Studies on the Quality Control Apparatus of Glycoprotein Folding in the Endoplasmic Reticulum

Marc-François Pelletier

Department of Biology
McGill University, Montréal
Date: August, 2001

A thesis submitted to the Faculty of Graduate Studies and Research in partial fulfillment of the requirements for the degree of Ph. D.

ABSTRACT

As nascent secretory and membrane proteins are inserted into the endoplasmic reticulum (ER), they are maintained in folding and/or assembly competent states by molecular chaperones including the Hsp70 and Hsp90 homologues, BiP and GRP94, and the lectin-like chaperones calnexin (CNX) and calreticulin (CRT). Folding is catalyzed by protein disulfide isomerase (PDI), its CNX (and CRT) associated homologue, ERp57, and protein prolyl isomerase (PPI). Moreover, N-linked glycoproteins benefit from a lectin-based "quality control apparatus" that ensures their correct folding or oligomeric assembly. Binding to these lectins occurs through oligosaccharide trimming from $Glc_3Man_9GlcNAc_2$ to the monoglucosylated form ($Glc_1Man_9GlcNAc_2$). Release and subsequent rebinding occurs though the hydrolysis and reglucosylation of the innermost glucose by glucosidase II and UDP-glucose glycoprotein:glucosyltransferase, respectively. This cyclical process, termed the "Calnexin Cycle", continues until their correct conformation is achieved.

The cloning and characterization of human glucosidase II is reported here. cDNAs for two splice variants of the catalytic α subunit and the β subunit were isolated. Expression of the β subunit was shown to be required for enzymatic activity, solubility and/or stability, and ER retention of the enzyme. Detailed kinetic analysis on recombinant α1/β and α2/β isoforms, using *p*-nitrophenyl α-D-glucopyranoside as a substrate, reveals that both exhibit kinetic profiles of a two binding site model, and share properties of catalysis and inhibition on this substrate. Moreover, similar rates of hydrolysis of the oligosaccharide substrates rules out the possibility that the two binding site kinetic model, first proposed by Alonso *et al.* (1999, *Biochem J.* **278**:721-7), is the result of co-purified isoforms of glucosidase II that have different substrate specificities.

Also, an ER protein two-hybrid system, based on Ire1p and the unfolded protein response (UPR) pathway in *Saccharomyces cerevisiae*, was developed to examine and map the interactions between CNX/CRT and ERp57. Ire1p fusions with CNX and CRT were shown to interact specifically with ERp57, and as expected, PDI did not. Through deletion analysis, new roles were assigned to the proline-rich loop domains of CNX and CRT, and the non-catalytic B thioredoxin domain of ERp57 in mediating their heterodimerization.

RÉSUMÉ

Lorsque les protéines secrétoires et membranaires sont inserrées dans le réticulum endoplasmique (RE), elles sont maintenues en état de compétences au rempliement et à l'assemblage oligomérique par les chaperons moléculaires incluant les homologues de la Hsp70 et la Hsp90, BiP et GRP94, et les chaperones-lectines, la calnexin (CNX) et la calreticuline (CRT). Le repliement est catalysé par l'isomérase de protéines disulfure (PDI), son homologue associé à la CNX et la CRT, la ERp57, et par l'isomérase de protéines prolyl. De plus, les glycoprotéines avec oligosaccharides liées à l'asparagine bénéficient d'un système de contrôle de qualité qui assure le repliement et l'assemblage oligomérique correct. L'intéraction avec les chaperons-lectines est induite par l'hydrolyse de l'oligosaccharide de la forme $Glc_3Man_9GlcNAc_2$ à la forme monoglucosylée ($Glc_1Man_9GlcNAc_2$). Les glycoprotéines sont relâchées et intéragissent à nouveau, par un procédé cyclique appelé "le Cycle Calnexine". Ce cycle est composé de qui l'hydrolyse et de la reglucosylation dû à l'activité des enzymes glucosidase II et la glucosyl transférase de UDP-glucose aux glycoprotéines, respectivement.

Le clonage et la caractérisation de la glucosidase II humaine sont présentés dans cet ouvrage. Deux cDNA de la sous-unité catalytique α (α1 et α2) et la sous-unité β de la glucosidase II humaine ont été clonés. Nous avons trouvé que l'expression de la sous-unité β est requise pour l'activité enzymatique, la solubilité et/ou la stabilité et la rétention du complexe hétérodimérique dans le RE. Une analyse cinétique détaillée des complexes recombinants α1/β et α2/β a été effectuée avec le substrat *p*-nitrophényl-α-D-glucopyranoside. Les deux complexes hétérodimériques possèdent des profils cinétiques qui répondent à un modèle de deux sites actifs et partagent les mêmes propriétés de catalyse et d'inhibition. La même hydrolyse, observée sur des substrats oligosaccharides physiologiques exclue la possibilité que le modèle cinétique proposé à priori par Alonso *et al.* (1999, *Biochem. J.*, **278** :721-727) est le résultat de la co-purification de deux isoformes possédant des préférences de substrats différentes. Elle exclut aussi l'existence de deux formes distinctes de la glucosidase II avec des activités pré- et post-calnexine.

Finalement, un système de double-hybrides en levure, basé sur Ire1p et l'activation du système de réponse aux protéines dépliées du RE de *Saccharomyces cerevisiae*, a été développé pour étudier les intéractions entre ERp57 et la CNX et la CRT

calréticuline. Lorsque le domaine luminal de Ire1p est remplacé par ERp57, l'hétérodimérisation est observée avec des fusions qui expriment la CNX ou la CRT. Tel que prévu, PDI qui est un proche homologue de ERp57, n'intéragit ni avec la CNX ou la CRT dans notre système. Utilisé pour identifier les sites d'intéractions entre ERp57 et la CNX et la CRT, le système démontre que le domaine riche en proline "en forme de crochet" de la CNX et la CRT et le domaine thiorédoxine B de ERp57 sont requis pour l'intéraction entre ces protéines. Nous proposons ainsi un nouveau rôle pour le domaine riche en proline de la CNX et la CRT, et du domaine thioredoxine B de la ERp57, dans l'intéractions de ces composants.

PREFACE

This thesis has been written and assembled in accordance with the guidelines of the Faculty of Graduate Studies and Research of McGill University. The thesis is comprised of an Abstract/Résumé, an Introduction (Chapter 1), a Results section (Chapters 2 and 3), a Conclusion (Chapter 4), and a Reference section. As chapters 2 and 3 have been published or submitted for publication, they each have an abstract, material and methods, results, and discussion section (or results/discussion, chapter 3).

An original contribution is made here, to our understanding of the structural determinants for enzymatic activity of the human form of the ER lumenal enzyme glucosidase II. Glucosidase II mediates N-glycan trimming in the endoplasmic reticulum, and also the entry/exit from the "Calnexin Cycle". In addition, a detailed kinetic analysis was performed which demonstrates that genetically purified and expressed isoforms of the heterodimeric complex both have a complex active site with two substrate binding sites, and both exhibit identical properties of catalysis and inhibition. Lastly, a technology was developed to test and map the protein-protein interactions of the glycoprotein folding matrix, which includes calnexin and calreticulin, and the ER/monoglucosylation specific protein disulfide isomerase, termed ERp57, within the endoplasmic reticulum. Functions were assigned to the "long arm" domains of calnexin and calreticulin, and to the non-catalytic B thioredoxin domain of ERp57.

The work presented here is largely my own with the following exceptions: Anne Marcil performed the Western analysis and immunoprecipitations in Figure 2, generated Figures 3,4 and 6, and performed the immunoprecipitions presented in Figure 5.

Chapter 1 is currently in press and will be published as a chapter under the same title by Marc F. Pelletier, John J. M. Bergeron , and David Y. Thomas in *Molecular Chaperones in the Cell: The Frontiers in Molecular Biology Series* (Peter Lund, ed, Oxford University Press, Oxford, U.K) and is used with the permission of Oxford University Press. Chapter 2 has been published as a paper by Marc F. Pelletier, Anne Marcil, Guy Sevigny, Claude A. Jakob, Daniel C. Tessier, Eric Chevet, Robert Ménard, John J.M. Bergeron, David Y. Thomas in the journal *Glycobiology* **10**:815-827 and is also used with the permission of Oxford University Press. Chapter 3 has been submitted to *Nature Biotechnology* as a paper by Marc F. Pelletier, Gregor Jansen, John J.M.

Bergeron, and David Y. Thomas. These papers form an integral part of this thesis and are presented in a logical manner.

ACKNOWLEDGEMENTS

I am truly indebted to my supervisor Dr. Dave Y. Thomas, for his unwavering confidence, encouragement, support, patience, and vision. An excellent example of this was demonstrated during the second provincial referendum, when he stood his ground and gave us the confidence we needed to endure the uncertainty of the NRC's survival in Quebec. "The BRI is on wheels", he said, and our nerves were calmed with visions of life in beautiful Ottawa. Humor aside, his commitment to his students and to teaching is exceptional and an example of true scholarship. Lastly, I am grateful for his financial assistance throughout my graduate studies at McGill, and for the many scientific meetings I attended.

I would also like to acknowledge the support and encouragement of Dr. John J.M. Bergeron. His passion for science and knowledge transcends all levels of academic achievement and is an inspiration to us all. As well, I'd like to thank Dr. Howard Bussey for his invaluable academic and scientific guidance at my all too many graduate training committee meetings. I am truly fortunate to have had such a renowned and experienced team overseeing my research throughout my academic training.

Many thanks go out to my colleagues and friends of the Thomas Laboratory. Special thanks to Drs. Malcolm Whiteway and Andre Nantel for the many insightful scientific discussions, both on and off the basketball court. They are both astute and devoted scientists, generous with their time, and make not-so-bad BB teammates. As well, among others I must thank Daniel Tessier for the years of technical supervision, for which I am truly grateful, and also to Josée Ash for her support both in and out of the laboratory. The laboratory could not have functioned without Josée, and it would not be an exaggeration to say that her absence would have led to certain chaos, and most probably of my doing.

I would also like to thank Anne Marcil for her contributions to the glucosidase II project. It was a pleasure to work with her, and quite easy to be inspired by her dedication and talent. Also, I'd like to thank Dr. Gregor Jansen for his expertise in yeast genetics, and for his help in the development of the ER protein two-hybrid system. I must also acknowledge the support of my compatriot-colleague-fellow student-in-arms, soon to be

Dr. Sean Taylor. His exuberant energy, patience, and drive were generously shared, and at many times, greatly needed.

To my mom for her absolutely unwavering support and encouragement, (and motherly praise) and to my father for teaching me the joys of scientific inquiry and technical challenge, many thanks. And, thank you to my brother Phil for raising the bar of personal challenge and success. Thanks as well to Robin for his encouragement and friendship thoughout these long and arduous years of graduate studies.

Το Πελαγια, Διμιτρι, και Πανογιοτης, ευχαριστο πολη για την αγαπη, ψποστιριζ η, και κατανοισι. Εσεις με βοιθησατε να καταλαβο τη σιμενι οικογενεια.

Lastly and most of all, thanks to my wife Maria. There is no doubt in my mind that I would not be where I am today without her. She inspires me to live every day to its fullest extent, and provides me with an endless source of energy and encouragement. She certainly is deserving of a Nobel "Spouse" prize for patience and understanding, and unquestionably adds significance to the meaning of "trophy" wife.

Ευχαριστο πολη μπανη μου. Θα σ'αγαπο για παντα και ελπιζο οτι μπορο να δωσο σ'εσ ενα οτι εχεις δωδει εσει σε μενα

TABLE OF CONTENTS

TABLE OF CONTENTS (continued)

TABLE OF CONTENTS (continued)

TABLE OF CONTENTS (continued)

TABLE OF CONTENTS (continued)

LIST OF FIGURES

LIST OF FIGURES (continued) **PAGE**

LIST OF TABLES

PAGE

LIST OF ABBREVIATIONS

AcP	acid phosphatase
ALLM	N-acetyl-L-leucinyl-L-leucinyl-L-methionine
ALLN	N-acetyl-L-leucinyl-L-leucinyl-L-norleucine
AZC	L-azetidine-2-carboxylic acid
CFTR	cystic fibrosis transmembrane conductance regulator
CPY	carboxypeptidase Y
CS	citrate synthase
CST	castanospermine
DNJ	deoxynojirimycin
DTT	dithiothreitol
Endo-H	endoglycosidase H
ER	endoplasmic reticulum
ERAD	ER-associated degradation pathway
ES	embryonic stem cell
G0-AcP	acid phosphatase with N-linked $Man_9GlcNAc_2$
G1-AcP	acid phosphatase with N-linked $Glc_1Man_9GlcNAc_2$
Glc	glucose
GlcNAc	glucosamine
HCMV	human cytomegalovirus
HMG-R	3-hydroxy-3-methylglutaryl-CoA reductase
JNK	c-Jun activated NH_2–terminal kinase
Man	mannose
MDH	malate dehydrogenase
MG-132	N-carbobenzoxyl-L-leucinyl-L-leucinyl-L-leucinyl
MHC I	major histocompatibility complex class I
PAGE	polyacrylamide gel eletrophoresis
PCR	polymerase chain reaction
PDI	protein disulfide isomerase
p-NP-Glc	*p*-nitrophenyl α-D-glucopyranoside
p-NP-Glcase	*p*-nitrophenyl α-D-glucopyranosidase

PPI	protein prolyl isomerase
ppαF	prepro-α Factor
SBA	soybean agglutinin
SAPK	stress activated protein kinase
TCA	trichloroacetic acid
TM	tunicamycin
TNF	tumor necrosis factor
UBC	ubiquitin conjugating enzyme
UGGT	UDP- glucose glycoprotein:glucosyltransferase
UPR	unfolded protein response
VSV	vesicular stomatitis virus

Chapter 1

Molecular Chaperone Systems in the Endoplasmic Reticulum

1.1 Introduction

The endoplasmic reticulum (ER) is an organelle defined microscopically as network of saccular and tubular membranes in continuity with the nuclear envelope. This definition can be further refined into a distinction between the rough endoplasmic reticulum to which ribosomes are attached, and the smooth endoplasmic reticulum, which consists of a network of ribosome-free tubules and is the site of exit of secretory cargo vesicles. The ER can also be defined functionally as the site of lipid metabolism, carbohydrate metabolism, drug detoxification, and the entry point for proteins into the secretory pathway. Proteins that are secreted from cells and proteins which are destined for cellular membranes, are synthesized on the ribosomes of the rough endoplasmic reticulum and translocated into the lumen of the ER, where they are glycosylated, become folded, disulfide bonds are formed, and some multiprotein complexes are assembled. Before proteins are transported out of the ER they are subjected to a 'Quality Control' process which assesses their state of folding and then, if they are correctly folded, allows them to exit the ER by packaging into transport vesicles. Proteins that are not folded correctly can have one of several fates, they may be retained in the ER for further cycles of the quality control process, or they are retrotranslocated out of the ER and degraded in the cytosol by the proteasomal system, or they can accumulate in the ER as aggregates.

Many of these processes in the ER are controlled by proteins, which are collectively termed 'molecular chaperones'. Christian Anfinsen won the Nobel Prize in 1972 for his demonstration a decade earlier that the correct folding of ribonuclease A is determined solely by the primary sequence of the protein. That is, the purified enzyme can be denatured and then can refold *in vitro* to restore enzymatically active ribonuclease. However, he also showed subsequently that the rate of folding could be increased closer to physiological rates by the addition of the ER enzyme protein disulfide isomerase (PDI). Hence arose the idea of one protein assisting another to attain its correct folding. This concept was further refined by others and the term "molecular chaperone" was used by Lasky and coworkers to describe the activity of a protein involved in the assembly of nucleoplasmin (1). The current working definition of molecular chaperones is that they can recognize unfolded proteins and promote protein folding by preventing irreversible

aggregation with other proteins (2). They associate with a protein during the folding process but during folding or when folding is complete, the chaperones disassociate and participate in the folding of new proteins. Another implication of the recognition of unfolded proteins by molecular chaperones is that they are indiscriminate and can interact with any unfolded protein and therefore any specialization arises from other mechanisms for example cellular location. The results of studies of molecular chaperone systems in the cytosol show that to accomplish their functions there is a requirement for different types of molecular chaperone. In the ER, in addition to a variety of molecular chaperones, there are also folding enzymes such as protein disulfide isomerases (PDI) and protein prolyl isomerases (PPI), and the lectin-like chaperones calnexin and calreticulin.

Reviewed here are some selected functions of the ER, the Calnexin Cycle and protein quality control, retrotranslocation of proteins out of the ER, and signalling to the nucleus of the presence of unfolded proteins in the ER. Figure 1 provides a flow chart and depicts the structural and dynamic protein interactions discussed.

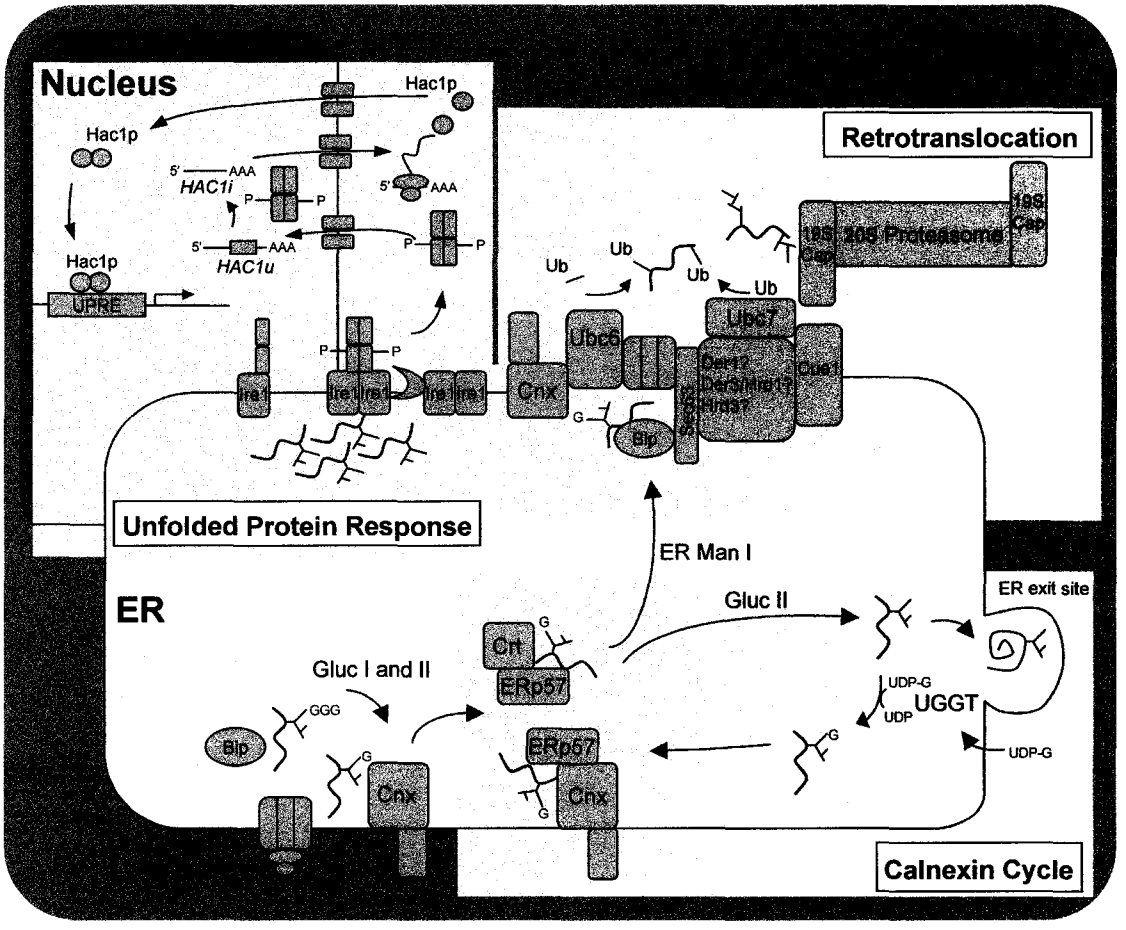
1.2 The Calnexin Cycle

1.2.1 Discovery of calnexin, and its characterization as an ER lectin-like chaperone

Calnexin provides a link between N-glycosylation and protein folding. The oligosaccharide $Glc_3Man_9GlcNAc_2$ is synthesized stepwise on an isoprenoid lipid dolichol-pyrophosphate and is transferred by oligosaccharyl transferase to nascent polypeptide chains on the asparagine residue in the sequence motif Asn-X-Ser/Thr. The glucose residues are then sequentially removed by ER glucosidases I and II (Gluc I and Gluc II in Figure 1) (3). The seemingly redundant process of glucose addition and removal is explained by the recognition of the $Glc_1Man_9GlcNAc_2$ glycoform of glycoproteins by calnexin and also by its ER lumenal homologue calreticulin. Some of the key components of the Calnexin Cycle have been identified and a model of how it functions is emerging. But there remain several questions concerning its mechanism.

Since the cDNA cloning in 1991, the study of calnexin has been a remarkably intensive area of endeavor. In an investigation of ER membrane proteins that are

Figure 1 Models for three biological processes that constitute the 'ER quality control apparatus'. N-linked glycosylation of nascent secretory or membrane proteins with a $Glc_3Man_9GlcNAc_2$ oligosaccharide occurs during co-translational translocation into the endoplasmic reticulum through the translocon channel. This subjects glycoproteins to a recently characterized and highly effective quality control apparatus. This 'molecular machine' prevents the progression of proteins that have not yet attained or cannot attain native conformation, through the secretory pathway. The interaction with the quality control apparatus is initiated by the trimming of the oligosaccharide to the $Glc_1Man_9GlcNAc_2$ form (through the action of the glucosidases I and II) which permits calnexin/calreticulin binding and thus entry into the Calnexin Cycle (Lower-right shaded panel). Prolonged and repeated interactions with calnexin/calreticulin permits the slow acting ER mannosidase I and II to hydrolyse the glycan to the Glc_1Man_8 form, preventing its release from calnexin/calreticulin by glucosidase II. Calnexin is then thought to bring the misfolded proteins back to the translocon for retrotranslocation, polyubiquitination by ubiquitin conjugating enzymes (UBCs), and degradation by the 26S proteasome (Upper-right shaded panel). If unfolded proteins accumulate in the ER, a signal is transmitted from the ER to the nucleus which initiates the transcription of ER chaperones that are under the transcriptional control of an unfolded protein response element (UPRE) (Upper-left shaded panel).


Nucleus
Hac1p
Hac1p
5' AAA
HAC1i
P P
5' AAA
5' AAA
HAC1u
P P
Hac1p
UPRE
Retrotranslocation
Ub
Ub
Ub
Ub
19S Cap
20S Proteasome
19S Cap
Ubc7
Ubc6
Cue1
Der1?
Derlin-1?
Hrd3?
Cnx
G
Bip
Ire1
Ire1
Ire1
Ire1
P P
Unfolded Protein Response
ER Man I
ER
ER exit site
Gluc II
Gluc I and II
Crt
G
ERp57
UDP-G
UDP
UGGT
GGG
Bip
G
Cnx
ERp57
G
Cnx
G
UDP-G
Calnexin Cycle

phosphorylated by endogenous kinases (4), Ikuo Wada isolated a complex of four proteins, one of which had similarity with an ER lumenal protein termed calreticulin. This had been originally identified as a high affinity calcium binding protein of the sarcoplasmic reticulum (5). This membrane protein shared some properties such as calcium binding with calreticulin, and the name calnexin was suggested by Professor A. Schachter from the History Department of McGill University. Studies on the kinetics of the assembly of MHC I β microglobulin in the ER had identified a protein of 88kDa which associates with unassembled MHC I heavy chain but is itself retained in the ER (6). Upon analysis this 88kDa protein proved to be calnexin and provided the first indication that it might function as a molecular chaperone (7). This was also tested in mammalian cells after the development of antibodies, which could immunoprecipitate calnexin. These experiments (8) used HepG2 cells,which have a well characterized set of secreted proteins. Pulse-chase experiments showed secretory glycoproteins with calnexin, and that during the chase they dissociated and continued through the secretory pathway. However, it was clear that not all proteins interact with calnexin, for example, human serum albumin, an abundant secretory non-glycoprotein in HepG2 cells does not interact with calnexin whereas the sequence related glycoprotein α-fetoprotein, which is also expressed in HepG2 cells, does bind. This suggested that calnexin binds only glycoproteins and this was confirmed with HepG2 cells labelled in the presence of the glycosylation inhibitor tunicamycin where no glycoproteins bind with calnexin. It was also established that calnexin recognizes glycoproteins from studies of the effects of the glucosidase II inhibitor bromoconduritol and the retention of glycoproteins in treated human hepatic cells, and a prescient suggestion was made that there may be a receptor in the ER which recognizes $Glc_1Man_9GlcNAc_2$ and prevents secretion of glycoproteins (9). *In vitro*, calnexin was shown to have a preference and weak affinity for free $Glc_1Man_9GlcNAc_2$ oligosaccharides (10).

However, the original experiments of Ou *et al.* (8) also used a standard experimental method of unfolding proteins *in vivo* by treating cells with a proline analogue L-azetidine-2-carboxylic acid (AZC) which is incorporated into proteins and prevents them from folding correctly. Analysis of secretory proteins in cells with AZC showed that they were retained for prolonged periods and thus calnexin was also assumed

to recognize unfolded proteins. On the basis of the results of *in vivo* experiments from many laboratories, there has been controversy on whether calnexin acts as a lectin or whether it can recognize unfolded proteins.

1.2.2 The lectin/true molecular chaperone debate

In vitro studies of the association of calnexin and calreticulin with defined substrates offer a way of answering the question of whether they recognize the oligosaccharide moiety, the protein, or both. Purified and well-characterized components for the *in vitro* systems are available. The form of calnexin used in these studies is the ER lumenal portion that has been shown to have properties similar to full length calnexin in that it interacts identically with HIV-1 gp120 in a reconstructed calnexin interaction system in insect cells, and *in vitro* in terms of calcium binding, protease resistance and ATP binding it behaves like the intact calnexin molecule (11,12). In addition, the lumenal domain of calnexin recognizes free $Glc_1Man_9GlcNAc_2$ oligosaccharide and not the $Glc_2Man_9GlcNAc_2$ or $Glc_3Man_9GlcNAc_2$ glycoforms (13). There is also additional support for the assumption that the lumenal domain of calnexin contains all the necessary elements for the recognition of glycoproteins from the experiments of Ikuo Wada who has shown that calreticulin, the closely related lumenal ER protein, can be attached to the transmembrane domain and cytosolic tail of calnexin and now binds some 'calnexin preferred' glycoproteins (14). Thus, by these criteria, the lumenal domain of calnexin used in these *in vitro* experiments has all the necessary elements for functionally binding glycoproteins.

The availability of suitable glycoprotein substrates is critical for the *in vitro* experiments. Most secreted glycoproteins have had their high mannose oligosaccharide trimmed and their outer chains extensively modified by enzymes during transit through the Golgi. There are some well-documented cases where a minor portion of a glycoprotein carries a high mannose oligosaccharide such as thyroglobulin and RNaseB. But recent results from studies on the glycosylation process in yeast and the knowledge that some of the genes responsible are not essential for growth in this organism, makes possible the engineering of glycoproteins with specific glycoforms. Thus, yeast strains have been constructed which secrete glycoproteins with uniform high mannose

oligosaccharides of various forms (15). For the experiments described below, preparations of RNaseB that were enriched in the $Man_9GlcNAc_2$ were used. The *in vitro* experiments took advantage of the considerable amount of information, which has accumulated on the folding of this protein and its non-glycosylated form RNaseA since the pioneering work of Anfinsen. To construct the $Glc_1Man_9GlcNAc_2$ glycoform, the ER enzyme UDP-glucose:glycoprotein glucosyltransferase (UGGT) was used with proteins bearing Man_9 N-linked oligosaccharides. This enzyme was first described by Armando Parodi and his colleagues as an activity that can transfer a glucose moiety from UDP-glucose to a $Man_9GlcNAc_2$ acceptor on an unfolded glycoprotein (15,16). Thus UGGT has a unique role in the cellular glycosylation process in that it is reversible and points to a potential role in the quality control of glycoproteins in that it is specific for unfolded proteins carrying $Man_9GlcNAc_2$ glycans. The experimental design was to glucosylate unfolded RNaseB with UGGT and UDP-glucose which is [^{3}H]-labelled in the glucose moiety and then produce the unfolded, the fully folded, and the partially folded intermediates by using the established refolding protocols for this enzyme. Measurements of binding to calnexin showed that fully folded, unfolded and partially folded forms of RNaseB all bind calnexin (17). Thus calnexin (and calreticulin) are defined by these *in vitro* experiments as lectins. The provisos of this seemingly definitive experiment are that although the affinity constants of calnexin for each of the folding forms of $Glc_1Man_9GlcNAc_2$-RNaseB were not individually measured, it can be inferred from the experimental design that they must differ less than an order of magnitude. In apparent contradiction, *in vivo* experiments (8,11) have shown that the form of proteins, in these cases transferrin and HIV gp120 that are bound to calnexin are unfolded. Confirming these results of Zapun *et al.* (17), Rodan *et al.* (18) showed that in an *in vitro* microsomal system that RNaseB which had been mutated to have two N-linked glycosylation sites bound to calnexin independently of the folding status of the enzyme. Thus the experimental results confirm that for RNaseB, a classical protein folding substrate, the sole determinant of calnexin binding is the correct oligosaccharide moiety Glc_1Man_9

There are some aspects of this result that are worth considering. One of the advantages of using RNaseB in protein folding studies is that it does not have extensive hydrophobic regions and does not aggregate when denatured, and perhaps studies using

proteins with more extensive hydrophobic regions will reveal other functions of calnexin. Recently, support for this view has been provided by the demonstration that calnexin can function as a molecular chaperone with classical substrates used in protein folding studies which are not glycoproteins. This surprising result comes from a series of experiments performed by David Williams and colleagues (19).

In these experiments calnexin, when added in stoichiometric amounts, protected the glycoprotein soybean agglutinin (SBA), and the nonglycosylated proteins citrate synthase (CS) and malate dehydrogenase (MDH) against thermal aggregation as measured by changes in light scattering. They examined the role of the oligosaccharide by assaying various glycoforms of SBA. Interestingly, they found that the monoglucosylated form ($Glc_1Man_9GlcNAc_2$) was best protected from thermal aggregation by calnexin while it required eight fold more calnexin to equivalently protect the $Man_9GlcNAc_2$ form, and twenty fold more for SBA with only one GlcNAc residue. However, definitive evidence for interaction with the polypeptide portions of the unfolded substrates came from experiments with non-glycosylated proteins CS and MDH. Not only did calnexin protect these unfolded proteins from aggregation, it maintained CS in a folding competent state which allowed for the reactivation of its enzymatic activity. The property of maintaining a protein in a folding competent state had previously been shown for the chaperones Hsp90, Hsp70 and the small heat shock protein chaperones (20-22). These results suggest that calnexin may also act as a 'bona fide' molecular chaperone. Moreover, calnexin associated only with the unfolded form of MDH to produce a complex that was hypersensitive to trypsin, which is most effective on unfolded proteins, while no association could be detected with enzymatically active fully folded MDH. Thus calnexin can discriminate between folded and non-folded substrates, satisfying a criterion of a molecular chaperone. The conclusion from these *in vitro* experiments, which were also found for calreticulin by Saito *et al.* (23), suggest that calnexin may interact with the protein portion of its substrate in addition to the glycan. There is some caution needed in interpreting these results as *in vivo* the proteins that have been found complexed with calnexin are always glycoproteins. Based on these results, the model predicts that the oligosaccharide is the ligand that initiates and stabilizes the subsequent polypeptide interactions with calnexin. Thus calnexin can exhibit the

properties of a lectin and may also act as a molecular chaperone.

