GRP50 to the cytoplasmic face of Golgi and vesicle membranes, the presence of a transmembrane domain in the nematode protein R31.2 might be taken as further confirmation of orthology. It seems reasonable to conclude that, in the course of evolution, the functional domains present in the nematode protein became physically separated, such that in the mammalian protein targeting to the membrane is achieved by other means. Both proteins would be anchored to the membrane by their carboxyl

termini, leaving their structurally related N-terminal domains available for binding via the ankyrin repeat and proline-rich domains. A comprehensive analysis of the genes that mediate membrane transport in C. elegans is currently underway (Koushika and Nonet, 2000), and may yet yield clues to the function of this protein in the nematode.

It is perhaps surprising that we were unable to identify GRP50-homologous sequences in the genome of *Saccharomyces cerevisiae*, which has long been the model system of choice for the study of Golgi function and membrane trafficking. This could be explained by the fact that the GRP50 sequence is comprised of structural domains which, aside from the ankyrin repeats, do not conform to a consensus that is easily detected in pairwise sequence analysis (Califano, 2001). The ever-expanding structural databases may eventually be useful for predicting these types of homologies (Teichmann et al., 2001). An alternative explanation could be that GRP50 functions in a pathway, or part of a pathway, for which there is no isomorphic counterpart in yeast. With regard to the biology of the Golgi apparatus, the primary difference between fungi and animal cells is structural; in fungi and plants, the multiple cisternae of the Golgi are not stacked together but instead are maintained as dispersed compartments (Warren, 1993). As a consequence, yeast do not exhibit the mitotic fragmentation of the Golgi apparatus that is observed in animal cells (Lowe et al., 1998; Lucocq and Warren, 1987). Presumably, animal cells contain proteins or protein complexes that contribute to the stability of the stacked cisternae which are not required in yeast.

When sequence analysis does not reveal orthologues of known function, expression profile analysis can be a useful tool in the characterization of a novel gene, since genes that function in a common biological process frequently display coordinate

regulation. The utility of this approach is underscored by the growing popularity of microarray expression data analyses in contemporary functional genomics (Altman and Raychaudhuri, 2001). Our analyses of the expression profile of GRP50 reveal several intriguing features. First, although GRP50 is ubiquitously expressed, its expression levels appear to be regulated in a cell- or tissue-type specific manner, and may be modulated in development. In addition, GRP50 is induced as part of the serum response in primary fibroblasts.

A recent study examined the expression profiles of close to 10,000 genes in serum-starved and serum-stimulated human primary fibroblasts (Iyer et al., 1999), and these data are available in a searchable format (genome-www.stanford.edu/serum). Comparison of the profile of GRP50 serum-induction to the dataset generated in this experiment revealed that fewer than fifty genes exhibited a serum-induction profile like that of GRP50. Among those genes that showed a similar profile of mRNA induction in response to serum were several that are implicated in tissue remodeling and wound healing (Iyer et al., 1999), but no genes encoding known Golgi-associated proteins. Moreover, we searched the 10K dataset for the names of known Golgi- and vesicle-associated genes and found at least twenty, none of which were serum-induced. In fact, all of these genes displayed expression patterns consistent with their role as "housekeeping" genes; that is, their mRNA levels were unchanged in the time course of this experiment. These results imply that the regulation of GRP50 expression is distinct from that of many other genes encoding Golgi- and vesicle-associated proteins, which may indicate a unique role for the GRP50 protein in specific physiological processes.

A number of studies have described global gene expression patterns in a variety of organisms using cDNA microarray-based methods (Cho et al., 1998; Spellman et al., 1998; White et al., 1999). In one recent example, the expression patterns of close to 19,000 nominally nonredundant genes were investigated in both embryonic and adult mouse tissues (Miki et al., 2001). However, we have not found GRP50 among the published mammalian cDNA datasets used in these studies, including the RIKEN 19K dataset derived from full-length selected mouse cDNAs (Miki et al., 2001) and the 10K dataset of human fibroblast-derived cDNAs described above (Iyer et al., 1999). This result was somewhat surprising, given the widespread expression of GRP50 indicated by our experiments, and may indicate a tendency of these analyses to exclude certain types of sequences. In any event, the fact that GRP50 is not included in these datasets limits our conclusions regarding its expression profile to the results of individual northern analyses.

Analysis of the structural features of the GRP50 polypeptide sequence indicates that the protein is composed of multiple modular domains, including an ankyrin repeat region, a polyproline-rich region, and a leucine-rich region that may form α-helical coiled coils (Burkhard et al., 2001). Because each of these types of domains has been shown in other proteins to function as a protein-protein interaction module, we conclude that GRP50 almost certainly participates in intermolecular interactions. The function of this protein is therefore probably mediated through the intermolecular interactions in which it participates. This assumption, if correct, implies that the biological role of GRP50 is determined by the function of its binding partner or partners, and by whether binding enhances or inhibits that function. Our data indicate that GRP50 is localized to the cytoplasmic face of Golgi and vesicle membranes, most likely though interaction with

other integral or peripheral membrane proteins. In addition, our results suggest that GRP50 is capable of associating with additional soluble cytoplasmic proteins to form a large macromolecular complex. These observations imply that GRP50 functions, in a complex with other proteins, to effect specific biological events at the cytoplasmic surface of the membrane.

A cursory examination of the structure of the Golgi apparatus and of the mechanisms that direct membrane trafficking in eukaryotic cells reveals that these processes are mediated by a host of proteins and multiprotein complexes (Schekman and Orci, 1996; Traub and Kornfeld, 1997). In addition, the key features of these processes at distinct subcellular sites are functionally similar but involve heterotypic protein complexes. It is therefore impossible to speculate which events might be mediated by GRP50 without knowledge of the proteins with which it interacts. The list of known proteins Golgi- and vesicle-associated proteins is long and getting longer, and involves both membrane- and cell-type specific components. However, our results offer a few hints as to which processes are more likely to involve GRP50-interacting proteins. For instance, our data show that GRP50 association with the Golgi apparatus is sensitive to the effects of Brefeldin A (BFA), and that the kinetics of this sensitivity resemble those of other proteins whose association with membranes is dependent on the actiivty of the ARF family of GTPases (Donaldson et al., 1992; Helms and Rothman, 1992). Because BFA is a specific inhibitor of a class of ARF-GEFs (Roth, 1999), these results imply that the association of GRP50 with the membrane requires the activity of a BFA-sensitive ARF or ARF-like protein.

