# Characterization of calnexin in Saccharomyces cerevisiae and Schizosaccharomyces pombe

Francesco Parlati

Department of Biology McGill University July, 1996

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

© Francesco Parlati 1996



National Library of Canada

Acquisitions and Bibliographic Services Branch

des services bibliographiques
395 rue Wetlington

du Canada

395 Wellington Street Ottawa, Ontario K1A 0IN4 395, rue Wellington Ottawa (Ontario) K1A NN4

Bibliothèque nationale

Direction des acquisitions et

Your lile Volre rélérence

Our lile Notre référence

The author has granted an irrevocable non-exclusive licence allowing the National Library of Canada to reproduce, loan, distribute or sell copies of his/her thesis by any means and in any form or format, making this thesis available to interested persons.

L'auteur a accordé une licence irrévocable et non exclusive à la Bibliothèque permettant nationale Canada du de reproduire, prêter, distribuer ou vendre des copies de sa thèse de quelque manière et sous quelque forme que ce soit pour mettre des exemplaires de cette disposition thèse à la des personnes intéressées.

The author retains ownership of the copyright in his/her thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without his/her permission. L'auteur conserve la propriété du droit d'auteur qui protège sa thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

ISBN 0-612-19762-X



### ABSTRACT

In eukaryotes, the endoplasmic reticulum is the site where folding of secretory proteins and the assembly of multimeric cell surface receptors take place. These processes are mediated by molecular chaperones that include the ER membrane bound chaperone calnexin and the sequence related calreticulin. Using a PCR strategy, a homologue for the mammalian calnexin/calreticulin family, CNE1, was isolated in S. cerevisiae. The CNE1 gene product, Cnelp, is an integral membrane glycoprotein of the ER. Disruption of the CNE1 gene did not lead to inviable cells or to gross effects on the levels of secreted wild type proteins. However, in CNE1 disrupted cells, there was an increase in the cell-surface expression of a normally intracellularly retained temperature sensitive mutant of the  $\alpha$ -pheromone receptor, Ste2-3p. In addition, an increase in the secretion of heterologously expressed mammalian  $\alpha_1$ -antitrypsin was also observed in *CNE1* disrupted cells. In order to study calnexin function in another genetically manipulable organism, a Schizosaccharomyces pombe calnexin homologue was sought. Using a similar PCR strategy, a S. pombe calnexin homologue,  $cnxl^+$ , was identified. The  $cnxl^+$  gene product, Cnxlp, was shown to be a calcium binding type I integral membrane glycoprotein. Unlike the sequence related S.cerevisiae CNEI gene, the cnx1<sup>+</sup> gene was essential for cell viability. Full length Cnx1p was able to complement the  $cnxl^+$  gene disruption but full length mammalian calnexin could not. The ER lumenal domain of Cnx1p, which was secreted from cells, was capable of complementing the cnx1::ura4<sup>+</sup> lethal phenotype. Both wild type PI M1 (Val 213)  $\alpha_1$ antitrypsin and the ER retained PI Z variant were expressed in S.pombe cells. As in mammalian cells, wild type  $\alpha_1$ -antitrypsin was normally secreted whereas the PI Z variant was retained intracellularly. Rescue of the secretion defective phenotype of the PIZ variant occurred in *S.pombe* strains which did not express the Cnx1p cytosolic tail. *S.pombe* cells expressing both full length and truncated mutants of Cnx1p did not secrete the PIZ variant, demonstrating that the presence of full length Cnx1p will maintain the integrity of the ER quality control machinery. Therefore, like mammalian calnexin, Cnelp and Cnxlp appear to function as constituents of the yeast ER protein quality control apparatus.

### RESUME

Le repliement des protéines sécrétées ainsi que l'assemblage des récepteurs multimériques s'effectue dans le réticulum endoplasmique. Ces processus sont assistés par les chaperons moléculaires, tel que la calnexine et la calreticuline. L'homologue de la calnexine et calreticuline, CNE1, chez Sacchomomyces cerevisiae a été identifié par la méthode d' amplification en chaine, 'PCR'. Le produit du gene CNEI, Cne1p, est une glycoprotéine transmembranaire du réticulum endoplasmique. La délétion du gène CNEI n'affecte ni la croissance des cellules, ni la sécrétion de protéines sauvages. Cependant, il y a une augmentation des niveaux sécrétés de la protéine mutante Ste2-3p et de la protéine hétérologue  $\alpha_1$ -antitrypsine. Afin d'étudier la fonction de la protéine calnexine dans une seconde espèce manipulable génétiquement, l'homologue chez Schizosaccharomyces pombe a été recherché. En utilisant la même technique de 'PCR', nous avons pu cloner le gène  $cnxl^+$ . Le produit du gène  $cnxl^+$ , Cnx1p, est une glycoprotéine de type I transmembranaire liant le calcium. Contrairement au gène CNEI de S. cerevisiae, le gène  $cnx1^+$  est essentiel pour la viabilité de S.pombe. La forme complète ainsi que le domaine lumenal de Cnx1p complémentent la délétion du gène  $cnx1^+$ , contrairement à la calnexine de mammifère. Les formes sauvages (PI M1 (Val 213)) et mutantes (PI Z) de l'  $\alpha_1$ antitrypsine ont été exprimées chez S.pombe. Chez les mammifères et S.pombe, la forme PI M1 (Val 213) est sécrétée normalement, alors que la forme PI Z est retenue dans la cellule. Dans une souche de S.pombe qui n'exprime pas le domaine cytoplasmique de Cnx1p, la secrétion de la forme PIZ a lieu. Une souche de S.pombe qui exprime la forme complète et la forme tronquée de Cnx1p ne restore pas la sécrétion de la forme PI Z de l' $\alpha_1$ antitrypsine. Ce qui démontre que l'expression de la forme complète de Cnx1p maintient l'intégrité de l'appareil de contrôle de qualité du réticulum endoplasmique. Finalement, la calnexine chez S.pombe et S.cerevisiae, comme chez les mammifères, est un constituant de l'appareil de contrôle de qualité du réticulum endoplasmique.

#### PREFACE

This thesis has been assembled in accordance with the regulations of the Faculty of Graduate Studies and Research of McGill University. This thesis includes an Abstract (*Résumé*), an Introduction (Chapter 1), a Results section (Chapters 2, 3,4), a Conclusion section (Chapter 5) and a Reference section. Sections of this thesis have been published or submitted for publication, therefore Chapters 2, 3 and 4 contain their respective Abstract, Introduction, Materials & Methods, Results and Discussion sections.

In this work I present my original contribution to the identification and characterization of the calnexin homologue CNEI in Saccharomyces cerevisiae and  $cnxI^+$  in Schizosaccharomyces pombe. Furthermore, evidence for a role in ER protein quality control for yeast calnexin is presented. This study provides new insight into the role of calnexin as a member of an ER protein quality control apparatus.

The work presented here is largely my own, with the following exceptions: i) Michel Dominguez performed the analytical and differential subcellular centrifugation, as well as the calcium overlay blot in Chapter 2, ii) David Y. Thomas performed the screening of a *S.pombe* library to obtain the full length  $cnxl^+$  genomic clone, iii) Daniel Dignard sequenced the  $cnxl^+$  genomic clone.

Chapter 2 has been published as a paper by F. Parlati, M. Dominguez, J. J. M. Bergeron and D. Y. Thomas in the *Journal of Biological Chemistry*, **270**: 244-253 in 1995, and is used by permission of the American Society for Biochemistry and Molecular Biology. Chapter 3 has been published as a paper by F. Parlati, D. Dignard, J. J. M. Bergeron and D. Y. Thomas in the *EMBO Journal*, **14**: 3064-3072 in 1995, and is used by permission of Oxford University Press. Chapter 4 has been submitted to the *Journal of Biological Chemistry* as a paper by F. Parlati, R. Sifers, J. J. M. Bergeron and D. Y. Thomas in June, 1996. The text of these papers are an integral part of this thesis, and they are presented in a logical manner.

## ACKNOWLEDGMENTS

I am indebted to my supervisor, Dr. Dave Thomas, whose vision and generosity have made this work possible. I am truly grateful for Dr. Thomas' encouragement and financial assistance for my participation at numerous scientific conferences. I am appreciative of his support and belief in my work.

I have been fortunate to have Dr. John Bergeron as a co-supervisor for the duration of this work. His constant and unwavering enthusiasm gave me confidence and strength to advance my work. His passion for science is insurmountable and an example for all.

I would also like to thank members of my supervisory committee, Dr. Howard Bussey and Dr. Peter Hechtman, for their guidance throughout the progress of this work.

Special thanks go to colleagues and friends in both Dr. Dave Thomas' and Dr. John Bergeron's lab. In particular, I am thankful to Josée Ash, Daniel Dignard, Pam Silver, Dr. Wei-Jia Ou and Daniel Tessier for their invaluable technical help. I would like to thank Dr. Martine Raymond and Dr. Malcolm Whiteway, whose scientific astuteness taught me a great deal throughout my graduate studies. Thanks to Michel Dominguez and Dr. Rick Hemming for endless scientific and not-so scientific conversations. Daily lab life would not be the same without them.

I would also like to thank Doreen Harcus, Dr. Csilla Csank, Dr. Cunle Wu, Dr. Thomas Leeuw, Dr. Dorris Germain for listening to my complaints all these years. Special thanks go to Dr. Miho Shida, Dr. Rick Hemming, Veena Sangwan and Vika Lytvyn for reading my thesis. I would like to acknowledge Michel Dominguez, Dr. Virginie Ansanay and Dr. André Zapun for helping in the French translation of the Abstract. I have made so many friends throughout these years, whose support I must acknowledge. To my dear friends Sophia Kazanis, Vika Lytvyn, Steve Hosein, Dr. Paola Domizio, Anita DeBellis, Tina Carriero and of course Veena Sangwan, thank you.

I don't know how to begin to thank Dr. Ana Paula Lima whose friendship was truly a life altering experience. She has always been a rock for me, even now, thousands of miles away. To Dr. Warren Winkelman, thank you for your friendship and understanding that has allowed me to keep my sanity through these last and very tough years.

To my brother and sister, thanks for putting up with me. Vorrei ringraziare miei genitori, Antonio e Angela. Grazie per vostro affeto, fierta e devozione per tutti questi anni.

# TABLE OF CONTENTS

P	Α	G	E
---	---	---	---

Abstract	•	•	•	•	•	•	•	•	•	1
Résumé	•	•								2
Preface		•		٠					•	3
Acknowled	gments	•					•		•	4
Table of Co	ontents			•			•		•	5
List of Figu	ires	•		•			•	•	•	8
List of Tabl	les .	•		•		•			•	11
List of Abb	reviations	•	•	•	•	•	•	•	•	12
Chapter 1:	Introducti	on to th	e mol	ecular c	haperon	e calne:	xin and s	sequenc	с	
	related cal	reticulii	n.	•			•			14
1.1 Protein	folding					•	•		•	15
1.1.	1 Protein	folding	in the	cell	•	•	٠	•	•	16
1.1.	2 Molecul	lar chap	erone	s: HSP	60 and	HSP 70	) familie	s.	•	16
1.2 Protein	modificati	ion and	foldir	ng in the	e ER		•		•	17
1.2.	I Modific	ations t	o the j	primary	structur	re.	•		•	17
1.2.	2 Folding	catalys	ts and	molecu	ılar cha <mark>ı</mark>	perones	in the E	R.	•	18
1.2.	3 ER qual	ity con	trol	٠	•		•	•		19
1.3 Calnexi	n mode of	action		•						21
1.3.	! Evidenc	e for ca	Inexi	ı bindin	g to N-l	inked				
	GlcNA	c <sub>2</sub> Mang	Glc <sub>1</sub> i	ntermed	liate	•	•		•	21
1.3.	2 Role of	UDP-g	lucose	: glycoj	protein	glucosy	ltransfe	rase		22
1.3.	3 Require	ment fo	or cale	ium	•		•		•	23
1.3.	4 Require	ment fo	or ATI	2	•	•	•	•	•	23
1.3.	5 Exceptio	ons to tl	he calı	nexin-ol	igosacc	haride l	oinding	model		25
1.4 Evidenc	e for caln	exin's a	and ca	Ireticuli	n's role	as mole	cular ch	aperon	es.	25
1.4.	l Calnexia	n assoc	iation	with m	onomeri	ic wild (	type pro	teins		26
1.4.	2 Calretic	ulin ass	ociati	on with	monom	eric wi	ld type p	oroteins		26
1.4.	3 Calnexi	n assoc	iation	with as	semblin	g protei	in comp	lexes	•	27
1.4.	4 The acti	on of B	iP and	i calnex	in in the	e foldin	g of wil	d type p	roteins	28

1.5 Stable calnexin	associatio	on wit	h misfo	ded pro	oteins	•	•	
1.5.1 Stable	calnexin	assoc	iation w	ith mist	folded r	nonomer	ie	
prote	ins		•			•		
1.5.2 Calne	xin assoc	iation	with the	e subuni	its of ur	assembl	ed	
prote	in comple	exes	٠	•		•		•
1.5.3 The as	ssociatior	n of Bi	P and c	alnexin	with m	isfolded	protein	s and
unass	sembled p	protein	subuni	ts .	•	•	•	
1.5.4 Protei	n retentio	n in th	ne ER ar	nd the ro	ble of ca	alnexin		
in pro	otein degr	adatio	m.	•	•	•	•	
	1 .4 14	e	ED					
1.6 The callexin/ca	ireticulin	family	y: an ER	chaper	one sup	erlamily	•	•
1.6.1 Calne	xin	•	•	•	•	•	•	•
1.6.2 Calne	xin signa	ture so	equence	s .	•	•	•	•
1.6.3 Calrel		•	•	•	٠	•	•	·
1.6.4 Calret	iculin sig	nature	e sequen	ces	٠	•	•	•
1.6.5 Calne	xin and c	alretic	ulin con	npared	•		•	•
1.7 Aim of this wo	rk.			•		•		
Chapter 2: Sacchar	romvees o	erevis	siae CNI	El enco	des an l	ER meml	orane r	rotein
with sec	mence si	milarit	v to cali	nexin an	id calret	iculin an	d fune	lions
as a cor	astituent c	of the I	ER quali	ity conti	rol anna	ratus	a rane	
							•	•
2.1 Abstract .			•	•			•	
2.2 Introduction				•	•	•	•	•
2.3 Materials and N	Aethods		•		•	•	•	•
2.4 Results .					•	•	•	

is an essential gene which	can b	e compl	lemento	ed by its	soluble	;	
ER domain .	•	•	•	•	•	•	78
3.1 Abstract			•				79
3.2 Introduction			•			•	80
3.3 Materials and Methods .	•		•	•	•	•	82
3.4 Results				•	•	•	85
3.5 Discussion	•		•		•	•	97
Chapter 4: Rescue of the secretion de	fectiv	e pheno	type of	nutant			
$\alpha_1$ -antitrypsin by C-termin	nal del	letions of	of calne	xin in			
Schizosaccharomyces por	nbe	•	•	•	•	•	101
3.1 Abstract	•	•	•			•	102
3.2 Introduction		•	•			٠	103
3.3 Materials and Methods .		•	•			•	105
3.4 Results	•	•	•		•	٠	108
3.5 Discussion	•		•		•	•	115
Chapter 5: Conclusions .	•	•	٠	•	•	•	119
5.1 Identity of S.cerevisiae CNE1		•		•	•	•	120
5.2 Identity of S. pombe cnx1+	•		•	•		٠	121
5.3 Phenotype of S.cerevisiae Cne1p	)	•	•	•	•	•	122
5.4 Phenotype of S. pombe cnx1+	•		•				123
5.5 S.cerevisiae and S.pombe calnex	in con	npared		•	•	•	124
5.6 Yeast calnexin mode of action	•	٠	•	•		•	125
5.7 Future experiments .	•				•	•	128
References			•	•	•	•	130

# Chapter 3: The calnexin homologue cnx1+ in Schizosaccharomyces pombe,

.

# LIST OF FIGURES

Figure 1	Interaction between calnexin/calreticulin and		
	N-linked oligosaccharide intermediates in the ER .		24
Figure 2	Conserved calcium binding and cysteine signature sequences in canine calnexin, mouse calreticulin		
	and S.pombe Cnx1p	•	<u>3</u> 9
Figure 3	Comparison of the conserved regions in calnexin and calreticulin homologues		45
Figure 4	The CNE1 PCR cloning strategy		57
Figure 5	Amino acid alignments of S. cerevisiae CNE1, canine		
	calnexin and mouse calreticulin		58
Figure 6	Hydrophobicity plot and topology of Cnelp	•	59
Figure 7	Cnelp is an integral membrane glycoprotein	•	61
Figure 8	Comparison of the distribution of the ER marker enzyme		
	NADPH cytochrome c reductase and Cne1p	•	62
Figure 9	Isopycnic sucrose density gradient centrifugation analysis of the distribution of Kar2p and Cne1p in the parent ML		
	fraction		63
Figure 10	Sucrose density gradient analysis	•	64
Figure 11	Double immunofluorescence of Cne1p and Kar2p in		
	S.cerevisiae by epifluorescence and confocal microscopy	•	66
Figure 12	Identification of <sup>45</sup> Ca binding proteins in S.cerevisiae ER		67

Figure 13	Gene disruption of <i>S.cerevisiae CNE1</i> and evaluation by			<b>60</b>
	Northern blot and Western blot	•	•	69
Figure 14	Acid phosphatase secretion	•	•	70
Figure 15	Halo assay for $\alpha$ -pheromone production			71
Figure 16	Effect of Cnelp on the secretion of $\alpha_1$ -antitrypsin	•	•	73
Figure 17	DNA sequence and deduced amino acid			
	sequence of the $cnxI^+$ gene	•	•	86
Figure 18	Hydrophobicity plot and topology of Cnx1p .			87
Figure 19	Induction of <i>cnx1</i> <sup>+</sup> by heat shock and calcium ionophore	:		
	A23187	•	•	89
Figure 20	Identification of S.pombe Cnx1p as a calcium binding			
	protein	•	•	90
Figure 21	Cnx1p is an integral membrane glycoprotein .	•	•	92
Figure 22	Gene disruption of $cnxI^+$	•	•	95
Figure 23	Complementation of the cnx1::ura4 + disruption in haplo	id		
	cells	•	•	96
Figure 24	Comparison of predicted high affinity calcium binding			
	motifs and cysteine motifs in <i>S.pombe</i> Cnx1p, A thaliana, calnexin, canine calnexin, <i>S.cerevisiae</i>			
	Cnelp and mouse calreticulin proteins	•	•	98
Figure 25	Rescued secretion of the PIZ variant expressed in			
	S.pombe cells			109

Figure 26: Intracellular expression of the wild type $\alpha_1$ -antitrypsin		
PI M1 (Val 213) variant, the PI Z variant and Cnx1p in	all	
four <i>cnx1</i> <sup>+</sup> backgrounds	•	112
Figure 27: Evaluation of resident ER lumenal protein BiP secretion	1.	114
Figure 28: Quality control regulation by Cnx1p	•	[18

•

# LIST OF TABLES

Table 1	Percent identity amongst animal, plant and fungi calnexins .	34
Table 2	Selected features for animal, plant and fungi calnexins	38
Table 3	Percentage identity amongst animal and plant calreticulins .	41
Table 4	Selected features for animal and plant calreticulins	43
Table 5	Effect of <i>CNE1</i> on the secretion of $\alpha_1$ -antitrypsin in <i>S.cerevisiae</i>	74
Table 6	Effect of <i>CNE1</i> on the secretion of Ste2-3 protein	74
Table 7	Description of constructs used in chapter 4	107

-.

# List of Abbreviations

ATP	adenosine triphosphate
BCIP	5-bromo-4-chloro-3-indolyl phosphate
BiP	binding immunoglobulin protein
BSA	bovine serum albumin
CFTR	cystic fibrosis transmembrane regulator
COS	CV-1 origin, SV40
CST	castanospermine
DAPI	4',6-diamidino-2-phenyl-indole
DJM	1-deoxynojirimycin
DMM	deoxymannojirimycin
DNA	deoxyribonucleic acid
DTT	dithiothreitol
E.coli	Escherichia coli
EGTA	ethylene glycol-bis(b-aminoethyl ether)N,N,N',N'-tetraacetic acid
EMM	Edinburgh minimal media
Endo-H	endoglycosidase H
ER	endoplasmic reticulum
FITC	fluorescein isothiocyanate
FH	familial hypercholesterolemia
Glc	glucose
GlcNAc	glucosamine
gp160	glycoprotein of 160 kilodaltons
GPI	glycosylphosphatidylinositol
GRP	glucose regulated protein
GST	glutathione-S-transferase
НА	hemagglutinin (influenza virus)
HSP	heat shock protein
kb	kilo bases
kDa	kilo daltons
Ig	immunoglobulin
IPTG	isopropylthio-β-D-galactoside
β <sub>2</sub> m	β <sub>2</sub> microglobulin
LDL	low density lipoprotein

Man	mannose
MBS	2-{N-Morpholino]ethane-sulfonic acid buffered saline
MDCK	Madien-Darby canine kidney
MHC I	major histocompatibility complex class 1
MHC II	major histocompatibility complex class 2
mlg	membrane immunoglobulin
ML	large granule
mRNA	messenger ribonucleic acid
N	nuclear pellet
NADPH	nicotinamide adenine dinucleotide phosphate
NBT	p-nitro blue tetrazolium chloride
Р	microsomal pellet
PAGE	polyacrylamide gel electrophoresis
PCR	polymerase chain reaction
PDI	protein disulfide isomerase
PMSF	phenylmethylsulphonyl fluoride
RNA	ribonucleic acid
S	supernatant
SC	synthetic complete
S.cerevisiae	Saccharomyces cerevisiae
SDS	sodium dodecylsulphate
S.pombe	Schizosaccharomyces pombe
SSR	signal sequence receptor
TBS	tris buffered saline
TCR	T-cell receptor
TRiC	TCP1 Ring Complex
UGGT	UDP-glucose: glycoprotein glucosyltransferase
VSV G	vesicular stomatitis virus G protein
YPD	yeast peptone dextrose

.

# **CHAPTER 1**

# Introduction to the molecular chaperone calnexin and sequence related calreticulin

# Introduction

In eukaryotic cells, secretory proteins and proteins targeted to various organelles within the secretory pathway are first translocated into the endoplasmic reticulum (ER). Upon translocation into the ER, some proteins receive post-translational modifications such as signal peptide cleavage, N- and O- linked glycosylation, disulfide bond formation, peptidyl proline *cis-trans* isomerization and GPI anchor addition. Once proteins are properly folded, they are permitted to exit the ER and continue to receive other post-translational modifications as they travel through the Golgi cisternae. From the Trans Golgi Network, proteins are subsequently transported to lysosomes (or vacuoles), storage vesicles or the plasma membrane (for reviews, see Palade, 1975; Huttner and Tooze, 1989; Rabouille and Nilsson, 1995).