1.2.3 The calnexin/ERp57 protein folding complex

What happens to glycoproteins while they are bound to calnexin is not clear but there are some clues to possible mechanisms. Glycoproteins bound to calnexin and calreticulin interact with a member of the protein disulfide isomerase family termed ERp57 (24). It was found using *in vitro* crosslinking, that as well as calnexin and calreticulin, ERp57 can be crosslinked in dog pancreatic microsomes to a glycoprotein precursor N-glycosylated with the $Glc_1Man_9GlcNAc_2$ form of oligosaccharide (24). An interpretation of this data is that ERp57 is itself either a lectin or that it associates with calnexin and calreticulin, which are themselves bound to glycoproteins. *In vitro* studies established that there was no enhancement of the rate of folding of monoglucosylated ($Glc_1Man_9GlcNAc_2$) RNaseB by ERp57 alone (25). Based on the crosslinking results, and the similarity of ERp57 with PDI a speculation was that it might be a disulfide isomerase for protein substrates bound to calnexin. Using an *in vitro* system that was developed to determine the functions of calnexin and calreticulin, using recombinant UGGT (15) and [^{3}H]UDP-glucose to produce labelled monoglucosylated RNaseB, it was shown that ERp57 does not recognize this oligosaccharide directly. However, the protein disulfide isomerase activity of ERp57 on the refolding of monoglucosylated RNaseB is greatly enhanced in the presence of calnexin, in contrast to the result with PDI whose activity on monoglucosylated RNaseB is decreased by the presence of calnexin (25). Calnexin had no effect on the refolding of non-monoglucosylated RNaseB catalyzed by either PDI or ERp57. Thus, it was concluded that ERp57 is a protein disulfide isomerase dedicated to glycoproteins that are bound to calnexin and calreticulin. This result provides an interesting new insight into calnexin function and we can speculate that other PDI family members such as ERp72 and PDI itself may each have their ER molecular chaperone partners.

1.2.4 Quality control of protein folding for N-linked glycoproteins in the ER

Quality control within the Calnexin Cycle takes place upon the hydrolysis of the last glucose on the $Glc_1Man_9GlcNAc_2$ oligosaccharides of glycoprotein substrates by glucosidase II, which prevents binding to calnexin. The process in itself is defined by the ability of the quality control apparatus to recognize and selectively bind proteins that are misfolded or have not yet attained their native conformation, and prevent their premature progression though the secretory pathway. Attaining native conformation or proper subunit assembly for protein complexes within the lumen of the ER has been shown in a number of cases to require cyclical interactions with both calnexin and calreticulin.

The protein folding sensor of the Calnexin Cycle is the ER enzyme UDP-glucose:glycoprotein glucosyltransferase (UGGT), which acts only on unfolded glycoproteins, reglucosylating them allowing for further binding to calnexin and calreticulin. Demonstrating this cyclical process under physiological conditions was achieved by using novel and innovative experimental approaches by both Wada *et al.* (26) and Cannon and Helenius (27).

1.2.4.1 The role of de/reglucosylation in quality control and protein folding

The importance of deglucosylation and reglucosylation in the quality control of protein folding for a glycoprotein has been demonstrated in a series of ingenious experiments (28). As a glycoprotein substrate, the iron binding protein transferrin, which has two homologous domains, nineteen pairs of cysteines that form disulfide bridges and two N-linked glycans was selected. HepG2 cells were pulse labelled with [^{35}S]-methionine, and microsomes were prepared on a discontinuous sucrose gradient. The process of transferrin folding was followed by non-reducing SDS-PAGE. In this gel system, reduced unfolded transferrin migrates slowly as a diffuse band, while fully folded transferrin moves more rapidly as a sharper band. Raising the temperature of the cell preparations to 43°C resulted in an increase in high molecular weight disulfide-linked transferrin aggregates at the top of the gel, and they could control transferrin misfolding. After labeling microsomes with [^{3}H]UDP-glucose at 37°C, followed by an unlabelled UDP-glucose chase and the association of transferrin with both calnexin and calreticulin was determined by co-immunoprecipitation. The half-lives of glucose removal from the

calnexin and calreticulin complexes were similar at 5.8 minutes and 4.7 minutes, respectively. Alternatively, by examining the kinetics of substrate dissociation from calnexin and calreticulin using [^{35}S]-methionine labeled microsomes (which labels the polypeptide moiety of the glycoprotein), it was shown that the addition UDP-glucose prolonged their association with transferrin. This result demonstrated that transferrin undergoes cyclical interactions with these lectin-like chaperones. It was also noticed that at increased temperature the addition of UDP-glucose suppressed the formation of disulfide-linked aggregates and promoted transferrin folding. Thus at 43°C, folding intermediates and folded transferrin were only faintly detectable in the absence of UDP-glucose. The significant contribution to folding efficiencies by the addition of UDP-glucose at 43°C, demonstrated that, not only are cyclical interactions occurring under these physiological conditions, but that the actions of de- and reglucosylation of the ER enzymes glucosidase II and UGGT promote multiple interactions with calnexin and calreticulin, a process that both prevents protein misfolding and facilitates the attainment of native protein conformation.

While work by Wada *et al.* (28) and Zapun *et al.* (17) also gave insight on the role of glucose trimming of N-linked glycans, and its affect on mediating cyclical interactions with the lectin-like chaperones calnexin and calreticulin, they used either a reconstituted system involving microsomal preparations or a completely *in vitro* system. Cannon and Helenius (27) developed a system that used a mutant form of glycoprotein from the ts045 vesicular stomatitis virus (VSV). The mutant VSV G protein has a thermoreversible folding defect that can be manipulated *in vivo* by shifting the temperature from permissive (30°C) to nonpermissive temperatures (40°C). VSV G protein is a type I transmembrane glycoprotein with two N-linked glycans, and while the mutant form folds rapidly and moves to the Golgi apparatus at 30°C, at 40°C it fails to acquire a complete set disulfide bonds and is retained in the ER. In these experiments, VSV G folding was monitored by immunoprecipitation with conformation specific antisera.

To examine the effects of reglucosylation on quality control, cells were infected with ts045 VSV, pulsed labeled, and then chased at permissive temperature to allow G protein to fold correctly and be released from calnexin. Upon a shift to nonpermissive temperature, VSV G protein unfolded and rebound to calnexin. The rates of unfolding,

reglucosylation, and rebinding to calnexin at nonpermissive temperature were examined. Similar kinetics were observed for VSV G unfolding and reglucosylation by UGGT, while, the rebinding to calnexin was bimodal, with rapid interaction phase followed by a slower second phase. This could be explained by the intracellular distribution of VSV G. Under permissive conditions, VSV G exhibited a faintly reticular distribution with punctate spots that co-localized with UGGT at putative ER exit sites. Alternatively, when under conditions of shift to non-permissive temperature, the ER staining was more prominent with a diminished spotty pattern. Hence, as VSV G protein changed from a folded to unfolded state, it returned from the ER exit sites. Cannon and Helenius (27) showed *in vivo* that glycoproteins are redirected back for further binding to the lectin-like chaperones under conditions that induce protein unfolding. It was also noted that the putative ER exit sites were devoid of calnexin, and thus, they ascribed the quality control function of the Calnexin Cycle to a late ER-exit site location. Moreover, both Wada *et al.* (28) and Cannon and Helenius (27) demonstrated the importance of the $Glc_1Man_9GlcNAc_2$ form of the oligosaccharide in mediating calnexin-substrate interactions, suggesting that calnexin acts primarily as a lectin.

1.3 Retrotranslocation of proteins from the ER

1.3.1 ER-associated degradation and genetic disease

There are components of the Calnexin Cycle that can discriminate between folded and unfolded proteins. There also exists an associated quality control system that allows correctly folded proteins to be exported to subsequent cellular compartments. Misfolded secretory proteins, however, which are usually the consequence of mutations or their over-expression, can have several fates. The α1PiZ allele of α1-antitrypsin is the underlying genetic defect of the most common genetic liver disease in children and results in a protein that can accumulate together with other ER resident proteins within distinct regions of the ER (29). This accumulation in the form of loop-sheet polymers may be hepatotoxic causing liver cirrhosis, while the ensuing serum depletion of α1-antitrypsin can lead to emphysema, in which alveolar destruction occurs as the result of the inability to block neutrophil elastase activity in the lungs (29). In the case of Cystic Fibrosis, the most common mutation (CFTRΔF508) in the gene that encodes the cystic

fibrosis transmembrane conductance regulator (CFTR), which functions as a plasma membrane chloride channel, leads to a prolonged association with calnexin and subsequent degradation of almost all the mutant CFTR protein. The maturation of wild-type CFTR is also relatively inefficient with 75% of the wild-type form being degraded, versus 99% of the mutant form. However, the mutant protein has been shown to form an active channel both *in vitro* and in cells cultured under conditions of reduced temperatures or overexpression, suggesting that the pathology of the disease may be based largely on trafficking (30-33). Many other protein trafficking diseases such as familial hypercholesterolemia, Tay-Sachs, Congenital Sucrase-Isomaltase Deficiency have similar pathologies resulting in the quality control of N-linked glycoproteins. In these cases the corresponding proteins are unable to fold or assemble correctly and become degraded by the ER-associated degradation pathway (ERAD).

1.3.2 Where does ER-associated degradation occur, in the ER or cytosol?

The degradation pathway used to eliminate folding incompetent proteins and alleviate the cellular stresses incurred by their accumulation was originally thought to reside in the ER (34). But as suggested by Ron Kopito in a later review: "The abundance of unfolded and partially folded polypeptide chains that are highly susceptible to proteolysis is difficult to reconcile with the presence in the same compartment of an aggressive proteolytic apparatus" (35). Experimental evidence of a cytosolically located proteolytic apparatus became apparent with the work of Sommer and Jentsch (36) who quite serendipitously made the first link between the degradation of an ER membrane protein and the ubiquitin-proteasome degradation pathway.

Sommer and Jentsch characterized a novel integral membrane ubiquitin-conjugating enzyme (UBC6p) from *Saccharomyces cerevisiae* that was localized on the ER membrane, with its catalytic domain in a cytosolic orientation. Ubiquitin-conjugating enzymes (UBCs) target proteins for degradation by the 26S proteasome by covalently attaching the 76 residue ubiquitin protein to lysine residues on substrate proteins (37). The initial ubiquitin molecule itself becomes ubiquitinated and this process continues leading to multiple ubiquitin chains that form a degradation signal. The 26S proteasome is composed of the 20S core catalytic complex which forms a multisubunit cylindrical

structure with multiple proteolytic activities that include tryptic-, chymotryptic- and post-glutamyl peptidyl hydrolytic-like specificities. The cylinder is capped by two 19S regulatory complexes at each end that are thought to mediate substrate recognition through the interaction of specific subunits with the polyubiquitin chain.

Based on the localization of the UBC6p protein, Sommer and Jentsch sought synthetic lethal phenotypes of UBC mutants with mutants of the secretory pathway. Surprisingly, they found that the *ubc6Δ* mutant suppressed the sec61 mutant phenotype. *SEC61* encodes a multispanning ER membrane subunit of heterotrimeric complex that also includes Sbh1p and Sss1p contributing to the translocon channel (38). In the *sec61 ubc6Δ* strain both translocation and glycosylation levels of the secretory proteins carboxypeptidase Y (CPY) and α-factor were restored to near wild-type levels. The sec61 defect was thus shown to require active UBC6p. Since the mutant sec61p was found to be stable in *sec61 ubc6Δ* cells, it was concluded that the substrates for UBC6p-mediated degradation were membrane proteins including proteins that interacted directly with the Sec61p. These experiments provided the first conclusive evidence of a cytosolic degradation pathway associated with ER, at least for membrane proteins.

1.3.3 ERAD of membrane and soluble ER protein

1.3.3.1 The proteasome degrades ER membrane proteins

A large number of reports from 1995 though 1996, showed that ER protein degradation used the ubiquitin-proteasome pathway in the cytosol (39-43). Two illuminating papers published together in the same edition of *Cell* (39,41) made significant contributions to our understanding of the pathology of Cystic Fibrosis by determining the protein degradation pathway that affects CFTR.

Jensen *et al.* (39) characterized the rapid turnover of the CFTR precursor and its maturation using peptide aldehyde protease inhibitors N-acetyl-L-leucinyl-L-leucinyl-L-norleucinal (ALLN), N-acetyl-L-leucinyl-L-leucinyl-L-methionyl (ALLM), and N-carbobenzoxyl-L-leucinyl-L-leucinyl-L-leucinyl (MG-132). ALLN was more effective at blocking the rapid degradation than ALLM, which agreed with their relative abilities to block protein degradation by purified proteasome preparations (44). Moreover, MG-132 completely blocked the ATP-dependent conversion of the precursor CFTR to the native

folded form capable of transport from the ER, a process that was not affected by the other proteasome inhibitors suggesting that there is a second proteolytic pathway. Lactacystin, which is highly specific for the multiple activities of the proteasome (45) was the most effective in blocking turnover, and demonstrated that the proteasome is responsible for the degradation of the CFTR precursor.

While Ward *et al.* (41) demonstrated the inhibitory effects of ALLN and lactacystin, they also noted the accumulation of polyubiquitinated CFTR. Upon examination of immunoblots of insoluble CFTR, they noticed a ladder of bands spaced at about 6.5 kDa, which is diagnostic of polyubiquitination. Several methods were used to characterize the polyubiquitinated CFTR. Firstly, they immunoprecipitated ΔF508 CFTR, ran a gel with the immunoprecipitate, and then immunoblotted with anti-ubiquitin. They only found ubiquitin immunoreactivity in the precipitations from cells expressing CFTR, and with pronounced accumulation of polyubiquitinated species in the presence of lactacystin. They purified polyubiquitinated CFTR by Ni^{2+} chelation chromatography of cells expressing a His_6 tagged form of ubiquitin. Ubiquitinated CFTR was observed in the eluate of columns loaded with either soluble or insoluble material from cells expressing both ΔF508 CFTR and His_6M-Ub. They also showed that the expression of a dominant negative ubiquitin (Ub_{K48R}), which prevents polyubiquitin chain formation on the Lys-48, attenuated the ubiquitination of CFTR. Together, these experiments demonstrated the polyubiquitination of CFTR *in vivo*, and are in agreement with the result of Jensen *et al.* (39) on its degradation by the ubiquitin-linked pathway.

1.3.3.2 Soluble lumenal proteins are also degraded by a cytosolic degradation pathway

Many membrane proteins of different cellular compartments share the same topology and their degradation appears to follow a common pathway. Non-membrane proteins of the ER (lumenal proteins) are compartmentalized such that they are apparently inaccessible to the 26S proteasomal system. The demonstration that lumenal and soluble secretory proteins follow the same degradation pathway as membrane proteins comes from experiments done *in vitro* and *in vivo* using both yeast and mammalian cells (40,42,43).

Werner *et al.* (43) examined the ER associated degradation of the mutant form of α1-antitrypsin (α1-ATZ) heterologously expressed in S. cerevisiae mutants. When expressed in strains with a proteasomal mutant genetic background (*pre1-1 pre2-2*) deficient in chymotrypsin-like activity, the intracellular half-life of α1-ATZ increased 3 fold from 56 to 163 minutes compared with the wild type. McCracken and Brodsky (46) developed an *in vitro* assay with purified ER (microsomes) and the mating pheromone precursor of S. cerevisiae (prepro-α factor, ppαF). When ppαF was added to these microsomal preparations, it was post-translationally translocated into the ER lumen and its location was verified by resistance to proteinase K digestion. A mutant form that had its N-linked glycosylation sites mutated (ΔGppαF) and was known to be subjected to ERAD was also translocated into the ER lumen, however, it was subsequently retrotranslocated out of the microsomes. It should be noted that in that same study, McCracken and Brodsky (46) found that degradation required the addition of cytosol, ATP, and the presence of calnexin. Using a similar system, Werner *et al.* (43) found that when cytosol of proteasomal mutant *pre1-1 pre2-2* was added, the degradation of ΔGppαF was significantly decreased. The role of the proteasome in degradation was confirmed with the proteasome specific inhibitors, decreasing degradation at least two fold.

Genetic screens using endogenous yeast misfolded proteins have been used to identify other components of the ERAD pathway (42). Hiller *et al.* (42) used a mutant form of soluble vacuolar protein carboxypeptidase Y (CPY*) that fails to reach the vacuole and is retained in the ER, and then degraded. In a screen for mutants defective in the ERAD of intracellular CPY*, they isolated the *der2-1* mutant which was found to be identical to a known gene *UBC7*. *UBC7* encodes an ubiquitin-conjugating enzyme responsible for the ATP dependent addition of ubiquitin from an intermediate to a target protein. This prompted an investigation of other genes in the ubiquitination pathway and they demonstrated the role of Ubc6p, Ubc7p, polyubiquitination, and a functional proteasome in the degradation of CPY*.

1.3.4 Membrane and soluble secretory proteins must cross the ER membrane barrier

Although the role of the ubiquitin-proteasomal degradation pathway in ERAD has been established, there remains a question of how misfolded proteins, selected by the ER quality control apparatus, traverse the ER membrane? One possibility was that the Sec61 translocon channel may work in both translocation into the ER for unfolded glycoproteins and that they can be retrotranslocated through the same channel. This was demonstrated with surprising results from the study of two human cytomegalovirus (HCMV) proteins that participated in an intriguing host immunity response evasion strategy (47,48). Expression of the HCMV genes *US2* or *US11* in cells lead to the rapid degradation ($t_{1/2}$ < 1 min) of newly synthesized class I heavy chains (MHCs) in a lactacystin dependent manner. That is the MHCs were dislocated to the cytosol where deglycosylation occurred by a cytosolic N-glycanase, followed by proteasomal degradation (47). Breakdown intermediates found associated with the Sec61 complex suggested that retrograde transport occurred through the same protein conducting channel that allowed the original insertion of the MHCs. This provided the first evidence that the Sec61 translocon channel participates in both translocation of proteins into the ER and their retrotranslocation into the cytosol.

Genetic studies in yeast with CPY* confirm the role of the Sec61 translocon as the channel for retrotranslocation of proteins from the ER and also a role of BiP and its co-chaperone Sec63p (49). CPY* was stabilized in the three mutants of *SEC61*, *KAR2*, and *SEC63*, and different mutations were found that blocked either the import function of the Sec61 translocon channel, or the export of misfolded proteins to the cytosol. For example, the allele *sec61-2* had a two fold stabilization effect on ERAD substrates which was the result of blocked export from the ER, while a mutation in *SEC62* (which encodes a subunit of the tetrameric complex including Sec62p, Sec63p, Sec71p and Sec72p associated with the Sec61 complex) led to an import defect. The *sec61-2* allele which encodes Sec61p, only expressed a mutant phenotype with respect to retrotranslocation. Another example is the *kar2-113* allele (BiP) which is fully functional in CPY* import at 25°C. Expression of this mutation led to a two fold increase in the CPY* half-life under these same conditions again suggesting that the processes of import and export from the

ER defined by BiP activity can be functionally independent.

1.3.5 Some possible mechanisms for retrotranslocation through the Sec61 translocon channel

While the experiments by Plemper *et al.* (49) provide some interesting possibilities for the mechanisms of retrotranslocation of ERAD substrates involving the Sec61 translocon channel and BiP, a recent paper (50) reported that the import and export from the ER are mechanistically distinct. Pulse-chase experiments were performed *in vivo* with two alleles of *KAR2* (*kar2-113* and *kar2-133*). These showed that these alleles are functional for import but incapable in the *in vitro* retrotranslocation assay (Section 1.3.3.2) of export from the ER, using ΔGppαF and α1-ATZ. Conversely, export could be demonstrated in the presence of cytosol from ssa1 mutants, which are defective for import into the ER and ΔGppαF as a substrate. Thus, there is additional support for the mechanistically distinct processes using some common components.

It was also noted by Plemper *et al.* (49) that with the pore size of translocon channel of approximately 20Å, CPY* unfolding would be required, be it partially or fully, prior to retrotranslocation, also suggesting a role for BiP in retrotranslocation. Interestingly, Tsai *et al.* (51) demonstrated that PDI acts as a redox driven chaperone, which unfolds cholera toxin, before a fragment of it is translocated to the cytosol. Thus PDI is a likely candidate for protein unfolding prior to retrotranslocation.

While BiP has been shown in yeast to required for import and may act as a ratchet or motor (52), the driving force mediating export from the ER has not yet been elucidated. An interesting development emerged when an ER integral membrane protein termed Cue1p (factor for coupling of ubiquitin conjugation to ER degradation) was isolated (53). Cue1p protein was found to recruit Ubc7p to the ER membrane, and to be essential for its function. Ubiquitination by both Cue1p-assembled Ubc7p and Ubc6p is necessary for retrograde transport, demonstrating that ubiquitination and export are mechanistically coupled.

Proteasomal activity has been implicated as a driving force for retrotranslocation (54). The evidence came from experiments which used fusion proteins consisting of an N-terminal degradation signal with the doubly membrane spanning Sec62p. In cells with

proteasomal mutations, it was demonstrated that the rapid proteolytic activity of the proteasome on the N-terminal cytosolic domain was followed by slowed proteolysis at the point when membrane extraction is necessary. Retrotranslocation of Sec62p from the membrane, and proteasomal activity on transmembrane ERAD substrates appeared to be coupled.

1.3.6 Proteins with unknown function

Other proteins involved in selective retrotranslocation have been found primarily through genetic screens in *S. cerevisiae*, that used the accumulation of the ERAD substrates such as 3-hydroxy-3-methylglutaryl-CoA reductase (HGM-R), CPY*, and α1-ATZ. All three screens looked for mutants that are inefficient in degrading these substrates. HMG-R is an integral membrane protein of the ER and an essential enzyme involved in sterol synthesis. Hampton *et al.* (54,55) expressed a Myc-tagged version of HMR-G (6myc-Hmg2p) and selected for cells with increased steady state levels of the protein on lovastatin, an inhibitor of HMG-R that induces its more rapid degradation. They isolated three mutant alleles designated *hrd1* to *hrd3*. The *HRD1* (also known as DER3, see below) and *HRD3* genes, which encode proteins predicted to be membrane bound, are also involved in ERAD of CPY* (56). While the function of these proteins is still unclear, HRD2 is homologous to p97 (TRAP-2) and encodes a component of the mature 26S proteasome.

In similarly designed experiments (55), three mutant alleles were found that result in the accumulation of CPY*, termed *der1* to *der3*. *DER1* encodes a membrane protein of still unknown function (57), *DER2* was found to be identical to *UBC7* (42), and *DER3* is identical to *HRD1*.

1.3.7 The mannosidase clock: a model for the selective process associated with ERAD

It is still unclear how proteins are selected for ERAD. The Calnexin Cycle through de/reglucosylation of immature glycoproteins ensures that unfolded proteins do not progress along the secretory pathway. The selective mechanism by which misfolded proteins are brought to the translocon still has not been determined. Experiments in *S.*

cerevisiae and mammalian cells have examined how N-linked oligosaccharide structure can influence ERAD. Slight differences exist between the two systems, however, a common denominator does arise. Both Knop *et al.* (57) and Jakob *et al.* (58) examined the turnover of CPY* in yeast. Knop showed that cells devoid of the ER-processing α-1,2-mannosidase displayed reduced degradation of CPY*, while Jakob demonstrated that the glycoproteins with N-linked $Man_8GlcNAc_2$ oligosaccharides were most effectively degraded when compared to glycoproteins with $Man_9GlcNAc_2$, $Man_7GlcNAc_2$, $Man_6GlcNAc_2$ glycan structures.

It was also shown recently that ER mannosidase I activity was required for ERAD in mammalian cells using truncated α1-antitrypsin (59). ER associated degradation of misfolded α1-antitrypsin was found to be dependent on calnexin binding. Moreover, ER mannosidase trimming resulted in prolonged association with calnexin, a prerequisite for ERAD to occur. This was to be expected since glucosidase II, which mediates release from calnexin by trimming the last glucose from $Glc_1Man_9GlcNAc_2$, exhibits much slower kinetics with $Glc_1Man_8GlcNAc_2$ (60).

Therefore, the model suggests that the ER mannosidase I induces prolonged association of misfolded glycoproteins with calnexin, which then directs them to the translocon and subsequently to the proteasome for degradation. The model itself is stochastic, whereby the rates of mannose trimming (on the middle branch of the N-linked oligosaccharides) on misfolded glycoproteins, combined with the residency time within the Calnexin Cycle, influence their interactions with the ER chaperones determining the selective process underlying ERAD.

Alternatively, it was shown by Cabral *et al.* (61) that the Pi Z variant of α1-antitrypsin, which polymerizes in the ER, undergoes additional oligosaccharide trimming by mannosidase II. This partitions the Pi Z variant away from the proteosome-mediated disposal mechanism, due to an arrested posttranslational interaction with calnexin (61). Moreover, Fagioli and Sitia (62) show that the disposal of unassembled IgM and J chains occurs in a mannosidase I dependent, glucosidase independent fashion exluding a possible role for calnexin.

Results from the mammalian cells model contrast with those from the yeast model with respect to the role of calnexin. Knop *et al.* (57) demonstrated that strains deleted for

CNE1 (*the S. cerevisiae* calnexin homologue (63)) had little effect on CPY* degradation, while Jakob *et al.* (58) also reported the optimal glycan structure for ERAD as being $Man_8GlcNAc_2$. Jakob proposed the existence of a Man_8 lectin that selectively targets misfolded proteins for retrotranslocation and degradation.

1.4 Signalling from the ER

The accumulation of unfolded proteins in the ER leads to a stress response termed the unfolded protein response, UPR. The UPR is the defined by the signalling of the presence of unfolded proteins in the ER to the nucleus and leads to increased transcription of some ER proteins and folding enzymes. Experimentally, there are several ways in which this response can be generated. For example, compounds such as dithiothreitol (DTT) which inhibit disulfide bond formation, or tunicamycin which inhibits *N*-glycosylation are potent inducers of the transcription of ER proteins such as BiP. This upregulation of the synthesis of molecular chaperones and folding enzymes leads to an increased capability of the ER to process unfolded proteins. The UPR is found in both mammalian cells and yeast, and recent results show that the components responsible for the response are remarkably conserved between species.

1.4.1 The unfolded protein response in yeast

In *S. cerevisiae* the *IRE1* gene was identified in a screen for genes involved in the UPR (64). The screen was to use conditions which lead to unfolded proteins in the ER and to use fusion of the *KAR2* promoter (the yeast BiP) with a *LacZ* gene acting as a reporter. The screen was for mutants which no longer upregulated the *KAR2* promoter. Several mutants were isolated and the genes responsible were cloned by complementation of the mutants using a yeast plasmid library. The remarkable finding was that the gene responsible, *IRE1*, had been previously identified as an inositol auxotroph (65). The *IRE1* gene codes for a predicted type 1 transmembrane protein, Ire1p, with an N-terminal ER lumenal sequence, a transmembrane domain, and a C-terminal cytosolically oriented domain which has both a serine/threonine kinase domain and an endoribonuclease domain. Genetic analysis has shown that the Ire1p protein

dimerizes due to the accumulation of unfolded proteins in the ER but the ligand which is responsible for this dimerization remains unknown (66).

Oligomerization of Ire1p has been shown to lead to *trans*-autophosphorylation of the kinase domain and apparently activation of the ribonuclease activity (67). The ribonuclease domain of Ire1p has sequence similarity with mammalian RNaseL, that is activated upon treatment of cells with interferon. In the case of RNaseL, the ligand for dimerization is a 2'-5'-linked oligoadenylate and the consequence of this binding is the activation of the ribonuclease activity (68). In the case of Ire1p, deletion of the predicted ribonuclease domain leads to a defective UPR but the kinase activity is not affected. The involvement of a ribonuclease activity in the UPR was rationalized by the isolation in a genetic overexpression screen of another member of the UPR pathway, the gene for *HAC1* (69). The protein Hac1p is a bZIP transcription factor and deletion of this gene also abolishes the UPR. The surprising result is that *HAC1* mRNA is a substrate for the site specific endoribonuclease activity of Ire1p. *HAC1* itself, is not a gene under the control of the transcriptional consequences of the UPR but its mRNA is constitutively transcribed, albeit poorly translated. This defect was found to be due to an internal 252 base pair sequence originally thought to be a region which affected the stability of the Hac1p protein, but subsequently it was shown that both forms of mRNA give rise to Hac1p proteins that are unstable. Another gene identified in the original screen as required for the UPR was *RLG1,* which codes for the ligase responsible for processing the introns in tRNAs (70). Although this was initially difficult to reconcile with the initial discovery of *IRE1*, with the later realization that it has site-specific endoribonuclease activity a rational and novel mechanism of control of the UPR was proposed (70).

Thus the current model for the UPR in yeast is that the Ire1p is oligomerized by an unknown ligand in the ER, and this oligomerization permits phosphorylation of the kinase domains, which in turn leads to activation of the intrinsic ribonuclease activity. *HAC1* mRNA is spliced and ligated, exported to the cytoplasm, and translated. The translated Hac1p protein is imported into the nucleus and combines with other transcription factors to increase the rate of transcription of ER molecular chaperones and folding enzymes (Figure 1).

One requirement of this model is that the *HAC1* mRNA is accessible to the processing activity of Ire1p and the Rlg1p. The latter was known to be restricted to the nucleus and recent results with a mammalian hIrep homologue have shown that it is located on the inner nuclear membrane which is contiguous with the ER and needs an intact kinase domain to signal the UPR (71). Other members of the UPR signalling pathway that have been identified by yeast two-hybrid as interacting with both Ire1p and with Hac1p are the members of the yeast transcriptional coactivator complex Gcn5p/Ada, containing Gcn5p, Ada2p, Ada3p, and Ada5p (72). In support of their functional role, the disruption of the *GCN5*, *ADA2*, and *ADA3* genes leads to a reduced UPR and deletion of the *ADA5* gene abolishes it (72). Lastly, a serine/threonine phosphatase, Ptc2p, was identified as a direct interactor with Ire1p and can dephosphorylate it (73).