A number of ARF proteins and of ARF effector proteins have been identified, and new members are added to this list frequently. To date, in mammals, there are six known ARF proteins, at least eight ARF-GEFs, and at least seven ARF-GAPs (Donaldson and Jackson, 2000; Roth, 1999). The ARFs are divided into three classes, one of which (the class I ARFs: ARF1, ARF2, and ARF3) is required for the recruitment of coat proteins to budding vesicles at the Golgi and the TGN. The class II ARFs are proteins whose specific functions are not known, and the class III ARF, ARF6, regulates membrane traffic and cytoskeletal organization at the plasma membrane (Roth, 1999). The ARF-GEFs contain numerous protein binding domains and are divided into two classes: one class is BFA sensitive, the other is BFA resistant. The ARF-GAPs similarly contain numerous protein interaction domains, including ankyrin repeats and SH3 domains (Donaldson and Jackson, 2000), and are acutely sensitive to BFA. The observed BFA-sensitivity of GRP50 could be explained by GRP50 association with one or more of these proteins, and future experiments aimed at determining the identity of GRP50-interacting proteins should address this possibility.

Finally, preliminary results from a two-hybrid screen for interacting proteins (see Appendix) suggest a possible role for GRP50 in cytoskeletal changes associated with cell growth and motility and with neurite outgrowth. The results of this screen are very preliminary and have yet to be confirmed biochemically, but they provide intriguing possibilities for speculation. A proposed role for GRP50 in cell motility or in cytoskeletal remodeling could fit nicely with the expression data discussed above, since GRP50 displays a serum-induction profile similar to that of many genes involved in these processes.

In this era of fully sequenced genomes, it often seems as though there are few remaining biologically important molecules that have not been described in at least one organism. The combination of the development of increasingly sophisticated bioinformatics algorithms and the rapidly growing number of proteins with solved structures has permitted the classification of apparently disparate sequences as structurally related. Nevertheless, for a newly identified protein without biologically well-characterized orthologues, the road from sequence to function is long. Ultimately, elucidating the biological role of proteins like GRP50 may lead to a greater understanding of the molecular processes that mediate vesicle formation and fusion, cytoskeletal reorganization, and signal transduction at the cytosolic face of membranes in eukaryotic organisms.

REFERENCES

Altman, R.B., and S. Raychaudhuri. 2001. Whole-genome expression analysis: challenges beyond clustering. *Curr Opin Struct Biol.* 11:340-347.

Burkhard, P., J. Stetefeld, and S.V. Strelkov. 2001. Coiled coils: a highly versatile protein folding motif. *Trends Cell Biol.* 11:82-88.

Califano, A. 2001. Advances in sequence analysis. *Curr Opin Struct Biol.* 11:330-333.

Cho, R.J., M.J. Campbell, E.A. Winzeler, L. Steinmetz, A. Conway, L. Wodicka, T.G. Wolfsberg, A.E. Gabrielian, D. Landsman, D.J. Lockhart, and R.W. Davis. 1998. A genome-wide transcriptional analysis of the mitotic cell cycle. *Mol Cell.* 2:65-73.

Donaldson, J.G., D. Finazzi, and R.D. Klausner. 1992. Brefeldin A inhibits Golgi membrane-catalysed exchange of guanine nucleotide onto ARF protein. *Nature.* 360:350-352.

Donaldson, J.G., and C.L. Jackson. 2000. Regulators and effectors of the ARF GTPases. *Curr Opin Cell Biol.* 12:475-482.

Helms, J.B., and J.E. Rothman. 1992. Inhibition by brefeldin A of a Golgi membrane enzyme that catalyses exchange of guanine nucleotide bound to ARF. *Nature.* 360:352-354.

Iyer, V.R., M.B. Eisen, D.T. Ross, G. Schuler, T. Moore, J.C. Lee, J.M. Trent, L.M. Staudt, J. Hudson, Jr., M.S. Boguski, D. Lashkari, D. Shalon, D. Botstein, and P.O. Brown. 1999. The transcriptional program in the response of human fibroblasts to serum. *Science.* 283:83-87.

Koushika, S.P., and M.L. Nonet. 2000. Sorting and transport in C. elegans: aA model system with a sequenced genome. *Curr Opin Cell Biol.* 12:517-523.

Lowe, M., N. Nakamura, and G. Warren. 1998. Golgi division and membrane traffic. *Trends Cell Biol.* 8:40-44.

Lucocq, J.M., and G. Warren. 1987. Fragmentation and partitioning of the Golgi apparatus during mitosis in HeLa cells. *Embo J.* 6:3239-3246.

Miki, R., K. Kadota, H. Bono, Y. Mizuno, Y. Tomaru, P. Carninci, M. Itoh, K. Shibata, J. Kawai, H. Konno, S. Watanabe, K. Sato, Y. Tokusumi, N. Kikuchi, Y. Ishii, Y. Hamaguchi, I. Nishizuka, H. Goto, H. Nitanda, S. Satomi, A. Yoshiki, M. Kusakabe, J.L. DeRisi, M.B. Eisen, V.R. Iyer, P.O. Brown, M. Muramatsu, H. Shimada, Y. Okazaki, and Y. Hayashizaki. 2001. Delineating developmental and metabolic pathways in vivo by

expression profiling using the RIKEN set of 18,816 full-length enriched mouse cDNA arrays. *Proc Natl Acad Sci U S A.* 98:2199-2204.

Roth, M.G. 1999. Snapshots of ARF1: implications for mechanisms of activation and inactivation. *Cell.* 97:149-152.

Schekman, R., and L. Orci. 1996. Coat proteins and vesicle budding. *Science.* 271:1526-1533.

Sedgwick, S.G., and S.J. Smerdon. 1999. The ankyrin repeat: a diversity of interactions on a common structural framework. *Trends Biochem Sci.* 24:311-316.

Spellman, P.T., G. Sherlock, M.Q. Zhang, V.R. Iyer, K. Anders, M.B. Eisen, P.O. Brown, D. Botstein, and B. Futcher. 1998. Comprehensive identification of cell cycle-regulated genes of the yeast Saccharomyces cerevisiae by microarray hybridization. *Mol Biol Cell.* 9:3273-3297.

Teichmann, S.A., A.G. Murzin, and C. Chothia. 2001. Determination of protein function, evolution and interactions by structural genomics. *Curr Opin Struct Biol.* 11:354-363.

Traub, L.M., and S. Kornfeld. 1997. The trans-Golgi network: a late secretory sorting station. *Curr Opin Cell Biol.* 9:527-533.

Warren, G. 1993. Membrane partitioning during cell division. *Annu Rev Biochem.* 62:323-348.

White, K.P., S.A. Rifkin, P. Hurban, and D.S. Hogness. 1999. Microarray analysis of Drosophila development during metamorphosis. *Science.* 286:2179-2184.