# 1.1 Protein folding

Within the cytosol and intracellular organelles, newly synthesized proteins are permitted to fold in order to achieve a mature conformation. For a simple monomeric protein, the information required for proper folding is present within its primary sequence, also known as a protein's primary structure. Segments within the primary structure form  $\alpha$ -helices,  $\beta$ -sheets or multiple turns which define the secondary structure of a polypeptide. Specific interactions between the secondary structures define the tertiary structure. For many proteins, several polypeptides must interact in order to form a biologically active complex, known as the quaternary structure of a protein (for reviews see Hartl and Martin, 1995; Gething and Sambrook, 1992).

In a classical experiment, Anfinsen showed that *in vitro*, ribonuclease could be denatured into a random structure and refolded into an active structure (Anfinsen, 1973). *In vitro* refolding has also been observed for other small monomeric proteins (Gething and Sambrook, 1992). However, *in vitro* refolding is generally slow and inefficient, requiring protein concentrations and physical conditions normally not present in cells. Moreover, *in vitro* refolding for many proteins follows less productive pathways which often leads to incompletely folded protein aggregates. Therefore, *in vitro* protein folding may not accurately mimic protein folding *in vivo*.

### 1.1.1 Protein folding in the cell

Although the information required for a polypeptide to fold into a mature state is present in its primary sequence (Anfinsen, 1973), *in vivo*, proteins exist in a concentrated solution and are prone to form aggregates while hydrophobic segments are exposed (Gething and Sambrook, 1992). Several classes of cellular proteins assist protein folding, thereby preventing newly synthesized proteins from aggregating. These proteins are collectively known as chaperones and are present in the cytosol and several organelles. The most encompassing definition of a molecular chaperone has been given by Hendrick and Hartl and it states: '(a molecular chaperone is) a protein that binds to and stabilizes an otherwise unstable conformer of another protein, and by controlled binding and release of the substrate, facilitates its correct fate *in vivo* be it folding, oligomeric assembly, transport to a particular compartment, or controlled switching between active/inactive conformation.' Two major families of chaperones, the HSP 60 and HSP 70 families as well as their comembers, have been identified (Hendrick and Hartl, 1993; Hartl and Martin, 1995).

### 1.1.2 Molecular chaperones: HSP 60 and HSP 70 families

Many molecular chaperones were initially identified as stress induced or <u>heat-shock</u> proteins (HSP). However, chaperones are involved in cellular processes under normal and stress conditions. In eukaryotes, HSP 60 is present within mitochondria and chloroplasts and is termed chaperonin 66 (cpn60). In bacteria, the HSP 60 homologue is cytosolically located and is termed GroEL. GroEL consists of 14 identical subunits, forming two backto-back cavities with a seven fold symmetry. Each cavity can accommodate an entire protein of less than 50 kDa or a segment of a polypeptide chain in order to facilitate its folding. GroEL has no obvious protein specificity. Once occupied, the cis-ring then binds the 7-mer GroES (HSP 10 or cpn10 in eukaryotes). A cycle of GroES binding and ATP hydrolysis on the cis-ring and subsequent ATP hydrolysis on the trans-ring is required for productive binding and release of substrate by GroEL. In the eukaryotic cytosol, the distantly related TRiC (TCP1 Ring Complex) has also been ascribed a chaperonin function. TRiC's chaperonin activity is also ATP dependent and its function appears to be specialized for the folding of cytoskeletal proteins, such as tubulin and actin. Since TRiC is specific for cytoskeletal proteins, it is not clear if cytosolic proteins require different cytosolic chaperonins for folding (Hendrick and Hartl, 1993; Sailbil, 1996; Clark, 1996).

The second chaperone family, HSP 70 (DnaK in bacteria), binds polypeptides as they emerge from the ribosomes in order to prevent protein aggregation (see Hendrick and Hartl, 1993; Brodsky, 1996). HSP 70 exists as a monomer, and likely binds to hydrophobic stretches of amino acids (Flynn *et al.*, 1991). Substrates for cytosolic HSP 70 include denatured cytosolic proteins and mitochondrial proteins. Proteins that are posttranslationally translocated into the ER are maintained in a denatured state by cytosolic HSP 70 prior to import (Clark, 1996; Brodsky, 1996). Upon translocation, ATP hydrolysis dependent polypeptide release from HSP 70 is mediated by a cofactor, HSP 40 (DNAJ in bacteria, Sis1p and Ydj1p in the yeast cytosol) (Rassow *et al.*, 1995; Schlenstedt *et al.*, 1995; Brodsky, 1996).

Other members of the HSP 70 family have been found in the lumen of the mitochondria (mt-HSP 70), chloroplasts (cl-HSP 70) and the ER (BiP or binding immunoglobulin protein) (Brodsky, 1996). Lumenal HSP 70 is required to drive translocation of nascent polypeptides (Sanders *et al.*, 1992; Brodsky *et al.*, 1993; Brodsky, 1996). Lumenal DNAJ homologues have also been implicated in the translocation of polypeptides, which include the soluble Isp45p in the yeast mitochondrial lumen and membrane bound Sec63p in the yeast ER. (Brodsky, 1996). Lumenal HSP 70 also mediates the folding of newly translocated proteins and is discussed below.

# 1.2 Protein modification and folding in the ER

Proteins traversing the secretory pathway are translocated into the ER in an unfolded state where they undergo protein modification, folding and quality control. The ER contains a concentrated pool of factors which are involved in these processes.

### 1.2.1 Modifications to the primary structure

Proteins undergo modifications to their 'primary structure' upon translocation into the ER. First, signal peptidase removes the N-terminal hydrophobic ER targeting sequence. Second, asparagines present within the consensus Asp-X-Ser/Thr motif are generally covalently modified by the addition of a GlcNAc<sub>2</sub>Man<sub>9</sub>Glc<sub>3</sub> sugar moiety. This moiety is then modified within the ER and Golgi (Figure 1). Third, O-glycosylation is initiated in the ER with the addition of 2 to 5 mannose residues on some serine and/or threonine residues. The presence of large amounts of serine or threonine residues is the only conserved feature in O-glycosylated proteins (Herscovics and Orlean, 1993). Fourth, addition of a glycosylphosphatidylinositol (GPI) anchor at the C-terminal domain of some secretory proteins also occurs in the ER. Proteins subjected to this modification have a hydrophobic stretch of amino acids at their C-terminus. The hydrophobic character and not the primary amino acid sequence of the C-terminus dictates whether proteins receive this modification. Addition of the GPI anchor is followed by removal of amino acids at the C-terminal side of the addition site resulting in a membrane anchored protein (Nuoffer *et al.*, 1993; Cross, 1990).

## 1.2.2 Folding catalysts and molecular chaperones in the ER

Modification by signal peptidase, glycosylation and GPI addition are the first steps in the folding of a protein in the ER. The protein folding process is aided by two classes of resident ER proteins: folding catalysts and molecular chaperones. For many proteins, the rate limiting step in folding is the formation of disulfide bonds or peptidyl proline *cis-trans* isomerization. In eukaryotes, protein disulfide isomerase (PDI), and closely related ERp 72 will accelerate the rate of disulfide bond formation (Gething and Sambrook, 1992; Freedman *et al.*, 1994; Freedman, 1995). Other ER homologues of PDI that have been identified are ERp 60 and P5 in mammals and Pdi1p, Eug1p, Mrp1p and Mrp2p in yeast (Bardwell and Beckwith, 1993; Tachikawa *et al.*, 1995). For other proteins, isomerization of peptidyl-proline bonds is a limiting step in their folding. Peptidyl proline *cis-trans* isomerases (PPIs) are responsible for isomerizing certain proline-peptidyl bonds into the *cis* or *trans* conformation. Cyclophilins and FK binding proteins (FKBP) carry out this function in the cell. PPIs are ubiquitous and have been identified in the cytosol, mitochondria and ER (Frigerio and Pelham, 1993; Schmid, 1993; Freedman, 1995).

In addition to its role in protein translocation, BiP has been implicated in protein folding in the ER. *In vitro* studies have shown that BiP binds a variety of hydrophobic peptides, with a preference for heptapeptides (Flynn *et al.*, 1991). It was first shown that BiP associated with unassembled IgG subunits (Hendershot *et al.*, 1987; Hendershot, 1990; Knittler and Haas, 1992) and has subsequently been shown to bind other proteins such as thyroglobulin, HA, viral VSV G protein, acetylcholine receptor and MHC I heavy chain (Hurtley *et al.*, 1989; Forsayeth *et al.*, 1992; Kim *et al.*, 1992; Knittler and Haas, 1994, Pind *et al.*, 1994; Hammond and Helenius, 1994b; Hammond and Helenius, 1995; Nobner and Parham, 1995; Williams and Watts, 1995). In yeast, BiP transcription levels are induced in response to environmental stresses which lead to protein misfolding (Mori *et al.*, 1992; Kohno *et al.*, 1993). Under these conditions, BiP

has been shown to form complexes with the aggregation-prone proteins (Marquardt and Helenius, 1992; Simons *et al.*, 1995). BiP's role in translocation, protein folding and quality control suggests that these events are tightly coupled in the ER.

GRP 94 is an ER resident protein which also acts as a chaperone and is involved in immunoglobulin and MHC I assembly (Melnick *et al.*, 1992; Li and Srivastava, 1993; Melnick *et al.*, 1994).

Members of the calnexin/calreticulin family have also been implicated in protein folding and quality control in the ER. Many studies have shown that in mammals, calnexin and calreticulin are present in a transient complex with newly synthesized but not yet folded wild type proteins or not yet assembled subunits of protein complexes (Bergeron *et al.*, 1994; Hammond and Helenius, 1995). Calnexin is also found in stable complexes with mutant misfolded protein and unassembled protein complex subunits (Bergeron *et al.*, 1994; Hammond and Helenius, 1995; Williams and Watts, 1995). Thus, based on these observations a chaperone function has been proposed for calnexin and sequence related calreticulin.

## **1.2.3 ER quality control**

The presence of misfolded proteins due to mutations or environmental stresses is likely to be detrimental for a cell. Cells have mechanisms which ensure that proteins are properly folded and active. Chaperones not only mediate the proper folding of cellular proteins (Gething and Sambrook, 1992), but chaperones also have the ability to recognize misfolded proteins and target them for degradation. This process has been termed 'protein quality control' (Hammond and Helenius, 1995). For secretory proteins, this process is best exemplified in the ER, where secretory proteins are modified and folded prior to export to the Golgi. Misfolded proteins are retained in the ER and subsequently degraded. The site of their degradation is the subject of intense investigation.

Diseases due to ER protein retention of secretory proteins are collectively known as 'protein trafficking diseases' (Amara *et al.*, 1992). The following is a brief description of mutations in three different proteins which lead to 'protein trafficking diseases' (reviewed in Amara *et al.*, 1992; Thomas *et al.*, 1995; Hammond and Helenius, 1995):

Cystic fibrosis is caused by the accumulation of mucus in the lung, most probably due to a defect in the chloride channel encoded by the cystic fibrosis transmembrane regulator (CFTR) (Cheng *et al.*, 1990). The most frequent allele,  $\Delta$ F508, leads to the ER retention of the CFTR protein and its subsequent degradation. This defect leads to diminished amounts of CFTR at the apical membrane, causing decreased chloride conductance.

Alpha<sub>1</sub>-antitrypsin is a serine antiprotease and its predominant physiological role is to protect lung alveoli from proteolytic damage by the protease elastase. Thus reduced serum levels of  $\alpha_1$ -antitrypsin lead to pulmonary emphysema in humans. In the human population, several mutations in  $\alpha_1$ -antitrypsin are found which lead to low secreted levels of  $\alpha_1$ -antitrypsin because of its ER retention and degradation (Cox, 1989). The PI Z variant is a common  $\alpha_1$ -antitrypsin mutant which is largely retained in the ER but can still function as a protease inhibitor. Patients homozygous for the PI Z allele develop emphysema. In some cases, ER accumulation of this variant leads to hepatic cirrhosis (Brantly, 1969). Thus, ER retention of this otherwise functional protein is the cause of two known diseases in humans.

Familial hypercholesterolemia (FH) is another example of a quality control disease. The low density lipoprotein (LDL) receptor regulates serum cholesterol levels, by binding and internalizing LDL. LDL is eventually intracellularly degraded and the LDL receptor recycles to the plasma membrane. Class 2 mutations found in the LDL receptor are responsible for impaired trafficking of the LDL receptor. In this case, LDL receptors are retained in the ER and subsequently degraded, causing elevated serum cholesterol levels. These mutations are found in 50% of naturally occurring FH-associated mutations.

Molecular chaperones, including members of the calnexin/calreticulin family and BiP have been found to associate with defective proteins. Calnexin, calreticulin and BiP are likely to be responsible for the ER retention and the eventual degradation of defective proteins.

# 1.3 Calnexin mode of action

In general, calnexin interacts with proteins that are N-glycosylated. Ou and colleagues first demonstrated in HepG2 cells that glycoproteins  $\alpha_1$ -antitrypsin, transferrin,  $\alpha$ -fetoprotein and apolipoprotein B-100 were found to transiently associate with calnexin. Nonglycosylated serum albumin was not found to be associated with calnexin (Ou *et al.*, 1993). In cells treated with the glycosylation inhibitor tunicamycin,  $\alpha_1$ -antitrypsin, transferrin,  $\alpha$ -fetoprotein and apolipoprotein B-100 failed to associate with calnexin. These observations suggested that calnexin had an affinity for N-linked glycoproteins (Ou *et al.*, 1993).

# **1.3.1 Evidence for calnexin binding to N-linked GlcNAc<sub>2</sub>Man<sub>9</sub>Glc<sub>1</sub> intermediate**

Following the addition of the oligosaccharide intermediate GlcNAc<sub>2</sub>Man<sub>9</sub>Glc<sub>3</sub> to asparagine, the terminal glucose is removed by glucosidase I. The next two glucose residues are then removed by glucosidase II (Chen, *et al.*, 1995; see Figure 1). The removal of a single mannose by an ER-mannosidase is the final modification of oligosaccharides which takes place in the ER. It is possible that calnexin can recognize one or more of these oligosaccharide intermediates.

Subsequent experiments have shown that calnexin recognizes and binds proteins with the GlcNAc<sub>2</sub>Man<sub>9</sub>Glc<sub>1</sub> oligosaccharide intermediate. When cells were treated with the glucosidase inhibitors castanospermine (CST) or 1-deoxynojirimycin (DJM), calnexin association with substrate proteins, including hemagglutinin (HA) and VSV G ts045 was abolished. In cells treated with the  $\alpha$ -mannosidase inhibitor deoxymannojirimycin (DMM), calnexin association with HA and VSV G ts045 was not abrogated. Therefore, calnexin must recognize the GlcNAc<sub>2</sub>Man<sub>9</sub>Glc<sub>0</sub>, the GlcNAc<sub>2</sub>Man<sub>9</sub>Glc<sub>1</sub> or the GlcNAc<sub>2</sub>Man<sub>9</sub>Glc<sub>2</sub> intermediates (Hammond *et al.*, 1994).

Further evidence for the binding of calnexin to the  $GlcNAc_2Man_9Glc_1$  intermediate was obtained using an *in vitro* translation system. Prior to HA translation, microsomes were treated with CST, which prevents formation of the  $Glc_2$  or  $Glc_1$  intermediate or pH 9 buffer, which selectively removes glucosidase II, preventing  $Glc_1$  formation. No calnexin binding was observed in either case. These observations are consistent with a model whereby calnexin. recognizes and binds to proteins with the  $GlcNAc_2Man_9Glc_1$  intermediate. However, these experiments do not rule out calnexin association with the  $GlcNAc_2Man_9Glc_0$  intermediate. Using isolated calnexin-HA immune complexes, the amount of HA complexed to calnexin increased when dissociation of the complex was prevented by the inhibition of glucosidase II. This result was likely due to the reglucosylation of the HA-GlcNAc\_2Man\_9Glc\_0 intermediate by UDP-glucose: glycoprotein glucosyltransferase (UGGT). These experiments lend support to the hypothesis that calnexin binds N-linked GlcNAc\_2Man\_9Glc\_1 (Hebert *et al.*, 1995).

Direct support for binding to the GlcNAc<sub>2</sub>Man<sub>9</sub>Glc<sub>1</sub> oligosaccharide and not to the GlcNAc<sub>2</sub>Man<sub>9</sub>Glc<sub>0</sub> oligosaccharide came from testing the association of calnexin with isolated oligosaccharides *in vitro*. A purified recombinant lumenal calnexin domain bound the GlcNAc<sub>2</sub>Man<sub>9</sub>Glc<sub>1</sub> oligosaccharide but not the GlcNAc<sub>2</sub>Man<sub>9</sub>Glc<sub>0</sub>, GlcNAc<sub>2</sub>Man<sub>9</sub>Glc<sub>2</sub> or GlcNAc<sub>2</sub>Man<sub>9</sub>Glc<sub>3</sub> oligosaccharides (Ware *et al.*, 1995). Calnexin bound poorly to the oligosaccharides GlcNAc<sub>2</sub>Man<sub>5-7</sub>Glc<sub>1</sub>, suggesting that the GlcNAc<sub>2</sub>Man<sub>9</sub>Glc<sub>1</sub> oligosaccharide is likely to be the endogenous substrate for calnexin.

Glucose trimming is also essential for calreticulin binding to nascent glycoproteins. Preincubation with castanospermine abrogates calreticulin association with gp160, HA and transferrin (Peterson *et al.*, 1995; Wada *et al.*, 1995; Otteken and Moss, 1996). It is likely that calreticulin also recognizes and binds proteins with the N-linked GlcNAc<sub>2</sub>Man<sub>9</sub>Glc<sub>1</sub> intermediate. Recently, Spiro *et al.*, have shown that calreticulin can bind sugar oligosaccharides GlcNAc<sub>2</sub>Man<sub>5-9</sub>Glc<sub>1</sub>, but not oligosaccharides which lack the terminal glucose residue (Spiro *et al.*, 1996). Therefore, it is possible that calreticulin and calnexin have similar substrate specificity.

## 1.3.2 Role of UDP-glucose:glycoprotein glucosyltransferase

The ER enzyme UGGT reglucosylates N-linked high mannose sugars [i.e. GlcNAc<sub>2</sub>Man<sub>7</sub>, GlcNAc<sub>2</sub>Man<sub>8</sub>, GlcNAc<sub>2</sub>Man<sub>9</sub>] (Sousa *et al.*, 1992). It has been known for some time that UGGT recognizes unfolded proteins and reglucosylates them, but completely folded glycoproteins are not substrates for UGGT (Sousa and Parodi, 1995; Hebert *et al.*, 1995; see Fig. 1). Therefore, incompletely folded proteins that have been deglucosylated by glucosidase II can be reglucosylated and again bind calnexin or calreticulin (Sousa and Parodi, 1995). UGGT activity is present in almost all eukaryotes including mammals, *Drosophila*, Trypanosomes and the yeast *S.pombe* (Trombetta *et al.*, 1989; Sousa *et al.*, 1992; Fernandez *et al.*, 1994; Labriola *et al.*, 1995; Fernandez *et al.*,

1996). Remarkably, *S.cerevisiae* cell lysates have no UGGT activity (Fernandez *et al.*, 1994). Inspection of the complete *S.cerevisiae* genome has shown that the closest match to UGGT is the ER lumenal protein Kre5p, thought to be involved in cell wall synthesis (Meaden *et al.*, 1990). In *S.pombe*, the UGGT transcript levels are induced by heat shock and other factors which induce protein misfolding (Fernandez *et al.*, 1996). Therefore, UGGT acts as a sensor for unfolded proteins and increases the pool of unfolded proteins capable of binding calnexin or calreticulin.

#### **1.3.3 Requirement for calcium**

Both calnexin and calreticulin have been shown *in vitro* to be calcium binding proteins (Wada *et al.*, 1991; Baksh and Michalak, 1991; Michalak *et al.*, 1992; Tjoelker *et al.*, 1994). Both contain a lumenal high affinity, low capacity calcium binding site and a C-terminal low affinity, high capacity calcium binding site (which is present in the C-terminal cytosolic tail of calnexin). In the presence of calcium, calnexin is more resistant to protease K digestion (Ou *et al.*, 1995) On the other hand, in the presence of the calcium chelator EGTA, calnexin no longer associates with nascent proteins in immunoprecipitates from HepG2 cells (Ou *et al.*, 1993) or  $\alpha_1$ -antitrypsin transfected into hepatoma cells (Le *et al.*, 1994). Thus, calcium appears to be important in preserving calnexin's structure which may affect calnexin-oligosaccharide interactions.

## **1.3.4 Requirement for ATP**

Depletion of cellular ATP also disrupts calnexin interactions with nascent proteins (Ou *et al.*, 1993; Wada *et al.*, 1994). Calnexin and calreticulin are ATP binding proteins, although neither shows any intrinsic ATPase activity (Ou *et al.*, 1995; Nigam *et al.*, 1994). The exact role for ATP binding has not yet been found in calnexin and calreticulin mediated glycoprotein folding. In comparison to other chaperones (HSP 60 and HSP 70), it is probable that ATP is used for a cyclical process preventing inappropriate protein folding (Hendrick and Hartl, 1993). Therefore, hydrolysis of ATP by a cofactor may effect calnexin-oligosaccharide association and/or dissociation.

# Figure 1: Interaction between calnexin/calreticulin and N-linked oligosaccharide intermediates in the ER.

Following the transfer of GlcNAc<sub>2</sub>Man<sub>9</sub>Glc<sub>3</sub> from the dolichol precursor to an asparagine residue on a secretory protein, two glucose residues are removed by the consecutive action of glucosidase I and glucosidase II. Proteins containing Asn-GlcNAc<sub>2</sub>Man<sub>9</sub>Glc<sub>1</sub> can bind to calnexin or calreticulin. Incompletely folded proteins which possess an Asn-GlcNAc<sub>2</sub>Man<sub>7-9</sub> oligosaccharide are substrates for the enzyme UDP-glucose:glycoprotein glucosyltransferase (UGGT). This enzyme adds a single glucose residue to regenerate the Asn-GlcNAc<sub>2</sub>Man<sub>7-9</sub>Glc<sub>1</sub> intermediate which can then reassociate with calnexin or calreticulin. Presumably, protein folding takes place while proteins remain bound to calnexin and calreticulin. Whether calnexin (or calreticulin) itself or other factors mediate folding remains to be tested.