1.4.2 The mammalian UPR

Two isoforms of Ire1 have been cloned from human and mouse, and termed Ire1α and Ire1β (71,74). A remarkable set of experiments (71,75) characterizing the human homologues were based on the result that demonstrated the endoribonuclease activity of the yeast Ire1p on *HAC1* mRNA (67). Tirasophon and colleagues tested for the endoribonuclease activity of human Ire1 (Ire1αp), however, neither a mammalian homologue for *HAC1*, nor a putative substrate for the enzyme were known. Therefore, they tested its activity on the yeast *HAC1* mRNA prepared *in vitro*. They expressed hIre1αp in COS-1 cells, and demonstrated the cleavage of *HAC1* mRNA at the 5' splice site junction into two species, a 224-nucleotide 5' exon and a 326-nucleotide intron/3' exon. The precise location of the splice site was determined by primer extension analysis and confirmed to be identical to the 5' splice site junction processed by a GST-yeast Ire1p fusion protein (which itself however could cleave at both 5' and 3' junctions of the intron). Tirasophon suggested that Ire1α and Ire1β may each have evolved with specificities for 5' and 3' splice site junctions respectively. They also tested the endoribonuclease activity of a kinase defective form of hIre1αp in which a conserved lysine residue in the ATP binding site had been mutated (K599A). The mutant form failed to cleave the *HAC1* mRNA substrate suggesting that the kinase activity is a prerequisite for endoribonuclease activity. Moreover, while overexpression of wild-type Ire1p

constituitively activated a *LacZ* reporter gene under the control of the rat BiP promoter, a result also observed in *S. cerevisiae* (67), they found that the mutant hIre1αp-K599A effectively blocked the unfolded protein response thus acting in a *trans*-dominant manner. Alternatively, Wang *et al.* (74) confirmed the functional homology of the murine Ire1p through its overexpression in 293T cells, leading to an increase in the levels of GRP78/BiP, but not to the same extent as with tunicamycin.

Recently, Niwa *et al.* (75) examined the capacity of mammalian cells to *in vivo* splice yeast *HAC1* mRNA. They transfected HeLa cells with a plasmid containing the yeast *HAC1* gene, and the UPR was induced with tunicamycin. PCR was performed on cDNAs generated from both control and induced cells. The PCR product of the cDNA from control cells was identical in size to the uncleaved *HAC1*, while a 370 bp product was retrieved from cells induced for the UPR. Sequencing the PCR product confirmed the accurate splicing of the *HAC1* mRNA. This result confirms the functional conservation of a highly specific mRNA splicing apparatus from yeast to mammalian cells. The question was how the yeast mRNA was spliced, and the model in which the hIre1αp and hIre1βp homologues each cleave specific splice site junctions of the *HAC1* mRNA was tested. Both hIre1p isoforms were made in baculovirus/insect cell system, and [α-^{32}P] UTP labeled and *in vitro* transcribed yeast HAC1 mRNA was prepared and used as a substrate to test the endoribonuclease specificities for each isoform. Both were highly specific and exhibited identical ribonuclease activities to the yeast Ire1p.

Niwa *et al.* (75) found that both hIre1αp and hIre1βp were localized to the ER membrane, but were translocated to the nucleus upon stimulation of the UPR. By indirect immunofluorescence with anti-peptide antibodies that were isoform specific, they found a localization pattern identical to the β subunit of the signal recognition particle and mp30, which are both ER-resident membrane proteins. It was noted that 24% (hIre1αp) and 54% (hIre1βp) of cells exhibited nuclear staining and upon the addition of tunicamycin, the number of cells with nuclear localization patterns of hIre1βp increased to 80% and 100%. Moreover, the pattern was diffuse throughout the nucleus, with no nuclear rim staining, which is inconsistent with them being integral membrane proteins. This anomaly was resolved by Western blot analysis that revealed the 140 kDa predominant form of Ire1p underwent proteolysis upon UPR induction which led to an increase in a 60

kDa band. Based on the size of the fragment and localization of the epitope, it was suggested that proteolysis occurred cytosolically at only the C-terminal of the transmembrane domain. Thus, a model has been developed which can be summarized as follows: unfolded proteins promote the oligomerization of Ire1p, then *trans*-autophosphorylation occurs on the cytoplasmic domain activating an endoribonuclease domain, which is proteolytically cleaved from the cytosolic surface of the ER and is then targetted to the nucleus, where it processes a yet unknown substrate similar to *HAC1*. This eliminates the transcriptional attenuation caused by the presence of the intron, allowing for the translation, and subsequent binding of the transcriptional activator to the UPR upstream *cis*-element (UPRE), transcriptionally activating the UPR induced genes (Figure 1).

One known protease that cytosolically cleaves integral membrane proteins is γ-secretase. PS1, a gene thought to encode a protein that either exhibits γ-secretase activity or is involved in activating γ-secretase, was a potential candidate for the proteolysis of hIrep. It has been shown that γ-secretase cleaves amyloid precursor protein (APP) generating Aβ, the major component of amyloid plaques seen in Alzheimer's patients (76). To test the involvement of PS1, Niwa's group examined the localization of both hIre1αp and hIre1βp in fibroblast cells derived from homozygous knockout mice ($PS1^{-/-}$). Localization of both forms of Ire1p upon UPR induction remained unchanged in $PS1^{-/-}$ cells, while 100% of Ire1αp and 95% of Ireβp displayed nuclear distribution patterns of in $PS1^{+/+}$ cells after treatment with tunicamycin. Moreover, they examined UPR induction in $PS1^{-/-}$ cells in a 7 hour time course monitoring changes in BiP mRNA levels. BiP was induced to three-fold that of basal levels, while $PS1^{+/+}$ cells exhibited a five-fold induction. It was clear from these experiments that PS1 was involved in UPR-dependent nuclear localization of both hIre1αp and hIre1βp, and also an important mediating factor in the mammalian UPR (75).

Interestingly, the RNase activity of hIre1αp has recently been shown to downregulate its own mRNA expression (77). Although, when tested under the same conditions as with *HAC1* as a substrate, hIre1αp could not cleave its own mRNA (77), suggesting that this regulatory control is indirect.

1.4.3 Ire1p and PERK in the regulation of other cellular responses to ER stress

While stress signaling from the lumen of the ER is comprehensively reviewed elsewhere (78), recent and exciting work has been performed by the Ron laboratory at the New York University Medical School, on Ire1p and the related PERK kinase, and their role in mediating the activity of stress-activated protein kinases (SAPKs) and translational attenuation in response to ER stress. SAPKs, also known as JNKs for cJUN NH_s-terminal kinases, are activated in response to ER stress induced by perturbations that promote the UPR (79,80). These kinases then activate transcription through the phosphorylation of transcriptional activators such as cJUN and ATF2 (81,82). The Ron group disrupted m*IRE1*α in mouse embryonic stem (ES) cells (83). Fibroblasts derived from wild-type embryos exhibited a two-fold increase of JNK activity upon the induction of ER stress, whereas *IRE1*$\alpha^{-/-}$ embryos not only failed to induce JNK activity, but displayed reduced levels of activity. Thus, mIre1αp was found to mediate JNK activation in embryonic cells in response to ER stress.

They then searched for the proteins that couple the UPR receptor to SAPK pathway activation using the yeast two hybrid screen with mIre1αp as a bait and isolated TRAF2. TRAFs are adapter proteins that are recruited to the cytosolic domains of ligated receptors, activating downstream JNKs. This has been best demonstrated for the tumor necrosis factor (TNF) receptor in the activation of the stress-activated pathway (84,85). The interaction was confirmed *in vivo* by co-immunoprecipitation experiments. A model was then proposed that Ire1p mediated activation of the JNK stress activated pathway directly through the recruitment of TRAF2, in a mechanism that is similar to JNK activation by cell surface receptors such as the TNF receptor.

Coincidentally with the upregulation of ER resident proteins by upstream cis-acting elements (UPRE) in response to accumulation of unfolded proteins, cells also down regulate translation through the phosphorylation of eIF2α. The phosphorylated form of eIF2α interferes with the assembly of the 43S translation-initiation complex (86). One of the kinases that phosphorylates eIF2α, the interferon-inducible RNA-dependent protein kinase (PKR) has been shown to be activated in response to treatments that lead to Ca^{2+} depletion in the ER (87). However, the Ron group found that the inhibition of protein synthesis in *PKR*$^{-/-}$ cells was close to that of wild-type cells in response to ER stress.

Using a sequence homology based approach, they searched for kinases related to PKR and HRI (heme-regulated eIF2α kinase) that could be involved in the response to ER stress. They identified a *Caenorhabditis elegans* cosmid clone (CEF46C3) that encodes a predicted type I transmembrane protein with a kinase domain highly similar to PKR and HRI, and a lumenal domain homologous to Ire1p. Using a human expressed sequence tag (EST) clone that encoded a peptide fragment similar to the predicted sequence to the *C. elegans* protein, they isolated the mouse PERK cDNA, that encoded a protein with 20% homology to the lumenal domain of Ire1p and 40% homology of the cytosolic domain to PKR. *In vitro* experiments demonstrated that bacterially expressed GST-PERK fusion (for PKR-like ER kinase) could phosphorylate eIF2α, while translation-competent reticulocyte lysates were profoundly inhibited by the addition of the fusion proteins. Lastly, they expressed the lumenal domain of PERK. This had a dominant-negative effect on ER stress induction of the UPR marker CHOP, suggesting a similar dominant negative effect on Ire1p as seen in yeast and mammalian cells with the expression of the lumenal domain of Ire1p (66,74). This suggested that the lumenal domains of Ire1p and PERK are functionally conserved, and share similar ligands or upstream transducers. Nevertheless, they mediate different signaling pathways, one upregulating the expression of ER resident chaperones and folding enzymes (Ire1p), while the second attenuates translation (PERK). This model is supported with the recent results of Bertolotti *et al.* (88). They found that the lumenal domains of Ire1p and PERK were interchangeable, and that both form stable complexes with BiP in unstressed cells. It was also found that BiP dissociates from the lumenal domains of both Ire1p and PERK upon ER stress, leading to the activation and oligomerization of both Ire1p and PERK. Thus, BiP is likely to be the ligand that regulates the activation of both ER stress receptor kinases.

1.5 Concluding remarks

It is clear that for the processes of the Calnexin Cycle, ERAD, and the unfolded protein response, most of the key molecules appear to be known but the details of their mechanisms remain to be elucidated. Similarly, for some of the ER processes such as protein translocation into the ER, glycosylation, vesicular transport, some of the components are known and mechanisms are being elaborated. The role of the ER in a

variety of diseases has been known for some time, and we are becoming increasingly aware of the underlying mechanisms.

1.6 Aim of work

1.6.1 Studies on human glucosidase II

Since the early report by Ou *et al.* (8), which was the first to describe a link between N-glycan trimming and protein folding in the ER, the importance of glucosidase II has been clear. The processing of N-linked oligosaccharides on nascent glycoproteins by the ER α-glucosidases, which rapidly trim the oligosaccharide to the $Glc_1Man_9GlcNAc_2$ form, is required to initiate binding to the lectin-like chaperones calnexin and calreticulin. More specifically, glucosidase II mediates entry into the Calnexin Cycle by trimming the oligosaccharide from the Glc_2 to the $Glc_1Man_9GlcNAc_2$ form. Moreover, through the removal of the last Glc, glucosidase II also mediates the exit from this cycle.

At the time of this study, there were two main questions that needed to be addressed with respect to glucosidase II. The first question was one of genetic heterogeneity of the catalytic α subunit. With the recent finding of splice variants of the catalytic α subunit (89), and in our study (90), it was unclear whether the complex kinetic model for glucosidase II (91) represented the activities of heterogeneous preparations of glucosidase II, or whether the splice variants each exhibit this complex activity. This question gives rise to two models for glucosidase II. The first is one of splice variants with either pre- and post-calnexin binding activities. Hence, this model would suggest that the splice variants are catalytically different, with splice variant(s) that hydrolyse the $Glc_2Man_9GlcNAc_2$ oligossacharide permitting entry into the Calnexin Cycle, and other splice variant(s) trimming the $Glc_1Man_9GlcNAc_2$ form, releasing the glycoprotein from calnexin/calreticulin. The alternative and simpler model suggests that the splice variants are catalytically indistinguishable and can hydrolyse both Glc_2- and $Glc_1Man_9GlcNAc_2$ forms of N-linked oligosaccharide.

The second question to be addressed was regarding the putative heterodimeric structure of glucosidase II. The requirement of the β subunit of glucosidase II for enzymatic activity was still unclear as reports supported two models. The first model

proposed by Trombetta *et al.* (92) suggested a heterodimeric structure for glucosidase II, since they were unable to separate the two subunits and maintain activity. They also found, upon cloning of the rat cDNA for the β subunit, that it encoded an ER retrieval sequon (HDEL), and thus proposed a role of the β subunit in the retention of the catalytic subunit. This model was confirmed by the report of D'Allesio *et al.* (93) on *Schizosaccharomyces pombe* glucosidase II. They found that monoglycosylated oligosaccharides ($Glc_1Man_9GlcNAc_2$) were severely diminished in strains with a β subunit deletion.

The second model, which suggests that the catalytic α subunit is sufficient for enzymatic activity, was proposed by Hentges and Bause (94) and Flura *et al.* (95). Hentges and Bause (95) observed glucosidase II preparation from pig liver with and without β subunit. As well, Flura *et al.* (94) cloned and expressed a cDNA derived from pig liver, and found increased glucosidase II activities when it was transiently expressed in CHO cells. Both these reports support the latter model.

Chapter 2 of this thesis addresses these two main questions. How does the genetic heterogeneity relate/alter to the current kinetic model for glucosidase II? Secondly, what is the functional nature of the β subunit? Is it required for enzymatic activity, moreover does the heterodimeric structure play a role in glucosidase II localization?

1.6.2 ER specific two-hybrid interactions: ERp57, a monoglucosylation specific protein disulfide isomerase, associated with calnexin and calreticulin

Our initial intention, upon beginning the experiments described in Chapter 3, was to characterize the subunit interactions of human glucosidase II. We believed it was possible to perform these experiments *in vivo* using an Ire1p-based yeast two-hybrid system. Ire1p, as described earlier, is an ER type I transmembrane kinase that requires oligomerization to activate the unfolded protein response pathway, which in turn upregulates the expression of genes that are downstream of an unfolded protein response element (UPRE).

Yeast strains were constructed with the *LacZ* gene from *E. coli* under the control of a yeast minimal UPRE, and parental constructs were prepared with *IRE1*, with the region encoding the lumenal domain deleted. However, constructs expressing Ire1p

fusions with the α subunit of glucosidase II as an ER lumenal domain, autoactivated the unfolded protein response pathway. Alternatively, fusions with calnexin interacted specifically with ERp57 fusions, and as expected, did not interact with PDI. With the available calnexin crystal structure (96), we used the ER protein two-hybrid system to map the regions that mediate the protein-protein interactions between both calnexin and calreticulin, with ERp57, and confirmed our findings *in vitro*. Chapter 3 describs the develpment of our ER protein two-hybrid assay, the validation of this system, and the results from our mapping experiments.

Chapter 2

The Heterodimeric Structure of Glucosidase II is Required for its Activity, Solubility, and Localization *in vivo*

2.1 Abstract

Glucosidase II is an ER heterodimeric enzyme that cleaves sequentially the two innermost α–1,3-linked glucose residues from *N*-linked oligosaccharides on nascent glycoproteins. This processing allows the binding and release of monoglucosylated ($Glc_1Man_9GlcNAc_2$) glycoproteins with calnexin and calreticulin, the lectin-like chaperones of the endoplasmic reticulum. We have isolated two cDNA isoforms of the human α subunit (α1 and α2) differing by a 66 bp stretch, and a cDNA for the corresponding β subunit. The α1 and α2 forms have distinct mobilities on SDS-PAGE and are expressed in most of the cell lines we have tested, but were absent from the glucosidase II-deficient cell line PHA^R 2.7. Using COS7 cells, the co-expression of the β subunit with the catalytic α subunit was found to be essential for enzymatic activity, solubilization and/or stability, and ER retention of the α/β complex. Transfected cell extracts expressing either α1 or α2 forms with the β subunit showed similar activities, while mutating the nucleophile (D542N), predicted from the glycoside hydrolase Family 31 active site consensus sequence, abolished enzymatic activity.

To compare the kinetic parameters of both α1/β and α2/β forms of human glucosidase II, the protein was expressed with the baculovirus expression system. Expression of the human α or β subunit alone led to the formation of active human/insect heteroenzymes, demonstrating functional complementation by the endogenous insect glucosidase II subunits. The activity of both forms of recombinant human glucosidase II was examined with a *p*-nitrophenyl α-D-glucopyranoside substrate, and a two binding site kinetic model for this substrate was shown. The $K_M1\text{-}2$ values and apparent $K_i1\text{-}2$ for deoxynojirimycin and castanospermine were determined and found to be identical for both isoforms suggesting they have similar catalysis and inhibition characteristics. The substrate specificities of both isoforms using the physiological oligosaccharides were assessed and found to be similar.

2.2 Introduction

Nascent glycoproteins entering into the ER are substrates for oligosaccharyltransferase, which transfers from a dolichol-phosphate precursor a $Glc_3Man_9GlcNAc_2$ glycan to a Asn-X-Ser/Thr sequon. The α-glucosidases of the

endoplasmic reticulum (ER) then sequentially hydrolyze the three terminal glucoses on the N-linked oligosaccharide. Glucosidase I cleaves the first α-1,2-linked glucose of the glycan structure, then glucosidase II sequentially cleaves the two inner α-1,3-linked glucoses to yield $Man_9GlcNAc_2$. The intermediate of this reaction ($Glc_1Man_9GlcNAc_2$) is specifically bound by the ER lectin-like chaperones calnexin and calreticulin (17,18,27,97-99). Then ERp57, a protein disulfide isomerase which is specific for monoglucosylated glycoproteins, associates with calnexin and calreticulin, and catalyses disulfide bond exchange on the lectin bound glycoproteins (25). The interaction between the glycoprotein and the lectin is abrogated upon cleavage of the last glucose residue by glucosidase II. If the protein has not attained its native conformation, it is then recognized and reglucosylated by the ER enzyme UDP-glucose glycoprotein:glucosyl transferase (UGGT) (15,100,101), which allows rebinding to calnexin and calreticulin. However, if correct folding is achieved, the nascent glycoprotein escapes recognition by UGGT and leaves the ER to continue its maturation along the secretory pathway. This process has been termed "The Calnexin Cycle" (102,103).

Inhibitors of the ER α-glucosidases have been useful in elucidating the effects of glycan processing on the maturation of secretory and membrane glycoproteins (94,97,104-107). While not all glycoproteins have been shown to interact with calnexin/calreticulin, impaired association with these lectins can lead to minor structural changes and loss of function, or ultimately to retro-translocation through the translocon and degradation in the cytosol by the ubiquitin-proteasome pathway (35,108,109). In earlier studies with cultured cells (110-114), and more recently with animal model systems (115,116), α-glucosidase inhibitors have shown also to inhibit viral propagation by blocking viral envelope glycoprotein-ER lectin associations (116,117).

Glucosidase II has been described both biochemically and genetically as a heterodimer with the α subunit displaying sequence homology with the glycoside hydrolase Family 31 (92,93,95,118). The α subunit contains the catalytic domain, which is characterized by the consensus sequence (G/F)-(L/I/V/M)-W-X-D-M-N-E (119), while the β subunit is presumably involved in ER localization through its C-terminal HDEL sequence. Although mostly localized to the ER, Brada *et al.* (120) found glucosidase II in endocytic structures beneath the plasma membrane as well as associated with vacuoles in

kidney brush border cells. These authors were the first to postulate that there might be two forms of the enzyme. Recently, Arendt and Ostergaard (118) found glucosidase II associated with CD45, a transmembrane protein-tyrosine phosphatase (PTP) in SAKR mouse T-lymphocyte cell line. In the same study, they cloned a cDNA for the α subunit that encodes an additional 22 amino acids stretch not found in the available human sequence (GenBank accession number D42041). More recently, the same authors reported the existence of several isoforms of the murine α subunit originating from alternative splicing (89). It was suggested that these isoforms differ in their activity, regulation, and possibly in their protein-protein interactions (89).

The β subunit of glucosidase II was originally identified as the protein kinase C substrate 80 K-H (121). It is characterized by its own ER signal sequence, EF-hand high-affinity calcium binding loops, two glutamic acid repeats in tandem, and a putative ER retention motif HDEL (89,92,118). Work by Trombetta *et al.* (122) has shown that the α and β subunits co-purify while attempts to separate the α and β chains under conditions that preserve enzymatic activity were unsuccessful. Alternatively however, Flura *et al.* (94) observed a variable increase in glucosidase II activity when they expressed a pig liver cDNA encoding the α subunit alone in CHO cells. Furthermore, Hentges and Bause (95) had pig liver glucosidase II preparations both with and without a detectable 60 kDa component. Thus the requirement of the β subunit and the functional nature of the heterodimeric glucosidase II complex remained to be established.

The current kinetic model for *p*-nitrophenyl α-D-glucopyranosidase (*p*-NP-Glcase) and maltase activities of glucosidase II, proposed by Alonso *et al.* (91,123), suggests the existence of both high and low affinity substrate binding sites. This model was supported by the earlier work of Hubbard and Robbins (60), which reported different rates of glucose hydrolysis from N-linked Glc_2- versus $Glc_1Man_9GlcNAc_2$ oligosaccharides. Moreover, bromoconduritol has been shown to only inhibit the hydrolysis of the Glc_1 form of the oligosaccharide (124,125). There are two possible explanations for this model, the first being based on the proposed two substrate binding sites. However, with the discovery of α subunit splice variants, the possibility that these activities are properties of distinct isoforms is raised. Previous results could be due to heterogeneous

enzyme preparations which are responsible for the differential activities. We can now test this hypothesis with genetically purified isoforms of glucosidase II.

We report here the isolation of human cDNAs encoding for the α1 (short form), α2 (long form) and β subunits of glucosidase II. A role of the β subunit was sought through expression in COS7 cells, while a heterologous expression system was developed to perform a detailed comparison of the recombinant α1/β and α2/β isoforms, based on their kinetic parameters for the *p*-nitrophenyl α-D-glucopyranoside substrate, and their substrate specificities for the physiological oligosaccharides.

2.3 Materials and Methods

2.3.1 Reagents

All reagents were purchased from Sigma (St. Louis, MO) except where otherwise mentioned. Dulbecco's modified Eagle's medium (DMEM), Opti MEM I, Dulbecco's PBS (D-PBS) and Lipofectamine were purchased from Gibco/Life Technologies Inc. (Rockville, MD). Fetal bovine serum was purchased from Hy-Clone (Logan, UT), deoxynojirimycin from Calbiochem (San Diego, CA) and octyl glucoside from Boehringer Mannheim Canada (Laval, QC). Restriction and modification enzymes were from Amersham Pharmacia Biotech Inc. (Uppsala, Sweden) and New England Biolabs (Beverly, MA).

2.3.2 Cell lines

RPMI8402 cells were obtained from Dr. C. Milstein (Cambridge, UK), Peer cell lines from Fujisaki Cell Center (Tokyo, Japan), mouse ES cells from Genome Systems Inc. (St-Louis, MO) and $PHA^R2.7$ from Dr. I. Trowbridge (La Jolla, CA). All other cell lines were purchased from American Type Culture Collection (ATCC, Manassas, VA).

2.3.3 Cloning of human glucosidase II cDNAs.

2.3.3.1 α-subunit

Rat glucosidase II was purified as described previously (17) electroblotted onto a PVDF membrane and the major 120 kDa protein was N-terminally sequenced by using an Applied Biosystems 470A protein sequencer equipped with an on-line

phenylthiohydantoin amino acid analyzer. The N-terminal sequence was determined as VDRSNFKTCEESSFCKRQRS. A partial ORF sequence of human immature myeloid cell line KG1 (accession number D42041) displayed homology to the N-terminal amino acid sequence. mRNA was isolated from HeLa cells using the Quick Prep mRNA Purification Kit (Amersham Pharmarcia Biotech,) and cDNA amplified using the Marathon cDNA Amplification Kit (Clontech, Palo Alto, CA). Oligonucleotides 5' CAT GCT CAG GGA TCC CAT AGA CAT GCT (anti-sense) and 5' AAG CAT ACG GCC AGG CCT CTC TCC ATA CCG (sense) were used in conjunction with the adapter primer AP1 supplied with the kit (underlined nucleotides indicate *Bam*HI and *Stu*I restriction sites respectively). PCR conditions were as described by the manufacturer (all PCR reactions were performed with a Perkin Elmer 9600 instrument). The amplicons of ~0.8 Kb (doublet) and ~3.6 Kb respectively were mixed and reamplified using AP2 primer and Expand Long Template PCR (Boehringer Mannheim, Canada). The resulting ~3.9 kb fragment was digested with *Not*I and ligated in the mammalian expression vector pcDNA3 (Invitrogen, San Diego, CA). The *EcoR*V/*Hind*III fragment of the polylinker was deleted to remove a hairpin structure to enhance translation and protein production. Full-length cDNAs for human glucosidase II were sequenced. The pointmutated α subunit was generated using the primer pairs 5' GTC TGG AAT **AAC** ATG AAC GAA CC (sense) (bold face indicates the mutated codon) / 5' GCT GAA GCT TAT CGC AGG TGA ATA CTC CAA TC (anti-sense) and 5'-GGT TCG TTC AT**G TT**A TTC CAG AC (anti-sense) (bold face indicates the mutated codon) / 5'-AAG CAT ACG GCC AGG CCT CTC TCC ATA CCG (sense) (underlined nucleotides indicate a *Stu*I restriction site). The fragments obtained were purified and reamplified using the flanking primers 5' GCT GAA GCT TAT CGC AGG TGA ATA CTC CAA TC (anti-sense) and 5'-AAG CAT ACG GCC AGG CCT CTC TCC ATA CCG (sense). An *Aat*II/*Kpn*I fragment (901 bp) was replaced in the original full-length clones and resequenced. The final 3.9 Kb amplicon contains original 5' and 3' untranslated regions.

2.3.3.2 β subunit

The β subunit of human glucosidase II sequence was previously identified (GeneBank accession #J03075) (92,126). This gene was amplified using normal human

lymphocytes mRNA as template. cDNA was obtained using the TITAN RT-PCR kit (Boehringer Mannheim) using oligos 5' GCT CGA GAA TTC GGT GAG ATG CTG TTG CCG CTG (sense) (underlined nucleotides indicate the *EcoR*I restriction site used for cloning) and 5' CGC GGT CTA GAT TAC TAG AGC TCG TCA TGG TCG TCT TC (anti-sense) (underlined nucleotides indicate the *Xba*I restriction site used for cloning). The 1.6 kb amplicon obtained was digested with *EcoRI/Xba*I and cloned in pcDNA3. Both DNA strands were sequenced. A 135 bp *Nco*I/*Xba*I fragment of sequence encoding the C-terminal HDEL was removed and replaced by a mutagenic PCR fragment amplified with the primer 5' CG CGG TCT AGA TTA CTA **ATG GTG ATG GTG ATG ATG** GTC GTC TTC GGT GGG TGC (anti-sense) (underlined nucleotides indicate the *Xba*I restriction site used for cloning, and bold represents the $(His)_6$-tag) and 5' GTC CGA GAA TTC GGT GAG ATG CTG TTG CCG CTG (sense) .

2.3.4 Southern blot analysis of glucosidase II

Human lymphocyte or mouse ES cell genomic DNA (10 μg) was digested overnight at 37°C with 100U of *BamH*I and/or *Xmn*I, separated on a 0.7 % agarose gel and transferred to Hybond N^+ membrane (Amersham Pharmacia Biotech). Blotted DNA was probed with a [^{32}P]-labeled *Xmn*I/*BamH*I cDNA fragment (220 bp). The blot was washed twice in 2X SSC, 0.1% SDS at 65°C for 30 minutes and exposed with Kodak XAR-5 film.

2.3.5 Transfections

COS7 African Green Monkey kidney cells were cultured in D-10 (DMEM supplemented with 10% FBS inactivated at 56 °C for 30 minutes). Cell monolayers at 40% confluency were transfected with 9.5 μg DNA and 40 μl Lipofectamine in 100 mm plates or 1.2 μg DNA and 5 μl Lipofectamine in 35 mm plates in Opti-MEM I and incubated for 16 hours. Medium was replaced with fresh D-10 medium and cells were collected 24 hours later. Large-scale co-transfections were performed with 4 μg pcDNA3 and 4 μg α subunit, with 4 μg pcDNA3 and 4 μg β subunit, or with 4 μg α and 4 μg β subunit. Small-scale transfections were performed with combinations of DNA as

described in Figure 8. Transfections included either 1.5 μg (large-scale) or 0.2 μg (small-scale) of the vector pQBI25 (Quantum Biotechnologies Inc, Montreal, QC) coding for Green Fluorescent Protein in order to evaluate transfection efficiency.

2.3.6 Mammalian cell lysate preparation

Cells were scraped and washed three times in D-PBS. They were then lysed on ice for 20 minutes in 10 mM Tris-HCl pH 6.8, 300 mM NaCl, 2 mM $CaCl_2$ and 1% Triton X-100 supplemented with Complete™ EDTA-free Protease Inhibitors (Boehringer Mannheim). Samples were homogenized using a Potter Elvejheim for 30 seconds on ice. The lysates were diluted 10-fold (0.1 % Triton X-100, final concentration) and cleared by centrifugation at 13000 x *g* for 5 minutes. The supernatant was collected and protein concentration measured using the Bio-Rad Protein Assay (Bio-Rad, Hercules,CA).

2.3.7 Glucosidase II assays on N-linked [^{3}H]Glc1Man9GlcNAc2

Acid phosphatase (AcP) was purified from *Saccharomyces cerevisiae* (DT111 strain that secretes glycoproteins with N-linked $Man_9GlcNAc_2$ glycans) and glycosylated *in vitro* with recombinant rat liver UDP-glucose:glycoprotein glucosyltransferase (UGGT) (15). Briefly, 10 μg of purified $Man_9GlcNAc_2$-acid phosphatase (G_0-AcP) in 20mM Tris-HCl pH 7.5, 10 mM $CaCl_2$ was labeled with 50 μCi UDP-[^{3}H]glucose (30 Ci/mmol, Amersham Pharmacia Biotech) using 100 nM of purified recombinant UGGT for 4 hours at 37°C. Unincorporated UDP-[^{3}H]glucose was removed using G-50 Sepharose spin columns (Amersham Pharmacia Biotech). Typical labeling obtained was 2-4 X 10^5 cpm/μg AcP. Cell lysates adjusted to a protein concentration of 500 μg/ml were incubated at 37 °C with 5 X10^5 cpm [3H]Glc1-AcP/ml, with or without 100 μM or 500 μM DNJ. Aliquots were removed every 15 minutes and immediately diluted in 3x SDS sample buffer with DTT (New England Biolabs) and boiled for 5 minutes. Half of each of the sample was resolved on 8% SDS-PAGE, followed by fluorography (Amersham Pharmacia Biotech). Dry gels were exposed with Kodak XAR-5 film for 60 h at –80°C. Films were scanned and image analyzed using the NIH Image Software (public domain).

2.3.8 Blue Native-Polyacrylamide Gel Electrophoresis

HeLa cell exctracts were solubilized in 2% CHAPS, 1% Triton X-100, or 2% octyl-glucoside for 30 minutes on ice. Unsolubilized proteins were removed by centrifugation for 30 minutes at 100 000 x g. Samples were mixed with Serva Blue and E-amino-caproic acid and separated on a linear 5 to 10 % acrylamide gradient as previously described (127). Separated proteins were transferred to Immobilon-P membrane and detected by Western blot analysis (see below).

2.3.9 In gel 4-methylumbellifeyl α-D-glucopyranosidase assay

Glucosidase II assays using the fluorescent substrate 4-methylumbelliferyl α-D-glucopyranoside were performed as described elsewhere (128).