APPENDIX

Yeast Two-Hybrid Screen for GRP50-interacting Proteins

INTRODUCTION

Elucidation of the molecular function of a novel protein can be facilitated by the identification of putative interacting proteins. If the target proteins have defined biological roles, this approach can reveal whether the protein of interest functions in a known pathway. When sequence analysis or experimental evidence suggest that a novel protein participates in a multiprotein complex, several approaches can be employed to identify other components of the complex. Biochemical techniques such as protein crosslinking, co-immunoprecipitation, and chromatographic fractionation may permit the isolation of proteins that physically associate with the protein of interest; however, determining the identity of these proteins can be challenging (McNabb and Guarente, 1996). The yeast two-hybrid system provides an alternative strategy for the identification of interacting proteins, and allows direct isolation of cDNAs encoding candidate binding partners.

The yeast two-hybrid screen permits rapid identification of novel protein interactions using an *in vivo* assay for reporter gene activation (Bartel and Fields, 1997; Chien et al., 1991; Fields and Song, 1989). Transcription of the reporter gene is regulated by the GAL4 transcription factor, which is composed of two physically and functionally separable domains, a DNA-binding domain (BD) and a *trans*-activation domain (AD). Both domains are required for activation of the GAL4-responsive promoter. The two domains of the GAL4 transcription factor can be expressed in yeast as separate proteins; in this case, transcription of a GAL4-driven reporter gene will proceed only if the two domains are brought into close physical proximity at the GAL4 promoter. In a two-hybrid

screen, the protein of interest (bait) is expressed as a fusion with the GAL4 BD, and a library of cDNA sequences is expressed as fusions to the GAL4 AD. Interaction of any of the proteins encoded by the cDNA library with the bait protein brings the GAL4 BD and AD together and permits transcription of the reporter gene(s).

As discussed in Chapters 2 and 3, the GRP50 protein contains multiple putative protein-binding domains and can be isolated from mammalian cells as part of a large multiprotein complex. In an effort to identify putative GRP50-interacting proteins, we screened a human fetal brain cDNA library using the Matchmaker™ Pretransformed cDNA Library System (Clontech). This library was chosen because, as shown in Chapter 2, GRP50 mRNA is abundantly expressed in fetal tissues and in adult brain. The pretransformed library system permits screening of a cDNA library by a simple yeast mating procedure (Bendixen et al., 1994). As shown below, the GRP50 protein exhibits both toxic and transcriptional transactivation effects when expressed as a GAL4 BD-fusion protein in yeast. As a result of these intrinsic properties of the GRP50 protein, the two-hybrid screen yielded a large number of clones, many of which are likely false positives. This section describes the results of this screen, and summarizes the ongoing efforts to identify and verify GRP50 binding partners.

RESULTS

An overview of the Matchmaker™ two-hybrid system used in this screen is depicted in Figure 1. In this system, GAL4 BD-bait fusion proteins are expressed in the haploid yeast strain PJ69-2A (MATa), and the library of GAL4 AD fusion proteins is expressed in the haploid yeast strain Y187 (MATα). Mating of a PJ69-2A(bait) strain with the Y187(library) strain yields a diploid strain which can be assayed for activation of the GAL4-responsive reporter genes *HIS3*, *ADE2*, and *lacZ* (Figure 1A). *HIS3* and *ADE2* are nutritional reporters, which permit selection of cells that have gained the ability to grow on medium lacking histidine and adenine (PJ69-2A and Y187 are auxotrophic for these amino acids). The *LacZ* reporter permits colorimetric detection and quantitation of GAL4 promoter activation, and can be used to assay the relative strengths of bait-target protein interactions (Aho et al., 1997; Estojak et al., 1995). A key feature of this system is that the *HIS3* and *ADE2* reporter genes are driven by heterologous promoters, whose only common feature is the small (17 nucleotide) consensus sequence that is bound by the GAL4 transcription factor (Figure 1B). These two distinct promoters are designed to limit the class of false positives that result from nonspecific interaction of a library protein with sequences in the reporter gene promoter.

Although the two-hybrid system is a powerful tool for the detection and identification of novel protein-protein interactions, there are two principal limitations of this technique (Bartel et al., 1993; Golemis et al., 1996). First, expression of some bait proteins (as GAL4 fusions) can be toxic in yeast, resulting in reduced viability or slow growth of bait strains. A second and more common difficulty arises when expression of

the GAL4-bait fusion protein results in inappropriate activation of GAL4-responsive genes, in the absence of any GAL4 AD protein. The basis for the two-hybrid system is a requirement for interaction of the bait protein (fused to the GAL4 BD) with a library protein (fused to the GAL4 AD) for reporter gene activation. It is therefore critical that expression of the bait protein alone does *not* lead to reporter gene activation, or the screen cannot be performed.

For these reasons, GAL4-bait expressing strains must be tested prior to library screening to ensure that they exhibit the expected phenotypes. Phenotype verification is aided by comparison with the control strains provided; in this system, the bait control is strain PJ69-2A containing plasmid pVA3-1, which encodes a GAL4 BD/p53 fusion protein, and the library control is strain Y187 containing plasmid pTD1-1, which encodes GAL4 AD/SV40 large T-antigen fusion protein. The p53 and SV40 T-Ag proteins interact in a two-hybrid system and bring the GAL4 BD and AD together to activate the reporter gene promoter (Iwabuchi et al., 1993; Li and Fields, 1993), such that mating of PJ69-2A(pVA3-1) with Y187(pTD1-1) yields a diploid strain which is His+, Ade+, and LacZ+. Importantly, the bait strain PJ69-2A(pVA3-1) can be used as a negative control when assaying experimental bait strains for toxicity or inappropriate transactivation of reporter genes.

A series of two-hybrid bait constructs, in which portions of mouse and human GRP50 were fused downstream of the GAL4 BD, were generated in the yeast expression vector pAS2-1 (Figure 2). These constructs were introduced into the haploid yeast strain PJ69-2A and the resultant bait strains were tested for toxicity and for reporter gene transactivation. Figure 2 summarizes the phenotypes of the GAL4 BD-GRP50 fusion

proteins. Two constructs, mGRP(full) and mGRP(C), exerted toxic effects on PJ69-2A; the PJ69-2A(mGRP(full)) strain grew poorly in liquid culture and on agar plates, and the PJ69-2A(mGRP(C)) transformants were not viable. In addition, the PJ69-2A(mGRP(full)) strain exhibited 10-fold reduction in mating efficiency (likely due to its slow growth rate) as compared to control strains (not shown).