### 1.3.5 Exceptions to the calnexin-oligosaccharide binding model

Some reports have suggested that calnexin can bind to unglycosylated proteins, leading to the hypothesis that calnexin may also recognize protein motifs. For example, removal of the single glycosylation site in MHC I heavy chain did not affect its association with calnexin (Margolese *et al.*, 1993; Carreno *et al.*, 1995). In addition, removal of oligosaccharide by Endo-H treatment of glycosylated MHC I heavy chain molecules bound to calnexin does not interrupt their association with calnexin (Ware *et al.*, 1995). The nonglycosylated T-cell receptor (TCR) subunit, CD3 $\epsilon$ , has also been shown to bind calnexin (Rajagopalan *et al.*, 1994). A similar observation has been made for calnexin binding with nonglycosylated fragments of P-glycoprotein (Loo and Clarke, 1995) and nonglycosylated thyroglobulin (Kim and Arvan, 1995). Recently, Helenius and coworkers have proposed that nonglycosylated proteins associate with calnexin by a mechanism that is different than the mechanism used by glycoproteins (Cannon *et al.*, 1996). They have proposed that removal of the N-linked sugars on normally glycosylated proteins leads to the formation of protein aggregates which trap calnexin. It remains to be determined if this association is specific.

# 1.4 Evidence for calnexin's and calreticulin's role as molecular chaperones

In mammals, calnexin and calreticulin have been assigned the roles of ER molecular chaperones (reviewed in Bergeron *et al.*, 1994; Hammond and Helenius, 1995; Williams and Watts, 1995). Various experiments by several groups have demonstrated the binding of calnexin/calreticulin to incompletely folded proteins and to incompletely assembled protein complexes. Once these proteins are folded and/or assembled, they no longer associate with calnexin or calreticulin (Ou *et al.*, 1993; David *et al.*, 1993; Nauseef *et al.*, 1995). Moreover, mutant proteins or unassembled protein complexes which fail to exit the ER are found in stable complexes with calnexin (Hammond and Helenius, 1995). As discussed below, this points to a chaperone function for calnexin, that is, transient association with folding proteins and assembling complexes as well as stable association with misfolded proteins and unassembled protein subunits in the ER. However, there is no proof that calnexin directly 'senses' the folding state of a protein, as do other chaperones in the HSP 60 and HSP 70 family.

#### 1.4.1 Calnexin association with monomeric wild type proteins

Upon protein translocation into the ER, several post-translational protein modifications take place. The time taken to perform these tasks is a good indication of the ER residency time of a non-ER protein. Calnexin is found associated with incompletely folded glycoproteins in the ER. Ou and colleagues demonstrated that in hepatocytes, incompletely folded  $\alpha_1$ -antitrypsin (see also Le *et al.*, 1994), apolipoprotein B-100, C3 and transferrin are found in association with calnexin with a half life of 5 min., 25 min., 30 min. and 40 min. respectively. Alpha<sub>1</sub>-antitrypsin spends the least time in the ER followed by apolipoprotein B-100, C3 and transferrin (Ou *et al.*, 1993). Although, calnexin association did not account for the total ER residency time for these proteins, association with calnexin may be the rate limiting step in export from the ER. Calnexin is also associated with other monomeric molecules including HA, VSV G protein, HIV gp160, MDCK gp80, CFTR, P-glycoprotein and thyroglobulin (Hammond *et al.*, 1994; Hammond and Helenius, 1994a; Otteken and Moss, 1996; Wada *et al.*, 1994; Pind *et al.*, 1994; Loo and Clarke, 1995; Kim and Arvan, 1995)

### 1.4.2 Calreticulin association with monomeric wild type proteins

Calreticulin has been proposed to have a number of functions in various cellular locations, including the nucleus, cytoplasm and the ER (Michalak, 1996). In addition, a secreted form of calreticulin has been found in various animal species (Jaworski *et al.*, 1995; Michalak, 1996). However, only the role of calreticulin as an ER molecular chaperone will be discussed here.

Calreticulin has been reported to associate with secretory proteins including myeloperoxidase, HA and HIV gp160 in a transient fashion (Otteken and Moss, 1996; Peterson *et al.*, 1995; Nauseef *et al.*, 1995). Like calnexin, calreticulin recognizes incompletely folded but not mature proteins. In the ER, calreticulin is found associated with the immature apopro-myeloperoxidase but not with the heme-containing pro-myeloperoxidase (Nauseef *et al.*, 1995). Calreticulin and calnexin associate with the incompletely disulfide bonded forms of HA rather than the mature forms (Peterson *et al.*, 1995). Thus far, proteins which bind only to calreticulin have not been reported.

It is not clear whether calnexin and calreticulin have different substrate specificities or whether they function in a cooperative fashion in folding glycoprotein. For example, calnexin and calreticulin have been shown to bind different pools of HIV gp120 indicating some differences in specificity (Otteken and Moss, 1996). Moreover, calnexin appears to have a higher affinity than calreticulin for unfolded glycoproteins (e.g. HA, VSV G,  $\alpha_1$ antitrypsin,  $\alpha$ -fetoprotein and HIV gp160), even though calreticulin is more abundant than calnexin (Peterson *et al.*, 1995; Wada *et al.*, 1995; Otteken and Moss, 1996). In all cases, the overall association time of glycoproteins with calnexin or calreticulin is very similar.

Topological differences between calreticulin and calnexin may also reflect different roles for these two chaperones. Most molecular chaperones in the ER are lumenal proteins and there may be a role for calreticulin as opposed to calnexin in ensuring that glycoproteins encounter these other chaperones. Furthermore, the ER lumenal calreticulin may have a role in retrieving incompletely folded glycoproteins which escape the ER. In either model, calreticulin would be important for presenting proteins to other chaperones (BiP, calnexin). In order to disassociate from calreticulin and bind to calnexin or BiP, glycoproteins would have a higher affinity for these chaperones as compared with calreticulin.

# 1.4.3 Calnexin association with assembling protein complexes

In addition to binding incompletely folded glycoproteins, calnexin has also been shown to bind protein complexes assembling in the ER. TCR, mIg, MHC I and MHC II (Hochstenbach *et al.*, 1992; Williams and Watt, 1995), acetylcholine receptor (Gelman *et al.*, 1995) and integrin (Lenter and Vestweber, 1994) are heteroligomers which undergo calnexin-mediated assembly in the ER.

MHC I is composed of a heavy chain, the invariant chain  $\beta_2$  microglobulin ( $\beta_2$ m) and a peptide antigen. Calnexin remains associated with MHC I heavy chain until a heterodimer is formed with  $\beta_2$ m. The heterodimer then binds intimately with the transporter associated with antigen processing (TAP). TAP delivers peptides from the cytosol to the ER. Upon peptide binding, the heavy chain- $\beta_2$ m-peptide trimer dissociates from TAP, exits the ER and subsequently traverses the secretory pathway at a normal rate (Degen and Williams 1991; Degen *et al.*, 1992; Jackson *et al.*, 1994; Ortmann *et al.*, 1994; Suh *et al.*, 1994).

MHC II heteroligomers consist of variant  $\alpha$  and  $\beta$  chains, and invariant Ii chain. Three pairs of  $\alpha\beta$  dimers are first assembled and sequentially added to the Ii trimer to form the ER-export-competent MHC II complex. Calnexin is found associated with  $\alpha$ ,  $\beta$  and Ii monomers, as well as with  $\alpha\beta$  dimers and partially assembled  $\alpha\beta$ Ii complexes. Calnexin dissociates when the final  $\alpha\beta$  dimer is added to form the ER-export-competent MHC II complex (Anderson and Cresswell, 1994; Schreiber *et al.*, 1994).

The T-Cell Receptor is a heteroheptamer and consists of a TCR  $\alpha\beta$  dimer in association with  $\gamma$ ,  $\delta$  and  $\epsilon$  and the  $\zeta$ - $\zeta$  or  $\zeta$ - $\eta$  dimer. Addition of the  $\zeta$ - $\zeta$  or  $\zeta$ - $\eta$  dimer is the last step in TCR assembly. Calnexin remains associated with various subunits and the partially assembled TCR. TCR complexes containing  $\zeta$  subunits are no longer found in association with calnexin (Hochstenbach *et al.*, 1992).

The B-cell receptor (mIg), consists of two heavy chains ( $\mu$ ) which are covalently linked with two light chains ( $\kappa$  or  $\lambda$ ). Only unassembled heavy chains and incompletely assembled heavy-light chain complexes are found associated with calnexin (Hochstenbach *et al.*, 1992).

Integrins are  $\alpha\beta$  heteromers and calnexin is found associated with the  $\alpha_6$  or  $\beta1$  monomers, but not with the assembled  $\alpha_6\beta1$  heterodimers (Lenter and Vestweber, 1994).

The role of calnexin in assembly of complexes is not clear. If calnexin functions as a lectin, it may serve to retain the unassembled subunits until they assemble and presumably attain a conformation which is recognized by UGGT as folded. Thus, before protein subunits are assembled into a mature protein complex, calnexin may be important in preventing the aggregation of unassembled subunits. Once protein subunits are assembled, calnexin binding is no longer required for protein stabilization.

## 1.4.4 The action of BiP and calnexin in the folding of wild type proteins

The division of labor between calnexin and BiP is poorly understood. Both BiP and calnexin associate with folding intermediates of VSV G (Hammond *et al.*, 1994; De Silva *et al.*, 1990; Hammond and Helenius, 1994b) and unassembled human MHC I heavy chains (Nobner and Parham, 1995). On the other hand, only calnexin was found in association with unassembled mouse MHC I heavy chain (Nobner and Parham, 1995) and folding intermediates of  $\alpha_1$ -antitrypsin (Graham *et al.*, 1990, Le *et al.*, 1994). In the case of thyroglobulin, calnexin and BiP are thought to act sequentially. Calnexin binds to newly synthesized thyroglobulin before BiP binding occurs (Kim and Arvan, 1995). Possibly, some proteins expose hydrophobic domains during 'calnexin mediated folding' and thus allow BiP association with these regions.

# 1.5 Stable calnexin association with misfolded proteins

Improper protein trafficking due to retention of misfolded protein in the ER is responsible for several human diseases (Thomas *et al.*, 1995). Calnexin has been implicated in the ER retention of some of these proteins, including mutants of  $\alpha_1$ -antitrypsin and CFTR. Calnexin has also been found in stable complexes with other ER retained misfolded monomeric proteins or unassembled subunits of multimeric proteins, thus supporting a molecular chaperone role for calnexin (Thomas *et al.*, 1995; Hammond and Helenius, 1995; Williams and Watts, 1995).

## 1.5.1 Stable calnexin association with misfolded monomeric proteins

Mutant misfolded proteins, such as the PI Z and Hong Kong  $\alpha_1$ -antitrypsin variants, CFTR  $\Delta$ F508 mutant, mutants of P-glycoprotein and VSV G ts045 are retained in the ER and subsequently degraded (Wu *et al.*, 1994; Le *et al.*, 1994; Pind *et al.*, 1994; Loo and Clarke, 1994; Hammond *et al.*, 1994). Calnexin can be found in stable association with these mutant proteins prior to their degradation. Both calnexin and calreticulin are found in association with an ER retained form of HA that is unable to obtain a mature conformation (Peterson *et al.*, 1995).

# **1.5.2** Calnexin association with the subunits of unassembled protein complexes

Calnexin is also found in association with various unassembled subunits of protein complexes. Some examples are discussed below.

The TCR complex is assembled in the ER. Partial complexes that lack any subunit are retained in the ER and are subsequently degraded (Bonifacino and Lippincott-Schwartz, 1991). Human CD3 $\epsilon$  was retained in the ER when coexpressed with human calnexin in COS cells (Rajagopalan *et al.*, 1994). When C-terminal truncated human calnexin mutants were coexpressed with CD3 $\epsilon$ , CD3 $\epsilon$  was found to be mislocalized with human calnexin in COS cells. Similarly, in a cell line which lacks the TCR  $\alpha$  subunit, calnexin remained associated with the TCR  $\beta$  subunit for a prolonged time. In mammalian cells which did not express  $\beta_2m$ , the MHC I heavy chain was ER retained and remained in association with calnexin (Degen *et al.*, 1992; Rajagopalan and Brenner, 1994). Further evidence for calnexin association and ER retention of unassembled MHC I heavy chain was demonstrated using a heterologous *Drosophila* system. Transfection of mammalian MHC I heavy chain into *Drosophila* cells resulted in the expression of this normally ER retained protein at the cell surface. Coexpression of the heavy chain and  $\beta_2m$  resulted in peptide free heavy chain/ $\beta_2m$  dimers and heavy chain monomers exiting the ER. This aberrant cell surface expression of unassembled MHC I was due to lack of endogenous *Drosophila* cells, it was found to be associated with heavy chain monomers and partially assembled heavy chain/ $\beta_2m$  dimers. Subsequent transport of heavy chain monomers and partially assembled heavy chain/ $\beta_2m$  dimers to the cell surface was retarded in this case (Jackson *et al.*, 1994). This experiment strongly supports the hypothesis that calnexin is able to retain unassembled protein subunits in the ER.

Integrin  $\beta$ -subunits that do not assemble with  $\alpha$  subunits to form the mature  $\alpha\beta$  heterodimeric integrin complex are also found in stable complexes with calnexin in the ER (Lenter and Vestweber, 1994).

Unassembled protein subunits are probably retained in the ER for the same reasons that mutant monomeric proteins are retained. In both cases, proteins expose hydrophobic segments and are thus considered to be misfolded by the ER quality control apparatus.

# 1.5.3 The association of BiP and calnexin with misfolded proteins and unassembled protein subunits

Both BiP and calnexin are found in complexes with misfolded proteins, including CFTR  $\Delta$ F508, VSV G ts045 and HA (Hurtley *et al.*, 1989; Hammond and Helenius, 1994b; Hammond *et al.*, 1994; Pind *et al.*, 1994). For other misfolded proteins, there seems to be a clear preference for association with either calnexin or with BiP. Only BiP is expected to be found in association with unglycosylated proteins. This association is observed when cells were treated with the glycosylation inhibitor tunicamycin (Dorner *et al.*, 1987). However, exceptions to this rule have been found (see section 1.3.5). BiP also associates with glycoproteins and is found associated with ER retained MHC II subunits (Bonnerot *et al.*, 1994; Nijenhuis and Neefjes, 1994). However, ER retained PI Z and Hong Kong  $\alpha$ 1-antitrypsin variants remain associated with calnexin and not BiP (Le *et al.*, 1994, Wu *et al.*, 1994).

The reasons for this specificity are unclear. Reglucosylation of misfolded proteins by UGGT (thus maintaining N-linked oligosaccharides in a GlcNAc<sub>2</sub>Man<sub>9</sub>Glc<sub>1</sub> state), may explain the stable association of misfolded and unassembled proteins with calnexin. MHC II subunits are glycosylated and transiently associate with calnexin. However, when unassembled MHC II subunits are retained in the ER, they are not found in stable complexes with calnexin (Anderson and Cresswell, 1994; Nijenhuis and Neefjes, 1994; Bonnerot *et al.*, 1994). BiP has an affinity for hydrophobic peptides and probably binds to exposed hydrophobic domains in incompletely folded polypeptides (Flynn *et al.*, 1991). Thus BiP binding to MHC II subunits may abrogate calnexin binding.

# **1.5.4** Protein retention in the ER and the role of calnexin in protein degradation

Misfolded protein mutants or unassembled subunits of protein complexes are retained in the ER. In some cases, these mutant proteins are subsequently degraded in a pre-Golgi compartment (Bonifacino and Lippincott-Schwartz, 1993). As described above, calnexin plays a role in the ER retention of mutant or unassembled proteins. Whether calnexin plays a direct role in mutant protein degradation is not well understood.

For some proteins, calnexin has been implicated in their ER retention and stabilization. Expression of MHC I heavy chain was found to be degraded rapidly in

*Drosophila* cells, whereas MHC I heavy chain coexpressed with mammalian calnexin was protected from degradation (Jackson *et al.*, 1994). Furthermore, castanospermine, which abolishes glycoprotein calnexin interactions, enhanced the degradation of the newly synthesized calnexin substrate proteins TCR  $\alpha$  (Kearse *et al.*, 1994) and MHC I heavy chain (Vassilakos *et al.*, 1996) Similarly, nonglycosylated MHC II invariant chain molecules (Romagnoli and Germain, 1995) no longer associated with calnexin and were rapidly degraded.

However, other misfolded proteins which associate with calnexin for prolonged times, such as VSV G ts045,  $\alpha_1$ -antitrypsin PI Z variant and CFTR  $\Delta$ F508 are eventually ER degraded. Calnexin and BiP associate with VSV G ts045 mutant and retain this mutant protein in the ER. Furthermore, BiP selectively retrieves any VSV G ts045 which escapes to the *cis*-Golgi network (Hammond and Helenius, 1994b). This mutant protein is eventually degraded with a half life of four hours (Hammond and Helenius, 1994b). The folding defect of the  $\alpha_1$ -antitrypsin PI Z variant leads to an accumulation of an intermediate which is prone to aggregation and is retained and degraded in the ER (Le *et al.*, 1992; Sifers, 1995). The  $\alpha_1$ -antitrypsin PI Z variant forms stable complexes with calnexin in the ER and is degraded within 30-45 minutes after synthesis (Le *et al.*, 1990). One hundred percent of CFTR  $\Delta$ F508 mutant is retained and degraded in the ER. Prolonged association of the CFTR  $\Delta$ F508 mutant with calnexin is observed. One pathway which is responsible for CFTR  $\Delta$ F508 degradation includes the cytosolic proteosome since proteosomal inhibitors prevent the degradation of CFTR  $\Delta$ F508 (Ward *et al.*, 1995; Jensen *et al.*, 1995).

It is not clear if calnexin promotes or inhibits ER degradation of mutant or unassembled proteins. In the short-term, calnexin may protect proteins which are incompletely folded or incompletely assembled. In the case of protein complexes, calnexin may indirectly stabilize protein subunits by promoting their association into a mature complex (Jackson *et al.*, 1994). However, mutant proteins that can not be properly folded are eventually targeted for proteolytic degradation, as is the case for the PI Z variant and VSV G ts045 (Le *et al.*, 1990; Hammond and Helenius, 1994b) Likewise, protein subunits that do not assemble into mature complexes are eventually degraded, regardless of whether they bind calnexin (Degen *et al.*, 1992). Hence, long-term association of proteins with calnexin may be responsible for their proteolytic degradation. Whether calnexin plays a direct role in 'presenting' proteins to the degradation pathway remains to be tested.

# 1.6 The calnexin/calreticulin family: an ER chaperone superfamily

# 1.6.1 Calnexin

Calnexin and calreticulin are remarkably conserved in eukaryotes. In mammals, calnexin and calreticulin clearly play a role in protein folding and quality control. However, calnexin/calreticulin homologues have also been found in a variety of animal, plant and yeast species. In non-mammalian species, the role of calnexin/calreticulin homologues remains to be tested. There are several features conserved in calnexin and calreticulin homologues and are likely important for the structure-function of calnexin and calreticulin. These features have been dubbed calnexin (calreticulin) signature sequences.

### Calnexin homologues in animal species

Calnexin was serendipitously identified by John Bergeron and colleagues as part of an ER membrane protein complex isolated from canine pancreas (Wada *et al.*, 1991). All calnexins have the overall topology of a type I integral membrane protein. In mammals, calnexin is ubiquitous and its amino acid sequence reveals a 571-573 amino acid protein, in addition to a 20 amino acid cleavable N-terminal signal sequence (Table 2). *Canis familiaris* [canine], *Mus musculus* [mouse], *Rattus rattus* [rat] and *Homo sapiens* [human] calnexins share 90% amino acid identity and have a 94% similarity overall. The exception is a mouse testis-specific calnexin homologue (calnexin-t) which shares 54% amino acid identity with mouse calnexin (Ohsako *et al.*, 1994).

Calnexin homologues have also been cloned from two worm species, the nematode *Caenorhabditis elegans* (Sulston *et al.*, 1992) and the trematode *Schistosoma mansoni* (Hawn *et al.*, 1993). The two worm calnexins share 37% amino acid identity. *C.elegans* calnexin and *S.mansoni* calnexin share 39% and 49% amino acid identity respectively with the mammalian calnexins (Table 1).
# Calnexin homologues in plant species

In plants, calnexin homologues have been cloned from Arabidopsis thaliana (Huang et al., 1993; Boyce et al., 1994) as well as Glycine max [soybean] (Goode et al., 1995) and Helianthus tuberosus [Jerusalem artichoke]. Plant calnexins are shorter than their animal counterparts and have a smaller cytoplasmic tail (46-55 amino acids) (Table 2). Two A.thaliana calnexin homologues have been identified and share the highest degree of identity (81%). Overall, plant calnexins share 68% to 73% amino acid identity. Plant and animal calnexins share 30-39% amino acid identity (Table 1).

# Calnexin homologues in fungi

In yeast, two calnexin homologues have been cloned, one from *Schizosaccharomyces pombe* (Jannatipour *et al.*, 1995; Parlati *et al.*, 1995a) and the other from *Saccharomyces cerevisiae* (De Virgilio *et al.*, 1993; Parlati *et al.*, 1995b). Mature *S.pombe* calnexin is 538 amino acids long, and has a shorter cytosolic tail than animal calnexins (48 amino acids). *S.pombe* calnexin shares between 31% to 38% amino acid identity with plant and animal calnexins (Table 1). On the other hand, *S.cerevisiae* calnexin is only 485 amino acids long and lacks a cytosolic tail. It shares 20% to 28% amino acid identity with other cloned calnexins (Table 1). From the sequence of the complete yeast genome, it is possible to unequivocally state the Cne1p is the only calnexin/calreticulin homologue in *S.cerevisiae*.

	Mammalian	C.elegans	S.mansoni	Plant	S.pombe	S.cerevisiae
Mammalian	-	39%	49%	37-38%	34%	24%
C.elegans	-	•	37%	30-35%	33%	20%
S.mansoni	-	-	-	35-39%	31%	n.d.
Plant	-	-	-	-	36-38%	25-28%
S.pombe	•	-	-	-	-	22%
S.cerevisiae	-		-	-	-	-

Table 1: Percent identity amongst animal, plant and fungi calnexins

Percent identities are calculated using the GeneWorks program. For mammalian and plant calnexins, the percent identities are listed as a range and cover all known mammalian or plant homologues.

#### 1.6.2 Calnexin signature sequences

There are several amino acid motifs present in calnexin homologues and they are summarized below.

# Putative ER lumenal calcium binding repeats

All calnexins encode three copies of the conserved sequence **KPEDWDE** and one copy of the closely related sequence **KPEGWLDD** (or slight variations thereof), for a total of four repeats [in some cases, glutamates (**E**) are exchanged for aspartates (**D**)]. These repeat sequences are much less conserved in *S.cerevisiae* calnexin, Cne1p. Upon closer inspection, it is observed that the spacing betwee the repeats is also conserved. The first and second repeat are separated by 10 amino acids, the second and third repeat are separated by 12 amino acids, and the third and fourth repeat are separated by 11 amino acids. It has been postulated that these repeat motifs form a high affinity low capacity calcium binding site (Baksh and Michalak, 1991; Michalak *et al.*, 1992; Tjoelker, *et al.*, 1994). The conserved spacing between the repeats suggests a conserved structure (Figure 2).