2.3.10 Antibodies

Map-peptide antibodies were raised in New Zealand rabbits. They correspond to amino acid sequence SFQHDPETSVLVLRK (α subunit) and DHDKSFAMKYEQGTG (β subunit). The antisera also recognized the mouse glucosidase II homologue. Affinity-purified antibodies were prepared as described in The Xenopus Molecular Marker Resource (http://vize222.zo.utexas.edu) using the approriate immobilized peptide Sepharose. Anti-β antiserum was raised by injecting human β subunit, electroeluted from SDS-PAGE, into New Zealand rabbits. Mouse anti-human glucosidase II antiserum was raised by injecting BALB/c mouse with purified recombinant glucosidase II.

2.3.11 Western blot analysis

Protein samples were resolved on 8% SDS-PAGE and transferred on Immobilon-P membrane (Millipore). The membranes were first probed with α subunit anti-peptide antiserum or affinity-purified antibodies, incubated with anti-rabbit HRP (Bio-Rad or Santa-Cruz) and developed using Lumilight (for overexpressed proteins) or Lumilight Plus (endogenously expressed proteins) chemiluminuescent substrate (Boehringer Mannheim). They were stripped according to the manufacturer instructions and reprobed with anti-β subunit antibodies.

2.3.12 Immunoprecipitations

HeLa cells were labeled with [^{35}S]-methionine ([^{35}S] Promix Protein Labeling Mix, Amersham Pharmacia Biotech), and immunoprecipitations were performed with either protein A (Figure 1B) or protein G Sepharose (Figure 4C) as described in Current Protocols in Immunology (129) using rabbit anti-eletroeluted β subunit or mouse anti-glucosidase $II_{\alpha/\beta}$.

2.3.13 Baculovirus production and protein expression

Human cDNAs encoding both α1 (D42041 with Y850H) and α2 subunits (AF144074) with $(His)_6$-tags at the C-terminus and the β subunit of glucosidase II, both with and without a $(His)_6$-tag (AF144075) were subcloned into pFastBac1 and transformed into *E. coli* DH10Bac to produce recombinant bacmids. *Sf*9 cells were transfected with bacmid DNAs and the virus stocks amplified as described by the manufacturer (Gibco/Life Technologies). Cells were grown in Sf900 II SFM serum-free medium at 27°C for 3 days to a cell density of 1.5 - 2.5 x 10^6/ml and infected with either one or both viruses for an α subunit and β subunit at an M.O.I. of 2-5 pfu/cell. Only the β subunit without a $(His)_6$-tag was used in α/β co-infections. Infected cells were incubated for 3 days, harvested by centrifugation at 3000 x *g* and stored at -80°C.

2.3.14 Recombinant glucosidase II purification

Infected cells (approx. 7.5 x 10^8) were lysed in a Dounce homogenizer in 20 ml of Buffer A (250 mM NaCl, 20 mM Tris-HCl, pH 7.5) supplemented with 1% Triton X-100 and Complete™ EDTA-free protease inhibitors (Boehringer Mannheim). The homogenate was treated with bovine pancreas deoxyribonuclease I (Amersham Pharmacia Biotech) for 30 minutes on ice, and centrifuged at 10 000 x *g* for 10 minutes at 4°C. The supernatant was loaded by gravity flow onto a 1 ml Ni^{2+}-NTA Superflow column (Qiagen Inc., Valencia, CA). 2x Buffer A with 1% Triton X-100 and 20 mM imidazole was added to the flow through, and was then loaded again onto the Ni^{2+}-NTA column. The column was washed with 50 ml of 2x Buffer A with 10 mM imidazole, followed by 4 ml each of 2x Buffer A with 25 mM imidazole, and 50 mM imidazole. Glucosidase II was then eluted in 5x 1ml fractions of 2x Buffer A with 200 mM

imidazole. The eluate was desalted on a Sephadex G-25M column (PD-10; Amersham Pharmacia Biotech) in Buffer B (20 mM Tris-HCl, pH 7.5, 5 mM β-mercaptoethanol).

2.3.15 Anion exchange chromatography

The desalted Ni^{2+}-NTA fractions were loaded onto a Mono Q HR5/5 column (Amersham Pharmacia Biotech) equilibrated in Buffer B on a BioCAD Perfusion Chromatography Workstation (PerSeptive BioSystems Inc., Farmingham, MA). A linear gradient of 50 ml from 0 to 0.5 M NaCl in Buffer B was performed and the eluate collected in 1 ml fractions. The fractions were tested for activity and protein concentrations determined using a Bio-Rad Protein Assay (Bio-Rad, Hercules, CA). All fractions for each step of purification were analyzed by SDS-PAGE and Western blot.

2.3.16 *p*-Nitrophenyl α-D-glucopyranosidase assays and analysis of kinetic parameters

The activity for *p*-nitrophenyl α-D-glucopyranosidase (*p*-NP-Glcase) was tested in a 100 µl reaction mixture consisting of 5 mM *p*-NP-Glc (Sigma) in 50 mM PIPES pH 6.5 for 30 minutes at 37°C. Assayed culture medium (COS7 cells) (Figure 4C) was precleared with activated charcoal (BDH, Ville St-Laurent, Canada) to remove traces of phenol red. The reactions were stopped with an equal volume of 1 M glycine/NaOH and the OD_{405} read immediately. One unit is the production of 1 µmole/minute of *p*-nitrophenol at 37°C. Progress curves were monitored for both isoforms by continuous assay (n=5) for 30 minutes at 37°C in a SpectraMax 250 spectrophotometer (Molecular Devices Corp., Sunnyvale, CA). Continuous assay reactions were performed in 200 µl of 50 mM PIPES pH6.5 with a substrate concentration ranging from 0.1 to 14 mM *p*-NP-Glc. The slopes of the progress curves were determined by linear regression analysis using SoftMax Pro software version 1.1.1 (Molecular Devices, Corp), and kinetic parameters determined using ENZFITTER version 1.05 (Biosoft, Cambridge, UK). The two-binding site adaptations of the Michaelis-Menten and the Eadie-Hofstee equations described by Alonso *et al.* (91), were used to analyze the data.

2.3.17 *p*-nitrophenyl α-D-glucopyranosidase inhibition

Inhibition experiments were also performed by continuous measurement of *p*-NP-Glcase activity. Deoxynojirimycin (DNJ) and castanospermine (CST) (Calbiochem) were tested between 1.25-10 μM and 20-80 μM respectively. The velocity was measured for the substrate concentration range of the high affinity site (0.25-1 mM) and the low affinity site (8-14 mM). Apparent K_i values were determined from plots of reciprocal activity versus inhibitor concentration.

2.3.18 Purification and Glucosidase II hydrolysis of $[^3H]Glc_{1-3}Man_9GlcNAc_1$

Metabolic labeling with [^{3}H]glucose and extraction of lipid-linked oligosaccharides was performed as described by Zufferey *et al.* (130) from the following strains of *Saccharomyces cerevisiae:* YG424 (MAT**a**, *ade2-101, his3Δ200, lys2-801, Δalg8::HIS3, Δgls2::KanMX*) (58); YG491 (MAT**a**, *ade2-101, his3Δ200, Δalg10::KanMX, Δgls2::KanMX*) (58); SS328 (MATα, *ade2-101, his3Δ200, lys2-801, ura3-52*) (131), to yield [^{3}H]$Glc_1Man_9GlcNAc_2$, [^{3}H]$Glc_2Man_9GlcNAc_2$, and [^{3}H]$Glc_3Man_9GlcNAc_2$, respectively. The oligosaccharides were quantified by liquid scintillation counting, pooled, and treated overnight with 5 x 10^3 units Endo H (New England Biolabs). Glucosidase II assays were performed with both purified recombinant isoforms with enzyme concentrations normalized for *p*-NP-Glcase activity. The reactions were in 840 μl with 210 μl of enzyme, 50 mM PIPES pH 6.5 and 2.0 x 10^4 cpm of [^{3}H]$Glc_{1-3}Man_9GlcNAc_2$ pooled oligosaccharides. The time points used for the assays were 0, 15, 30 and 60 minutes, and the reactions were terminated by boiling for 2 minutes. Glucose hydrolysis from the oligosaccharides was monitored by HPLC as described by Romero and Herscovics (132).

2.4 Results

2.4.1 Cloning of human α1 and α2 subunits of glucosidase II

The cDNAs encoding the α and β subunits of glucosidase II were respectively isolated from HeLa cells and human lymphocytes. The full-length cDNA obtained for the human α subunit was essentially identical to the human partial open reading frame (GenBank# D42041), though it included a start codon and a tyrosine to histidine

substitution at position 850. Furthermore, an isoform with a 66 base pair insertion (corresponding to 22 amino acids, GenBank# AF144074) was isolated. This insertion encodes the amino acid sequence FSDKVNLTLGSIWDKIKNLFSR and has 86% identity to the previously published mouse sequence (GenBank# U92793) (118). Also found in human lymphocytes, it is located in the N-terminal region of the α chain. The insertion contains a potential N-glycosylation site that is not present in the mouse gene. Based on the human 80K-H cDNA sequence (GenBank# J03075), the cDNA encoding the β subunit was amplified from human lymphocyte mRNA. The amplified DNA fragment was identical to the published sequence except for an additional glutamic acid at position 314 (AF144075).

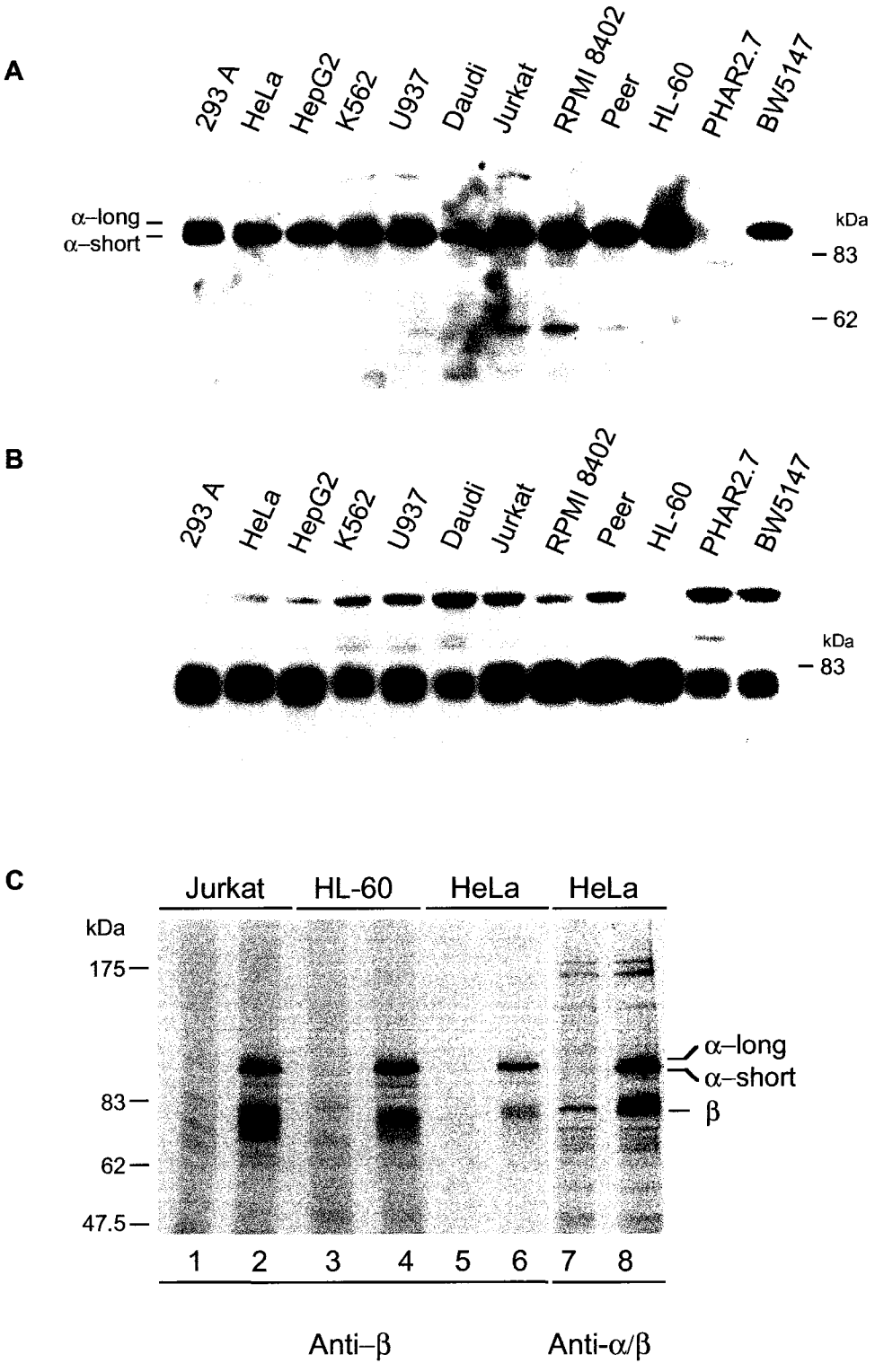
2.4.2 α1 and α2 isoforms originate from a single gene

The 66 nucleotide insertion found in the α2 isoform provides an additional *Xmn*I restriction site. Thus a 220 bp *Xmn*I/*Bam*HI fragment from the human α2 cDNA was used to probe *Xmn*I/*Bam*HI digested genomic DNA, isolated from human lymphocytes or mouse ES cells. If more than one gene had been responsible for the isoforms of this enzyme, a *BamHI*/*Xmn*I double digestion would have resulted in at least two bands: a higher molecular weight band for the short form and a lower one for the long form. Since only one band could be detected (data not shown), we conclude that both isoforms originate from a single gene. Alternative splicing of the mRNA is probably responsible for the expression of the isoforms of the α subunit.

2.4.3 Two α/β complexes of human glucosidase II are expressed *in vivo*

Western blot analysis was performed to study differential expression of α and β subunits in various human cell lines. Although the longer α subunit was expressed at lower levels, both isoforms were present in all human cell lines tested (Figure 2A) except for Peer T-cells. PHA^R2.7, a mouse glucosidase II deficient cell line was shown to be completely devoid of α subunit whereas only one form of α subunit could be detected in

Figure 2. Endogenous expression of two α/β complexes of human glucosidase II. (A) Triton X-100 lysates were prepared as described in the Experimental Procedures. 50 μg of each sample was resolved by reducing SDS-PAGE, electroblotted, and subjected to immunodetection with peptide-derived and affinity-purified anti-α (A) or anti-β (B) antisera. (C) Jurkat, HL-60, and HeLa were labeled with [^{35}S]-Promix and immunoprecipitations were performed on cell extracts with pre-immune (lanes 1,3,5) and immune (lanes 2,4,6) rabbit antiserum raised against electroeluted β subunit or pre-immune (lane 7) and immune (lane 8) mouse antiserum raised against purified recombinant glucosidase $II_{\alpha/\beta}$.


A
293 A
HeLa
HepG2
K562
U937
Daudi
Jurkat
RPMI 8402
Peer
HL-60
PHAR2.7
BW5147
α–long
α–short
kDa
83
62
B
293 A
HeLa
HepG2
K562
U937
Daudi
Jurkat
RPMI 8402
Peer
HL-60
PHAR2.7
BW5147
kDa
83
C
Jurkat
HL-60
HeLa
HeLa
kDa
175
83
62
47.5
α–long
α–short
β
1 2 3 4 5 6 7 8
Anti–β
Anti-α/β

its parental BW5147 cell line. The β subunit was present in all cell lines tested, even in the glucosidase II deficient cell line $PHA^R2.7$, where it appears as an abundant protein of ~80 kDa (Figure 2B).

Immunoprecipitations were performed from [^{35}S]-labeled Jurkat, HL-60 and HeLa cell extracts with rabbit antiserum raised against electroeluted β subunit (Figure 2B, lanes 1-6) and on HeLa cell extracts with mouse antiserum raised against recombinant purified human glucosidase II (lanes 7 and 8). Preimmune antiserum was used as a control for both sets of immunoprecipitations (lanes 1, 3, 5, and 7). The α and β subunits were immunoprecipitated together equally well by both antisera, and while the long α subunit isoform was precipitated by both antisera, the doublet at ~110 is more visible upon immunoprecipitation with the mouse anti-recombinant human glucosidase $II_{\alpha\beta}$ (lane 8).

2.4.4 Expression of glucosidase II β subunit promotes expression of soluble α subunit, and yields enzymatically active glucosidase II

Glucosidase II α and β subunits were transiently transfected individually or together in COS7 cells. Endogenous glucosidase II subunits were visible in the Triton X-100 soluble fraction but barely detectable in the pellet fraction (Figure 3, lane 1). However, when α1 or α2 subunits were overexpressed (Figure 3A lanes 2,5 and B, lanes 3,5), there was a slight increase in the levels of soluble protein but most of the expressed α subunit was found in the Triton X-100 pellet. Overexpressed β subunit (Figure 3A, lane 3) was generally soluble when expressed alone, but traces of insoluble material (pellet) were detected. Co-expression of the β subunit with any of the α subunits (Figure 3A, lanes 4, 6, and B, lanes 2,4,6) significantly reduced the extent of the insoluble α subunits but did not significantly increase the total amount of soluble α/β enzyme.

Glucosidase II activity in the COS7 lysates was determined on [^{3}H]G1-AcP as a substrate. Glucosidase II α1 and α2 subunits, with and without β subunit, were expressed and the activity in the lysates determined (Figure 4). Disappearance of radiolabel was

Figure 3. Western blot analysis of transfected subunits. COS7 cells were transfected with glucosidase II subunits as indicated on top of each lane. Expression of α and β subunits was analyzed in the soluble and pellet fractions of 1% Triton X-100 lysates and probed with anti- α or β anti-peptide antiserum. Panels (A) and (B) represent two sets of experiments.

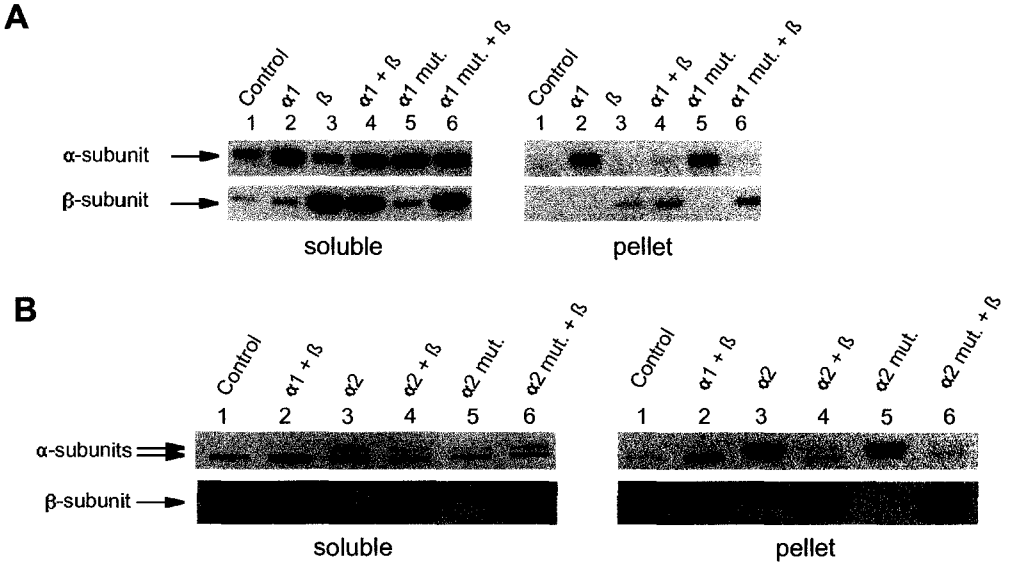
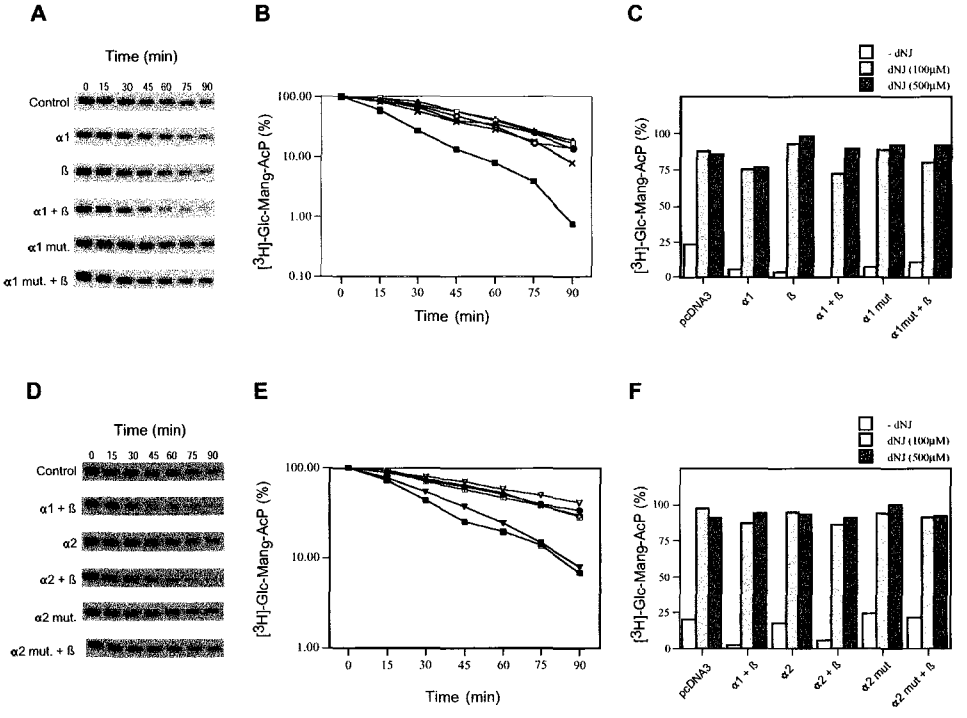

A
Control
α1
ß
α1 + ß
α1 mut.
α1 mut. + ß
1 2 3 4 5 6
α-subunit
β-subunit
soluble
pellet
B
Control
α1 + ß
α2
α2 + ß
α2 mut.
α2 mut. + ß
1 2 3 4 5 6
α-subunits
β-subunit
soluble
pellet

Figure 4. Glucosidase II activity in COS7 cell lysates with transfected subunits. Panels A, B and C illustrate the activities of the α1 or α1 mutant (mut.) subunits with or without co-expression of the β subunit. Panels D, E and F illustrate the activities of α2 or α2 mutant (mut.) subunits co-expressed or not with the β subunit and the comparative activity of α1 and α2 co-expressed with the β subunit. Lysates were prepared as described in the Experimental Procedures then incubated at 37°C in presence of [^{3}H]G1-AcP substrate. Samples were removed at the indicated time and separated by reducing SDS-PAGE, amplified and autoradiographed. Films were scanned (Panels A and D), analyzed using NIH image software and plotted (Panel B and E) as a percentage of initial substrate (time 0). Control (solid circle), α1 (open circle), β (multiplication sign), α1 + β (solid square), α1mut. (open square), α1mut. + β (solid triangle), α2 (open triangle), α2 + β (solid inverted triangle), α2mut. (open inverted triangle), α2mut. + β (crosshatched square). Panels C and F represent inhibition of specific deglucosylation activity by DNJ. Columns indicate percentage of residual substrate after 120 minutes of incubation at 37°C of lysates without or with 100 μM or 500 μM DNJ.


A
Time (min)
0 15 30 45 60 75 90
Control
α1
ß
α1 + ß
α1 mut.
α1 mut. + ß
B
[3H]-Glc-Mang-AcP (%)
100.00
10.00
1.00
0.10
0 15 30 45 60 75 90
Time (min)
C
- dNJ
dNJ (100μM)
dNJ (500μM)
[3H]-Glc-Mang-AcP (%)
0 25 50 75 100
pcDNA3
α1
ß
α1 + ß
α1 mut
α1mut + ß
D
Time (min)
0 15 30 45 60 75 90
Control
α1 + ß
α2
α2 + ß
α2 mut.
α2 mut. + ß
E
[3H]-Glc-Mang-AcP (%)
100.00
10.00
1.00
0 15 30 45 60 75 90
Time (min)
F
- dNJ
dNJ (100μM)
dNJ (500μM)
[3H]-Glc-Mang-AcP (%)
0 25 50 75 100
pcDNA3
α1 + ß
α2
α2 + ß
α2 mut
α2 mut + ß

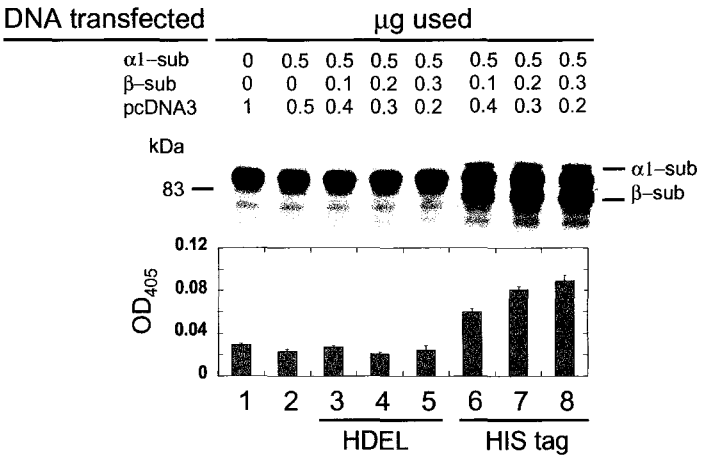
monitored over time and reflected the hydrolysis of the terminal glucose from the *N*-linked oligosaccharides on acid phosphatase, by both endogenous and conditionally overexpressed glucosidase II. Control transfections with pcDNA3 plasmid represent the endogenous levels of glucosidase II activity found in COS7 cells. Transfection of the catalytic α1 or α2 subunit alone did not increase the enzymatic activity over that observed in the control. As expected, expression of the β subunit alone did not increase endogenous glucosidase II activity. Increased rates of hydrolysis of radiolabeled substrate occurred only when α1 or α2 were co-transfected together with the β subunit. Glucosidase II activity of the lysates did not increase upon transfection of $\alpha 1_{mut}$ or $\alpha 2_{mut}$ (D542N and D564N, respectively) subunits together with β subunit despite comparable protein yield (see Figure 3A, lanes 4,6 and B, lanes 4,6). This result demonstrates that the glycoside hydrolase Family 31 putative active nucleophile (56), here D542 (α1) and D564 (α2) is required for catalytic activity. Interestingly, expression of either mutant alone had a dominant negative effect possibly competing for endogenous β subunit for heterodimeric assembly.

Deoxynojirimycin (DNJ), a competitive inhibitor of glucosidases I and II was added to the assays at concentrations of 100 μM or 500 μM. Under these conditions, deglucosylation was inhibited by 70 to 100% (Figure 4C and F). This residual activity may result from an endo-α-mannosidase present in the crude lysates, which can cleave $[^3H]Glc_1Man_1$- from $[^3H]Glc_1Man_9GlcNAc_2$-AcP to leave $Man_8GlcNAc_2$-AcP. A concentration of 100 μM inhibited most of the glucosidase II activity.

2.4.5 The α subunit HDEL-ER retrieval motif is responsible for the localization of the heterodimeric glucosidase II

The β subunit of glucosidase II contains a C-terminal HDEL putative ER retrieval motif. We replaced this amino acid sequence with a $(His)_6$-tag and transfected COS7 cells with an increasing amount of each β subunit, together with a constant level of α1 subunit of glucosidase II (Figure 5). From these cells (Figure 5, lanes 6-8), but not in the control cells (Figure 5, lanes 3-5), glucosidase II could be detected extracellularly by immunoprecipitation and by p-nitrophenyl α-D-glucopyranosidase (p-NP-Glcase) activity

Figure 5. The HDEL/ER-retrieval motif found in the β subunit is responsible for glucosidase II localization. Immunoprecipitations were performed from the medium of [^{35}S]-labeled COS7 cells cultures upon transfections with equal amounts of α subunit DNA and a range of 0 to 0.3 μg of β subunit DNA, with the C-terminal HDEL (lanes 3-5) and with this sequence replaced by a $(His)_6$-motif (lanes 6-8). Medium from COS7 cultures also was tested for *p*-nitrophenyl α-D-glucopyranosidase activity. Empty vector (pcDNA3) was used in both experiments to bring the total amount of DNA used in the transfections to 1 μg.

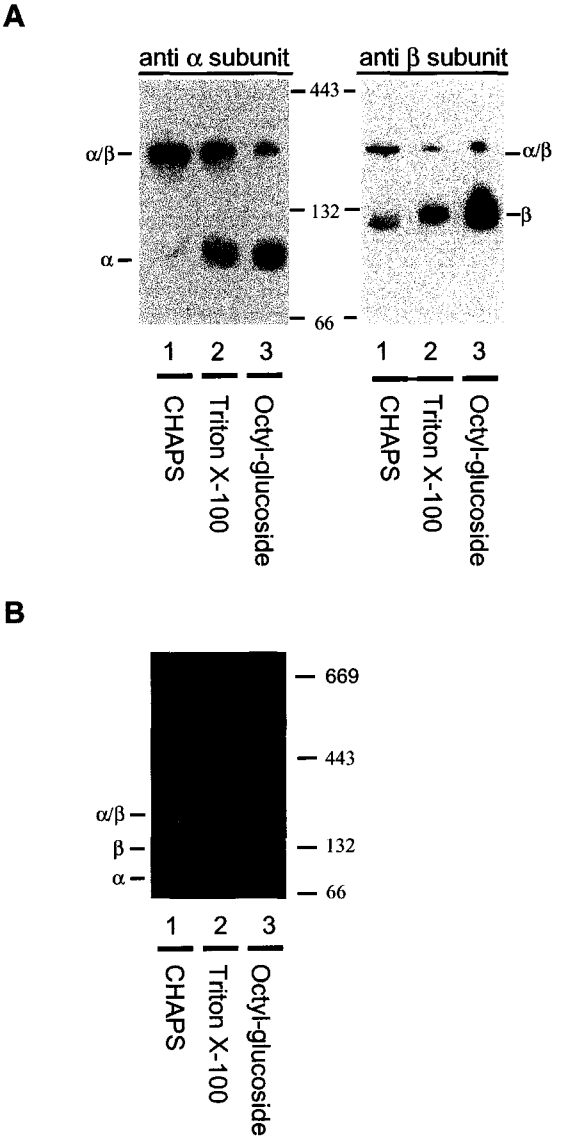

DNA transfected
μg used
α1–sub 0 0.5 0.5 0.5 0.5 0.5 0.5 0.5
β–sub 0 0 0.1 0.2 0.3 0.1 0.2 0.3
pcDNA3 1 0.5 0.4 0.3 0.2 0.4 0.3 0.2
kDa
83
α1–sub
β–sub
0.12
0.08
0.04
0
OD405
1 2 3 4 5 6 7 8
HDEL
HIS tag

Equal loading of the immunoprecipitations can be assessed by the band at 90 kDa which is the result of non-specific protein G binding (Figure 5). This data confirms that the C-terminal HDEL/ER-retrieval motif of the β subunits is required to prevent the secretion of the active enzyme into the extracellular space. Optimal activity was observed with 0.4 μg of β subunit DNA, followed by a decline which may indicate deleterious effects of the overexpression of this subunit (data not shown).