All of the GAL4 BD-GRP50 bait strains exhibited transactivation of the *ADE2* reporter, as measured by robust growth on medium lacking adenine (Figure 2). However, the bait strains expressing hGRP(1-330) and mGRP(-pro) exhibited the expected His- and LacZ- phenotypes, indicating that these fusion proteins could be used in a two-hybrid screen using the *HIS3* nutritional reporter alone. Using only one of the two nutritional reporters was expected to reduce the stringency of the two-hybrid screen, since this approach fails to exploit the additional selectivity conferred by the distinct *ADE2* and *HIS3* promoters (see Figure 1B). Nonetheless, the GAL4 BD-hGRP(1-330) protein was the longest GRP50 construct that did not convert the yeast to His+, and did not exert negative effects on yeast growth. This protein was used as the bait in our two-hybrid screen.

The GAL4 BD-hGRP(1-330) protein was used to screen the Matchmaker™ pretransformed human fetal brain library for two-hybrid interactions, and the results of this screen are summarized in Figure 3. Diploid His+ colonies were classified as strong or weak positives according to the number of days required for growth (on medium lacking histidine) to at least 2 mm in diameter. 47 strong positives were obtained after 8 days, and 144 weak positives were obtained after 15 days of growth on selective medium. All positive clones were then analyzed for expression of the *LacZ* reporter gene using a

colony-lift filter assay for β-galactosidase activity (see Material and Methods). Of the 191 His+ clones obtained in this screen, 57 were also LacZ+ (Figure 3) and were investigated further.

In order to identify the two-hybrid positive clones, cDNA inserts from 40 His+/LacZ+ positives were directly amplified from yeast using the polymerase chain reaction (see Materials and Methods) and vector-derived primers. Restriction digests were used to sort clones and eliminate duplicates, and unique cDNAs were sequenced and identified by database searching. Table 1 lists the identities of 29 cDNAs isolated in the two-hybrid screen. Several cDNAs were isolated more than once, as indicated by the number in the left column of the table. In addition, some clones contained only 3-prime untranslated sequences and no coding sequences, as indicated in the table (second column). Of the proteins identified in this screen, two (SCG10 and Rab6) are known Golgi-associated proteins (Lutjens et al., 2000; Van Wye et al., 1996), and five (α-spectrin, SCG10, SUN2, zyxin, and β-tubulin) are associated with the cytoskeleton and affect motility (Charbaut et al., 2001; Cianci et al., 1999; Malone et al., 1999; Nix et al., 2001). Because our data have shown that GRP50 is a membrane-associated protein, these proteins seem particularly intriguing as candidates for GRP50 interacting proteins.

Some of the cDNA clones obtained in our screen did not appear to contain protein-coding sequences immediately downstream of the GAL4 AD, and many that did contain protein-coding sequences were out of frame with the GAL4 AD. However, translational read-through across non-translated gaps can occur in *S. cerevisiae*, and translational frameshifting is common in this organism (Dinman, 1995). We therefore tested the two-hybrid positive yeast clones listed in Table 1 for the presence of GAL4

AD-fusion proteins of the expected sizes (based on cDNA sequence). Yeast protein extracts were prepared from two-hybrid positive strains and analyzed by immunoblotting (data not shown, see Materials and Methods). Based on sequence and protein expression results, the eight proteins listed above the solid line in Table 1 were selected as the first set of clones to be analyzed further.

We sought to confirm the two-hybrid results by alternative methods. The cDNAs encoding these eight proteins were subcloned into the mammalian expression vector pCMV-Tag2 (Stratagene) and coexpressed with myc_6-GRP (see Chapter 3) in cultured mammalian cells. Transfected cells were analyzed by immunofluorescence microscopy and immunoprecipitation to determine whether these proteins co-localize and/or co-purify with myc_6-GRP (data not shown). To date, we have been unable to confirm the interaction of any of the two-hybrid positives with GRP50. Efforts to characterize and, if possible, confirm all of the positives obtained in the two-hybrid screen are ongoing.

DISCUSSION

This section describes the results of a two-hybrid screen for GRP50-interacting proteins. Our efforts to carry out this screen were complicated by the fact that (over)expression of the full-length GRP50 protein (fused with the GAL4 AD) is toxic in yeast. In addition, all of the GAL4 BD-GRP50 bait constructs were capable of transactivating one or more of the reporter genes used in this screen. As a result of these properties of the GRP50 protein, the selection for two-hybrid interacting proteins was not sufficiently stringent to prevent isolation of a large number of positives. For this reason, many if not most of the clones isolated are likely to be false positives (Bartel et al., 1993; Serebriiskii et al., 2000).

The transactivating properties of the GAL4 BD-GRP50 fusion protein could be mediated by several factors. First, the GRP50 protein contains multiple regions containing acidic residues that might mimic transactivation domains; short amphipathic acidic domains have been shown in other proteins to result in inappropriate transcriptional activation in the two-hybrid system (Ruden, 1992; Ruden et al., 1991). In addition, the presence of multiple protein-interacting domains in the GRP50 protein could result in recruitment, to the reporter gene promoters, of cellular factors that activate transcription. The observation that the GAL4 BD-mGRP(full) fusion protein transactivates the *HIS3* reporter, while the GAL4 BD-hGRP(1-330) does not may be explained by the differences between these two proteins. These proteins differ in the carboxyl-terminal 38 residues of GRP50 (deleted in hGRP(1-330) and in the proline-rich central domain, which exhibits the greatest degree of divergence between the mouse and

human GRP50 proteins (See Chapter 2, Figure 6). Transactivation of the HIS3 reporter could be due to residues in the central, proline-rich domain of GRP50, since the strain expressing the fusion protein lacking this domain, GAL4 BD-mGRP(-pro), is His- and is phenotypically identical to the GAL4 BD-hGRP(1-330) strain.

Yeast two-hybrid screens frequently yield a number of false positive results (Serebriiskii et al., 2000). Because of the toxic and transactivating effects that result from expression of GRP50 as a GAL4-fusion protein, it was particularly difficult to rule out false positives among the large number of clones isolated in this screen. To date, only the strongest positives (from the β-galactosidase assay) have been characterized, and we are continuing in our efforts to identify the remaining clones. Still, some of the positive clones we have identified may represent genuine GRP50-interacting proteins. The identification of multiple proteins that function in cytoskeletal changes associated with cellular motility and neurite outgrowth is particularly intriguing, given the architectural structure and membrane-associated localization of the GRP50 protein. If confirmed, these types of protein interactions might point to a role for GRP50 in linking the membranes of the Golgi and/or vesicles with proteins of the cytoskeleton. Identification of GRP50 binding partners is critical to understanding the molecular function of this protein, and future efforts will continue to address this problem.