# **Conserved cysteine residues**

All calnexins encode 4 conserved cysteine residues (Figure 2). In mammals, the first and second cysteine are separated by 33 amino acids, and the degree of separation varies from 30 to 34 amino acids in all calnexins. The second and third cysteines are separated by 165 amino acids in mammals, and this varies between 163 and 170 amino acids for all calnexins. Surprisingly, the third and fourth cysteines are always separated by 5 amino acids, except for mouse calnexin-t where there is only a three amino acid separation. In all calnexins cloned thus far amino acids flanking the cysteine residues are well conserved. It has been established that canine calnexin has at least one and possibly two disulfide bonds (Ou *et al.*, 1995).

# C-proximal aspartic/glutamic acid stretch

A third motif consists of several stretches of acidic amino acids present in the cytosolic tail of calnexin, which are often separated by basic amino acids. This region has been identified in calreticulin as a high capacity, low affinity calcium binding site (Michalak

et al., 1992). This domain is clearly present in mammalian calnexin and to a lesser extent in mouse calnexin-t an <sup>•</sup> other animal calnexins. The cytosolic tail encoded by plant calnexins is more basic, and *S.pombe* calnexin does not contain stretches of acidic amino acids in its cytosolic tail region (Figure 3).

#### **C-terminal ER retention/retrieval sequence**

Lumenal ER proteins often have the tetra-peptide KDEL (HDEL, in plants and yeast) at their C-terminus. ER membrane proteins often have the sequence KKXX or variations thereof at their cytoplasmically oriented C-terminus (Nilsson and Warren, 1994). These sequences function to maintain proteins in the ER. In mammalian calnexins, the retrieval sequence has been identified as RKPRRE, present at the C-terminus (Rajagopalan *et al.*, 1994). Deletion of this sequence in human calnexin and expression in COS cells causes it to localize in the Golgi and vesicular compartments. This motif is almost identical in all mammalian calnexins with slight variations in mouse calnexin-t and other animal calnexins (Table 2). In plants, a similar motif rich in basic residues is present at the C-terminus of calnexin (Table 2). The *S.pombe* calnexin C-terminus codes for a different sequence, PTAKNED. However, alignment of the five amino acids at the C-terminus of all calnexins (excluding *S.cerevisiae* calnexin) reveals the consensus sequence (Al)(K/R)( $X_1$ )(D/E)( $X_{0/1}$ ), where Al is an aliphatic amino acid. This is reminiscent of the ER retrieval sequence KDEL, where both acidic, basic and aliphatic amino acids are important for ER protein retrieval (Pelham, 1990).

#### **Casein kinase II phosphorylation sites**

Canine calnexin was originally identified as a phosphoprotein of the ER membrane (Wada *et al.*, 1991). It was latter shown that *in vitro*, casein kinase II could phosphorylate calnexin (Ou *et al.*, 1992). Using calnexin from dog sarcoplasmic reticulum, it was shown that a single serine (either Ser<sup>535</sup> or Ser<sup>545</sup>) was phosphorylated in the cytoplasmic domain (Cala *et al.*, 1993). All calnexins (with the exception of *G.max* and *S.cerevisiae* calnexins), contain potential casein kinase II sites in their cytosolic domains. Thus, phosphorylation may play a role in calnexin regulation.

# Other characteristics of calnexin

Glycosylation sites are generally absent from mammalian calnexins, whereas calnexins in plants, *C.elegans* and *S.pombe* have a single glycosylation site in the lumenal domain. *S.cerevisiae* calnexin, Cnelp, has five N-linked glycosylation sites.

It has been observed that mammalian calnexins migrate on SDS-PAGE with apparent molecular weights between 88-92 kDa, values which are higher than predicted (65 kDa). This retarded migration has been attributed to the low pI values for mammalian calnexin. The calculated pI values for mammalian and *C.elegans* calnexins ranges from 4.17 to 4.2, whereas the calculated pI values for plant calnexins ranges between 4.4 and 4.6. *S.pombe* calnexin has the lowest calculated pI value (4.08) whereas *S.mansoni* and *S.cerevisiae* calnexin have the highest calculated pI values (4.82 and 4.84 respectively). Similar to mammalian calnexin, mouse calnexin-t, *A.thaliana*, *S.pombe* and *S.cerevisiae* calnexins have higher than predicted apparent molecular weights on SDS-PAGE (Huang *et al.*, 1993; Ohsako *et al.*, 1994; Bergeron *et al.*, 1994; Parlati *et al.*, 1995a; Parlati *et al.*, 1995b).

Species	Mature	Cytosolic	C-terminal	Presence of an
	Protein	Length	Sequence	Acidic
	Length			Tail
Animalia				
Mammalia	571-573	89-121	N <sub>/K</sub> R <sup>K</sup> /R <sup>P</sup> /vR <sup>K</sup> /R <sup>E</sup> /D	YES
H.sapiens	572	91	NRKPRRE	YES
C.familiaris	573	91	NRKPRRE	YES
R.rattus	571	89	NRKPRRE	YES
M.musculus	571	89	NRKPRRE	YES
M.musculus-t	592	121	KRRVRKD	YES
Nematoda				
C.elegans	600	120	RTARRGD	YES
Trematoda				
S.mansoni	563	98	KRRSRKE	NO
Plantae	508-523	46-55	R <sup>R</sup> /Q <sup>P</sup> /TRR <sup>E</sup> /D <sup>N/</sup> T/S	NO
H.tuberosus	519	55	RRPRRDT	NO
G.max	523	55	RRPRRET	NO
A.thaliana	511	48	RQPRRDN	NO
A.thaliana	508	46	RQTRRES	NO
Fungi				
S.pombe	538	48	PTAKNED	NO
S.cerevisiae	485	1		

# Table 2: Selected features for animal, plant and fungi calnexins.

The 'mature protein length' indicates the number of amino acids after signal peptide cleavage. The 'cytosolic length' indicates the number of amino acids in the cytosolic tail. The 'C-terminal sequence' is the sequence of the C-terminal seven amino acids, which encodes the putative ER retention sequence. *M.musculus-t* refers to mouse testis specific calnexin or calnexin-t.

Figure 2: Conserved calcium binding and cysteine signature sequences in canine calnexin, mouse calreticulin and *S.pombe* Cnx1p.

Conserved calcium binding repeats are designated 1 to 4, conserved cysteines are designated A, B, C. The signal sequence  $(\square)$ , a highly conserved central domain  $(\square)$ , the transmembrane domain  $(\square)$  and the C-proximal stretch of acidic  $(\square)$  amino acids are shown.



	A		В	C	
Canine calnexin	<sup>161</sup> CGGAYVKLL	161 CGGAYVKLL 188 IMFGE		358 PKCESAPGCG	
Mouse calreticulin	<sup>105</sup> CGGGYVKLF	<sup>105</sup> CGGGYVKLF <sup>130</sup> IMFGPI			
S.pombe Cnx1p	<sup>132</sup> CGGAYLKLL	132 CGGAYLKLL 156 IMFO		324 PLCIEGAGCG	
	1	2	3	4	
Canine calnexin	284 KPEDWDE	<sup>301</sup> KPDDWNE	<sup>320</sup> KPDGWLD	<sup>339</sup> KPEDWDE	
Mouse calreticulin	<sup>215</sup> KPEDWDE	<sup>232</sup> KPEDWDK	<sup>249</sup> KPEDWDE		
S.pombe Cnx1p	<sup>249</sup> KPADWVD	266 KPDDWDE	<sup>285</sup> KPEDWLE	<sup>304</sup> KPEDWDD	

Calreticulin is a major calcium binding ER resident lumenal protein (Smith and Koch, 1989; Michalak *et al.*, 1992) and shares a high degree of amino acid identity with calnexin.

# Calreticulin homologues in animal species

Calreticulin has been cloned from a variety of mammals including *H.sapiens* [human], *Oryctolagus cuniculus* [rabbit], *M.musculus* [mouse] and *R.rattus* [rat] (Michalak *et al.*, 1992). Mammalian calreticulins are highly conserved and share 90% amino acid identity and a 94% overall amino acid similarity. No tissue specificity has been observed for calreticulin other than for *Bos taurus* [bovine], which encodes a brain specific isoform and shares a 70% amino acid identity (75% total similarity) with other ubiquitous mammalian calreticulins (Liu *et al.*, 1993).

In addition, two Xenopus laevis brain variants have also been found, which share over 93% amino acid identity (Treves et al., 1992). X.laevis calreticulins share 78% identity with mammalian calreticulin but only 62% identity with the bovine brain variant. Calreticulins from Drosophila melanogaster [fruit fly] (Smith, 1992a), Amblyomma americanum [tick] (Jaworski et al., 1995), as well as from Aplysia californica [marine snail] (Kennedy et al., 1992) have been identified. Calreticulin has been also identified in two nematodes, C.elegans and Onchocerca volvulus (Unnasch et al., 1988; Smith, 1992b) and a trematode, Schistosoma japonicum (Huggins et al., 1995). The percent identities between calreticulins from different species are summarized in Table 3. Except for S.japonicum they all share 59% or greater amino acid identity, indicating that the calreticulin family is highly conserved (Table 3).

# Calreticulin homologues in plant species

In plants, calreticulin homologues have been found in at least five angiosperm species, including *Hordeum vulgare* [barley] (Chen *et al.*, 1994), *Nicotiana tabacum* [tobacco] (Denecke *et al.*, 1995), *A.thaliana* (Benedetti and Turner, 1995), *Capsicum annum* [pepper] (Hugueney *et al.*, 1995) and *Zea mays* [maize] (Kwiatkowski *et al.*, 1995). They share 75% or greater amino acid identity and show 45-50% identity with animal calreticulins (Table 3).

	Mammalian	X.laevis	D.melanogaster	C.elegans	O.volvulus	Plant
Mammalian	-	78%	64%	62%	59%	48-49%
X.laevis	-	-	62%	62%	59%	49-50%
D.melanogaster	-	-	-	58%	58%	48%
C.elegans	•	-	-	-	65%	48%
O.volvulus	-	-	-	-	-	45-47%
Plant	-	-	•	-	-	-

# TABLE 3: Percentage identity amongst animal and plant calreticulins.

Percent identities are calculated using the GeneWorks program. For mammalian and plant calnexins, the percent identities listed covers all sequenced mammalian and plant calreticulins.

# 1.6.4 Calreticulin signature sequences

Some amino acid motifs present in calnexins are also present in calreticulins and are discussed below.

#### **Calcium binding repeats**

There are two conserved **KPEDWDE** repeats as well as a closely related **KPEDWDK** (in animals) or **KPEGYDD** in plants (n. b., sometimes **E** and **D** are substituted). Hence, three repeated motifs are present in calreticulin, compared to four found in calnexins. The amino acid spacing between these motifs is also conserved as is the case for calnexin. In all calreticulins, there are 10 amino acids between the first two repeats and 10 amino acids between the second and third repeat (11 amino acids in plant) (Figure 2). As observed for calnexin, the conserved spacing between these motifs suggests a conserved structure which may be possibly important for calcium binding (Michalak *et al.*, 1992).

# **Conserved cysteine residues**

Two cysteine residues are conserved in all calreticulins and correspond to the first two conserved cysteine residues in calnexin (Figure 2). The amino acid spacing between these cysteine residues is also conserved (31 residues). Studies with purified bovine calreticulin have shown that they form a disulfide bond (Matsuoka *et al.*, 1994).

# C-Proximal aspartic/glutamic acid stretch

Acidic stretches of amino acids are also found near the C-terminus of some calreticulins. Mammalian, X.laevis, D.melanogaster, A.americanum, A.californica, C.elegans, and plant calreticulins show significant stretches of acidic amino acids near the C-terminus (Figure 3). On the other hand, S.japonicum does not have an aspartic/glutamic acid rich domain and the C-terminal domain of O.volvulus has stretches of basic amino acids (Unnasch et al., 1988).

#### ER retrieval sequences

ER resident lumenal proteins encode the C-terminal KDEL sequence or slight variations thereof. This motif acts as a signal to recycle ER proteins, via a receptor mediated mechanism, which may escape the ER lumen (Pelham, 1990). All calreticulins have C-terminal ER retrieval sequences which are summarized in Table 4. The only exception is *O.volvulus* calreticulin, which has the sequence KKKK at its C-terminus.

# Other characteristics of calreticulin

Consensus sequences for N-linked glycosylation sites are present in all known plant calreticulins (except *A.thaliana*). One N-linked glycosylation consensus sequence is present in mammalian and frog calreticulin, and glycosylated forms of calreticulin have been reported in mammals (Van *et al.*, 1989; Peter *et al.*, 1989; Matsuoka *et al.*, 1994).

Like calnexin, calreticulin migrates with a higher than expected molecular weight on SDS-PAGE. The predicted molecular weight of mammalian calreticulin is approximately 46 kDa, whereas it migrates with an apparent molecular weight of 55 kDa (Michalak *et al.*, 1992). Here again, a low pI value is likely to be responsible for this aberrant SDS-PAGE migration. The calculated pI values range from 4.0 to 4.07 in mammalian, *A.californica* and *D.melanogaster* calreticulin. The calculated pI values range from 4.21 to 4.32 for *A.americanum*, *C.elegans* and *S.japonicum* calreticulins. For plant calreticulin, the pI values range from 4.05 to 4.14. Surprisingly, the calculated pI value for *O.volvulus* calreticulin is 6.16.

Species	Mature	C-terminal	Presence of an
	Protein	Sequence	Acidic Tail
	Length		
Animalia			
Mammalia	399-403	KDEL	YES
H.sapiens	400	KDEL	YES
O.cuniculus	401	KDEL	YES
R.rattus	399	KDEL	YES
M.musculus	399	KDEL	YES
B.taurus Brain	403	KDEL	YES
Amphibia			
X.laevis	399	KDEL	YES
Arachnida			
A.americanum	INCOMPLETE	HEEL	YES
Insecta			
D.melanogaster	387	HDEL	YES
Gastropoda			
A. californica	390	KDEL	YES
Nematoda			
C.elegans	371	HDEL	YES
O.volvulus	380	KKKK	NO
Trematoda			
S.japonicum	INCOMPLETE	HDEL	NO
PLANTAE			
H.vulgare	394	HDEL	YES
Z.mays	396	HDEL	YES
N.tabacum	388	HDEL	YES
A.thaliana	405	HDEL	YES
C.annum	390	HDEL	YES

# TABLE 4: Selected features for animal and plant calreticulins.

The 'mature protein length' indicates number of amino acids after signal peptide cleavage. The 'C-terminal sequence' refers to the four C-terminal amino acids, which encodes the putative ER retrieval sequence.

# 1.6.5 Calnexin and calreticulin compared

In humans, mouse, rat, *C.elegans* and *A.thaliana*, complete cDNAs for calnexin and calreticulin have been identified. Calnexin and calreticulin proteins from the same species show 28% amino acid identity, except in *C.elegans*, where they share 24% amino acid identity. The overall amino acid identity, conserved motifs, in addition to ER localization signals argues that calnexin and calreticulin form a superfamily of proteins with possibly similar functions. Figure 3 depicts conserved domains amongst members of the calnexin and calreticulin family.

# Figure 3: Comparison of the conserved regions in calnexin and calreticulin homologues.

The signal sequence ( $\blacksquare$ ), a highly conserved central domain ( $\boxtimes$ ), the transmembrane domain ( $\boxtimes$ ) and the C-proximal stretch of acidic ( $\boxtimes$ ) & basic ( $\boxtimes$ ) amino acids are shown. The highly conserved central domain encodes the putative calcium motifs and conserved cysteine residues. The amino acid length of calnexin and calreticulin (including the signal sequence) is noted. For calreticulin, the C-terminal amino acid sequence is also noted.



# 1.7 Aim of this work

The ER is the site of folding and modification of many organellar and secretory proteins. In addition, there are mechanisms in the ER for the surveillance and retention of misfolded proteins or unassembled complexes, known as the ER protein quality control apparatus. The ER resident proteins BiP and calnexin have been implicated in the folding and quality control of newly synthesized and translocated proteins. Calnexin has a role in the folding of monomeric proteins, the assembly of protein complexes, as well as the retention of misfolded or unassembled proteins in the ER. Recognition of the N-linked sugar intermediate GlcNAc<sub>2</sub>Man<sub>9</sub>Glc<sub>1</sub> is the principal mechanism by which calnexin and calreticulin bind transiently to proteins. The implications of this binding are not completely understood. Misfolded or unassembled proteins remain in stable association with calnexin presumably because the N-linked sugar moiety GlcNAc<sub>2</sub>Man<sub>9</sub>Glc<sub>1</sub> is not trimmed. The long-term consequence of a stable association with calnexin is likely to be proteolytic degradation.

Proteins that carry out basic cellular functions are often present in most, if not all, eukaryotic species. As described above, calnexin homologues are present in a wide variety of eukaryotic species including yeasts. Yeasts are a useful experimental system for determining the effects of deleting or overexpressing a gene. One can also assign functional significance to various protein domains by characterizing the phenotype upon expression of mutant forms of the protein.

The aim of this work was to identify and characterize the calnexin and/or calreticulin homologue in the yeast species *S.cerevisiae* and *S.pombe*. Once identified, the biochemical properties of the genes and gene products of the calnexin/calreticulin yeast homologues were characterized. These properties included calcium binding, the glycosylation state, intracellular localization and membrane affinity of the gene products and stress induciblity of the gene transcript. The potential role of yeast calnexin and/or calreticulin in ER quality control was assessed using gene disruption and the expression of various calnexin/calreticulin domains.

# **CHAPTER 2**

Saccharomyces cerevisiae CNE1 encodes an ER membrane protein with sequence similarity to calnexin and calreticulin and functions as a constituent of the ER quality control apparatus

# 2.1 Abstract

We have used a PCR strategy to identify in the yeast Saccharomyces cerevisiae, genes of the mammalian calnexin/calreticulin family, and we have identified and isolated a single gene, CNE1. The protein predicted from the CNE1 DNA sequence shares some motifs with calnexin and calreticulin, and it is 24% identical and 31% similar at the amino acid level with mammalian calnexin. On the basis of its solubility in detergents and its lack of extraction from membranes by 2.5 M urea, high salt and sodium carbonate at pH 11.5, we have established that Cnelp is an integral membrane protein. However unlike calnexins, the predicted carboxy-terminal membrane-spanning domain of Cnelp terminates directly. Furthermore, based on its changed mobility from 76 kDa to 60 kDa after endoglycosidase H digestion, Cnelp was shown to be N-glycosylated. Localization of the Cnelp protein by differential and analytical sub-cellular fractionation as well as by confocal immunofluorescence microscopy showed that it was exclusively located in the ER, despite the lack of known ER retention motifs. Although six Ca<sup>2+</sup> binding proteins were detected in the ER fractions, they were all soluble proteins and Ca<sup>2+</sup> binding activity has not been detected for Cnelp. Disruption of the CNE1 gene did not lead to inviable cells or to gross effects on the levels of secreted proteins such as  $\alpha$ -pheromone or acid phosphatase. However, in *CNE1* disrupted cells, there was an increase of cell-surface expression of an ER retained temperature sensitive mutant of the  $\alpha$ -pheromone receptor, Ste2-3p, and also an increase in the secretion of heterologously expressed mammalian  $\alpha_1$ -antitrypsin. Hence, Cnelp appears to function as a constituent of the S.cerevisiae ER protein quality control apparatus.

# 2.2 Introduction

Calnexin is an integral membrane calcium binding phosphoprotein found in the ER of mammalian cells (Wada *et al.*, 1991; Bergeron *et al.*, 1994). Closely related DNA sequences have been found in plants and nematodes (Sulston *et al.*, 1992; Huang *et al.*, 1993). A function for calnexin as a molecular chaperone has been identified (Bergeron *et al.*, 1994). It associates transiently with several membrane glycoproteins during their maturation in the endoplasmic reticulum including MHC I heavy chain (Degen *et al.*, 1991; Ahluwalia *et al.*, 1992; Degen *et al.*, 1992; Galvin *et al.*, 1992), MHC II (Anderson *et al.*, 1994), the T cell receptor, membrane Ig (Hochstenbach *et al.*, 1992), the viral membrane glycoproteins influenza HA and the "G" protein of vesicular stomatitis virus (Hammond *et al.*, 1994) as well as the cystic fibrosis transmembrane conductance regulator, CFTR (Pind *et al.*, 1994) and integrin (Lenter and Vestweber, 1994). In addition, calnexin associates transferrin,  $\alpha_1$ -antitrypsin, complement C3, apoB-100 (Ou *et al.*, 1993; Le *et al.*, 1994), as well as the major secreted glycoprotein of MDCK cells, gp80 (Wada *et al.*, 1994).

A second related function has been proposed for mammalian calnexin as a constituent of a protein quality control apparatus in the ER recognizing and retaining some mutant proteins and components of unassembled complexes. For example, when soluble secretory glycoproteins are synthesized in the presence of the proline analog azetidine 2-carboxylic acid, they are retained in the ER and remain associated with calnexin for a prolonged period (Ou *et al.*, 1993). Similarly, mutant proteins such as VSV G ts 045 glycoprotein are retained in the ER by their association with calnexin (Hammond *et al.*, 1994). Components of unassembled complexes are also retained in the ER in association with calnexin, for example, the MHC class I heavy chain synthesized in the absence of  $\beta_2$ -microglobulin (Degen *et al.*, 1991; Ahluwalia *et al.*, 1992, Degen *et al.*, 1992; Jackson *et al.*, 1994) and the T cell receptor synthesized in the absence of the  $\alpha$ -chain (David *et al.*, 1993; Rajagopalan *et al.*, 1994). Thus calnexin has the properties expected of a component of such a quality control mechanism.

Recently, the sequence of a gene in *S.cerevisiae* with similarity to mammalian calnexin has been reported (De Virgilio *et al.*, 1993). The isolation of a calnexin homologue from yeast will help elucidate the molecular mechanisms whereby calnexin carries out its roles as a molecular chaperone and the retention of proteins in the ER membrane. We have identified by a PCR strategy, a candidate calnexin gene in *S.cerevisiae CNE1*. We have

characterized and localized the Cne1p protein and have determined by gene disruption some of its functions.

# 2.3 Materials and Methods

# Strains and media

The S.cerevisiae diploid strain W303D (MAT  $a/\alpha$  ade2-1/ade2-1 can1-100/can1-100 ura3-1/ura3-1 leu2-3,112/leu2-3,112 trp1-1/trp1-1 his3-11,15/his 3-11,15), W303-1a (MAT a ade2-1 can1-100 ura3-1 leu2-3,112 trp1-1 his3-11,15), W303-1b (MAT  $\alpha$  ade2-1 can1-100 ura3-1 leu2-3,112 trp1-1 his3-11,15), DC 17 $\alpha$  (MAT  $\alpha$  his1) and M200-6C (MAT a sst1 sst2) strains were grown at 30°C in YPD medium containing 1% yeast extract (Difco), 2% Bacto-peptone (Difco), and 2% dextrose (BDH) or synthetic media (SC) with the appropriate amino acid supplements and either 2% glucose or 2% sucrose. The E.coli strain MC1061 was used (Maniatis et al., 1982). Yeast synthetic media was previously described (Sherman et al., 1979).