2.4.6 Blue-native electrophoresis separation of glucosidase II shows only heterodimeric structure

Protein complexes can be electrophoretically separated using blue-native polyacrylamide gel electrophoresis (BN-PAGE) (127). Upon solubilization, protein complexes are coated with Coomassie Brilliant Blue G dye which confers negative charge without disrupting protein-protein interactions therein, allowing for their migration and separation in a polyacrylamide gel matrix. Proteins from HeLa cell lysates were solubilized with 2% CHAPS, 1% Triton X-100, or 2% Octyl-glucoside, subjected to BN-PAGE, and transferred to Immobilon-P membrane for Western blot analysis. Using specific antisera for the α and β subunits of glucosidase II, we observed that the α/β heterodimer complex migrated at approximately 250 kD (Figure 6A) while no higher molecular weight complexes were found. This suggests that glucosidase II is not permanently or strongly associated with other components of the ER lumen. Varying amounts of non-complexed α or β subunits were detected depending on the detergent employed. Octyl-glucoside appeared to solubilize the individual subunits most efficiently. Moreover, an in gel glucosidase II assay using the the fluorescent substrate 4-methylumbelliferyl α-D-glucopyranoside confirmed that only the heterodimeric complex exhibited enzymatic activity (Figure 6B). Slow migration of the β subunit might be explained by poor binding of the Coomassie Blue dye described in the work of Trombetta *et al.* (92).

Figure 6. Structural analysis of human glucosidase II by blue native-polyacrylamide gel electrophoresis. (A) Western analysis was performed on HeLa cell extracts made with 2% CHAPS, 1% Triton X-100, or 2% octyl-glucoside (lanes 1,2, and 3, respectiveley). Protein samples were then separated by blue-native PAGE (see Materials and Methods) and transferred to Immobilon-P membrane. Peptide-derived antisera against either the α or β subunit were used as indicated. (B) An *in gel* glucosidase II assay using the fluorescent substrate 4-methylumbelliferyl α-D-glucopyronoside was performed with HeLa extracts prepared and separated as descibed above.


A
anti α subunit
anti β subunit
443
132
66
α/β
α
α/β
β
1 2 3
CHAPS
Triton X-100
Octyl-glucoside
1 2 3
CHAPS
Triton X-100
Octyl-glucoside
B
669
443
132
66
α/β
β
α
1 2 3
CHAPS
Triton X-100
Octyl-glucoside

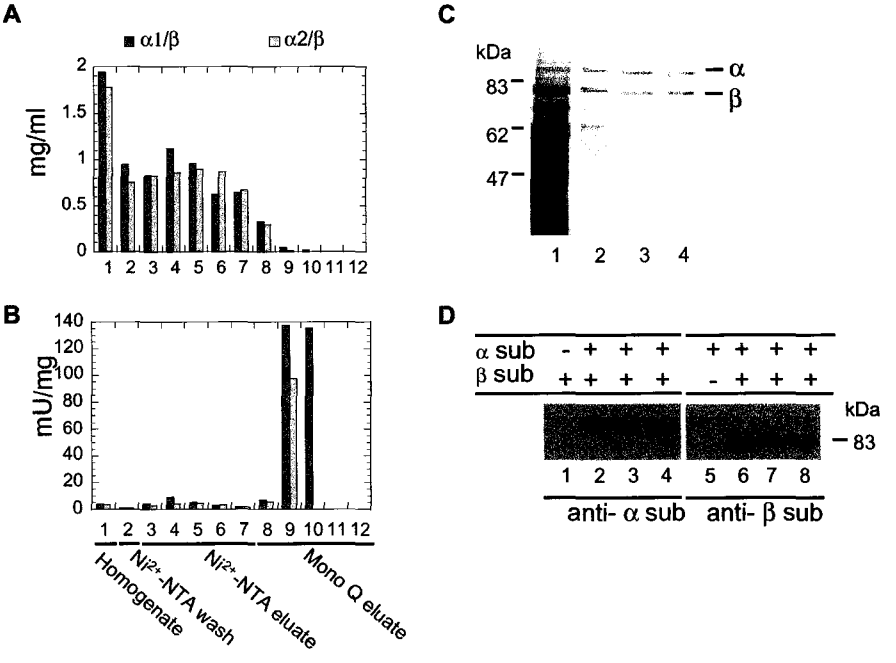
2.4.7 Expression of the Human α and β Subunits of Glucosidase II in Insect Cells, and Purification of the Recombinant Protein Complex

The baculovirus/insect cell expression system was used to produce the α1/β and α2/β isoforms of human recombinant glucosidase II. Initially, we expressed $(His)_6$-tagged subunits individually in *Sf*9 insect cells using recombinant baculovirus. Although the α1 and α2 subunits were expressed at high levels, they were present as inactive insoluble aggregates that could only be partially solubilized in 8 M urea (data not shown). The minor portion of α subunit (<1%) that was soluble and active, remarkably co-purified with β subunit, which is presumably the endogenous one present in *Sf*9 cells. While the recombinant human β subunit was soluble, it co-purified with a putative α subunit from insect cells and the resulting heteroenzyme exhibited *p*-NP-Glcase activity (data not shown). This confirmed that the β subunit is essential for solubility of glucosidase II and demonstrated a surprising functional complementation of the human α and β subunits of glucosidase II by their insect homologues.

Next, either α1 or α2 subunits were simultaneously co-expressed with the β subunit. This co-expression lead to a significant increase in soluble human recombinant α subunit (data not shown). In this case, only the α subunit was $(His)_6$-tagged to allow for separation of the human recombinant glucosidase II from the endogenous insect enzyme. The first step of purification consisted of Ni^{2+}-NTA chromatography. This separation was followed by anion exchange chromatography and contributed to a significant improvement in purification factor (Figure 7A) thus increasing the specific activity of recombinant glucosidase II (Figure 7B). The relative specific activity levels obtained for each isoform throughout purification correlated with the expression levels, which were consistently higher for α1/β than α2/β (Figure 7B).

The α1 subunit (110 kDa) (Figure 7C, lane 4) displayed a slightly higher electrophoretic mobility than the α2 subunit due to the additional 22 amino acid insertion of the latter (Figure 7C, lane 3). Furthermore, the β subunit (80 kDa) (Figure 7C), later determined as human by Western analysis (Figure 7D), co-purified equally well with both isoforms. Neither the α nor β subunits of *Sf*9 glucosidase II cross-reacted with either

Figure 7. Purification profiles of α1/β and α2/β recombinant isoforms of human glucosidase II from insect cells (A) Protein concentrations (mg/ml) of the α1/β (dark) and α2/β (light) isoforms are described for the cell homogenate (lane 1), Ni^{2+}-NTA wash (lane 2) and eluate (lanes 3-7) and fractions near the Mono Q activity peak (lanes 8-12). (B) *p*-nitrophenyl α-D-glucopyranosidase specific activity (mU/mg) for the α1/β and (dark) α2/β (light) isoforms at 5 mM *p*-NP-Glc, in 50 mM PIPES pH 6.5, for 30 minutes at 37°C. One unit of *p*-NP-Glcase activity is defined as the production of 1 μmole *p*-nitrophenol/minute at 37°C. (C) Samples were separated on an 8% SDS-PAGE. The α2/β isoform in the cell homogenate (lane 1) and from the peaks of activity of the Ni^{2+}-NTA (lane 2) and in the Mono Q fractions (lane 3) is shown. Lane 4 shows the Mono Q activity peak from the purification of the α1/β isoform. (D) Western analysis was performed with peptide-derived antisera specific for the human α (lanes 1-4) and β subunits (lanes 5-8). Extracts were made from cells expressing only the β subunit and only the α subunit (lanes 1,5). Represented are cell extracts from co-expression of the α2 subunit with the β subunit (lanes 2,6), activity peaks of the Ni^{2+}-NTA (lanes 3,7) and the mono Q fractions (lanes 4,8) respectively.


A
α1/β
α2/β
mg/ml
2
1.5
1
0.5
0
1 2 3 4 5 6 7 8 9 10 11 12
B
mU/mg
140
120
100
80
60
40
20
0
1 2 3 4 5 6 7 8 9 10 11 12
Homogenate
Ni2+-NTA wash
Ni2+-NTA eluate
Mono Q eluate
C
kDa
83
62
47
α
β
1 2 3 4
D
α sub - + + + + + + +
β sub + + + + - + + +
kDa
83
1 2 3 4 5 6 7 8
anti- α sub
anti- β sub

antisera (Figure 7D, lanes 1 and 5 respectively). This demonstrates the specificity of the peptide antibodies for the human recombinant subunits of glucosidase II and confirms that the human recombinant β subunit was indeed co-purified from co-infected insect cells.

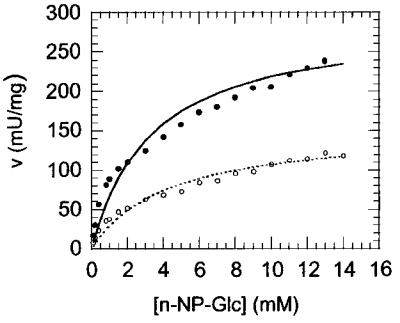
2.4.8 Comparison of the kinetic parameters for *p*-nitrophenyl α-D-glucopyranosidase activity for recombinant glucosidase II isoforms.

Reaction rates of glucosidase II with the substrate *p*-NP-Glc did not obey the simple hyperbolic form of the Michaelis-Menten equation (Figure 8A), but can be described by a double hyperbolic form corresponding to a model that takes into account the two postulated binding sites (Figure 8B) (91). This can also be clearly visualized using an Eadie-Hofstee plot (Figure 8C). These findings agreed with the results of Alonso *et al.* (91) for rat liver glucosidase II activity on *p*-NP-Glc andfor maltase activities (123). The $K_{M1\text{-}2}$ values for both the high and low affinity sites were indistinguishable for both isoforms using the *Student's t*-test (Table 1). The observed K_{M1} for the high affinity sites of α1/β and α2/β were slightly lower at 0.50 and 0.52 mM, than for the K_{M1} of the high affinity site of the rat liver enzyme at 0.78 mM (91). They also approach the rat liver glucosidase II high affinity K_{M1} for maltose at 0.43 mM (123). The K_{M2} for the low affinity sites of α1/β and α2/β were 61.2 and 68.5 mM, which are lower than the reported K_{M2} for rat liver glucosidase II for *p*-NP-Glc at 481 mM (91) but approaches however the K_{M2} for maltose, at 57.7 mM (123).

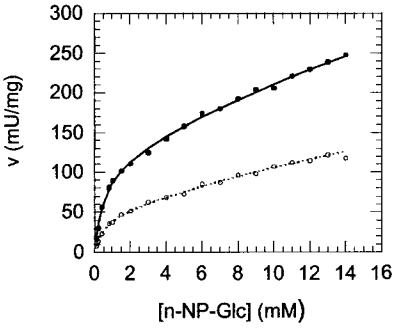
The velocities for both isoforms may reflect the different active enzyme concentrations found in our preparations. With higher specific activity, α1/β yielded $V_{max1\text{-}2}$ of 112 and 744 mU/mg versus 49 and 466 mU/mg for α2/β (Table 1). A difference in enzyme concentration would also translate to a downward shift on the Eadie-Hofstee plot with the shape of the curve remaining unaltered (Figure 8C). Alternatively, the isoforms may have different k_{cat} values which could also lead to a similar downward shift. Nevertheless V_{max}/K_M ratios, which represent the efficiency of the active sites, suggested differences of 18.3 and 13.6 fold for site 1 versus site 2 (Table 1), for α1/β and α2/β respectively, and agreed well with previous results for rat liver glucosidase II which gave an approximately 20 fold difference in efficiency between sites (91).

Figure 8. Kinetic profiles for α1/β and α2/β recombinant isoforms of human glucosidase II. Data for the α1/β (●) and α2/β (○) is presented. (A) Curve fit obtained using a simple single hyperbolic Michaelis-Menten model. (B) Velocity (v, mU/mg) plotted against [*p*-NP-Glc] demonstrates fitting to the double hyberbolic equation for both isoforms. (C) Eadie-Hofstee plots (v x v/[*p*-NP-Glc]) also support the two binding site model for *p*-NP-Glcase activity (91).

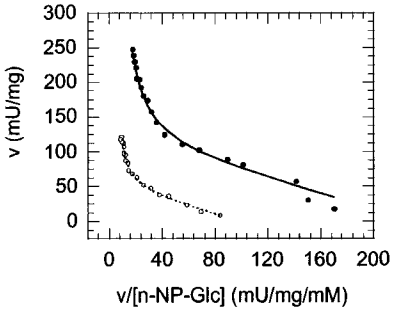
A


v (mU/mg)
300
250
200
150
100
50
0
0 2 4 6 8 10 12 14 16
[n-NP-Glc] (mM)

B


v (mU/mg)
300
250
200
150
100
50
0
0 2 4 6 8 10 12 14 16
[n-NP-Glc] (mM)

C


v (mU/mg)
300
250
200
150
100
50
0
0 40 80 120 160 200
v/[n-NP-Glc] (mU/mg/mM)

2.4.9 Inhibition of p-NP-Glcase activity

Inhibition at both the high and low affinity sites was tested for the two ranges of substrate concentration, 0.25 - 1.0 mM and 8.0-14 mM *p*-NP-Glc respectively. Plots of reciprocal activity versus inhibitor concentration at these ranges suggested similar inhibition of *p*-NP-Glcase activity for both isoforms. We report here $K_{i1\text{-}2}$ values for deoxynojirimycin of 3.0 and 18.8 μM (α1/β) and 2.9 and 22.0 μM (α2/β) for the high and low affinity sites respectively (Table I). Similarly with castanospermine we measured $K_{i1\text{-}2}$ values of 17.0 and 45.3 μM (α1/β), and 15.0 and 39.8 μM (α2/β) for the high/low affinity sites (Table I). These values were indistinguishable between isoforms demonstrating similar inhibition by deoxynojirimycin and castanospermine of *p*-NP-Glcase activity for both isoforms. It should be noted that neither α1/β nor α2/β were affected by EDTA nor EGTA at concentrations up to 10 mM (data not shown). Hence, recombinant glucosidase II activity was not affected by the depletion of cations, including calcium

2.4.10 Glucose hydrolysis from $Glc_{1\text{-}3}Man_9GlcNAc_1$ oligosaccharides by α1/β and α2/β forms of glucosidase II

Glucosidase II assays were performed for both α1/β and α2/β forms of recombinant enzyme after normalizing for *p*-NP-Glcase activity. [^{3}H]-Mannose labeled lipid-linked $Glc_{1\text{-}3}Man_9GlcNAc_1$ oligosaccharides were extracted, purified and pooled from *S. cerevisiae* Δ*alg8*Δ*gls2*, Δ*alg10*Δ*gls2*, and Δ*gls2* mutants (58) and used as substrates. The extent of hydrolysis was determined by following changes in HPLC elution profiles (Figure 9). While the $Glc_3Man_9GlcNAc_1$ oligosaccharide remained uncleaved after 30 minutes, glucose hydrolysis from both Glc_2- and $Glc_1Man_9GlcNAc_1$ was shown for both enzyme isoforms. Hydrolysis from the $Glc_2Man_9GlcNAc_1$ was observed as a downward shift for the Glc_2 peak with a concomitant rise in the Glc_1 peak. Moreover, the rise in the $Man_9GlcNAc_1$ demonstrated the subsequent cleavage of the $Glc_1Man_9GlcNAc_1$ oligosaccharides. The profiles for both forms of recombinant glucosidase II not only suggested that they share substrate specificities, but with similar vectorial changes in peak height comfirm that they are catalytically indistinguishable for the substrates examined.

TABLE I

Kinetic parameters and inhibition of p-nitrophenyl α-D-glucopyranosidase activity for the α1/β and α2/β recombinant isoforms of human glucosidase II

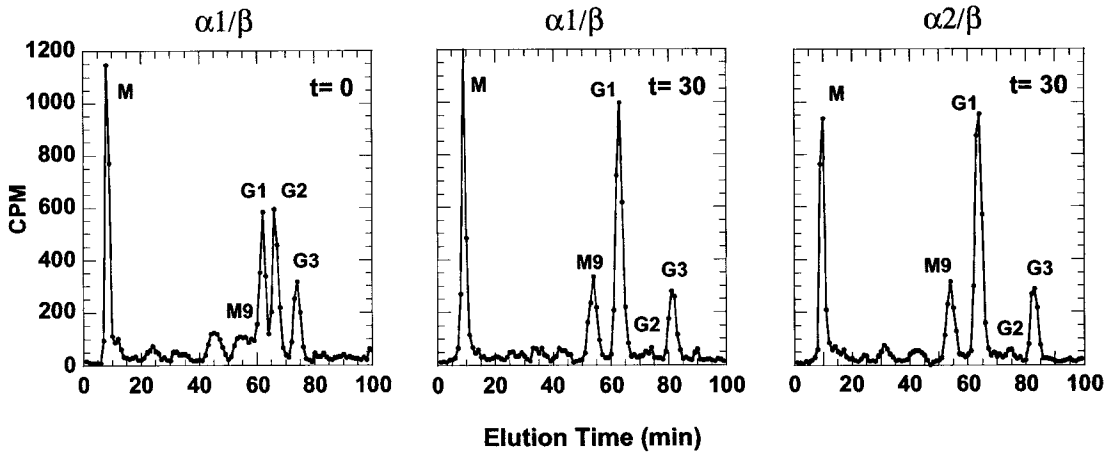
Curve fit was performed to the Michaelis-Menten equation modified for a two binding site model. App. K_i were determined from plots of reciprocal activity versus inhibitor concentration, for the high and low affinity binding sites, at substrate concentration ranges of 0.25 - 1.0 mM and 8.0-14 mM *p*-NP-Glc, respectively. Values are means ± S.E.M.

Isoform	Binding site	K_M (mM)	Vmax (mU[a]/mg)	$Vmax/K_M$ (ml/min per g)	Site1/Site2	DNJ[b] (μM app. K_i)	CST (μM app. K_i)
α1/β	**1 (high affinity)**	0.50 ± 0.05	112 ± 4	222 ± 24	18.3 ± 5.0	3.0 ± 0.3	17.0 ± 0.8
	2 (low affinity)	61.2 ± 12.8	744 ± 113	12.1 ± 3.1		18.8 ± 1.3	45.3 ± 3.5
α2/β	**1 (high affinity)**	0.52 ± 0.06	49 ± 2	92 ± 11	13.6 ± 3.5	2.9 ± 0.7	15.0 ± 1.5
	2 (low affinity)	68.5 ± 12.2	466 ± 65	6.8 ± 1.5		22.0 ± 1.4	39.8 ± 1.2

[a]One unit of activity is defined as 1 μmole of nitrophenol formed/min, at 37°C.

[b]Abbreviations: DNJ, deoxynojirimycin; CST, castanospermine

Figure 9. Comparison of α1/β and α2/β glucosidase II activity on physiological substrates. Glucosidase II assays were performed on [^{3}H]$Glc_{1-3}Man_9GlcNAc_1$ oligosaccharides as described in the Experimental Procedures. HPLC profiles of oligosaccharides after hydrolysis for 0 and 30 minutes are shown for the α1/β glucosidase II, and 30 minutes for the α2/β form. The positions of mannose (M), $Man_9GlcNAc_1$ (M9), $Glc_1Man_9GlcNAc_1$ (G1), $Glc_2Man_9GlcNAc_1$ (G2), $Glc_3Man_9GlcNAc_1$ (G3), are indicated.


α1/β
α1/β
α2/β
CPM
1200
1000
800
600
400
200
0
M
t= 0
G1
G2
G3
M9
t= 30
0
20
40
60
80
100
Elution Time (min)

2.5 Discussion

Our results demonstrated the expression of two forms of human glucosidase II *in vivo*. We have identified two α/β complexes through Western analysis and immunoprecipitations and also isolated two forms of the α subunit that differed by the inclusion of a 66 bp stretch. Brada and co-workers (120) were first to suggest that there might be two forms of glucosidase II based on different subcellular localization and also different enzyme species carrying Endo H-sensitive high-mannose, as well as sialylated Endo H-resistant oligosaccharide chains in pig kidney epithelial cells. Later, Hentges and Bause (95) demonstrated the purification of a tight glucosidase II α subunit doublet estimated at 107/112 kDa, with the 107 kDa more abundant form ascribed as a cross-reacting degradation product. However, it was Arendt and Ostergaard (118) that identified a murine α subunit cDNA highly conserved with the human cDNA (D42041) that also encoded for an additional 22 amino acids. In more recent studies, the same authors ascribed the domain as a splice variant termed box A1 (66 bp) and also identified a second domain by PCR, referred to as A2 (27 bp) (89). Complementary DNAs with three of four possible combinations were isolated; i.e. $A1^+$ $A2^-$ (form1), $A1^-$ $A2^+$ (form 2), and $A1^-$ $A2^-$ (form 3) while the existence of a cDNA encoding an isoform with both A1 and A2 has yet been confirmed (89). It is possible, based on a comparison with mouse splice variants, that the longer α subunit observed by Western analysis and by immunoprecipitation corresponds to the cloned α2 subunit which could be of the $A1^+$ $A2^-$ form. This assumption rests on the similar migration of the observed long α subunit with the recombinant α2 subunit (data not shown) and the respective difficulty of resolving the alternative $A1^-$ $A2^+$ splice variant on a SDS-PAGE. Nevertheless, other splice variants may exist that have not yet been identified that encode for proteins with molecular weights similar to the α2 subunit.

Different levels of expression have previously been observed for the α subunit isoforms (89,95). Hentges and Bause (95), in the purification of pig liver glucosidase II, found lower yields for the slower mobility form in their 107/112 kDa α subunit doulet. Also, cloning experiments by Arendt *et al.* (89) suggest a five fold lower expression of the form 1 ($A1^+$ $A2^-$) than form 3 ($A1^-$ $A2^-$) using a PCR cloning strategy. We also found lower yields for α2 subunit, when expressed in COS7 cells (Figure 3) and *Sf*9 insect cells

(Figure 7B). While factors that lead to differential expression of the α subunits remain unknown, protein complex stability may be ruled out since purified recombinant α1/β and α2/β display equal stability in continuous enzymatic assays, with linear progress curves proceeding for close to one hour (data not shown).

These studies also sought to clarify the role of the β subunit in the expression of active human glucosidase II in mammalian cells. Trombetta *et al.* (92) first characterized glucosidase II as a heterodimer. They noted that the lower molecular weight subunit could not be removed while maintaining glucosidase II activity. More recently, D'Alessio *et al.* (93) genetically demonstrated the heterodimeric structure of *Schizosaccharomyces pombe* glucosidase II through the loss of endogenous glucosidase II activity in cells disrupted for GII_{β}. It was postulated in both studies that the β subunit was involved in maintaining the integrity of catalytically active glucosidase II and the ER localization though its HDEL-ER retrieval signal. Flura *et al.* (94) on the other hand observed a slight increase in glucosidase II activity in CHO cells upon the lone expression of a pig liver derived cDNA encoding the α subunit. Moreover, while Hentges and Bause (95) identified a 60 kDa component, it was undetectable by Coomassie staining in some of their preparations. We found however that the α1 or α2 subunits were largely insoluble when expressed alone but this could be counteracted through the co-expression of the β subunit, which concomitantly produced active glucosidase II. As well, we demonstrate here the requirement of heterodimeric structure for enzyme activity. Thus, the β subunit contributes to the solubility/stability of the α/β complex and is a prerequisite for the formation of the active enzyme. This rules out the hypothesis that the β subunit is simply required for α subunit folding and that the α subunit then remains active upon dissociation. Moreover, the detection of active glucosidase II in the culture medium upon the replacement of the C-terminal HDEL sequon with a $(His)_6$-tag, establishes a role of the β subunit in retaining the α/β glucosidase II dimer in the ER. It is possible however that the variable-fold increase in enzyme activity observed by Flura *et al.* (94) may be attributable to heterodimer assembly with endogenously expressed β subunit. While abundantly expressed, the level of endogenous β subunit does in fact vary according to cell line, as observed in our study (Figure 2B).

Heterologous expression of the α subunit in *Sf*9 insect cells confirmed these functions of the β subunit, but also by the formation of active human/insect heteroenzymes remarkably demonstrates the wide conservation of its properties. While the yields of α subunit were high when expressed alone, most of the recombinant protein was insoluble. The small fraction of α subunit that was soluble always co-purified with an insect protein with a molecular weight similar to that of the human β subunit. Conversely, purified recombinant human β subunit, when expressed alone, had an associated insect protein at approximately 110 kDa. These preparations also displayed *p*-NP-Glcase activity characteristic of glucosidase II. Protein alignments of the human β subunit with homologues in *Caenorhabditis elegans* (Z47356) and *S. pombe* (D89245) suggest that it is a highly conserved protein. Both have putative ER targeting and retrieval signals, and overall homologies of 36% and 22% respectively, particularly concentrated in the N-terminal region (Q50-E105). This region of high conservation, with 71% and 79% for *C. elegans* and *S. pombe* respectively, may play a role in mediating α/β interactions. Interestingly, the open reading frame YDR221w in *S. cerevisiae* (S59428) with an homology of 21% has both a putative ER targeting signal peptide and a transmembrane domain. Moreover, most of its homology is found within the conserved domain (61%). This open reading frame probably encodes a likely candidate for the glucosidase II β subunit for this species. While it remains to be demonstrated, we propose that this highly conserved region acts as the major α/β protein-protein interaction domain.

One hypothesis for the existence of the α1and α2 forms was that they each may display different specificities for the $Glc_2Man_9GlcNAc_1$ and $Glc_1Man_9GlcNAc_1$ substrates. The current kinetic model for *p*-NP-Glc and maltose, which predicts both high and low affinity substrate binding sites (91) could have easily been derived from the co-purification of two isoforms of glucosidase II with such substrate specificities. The first evidence of this differential activity was demonstrated by Hubbard and Robbins (60). They showed that the sequential hydrolysis of glucose from the N-linked oligosaccharides occured at different rates. Alonso *et al.* (124) later assigned the cleavage of the innermost α-1,3-linked glucose ($Glc_1Man_9GlcNAc_2$) to the low affinity site since bromoconduritol, which selectively inhibited both *p*-NP-Glc and maltose binding to this site, had also been shown to inhibit glucose hydrolysis from $Glc_1Man_9GlcNAc_2$

(124,125). However, we have found, through the expression and purification of both recombinant α1/β and α2/β complexes of glucosidase II and through the determination of their kinetic parameters for the *p*-NP-Glc substrate, that both forms displayed activity that fit well to the two binding site model proposed by Alonso *et al.* (91). Hence, the kinetics displayed are truly representative of an intricate active site and not the result of co-purified isoforms with different substrate affinities. Furthermore, we have shown that the α1/β and α2/β forms of glucosidase II share similar properties of *p*-NP-Glcase hydrolysis and inhibition while exhibiting similar substrate specificities for the physiological oligosaccharide substrates.

A role for two or more isoforms of glucosidase II still remains to be determined. However, as an early step in glycan processing of nascent glycoproteins and the entry point into the Calnexin Cycle, glucosidase II is widely used by most, if not all secreted glycoproteins that may or may not require ER quality control. It is not without precedent to find biological complexity in such a pivotal step. And while glucosidase II was not found complexed with other ER constituents, transient interactions with the translocon, ER chaperones or even proteins of the ER matrix may exist that are isoform specific. Differential regulation cannot be ruled out and may be mediated by yet unknown interactions. While it remains to be determined, any new found complexity in the early steps of glycan trimming may provide us with novel and very specific anti-viral strategies, as inhibitors of the ER α-glucosidases have already shown great promise against viruses that require the host ER lectin-like chaperone protein folding apparatus (111).

Chapter 3

Specific interactions of the ER proteins ERp57, calnexin, and calreticulin, established with an ER specific yeast two-hybrid system

3.1 Abstract

Proteins entering the endoplasmic reticulum (ER) become N-glycosylated at specific sites. This glycosylation is linked to protein folding through the action of the lectin-type molecular chaperone calnexin (25). Glycoproteins bound to calnexin are specific substrates for one member of the protein disulfide isomerase family, ERp57. There is biochemical evidence of a physical interaction between ERp57 and calnexin (25,133), however genetic methods for investigating the interactions of ER proteins have been hampered by the unique environment of the ER of high protein concentrations, high calcium content, and redox state not found in the nucleus, where conventional two-hybrid interactions occur. We have developed an ER protein two-hybrid system based on the Ire1p unfolded protein response signaling kinase. We have used this system to map the regions of protein-protein interaction of calnexin and its soluble homologue calreticulin with ERp57, and to show that protein disulfide isomerase (PDI) does not interact. The system was used to define domains of calnexin/calreticulin and ERp57 that interact and the interaction of these fragments was verified by their physical association *in vitro*. The ER protein two-hybrid system will have a wide application in an area where there has been a paucity of genetic tools.

3.2 Introduction

The elegant simplicity of the two-hybrid system, as developed originally and its subsequent variations, links protein interactions to changes in transcription. These systems have detected a plethora of intracellular protein interactions (134-138), and many of these predicted interactions have been authenticated by biochemical techniques, such as co-immunoprecipitation. The two-hybrid system in addition to permitting the rapid identification of binding partners, can also be used to define residues critical to the interaction, and more recently the construction of maps of large scale protein interaction networks (138-141).

The two-hybrid system has been less used for membrane proteins, but some systems have become available. These include the split-ubiquitin system (136,142), the SOS and Ras recruitment systems (SRS and RRS) (137,143), and the G-protein fusion system (144). These techniques have broadened the spectrum of proteins that can be

analyzed to include integral membrane and membrane-associated proteins which are topologically restricted from the nucleus (142-144).

We report here, and demonstrate the reliability and utility thereof, a new two-hybrid system that detects protein interactions occurring within the lumen of the endoplasmic reticulum (ER). We have used this ER two-hybrid system to map the interactions between the proteins that constitute part the glycoprotein folding complex within the ER.

3.3 Materials and methods

3.3.1 Manipulations, yeast strains, and plasmids

Standard protocols were used for yeast growth and transformations (148). The reporter strain was constructed by integrating *UPR-Y::CYC1::LacZ*, derived from pLG-Δ178 (149) kindly provided by Claude A. Jakob (Zurich, Switzerland) into the *ura3-52* locus of W303a (MAT**a**; *ura3-52*; *trp1*; *leu2*; *his3*; *ade2*; *can1-100*). This strain was then crossed to BY4742 Matα *ΔIRE1::KanMX* (ATCC # 4011907). The diploids were sporulated, and haploids of both mating types were isolated (yLJ29 MAT**a**; *trp1*; *leu2*; *his3; ΔIRE1::KanMX; ura3-52::UPR-Y::LacZ-URA3*, and yLJ31 MATα *ura3*; *trp1*; *leu2*; *his3; ade2; ΔIRE1::KanMX; UPR-Y::CYC1::LacZ*). Unless otherwise mentioned, all plasmids were made by recombinational cloning directly in yeast as described elswhere (139). The parent plasmid pLJ89 was made in three steps: (1) a 2μ yeast plasmid with a *LEU2* marker (pGreg505) was linearized and recombined with a 477 bp PCR product (amplified using Expand High Fidelity™ system (Roche, Laval, Canada)) that spanned from -411 to +66 bp coding region of *IRE1* (i.e. *IRE1* promoter and ER signal sequence). (2) A 66 bp linker that has both *Not*I and *Sal*I sites was added 3' to the sequence encoding the ER signal sequence, without disrupting the reading frame. (3) The region encoding the transmembrane domain, kinase, and endoribonuclease was amplified and was similarly introduced (Figure 11A). To produce pLJ96, the *IRE1* cassette was removed from pLJ89 by *Pme*I and *Xba*I, (New England Biolabs, Mississauga, Canada) and subcloned into pGreg503, that contains a *HIS3* marker. All constructs used for testing protein interactions were derived from pLJ89 and pLJ96. They were linearized with *Sal*I,

thegenes of interest amplified by high fidelity PCR with, then recombined into pLJ89 or pJ96, in the strains yLJ29 or yLJ31.