MATERIALS AND METHODS

Expression Plasmids and Yeast Strains

Bait constructs were generated in the plasmid vector pAS2-1 (Clontech), which encodes the GAL4 DNA-BD as an N-terminal fusion with the protein of interest. pAS2-1 contains the *TRP1* gene for use as a selectable marker; yeast carrying the plasmid are selected for the ability to grow on medium lacking tryptophan. The cDNAs encoding the various domains of the murine or human GRP50 depicted in Figure 2 were cloned into BamHI/XhoI digested pAS2-1, and junctions were verified by sequencing with the pAS2-1 forward primer 5'-TCGGAAGAGAGTAGTAAC-3' and the pAS2-1 reverse primer 5'–CCTGAGAAAGCAACCTGACC-3'. Bait constructs were introduced into the haploid yeast strain PJ69-2A (MATa) by electroporation according to standard procedures (Ausubel et al., 2001) and bait strains were isolated by selecting for Trp+ colonies.

Four independent clones of each bait strain were isolated, and His, Ade, and LacZ phenotypes were verified by comparison to the control bait strain PJ69-2A(pVA3-1). Bait strains were tested for toxicity by measuring the growth rate in liquid cultures (by OD_{600}) of bait and control strains. Mating efficiencies of bait and control strains were compared by performing a test mating with the AD-library control strain Y187(pTD1-1). Yeast protein extracts were prepared from bait strains, and expression of fusion proteins was verified by immunoblotting with anti-GAL4 BD antibody (not shown).

The human fetal brain AD-library, pretransformed into the haploid yeast strain Y187 (MATα), was purchased from Clontech. Human fetal brain cDNAs were cloned

into the plasmid vector pACT2, which encodes the GAL4 AD as an N-terminal fusion with a library of cDNA sequences. pACT2 contains the *LEU2* gene for use as a selectable marker; yeast carrying the plasmid are selected for the ability to grow on medium lacking leucine. The original cDNA library was estimated to contain 3.5 x 10^6 independent cDNA clones prior to amplification, and the titer of the library (in Y187) was given as >5 x 10^7 cfu/ml (according to manufacturer's product data sheet).

Growth, Maintenance, and Selection of Yeast Strains

Yeast strains were maintained according to the manufacturer's instructions and standard techniques (Ausubel et al., 2001). Yeast media and supplements were purchased from Clontech (except where indicated) and prepared according to the manufacturer's instructions, except that rich medium (YPD) was supplemented with 0.003% adenine hemisulfate (Sigma) and 15 μg/ml kanamycin. Minimal selection medium (SD) was supplemented with the appropriate "dropout" powder, (e.g., -Trp, -Trp/-His, -Trp/-Leu/-His/-Ade) and 0.003% adenine hemisulfate (except -Ade media). All SD media lacking histidine (-His formulations) were supplemented with 5 mM 3-amino-1,2,4-triazole (3-AT) to suppress background growth due to leaky *HIS3* reporter gene expression (Fields, 1993). The optimal 3-AT concentration necessary to prevent growth of the bait strain PJ69-2A(hGRP-1-330) on SD –His plates was determined empirically.

Two-hybrid library screening

The bait strain PJ69-2A(hGRP-1-330) was mated with 1 ml of the pretransformed Y187(library) strain according to the manufacturer's instructions. Briefly, SD -Trp liquid

medium was inoculated with strain PJ69-2A(hGRP-1-330) and grown overnight at 30°C with shaking. The bait culture (OD_{600} = 1.1) was harvested by centrifugation, resuspended, and combined with 1 ml pretransformed library in a 2 L flask. YPDA/Kan (YPD plus adenine and kanamycin) medium was added to a final volume of 50 ml and the culture was incubated overnight at 30°C with gentle swirling (50 rpm). The mating culture was harvested by centrifugation, resuspended in ~10 ml YPDA/Kan, and plated onto 54 15 cm plates containing SD-agar -Trp/-Leu/-His/-Ade (QDO medium), plus 5 mM 3-AT. Plates were incubated at 30°C; strong positives were picked and re-streaked onto QDO medium after 8 days, and weak positives were picked and re-streaked after 15 days. The mating efficiency in this experiment was low (1.5% vs. expected >5%), and the total number of clones screened was ~8 x 10^5.

Filter-lift assay for β-galactosidase activity

Two-hybrid positive clones that survived the nutritional selection were streaked onto SD-agar (-Trp/-Leu) plates and assayed for β-galactosidase activity using a previously described method (Breeden and Nasmyth, 1985). The positive control diploid PJ69-2A(pVA3-1) x Y187(pTD1-1) and the negative control diploid PJ69-2A(pAS2-1) x Y187(pTD1-1) were always included in these experiments for comparison. Freshly grown colonies were transferred to filter paper (Grade 410, VWR) using the colony-lift method, and cells were permeabilized by immersing the filters in LN_2 for 10 seconds. Following freeze/thaw, filters were transferred (colony side up) to filter paper that had been pre-soaked in assay buffer Z (150 mM $NaHPO_4$, 10 mM KCl, 1 mM $MgSO_4$, pH 7), plus 40 mM β-mercaptoethanol, and 0.33 mg/ml 5-bromo-4-chloro-3-indolyl-β-D-

galactopyranoside (X-gal). Filters were incubated at 30°C and monitored hourly for the appearance of blue color; yeast strains were scored as positive in this assay if colonies turned blue within ~4 hours. At later time points (between 6 and 8 hours), false positives appeared (as judged by phenotypes of negative control strains).

Identification and analysis of positive cDNA clones

cDNA inserts were amplified directly from two-hybrid positive yeast strains using the polymerase chain reaction and the following pairs of nested primers (derived from the pACT2 AD-library vector):

pACT2 set 1	forward 5'-TCTATTCGATGATGAAGATACCC-3'
	reverse 5'- GTGCACGATGCACAGTTG-3'
pACT2 set 2	forward 5'-CCCACCAAACCCAAAAAAAG-3'
	reverse 5' GAACTTGCGGGGTTTTTCAG

For each positive strain, a single large (>2 mm), fresh yeast colony was scraped into 100 µl of 10 mM Tris (pH 7.5), 0.1 mM EDTA. Samples were vortexed vigorously and boiled for 10 minutes, then vortexed again. Samples were centrifuged for 5 minutes to remove debris, and 5 µl of the supernatant was removed for use as the template for amplification using the proofreading thermostable polymerase *PfuTurbo* (Stratagene). The "set 1" primers listed above were used in the first amplification step; upon completion of the reaction, 1 µl of the product was removed to a new plate for the second amplification, using the set 2 (internal) primers.