# **PCR** amplification

To identify and isolate genes similar to calnexin from *S.cerevisiae* genomic DNA, degenerate oligonucleotides for the sequences KPEDWDE and YKG<sup>K</sup>/EWKP with all possible codons at each position were synthesized using a BioSearch series 8000 DNA synthesizer (see Fig 4B). Amplification was performed using a Perkin-Elmer Cetus Thermocycler (Hann *et al.*, 1989). Samples were then electrophoresed on a 2% agarose gel and visualized by ethidium bromide staining. The band migrating at approximately 300 bp was purified by electroelution and cloned into the *Sma* 1 site of plasmid pTZ19R and sequenced by the dideoxy protocol using T7 DNA Polymerase (Pharmacia).

### Cloning of calnexin in S.cerevisiae

In order to clone the entire sequence of gene we identified, YEp24 genomic *S.cerevisiae* DNA libraries were screened using the isolated calnexin PCR fragment as a probe labeled by nick translation (Maniatis *et al.*, 1982). Two independent clones were isolated and mapped using restriction enzyme analysis. By Southern analysis, a 3.8 kb *Sph* 1 fragment was found to hybridize to the PCR probe and was subcloned into the *Sph* 1 site of pTZ19R (Pharmacia) to generate plasmid pFP10.1. Based on sequence information provided by the PCR fragment, oligonucleotides were synthesized and used to sequence the gene as previously described (Maniatis *et al.*, 1982). The entire *CNE1* gene was sequenced and the *CNE1* sequence was released to GenBank (accession number: L11012).

# Antibody production

Polyclonal antibodies recognizing calnexin were obtained by immunizing rabbits with GST::Cne1p fusion proteins expressed in *E.coli*. The fusion was made by inserting a *Bam* H1-Sph 1 fragment (*CNE1*) into pGex-2T (Smith and Johnson, 1988). GST-calnexin was expressed by IPTG induction and purified (Smith and Johnson, 1988).

# Membrane extraction and Endo-H digestion

Extracts of post-nuclear supernatants were mixed with 1 volume of 1 M NaCl, 0.2 M sodium carbonate pH 11.5, 2.5 M urea, 2% Triton X-100, 0.2% Triton X-100, 2% deoxycholate or 0.2% SDS and were subsequently analyzed as previously described (Feldheim *et al.*, 1992). Cne1p antiserum was used at 1:2000 dilution. Endo-H digestions were performed by incubating 50  $\mu$ g of ML fraction proteins in 100 mM sodium acetate pH 4.9, 150 mM NaCl, 10 mM DTT, 1% Triton X-100 + inhibitors (1 mM PMSF, 1  $\mu$ g/ml pepstatin, 1  $\mu$ g/ml leupeptin and 1  $\mu$ g/ml aprotinin) and incubating with 2  $\mu$ g of Endo-H for 16 hours at 37°C.

# Yeast fractionation

S.cerevisiae strain W303-1a was grown at 30°C in YPD medium to a density 2-4 OD<sub>600</sub>/ml, cells were harvested by centrifugation and washed in water. Spheroplasts (100 OD<sub>600</sub>/ml) were generated by a 60 minute incubation at 30°C in 0.7 M sorbitol, 1.5% peptone, 0.75% yeast extract, 0.5% glucose, 10 mM Tris, 1 mM DTT and Zymolyase T100 1 mg/g wet weight yeast, and homogenized with a Potter-Elvejhem homogenizer in 0.1 M sorbitol, 20 mM Hepes, 50 mM potassium acetate pH 7.4, 1 mM PMSF, 5  $\mu$ g/ml aprotinin. The homogenate was then subjected to differential centrifugation at 4°C. Three different fractions: i.e. nuclear (N), large granule (ML) and microsomal (P) and a final supernatant (S) were separated by successive centrifugation at a square angle velocity of 8.2 x 10<sup>5</sup>, 1.8 x 10<sup>9</sup> and 1.2 x 10<sup>11</sup> rad<sup>2</sup>s<sup>-1</sup>. For isopycnic sucrose gradient centrifugation, the large granule (ML) fraction was loaded on a sucrose density gradient (0.5 to 2.3 M sucrose, 20 mM Hepes, pH 7.4) and centrifuged for 8h at 7.6 x 10<sup>10</sup> rad<sup>2</sup>s<sup>-1</sup>. (SW40 Beckman Instruments), fractions were collected and analyzed for activity of the marker enzymes, ATPase (Bowman *et al.*, 1979), NADPH cytochrome c reductase (Kubota *et al.*, 1977), GDPase (Abeijon *et al.*, 1989) and monoamine oxidase (Bandlow, 1972). Kar2p and Cne1p were detected by

immunoblot and subsequently quantitated by densitometry. Anti-Kar2p and anti-Cne1p antisera were used at 1:2000 and 1:1000 dilution respectively.

# 45Ca overlay

Samples were electrophoresed by SDS-PAGE and evaluated for <sup>45</sup>Ca overlay exactly as previously described (Wada *et al.*, 1991). Yeast ER fractions were recovered from analytical isopycnic gradients of ML fractions at densities greater than 1.151g/ml ( $\rho$ > 1.151 g/ml). Membrane and soluble proteins were separated by Triton X-114 extraction as described by Bordier (Bordier, 1981). Control experiments were carried out with dog pancreatic ER membranes also extracted with Triton X-114 exactly as previously described (V/ada *et al.*, 1991).

# Immunofluorescence

Staining of *S.cerevisiae* was performed essentially as previously described (Pringle *et al.*, 1991) with the following incubations: 1) anti-Cne1p antisera (1:1000) for 60 min., 2) rhodamine conjugated Fab (1:50 Jackson Immunochemicals) for 45 min., 3) anti-Kar2p antisera (1:2000) for 60 min., 4) FITC conjugated IgG (1:50, Jackson Immunochemicals) and DAPI (2 mg/ml, Sigma) for 45 min.. Cells were viewed using epifluorescence (Aristoplan, Leitz) and by confocal microscopy (Molecular Dynamics).

# Disruption of the yeast CNE1 gene

Plasmid pFP10.1 was digested with Sph 1 and religated to clone the insert in the opposite orientation, creating plasmid pFP10.11. A 1 kb Eco R1 fragment was removed from pFP-10.11. This effectively removes the multiple cloning site to create the plasmid pFP-10.12. A 750 bp Bam H1-Pst 1 internal to CNE1 was replaced with a 2 kb Bam H1-Pst 1 fragment containing the LEU2 gene from pJJ250 (Jones and Prakash, 1990). The resulting plasmid, pFP10.13, was cut with Sca 1 and Sph 1 to linearize the plasmid and transformed into the leu2<sup>-</sup> diploid yeast strain W303D (Ito et al., 1983). Transformants were selected on SC glucose minus leucine plates. Disruption of the CNE1 gene was confirmed by Southern blots. For further genetic analysis, diploids were sporulated and tetrad dissection was performed by standard procedures and the presence of the disruption in parent cells and spores confirmed by Southern blot analysis.

#### Acid phosphatase assay

Plasmid pRS306-*CNE1* was constructed by inserting the *Sca* 1- *Hpa* 1 fragment containing the open reading frame of *CNE1* into the *Bam* H1 site (3' to the Gal promoter) of vector pRS306 Gal (Sikorski and Hieter, 1989) and transformed into strain W303-1b  $\Delta cne1$ ::*LEU2*. This strain was subsequently grown at 30°C in SC sucrose-uracil (Sherman *et al.*, 1979) to OD<sub>600</sub> of 1. Cultures were then divided in three and glucose, galactose (2% w/v final concentration) or sucrose (4% w/v final concentration) added. Cell surface acid phosphatase activity was determined as described (Tohe-e *et al.*, 1973).

# Heterologous expression of a<sub>1</sub>-antitrypsin

The cDNAs for wild type and the PI Z  $\alpha_1$ -antitrypsin variant were cloned into the *Pvu* II site of pVT-101U (Vernet *et al.*, 1987). Plasmids pVT-AlPi (wild type) and pVT-AlPz (mutant) were transformed into W303-1b and W303-1b  $\Delta cne1::LEU2$ , these were grown overnight in SC glucose -uracil at 30°C. Equal numbers of cells were spotted onto SC glucose-uracil plates and overlaid with a nitrocellulose filter (BA85 Schleicher and Schuell). Plates were incubated at 30°C overnight and nitrocellulose was subsequently washed to remove yeast cells and immunoblotted with  $\alpha_1$ -antitrypsin antiserum at 1:1000 dilution (Calbiochem). For detection, either a secondary antibody linked to alkaline phosphatase or Protein A linked to <sup>125</sup>I was used.

#### Halo assay for $\alpha$ -pheromone production

20 ml cultures of wild type strain (W303-1b pVT), calnexin deleted strain (W303-1b  $\triangle cne1::LEU2$  pVT) or calnexin overproducing strain (W303-1b  $\triangle cne1::LEU2$  pVT-*CNE1*) were grown in SC glucose -uracil to OD<sub>600</sub> of 1 and centrifuged at 1000 x g for 5 minutes. Cells were resuspended in 250 µl of water, and 5 µl was spotted onto a lawn of M200-6C cells on YPD agar. Agar plates were incubated at 30°C for 48 hours.

# Quantitative mating assay

Assays were performed as described (Sprague, 1991). A 3.0 kb *Eco* R1-*Sph* 1 fragment containing calnexin was cloned into the *Eco* R1 site of pAD13, a low copy number plasmid (Costigan *et al.*, 1992). Mating ficiency for strains DJ 283-7-1a (*Mat* **a** *ste2-3ts can1ts bar1-1 ade2<sub>oc</sub> his4<sub>am</sub> lys2<sub>oc</sub> le*  $\pm$  *trp1<sub>am</sub> ura3 cry1 SUP4-3<sub>am</sub>ts*  $\triangle$ *cne1::LEU2*)

transformed with pAD13 or pAD13-*CNE1* were measured at 23°C and 37°C using tester strain DC17 $\alpha$ . Strains were grown to an OD<sub>600</sub> of 1 at either 23°C or 37°C and then mixed with confluent DC 17 $\alpha$  for three hours at 23°C of 37°C. Mating efficiency is defined as the number of diploids formed per input haploid. Relative mating efficiency was standardized for each experiment. The mating efficiency of DJ 283-7-1a pAD13, with DC 17 $\alpha$  at 23°C was set at 100.

# 2.4 Results

We used a specific PCR approach to clone genes with sequence similarity to mammalian calnexin and calreticulin from *S.cerevisiae*. Degenerate oligonucleotide primers were designed which corresponded to the amino acid sequence motifs shared between mammalian calnexin and calreticulin (Fig. 4A, 4B). Using S.cerevisiae DNA as a template, an amplified DNA fragment of approximately 300 bp was identified, cloned and sequenced. The sequence corresponded most closely to that of mammalian calnexin (nucleotides 1073-1424) (38%) and of mammalian calreticulin (24%). This DNA fragment was then used to probe a *S.cerevisiae* genomic library in the yeast vector YEp24. Two independent clones with an overlapping common region were isolated from 4 x 10<sup>4</sup> colonies screened. The yeast DNA insert was subcloned on the basis of its hybridization with the DNA probe and its nucleotide sequence was determined. The DNA sequence predicts a protein of 502 amino acids and was found to be identical to a previously reported gene sequence, CNE1 (De Virgilio et al., 1993). The overall sequence identity to canine calnexin was 24% and mouse calreticulin was 21% (Fig. 5). The predicted protein (Fig. 6) contains a signal sequence, Nlinked glycosylation sites and a carboxy terminal transmembrane domain (Kyte and Doolittle, 1982; von Heijne, 1986; Argos et al., 1982). Unlike calnexin, the predicted Cne1p sequence did not contain a carboxy terminal cytosolic domain, and unlike calreticulin it did not have a carboxy terminal ER retention motif (HDEL in yeast) (Pelham, 1990). Thus, on the basis of overall predicted structure, the sequence we identified did not closely resemble either known calreticulin or known calnexin sequences.

# Figure 4: The CNE1 PCR cloning strategy.

Panel A, the central domains of calnexin and calreticulin are aligned between amino acids 254 to 389 and 185 to 281 respectively. Amino acid sequences used to design sense and antisense oligonucleotides are indicated in bold type. Panel B, amino acid sequences and corresponding nucleotide sequences (PCR S, sense; PCR A, antisense) used as primers for PCR amplification of yeast (Y=C or T; R=A or G; N=A, C, G, T).

CALNEXIN	254 SREIEDPEDQ <b>KPEDWDE</b> RPKIPDPDAV	VKPDDWNEDAPAKIPDI	SEATKPDGWLDDEPEYVPDPDAEKP
CALRETICULIN	FLPPKKIKDPDAAKPEDWDERAKIDDPTDS	SKPEDWDK	

			3
CALNEXIN	EDWDEDMDGEWEAPQIANPKCES	APGCGVWQRPMIDNPNYKGKWKPPMIDNPNYQGIWKPR	KIPNPDFF
CALRETICULIN	EDWDEEMDGEWEPP	VIQNPEYKGEWKPRQIDNPDYKGTWIHP	EIDNPE
	247	248	281

в

PCR S	PCR A
K P E D W D E	YKG <sup>K</sup> WKP E
5' AARCCNGARGAYTGGGAYGA 3'	3' ATRTTYCCNYTYACCTTYGG 5'

Α

Figure 5: Amino acid alignments of *S.cerevisiae CNE1*, canine calnexin and mouse calreticulin.

Amino acids conserved in at least two sequences are shaded. The alignment was performed using the GeneWorks program.

Yeast CNE1	KERSAYE-WARFLINL LUKKG-TSLLSNVTL AEDSFWEHFQAYTNT KHLINDEWI-TBEADN NEGSKIYGAQWREGOGR	-4
Drg calnexin	EGKWLECHLEVLGT TIVQAHEGHDDDMID IEDDLDDVIEEVEDS KSKPETSAPTEPKET YKAFVPSGEVYFADS FDRGTLSGWILEKAK	<u>ð</u> 0
Mouse calreticulin	LLSVPLLGLGLGLA ADPAIYFKEQFLDGRAWGRAW	55
Consensus	₩S.L.L. X.L	90
Yeast CNE1	LQGSAW DWGIAVRTGNA AAMIGHLLENP INVSETDTERMONYEIKLDHSLTCOGAF IKUMSGFMARHALKH	144
Dog calnexin	KDDTDDEIANYDGKW WVDENEETKLPGDKG LVLMSRAKHHAISAK LUKPFLFDTKPLFDTKPLFDGIECOGAY WALSKTPELH-LOO	179
Mouse calreticulin	FYGDL	121
Consensus	GWRL.T AHALSAKPFT.T.T.T.Y.Y.Y.K.M.I.COGAY VKC.SLLK.	180
Yeast CNE1	YAPDTEGVELVITTE YEAPEIN-GVQFAIN KVDKITHES LRYLQ MPLSKLTDTSQUELYTTY IDESADEFDIL IDGKYVMVREHIEDK	233
Dog calnexin	F-HDKTPYTIME SEDYKLHFIFRHK NPKTGVYEE HAKRP ADLKTYFTDKK AT AL-PDEFET UV-CHIVNSGNLL	265
Mouse calreticulin	M-HGDSEYNIM CHIIGPGTK-KVHVIFNYKGINV-LI XDIR-CKDDEF	198
Consensus	-HDY.IMPORD. GGP.KV.F.NYELDDDDDDPD. FELL 10S.V.SG.LED-	270
Yeast CNE1	KKVNFEPPITNPLMARPNVSVANNH HENRI REPEREAVKLSDRDE RDPLMIPHPDGRENP ENNSIGNELDER SVRKELEHANNI	323
Dog calnexin	MTPPVN SRE 12 PEDQ E COMP P K P DAVKPDOWNE DAPAKIPDEEAKKO GILDDE YYPP CERED ODDODIE	352
Mouse calreticulin		262
Consensus		360
Yeast CNE1	PEMIKNPLCTAERGC GQQIPGLE AK BER PGELNE I HER ME HPPEZER LYYER HPLRIENVIE GVILLAR SSSPNMLISHIY	413
Dog calnexin	A QIANPKCESAPGC GVWQRPM PN WWKPPM D V QII KPRK P DFFEDL EPFKMTPF-AAIGLE LISMTSDIFFD FIV	441
Mouse calreticulin	PPVIEPE EWKPRO DE WKPRO DE WKRON DE W	332
Consensus	РӇ.І.NP.СGC G Д. Р	450
Yeast CNE1	GKUVTEANIIG KTULMRDRAFRGSDGPTE RKFMNSRLGNLQTTF HNERESPNPFDRIID RILEQPLKFVLTAAV VLLTTSVLCCVVFTX	503
Dog calnexin	CCERRVVDDWAIDG /GLKK	509
Mouse calreticulin		دود
Consensus		540
Yeast CNE1		503
Dog calnexin	SSPVEYKKTDAPQPD VKEEEEEKEEEKDKGDEEEEGEEKL EEKQKSDAEEDGGTA SQEEDDRKPKAEEDE ILNRSPRNRKPRREX	594
Mouse calreticulin	EKQNKDKQDEEQR LKESEEDKKRKEEEE AEDK-EDDDDRDEDE DEEDEKEEDEEE-SP GQAKDEL	416
Consensus	K	630

# Figure 6: Hydrophobicity plot and topology of Cne1p.

Hydrophobicity plot (A) and predicted topology (B) of Cne1p showing the 5 predicted sites of N-linked glycosylation and the single transmembrane domain at the extreme carboxyl terminus. The predicted signal sequence cleavage is at residue threonine 20 (T20).

.



Residue Number

В

Α



# Identification of Cne1p as an integral membrane protein

Antibodies were raised to Cne1p which was expressed in *E.coli* as a fusion protein with GST and purified by affinity chromatography on glutathione beads. This antiserum recognized a protein in yeast of 76 kDa which was present in a particulate cell fraction. To determine if Cne1p is an integral membrane protein, membrane preparations were solubilized in SDS, sodium deoxycholate or Triton X-100. No significant extraction of calnexin was observed with either sodium carbonate at pH 11.5, 0.5 M NaCl or 2.5 M urea (Fig. 7A). By these criteria, the properties Cne1p correspond to those expected of an integral membrane protein .

The identification of Cne1p as a doublet at a molecular mass of approximately 76 kDa on SDS-PAGE is higher than that expected from the predicted sequence. In order to determine if the protein was N-glycosylated, solubilized membranes were digested with Endo-H and analyzed by SDS-PAGE. This treatment resulted in an increased mobility of the protein with an apparent molecular mass of 60 kDa (Fig. 7B). This change corresponds to that predicted if all 5 potential sites of glycosylation were modified by the addition of core sugars (ca. 3 kDa for each site). However, the predicted molecular mass of the nonglycosylated protein is 56 kDa.

# Subcellular localization of Cne1p

We determined the subcellular location of Cne1p by differential and analytical subcellular fractionation as well as by fluorescence microscopy. Differential centrifugation identified most of Cne1p in the large granule (ML) fraction of *S.cerevisiae* homogenates. The ML fraction was enriched in NADPH cytochrome c reductase activity as determined by de Duve plots which reveal the quantitative distribution of this marker enzyme for the ER (Fig. 8). Analytical centrifugation was then carried out with the ML fraction. Density gradient centrifugation revealed a similar distribution of the ER lumenal protein Kar2p and Cne1p (Fig. 9). This distribution corresponded to median densities of 1.195 g/cc for both proteins (Fig. 10) which was also that of NADPH cytochrome c reductase (1.195 g/cc). However, these distributions were clearly different than those of the Golgi marker enzyme GDPase (median density 1.138 g/cc), the plasma membrane marker ATPase (median density 1.156 g/cc) and the mitochondrial marker monoamine oxidase (median density 1.177 g/cc). Cne1p is not localized to the vacuole since the antibodies for carboxypeptidase Y revealed it to be principally in the N fraction, with very little in the ML or P fractions (data not shown).

# Figure 7: Cne1p is an integral membrane glycoprotein.

Panel A, spheroplasts were prepared and extracted with SDS (0.1%), sodium deoxycholate or DOC (1%), 0.1 M sodium carbonate, pH 11.5, Triton X-100, 0.5 M NaCl (high salt), 2.5 M urea or Tris buffered saline pH 7.5 (mock) followed by centrifugation (30 min. at 100,000 x g to give a pellet (P) and supernatant (S) fraction. Molecular mass markers are indicated on the left. Panel B, a total particulate fraction of homogenized spheroplasts was digested with Endo-H giving a change in mobility of calnexin from a doublet at ca. 76 kDa to 60 kDa. Molecular mass markers are indicated on the right.

CARBONNTE PH 11.5 WHEELUAR WEEK TRIONO.100 TRIONIS HIGH SALT MOL UREA హ l П ΡS P S PS PS PS ≢nj ΡS ΡS ΡS ΡS **17** ŧ. 106 — 0.5 80 — 49.5-. . 14 **k** -- • .

B



Α

# Figure 8: Comparison of the distribution of the ER marker enzyme NADPH cytochrome c reductase and Cne1p.

Panel A: Differential centrifugation of *S.cerevisiae* homogenates into nuclear (N), large granule (ML), microsomal (P) and cytosolic (S) fractions with the distribution of NADPH cytochrome c reductase expressed as a de Duve plot (de Duve, 1975). Panel B: The distribution of *S.cerevisiae* Cne1p in the same fractions (30  $\mu$ g protein was applied to each lane except for P, to which 60  $\mu$ g of protein was applied and detected by immunoblotting with anti-Cne1p antiserum). The ML fraction contains the highest specific activity of NADPH cytochrome c reductase (Panel A) as well as Cne1p (Panel B).


# Figure 9: Isopycnic sucrose density gradient centrifugation analysis of the distribution of Kar2p and Cne1p in the parent ML fraction.

ML fractions were centrifuged on linear sucrose gradients as described in Materials and Methods and equal volumes of each fraction were examined for their content of Kar2p and Cne1p determined by immunoblotting with their respective antibodies. The median density of the Kar2p containing compartment was 1.1951 g/cc and that for calnexin was 1.1955 g/cc.



fractions

### Figure 10: Sucrose density gradient analysis.

The distribution of marker enzymes for the Golgi marker enzyme GDPase, the plasma membrane marker ATPase, the mitochondrial marker monoamine oxidase, the ER markers NADPH cytochrome c reductase, the ER lumenal protein Kar2p and the membrane protein Cne1p as determined by analysis of sucrose density gradient. The quantitative distribution of enzyme activities was evaluated as described in Materials and Methods and that of Kar2p and calnexin by densitometric evaluation of the data of Fig. 9. The median densities for the distribution of the respective constituents are indicated.