3.3.2 β-galactosidase assays

Filter assays were performed on strains that were grown for two days, mated on YPD for 12 hours, and diploids isolated and grown on synthetic dropout medium (-Ura -Leu -His) for 24 to 48 hours. The cell patches were transferred to nitrocellulose membrane, dipped in liquid N_2 for 5 seconds, and then placed on Whatman filter paper imbibed with Z buffer (60 mM $Na_2HPO_4 \cdot 7H2O$, 40 mM $NaH_2PO_4 \cdot H2O$, 10 mM KCl, 1 mM MgSO4·7H2O, 50 mM β-mercaptoethanol, pH7.0) with 4 mg/ml X-Gal. The filters were then incubated for 30-60 minutes at 30°C. Quantitative β-galactosidase assays were performed as described elsewhere (150) and repeated three times, each with n=3.

3.3.3 Production of GST fusion proteins and GST-pull down assay

Plasmids encoding GST-$CNX_{K46\text{-}M417}$ and GST- $CNX_{M267\text{-}L212}$ were a kind gift from Robert Larocque (Montreal, Canada). Expression and purification of the GST fusions were performed as previously described (151). Approximately 2 μg of GST-$CNX_{K46\text{-}M417}$, GST- $CNX_{M267\text{-}L212}$, or GST were loaded onto a column with 200 μl of Glutathione Sepharose 4B (Amersham Pharmacia Biotech, Piscataway, NJ) pre-equilibrated with TBS 1% Triton X-100 pH 7.0. The resin was washed with 5 ml of the same buffer, and 5 μg of ERp57 was loaded onto the resin. The column was washed again with 10 ml of TBS, and then the proteins were eluted in 100 μl of 10mM reduced glutathione. Western analysis was performed as described elsewhere (90), with rabbit polyclonal anti-ERp57 antiserum.

3.4 Results/Discussion

3.4.1 The Calnexin/Calreticulin Cycle and the Model for the ER protein two-hybrid system

The glycoprotein folding complex of the ER complex links protein folding and their N-glycosylation, through the lectin-like chaperones calnexin (CNX) and calreticulin (CRT). We have shown these to functionally interact with ERp57 (25,133), a member of

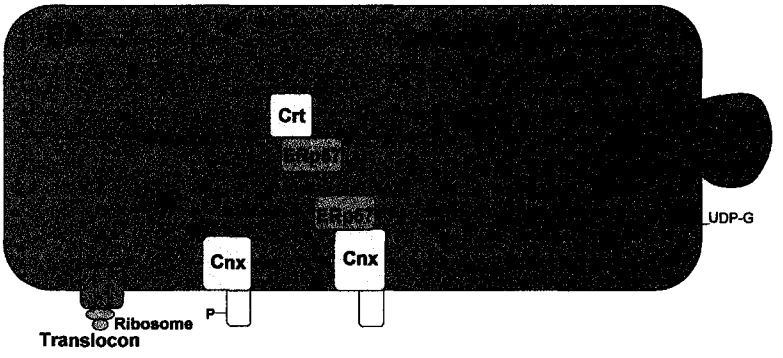
the protein disulfide isomerase family. These associations (CNX/ERp57 and CRT/ERp57) assist in the folding and oligomeric assembly of nascent glycoproteins, and are constituents of a quality control apparatus that ensures their correct folding and/or multimeric assembly prior to their progression through the secretory pathway (145,146). Binding of glycoproteins to CNX or CRT is initiated as nascent glycoproteins are translocated into the lumen of the ER, where high mannose oligosaccharides ($Glc_3Man_9GlcNac_2$) are transferred by the oligosaccharyl transferase complex (OST) from the dolichol-pyrophosphate precursor to the asparagine residues within the amino acid sequon N-X-S/T. These glycans are then rapidly trimmed to the monoglucosylated form ($Glc_1Man_9GlcNAc_2$) by glucosidase I and II, and then more slowly to $Man_9GlcNAc_2$ (147). Glycoproteins bearing $Glc_1Man_9GlcNAc_2$ glycans bind specifically to CNX (17) or to CRT (18), and their release is catalysed by glucosidase II, which removes the remaining glucose. If the native conformation or oligomeric assembly of the glycoprotein has not been achieved, the enzyme UDP-glucose:glycoprotein glycosyltransferase (UGGT) reglucosylates the $Man_9GlcNAc_2$ oligosaccharide, and the glycoprotein rebinds to either CNX/ERp57 or CRT/ERp57 complex. This process has been termed "the Calnexin/Calreticulin Cycle" (102,145,146) (Figure 10A).

Our ER protein interaction system uses the unfolded protein response (UPR) in *Saccharomyces cerevisiae*, to signal the interaction of proteins in the ER (Figure 10B). In wild-type cells this response is initiated by the type I membrane protein Ire1p, that has its N-terminus in the lumen of the ER, and senses and dimerizes in response to the accumulation of misfolded proteins within the ER (66). The Ire1p kinase domains in the cytosol auto-*trans*phosphorylate, and activate an endoribonuclease domain (66) that specifically cleaves the mRNA of *HAC1*. Hac1p is a bZIP transcriptional activator that binds the UPR element (UPRE) upstream of the genes that encode many ER proteins. The intron within *HAC1* mRNA has been shown to attenuate the translation of this mRNA (152), and its removal leads to the upregulation of the expression of many molecular chaperones, and folding enzymes of the ER, as well as many components of the ER associated degradation pathway (ERAD) (153).

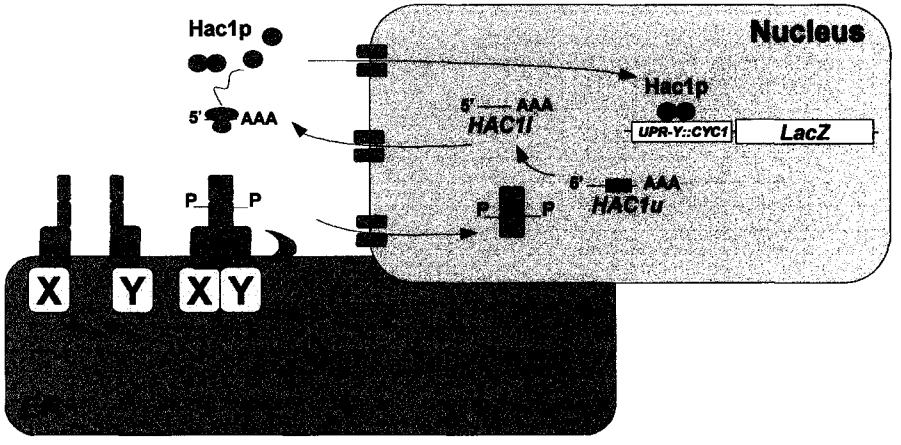
We have used some elements of the UPR to construct a system that reports protein interactions in the ER. To use this system, we first integrated a cassette with the

Figure 10. Models of the "Calnexin Cycle" (A), and for the ER protein two-hybrid system (B). (A) As nascent glycoproteins with N-linked oligosaccharides enter the secretory pathway, they undergo oligosaccharide trimming by the ER glucosidases I and II (Gluc I and II). Trimming to the $Glc_1Man_9GlcNAc_2$ form permits binding to the lectin-like chaperones calnexin (CNX) and calreticulin (CRT). Their associated protein disulfide isomerase (ERp57) then catalyzes protein folding of glycoprotein substrates bound to CNX and CRT. Gluc II mediates the release of substrates from the glycoprotein folding system, and if native conformation or proper oligomeric assembly is not achieved, the protein folding sensor, termed UGGT, reglucosylates the unfolded protein for further binding to the lectin-like chaperones. (B) Fusions (X and Y) with the transmembrane kinase/endonuclease Ire1p that lead to oligomerization, induce auto-*trans*phosphorylation of the Ire1p kinase domain, nuclear targeting of the Ire1p endoribonuclease domain, which then processes the mRNA of the UPR transcriptional activator, Hac1p. With the intron removed, Hac1p is expressed as the active form ($Hac1p_i$) which binds to the minimal yeast UPRE upstream of *LacZ*.

A


Crt
UDP-G
Cnx
Cnx
P
Ribosome
Translocon

B


Hac1p
Nucleus
5'
AAA
Hac1p
5'
AAA
HAC1i
UPR-Y::CYC1
LacZ
6'
AAA
HAC1u
P
P
P
P
X
Y
X
Y

LacZ gene from *E. coli* under the control of a chimeric promoter into *Saccharomyces cerevisiae* (W303a). This promoter consists of a minimal yeast UPRE fused to a truncated *CYC1* promoter. The reporter strain was then created by crossing this strain with BY4742 (*Δire1*), and haploids were isolated that carried both the integrated reporter with the *Δire1* genotype. This strain cannot respond with a UPR to agents that induce this in the wild-type (results not shown).

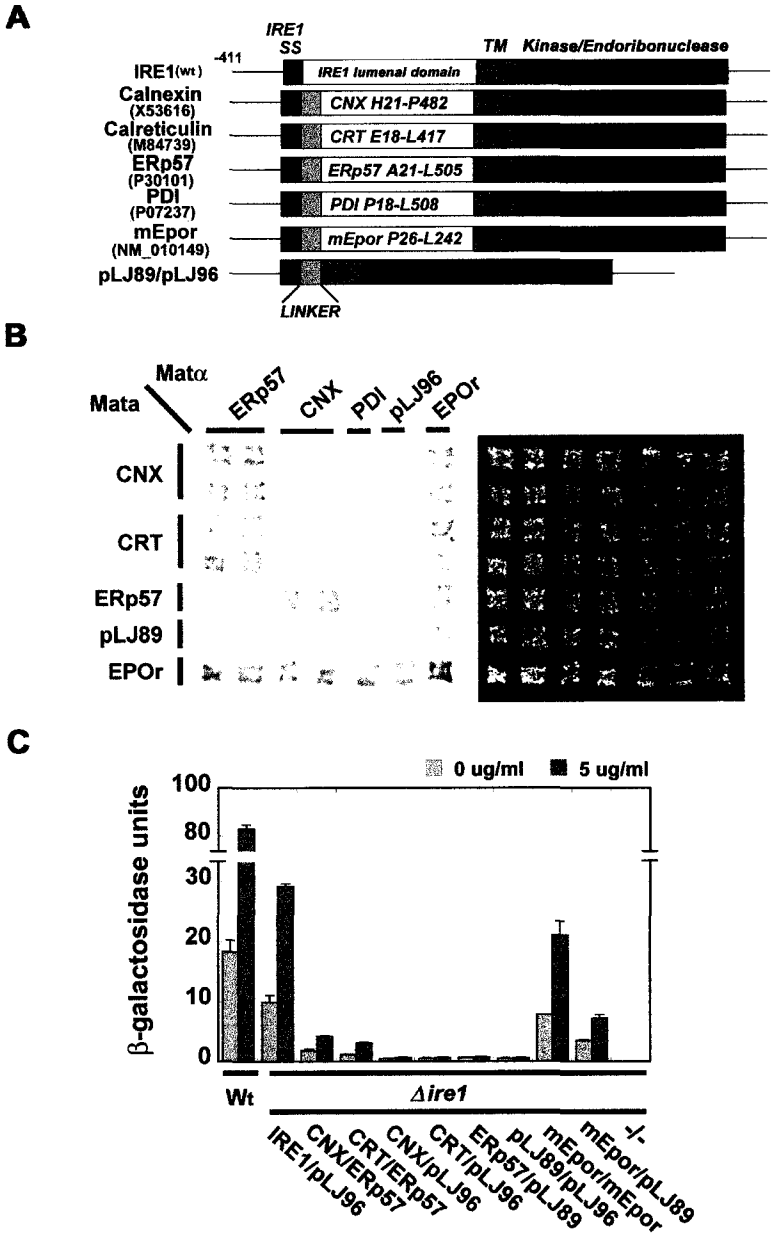
The plasmids pLJ89 and pLJ96 (Figure 11A), with *LEU2* and *HIS3* markers respectively, encode Ire1p with the N-terminal lumenal domain deleted up to the transmembrane domain (TM). The Ire1p signal sequence is intact and followed by a 23 amino acid linker to ensure that the signal peptidase does not cleave the nascent protein subcloned in place of the Ire1p ER lumenal domain. The genes to be inserted were amplified by high fidelity PCR using primers with plasmid homology regions, and subcloned directly into yeast by recombinational cloning between the homologous regions of the linker and the TM domains. The endogenous promoter of *IRE1* was subcloned and used to express the fusion proteins since overexpression has been shown, at least for the mammalian homologue, to induce UPR (74).

3.4.2 Method validation and the ER specific interactions between ERp57 and calnexin/calreticulin

To test our approach, MAT**a** strains expressing a first set of Ire1p fusions were crossed with MATα strains expressing a second set of fusions (Figure 11B). The first set consisted of fusions containing the lumenal domain of CNX, complete CRT, complete ERp57, the parental vector pLJ89 (negative control), or the extracellular domain of the murine erythropoietin receptor (EPOr), as a positive control, respectively (Figure 11B). The second set of fusions were ERp57, the lumenal domain of CNX, the ER protein disulfide isomerase (PDI), the parental vector pLJ96 (negative control), or again the extracellular domain the EPOr (positive control), respectively (Figure 11B). Diploid strains expressing both sets of fusion proteins were isolated, transferred to nitrocellulose, and tested for β-galactosidase activity (Figure 11B).

We have shown previously that ERp57 functionally interacts with CNX and CRT to specifically promote folding of bound glycoproteins ERp57 (25). The homologous ER

Figure 11. Interactions between calnexin and calreticulin, with ERp57, using the Ire1p based two-hybrid system, and functional complementation of the lumenal domain of Ire1p, by fusions with Ire1p that lead to mono/heterodimerization. (A) Schematic diagram of the constructs used to test the ER protein two-hybrid system are shown. The parental vectors pLJ89 (*LEU2*) and pLJ96 (*HIS3*) encode the Ire1p signal sequence (SS), a 23 aa linker in frame with the Ire1p transmembrane (TM) and kinase/endoribonuclease domains, under the control the *IRE1* promoter. Wild-type *IRE1* was also subcloned into the plasmid that was parental to pLJ89 (Wt). Clone accession numbers are indicated in parenthesis. (B) A β-galactosidase filter assay was performed on *S. cerevisiae* diploid strains expressing Ire1p fusion proteins. Yeast cell patches prior to transfer are shown on right. Crosses were made with strains carrying pLJ89 and pLJ96 as negative controls, while the extracellular domain of the murine erythropoietin receptor (EPOr) served as a positive control. (C) Quantitative permeable cell/β-galactosidase assays were performed on yeast cells in the absence or presence of 5 μg/ml of tunicamycin for 1 h. β-galactosidase activity is reported in the absence of tunicamycin (lightly shaded bars) and for treated cells (dark bars). Plasmid combinations are indicated, and the parental W303a strain (with the integrated UPR reporter) was used as a positive control (Wt). β-galactosidase units are defined as $[A_{420} \times 1000]/[A_{600}$ of cells x culture vol. (ml) x reaction time (min)], and error bars represent standard deviation, $n = 3$.


A
IRE1
SS
TM
Kinase/Endoribonuclease
-411
IRE1(wt)
IRE1 lumenal domain
Calnexin
(X53616)
CNX H21-P482
Calreticulin
(M84739)
CRT E18-L417
ERp57
(P30101)
ERp57 A21-L505
PDI
(P07237)
PDI P18-L508
mEpor
(NM_010149)
mEpor P26-L242
pLJ89/pLJ96
LINKER
B
Matα
Mata
ERp57
CNX
PDI
pLJ96
EPOr
CNX
CRT
ERp57
pLJ89
EPOr
C
0 ug/ml
5 ug/ml
β-galactosidase units
100
80
30
20
10
0
Wt
Δire1
IRE1/pLJ96
CNX/ERp57
CRT/ERp57
CNX/pLJ96
CRT/pLJ96
ERp57/pLJ96
pLJ89/pLJ96
mEpor/mEpor
mEpor/pLJ89
-/-

protein PDI does not functionally interact in this way. Also, it has been shown that ERp57 but not PDI can be cross-linked to CNX and CRT (133), implying a physical as well as a functional interaction. With these proteins in the ER protein two-hybrid system, we show that this specificity is maintained. That is we detect the interaction of CNX and CRT with ERp57 and not with PDI (Figure 11B). Also, when CNX and ERp57 were subcloned into the plasmids pLJ89 and pLJ96 in a reciprocal combination, the interaction was still observed (Figure 11B). From these results, we also conclude that CNX does not interact with CRT, nor does ERp57 interact with PDI. In addition, we did not find homodimerization for any of these fusion proteins (Figure 11B). We used the extracellular domain of the murine erythropoietin receptor (EPOr) as a positive control, as it does mediate homodimerization of this receptor in mammalian cells. It has been shown for the EPOr that the extracellular domains are already dimerized and that ligand binding causes a change in the orientation of the cytosolic JAK-2 kinase domains, which then become activated through a *trans*-phosporylation event (154-156). Thus this interaction is seen in the ER protein two-hybrid system as *LacZ* positives in the diploid cells (Figure 11B and C) and also in haploid cells (results not shown).

3.4.3 Confirmation of ER topology through the functional complementation of Ire1p

A ligand-independent activation model for Ire1p was recently proposed by Liu *et al.* (157), who found that a leucine zipper homodimerization domain could replace the lumenal domain of Ire1p, and the fusion protein could also sense the accumulation of unfolded proteins within the ER. We tested whether this was also true for heterodimerizing fusions of Ire1p. We performed quantitative β-galactosidase assays on cells from the previous experiment (Figure 11C). To induce the UPR, cells were grown in the presence of 5 μg/ml of the glycosylation inhibitor tunicamycin. In both cases, either homo- and heterodimerization, the UPR was activated, when compared to the control strains (Figure 11C). This supports the finding that the response to unfolded proteins does not specifically require the Ire1p lumenal domain (157). It also suggests that fusions that lead to both homo- and heterodimerization will functionally complement Ire1p, and also that our fusions are correctly localized.

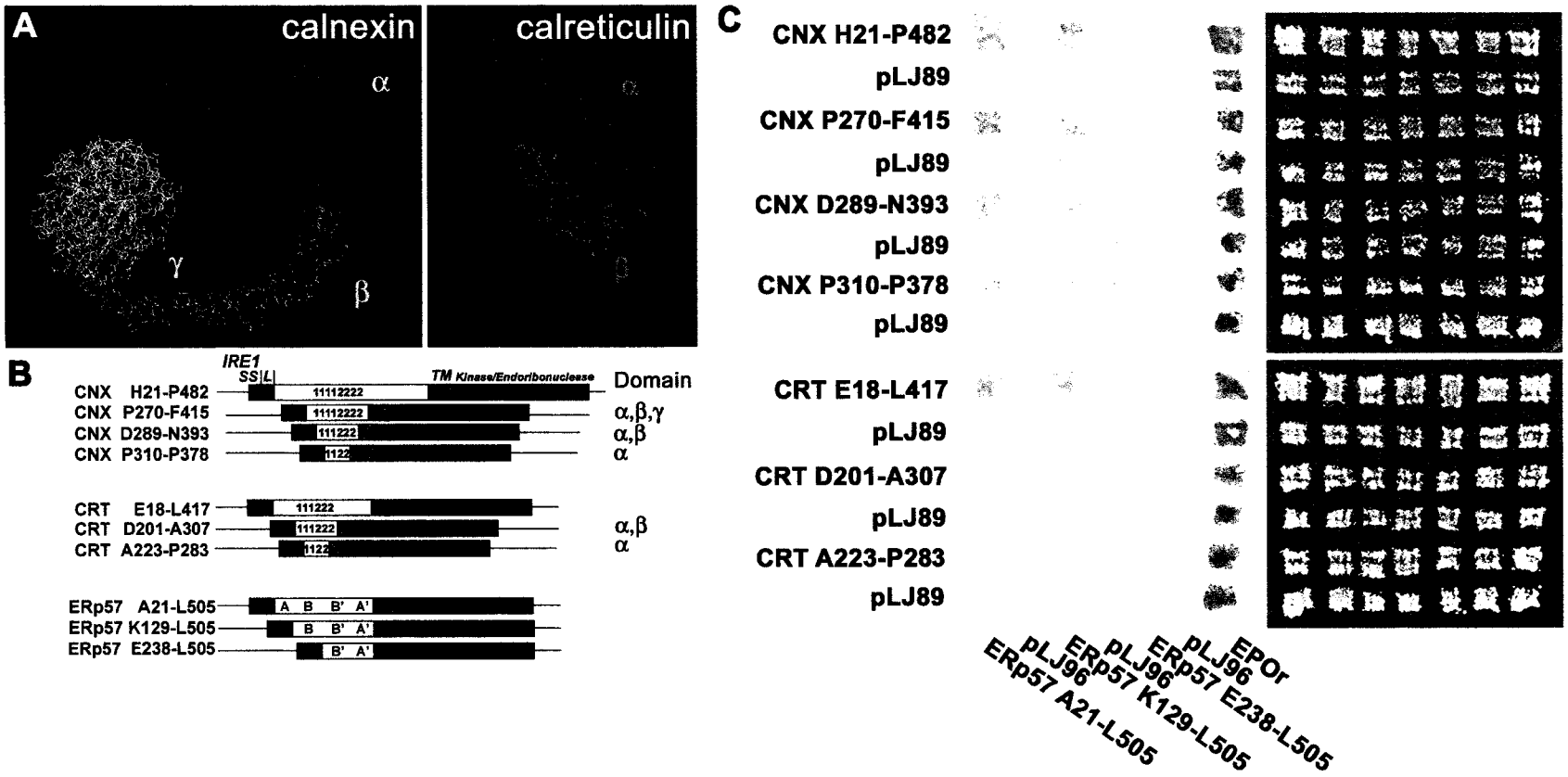
Interestingly, a dose dependent response was seen for the diploid strains (and haploid strains, results not shown) expressing the murine EPOr, with a level of response similar to the strain expressing Ire1p (Figure 11C). However, the tunicamycin induced UPR in the strain with the plasmid borne Ire1p is weaker than found in the wild-type W303a strain (Figure 11C). To account for this difference, it is possible that 411 bp 5' of the ATG codon of *IRE1* is an incomplete promoter, insufficient for the expression of a full complement of *IRE1*. To confirm that these parental plasmids target the fusion proteins correctly to the ER, we subcloned GFP as the lumenal domain of Ire1p and placed this fusion under the control of much stronger *GAL1* promoter, and even at high expression levels the fusion protein displayed a perinuclear localization typical of ER membrane proteins in *S. cerevisiae* (data not shown).

3.4.4 Mapping the protein interactions between ERp57 and calnexin/calreticulin

Next, we mapped the protein interactions between CNX and CRT, with ERp57. We based our constructs on the structure of CNX that has recently been solved at a 2.9 Å resolution (Figure 12A) (96). CNX is characterized by two main structural components, a globular lectin domain and an extended region or loop domain (P-domain (4)). This proline-rich region (P-domain) leaves the lectin domain at residue P_{270}, with four copies of repeat motif (termed type 1), and then returns to the lectin domain at residue F_{415}, with four copies of another repeat motif (type 2) in a "11112222' configuration, thus forming a hook-like arm. This topology of the CNX loop domain is in agreement with that of CRT, recently proposed by Ellgaard *et al.* (158). The CRT loop is shorter than that of CNX, as it is missing the first type 1 repeat, as well as the last type 2 repeat (Figure 12A and B). Thus it resembles a shortened CNX loop with a "111222" configuration.

When tested in our two-hybrid system, the CNX lectin domain failed to mediate a specific interaction with ERp57 (data not shown), while its loop domain (P_{270}-F_{415}) did confer this specificity. To map the region of interaction, we therefore designed constructs of the loop domains of CNX and CRT (Figure 12B). To summarize, in addition to the full lumenal domains of CNX and CRT, we made three additional CNX, and two additional CRT constructs. We subcloned (1) the region encoding the tip of the CNX loop, which

Figure 12. The loop domains of calnexin and calreticulin are suffient, and the B thioredoxin domain of ERp57 is required for the heterodimerization of both calnexin and calreticulin, with ERp57. (A) The proposed crystal structure for the lumenal domain of CNX and the NMR derived structure of the CRT loop domain are shown (96,158). For mapping purposes, the loop domain of CNX was divided into three domains: α (red), β (blue), and γ (gold), and two for CRT, α (red), β (blue) respectively (B) Schematic diagram of the constructs used to map the protein-protein interaction domains of calnexin (CNX), calreticulin (CRT), and ERp57. The loop domains of CNX and CRT, with their corresponding type 1 and type 2 repeat configuration, and the thioredoxin domains of ERp57 (A,B,B',A') are shown. Also, corresponding structural domains (α,β, and γ) are indicated. (C) β–galactosidase filter assays on strains expressing the mapping fusion proteins described in A are shown. pLJ89/pLJ96 and EPOr were used respectively as negative and positive controls.


A
calnexin
calreticulin
α
γ
β
B
IRE1
SS L
TM Kinase/Endoribonuclease
Domain
CNX H21-P482
CNX P270-F415
CNX D289-N393
CNX P310-P378
α,β,γ
α,β
α
CRT E18-L417
CRT D201-A307
CRT A223-P283
α,β
α
ERp57 A21-L505
ERp57 K129-L505
ERp57 E238-L505
A B B' A'
B B' A'
B' A'
C
CNX H21-P482
pLJ89
CNX P270-F415
pLJ89
CNX D289-N393
pLJ89
CNX P310-P378
pLJ89
CRT E18-L417
pLJ89
CRT D201-A307
pLJ89
CRT A223-P283
pLJ89
pLJ96 ERp57 A21-L505
pLJ96 ERp57 K129-L505
pLJ96 ERp57 E238-L505
EPOr

contains two of each repeat motif (repeats 1122,P_{310}-P_{378}, domain α,), (2) a longer loop with three of each repeat (repeats 111222, D_{289}-N_{393}, domains αβ), (3) and one which has four of each and forms the entire loop domain (repeats 11112222, P_{270}-F_{415}, domains αβγ,) (Figure 12A and B). Similarly for CRT, we subcloned the tip region (repeats 1122, A_{223}-P_{283}, domain α), and the full CRT loop (repeats 111222, D_{201}-A_{307}, domain α) (Figure 12A and B).

The constructs for mapping ERp57 were based on the proposed domain structure of its sequence-related homologue PDI (159). Both PDI and ERp57 contain four thioredoxin domains in tandem (A-B-B'-A'). The A and A' share high similarity to thioredoxin, and each contain a copy of the active site consensus sequence -C-G-H-C-, with the N-teminal cysteine being reactive, while the B and B' domains, which are less conserved, have lost this active site consensus sequence (159). In addition to the full-length ERp57 fusion construct, we made two deletion constructs. The first had the first thioredoxin domain removed (leaving B-B'-A', K_{129}-L_{505}), and the second had two removed (leaving B'-A', E_{238}-L_{505}) (Figure 12B). To test for protein interactions, the CNX and CRT constructs were transformed into the MAT**a** reporter strain, while the ERp57 constructs, into the MATα reporter strain. The haploid strains were streaked, the diploids selected, and the interactions were verified with the β-galactosidase filter assay (Figure 12C). The parental plasmids pLJ89 and pLJ96, and the extracellular domain of EPOr were used as negative and positive controls, respectively. The results of this experiment show that two sets of repeat motifs, hence the tip of loop domain (with a "1122" repeat configuration) is sufficient to mediate the interaction of CNX and of CRT, with ERp57 (Figure 12C). Moreover, the results show that the second thioredoxin domain of ERp57 (B) is required to interact with both CNX and CRT (Figure 12B). This result further confirms the specificity of our system, and also demonstrates its sensitivity for detecting the interaction of small domains (60 aa in the case of CRT). In addition to the functional assay (25) and the cross-linking results (133), our results show that there are specific interactions between the regions of CNX and CRT, and ERp57, that we have mapped. While ERp57 and PDI are homologous (33% identity), the conserved regions are concentrated in the A and A' thioredoxin domains, which each contain an active-site

sequence. The B and B' domains of ERp57 and PDI show a low degree of similarity (159), and explain the specificity that we detected here.

3.4.5 Direct interaction of the loop domain of calnexin with ERp57 *in vitro*

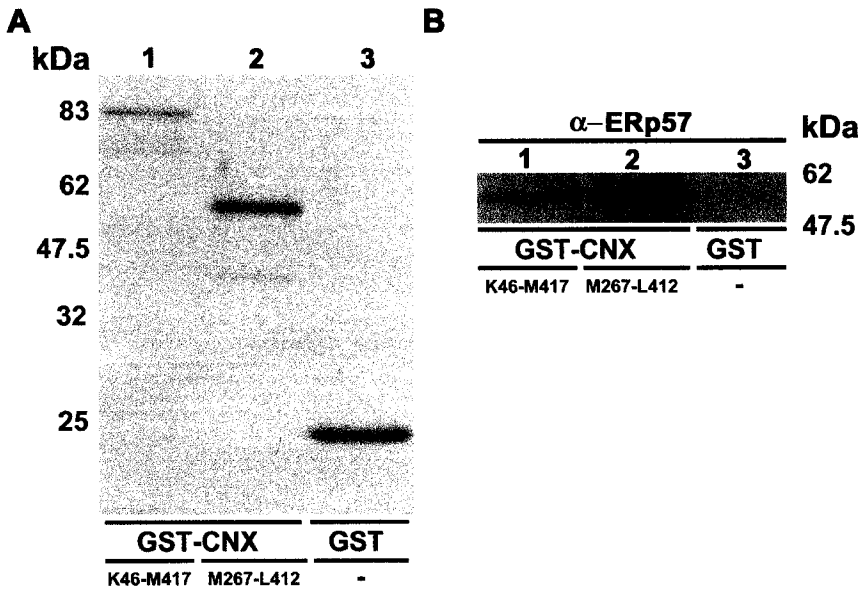
To confirm that these two-hybrid results can also be seen as a physical interaction, we used a GST fusion of the full lumenal domain of CNX (GST-$CNX_{K46\text{-}M417}$), and the loop domain (GST-$CNX_{M267\text{-}L412}$), using GST as a control, and tested for ERp57 binding (Figure 13A). The results show that ERp57 binds specifically to both GST-CNX fusions (Figure 13B). Thus the functional ERp57 (25) and crosslinking (133) results showing the interaction of CNX and ERp57 were confirmed, and the regions that promote this interaction defined.

3.5 Conclusions

This ER protein two-hybrid system provides a tool for mapping protein interactions for both membrane (as is the case for CNX) and soluble ER proteins within the ER, and potentially in other organelles. There is some evidence that the *S. cerevisiae* CNX (Cne1p) and UGGT (Kre5p) homologues have different functions from their mammalian homologues, which may explain why the heterologously expressed mammalian ER proteins do not appear to be interfering with their yeast counterparts (160). However, the yeast ER does provide the redox potential, calcium and ionic concentrations, and some of the enzymes and chaperones that are found in mammalian cells enabling the specific interactions to be detected.