PCR-amplified cDNA clones were digested with *AluI* and the resulting restriction fragment polymorphisms were used to eliminate duplicate clones. Unique clones were

gel-purified and sequenced using an ABI automated sequencer and the primers listed above. Clones were identified by scanning the nonredundant GenBank sequence database, using the BLAST server at the NCBI (ncbi.nlm.nih.gov/blast).

Selected two-hybrid positive clones were further characterized as follows: In order to determine whether GAL4 AD-fusion proteins of the expected sizes (based on cDNA sequence) were produced in these strains, protein extracts were prepared and analyzed by immunoblotting with the 12CA5 antibody (Santa Cruz Biotechnology, data not shown). 12CA5 recognizes the HA-epitope encoded downstream of the GAL4 AD in pACT2. Strains in which GAL4 AD-fusion protein were detected were then used to prepare plasmid DNA from yeast using the Yeastmaker™ Yeast Plasmid Isolation Kit (Clontech). Selected cDNA inserts were excised from pACT2 and subcloned into the mammalian expression pCMV-Tag2, which encodes an NH_2-terminal FLAG epitope tag. The FLAG-tagged proteins were coexpressed with myc_6-GRP (see Chapter 3) in cultured mammalian cells and assayed for co-localization (by indirect immunofluorescence microscopy) and/or co-immunoprecipitation with the myc_6-GRP protein (not shown).

ACKNOWLEDGEMENTS

I would like to thank Rebecca Scarr for advice regarding the two-hybrid screen in general and for experimental protocols used in this screen. I am especially indebted to Vanessa Cheung for assistance with the experiments described in this section.

Figure 1

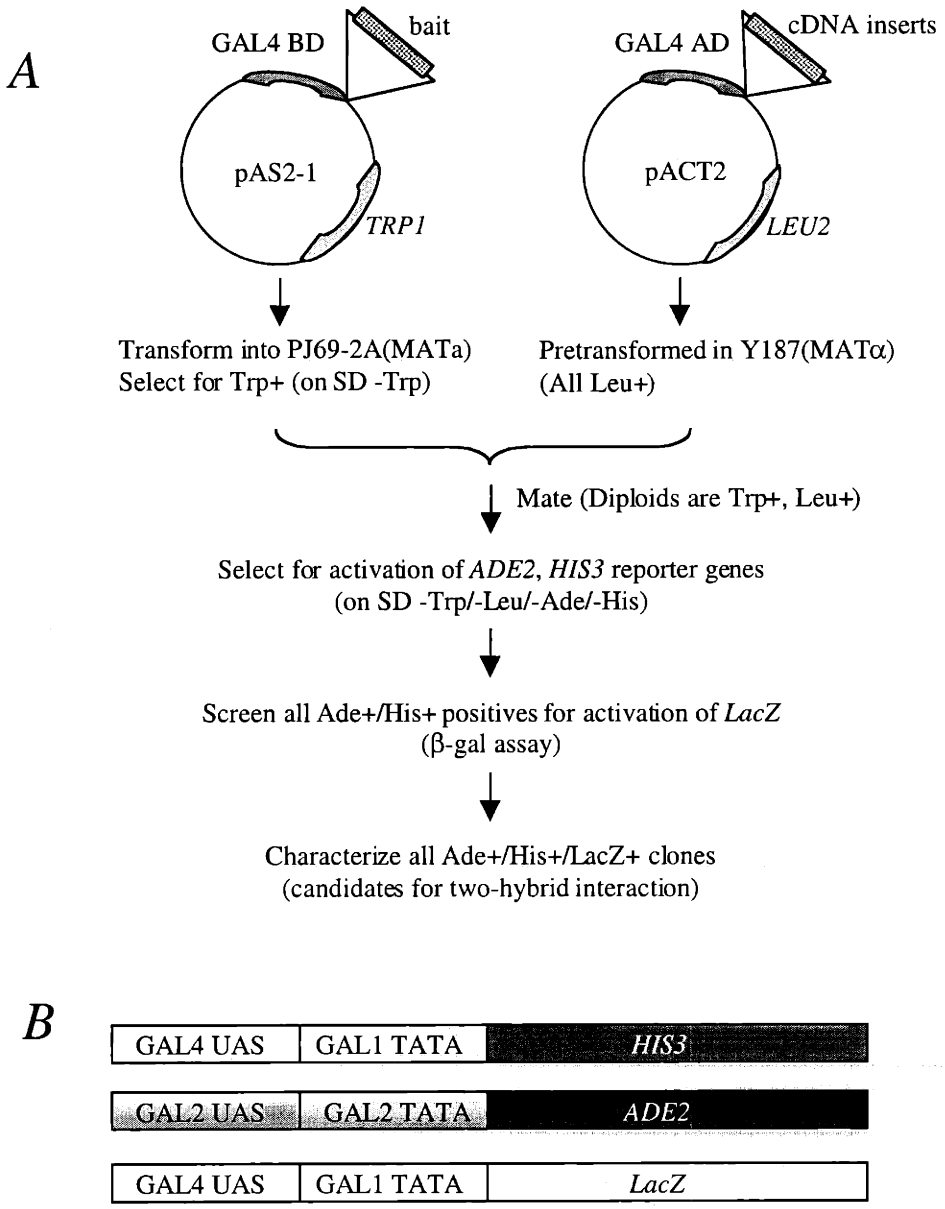


Figure 1. ***(A)*** Outline of the two-hybrid system used in this screen. A bait strain, constructed in PJ69-2A (MATa), is mated to the pretransformed library strain Y187 (MATα). PJ69-2A and Y187 are auxotrophic for Trp, Leu, Ade, and His. The diploid mating products are selected for activation of the nutritional reporters *ADE2* and *HIS3*, and positive clones are screened for activation of the *LacZ* reporter. ***(B)*** The diagram shows the structure of the *ADE2* and *HIS3* and *LacZ* reporter genes. *ADE2* and *HIS3* are driven by heterologous promoters to minimize selection of false positives.