Hence, the distribution of Cnelp corresponded most closely to that of the ER lumenal protein Kar2p.

Further examination was carried out by epifluorescence (Fig. 11A-C) and confocal immunofluorescence microscopy (Fig. 11D). Cne1p (Fig. 11C) was colocalized to a compartment identical to that for the ER lumenal protein Kar2p (Fig. 11B); i.e. perinuclear and in filamentous structures extending into the cytosol. DAPI staining of the nuclei is shown in Fig. 11A. Cells were analyzed by confocal microscopy (Fig. 11D) with a strong perinuclear staining pattern observed for Cne1p. In Fig. 11A to 11C a sandwich protocol was used (Schulze and Kirschner, 1986) whereby rhodamine fluorescence is specific for Cne1p, likewise FITC fluorescence is specific for Kar2p distribution.

#### Cne1p is not a prominent 45Ca binding protein of S.cerevisiae ER

We have previously demonstrated that mammalian calnexin and associated SSR $\alpha$  are the major integral membrane proteins of the ER which bind  $^{45}$ Ca in an overlay assay. As shown in Figure 12, two integral membrane proteins of dog pancreatic ER corresponding to canine calnexin (90 kDa) and SSRa (35 kDa) bound <sup>45</sup>Ca. An FR fraction from S.cerevisiae was isolated as pooled fractions 9-18 from Figure 9. Separation into peripheral and integral membrane proteins by the method of Bordier (Bordier, 1981) revealed that the 6 major <sup>45</sup>Ca binding proteins of the yeast ER fractionated into the aqueous phase. These proteins most likely correspond to lumenal ER proteins. As mitochondrial membrane contamination is notoriously associated with 'pure ER' fractions in yeast, the calcium binding proteins encountered in this analysis may well derive from mitochondria. <sup>45</sup>Ca binding to an integral membrane protein of the expected mobility of Cne1p was not detected. This conclusion was supported by further experiments using a GST::Cnelp fusion protein, expressed and purified in *E.coli*. This protein did not reveal detectable 45Ca binding by the 45Ca overlay protocol, although control proteins (parvalbumin, calmodulin) were reactive (data not shown). This is the first report identifying Ca<sup>2+</sup> binding proteins in S.cerevisiae ER although Cnelp is not one of them.

# Figure 11: Double immunofluorescence of Cne1p and Kar2p in S.cerevisiae by epifluorescence and confocal microscopy.

Field showing nuclear staining with DAPI (Panel A). Same field showing Kar2p distribution (Panel B) and Cne1p distribution (Panel C) by epifluorescence microscopy. ER localization of Cne1p by confocal immunofluorescence microscopy (Panel D). The bar represents 2 µm.









#### Figure 12: Identification of <sup>45</sup>Ca binding proteins in S.cerevisiae ER.

Integral membrane proteins (100  $\mu$ g) from dog ER (lane 1) and from *S.cerevisiae* ER (50  $\mu$ g protein) (lane 2) as well as from detergent (lane 3) and aqueous (lane 4) phases of Triton X-114 extracted *S.cerevisiae* ER (100  $\mu$ g protein) were electrophoresed on SDS-PAGE and transferred to nitrocellulose membrane. In the aqueous phase, six polypeptides of molecular masses 26, 35, 50, 59, 66, 72 kDa were identified as <sup>45</sup>Ca binding proteins of *S.cerevisiae* ER. Integral membrane proteins of 90 kDa and 35 kDa corresponding to mammalian calnexin and SSR $\alpha$  were identified in the Triton X-114 phase of dog pancreatic ER. Molecular mass markers as indicated on the left.



#### Deletion of the CNE1 gene

To determine the phenotype of CNE1, the CNE1 gene was deleted by inserting the LEU2 gene into an internal deletion of CNE1 creating plasmid pFP 10.13 (Fig. 13). The plasmid was linearized, transformed into strain W303D and  $LEU^+$  diploids were selected. The transformed diploid was then sporulated and 7 asci were dissected. For every tetrad, all four spores were viable showing that the gene is not essential for viability. CNE1 RNA was not detected in the  $leu2^-$  spore (Fig. 13A) and neither was Cne1p as determined by immunoblots of particulate and soluble fractions isolated from the CNE1 deleted strain (Fig. 13B). The 30 kDa band as compared to wild type protein found in lane 1 represents a fragment of Cne1p which was sometimes observed (Fig. 13 B, lane 1). The protein was not detected by double immune epifluorescence or confocal immunofluorescence examination of *S.cerevisiae CNE1* deleted strains with Cne1p specific antisera (not shown).

#### **Cne1p** and secretion

To test if Cne1p is a molecular chaperone for glycoproteins (Bergeron *et al.*, 1994) the secretion of the glycoproteins acid phosphatase (Fig. 14) and  $\alpha$ -pheromone (Fig. 15) was determined in *CNE1* deleted strain. The levels of secreted  $\alpha$ -pheromone in wild type, deleted or overexpressing *CNE1* strains are identical, as determined by halo assay. Likewise when *CNE1* expression is induced or repressed, levels of cell surface acid phosphatase remain constant.

# Figure 13: Gene disruption of *S.cerevisiae CNE1* and evaluation by Northern blot and Western blot.

Schematic representation of plasmid pFP10.12 containing the entire *CNE1* gene and pFP10.13 containing  $\triangle cne1::LEU2$ . The *CNE1* open reading frame is shaded in black. Restriction sites referred to in the text are shown. Panel A, total RNA from cells containing wild type copy and  $\triangle cne1::LEU2$  was prepared (Wise, 1991) and probed with labeled DNA containing the entire *CNE1* gene. 20 µg of total RNA was loaded per lane and transferred to nylon membrane. Lane 1 (-),  $\triangle cne1::LEU2$  spore disruptant and lane 2 (+) wild type spore for *CNE1*. *CNE1* RNA is not detected in  $\triangle cne1::LEU2$  disrupted cells. Panel B, Immunoblot detection of *S.cerevisiae* Cne1p. Total particulate (P) and cytosolic (S) fractions from yeast cell lysates from wild type *CNE1* (lanes 1, 2) or  $\triangle cne1::LEU2$  strains (lanes 3, 4) were analyzed by immunoblotting with anti-Cne1p antisera. 20 µg protein were applied to each lane. Molecular mass markers are indicated on the left.



Ρ

В



Α

# Figure 14: Acid phosphatase secretion.

Acid phosphatase content was evaluated in *CNE1* deleted strains transformed with a calnexin *GAL* promoter construct. Cells were grown in sucrose to an  $OD_{600}$  of 0.1 and then induced 2% with galactose, or repressed with 2% glucose. Sucrose was supplemented to 4% final concentration. Aliquots were taken at the indicated times for acid phosphatase as described in Materials and Methods.



Acid Phosphatase Activity (nmole p-nitrophenol/hr.)

# Figure 15: Halo assay for $\alpha$ -pheromone production.

Panel A, Wild type strain (W303-1b pVT), Panel B, *CNE1* deleted strain (W303-1b  $\triangle cne1::LEU2$  pVT) or Panel C, *CNE1* overexpressing strain (W303-1b  $\triangle cne1::LEU2$  pVT-*CNE1*) were spotted on a lawn of **a**-mating type cells (strain M200-6C, as described in Materials and Methods). Agar plates were incubated at 30°C for two days to allow haloes to develop.



Α

B

С

The soluble glycoprotein  $\alpha_1$ -antitrypsin is a substrate for mammalian calnexin (Ou *et al.*, 1993; Le *et al.*, 1994) and PI Z  $\alpha_1$ -antitrypsin variant has been shown to be retained by calnexin prior to its degradation or accumulation in the ER (Le *et al.*, 1994). When heterologously expressed in yeast, both wild type and the PI Z variant  $\alpha_1$ -antitrypsin are retained in the ER with the mutant form being degraded therein (McCracken and Kruse, 1993). Hence, we were interested to determine the role of Cne1p in the retention of wild type and PI Z variant of mammalian  $\alpha_1$ -antitrypsin. The amount of secreted  $\alpha_1$ -antitrypsin was tested in wild type and *CNE1* disrupted strains by growing the appropriate strain on agar plates overlaid with nitrocellulose and immunoblotting with antiserum to  $\alpha_1$ -antitrypsin. Both wild type (pVT-AIPi) and the PI Z-mutant (pVT-AIPz) of  $\alpha_1$ -antitrypsin were secreted to a higher extent in *CNE1* disrupted cells than in wild type cells (Fig. 16). Quantitation of the blots showed a 2 to 2.6 fold increase in secretion from calnexin disrupted cells (Table 5).

The evaluation of a possible retention function for *CNE1* was extended to an endogenous yeast seven transmembrane glycoprotein, the  $\alpha$ -pheromone receptor, Ste2p. This protein is normally present and functional in the plasma membrane of *S.cerevisiae*, but the Ste2-3ts mutant protein has been shown to be intracellularly retained (D. Jeness, pers. comm.) at restrictive temperature (37°C), resulting in a 100 fold decrease in mating frequency. To determine if Cne1p plays a role in the intracellular retention of Ste2-3ts protein, we evaluated its function at the cell surface with a quantitative mating assay. At the non-permissive temperature, the relative mating efficiency was 5 fold greater in *CNE1* deleted strains indicating increased transport and/or function of Ste2-3ts protein at the plasma membrane (Table 6).

#### Figure 16: Effect of Cne1p on the secretion of $\alpha_1$ -antitrypsin.

Wild type  $\alpha_1$ -antitrypsin (pVT-AlPi), PI Z variant  $\alpha_1$ -antitrypsin (pVT-AlPz) or vector alone (pVT) were transformed into W303-1a (*CNE1*) or W303-1a  $\Delta cne1::LEU2$  ( $\Delta cne1::LEU2$ ) cells. Equal numbers of cells were spotted onto agar plates, overlaid with nitrocellulose membrane and incubated overnight at 30°C. The nitrocellulose membrane was washed, immunoblotted with anti- $\alpha_1$ -antitrypsin antisera and revealed by the alkaline phosphatase method (see Materials and Methods).

# *∆* cne1:: LEU2 CNE1



Plasmid	Relative Amount Secreted <sup>a</sup> <u>Acne1::LEU2/CNE1</u>	
pVT-AlPi	2.0± 0.3	
pVT-AIPz	2.6 ± 0.6	

TABLE 5: Effect of *CNE1* on the secretion of  $\alpha$ 1-antitrypsin in *S.cerevisiae* 

<sup>a</sup>Densitometric evaluation of secreted  $\alpha_1$ -antitrypsin. Table lists the mean  $\pm$  standard deviation values of 7 experiments for pVT-AIPi and 3 experiments for pVT-AIPz.

TABLE 6: Effect of CNE1 on the secretion of Ste2-3ts protein.

Plasmid	Relative Mating Frequency <sup>a</sup>	
	23 <sup>0</sup> C	37°C
pAD13-CNE1	92 ± 6.5	1.1 <u>+</u> 0.2
pAD13	100	5.4 <u>+</u> 0.6

Relative mating frequency was tested for strain DJ-283-7-1a transformed with pAD13 or pAD13-*CNE1*. pAD13 is deficient for *CNE1* and pAD13-*CNE1* is wild type for *CNE1*.

<sup>a</sup>The frequency of mating was calculated as the ratio of the number of dipoids formed on selective media to the number of input cells. The reported values are the mean  $\pm$  standard deviation values of three experiments. The mating frequency for pAD13 at 23<sup>o</sup>C is normalized to 100.

### 2.5 Discussion

In mammalian cells, calnexin has been shown to have a central role in the retention of incompletely folded glycoproteins in the ER and in the assembly of multisubunit cell surface receptors (Bergeron *et al.*, 1994). The presence of a calnexin homologue would be of considerable interest as its function could be studied using the range of tools available in this organism. An important question is whether *CNE1* is the calnexin or calreticulin homologue in yeast. We have addressed this question in three ways: by a comparison of the sequences, by an analysis of the protein and by the phenotype of *CNE1* deleted cells.

The PCR strategy that we employed was expected to generate yeast DNA sequences which corresponded to calreticulin as well as calnexin. Although 11 separately cloned 250-350 bp products of the PCR reaction were sequenced, only the yeast *CNE1* sequence was detected as an open reading frame (5 out of 11 clones). All other clones sequenced did not have an open reading frame and did not contain internal similarities to calnexin or calreticulin. This PCR generated sequence was used as a probe to clone the complete *CNE1* gene from a yeast plasmid library. Of the two different plasmids recovered, both contained the same *CNE1* gene. Using the complete *CNE1* sequence as a probe, we further determined if there were related sequences in the yeast genome using the lambda clone grid filters. Using hybridization at low stringency on these filters and on a Southern blot of DNA from a *CNE1* disrupted strain, we were unable to detect any related sequences. Thus by hybridization criteria there do not appear to be genes in yeast which are closely related to *CNE1*. The *CNE1* gene we mapped by this technique is located on the left arm of chromosome 1, distal to genes *CDC24* and *CDC19* and to other known mapped genes (Riles *et al.*, 1993).

Mammalian calnexin and calreticulin have the motifs of KPEDWDE repeated three or four times. Only one related motif was found in *S.cerevisiae CNE1* at residues 255 to 261 consisting of KPHDWDD. Mammalian calnexin also reveals 3 repeats of GXW. Only 2 were found in *CNE1*. In the plant *A.thaliana*, a calnexin gene has been identified with greater sequence similarity to mammalian calnexin than that of *S.cerevisiae* (Huang *et al.*, 1993). All four KPEDWDE motifs are retained as well as the 3 GXW motifs and a cytosolic tail albeit without sequence identity to that of mammalian calnexins. In addition, the overall organization of Cne1p terminates in a hydrophobic sequence and lacks the carboxy terminal cytosolic domain found in other calnexins (Fig. 3). We also confirmed that there is not a motif for an RNA splice site present which could account for an alternative *CNE1* sequence. The sequence of the predicted *S.cerevisiae* Cne1p protein predicts an N-terminal hydrophobic signal sequence. N-linked glycosylation sites, and a carboxy terminal hydrophobic potentially membrane spanning sequence. We confirmed the localization of Cne1p in the yeast ER by differential and analytical subcellular fractionation and by epifluorescent and confocal immunofluorescence microscopy which showed a colocalization of Cne1p and the ER lumenal protein Kar2p. We confirmed that Cne1p is an integral membrane protein as it could not be extracted from membranes by treatment with 2.5 M urea, high salt and sodium carbonate at pH 11.5. This is a property that Cne1p shares with mammalian calnexin which is also an integral membrane protein, whereas calreticulin is a soluble ER lumenal protein. We also confirmed that Cne1p has N-linked glycosylation as predicted from the sequence. After Endo-H treatment, the relatively tight mobility of Cne1p in SDS-PAGE was altered by about 18 kDa, indicating that all potential N-glycosylation sites are utilized (Herscovics and Orlean, 1993).

An ER membrane protein such as Cne1p (depicted in Fig. 6A) is unusual because only one amino acid is predicted to be cytosolically exposed. Since we have demonstrated localization of Cne1p in the yeast ER, there is a question of how it is retained. We have confirmed that *S.cerevisiae* Cne1p was not GPI linked since no incorporation of <sup>3</sup>H-inositol was detected nor was the protein susceptible to digestion by PI specific phospholipase C. In mammalian calnexin, the cytosolically oriented sequence RKPRRE has been shown to act as retention and/or retrieval sequences, maintaining this type I integral membrane protein in the ER (David *et al.*, 1993). The lack of a cytosolic tail for *S.cerevisiae* Cne1p but its localization to the yeast ER implies that retention is effected by association with an unknown resident membrane or lumenal protein and hot by the cytosolic proteins interacting with a retention motif (Jackson *et al.*, 1994).

Mammalian calnexin has been shown to be one of two major calcium binding integral membrane proteins of the ER (Wada *et al.*, 1991). Similar experiments with yeast ER membranes showed that there do not appear to be any abundant calcium binding proteins present in the ER membrane (Fig. 12), although we did detect yeast ER lumenal calcium binding proteins. Indeed this is the first demonstration of calcium binding proteins in the ER of *S.cerevisiae*. Confirmation of the inability of yeast Cne1p to bind calcium *in vitro* was obtained with isolated *E.coli* produced GST::Cne1p fusion protein (not shown). Calcium has been demonstrated to be essential for the binding of mammalian calnexin with its protein substrates (Wada *et al.*, 1991; Ou *et al.*, 1993). Although Cne1p has sequence similarity with mammalian calnexin, it is atypical in that it is N-glycosylated, it is an integral ER

membrane protein but does not have a recognizable retention mechanism, and unlike mammalian caluexin it is not a strong calcium binding protein.

Calnexin genes from different organisms show a considerable conservation in their sequence suggesting that the function of the protein is similar and that the preservation of the sequence is important for that function. Mammalian calnexin has been identified as a molecular chaperone for newly synthesized soluble and membrane bound glycoproteins of the secretory apparatus (Bergeron *et al.*, 1994). Mammalian calnexin has also been identified as responsible for the ER retention of soluble and membrane bound proteins prior to their exit from the ER. These functions suggested that there would be an essential phenotype for yeast cells which lack calnexin. However, yeast strains carrying a deletion of the *CNE1* gene were viable, grew at normal rates and we were unable to identify any effect on the secretion of the glycoproteins  $\alpha$ -pheromone or acid phosphatase. From the results with some mammalian secretory proteins, there is evidence that they bypass the participation of calnexin in their folding (Ou *et al.*, 1993). This observation has been attributed to alternative, or back up, mechanisms for protein folding in the mammalian ER (Bergeron *et al.*, 1994).

We did observe an effect on i) the retention of heterologously expressed  $\alpha_1$ antitrypsin in *S.cerevisiae*, as well as ii) the function of a temperature sensitive mutant *ste2-3ts* of the  $\alpha$ -pheromone receptor in *CNE1* disrupted cells. The effect on *ste2-3ts* could be due to an effect of Cne1p on Ste2-3p intracellular trafficking or on Ste2-3p function at the plasma membrane. The latter explanation is less likely since Cne1p is clearly localized in the ER. Although these effects are small, they suggest that Cne1p is a constituent of the yeast quality control apparatus participating in the retention of heterologously expressed or incorrectly folded proteins.

There remains the question of whether the *CNE1* gene we have identified and its gene product, Cne1p, we have characterized represents the yeast calnexin homologue. Alternatively, there may be a closer relative of mammalian calnexin or calreticulin in the yeast genome. We obviously cannot totally exclude this possibility, but the genetic methods currently available in this organism provide an opportunity to identify genes whose function are synergistic with *CNE1*.

# **CHAPTER 3**

The calnexin homologue *cnx1*+ in Schizosaccharomyces pombe, is an essential gene which can be complemented by its soluble ER domain.

### 3.1 Abstract

Secretory proteins become folded by the action of a number of molecular chaperones soon after they enter the endoplasmic reticulum (ER). In mammalian cells, the ER membrane protein calnexin has been shown to be a molecular chaperone involved in the folding of secretory proteins and in the assembly of cell surface receptor complexes. We have used a PCR strategy to identify the *Schizosaccharomyces pombe* calnexin homologue,  $cnx1^+$ . The  $cnx1^+$  encoded protein, Cnx1p, was shown to be a calcium binding type 1 integral membrane glycoprotein. At its 5' end, the  $cnx1^+$  gene has consensus heat shock transcriptional control elements and was inducible by heat shock and by the calcium ionophore A23187. Unlike the sequence related *S.cerevisiae CNE1* gene, the *S.pombe*  $cnx1^+$  gene disruption but the full length mammalian calnexin could not. The ER lumenal domain of Cnx1p, which was secreted from cells, was capable of complementing the  $cnx1:ura4^+$  lethal phenotype. The equivalent region of mammalian calnexin has been shown to possess molecular chaperone activity. It is possible that the lethal phenotype is caused by the absence of this chaperone activity in the *S.pombe*  $cnx1^+$  gene disruption.

# 3.2 Introduction

In cukaryotes, the ER is the site of the folding of secretory proteins and the assembly of multimeric cell surface receptors. These processes are mediated by molecular chaperones (reviewed in Bergeron *et al.*, 1994). Some of these molecular chaperones have been identified and appear to be present in a wide variety of eukaryotic species. One such molecular chaperone, the ER membrane protein calnexin, interacts with secretory glycoproteins soon after they enter the ER (Ou *et al.*, 1993). Newly synthesized glycoproteins specifically associate with calnexin while they are monomeric, incompletely folded and their oligosaccharide modification is the GlcNAc<sub>2</sub>Man<sub>9</sub>Glc<sub>1</sub> intermediate (Ou *et al.*, 1993; Wada *et al.*, 1994; Le *et al.*, 1994; Hammond *et al.*, 1994). Recently, it has been demonstrated that the GlcNAc<sub>2</sub>Man<sub>9</sub>Glc<sub>1</sub> oligosaccharide interacts directly with calnexin (Ware *et al.*, 1995). In addition, tunicamycin (Ou *et al.*, 1993) and the glucosidases I and II inhibitors deoxynojirimycin and castanospermine (Hammond *et al.*, 1994) inhibit the association of incompletely folded glycoproteins with calnexin, supporting the observation that the GlcNAc<sub>2</sub>Man<sub>9</sub>Glc<sub>1</sub> intermediate is important for calnexin recognition.

Calnexin is associated with proteins while they are being folded by the action of other ER molecular chaperones and chaperonins. The time of association of a secretory protein with calnexin reflects the time required for its folding (Ou *et al.*, 1993). Secretory proteins that are not assembled correctly into complexes or do not fold correctly due to mutations or incorporation of amino acid analogues, are retained by calnexin in the ER (reviewed in Bergeron *et al.*, 1994). Thus calnexin also has a function as a constituent of an ER quality control apparatus.

In common with other molecular chaperones, calnexin genes have been identified in a wide variety of eukaryotes including mammals, nematodes and plants (Wada *et al.*, 1991; Sulston *et al.*, 1992; Huang *et al.*, 1993). They share sequence motifs with the ER lumenal protein calreticulin, which has similarly been found in a number of eukaryotes. Together they appear to form a gene family with a possible similarity in function (Wada *et al.*, 1991; Smith and Koch, 1989). How these proteins perform their function is a topic of considerable interest.

We have recently cloned the *CNE1* gene from *S.cerevisiae* which shares sequence similarity with mammalian calnexin and calreticulin. The *CNE1* gene codes for an integral membrane ER glycoprotein, Cne1p. Unlike mammalian calnexin, Cne1p does not bind

calcium in an *in vitro* assay, and it does not have a cytosolically directed domain at the carboxy terminus (Parlati *et al.*, 1995b). We demonstrated that *CNE1* does have an effect on the retention of mutant proteins in the ER (Parlati *et al.*, 1995b) but it is not essential for the viability of *S.cerevisiae*. Thus, not all of its properties correspond with those expected of a *bona fide S.cerevisiae* calnexin homologae. However, the *S.pombe cnx1+* gene which we characterized has several properties in common with mammalian calnexin and is most likely the *S.pombe* calnexin homologue.