The ER protein two-hybrid system is, in conjunction with our recombinational cloning approach, simple and amenable to high throughput technology, and hence could be used for the comprehensive analysis of the protein interaction networks of the ER.

Figure 13. The loop domain of calnexin interacts directly with ERp57 *in vitro*. (A) GST fusions consisting of the full-length lumenal domain of CNX ($CNX_{K46\text{-}M417}$, lane 1), the loop domain (or P-region, $CNX_{M267\text{-}L412}$, lane 2), and GST alone (lane 3) were purified and are shown on a 10% SDS-PAGE. (B) The purified fusion proteins were loaded onto columns with Glutathione Sepharose 4B, followed by purified ERp57 (25). The columns were washed, the proteins were then eluted with reduced glutathione, and a Western blot performed on the eluant with anti-ERp57 antiserum.


A
kDa
1
2
3
83
62
47.5
32
25
GST-CNX
GST
K46-M417
M267-L412
-
B
α–ERp57
kDa
1
2
3
62
47.5
GST-CNX
GST
K46-M417
M267-L412
-

Chapter 4

Conclusions

4.1 The biology of the Calnexin/Calreticulin Cycle from yeast to mammals, to viruses

The Calnexin/Calreticulin (CNX/CRT) Cycle is a process required for the folding and oligomeric assembly of many secretory and membrane glycoproteins. Numerous pathologies arise when nascent glycoproteins cannot attain their native conformation (145,161). These pathologies provide evidence of the activity of this process and allude to its importance for both protein folding and quality control in the secretory pathway.

The inherent dependence on the CNX/CRT cycle differs from yeast to mammals. While there is an apparent redundancy with two lectin-like chaperones in mammalian cells, transgenic mice lacking CRT die on embryonic day 18 (162). Similarly, a recent report of a human new born with glucosidase I deficiency suggests that blocking access to the cycle leads to a severe and progressive disease state characterized by hypotnia, dysmorphic features, hypoventilation, and here, death at age 74 d (163). While it is unclear whether the embryonic lethality from the CRT deletion is the result of the loss of a major calcium binding protein of the ER, or from a weakened glycoprotein folding machinery, the latter report provides evidence of the essential nature complex N-glycan processing in early mammalian development.

The calnexin system of the yeast *Schizosaccharomyces pombe* appears to be the most similar to the mammalian system, with all the constituents of the glycoprotein folding/quality control apparatus. The CNX homologue in this yeast, termed *cnx1+*, encodes an essential gene (164,165). Moreover, the loss of CNX-monoglucosylated oligosaccharide ($Glc_1Man_9GlcNAc_2$) interactions in an *Δalg6Δ/gpt1* double mutant (*gpt1+* encodes the *S. pombe* homologue of UDP-glucose:glycoprotein glucosyltransferase) leads to very slow growth and a rounded morphology at 28°C (166). However, as shown in *Δgls2/Δalg6* double-mutant cells that are unable to remove the glucose units added by Gpt1p, since they lack glucosidase II, the cyclical process of binding and release to calnexin is only required under conditions that lead to ER-related stress (166).

In contrast, the CNX/CRT Cycle appears to be absent in *S. cerevisiae*. While there is a CNX homologue, termed Cne1p, that seems to function as a constituent of yeast ER quality control apparatus (46,63), the *S. cerevisiae* homologue of UGGT (Kre5p) does not have glucosyltransferase activity, as does UGGT in the mammalian CNX/CRT

cycle (167,168). Moreover, disruption of the *CNE1* gene is not deleterious, nor does it lead to gross effects on the levels of secreted proteins (169). Nevertheless, in a study where mutant strains were constructed to produce non- (Glc_0), mono- (Glc_1), and diglucosylated (Glc_2) glycoproteins in the ER, the induction of the unfolded protein response upon treatment of the cells with DTT was shown to be significantly greater in the Glc_0 and Glc_2 strains (168). This result suggests, although to a lesser degree than for *S. pombe*, that the calnexin-monoglucosylated glycoprotein interaction in *S. cerevisiae* is an important compensatory mechanism of the accumulation of unfolded proteins within the ER.

As mentioned in Chapter 2, the Calnexin/Calreticulin Cycle is not restricted to host proteins. It was shown through studies on the role of N-linked glycosylation that some viral glycoproteins require this process for correct folding (94,97,104-107). This raises the exciting possibility that the ER α-glucosidases, which mediate entry into the CNX/CRT cycle, may be useful targets for anti-viral agents. The best understood examples are with HIV and Hepatitis B virus (HBV). The viral envelope glycoprotein of HIV gp120, which is non-covalently complexed to gp41, is regionally misfolded upon treatment of infected cells with the α-glucosidase inhibitor N-butyl-deoxynojirimycin (NB-DNJ). This misfolding prevents a conformational shift and cleavage of gp120 that normally occurs upon the interaction of gp120 and CD4. This cleavage normally exposes gp41 which mediates viral fusion with the cellular membrane (116).

In the case of HBV, there are three envelope glycoproteins: large (L), medium (M) and small (S), all of which share a common glycosylation site (N_{146}). The M glycoprotein, which has an additional glycosylation site at N_4, has been shown by site-directed mutagenesis to mediate its interaction with CNX. This interaction is critical to the proper folding of the M glycoprotein, which in turn is required for viral envelopment and secretion (116,170,171). To date this is the most promising use of the CNX/CRT Cycle as an anti-viral target, since this approach blocks the formation of full viral particles rather than blocking infectivity. Hence, lower doses of α-glucosidase inhibitor would be sufficient to block the viral life cycle. This is in fact the case as shown with the HBV animal model, Woodchuck Hepatitis Virus (WHV), where the viral titer was reduced to undetectable levels upon treatment of the animals with low doses of N-nonyl-

DNJ (115). Moreover, this treatment did not lead to major disturbances in the glycosylation state of endogenous serum glycoproteins (115).

4.1.2 Complexity at the entry of the Calnexin/Calreticulin Cycle

It is clear that understanding the mechanistic details of the CNX/CRT cycle has both biological and medical implications. Understanding these details provides (1) new insight into the relationship of N-linked glycan trimming and chaperone function in the early secretory pathway, and (2) possible therapeutic strategies against many protein trafficking diseases and hepatic viral infections, most notably the Hepatitis B and C viruses. These details present a significant set of challenges. The quality control apparatus of the secretory pathway is understandably complex considering its exquisite sensitivity, and the nature and complexity of protein folding itself. This is especially true for proteins and oligomeric complexes that are sequestered to a harsh extracellular environment. As such, this work sought to address these challenges.

The early biochemical characterization of the ER α–glucosidases led to a detailed understanding about N-linked oligosaccharide processing on nascent secretory and membrane glycoproteins. It was first shown that a 14-mer oligosaccharide was transferred from a dolichol-pyrophosphate precursor to asparagine residues in the sequon N-X-S/T (172,173). Next, it was found that newly transferred oligosaccharides of the form $Glc_3Man_9GlcNAc_2$, are rapidly processed to the $Glc_1Man_9GlcNAc_2$ form through the action of glucosidase I and II, and then to the $Man_9GlcNAc_2$ form by glucosidase II (60,174,175). Detailed experiments on the kinetics of glucosidase II activity agreed with the *in vivo* results, demonstrating that glucosidase II exhibits the activity of a glucoside hydrolase with both a high and low affinity substrate binding site (91,123,124). Lastly, a link was made between N-linked glycosylation and chaperone function in the ER, upon the characterization of calnexin as a lectin-like chaperone that binds monoglucosylated glycoproteins (8). The complex activity of glucosidase II in fact supported a model whereby the slow hydrolysis of the last glucose residue favors calnexin-monoglucosylated glycoprotein interactions.

Later experiments on glucosidase II raised several questions. Firstly, upon purification of glucosidase II to homogeneity, a non-catalytic β subunit with unknown

function was revealed (92,95,128). Trombetta *et al.* (92) found that enzymatic activity could not be maintained upon separation of the tightly bound α and β subunits. Peptide sequencing allowed them to identify corresponding human cDNA sequences, which revealed that cDNA for the β subunit encoded a putative ER retention signal (HDEL). They proposed that the β subunit was required for enzymatic activity and responsible for the ER localization of glucosidase II. Alternatively, Henteges and Bause (95) had active pig liver glucosidase II preparations both with and without a β subunit. Moreover, Flura *et al.* (94) expressed cDNA for the α subunit alone in CHO cells and observed glucosidase II activity. Recently, Treml *et al.* (176) expressed cDNAs for both the α and β subunits, and found that the expression of the β subunit was required to yield active glucosidase II. They proposed a model for the role of the β subunit based on the results of Henteges and Bause (95). They suggested that the β subunit was involved in α subunit maturation rather than being required for enzymatic activity once the α subunit has aquired its mature form (176).

Our results agree with those of Trombetta *et al.* (92). We show (1) that the expression of the β subunit is not only required for enzymatic activity, but contributes to the solubility/stability of the catalytic α subunit, (2) that active glucosidase II is exclusively heterodimeric, and (3) the removal of the HDEL sequence from the β subunit leads to the secretion of the α subunit.

Nevertheless, we found that the β subunit bound Coomassie Blue dye poorly, which could explain the results of Henteges and Bause (95). Moreover, we found that the expression of the β subunit to be quite variable and cell line dependent (Figure 2B), while the expression of the α subunit was uniform. If the COS1 cell line used by Flura *et al.* (94) to express the α subunit cDNA had an excess of β subunit, an increased enzymatic activity should be observed. We also demonstrated through the formation of heteroenzymes upon expression in *Sf*9 cells, that the interactions between the α and β subunits are conserved from insects to humans. Considering this highly conserved function, we performed sequence alignments on human β subunit with *C. elegans* and *S. pombe*. A highly conserved N-terminal region (Q50-E105) was identified, and proposed as a major α/β protein-protein interaction domain. This was recently confirmed by Arent

and Ostergaard (177). This domain was also found in *S. cerevisiae* (ORF YDR221w) but it remains to be shown that this is the β subunit homologue in this yeast.

Cloning the mammalian α and β subunits of glucosidase II in this study (90) and that of Arendt *et al.* (89) added complexity to the CNX/CRT cycle. These cloning experiments revealed the existence of splice variants for both the α and β subunits. With the existence of splice variants of the catalytic α subunit, the questions raised are the following: (1) do all the splice variants encode active subunits? (2) do they share substrate specificities? (3) are their kinetic properties similar? It was possible that the complex two binding site kinetic model described in earlier biochemical studies (91,123,124) was the result of co-purified splice variants with differing substrate specificities. The existence of multiple forms of glucosidase II could suggest forms with pre- and post-calnexin activities. To address these questions, we established a methodology to express and purify recombinant isoforms of heterodimeric glucosidase II using the baculovirus system.

Here, a detailed kinetic analysis on the properties of the recombinant isoforms of glucosidase II, α1/β and α2/β, revealed that they were catalytically indistinguishable. Both exhibited complex activities that fit the two binding site kinetic model proposed by Alonso *et al.* (91). Similar K_Ms were found for *p*-NP-Glc, and K_is for deoxynojirimycin (DNJ) and castanospermine (CST). When tested on the physiological oligosaccharide substrates, both recombinant isoforms also exhibited comparable substrate specificities and activities. These findings clearly addressed the questions one to three, raised above, and ruled out the model pre- and post-calnexin forms of glucosidase II.

The question remains as to why there are splice variants of the glucosidase II. Baldwin *et al.* (178) recently demonstrated a specific interaction between isoforms of glucosidase II expressing Box A1 (178) and CD45. CD45 is an abundant, transmembrane, protein-tyrosine phosphatase expressed in cells of hematopoietic origin, with eleven to eighteen putative glycosylation sites, and is involved in signal transduction events leading to thymocyte maturation and T cell activation (179). The interaction appears to be specific since glucosidase II has not been found associated with other N-linked glycan containing proteins such as Class I major histocompatability complex, LFA-1, of CD44 (178). Baldwin *et al.* (178) suggested that Box A1 functions to stabilize

the interaction, but not through modifying the enzymatic properties, since they also found that the isoforms displayed similar activities. They found that enzymatic inhibition by DNJ blocked glucosidase II-CD45, and that mannose, but not glucose, blocked the association of glucosidase II with CD45. They concluded that the active site and lectin activity of Box A1 are required to mediate this specific interaction. These results suggest a model where the glucosidase II form with Box A1 forms an enzyme-substrate complex with CD45, then retains CD45 specifically with a lectin domain encoded or exposed by Box A1. This interaction could have numerous implications for CD45. Glucosidase II could be acting as a molecular chaperone to retain CD45 for further post-translational modifications, to block the interaction with additional proteins, to regulate cell surface expression of CD45, or to regulate CD45 activity (178). Nevertheless, these findings provide interesting and plausible new roles for the existence of glucosidase II splice variants.

4.3 Studying protein interactions *in vivo*

4.3.1 Novel systems and post-genomic efforts

With the advent of the classical two-hybrid assay proposed by Fields and Song (134), an abundance of protein-protein interactions have been found and characterized. The genetically based system complemented the standard biochemical approaches, which include methods such as chemical cross-linking, co-immunoprecipitation, co-fractionation by chromatography, GST-pull down assays, and far Western analysis. The interactions found with the Fields and Song system, when compiled into networks, have provided the first detailed interaction maps, greatly enhancing our understanding of the inherent biological significance of each protein therein. Some of the limitations of the two-hybrid system, which is based on the re-assembly of the Gal4p transcriptional activator *in vivo*, through direct interactions of the fused proteins of interest, have been overcome through the development of alternative approaches. These include the split-ubiquitin system (136,142), the SOS and Ras recruitment systems (SRS and RRS) (137,143), G-protein fusions (144), and the oligomerization-assisted enzymatic complementation systems which result in the reassembly of murine dihydrofolate reductase (mDHFR) (180) or the β-galactosidase (181) from *E. coli*.

These techniques have extended the range of protein interactions that can be studied *in vivo*, which now include transcriptional activators that tend to self-activate the Gal4p based system (137), and integral membrane and membrane-associated proteins (142-144). They have also permitted the study of protein-protein interactions endogenously within mammalian cells. Regardless of the inherent constraints of the classical yeast two hybrid system, it has been used for the formidable endeavor of comprehensively mapping the protein-protein interactions within the nematode *C. elegans* (140) and yeast *S. cerevisiae* (138,139) for which complete genome sequences are available (182,183).

4.3.2 The ER protein two hybrid

The ER's unique environment of high protein concentrations, high calcium content, and redox state has precluded the use of some of the classical methods of studying protein-protein interactions that have been successful at defining the protein networks in the cytosol and other organelles. In an effort to map the protein interactions of both CNX and CRT with ERp57, we developed an ER protein two hybrid system that is based on the UPR sensor Ire1p, to report the interactions occurring within the lumen of the ER (Chapter 3). This system provides a similar redox potential, calcium and ionic concentrations, the signal sequence peptidase, and ER N-glycosylation and trimming enzymes (including OT, glucosidase I and II, and ER mannosidases) for proper folding and post-translational modifications that are specific to the ER. As such, we believe that it is the most suitable system for the characterization of the heterodimeric CNX/ERp57 and CRT/ERp57 complexes.

4.3.3 The CNX/ERp57 and CRT/ERp57 heterodimeric complexes

A number of recent experiments revealed the functional relationship between the lectin-like chaperones CNX and CRT, and their associated protein disulfide isomerase, ERp57. Cross-linking experiments performed by Oliver *et al.* (24) first demonstrated that ERp57 substrates were monoglucosylated glycoproteins. They subsequently found that ERp57, when *in vitro* translated and imported into dog pancreatic microsomes, cross-linked to endogenous calnexin and calreticulin, while radiolabeled PDI did not (133).

Zapun *et al.* (25) found that the disulfide isomerase activity of ERp57 and not PDI, on the refolding of monoglucosylated ribonuclease B, was enhanced upon its association with either calnexin or calreticulin. These results demonstrate that the CNX/ERp57 and CRT/ERp57 complexes act as folding scaffolds specific for monoglucosylated secretory and membrane glycoproteins. We confirmed the specific interactions of CNX/CRT with ERp57 in our ER protein two-hybrid system.

With the recently solved crystal structure of calnexin (96) and NMR-derived structure of the P-domain of calreticulin (158), we designed deletion constructs for CNX and CRT. The CNX structure is characterized by a globular lectin domain and a loop 'long arm' domain that corresponds to the central, proline rich P-domain of both CNX and CRT. Deletion constructs were also made for ERp57 based on its thioredoxin domain configuration (A B B' A') (159). We found that the two central tandem repeats (1122) of the CNX and CRT loop domain were sufficient, while the B thioredoxin domain of ERp57 was required to mediate the CNX/ERp57 and CRT/ERp57 interactions. We confirmed these results by demonstrating that specific interaction between the loop domain of calnexin and ERp57 *in vitro*. Thus, we defined a functional role for the loop domain of CNX and CRT, and the non-catalytic B' domain of ERp57, and lastly, demonstrated the specificity and sensitivity of our ER protein two-hybrid system.

4.3.4 The *trans*-phosphorylating ER protein two-hybrid system, ER protein network analysis, and viral-host protein interactions

The genetically-based ER protein two-hybrid assay, which uses a recombinational cloning approach, offers some interesting possibilities for high throughput protein interaction analysis. We propose however, to adapt the ER protein two-hybrid system to function exclusively in *trans*, to enhance the versatility of the system. This could be achieved by using two mutants of Ire1p described by Shamu and Walter (66). When expressed as fusion proteins, the inactive kinase mutant Ire1p-K702R and the phosphorylation site double mutant Ire1p-S840A/S841A should maintain their capacity to report heterodimerization through *trans*-phosphorylation of the Ire1p-K702R kinase domain, and not allow autoactivation by a fusion protein that forms homodimers (Figure 14).

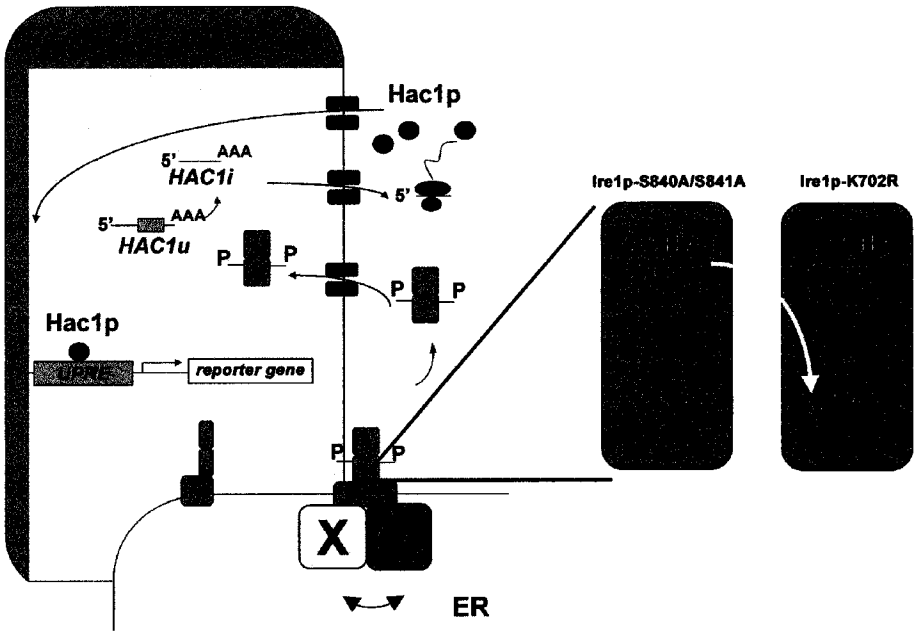
This adaptation should permit the construction of cDNA libraries within the Ire1p-based parental vectors, by eliminating the background reporter gene activity, that would be induced by proteins that homodimerize. A genetic screen would then be possible with a single bait and cDNA Ire1p-fusion library. Alternatively, libraries could be constructed upon data mining the human genome for ER proteins. These could then be tested using a matrix approach. This last approach is currently being used for the comprehensive analysis of the protein interaction networks of *S. cerevisiae* and C. *elegans* (138-140).

We also believe the system could be used to study host-virus protein interactions, for viruses such as hepatitis C virus (HCV). The HCV viral capsid is thought to be assembled in the ER (185), and would be an excellent candidate for this approach. Thus, the system could be used to screen for the hepatic cell viral receptor of the (HCV). Firstly, Ire1p fusion proteins expressing the HCV envelope glycoproteins E1 and E2 (184), could each be co-expressed with a human liver cDNA/Ire1p fusion library, to screen for hepatic cell membrane proteins that interact with these envelope glycoproteins. The system could also be used to characterize and map E1-E2 interactions (185), as well as E2-CD81 and E1/E2-LDLr first described by Wunschmann *et al.* (186), both which are thought to mediate infectivity. Lastly, host-virus protein interactions could be analyzed for the non-structural proteins of HCV and extend our knowledge of capsid formation and viral replication of HCV.

4.4 Summary

In this thesis, a detailed kinetic analysis was presented demonstrating the complex enzymatic activity of the recombinant isoforms of heterodimeric glucosidase II ($\alpha 1/\beta$ and $\alpha 2/\beta$). Both were shown to exhibit kinetic profiles of a two binding site model, and to share properties of catalysis and inhibition. Also, expression of the β subunit was found to be essential for enzymatic activity, solubilization and/or stability, and ER retention of the α/β complex. Moreover, an ER protein two-hybrid system was developed and used to map the interactions between CNX/CRT and ERp57. New roles were assigned to the "long arm" P-domains of CNX and CRT, and to the non-catalytic B thioredoxin domain of ERp57, in mediating these interactions.

Figure 14. Schematic diagram of the *trans*-phosphorylating ER protein two-hybrid system. Dimerization of proteins X and Y, expressed as fusions with the Ire1p mutant proteins Ire1p-S840A/S841A and Ire1p-K702R respectively, should lead to the trans-phosphorylation of the Ire1p-K702R kinase domain. This will subsequently activate the endoribonuclease domain, which will process the mRNA of the transcriptional activator Hac1p. Hac1p will then induce the expression of the reporter gene.


Hac1p
5'
AAA
HAC1i
5'
AAA
5'
HAC1u
P
P
P
P
Hac1p
UPRE
reporter gene
P
P
X
ER
Ire1p-S840A/S841A
Ire1p-K702R

References

1 Ellis, R.J. & Hemmingsen, S.M. Molecular chaperones: proteins essential for the biogenesis of some macromolecular structures. *Trends Biochem. Sci.* **14**, 339-342 (1989).

2 Hendrick, J.P. & Hartl, F.U. Molecular chaperone functions of heat-shock proteins. *Ann. Rev. Biochem.* **62**, 349-384 (1993).

3 Hubbard, S.C. & Ivatt, R.J. Synthesis and processing of asparagine-linked oligosaccharides. *Ann. Rev. Biochem.* **50**, 555-584 (1981).

4 Wada, I. *et al.* SSR alpha and associated calnexin are major calcium binding proteins of the endoplasmic reticulum membrane. *J. Biol. Chem.* **266**, 19599-19610 (1991).

5 Fliegel, L., Burns, K., MacLennan, D.H., Reithmeier, R.A., & Michelak, M. Molecular cloning of the high affinity calcium-binding protein (calreticulin) of skeletal muscle sarcoplasmic reticulum. *J. Biol. Chem.* **264**, 21522-21528 (1989).

6 Degen, E. & Williams, D.B. Participation of a novel 88kDa protein in the biogenesis of murine class I histocompatibility. *J. Cell Biol.* **112**, 1099-1115 (1991).

7 Ahluwalia, N., Bergeron, J.J., Wada, I., Degen, E., & Williams, D.B. The p88 molecular chaperone is identical to the endoplasmic reticulum membrane protein, calnexin. *J. Biol. Chem.* **267**, 10914-10918 (1992).

8 Ou, W.J., Cameron, P.H., Thomas, D.Y., & Bergeron, J.J. Association of folding intermediates of glycoproteins with calnexin during protein maturation. *Nature* **364**, 771-776 (1993).

9 Yeo, T.K., Yeo, K.T., & Olden, K. Accumulation of unglycosylated liver secretory glycoproteins in the rough endoplasmic reticulum. *Biochem. Biophys. Res. Commun.* **160**, 1421-1428 (1989).

10 Ware, F.E. *et al.* The molecular chaperone calnexin binds Glc1Man9GlcNAc2 oligosaccharide as an initial step in recognizing unfolded glycoproteins. *J. Biol. Chem.* **270**, 4697-4704 (1995).

11 Li, Y. *et al.* Effects of inefficient cleavage of the signal sequence of HIV-1 gp120 on its association with calnexin, folding, and intracellular transport. *Proc. Natl. Acad. Sci. U. S. A.* **93**, 9606-9611 (1996).

12 Ou, W.J., Bergeron, J.J., Li, Y., Kang, C.Y., & Thomas, D.Y. Conformational changes induced in the endoplasmic reticulum luminal domain of calnexin by Mg-ATP and Ca2+. *J. Biol. Chem.* **270**, 18051-18059 (1995).

13 Vassilakos, A., Michalak, M., Lehrman, M.A., & Williams, D.B. Oligosaccharide binding characteristics of the molecular chaperones calnexin and calreticulin. *Biochemistry* **37**, 3480-3490 (1998).

14 Wada, I., Imai, S., Kai, M., Sakane, F., & Kanoh, H. Chaperone function of calreticulin when expressed in the endoplasmic reticulum as the membrane-anchored and soluble forms. *J. Biol. Chem.* **270**, 20298-20304 (1995).

15 Tessier, D.C. *et al.* Cloning and characterization of mammalian UDP-glucose glycoprotein:glucosyltransferase and the development of a specific substrate for this enzyme. *Glycobiology* **10**, 403-412. (2000).

16 Trombetta, S.E., Bosch, M., & Parodi, A.J. Glucosylation of glycoproteins by mammalian, plant, fungal, and trypanosomatid protozoa microsomal membranes. *Biochemistry* **28**, 8108-8116 (1989).

17 Zapun, A. *et al.* Conformation-independent binding of monoglucosylated ribonuclease B to calnexin. *Cell* **88**, 29-38 (1997).

18 Rodan, A.R., Simons, J.F., Trombetta, E.S., & Helenius, A. N-linked oligosaccharides are necessary and sufficient for association of glycosylated forms of bovine RNase with calnexin and calreticulin. *EMBO J.* **15**, 6921-6930 (1996).

19 Ihara, Y., Cohen-Doyle, M.F., Saito, Y., & Williams, D.B. Calnexin discriminates between protein conformational states and functions as a molecular chaperone *in vitro*. *Mol. Cell* **4**, 331-341 (1999).

20 Ehrnsperger, M., Graber, S., Gaestel, M., & Buchner, J. Binding of non-native protein to Hsp25 during heat shock creates a reservoir of folding intermediates for

reactivation. *EMBO J.* **16**, 221-9 (1997).

21 Freeman, B.C. & Morimoto, R.I. The human cytosolic molecular chaperones hsp90, hsp70 (hsc70) and hdj-1 have distinct roles in recognition of a non-native protein and protein refolding. *EMBO J.* **15**, 2969-2979 (1996).

22 Lee, G.J., Roseman, A.M., Saibil, H.R., & Vierling, E. A small heat shock protein stably binds heat-denatured model substrates and can maintain a substrate in a folding-competent state. *EMBO J.* **16**, 659-671 (1997).

23 Saito, Y., Ihara, Y., Leach, M.R., Cohen-Doyle, M.F., & Williams, D.B. Calreticulin functions *in vitro* as a molecular chaperone for both glycosylated and non-glycosylated proteins. *EMBO J.* **18**, 6718-6729 (1999).

24 Oliver, J.D., van der Wal, F.J., Bulleid, N.J., & High, S. Interaction of the thiol-dependent reductase ERp57 with nascent glycoproteins. *Science* **275**, 86-88 (1997).

25 Zapun, A. *et al.* Enhanced catalysis of ribonuclease B folding by the interaction of calnexin or calreticulin with ERp57. *J. Biol. Chem.* **273**, 6009-6012 (1998).

26 Wada, I., Imai, S., Kai, M., Sakane, F., & Kanoh, H. Chaperone function of calreticulin when expressed in the endoplasmic reticulum as the membrane-anchored and soluble forms. *J. Biol. Chem.* **270**, 20298-20304 (1995).

27 Cannon, K.S. & Helenius, A. Trimming and readdition of glucose to N-linked oligosaccharides determines calnexin association of a substrate glycoprotein in living cells. *J. Biol. Chem.* **274** , 7537-7544 (1999).

28 Wada, I., Kai, M., Imai, S., Sakane, F., & Kanoh, H. Promotion of transferrin folding by cyclic interactions with calnexin and calreticulin. *EMBO J.* **16**, 5420-5432 (1997).

29 Perlmutter, D.H. Liver disease associated with alpha 1-antitrypsin deficiency. *Prog. Liver Dis.* **11**, 139-165 (1993).

30 Cheng, S.H. *et al.* Defective intracellular transport and processing of CFTR is the molecular basis of most cystic fibrosis. *Cell* **63**, 827-834 (1990).

31 Drumm, M.L. *et al.* Chloride conductance expressed by delta F508 and other mutant CFTRs in Xenopus oocytes. *Science* **254**, 1797-1799 (1991).

32 Dalemans, W. *et al.* Altered chloride ion channel kinetics associated with the delta F508 cystic fibrosis mutation. *Nature* **354**, 526-528 (1991).

33 Denning, G.M. *et al.* Processing of mutant cystic fibrosis transmembrane conductance regulator is temperature-sensitive. *Nature* **358**, 761-764 (1992).

34 Klausner, R.D. & Sitia, R. Protein degradation in the endoplasmic reticulum. *Cell* **62**, 611-614 (1990).

35 Kopito, R.R. ER quality control: the cytoplasmic connection. *Cell* **88**, 427-430 (1997).

36 Sommer, T. & Jentsch, S. A protein translocation defect linked to ubiquitin conjugation at the endoplasmic reticulum. *Nature* **365**, 176-179 (1993).

37 Ciechanover, A. The ubiquitin-proteasome pathway: on protein death and cell life. *EMBO J.* **17**, 7151-7160 (1998).

38 Rapoport, T.A., Rolls, M.M., & Jungnickel, B. Approaching the mechanism of protein transport across the ER membrane. *Curr. Opin. Cell Biol.* **8**, 499-504 (1996).

39 Jensen, T.J. *et al.* Multiple proteolytic systems, including the proteasome, contribute to CFTR processing. *Cell* **83**, 129-135 (1995).

40 Qu, D., Teckman, J.H., Omura, S., & Perlmutter, D.H. Degradation of a mutant secretory protein, alpha1-antitrypsin Z, in the endoplasmic reticulum requires proteasome activity. *J. Biol. Chem.* **271**, 22791-22795 (1996).

41 Ward, C.L., Omura, S., & Kopito, R.R. Degradation of CFTR by the ubiquitin-proteasome pathway. *Cell* **83**, 121-127 (1995).

42 Hiller, M.M., Finger, A., Schweiger, M., & Wolf, D.H. ER degradation of a misfolded luminal protein by the cytosolic ubiquitin-proteasome pathway. *Science*

273, 1725-1728 (1996).

43 Werner, E.D., Brodsky, J.L., & McCracken, A.A. Proteasome-dependent endoplasmic reticulum-associated protein degradation: an unconventional route to a familiar fate. *Proc. Natl. Acad. Sci. U. S. A.* **93**, 13797-13801 (1996).