Figure 2

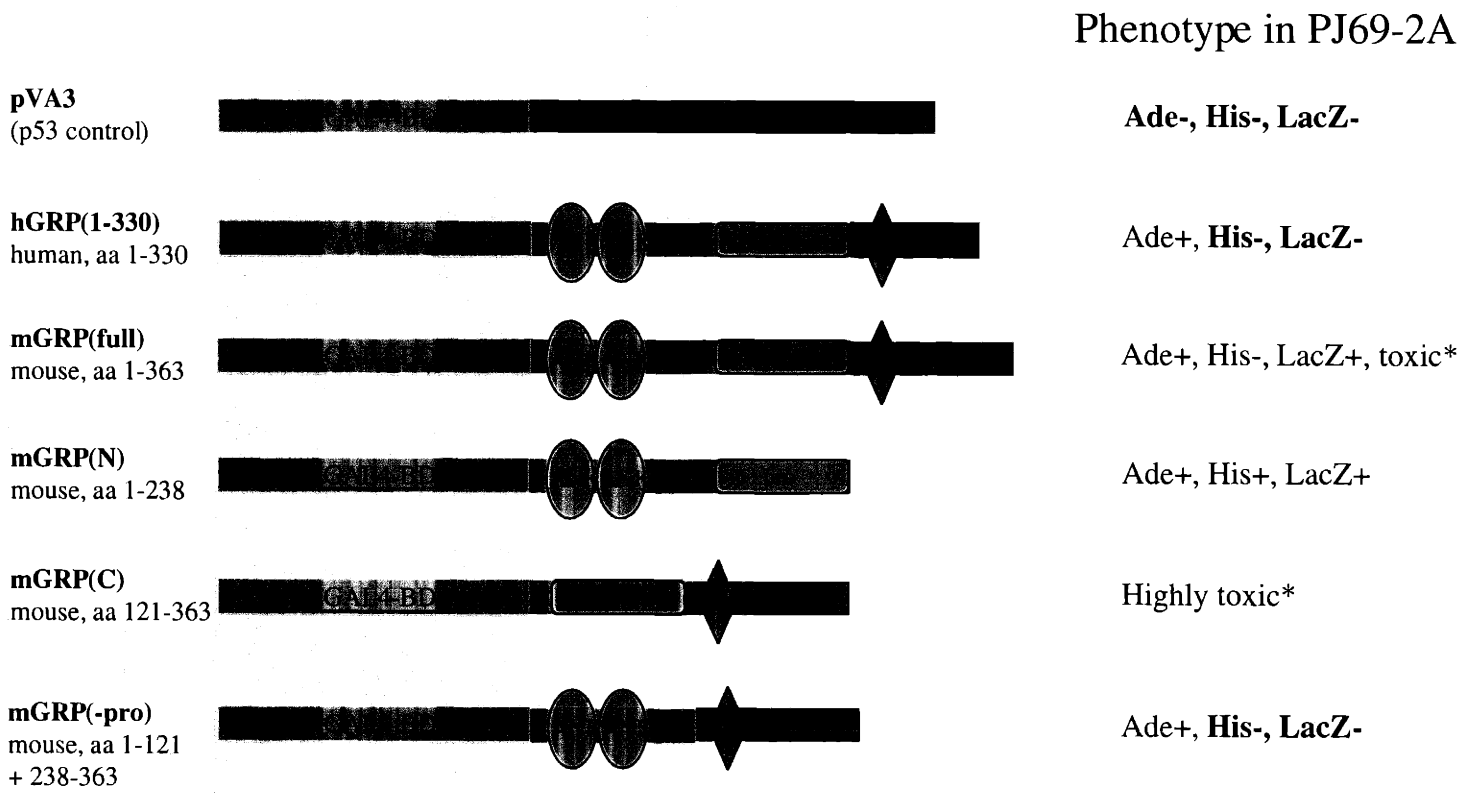


Figure 2. Phenotypic analysis of GAL4 BD-bait fusion proteins in yeast strain PJ69-2A. Cartoons depict the five GAL4 BD-GRP50 bait constructs generated, along with the control bait construct pVA3. Amino acids included in each GRP50 construct are indicated (numbering is from initiator methionine, see Chapter 1 for sequence). Constructs were introduced into PJ69-2A, and the resultant strains were assayed for growth on medium lacking adenine or histidine and for β-gal activity (see Materials and Methods). Expected phenotypes are in bold. *Transformants derived from the mGRP(full) construct grew poorly, and those derived from mGRP(C) were not viable.

Figure 3

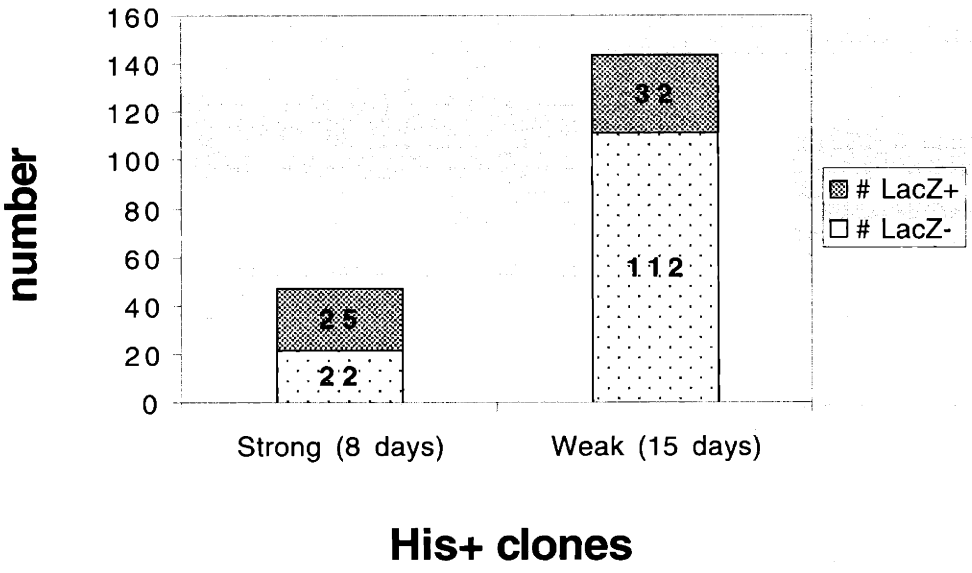


Figure 3. Summary of positive clones obtained in the two-hybrid screen. Clones that survived the nutritional selection are designated His+; strong positives grew on selective medium within 8 days, and weak positives grew within 15 days. All clones were tested for activation of the *LacZ* reporter using a filter-based β-galactosidase assay (see Materials and Methods), and the numbers of LacZ+ and LacZ- clones within each class are indicated.