### **3.3 Materials and Methods**

#### Strains and media

S.pombe strains Q358 ( $h^{-}$  leu1-32 ura4-D18 ade6-M210) and Q359 ( $h^{+}$  leu1-32 ura4-D18 ade6-M216) were used. The strains were grown at 30°C in YPD medium or EMM media supplemented with nutrient requirements as previously described (Moreno *et al.*, 1991). *E.coli* strain MC1061(Maniatis *et al.*, 1982) was used.

#### Cloning of S.pombe cnx1+

To amplify the calnexin/calreticulin gene equivalent from *S.pombe*, the degenerate oligonucleotides, PCR primer **S** (5' AARCCNGARGAYTGGGAYGA 3') and PCR primer **A** (3' ATRTTYCCNYTYACCTTYGG 5') were used to amplify genomic DNA. Amplification reactions were performed as previously described (Parlati *et al.*, 1995b). The products of the PCR reaction were electrophoresed on a 2% agarose gel and stained with ethidium bromide. A band at approximately 350 bp was purified and cloned into the *Sma* 1 site of pTZ-19R (Pharmacia). The fragment was sequenced and was found to encode a peptide with high amino acid sequence similarity with canine calnexin. The PCR fragment was then <sup>12</sup>P radiolabeled using the QuickPrime method (Pharmacia) and subsequently used as a probe to clone full length calnexin from a *S.pombe* genomic library in plasmid pWH5 (Wright *et al.*, 1986). This gene was sequenced by standard procedures and the sequence of *cnx1*<sup>+</sup> was released to GenBank (accession number: M98799) on December 31, 1993.

#### **RNA** extraction and transcript analysis

A 100 ml culture of strain Q360 was grown overnight at 23°C to an OD<sub>600</sub> of 1 and a 25 ml aliquot was transferred to a 39°C bath for 15 or 30 minutes. Cells were then rapidly collected by centrifugation and frozen immediately in a dry ice-methanol bath. Overnight cultures (25 ml) grown to an OD<sub>600</sub> of 1 were also treated with either 1 µg/ral tunicamycin (Sigma), 10 µM A23187 (Sigma) or 10 mM 2-deoxyglucose (Sigma) for three hours at 30°C. Cells were then harvested and quickly frozen in dry ice/methanol bath. RNA extraction was performed by the hot phenol method (Wise, 1991) and Northern analysis was performed as previously described (Maniatis *et al.*, 1982), using a <sup>32</sup>P-probe containing the entire open reading frame of *cnx1*<sup>+</sup>. Densitometric analyses used an LKB Ultrascan Laser Densitometer.

#### Antibody production

Polyclonal antibodies recognizing Cux1p were obtained by immunizing rabbits with GST::Cux1p fusion proteins expressed in *E.coli*. This fusion protein was made by inserting a PCR fragment encoding amino acid 23 to 492 into pGex-2T (Smith and Johnson, 1988).

#### Membrane extraction and Endo-H digestion

*S.pombe* membranes (ML fraction) were prepared and treated essentially as described (Parlati *et al.* 1995b). Membranes were mixed with 1 volume of either 1 M NaCl, 0.2 M sodium carbonate pH 11.5, 5 M urea, 2% Triton X-100, or 0.2% SDS and were subsequently analyzed as described (Feldheim *et al.*, 1992). The Cnx1p antiserum was used at a 1:4000 dilution. Endo-H digestions were performed by incubating 20  $\mu$ g of ML fraction proteins in 100 mM sodium acetate pH 4.9, 150 mM NaCl, 10 mM DTT, 1% Triton X-100, 0.1% SDS + inhibitors (1 mM PMSF, 1  $\mu$ g/ml pepstatin, 1  $\mu$ g/ml leupeptin and 1  $\mu$ g/ml aprotinin) and incubating with 2  $\mu$ g of Endo-H for 16 hours at 37°C.

#### <sup>45</sup>Ca overlay

The *S.pombe* GST::Cnx1p fusion (residues 23 to 492) was expressed in *E.coli* and purified (Smith and Johnson, 1988). Samples were electrophoresed by SDS-PAGE and prepared for  ${}^{45}$ Ca overlay exactly as previously described (Wada *et al.*, 1991). Control experiments were carried out with canine pancreatic ER membranes obtained by extraction with Triton X-114 (Wada *et al.*, 1991), and Cne1p as previously described (Parlati *et al.*, 1995b).

#### Gene disruption

A 4.3 kb *Pst* 1 DNA fragment containing the complete *cnx1*<sup>+</sup> gene was cloned into pTZ19R creating plasmid pFPP3. A PCR fragment containing nucleotides -490 to -1 (5' to the coding sequence) was amplified, cut with *Spe* 1 and cloned into the *Spe* 1/*Sma* 1 site of pBluescript KS+ creating plasmid pFPP3.1. The *Nsi* 1/*Eco* RV fragment (3' to the coding sequence) from pFPP3 was cut, purified and inserted into the *Eco* RV site of pFPP3.1 creating plasmid pFPP3.2. A 1.8 kb *Hind* III fragment containing the *ura4*<sup>+</sup> gene (Grimm *et al.*, 1988) was blunt ended and inserted into the similarly treated *Eco* R1 site of pFPP3.2

creating plasmid pFPP4 essentially disrupting the cnx1+ gene. Plasmid pFPP4 was digested with Xba I, Hind III (which cleave on either side of the insert) and Xmn I (which cuts in the vector) and approximately 5  $\mu$ g of the linear fragment containing the disrupted gene was isolated and purified by the GeneClean method. This fragment was used to disrupt the diploid strain h<sup>-</sup> leu1-32 ura4-D18 ade6-M210 /h<sup>+</sup> leu1-32 ura4-D18 ade6-M216 and uracil prototrophs were selected. Colonies were grouped into pools of 10 and chromosomal DNA was prepared (Moreno et al., 1991). PCR was used to screen for homologous recombinants using a sense oligonucleotide coding for nucleotides -510 to -495 (5' to the cnxl<sup>+</sup> ORF) and an antisense 20 bp oligonucleotide from within the ura4<sup>+</sup> gene. Among 20 pools screened, 15 were positive for the expected 550 bp PCR product. Four pools were chosen and the PCR reaction was repeated for each constituent colony. We found 4 colonies that were positive and chromosomal DNA was prepared from each. Southern blots were performed by digesting these DNAs with either Eco RI or Eco RV, followed by electrophoresis on 1% agarose gels and transfer to nylon membranes. The <sup>32</sup>P-labeled probes used were: i) the entire ura4<sup>+</sup> gene hybridized to genomic Eco RI and Eco RV digests; ii) PCR fragment from -490 to -1 hybridized to a genomic *Eco* RI digest; iii) *Nsi* VEco RV fragment of cnx1+ hybridized to a genomic Eco RV digest.

#### Complementation of the cnx1<sup>+</sup> gene disruption

Sequences encoding full length  $cnxI^+$  and  $cnxI^+$  terminated at amino acids 524, 484 and 474 were amplified using PCR (see conditions above). Oligonucleotides were designed in order to introduce a stop codon at amino acid 525, 485 and 475 respectively. These amplified sequences as well as full length canine calnexin (Wada *et al.*, 1991) were subcloned into the *Sma* 1 site of *S.pombe* vector pREP1, under the regulation of the *nmt1*<sup>+</sup> promoter (Maundrell, 1993). The constructs were subsequently transformed in the dipoid strain heterozygous for the  $cnxI^+$  deletion, and leu<sup>+</sup> transformants were selected. Random spore analysis of the leu<sup>+</sup> strains was done as previously described (Moreno *et al.*, 1991). Spores were plated onto phloxine B agar plates supplemented with adenine, in order to detect easily haploids that were both leu<sup>+</sup> (expressing the recombinant protein) and ura<sup>+</sup> (disrupted for  $cnxI^+$ ). These haploids were subsequently tested for the presence of Cnx1p recombinant protein intracellularly and extracellularly by immunoblotting.

# 3.4 Results

#### Cloning of a S. pombe member of the calnexin/calreticulin family, cnx1+

We have previously used sequence motifs that are shared between mammalian calnexin and the ER lumenal protein calreticulin to identify a calnexin related gene in S.cerevisiae, CNE1 (Parlati et al., 1995b). We used a similar strategy for cloning of a S.pombe member of the calnexin/calreticulin family. Degenerate oligonucleotides coding for the regions of amino acid similarity conserved between mammalian calnexin and calreticulin, YKG<sup>K</sup>/<sub>E</sub>WKP and the repeat motif KPEDWDE (Fig. 4) were used to prime a PCR reaction using S.pombe genomic DNA as a template (Fig. 4). From the organization of these motifs in calnexin and calreticulin, an amplified fragment of 350 bp was expected (Figure 4A) and double stranded DNA products of about this size were cloned into the plasmid pTZ19R. The DNA sequence of one of these clones and its derived amino acid sequence revealed that this PCR product shared high amino acid sequence similarity with both mammalian calnexin and calreticulin. Using this fragment as a probe, the entire gene was cloned and sequenced (Figure 17, see Materials and Methods). The  $cnxl^+$  genomic sequence identified has consensus heat shock elements at its 5' end (Figure 17). The coding sequence does not contain introns and predicts a type I integral membrane protein of 560 amino acids that has a similar overall arrangement to mammalian calnexin. There is a predicted N-terminal cleavable signal sequence, an N-glycosylation motif at residue 418, and a membrane spanning domain proximal to a cytoplasmic domain (Argos *et al.*, 1982; Kyte and Doolittle, 1982; von Heijne, 1986) (Figure 18). S. pombe cnx1+ encodes four repeats related to the motif KPEDWDE, which in calnexin have been shown experimentally to be high affinity, low capacity calcium binding sites (Tjoelker et al., 1994). These amino acid repeats are also present in A.thaliana, in mammalian calnexins and in mammalian calreticulins. Overall, S. pombe cnx1+ is 38% identical to A. thaliana calnexin and 34% identical to canine calnexin and 25% identical to mouse calreticulin, but only 22% identical to S.cerevisiae CNE1. Thus, the S.pombe  $cnx1^+$  gene is most likely to be a calnexin homologue since it has higher amino acid identity with calnexin than calreticulin. Additionally, it encodes a C-terminal domain after a predicted transmembrane domain that is present only in calnexin genes. We have looked for other calnexin and calreticulin homologues in S.pombe. Hybridization at low stringency of a cnx1+ probe with the entire genome, presented in the form of bacteriophage P1 clones, did not identify any other related sequences. We determined that the  $cnxl^+$  gene maps to a region on chromosome I, between probes 57b12 and 20h4 and its location is P1 phage clone

# Figure 17: DNA sequence and deduced amino acid sequence of the $cnxI^+$ gene

*S.pombe cnx1*<sup>+</sup> gene codes for a 560 amino acid protein. The predicted protein contains a signal sequence at the amino terminus (*italics*), a single N-glycosylation site at N418 (CHO), a transmembrane domain (**bold type**) and a 48 amino acid predicted cytosolically oriented sequence. Putative heat shock elements are indicated (<u>underlined</u>). Nucleotides matching the heat shock element (HSE) consensus sequences are marked by asterisks. Sequences encoding partial HSE-elements are also denoted by asterisks, but are not underlined. A putative TATA box is shaded. The sequence is available as GenBank #M98799.

TROTOTTTTTTTTTTTACTTRATOGO-BORGANAATTAAA DI WAATANDIA DIAGTADIAAA DI AAA DIGUDA DI DANAA DI DIGUDA DI DI DI AA ... . . . . . . 12 ATGAAGTACGGAAAGGTATCTTTTCTAGCGTTTTTGT%T%TFTTTATAT%TC/DEFFARTCA/TTEFTATC/DEFAATC/DEFAATC/DEFAATCA/AAAAA/TT M K Y G K V S F L A L L C S L Y V R J S L A P P E S E TTOOTITITAACCCAACTGAAGTGAAACCGCCCCTTTTTOAGGAAFTITUAAFTTUAEFTUAEFUTUAA STTUAEACCCCAACTGAAGTGAAACCGCCCCTTTTTOAACCC NPTEVEAPLVE2F2IAWBEEWII E H A F R F V N G I E E M S Y V J E W T V E E B B B B B A L F B E A GGTCTCGTTATGAAAGACGAAGCCGCTCATCATGCTATFTCTTATGAATTTCGACCCATAAACGAA CAAAAAA GAATTTCGAC LVMRDEAAHHAISYEEDEPINEPERI ..... QYEVNPEEGLNCGGAYLELLAETTHFEMIN 1.00 TCCATCGATTACAGAATCATGTTTGGA2CCGACAAATGTGGGGTTAATSIACNFFFFFFAATSIACNAFTTTTAA SAA AAAAAAAAAAA SIDYRIMFGPDECGVRDRVHFIFFHFNALCIM GGTGAGTATTCCGAGAAGCATCTCGACTCTAGACCCGCCACCTPTTTAAAACCTVPTATTACAAACTTPTTACACTCTTAA TETTAAACCCGCCGCCACCTPTTTAAACCTVPTATTACAACCTTTTTACACTCTTAA G E Y S E K H L D S R P A S L L E P G I T N L Y T L I V F F 1.10 Q T F E V R I N G D V V R Q G S L F Y D F I P P V L P P Ð -14.0 GANATTTACGATCCCGAAGATATCAAGCCTGCTGATIGGGTTGACGAACCTGAAATTCCTGAACCCTAAGACCTTAAGACTTCAAGACTGATFAG F - **I Y D P E D I E P A D W V D E P E I P P P A A V F P D U W** 12.10 GATGAAGACGCCCCGTATGATTCCTGACCCAGTGAAAACCTGACGATTGATGATGATGATGAACCTCTCTACACCTCTCTACACCTCCTGACCCAGTGAACCTCCTGACGCAGTGAACCTCCCGGAGGCAGTGACCCAGTGACCCAGTGACGACGAGTGACGCCAGTGACGACGCAGTGACGCAGTGACGCAGTGACGACGCAGTGACGACGAGTGACGAGTGACGACGAGTGACGAGTGACGAGTGACGACGAGTGACGACGAGTGACGAGTGACGAGTGACGACGAGTGACGAGTGACGAGTGACGAGTG EDAPRMIPDPDAVEPEDWLEDEPLYIPDP 10.1 GAAGCTCAGAAGCCTGAGGATTGGGATGATGAAGAAGATGGTGATTGGATTCCAAGTGAAATGATGATAATGCTAAATGCAAATGATGAAGATGCTGAG 1414.1 E A Q K P E D W D D E E D G D W I P S E I I N P Y LEGA 4.1.5 GGCTGTGGTGAATGGAAGCCACCGATGATTCGTAACCCTAACTATCGT95CCCTT96TCTCCTACTACTAATCCT9AATCCT9AATCCT9AATCCT95AATCCT95CCCT 1000 C G E W K P P M I R N P N Y R G P W J P P M I P N P E F J 41. . 11.0 GEWYPRRIPNPDYFDDDHF55F 5 Y G V ELWTMQPNIRFSNIYVGHSLEDAEFLGNET. 4.11 TTCCTACCTAAATTGAAAGCTGAAAGAGAATTACTTTCTAAGCAAGAGTCCATGGAAAAGCAATCCAATCCATGTAAGAGAGCTCTAAG 1.455 - L P K L K A E R E L L S K Q E S M E F Q J M H V D D E F B QILERFLDVYDIIKAKLPPNVAEYVUYYVE 4.81 11.45 T I I E T P E I G I A I V A V L G S L T A V I L T C Y P Y P 515 11.25 FASSSPASLSTGTTEAEKEQGEKFFGETET 545  ${\tt GAGAAGATAGACGTTTCTTACGCCGAAACTGAATCACCAACTGCGAAGAAGAACTAATGAAACTGCCAAATGAAATGTCAAAAGTTCTC = 1.225$ E K I L V S Y A P E T E S P T A K N E D 5,4, 5, 

ATTTEGTEFACAAAGTTEFGATTGEGACTTEFTGACTGAAAGTCAAAGTAGCGACTCTCTCTTAAAGCCTTA

. . .

. . .

RED BADA CEDA WARD STRATCH CONTACTANA CIMINEL CAAL IA

1880

### Figure 18: Hydrophobicity plot and topology of Cnx1p

Panel A, hydrophobicity plot and Panel B, predicted topology of Cnx1p showing the single predicted N-linked glycosylation site, the single transmembrane domain proximal to the carboxyl terminus and the cytoplasmic tail. The predicted signal sequence cleavage is at residue aspartate 23 (D 23).



A





3G10p (data not shown) (Hoheisel *et al.*, 1993). We tentatively conclude that  $cnxI^+$  is the only *S.pombe* calnexin gene.

#### The cnx1+ transcript is inducible by heat shock and a calcium ionophore

In eukaryotes, the consensus sequence for the heat shock element is characterized by three or more repeats of the sequence nGAAn in an alternating orientation (Pelham and Bienz, 1982; Amin *et al.*, 1988). Inspection of the upstream sequences of  $cnx1^+$  identified two heat shock elements (see Figure 17), which are in the correct position to act as elements controlling transcription of  $cnx1^+$ . To test this, exponentially growing *S.pombe* cells were subjected to a transient heat shock at 39°C and the RNA transcripts were analyzed by Northern blots. The  $cnx1^+$  transcript was found to be induced approximately 1.6 fold by a transient heat shock (Figure 19A, B). If the heat shock treatment was carried out for an extended time, the level of transcription diminished to the basal level (Figure 19A, B). Other agents that are known to cause stress were also tested for their effect on  $cnx1^+$  transcription. Treatment with the glycosylation inhibitor tunicamycin (Figure 19C, D) or with 2-deoxyglucose (data not shown) had no effect, but treatment with the calcium ionophore A23187 gave a marked increase in  $cnx1^+$  transcription (Figure 19C, D).

#### The cnx1+ gene product, Cnx1p, is a calcium binding protein

Calnexin in mammalian cells was originally described as one of the two major calcium binding proteins of the ER membrane (Wada *et al.*, 1991). All calnexins and calreticulins tested thus far bind calcium, except for *S.cerevisiae* Cne1p (Parlati *et al.*, 1995b). In order to determine if Cnx1p is a calcium binding protein, we used a GST fusion of Cnx1p expressed in *E.coli*. This fusion comprised amino acids 23 to 492 of Cnx1p, that is, it excludes the predicted signal sequence and the C-terminal transmembrane and cytoplasmic domains of Cnx1p. Tested in a calcium overlay assay, the GST::Cnx1p fusion protein bound calcium (Figure 20, lane 2) but as expected, GST alone, phosphorylase B and GST::Cne1p (*S.cerevisiae*) (Figure 20, lanes 1, 6 and 7 respectively) did not bind calcium. The positive controls, calmodulin (Figure 20, lane 4), parvalbumin (Figure 20, lane 5) and two bands corresponding to mammalian calnexin and pgp 35 (Figure 20, lane 3) obtained from Triton X-114 extracted canine pancreatic stripped rough microsomes, bound calcium. Thus, like mammalian calnexin and calreticulin, *S.pombe* Cnx1p is a calcium binding protein.
Figure 19: Induction of  $cnx1^+$  by heat shock and calcium ionophore A23187 Northern blot analysis of total *S.pombe* RNA using a probe from the complete ORF of the  $cnx1^+$  gene *S.pombe* strain Q 360 was grown at 23°C and shifted to 39°C for 0 min. (Panels A, B, Iane 1), 15 min. (Panels A, B, Iane 2) and 30 min. (Panels A, B, Iane 3) or grown at 30°C (Panels C, D, Iane 1) or treated with tunicamycin (Panels C, D, Iane 2) or the calcium ionophore, A23187 (Panels C, D, Iane 3). The results of 3 independent experiments evaluated by densitometry are illustrated in Panels A and C (± S.D.). Equal amounts of RNA were applied to each Iane and gels were stained with ethidium bromide in order to verify the quantities of 28S and 18S RNA.













A

#### Figure 20: Identification of S. pombe Cnx1p as a calcium binding protein

Lane 1 - GST alone (20 µg protein), lane 2 - GST::Cnx1p(*S.pombe*) fusion (20 µg protein), lane 3 - Triton X-114 extracted canine pancreas ER (100 µg protein) showing two polypeptides of 90 kDa (calnexin) and 35 kDa (pgp 35) known to bind Ca<sup>2+</sup>(Wada *et al.*, 1991), lane 4 - calmodulin (2 µg protein), lane 5 - parvalbumin (10 µg protein), lane 6 - phosphorylase B (10 µg protein) and lane 7 - GST::Cne1p (*S.cerevisiae*) fusion protein (20 µg protein) were electrophoresed on a 10% SDS-PAGE, transferred to nitrocellulose membrane and <sup>45</sup>Ca binding performed as described previously (Wada *et al.*, 1991). Molecular mass markers are indicated on the left.



#### Cnx1p is a integral membrane glycoprotein

Polyclonal antibodies were raised to the *E.coli* expressed GST::Cnx1p fusion protein. These antibodies recognize a protein in *S.pombe* of 91 kDa that is associated with a membrane fraction. To determine if Cnx1p is an integral membrane protein, *S.pombe* membrane preparations were treated with a regimen of sodium carbonate at pH 11.5 or 0.5 M NaCl or 2.5 M urea which did not release it from the membranes, and with 0.1% SDS and 1% Triton X-100 that led to solubilization of the 91 kDa protein (Figure 21 A). Thus, the properties of *S.pombe* Cnx1p correspond to those expected of an integral membrane protein. Membrane association is expected for a calnexin homologue but not a calreticulin homologue which is a soluble ER lumenal protein.

The apparent molecular mass of 91 kDa for *S.pombe* Cnx1p is higher than predicted from the Cnx1p sequence. To confirm that the predicted site of N-glycosylation at residue 418 is used (see Figure 18), *S.pombe* membrane preparations were digested with endoglycosidase H (Endo-H). Immunoblotting with antibodies to Cnx1p revealed an increase in mobility on SDS-PAGE to approximately 88 kDa (Figure 21B). This mobility change corresponds to that expected if a single potential site of glycosylation (ca. 3 kDa for each site) is modified by the addition of core sugar residues (Herscovics and Orlean, 1993). From the primary sequence, the predicted molecular mass of the nonglycosylated protein is 63 kDa (Figure 17). This anomalous migration on SDS-PAGE is also observed for the nonglycosylated canine calnexin that has a predicted molecular mass of 67 kDa but a mobility on SDS-PAGE corresponding to 90 kDa (Wada *et al*, 1991). As with mammalian calnexin, the low pI of Cnx1p (calculated as pI 4.13 ) is likely to be responsible for this discrepancy.