44 Rock, K.L. *et al.* Inhibitors of the proteasome block the degradation of most cell proteins and the generation of peptides presented on MHC class I molecules. *Cell* **78**, 761-771 (1994).

45 Fenteany, G. *et al.* Inhibition of proteasome activities and subunit-specific amino-terminal threonine modification by lactacystin. *Science* **268**, 726-731 (1995).

46 McCracken, A.A. & Brodsky, J.L. Assembly of ER-associated protein degradation *in vitro*: dependence on cytosol, calnexin, and ATP. *J. Cell Biol.* **132**, 291-298 (1996).

47 Wiertz, E.J. *et al.* The human cytomegalovirus US11 gene product dislocates MHC class I heavy chains from the endoplasmic reticulum to the cytosol. *Cell* **84**, 769-779 (1996).

48 Wiertz, E.J. *et al.* Sec61-mediated transfer of a membrane protein from the endoplasmic reticulum to the proteasome for destruction. *Nature* **384**, 432-438 (1996).

49 Plemper, R.K., Bohmler, S., Bordallo, J., Sommer, T., & Wolf, D.H. Mutant analysis links the translocon and BiP to retrograde protein transport for ER degradation. *Nature* **388**, 891-895 (1997).

50 Brodsky, J.L. *et al.* The requirement for molecular chaperones during endoplasmic reticulum- associated protein degradation demonstrates that protein export and import are mechanistically distinct. *J. Biol. Chem.* **274**, 3453-3460 (1999).

51 Tsai, B., Rodighiero, C., Lencer, W. I. & Rapoport, T. A. Protein disulfide isomerase acts as a redox-dependent chaperone to unfold cholera toxin. *Cell* **104**, 937-948 (2001).

52 Brodsky, J.L. Translocation of proteins across the endoplasmic reticulum

membrane. *Int. Rev. Cytol.* **178**, 277-328 (1998).

53 Biederer, T., Volkwein, C., & Sommer, T. Role of Cue1p in ubiquitination and degradation at the ER surface. *Science* **278**, 1806-1809 (1997).

54 Mayer, T.U., Braun, T., & Jentsch, S. Role of the proteasome in membrane extraction of a short-lived ER- transmembrane protein. *EMBO J.* **17**, 3251-3257 (1998).

55 Hampton, R.Y., Gardner, R.G., & Rine, J. Role of 26S proteasome and HRD genes in the degradation of 3-hydroxy-3-methylglutaryl-CoA reductase, an integral endoplasmic reticulum membrane protein. *Mol. Biol. Cell* **7**, 2029-2044 (1996).

56 Plemper, R.K. *et al.* Genetic interactions of Hrd3p and Der3p/Hrd1p with Sec61p suggest a retro-translocation complex mediating protein transport for ER degradation. *J. Cell Sci.* **112**, 4123-4134 (1999).

57 Knop, M., Hauser, N., & Wolf, D.H. N-Glycosylation affects endoplasmic reticulum degradation of a mutated derivative of carboxypeptidase yscY in yeast. *Yeast* **12**, 1229-1238 (1996).

58 Jakob, C.A., Burda, P., Roth, J., & Aebi, M. Degradation of misfolded endoplasmic reticulum glycoproteins in Saccharomyces cerevisiae is determined by a specific oligosaccharide structure. *J. Cell Biol.* **142**, 1223-1233 (1998).

59 Liu, Y., Choudhury, P., Cabral, C.M., & Sifers, R.N. Oligosaccharide modification in the early secretory pathway directs the selection of a misfolded glycoprotein for degradation by the proteasome. *J. Biol. Chem.* **274**, 5861-5867 (1999).

60 Hubbard, S.C. & Robbins, P.W. Synthesis and processing of protein-linked oligosaccharides *in vivo*. *J. Biol. Chem.* **254**, 4568-4576 (1979).

61 Cabral, C. M., Choudhury, P., Liu, Y. & Sifers, R. N. Processing by endoplasmic reticulum mannosidases partitions a secretion-impaired glycoprotein into distinct disposal pathways. *J. Biol. Chem.* **275**, 25015-25022 (2000).

62 Fagioli, C. & Sitia, R. Glycoprotein quality control in the endoplasmic reticulum. Mannose trimming by endoplasmic reticulum mannosidase I times the proteasomal

degradation of unassembled immunoglobulin subunits. *J. Biol. Chem.* **276,** 12885-12892 (2001).

63 Parlati, F., Dominguez, M., Bergeron, J.J., & Thomas, D.Y. *Saccharomyces cerevisiae CNE1* encodes an endoplasmic reticulum (ER) membrane protein with sequence similarity to calnexin and calreticulin and functions as a constituent of the ER quality control apparatus. *J. Biol. Chem.* **270**, 244-253 (1995).

64 Cox, J.S., Shamu, C.E., & Walter, P. Transcriptional induction of genes encoding endoplasmic reticulum resident proteins requires a transmembrane protein kinase. *Cell* **73**, 1197-1206 (1993).

65 Nikawa, J. & Yamashita, S. IRE1 encodes a putative protein kinase containing a membrane-spanning domain and is required for inositol phototrophy in *Saccharomyces cerevisiae. Mol. Microbiol.* **6**, 1441-1446 (1992).

66 Shamu, C.E. & Walter, P. Oligomerization and phosphorylation of the Ire1p kinase during intracellular signaling from the endoplasmic reticulum to the nucleus. *EMBO J.* **15**, 3028-3039 (1996).

67 Sidrauski, C. & Walter, P. The transmembrane kinase Ire1p is a site-specific endonuclease that initiates mRNA splicing in the unfolded protein response. *Cell* **90**, 1031-1039 (1997).

68 Zhou, A., Hassel, B.A., & Silverman, R.H. Expression cloning of 2-5A-dependent RNAase: a uniquely regulated mediator of interferon action. *Cell* **72**, 753-765 (1993).

69 Cox, J.S. &Walter, P. A novel mechanism for regulating activity of a transcription factor that control the unfolded protein response. *Cell* **87**, 391-404 (1996).

70 Sidrauski, C., Cox, J.S., & Walter, P. tRNA ligase is required for regulated mRNA splicing in the unfolded protein response. *Cell* **87**, 405-413 (1996).

71 Tirasophon, W., Welihinda, A.A., & Kaufman, R.J. A stress response pathway from the endoplasmic reticulum to the nucleus requires a novel bifunctional protein kinase/endoribonuclease (Ire1p) in mammalian cells. *Genes Dev.* **12**, 1812-1824 (1998).

72 Welihinda, A.A., Tirasophon, W., Green, S.R., & Kaufman, R.J. Gene induction in response to unfolded protein in the endoplasmic reticulum is mediated through Ire1p kinase interaction with a transcriptional coactivator complex containing Ada5p. *Proc. Natl. Acad. Sci. U. S. A.* **94**, 4289-4294 (1997).

73 Welihinda, A.A., Tirasophon, W., Green, S.R., & Kaufman, R.J. Protein serine/threonine phosphatase Ptc2p negatively regulates the unfolded-protein response by dephosphorylating Ire1p kinase. *Mol. Cell Biol.* **18**, 1967-1977 (1998).

74 Wang, X.Z. *et al.* Cloning of mammalian Ire1 reveals diversity in the ER stress responses. *EMBO J.* **17**, 5708-5717 (1998).

75 Niwa, M., Sidrauski, C., Kaufman, R.J., & Walter, P. A role for presenilin-1 in nuclear accumulation of Ire1 fragments and induction of the mammalian unfolded protein response. *Cell* **99**, 691-702 (1999).

76 Selkoe, D. J. The cell biology of beta-amyloid procursur protein and presinilin in Alzheimer's disease. *Trends Cell Biol.* **8**, 447-453 (1998).

77 Tirasophon, W., Lee, K., Callaghan, B., Welihinda, A. & Kaufman, R. J. The endoribonuclease activity of mammalian IRE1 autoregulates its mRNA and is required for the unfolded protein response. *Genes Dev.* **14**, 2725-2736 (2000).

78 Kaufman, R.J. Stress signaling from the lumen of the endoplasmic reticulum: coordination of gene transcriptional and translational controls. *Genes Dev.* **13**, 1211-1233 (1999).

79 Kyriakis, J.M. *et al.* The stress-activated protein kinase subfamily of c-Jun kinases. *Nature* **369**, 156-160 (1994).

80 Srivastava, R.K. *et al.* Bcl-2 and Bcl-X(L) block thapsigargin-induced nitric oxide generation, c-Jun NH(2)-terminal kinase activity, and apoptosis. *Mol. Cell Biol.* **19**, 5659-5674 (1999).

81 Gupta, S., Campbell, D., Derijard, B., & Davis, R.J. Transcription factor ATF2 regulation by the JNK signal transduction pathway. *Science* **267**, 389-93 (1995).

82 Sanchez, I. *et al.* Role of SAPK/ERK kinase-1 in the stress-activated pathway

regulating transcription factor c-Jun. *Nature* **372**, 794-798 (1994).

83 Urano, F. *et al.* Coupling of stress in the ER to activation of JNK protein kinases by transmembrane protein kinase IRE1. *Science* **287**, 664-666 (2000).

84 Lee, S.Y. *et al.* TRAF2 is essential for JNK but not NF-kappaB activation and regulates lymphocyte proliferation and survival. *Immunity* **7**, 703-713 (1997).

85 Yeh, W.C. *et al.* Early lethality, functional NF-kappaB activation, and increased sensitivity to TNF-induced cell death in TRAF2-deficient mice. *Immunity* **7**, 715-725 (1997).

86 Hinnebusch, A.G. The eIF-2 alpha kinases: regulators of protein synthesis in starvation and stress. *Semin. Cell Biol.* **5**, 417-426 (1994).

87 Prostko, C.R., Dholakia, J.N., Brostrom, M.A., & Brostrom, C.O. Activation of the double-stranded RNA-regulated protein kinase by depletion of endoplasmic reticular calcium stores. *J. Biol. Chem.* **270**, 6211-6215 (1995).

88 Bertolotti, A., Zhang, Y., Hendershot, L.M., Harding, H.P., & Ron, D. Dynamic interaction of BiP and ER stress transducers in the unfolded-protein response. *Nat. Cell Biol.* **2**, 326-332 (2000).

89 Arendt, C.W., Dawicki, W., & Ostergaard, H.L. Alternative splicing of transcripts encoding the alpha- and beta- subunits of mouse glucosidase II in T lymphocytes. *Glycobiology* **9**, 277-283 (1999).

90 Pelletier, M.F. *et al.* The Heterodimeric Structure of Glucosidase II is Required for its Activity, Solubility and Localization *in vivo*. *Glycobiology* **10**, 815-827 (2000).

91 Alonso, J.M., Santa-Cecilia, A., & Calvo, P. Glucosidase II from rat liver microsomes. Kinetic model for binding and hydrolysis. *Biochem. J.* **278**, 721-727 (1991).

92 Trombetta, E.S., Simons, J.F., & Helenius, A. Endoplasmic reticulum glucosidase II is composed of a catalytic subunit, conserved from yeast to mammals, and a tightly bound noncatalytic HDEL-containing subunit. *J. Biol. Chem.* **271**, 27509-27516 (1996).

93 D'Alessio, C., Fernandez, F., Trombetta, E.S., & Parodi, A.J. Genetic evidence for the heterodimeric structure of glucosidase II. The effect of disrupting the subunit-encoding genes on glycoprotein folding. *J. Biol. Chem.* **274**, 25899-25905 (1999).

94 Flura, T., Brada, D., Ziak, M., & Roth, J. Expression of a cDNA encoding the glucose trimming enzyme glucosidase II in CHO cells and molecular characterization of the enzyme deficiency in a mutant mouse lymphoma cell line. *Glycobiology* **7**, 617-624 (1997).

95 Hentges, A. & Bause, E. Affinity purification and characterization of glucosidase II from pig liver. *Biol. Chem.* **378**, 1031-1038 (1997).

96 Schrag, J. D. *et al.* The structure of calnexin, an ER chaperone involved in quality control of protein folding. *Submitted.*

97 Ou, W.J., Thomas, D.Y., Bell, A.W., & Bergeron, J.J. Casein kinase II phosphorylation of signal sequence receptor alpha and the associated membrane chaperone calnexin. *J. Biol. Chem.* **267**, 23789-23796 (1992).

98 Bergeron, J.J., Brenner, M.B., Thomas, D.Y., & Williams, D.B. Calnexin: a membrane-bound chaperone of the endoplasmic reticulum. *Trends Biochem. Sci.* **19**, 124-128 (1994).

99 Hammond, C. & Helenius, A. Folding of VSV G protein: sequential interaction with BiP and calnexin. *Science* **266**, 456-458 (1994).

100 Sousa, M.C., Ferrero-Garcia, M.A., & Parodi, A.J. Recognition of the oligosaccharide and protein moieties of glycoproteins by the UDP-Glc:glycoprotein glucosyltransferase. *Biochemistry* **31**, 97-105 (1992).

101 Sousa, M. &Parodi, A.J. The molecular basis for the recognition of misfolded glycoproteins by the UDP-Glc:glycoprotein glucosyltransferase. *EMBO J.* **14**, 4196-4203 (1995).

102 Zapun, A., Jacob, C.A., Thomas, D.Y., & Bergeron, J.J. Protein folding in a specialized compartment: the endoplasmic reticulum. *Struct. Fold. Des.* **7**, 173-182 (1999).

103 Hammond, C. & Helenius, A. Quality control in the secretory pathway. *Curr. Opin. Cell Biol.* **7**, 523-529 (1995).

104 Balow, J.P., Weissman, J.D., & Kearse, K.P. Unique expression of major histocompatibility complex class I proteins in the absence of glucose trimming and calnexin association. *J. Biol. Chem.* **270**, 29025-29029 (1995).

105 Hebert, D.N., Foellmer, B., & Helenius, A. Glucose trimming and reglucosylation determine glycoprotein association with calnexin in the endoplasmic reticulum. *Cell* **81**, 425-433 (1995).

106 Kearse, K.P., Williams, D.B., & Singer, A. Persistence of glucose residues on core oligosaccharides prevents association of TCR alpha and TCR beta proteins with calnexin and results specifically in accelerated degradation of nascent TCR alpha proteins within the endoplasmic reticulum. *EMBO J.* **13**, 3678-3686 (1994).

107 Ora, A. & Helenius, A. Calnexin fails to associate with substrate proteins in glucosidase- deficient cell lines. *J. Biol. Chem.* **270**, 26060-26062 (1995).

108 Sommer, T. & Wolf, D.H. Endoplasmic reticulum degradation: reverse protein flow of no return. *FASEB J.* **11**, 1227-1233 (1997).

109 Brodsky, J. & McCracker, A.A. ER-associated and proteasome-mediated protein degradation: how two topology restricted events came together. *Trends Cell Biol.* **7**, 151-156 (1997).

110 Karpas, A. *et al.* Aminosugar derivatives as potential anti-human immunodeficiency virus agents. *Proc. Natl. Acad. Sci. U.S.A.* **85**, 9229-9233 (1988).

111 Zitzmann, N. *et al.* Imino sugars inhibit the formation and secretion of bovine viral diarrhea virus, a pestivirus model of hepatitis C virus: implications for the development of broad spectrum anti-hepatitis virus agents. *Proc. Natl. Acad. Sci. U. S. A.* **96**, 11878-11882 (1999).

112 Block, T.M. *et al.* Secretion of human hepatitis B virus is inhibited by the imino sugar N-butyldeoxynojirimycin. *Proc. Natl. Acad. Sci. U. S. A.* **91**, 2235-2239

(1994).

113 Fischer, P.B. *et al.* The alpha-glucosidase inhibitor N-butyldeoxynojirimycin inhibits human immunodeficiency virus entry at the level of post-CD4 binding. *J. Virol.* **69**, 5791-5797 (1995).

114 Taylor, D.L. *et al.* Loss of cytomegalovirus infectivity after treatment with castanospermine or related plant alkaloids correlates with aberrant glycoprotein synthesis. *Antiviral Res.* **10**, 11-26 (1988).

115 Block, T.M. *et al.* Treatment of chronic hepadnavirus infection in a woodchuck animal model with an inhibitor of protein folding and trafficking. *Nat. Med.* **4**, 610-614 (1998).

116 Mehta, A., Zitzmann, N., Rudd, P.M., Block, T.M., & Dwek, R.A. Alpha-glucosidase inhibitors as potential broad based anti-viral agents. *FEBS Lett.* **430**, 17-22 (1998).

117 Elbein, A.D. Glycosidase inhibitors: inhibitors of N-linked oligosaccharide processing. *FASEB J.* **5**, 3055-3063 (1991).

118 Arendt, C.W. & Ostergaard, H.L. Identification of the CD45-associated 116-kDa and 80-kDa proteins as the alpha- and beta-subunits of alpha-glucosidase II. *J. Biol. Chem.* **272**, 13117-13125 (1997).

119 Henrissat, B. & Bairoch, A. New families in the classification of glycosyl hydrolases based on amino acid sequence similarities. *Biochem. J.* **293**, 781-788 (1993).

120 Brada, D., Kerjaschki, D., & Roth, J. Cell type-specific post-Golgi apparatus localization of a "resident" endoplasmic reticulum glycoprotein, glucosidase II. *J Cell Biol* **110**, 309-18 (1990).

121 Hirai, M. & Shimizu, N. Purification of two distinct proteins of approximate Mr 80,000 from human epithelial cells and identification as proper substrates for protein kinase C. *Biochem. J.* **270**, 583-589 (1990).

122 Trombetta, E.S. & Helenius, A. Lectins as chaperones in glycoprotein folding.

Curr. Opin. Struct. Biol. **8**, 587-592 (1998).

123 Alonso, J.M., Santa-Cecilia, A., Chinchetru, M.A., & Calvo, P. Characterization of the maltase activity of glucosidase II from rat liver. *Biol. Chem.* **374**, 977-982 (1993).

124 Alonso, J.M., Santa-Cecilia, A., & Calvo, P. Effect of bromoconduritol on glucosidase II from rat liver. A new kinetic model for the binding and hydrolysis of the substrate. *Eur. J. Biochem.* **215**, 37-42 (1993).

125 Datema, R., Romero, P.A., Legler, G., & Schwarz, R.T. Inhibition of formation of complex oligosaccharides by the glucosidase inhibitor bromoconduritol. *Proc. Natl. Acad. Sci. U.S.A.* **79**, 6787-6791 (1982).

126 Tabaczewski, P. & Stroynowski, I. Expression of secreted and glycosylphosphatidylinositol-bound Qa-2 molecules is dependent on functional TAP-2 peptide transporter. *J. Immunol.* **152**, 5268-5274 (1994).

127 Schagger, H. & von Jagow, G. Blue native electrophoresis for isolation of membrane protein complexes in enzymatically active form. *Anal. Biochem.* **199**, 223-231 (1991).

128 Brada, D. & Dubach, U.C. Isolation of a homogeneous glucosidase II from pig kidney microsomes. *Eur. J. Biochem.* **141**, 149-56 (1984).

129 Coligan, J.E., Kruibeek, A.M., Margulies, D.H., Shevach, E.M., and Strober, W. Isolation and Analysis of Proteins *in* Current Protocols in Immunology. John Wiley & Sons, Inc., New York (1997).

130 Zufferey, R. *et al.* STT3, a highly conserved protein required for yeast oligosaccharyl transferase activity *in vivo*. *EMBO J.* **14**, 4949-4960 (1995).

131 Vijayraghavan, U., Company, M., & Abelson, J. Isolation and characterization of pre-mRNA splicing mutants of Saccharomyces cerevisiae. *Genes Dev.* **3**, 1206-1216 (1989).

132 Romero, P.A. & Herscovics, A. Transfer of nonglucosylated oligosaccharide from lipid to protein in a mammalian cell. *J. Biol. Chem.* **261**, 15936-15940 (1986).

133 Oliver, J. D., Roderick, H. L., Llewellyn, D. H. & High, S. ERp57 functions as a subunit of specific complexes formed with the ER lectins calreticulin and calnexin. *Mol. Biol. Cell* **10**, 2573-2582 (1999).

134 Fields, S. & Song, O. A novel genetic system to detect protein-protein interactions. *Nature* **340**, 245-246 (1989).

135 Gyuris J, Golemis E, Chertkov H, & Brent R. Cdi1, a human G1 and S phase protein phosphatase that associates with Cdk2. *Cell* **75**, 791-803 (1993).

136 Johnsson, N. & Varshavsky, A. Ubiquitin-assisted dissection of protein transport across membranes. *EMBO J.* **13**, 2686-2698 (1994).

137 Aronheim, A., Zandi, E., Hennemann, H., Elledge, S. J. & Karin, M. Isolation of an AP-1 repressor by a novel method for detecting protein-protein interactions. *Mol. Cell Biol.* **17**, 3094-3102 (1997).

138 Schwikowski, B., Uetz, P. & Fields, S. A network of protein-protein interactions in yeast. *Nat. Biotech.* **18**, 1257-1261 (2000).

139 Uetz, P. *et al.* A comprehensive analysis of protein-protein interactions in Saccharomyces cerevisiae. *Nature* **403**, 623-627 (2000).

140 Walhout, A. J., Boulton, S. J. & Vidal, M. Yeast two-hybrid systems and protein interaction mapping projects for yeast and worm. *Yeast* **17**, 88-94 (2000).

141 Walhout, A. J. *et al.* Protein interaction mapping in C. elegans using proteins involved in vulval development. *Science* **287**, 116-122 (2000).

142 Stagljar, I., Korostensky, C., Johnsson, N. & te Heesen, S. A genetic system based on split-ubiquitin for the analysis of interactions between membrane proteins *in vivo*. *Proc. Natl. Acad. Sci. U. S. A.* **95**, 5187-5192 (1998).

143 Broder, Y. C., Katz, S. & Aronheim, A. The Ras recruitment system, a novel approach to the study of protein-protein interactions. *Curr. Biol.* **8**, 1121-1124 (1998).

144 Ehrhard KN, Jacoby JJ, Fu XY, Jahn R & Dohlman HG. Use of G-protein fusions to monitor integral membrane protein-protein interactions in yeast. *Nat. Biotech.* **18**, 1075-1079 (2000).

145 Pelletier, M.F., Bergeron, J.J.M., & Thomas, D.Y. Molecular Chaperone Systems in the Endoplasmic Reticulum . *In* Molecular Chaperones in the Cell: Frontiers in Molecular Biology Series. Peter Lund,ed Oxford University Press. Oxford U.K. (2001)

146 Ellgaard, L., Molinari, M. & Helenius, A. Setting the standards: quality control in the secretory pathway. *Science* **286**, 1882-1888 (1999).

147 Kornfeld, R & Kornfeld, S. Assembly of asparagine-linked oligosaccharides. *Annu. Rev. Biochem.* **54**:631-664

148 Guthrie, C. & Fink, G.R. Guide to yeast genetics and molecular biology. Academic Press. San Diego (1991).

149 Mori, K. *et al.* A 22 bp cis-acting element is necessary and sufficient for the induction of the yeast KAR2 (BiP) gene by unfolded proteins. *EMBO J.* **11**, 2583-2593 (1992).

150 Nantel, A., Mohammad-Ali, K., Sherk, J., Posner, B.I. & Thomas, D.Y. Interaction of the Grb10 adapter protein with the Raf1 and MEK1 kinases. *J. Biol. Chem.* **273**, 10475-10484 (1998).

151 Zheng C.F. & Guan K.L. Properties of MEKs, the kinases that phosphorylate and activate the extracellular signal-regulated kinases. *J. Biol. Chem.* **268**, 23933-23939 (1993).

152 Chapman, R.E. & Walter, P. Translational attenuation mediated by a mRNA intron. *Curr. Biol.* **7**, 850-859 (1997).

153 Travers, K. J. *et al.* Functional and genomic analyses reveal an essential coordination between the unfolded protein response and ER-associated degradation. *Cell* **101**, 249-258 (2000).

154 Livnah, O. *et al.* An antagonist peptide-EPO receptor complex suggests that receptor dimerization is not sufficient for activation. *Nat. Struct. Biol.* **5**, 993-1004 (1998).

155 Remy, I., Wilson, I. A. & Michnick, S. W. Erythropoietin receptor activation by a ligand-induced conformation change. *Science* **283**, 990-993 (1999).

156 Wilson, I. A. & Jolliffe, L. K. The structure, organization, activation and plasticity of the erythropoietin receptor. *Curr. Opin. Struct. Biol.* **9**, 696-704 (1999).

157 Liu, C. Y., Schroder, M. & Kaufman, R. J. Ligand-independent dimerization activates the stress response kinases IRE1 and PERK in the lumen of the endoplasmic reticulum. *J. Biol. Chem.* **275**, 24881-24885 (2000).

158 Ellgaard L *et al.* NMR structure of the calreticulin P-domain. *Proc. Natl. Acad. Sci. U. S. A.* **98**, 3133-3138 (2001).

159 Ferrari, D. M. & Soling, H. D. The protein disulfide-isomerase family unravelling a string of folds. *Biochem. J.* **339**, 1-10 (1999).

160 Shahinian, S. *et al.* Involvement of protein N-glycosyl chain glucosylation and processing in the biosynthesis of cell wall beta-1,6-glucan of *Saccharomyces cerevisiae*. *Genetics* **149**, 843-856 (1998).

161 Helenius A. Quality control in the secretory assembly line. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* **356**, 147-150 (2001).

162 Mesaeli, N. *et al.* Calreticulin is essential for cardiac development. *J. Cell Biol.* **144**, 857-868 (1999).

163 De Praeter, C. M. *et al.* A novel disorder caused by defective biosynthesis of N-linked oligosaccharides due to glucosidase I deficiency. *Am. J. Hum. Genet.* **66**, 1744-1756 (2000).

164 Parlati, F., Dignard, D., Bergeron, J. J. & Thomas, D. Y. The calnexin homologue *cnx1+* in *Schizosaccharomyces pombe*, is an essential gene which can be complemented by its soluble ER domain. *EMBO J.* **14**, 3064-3072 (1995).

165 Jannatipour, M. & Rokeach, L. A. The *Schizosaccharomyces pombe* homologue of the chaperone calnexin is essential for viability. *J. Biol. Chem.* **270**, 4845-4853 (1995).

166 Fanchiotti, S., Fernandez, F., D'Alessio, C. & Parodi, A. J. The UDP-Glc:Glycoprotein glucosyltransferase is essential for *Schizosaccharomyces pombe* viability under conditions of extreme endoplasmic reticulum stress. *J. Cell. Biol.* **143**, 625-635 (1998).

167 Fernandez, F.S., Trombetta, S.E., Hellman, U., & Parodi, A.J. Purification to homogeneity of UDP-glucose:glycoprotein glucosyltransferase from Schizosaccharomyces pombe and apparent absence of the enzyme fro Saccharomyces cerevisiae. *J. Biol. Chem.* **269**, 30701-30706 (1994).

168 Jakob, C.A., Burda, P., te Heesen, S., Aebi, M., & Roth, J. Genetic tailoring of N-linked oligosaccharides: the role of glucose residues in glycoprotein processing of Saccharomyces cerevisiae *in vivo*. *Glycobiology* **8**, 155-164 (1998).

169 Parlati, F., Dominguez, M., Bergeron, J. J. & Thomas, D. Y. *Saccharomyces cerevisiae CNE1* encodes an endoplasmic reticulum (ER) membrane protein with sequence similarity to calnexin and calreticulin and functions as a constituent of the ER quality control apparatus. *J. Biol. Chem.* **270**, 244-253 (1995).

170 Werr, M. & Prange, R. Role for calnexin and N-linked glycosylation in the assembly and secretion of hepatitis B virus middle envelope protein particles. *J. Virol.* **72**, 778-782 (1998).

171 Mehta, A., Lu, X., Block, T.M., Blumberg, B.S. & Dwek, R.A. Hepatitis B virus (HBV) envelope glycoproteins vary drastically in their sensitivity to glycan processing: evidence that alteration of a single N-linked glycosylation site can regulate HBV secretion. *Proc. Natl. Acad. Sci. U. S. A.* **94**, 1822-1827 (1997).

172 Waechter, C.J. & Lennarz, W.J. The role of polyprenol-linked sugars in glycoprotein synthesis. *Annu. Rev. Biochem.* **45**, 95-112 (1976).

173 Turco, S.J., Stetson, B. & Robbins, P.W. Comparative rates of transfer of lipid-linked oligosaccharides to endogenous glycoprotein acceptors *in vitro*. *Proc. Natl. Acad. Sci. U. S. A.* **74**, 4411-4414 (1977).

174 Grinna, L.S. & Robbins, P.W. Substrate specificities of rat liver microsomal glucosidases which process glycoproteins. *J. Biol. Chem.* **255**, 2255-2258 (1980).

175 Grinna, L.S. & Robbins, P.W. Glycoprotein biosynthesis. Rat liver microsomal glucosidases which process oligosaccharides. *J. Biol. Chem.* **254**, 8814-8818 (1979).

176 Treml, K., Meimaroglou, D., Hentges, A. & Bause, E. The alpha- and beta-subunits are required for expression of catalytic activity in the hetero-dimeric glucosidase II complex from human liver. *Glycobiology* **10**, 493-502 (2000).

177 Arendt, C. W. & Ostergaard, H. L. Two distinct domains of the beta-subunit of glucosidase II interact with the catalytic alpha-subunit. *Glycobiology* **10**, 487-492 (2000).

178 Baldwin, T. A., Gogela-Spehar, M. & Ostergaard, H. L. Specific isoforms of the resident endoplasmic reticulum protein glucosidase II associate with the CD45 protein-tyrosine phosphatase via a lectin-like interaction. *J. Biol. Chem.* **275**, 32071-32076 (2000).

179 Trowbridge, I. S. & Thomas, M. L. CD45: an emerging role as a protein tyrosine phosphatase required for lymphocyte activation and development. *Ann. Rev. Immunol.* **12**, 85-116 (1994).

180 Pelletier, J.N., Arndt, K.M., Pluckthun, A. & Michnick, S.W. An *in vivo* library-versus-library selection of optimized protein-protein interactions. *Nat. Biotech.* **17**, 683-690 (1999).

181 Rossi, F.M., Blakely, B.T. & Blau, H.M. Interaction blues: protein interactions monitored in live mammalian cells by beta-galactosidase complementation. *Trends Cell Biol.* **10**, 119-122 (2000).

182 Goffeau A *et al.* Life with 6000 genes. *Science* **274**, 543-547 (1996).

183 The *C. elegans* Sequencing Consortium. Genome sequence of the nematode *C. elegans*: a platform for investigating biology. *Science* **282**, 2012-2018 (1998).

184 Dubuisson, J. Folding, assembly and subcellular localization of hepatitis C virus

glycoproteins. *Curr. Top. Microbiol. Immunol.* **242**, 135-148 (2000).

185 Deleersnyder, V. *et al.* Formation of native hepatitis C virus glycoprotein complexes. *J. Virol.* **71**, 697-704 (1997).

186 Wunschmann, S., Medh, J. D., Klinzmann, D., Schmidt, W. N. & Stapleton, J. T. Characterization of hepatitis C virus (HCV) and HCV E2 interactions with CD81 and the low-density lipoprotein receptor. *J. Virol.* **74**, 10055-10062 (2000).