Table 1

# obtained	Clone ID	Function
4	AHCY, encodes S-adenosyl homocysteine hydrolase	48 kDa enzyme that promotes transmethylation reactions
1	BCRG2 (XM_031102), breakpoint cluster region protein 2	Uncharacterized protein, maps to breakpoint cluster from uterine leiomyoma
2	L-ferritin	Ferritin light chain subunit (iron binding protein)
1	Non-erythroid α-II-spectrin (α-fodrin)	Cytoskeletal protein, rain and muscle cell specific isoform
2	SCG10 (neuron-specific stathmin isoform)	Neuronal growth-associated protein (nGAP), destabilizes microtubules at growth cone, localizes to Golgi
1	Serine protease inhibitor, Kunitz type 2 (bikunin)	Plasma protease inhibitor, inhibits plasmin and proteases of the coagulation pathway
1	SUN2 (KIAA0668), sad1-unc84 homeodomain protein	Involved in nuclear migration, may link nuclear envelope to MTOC
2	zyxin	Low-abundance cytoskeletal phosphoprotein, localizes to focal adhesions, binds α-actinin
1	β-tubulin	Monomer component of microtubules
1	AC026888, from 14q24.3-31	Unknown
1	AK025855, unknown cDNA	Unknown
1	DKFZp761D2324	Hypothetical protein, contains ricin-like β-trefoil
1	FEZ1/LZTS1 (3' UTR)	Putative tumor suppressor gene encodes leucine zipper protein
1	Glycoprotein M6A (GPM6A), (3' UTR)	Membrane glycoprotein associated with erythroleukemia
2	L1-cell adhesion molecule (L1CAM) (3' UTR)	Extracellular protein involved in cell migration and neurite outgrowth
1	LERK-6/Ephrin-A2 (EPLG6) (intron)	Receptor tyrosine kinase involved in cell migration and axon guidance
2	NM_021932, hypothetical protein from 11p15.5 (3' UTR)	unknown
1	RAB6 cDNA (3' UTR)	Golgi-associated GTPase of the Ras superfamily, involved in membrane trafficking
3	Unrelated sequences containing Alu repeats	

Table 1. List of 29 two-hybrid positive clones and their identities. The number in the left column indicates the frequency with which a given sequence was isolated. Some clones contained only 3' UTR sequences and no coding sequences (indicated in parentheses).

REFERENCES

Aho, S., A. Arffman, T. Pummi, and J. Uitto. 1997. A novel reporter gene MEL1 for the yeast two-hybrid system. *Anal Biochem.* 253:270-272.

Ausubel, F.M., R. Brent, R.E. Kingston, D.D. Moore, J.G. Seidman, J.A. Smith, and K. Struhl. 2001. Current Protocols in Molecular Biology. John Wiley & Sons, Inc.

Bartel, P., C.T. Chien, R. Sternglanz, and S. Fields. 1993. Elimination of false positives that arise in using the two-hybrid system. *Biotechniques.* 14:920-924.

Bartel, P.L., and S. Fields. 1997. The Yeast Two-Hybrid System. Oxford University Press, Oxford.

Bendixen, C., S. Gangloff, and R. Rothstein. 1994. A yeast mating-selection scheme for detection of protein-protein interactions. *Nucleic Acids Res.* 22:1778-1779.

Breeden, L., and K. Nasmyth. 1985. Regulation of the yeast HO gene. *Cold Spring Harb Symp Quant Biol.* 50:643-650.

Charbaut, E., P.A. Curmi, S. Ozon, S. Lachkar, V. Redeker, and A. Sobel. 2001. Stathmin family proteins display specific molecular and tubulin binding properties. *J Biol Chem.* 276:16146-16154.

Chien, C.T., P.L. Bartel, R. Sternglanz, and S. Fields. 1991. The two-hybrid system: a method to identify and clone genes for proteins that interact with a protein of interest. *Proc Natl Acad Sci U S A.* 88:9578-9582.

Cianci, C.D., Z. Zhang, D. Pradhan, and J.S. Morrow. 1999. Brain and muscle express a unique alternative transcript of alphaII spectrin. *Biochemistry.* 38:15721-15730.

Dinman, J.D. 1995. Ribosomal frameshifting in yeast viruses. *Yeast.* 11:1115-1127.

Estojak, J., R. Brent, and E.A. Golemis. 1995. Correlation of two-hybrid affinity data with in vitro measurements. *Mol Cell Biol.* 15:5820-5829.

Fields, S. 1993. The two-hybrid system to detect protein-protein interactions. *METHODS: A Companion to Meth. Enzymol.* 5:116-124.

Fields, S., and O. Song. 1989. A novel genetic system to detect protein-protein interactions. *Nature.* 340:245-246.

Golemis, E.A., J. Gyuris, and R. Brent. 1996. Analysis of protein interactions; Interaction trap/two-hybrid systems to identify interacting proteins. *In* Current Protocols in Molecular Biology. John Wiley & Sons, Inc. Sections 20.21 and 20.22.

Iwabuchi, K., B. Li, P. Bartel, and S. Fields. 1993. Use of the two-hybrid system to identify the domain of p53 involved in oligomerization. *Oncogene*. 8:1693-1696.

Li, B., and S. Fields. 1993. Identification of mutations in p53 that affect its binding to SV40 large T antigen by using the yeast two-hybrid system. *Faseb J.* 7:957-963.

Lutjens, R., M. Igarashi, V. Pellier, H. Blasey, G. Di Paolo, E. Ruchti, C. Pfulg, J.K. Staple, S. Catsicas, and G. Grenningloh. 2000. Localization and targeting of SCG10 to the trans-Golgi apparatus and growth cone vesicles. *Eur J Neurosci.* 12:2224-2234.

Malone, C.J., W.D. Fixsen, H.R. Horvitz, and M. Han. 1999. UNC-84 localizes to the nuclear envelope and is required for nuclear migration and anchoring during C. elegans development. *Development*. 126:3171-3181.

McNabb, D.S., and L. Guarente. 1996. Genetic and biochemical probes for protein-protein interactions. *Curr Opin Biotechnol.* 7:554-559.

Nix, D.A., J. Fradelizi, S. Bockholt, B. Menichi, D. Louvard, E. Friederich, and M.C. Beckerle. 2001. Targeting of zyxin to sites of actin membrane interaction and to the nucleus. *J Biol Chem.* 276:34759-34767.

Ruden, D.M. 1992. Activating regions of yeast transcription factors must have both acidic and hydrophobic amino acids. *Chromosoma.* 101:342-348.

Ruden, D.M., J. Ma, Y. Li, K. Wood, and M. Ptashne. 1991. Generating yeast transcriptional activators containing no yeast protein sequences. *Nature*. 350:250-252.

Serebriiskii, I., J. Estojak, M. Berman, and E.A. Golemis. 2000. Approaches to detecting false positives in yeast two-hybrid systems. *Biotechniques*. 28:328-330, 332-326.

Van Wye, J., N. Ghori, P. Webster, R.R. Mitschler, H.G. Elmendorf, and K. Haldar. 1996. Identification and localization of rab6, separation of rab6 from ERD2 and implications for an 'unstacked' Golgi, in Plasmodium falciparum. *Mol Biochem Parasitol.* 83:107-120.