#### Figure 21: Cnx1p is an integral membrane glycoprotein

Panel A, membranes were prepared and extracted with 0.1% SDS, 1% Triton X-100, 0.1 M sodium carbonate, pH 11.5, 0.5 M NaCl (high salt), 2.5 M urea or Tris-buffered saline pH 7.5 (mock), followed by centrifugation (30 min. at 100,000 x g) to give pellet (P) and supernatant (S) fractions that were analyzed by immunoblotting with anti-Cnx1p antisera. Molecular mass markers are indicated on the left. Panel B, a total particulate fraction of homogenized spheroplasts was digested with Endo-H. Analysis by immunoblotting revealed a change in mobility for Cnx1p from 91 kDa to 88 kDa. After Endo-H treatment, a minor band of 74 kDa is found, most likely originating from the light 78 kDa band in lane 1. The significance of this band is unknown, although we speculate that it may represent a minor Cnx1p degradation product. Molecular mass markers are indicated on the right.



49.5 —

В



1 2

#### The cnx1+ gene is essential

Gene disruption was performed in order to determine if  $cnx1^+$  is an essential gene in *S.pombe*. Approximately 80% of the coding sequence was replaced with the *S.pombe*  $ura4^+$  gene. The sequences flanking  $cnx1^+$ , that is, 500 bp on the 5' of the ORF and 1.0 kb on the 3' side were retained in order to promote a good frequency of homologous recombination (Figure 22A). A diploid *S.pombe* strain (see Materials and Methods) was transformed with this linear construct and uracil prototrophs were selected. The DNA of some transformants was analyzed by PCR and Southern blots to confirm that the recombination had occurred at the  $cnx1^+$  locus (see below). Diploid transformants heterozygous for the  $cnx1^+$  gene disruption grew normally, were sporulated and 17 tetrads were dissected. Each tetrad gave the same segregation of 2 viable spores that grew to form visible colonies and 2 apparently inviable spores (Figure 22C). All the spores that grew were uracil auxotrophs. Thus the  $cnx1^+$  gene in *S.pombe* is essential for viability.

Southern blot analysis of the DNA from the parental diploid strain, the ura4+ heterozygous diploid, and ura<sup>-</sup> spores were probed with the following probes: i) the entire  $ura4^+$  gene; ii) a probe corresponding to nucleotides -490 to -1, 5' to the  $cnxl^+$  ORF; and iii) the Nsi I/Eco RV fragment of cnx1+ (see Figure 22A). These probes were used to distinguish between the wild type and the recombinant allele. DNA from the parental diploid strain (Figure 22B lanes 1, 5, 9, 13), an ura4+ diploid transformant (Figure 22B lanes 2, 6, 10, 14), and ura<sup>-</sup> spores (Figure 22B lanes 3, 4, 7, 8, 11, 12, 15, 16) were digested with Eco RI (Figure 22B lanes 1-4, 9-12) or Eco RV (Figure 22B lanes 5-8, 13-16). Hybridization with the probe of the entire  $ura4^+$  gene resulted in a 4.7 kb band when restricted with Eco RI (lane 2) and two bands of 2.3 kb and 6.0 kb when restricted with Eco RV (lane 6) in the heterozygous diploid. These bands were not present in the parental or ura<sup>-</sup> spores. Hybridization with the probe derived from a PCR fragment 5' to the  $cnxl^+$ gene (nucleotides -490 to -1) revealed a 2.9 kb band for the wild type allele and 4.7 kb band present only in the heterozygous diploid (lane 10). Hybridization with the probe encompassing the Nsi I/Eco RV fragment of cnx/+ resulted in a 1.7 kb band corresponding to the wild type allele and a 2.3 kb band only present in the heterozygous diploid (lane 14) corresponding to  $cnxl^+$  disrupted allele. Thus we have confirmed that the  $cnxl^+$  gene is disrupted by homologous recombination to yield a cnx1+/cnx1:: ura4+ heterozygous diploid. The spores that grew were all ura4<sup>-</sup> and contained the wild type  $cnx1^+$  allele. We conclude that the cnxl gene is essential for viability. Upon microscopic examination of the apparently nonviable spores we observed that they did germinate and divided to produce microcolonies of 20 to 50 cells. Therefore,  $cnxl^+$  is not essential for spore germination but the results suggest that after several cell divisions, Cnx1p becomes diluted and cells eventually stop dividing.

#### Complementation of the cnx1::ura4+ gene disruption in haploid cells

Since the *cnx1::ura4* + gene disruption is lethal in haploids, we could determine which region of the molecule was essential for growth. The *cnx1::ura4*+ gene disruption strain was transformed with a series of deletion plasmids lacking fragments of the  $cnxl^*$  Cterminus. Their ability to complement the lethal cnx1::ura4+ gene disruption, as well as the location of Cnx1p was determined (Figure 23A). Plasmid pCNX560, expressing full length Cnx1p, complemented the lethal phenotype of the *cnx1::ura4*+ gene disruption in haploid cells. The full length Cnx1p was detectable intracellularly in membrane fractions (Figure 23B). Plasmid pCNX524 expressed Cnx1p lacking the cytosolic tail and produced a protein of lower molecular weight, which was also detectable intracellularly in a membrane fraction (Figure 23B, lanes 1, 2, 3). This truncated protein was able to complement the cnx1::ura4 + gene disruption in haploids. Plasmid pCNX484 and pCNX474 both expressed Cnx1p lacking the C-terminal cytosolic domain and the putative transmembrane domain. Remarkably they both complemented the lethal phenotype of the gene disruption. For these constructs, less Cnx1p was detectable intracellularly (Figure 23B lanes 5-10) and the truncated Cnx1p could now be detected in the medium (not shown), thus they are not retained in the ER and are secreted (Figure 23A). We also attempted to complement the gene disruption strain with the full length canine calnexin expressed in the same plasmid. The mammalian protein (FL 90) was detected in diploid cells heterozygous for the  $cnxl^+$ disruption (Figure 23B, lanes 13, 14). However, we could not obtain any transformants where the mammalian calnexin complemented the lethal phenotype of the cnx1::ura4 + gene disruption in haploid cells.

#### Figure 22: Gene disruption of cnx1+

Panel A, schematic representation of plasmid pFPP3 containing the entire  $cnx1^+$  gene and pFPP4 containing cnx1: $ura4^+$ . The indicated restriction sites were used to evaluate by Southern blot analysis the presence or absence of  $cnx1^+$  in the parental and heterozygous diploids, and in the urat spores. Panel B, analysis of  $cnx1^+$  disruption by Southern blot. DNA from the parental diploid strain (lanes 1, 5, 9, 13), heterozygous diploid strain (lanes 2, 6, 10, 14) e: urat spores (lanes 3, 4, 7, 8, 11, 12, 15, 16) was restricted with either *Eco* RI (lanes 1-4, 9-12) or *Eco* RV (lanes 5-8, 13-16) and probed with: i) the entire  $ura4^+$  gene (lanes 1-8); ii) PCR fragment from nucleotides -490 to -1, 5' of the  $cnx1^+$  ORF (lanes 9-12); iii) *Nsi VEco* RV fragment of  $cnx1^+$  (lanes 13-16). The urat spores are from the two tetrads, i.e. lanes 3, 7, 11, 15 from spore 1, and lanes 4, 8, 12, 16 from spore 2. Panel C, results of the tetrad analysis showing two viable and two non-viable spores for each tetrad.









cnx1+



Α

Figure 23: Complementation of the cnx1::ura4 + disruption in haploid cells. Panel A, contructs used to complement the cnx1::ura4+ disruption are depicted with lumenal (dashed), transmembrane (filled) and cytosolic tails (open) rectangles. Their ability to complement (Comp) the cnx1::ura4 disruption as well as their intracellular (I) and extracellular (E) location (Loc) is noted. pCNX560 encodes the entire Cnx1p. pCNX524, pCNX484 and pCNX474 encode Cnx1p truncated at amino acid 524, 484 and 474 respectively. pFL90 encodes the full length canine calnexin gene. Panel B, Immunoblot detection of Cnx1p and canine calnexin in S. pombe cells. 10 µg of protein from a membrane fraction for 3 different cnx1::ura4+ haploids complemented with either pCNX560 (lanes 4, 11, 12), pCNX524 (lanes 1, 2, 3), pCNX484 (lanes 8, 9, 10) and pCNX474 (lanes 5, 6, 7) were electrophoresed on an 8% SDS-PAGE, transferred to nitrocellulose and blotted with anti-Cnx1p antisera. Canine calnexin was detected in a membrane fraction of diploid cells heterozygous for the *cnx1* deletion transformed with pFL90 (lanes 13, 14). 10 µg of membrane proteins were treated as above (lanes 1-12), but immunoblotted with anti-canine calnexin antisera. Molecular mass markers (kDa) are indicated on the left.



B

A



## 3.5 Discussion

It has been established that calnexin acts as a molecular chaperone that recognizes secretory glycoproteins and unassembled multimeric cell surface receptors (see Bergeron et al., 1994). In mammalian cells, its function as a molecular chaperone has been largely elucidated by *in vivo* studies showing the transient association of incompletely folded secretory proteins with calnexin. The molecular mechanism for the recognition of proteins by calnexin remains uncertain. It has been suggested that multiple mechanisms including Nlinked glycosylation and/or direct binding to peptide motifs of incompletely folded proteins are involved (Bergeron et al., 1994; Ware et al., 1995). It is certain that incompletely folded proteins are recognized. For example, transferrin expressed in human HepG2 cells is released from calnexin as its disulphide bonds form (Ou et al., 1993). Also, it is known that calnexin can recognize the free oligosaccharide GlcNAc<sub>2</sub>Man<sub>9</sub>Glc<sub>1</sub> (Ware *et al*, 1995) but that its affinity for this oligosaccharide on a polypeptide is probably higher (Araunachalam *et al.*, 1995). Less is known about features of calnexin that are necessary for its function. Its apparent ubiquity in mammalian cells makes experimental alteration of its function difficult, so the characterization of a yeast calnexin homologue would be an advantage. We have previously cloned CNE1 from S.cerevisiae by a similar strategy to that described here. We have shown that CNE1 has sequence similarities with the calnexin/calreticulin family and shares some of the properties of calnexins but does not bind calcium or have a C-terminal cytoplasmic domain (Parlati et al., 1995b). The S.pombe  $cnxI^+$  gene described here also has sequence similarity and properties in common with calreticulin/calnexin, but several of these characteristics argue that S.pombe cnx1+ codes for a homologue of calnexin whereas the identity of S.cerevisiae CNE1 is less certain.

Over its complete sequence, Cnx1p is 34% identical to canine calnexin and 25% identical to mouse calreticulin, compared with an overall identity of Cne1p of 24% with canine calnexin and 21% with mouse calreticulin. As for all mammalian and the plant calnexins, *S.pombe* Cnx1p has four repeats related to the sequence KPEDWDE, but *S.cerevisiae* Cne1p has a single related motif KPHDWDD with less conserved variations of this motif at the other equivalent positions (Figure 24A). This observation is noteworthy since this repeat motif has been suggested by previous studies to represent the high affinity calcium binding domain of calnexin as well as calreticulin (Michalak *et al.*, 1992; Tjoelker *et al.*, 1994). As shown in Figure 20, Cne1p does not bind calcium whereas Cnx1p does. Mammalian calnexins and *S.pombe* Cnx1p possess four cysteine residues that are

Figure 24: Comparison of predicted high affinity calcium binding motifs (A) and cysteine motifs (B) in *S.pombe* Cnx1p, *A.thaliana* calnexin, canine calnexin, *S.cerevisiae* Cne1p and mouse calreticulin proteins.

Amino acid sequences were aligned using the GeneWorks program. The numbers denote the place in the original sequence of the first residue in the motif.

# Α

S.pombe Cnx1	249	KPADWVD	266	KPDDWDE	285	KPEDWLE	304	KPEDWDD
A.thaliana calnexin	230	KPEDWDE	247	KPEDWDE	266	KPEGWLD	285	KPEDWDD
Canine calnexin	284	KPEDWDE	301	KPDDWNE	320	KPDGWLD	339	KPEDWDE
S.cerevisiae Cnelp	255	KPHDWDD	272	KLSDRDE	291	EPPEWNS	310	KPSWWKE
Mouse calreticulin	215	KPEDWDE	232	KPEDWDK			249	KPEDWDE

# В

.

132 CGGAYLKLL	156 IMFGPDK <b>C</b> G	324 PKCIEGAGCG
108 CGGAYLKYL	136 IMFGPDKCG	305 PKCEAAPGCG
161 <b>C</b> GGAYVKLL	188 IMFGPDKCG	358 PKCESAPGCG
124 CGGAFIKLM	154 LVFGPDY <b>C</b> A	330 PLCTAERGCG
105 CGGGYVKLF	130 IMFGPDICG	
	<ul> <li>132 CGGAYLKLL</li> <li>108 CGGAYLKYL</li> <li>161 CGGAYVKLL</li> <li>124 CGGAFIKLM</li> <li>105 CGGGYVKLF</li> </ul>	132CGGAYLKLL156IMFGPDKCG108CGGAYLKYL136IMFGPDKCG161CGGAYVKLL188IMFGPDKCG124CGGAFIKLM154LVFGPDYCA105CGGGYVKLF130IMFGPDICG

remarkably conserved amongst all calnexins including Cne1p, two of which are also conserved in calreticulin (Figure 24B).

The diagnostic difference between calnexins and calreticulins is that the former are type I ER membrane proteins, while calreticulins are ER lumenal proteins with carboxy terminal KDEL retrieval signals (Pelham, 1990; Michalak *et al.*, 1992). For *S.cerevisiae* we have shown that Cne1p is an integral membrane glycoprotein and is localized to the ER membrane (Parlati *et al.*, 1995b). We have also shown that *S.pombe* Cnx1p is an integral membrane glycoprotein. Although we have not yet demonstrated its ER localization, the tight banding of the protein on SDS-PAGE is consistent with ER glycosylation. Hence we expect Cnx1p to be localized to the ER. There is no obvious ER retention motif present in the Cnx1p cytosolic tail and the ER retention/retrieval motif for mammalian calnexin (an extreme carboxy terminal RKPRRE motif), is not present in Cnx1p (Rajagopalan *et al.*, 1994).

S.pombe  $cnx1^+$  is inducible by heat shock and by treatment with the calcium ionophore A23187. This latter feature has not been described for mammalian calnexin (Bergeron *et al.*, 1994) although it is well known for other ER chaperones such as BiP, GRP 94 and PDI both in mammalian cells and yeasts (Normington *et al.*, 1989; Rose *et al.*, 1989; Mori *et al.*, 1992; Pidoux and Armstrong, 1992). Calreticulin is also induced by heat shock (Conway *et al.*, 1995). The difference in mRNA abundance of  $cnx1^+$  in response to A23187 but not tunicamycin may be consistent with the suggested N-linked glycosylation specificity of calnexin (Ou *et al.*, 1993), that is, inhibition of the oligosaccharide glycosylation precursor is not sensed by calnexin. In addition, this result may suggest that Cnx1p has a role in Ca<sup>2+</sup> regulation or sequestration. A heat shock consensus sequence was also identified in the *S.cerevisiae CNE1* gene, but we were unable to find experimental conditions that altered its transcription (Parlati *et al.*, 1995b).

The essential nature of the *S.pombe*  $cnxl^+$  gene is in contrast to the non-essential nature of the *S.cerevisiae CNE1* gene (Parlati *et al.*, 1995b). In an *S.cerevisiae CNE1* deleted strain there was a small effect observed on the function of a temperature sensitive mutant of the  $\alpha$ -pheromone receptor (*ste2-3ts*) at the non-permissive temperature. Furthermore, an increase in the secretion of heterologously expressed mammalian  $\alpha_1$ -antitrypsin was observed (Parlati *et al.*, 1995b). To explain these small effects we speculated that there may be other systems for protein folding in the ER of *S.cerevisiae*. Despite a search by low stringency hybridization in both the *S.cerevisiae* and *S.pombe* 

genomes, we have not been able to demonstrate the presence of related calnexins or calreticulins (Parlati *et al.*, 1995b).

Complementation of the lethal cnx1::ura4+ disruption strain has shown that the Cterminal part of the molecule is not essential for the viability. Indeed deletion of the transmembrane as well as the predicted cytosolic domain led to a soluble truncated form of Cnx1p that was secreted but still complemented. We were able, however, to detect truncated Cnx1p intracellularly. Our previous studies have pointed to the lumenal domain of calnexin as important for its function as a molecular chaperone in mammalian cells (Ou et al., 1993), and we speculate that this function is carried out by the ER lumenal soluble Cnx1 protein while it is in the ER. A similar result has also been observed in *S. cerevisiae* for the ER lumenal molecular chaperone BiP. When the carboxy terminal tetra peptide ER retrieval motif, HDEL, is deleted from the essential gene BiP, the protein is secreted. However, a small amount of BiP remains intracellular and the cells are viable (Hardwick et al., 1990). By analogy, we propose that sufficient levels of truncated soluble Cnx1p remain in the ER in order to perform the essential function. We are unable to definitively prove that the lethal phenotype of the cnx1:: $ura4^+$  disruption is due to the lack of molecular chaperone function. However, studies with the lumenal domain of the mammalian calnexin have shown that it can interact with secretory glycoproteins (Ou *et al.*, unpublished results). Thus, it is likely that this domain has molecular chaperone properties but it may also have other functions.

The lack of complementation of the  $cnx1::ura4^+$  disruptant by mammalian (canine) calnexin is surprising. We have shown that the protein is made in *S.pombe* and is membrane localized. The ER lumenal region of  $cnx1^+$  (amino acids 1 to 474) that can complement the  $cnx1::ura4^+$  disruption has 50% sequence similarity with the equivalent region in mammalian calnexin. Our observation that the lumenal domain of Cnx1p can still complement the disruption even though it is secreted argues that if mammalian calnexin is not ER retained, enough would be present to perform its ER function. Although we expect that Cnx1p and mammalian calnexin can recognize the same secretory glycoprotein substrates, sequence differences between these proteins are of value in order to map the regions essential for molecular chaperone function.

# **CHAPTER 4**

Rescue of the secretion defective phenotype of mutant  $\alpha_1$ -antitrypsin by C-terminal deletions of calnexin in Schizosaccharomyces pombe.

## 4.1 Abstract

An inheritable predisposition towards pulmonary emphysema and in some cases, liver cirrhosis is associated with the inefficient secretion of the  $\alpha_1$ -antitrypsin PI Z variant. The role of the ER molecular chaperone calnexin in the regulation of PI Z secretion was directly assessed in the yeast *Schizosaccharomyces pombe*. The normal PI M1 (Val 213)  $\alpha_1$ -antitrypsin but not the PI Z variant (D342K) was efficiently secreted from wild type *S.pombe*. Secretion of the PI Z variant was restored in *S.pombe* strains which only expressed C-terminally truncated mutants of *S.pombe* calnexin. Thus, the cytosolic tail of calnexin plays an important role in mediating the interaction between the PI Z variant and the ER protein 'quality control' apparatus.

. .

### 4.2 Introduction

Several human genetic diseases termed 'protein trafficking diseases' result from the impaired intracellular transport of mutant proteins (Amara *et al.*, 1992; Thomas *et al.*, 1995). These diseases are related to mutant protein recognition and retention by molecular chaperones (Amara *et al.*, 1992; Thomas *et al.*, 1995; Hammond and Helenius, 1995). In humans, one such 'protein trafficking disease' results from homozygous expression of the  $\alpha_1$ -antitrypsin PI Z variant in hepatocytes (Cox, 1989). In affected patients, only 15-20% of normal  $\alpha_1$ -antitrypsin levels are present in serum. The low serum levels of this protease inhibitor results in elastolytic destruction of lung alveoli thereby causing emphysema (Cox, 1989). ER accumulation of the PI Z variant in hepatocytes is the basis of the pathogenesis of liver disease in affected children (Sharp *et al.*, 1969). The PI Z variant shows prolonged association with calnexin, is poorly secreted from mammalian cells and is subjected to proteolytic degradation in a pre-Golgi secretory compartment (Graham *et al.*, 1990; Le *et al.*, 1990; Sifers *et al.*, 1992b; Ou *et al.*, 1993; Le *et al.*, 1994; Wu *et al.*, 1994).

Calnexin is a type I ER transmembrane protein expressed in all eukaryotic cells (Bergeron *et al.*, 1994). In mammalian cells, calnexin interacts transiently with N-linked glycoproteins during normal protein maturation in the ER (Ou *et al.*, 1993; Hammond *et al.*, 1994; Hebert *et al.*, 1995). This molecular chaperone function is a consequence of lumenal interactions between calnexin and the GlcNAc<sub>2</sub>Man<sub>9</sub>Glc<sub>1</sub> sugar intermediate on substrate proteins (Hebert *et al.*, 1995; Ware *et al.*, 1995). There is evidence that calnexin is also a constituent of the ER quality control apparatus. Calnexin associates for a prolonged time with misfolded or unassembled secretory proteins and is likely to be responsible for their ER retention (Hammond and Helenius, 1995; Williams and Watts, 1995).

The related yeast calnexin homologue *CNE1* also has an effect on ER quality control (Parlati *et al.*, 1995b; McCracken and Brodsky, 1996). In *S.cerevisiae*, Cne1p is partially responsible for the intracellular retention of  $\alpha_1$ -antitrypsin, but *CNE1* is not essential for viability (Parlati *et al.*, 1995b). In comparison with *S.cerevisiae CNE1*, *Schizosaccharomyces pombe* calnexin (*cnx1*+) is more closely related to mammalian calnexin. Unlike *S.cerevisiae* Cne1p, *S.pombe* Cnx1p is a high affinity calcium binding protein and encodes a cytosolic tail. In *S.pombe*, the *cnx1*+ gene is essential for viability, which may signify a more important role for calnexin in ER quality control (Jannatipour and Rokeach, 1995; Parlati *et al.*, 1995a). All calnexins except for Cne1p code for a C-terminal cytosolic tail (Parlati *et al.*, 1995b). The cytosolic tail is responsible for ER retention of calnexin, and previous studies suggest that the calnexin cytosolic tail contributes to ER retention of unassembled protein subunits (Rajagopalan *et al.*, 1994). Using human  $\alpha_1$ -antitrypsin and its PI Z variant, we have determined the contribution of the cytosolic tail of *S.pombe* calnexin in ER quality control. We demonstrate that the normally retained PI Z variant is specifically and efficiently secreted from *S.pombe* strains which do not express the calnexin cytosolic tail.

## 4.3 Materials and Methods

#### **Plasmid constructs**

cDNA encoding the wild type  $\alpha_1$ -antitrypsin PI M1 (Val 213) or the PI Z variant were obtained by digesting plasmid phAT85 and phAz respectively (Sifers *et al.*, 1989) with *Eco* RI and creating blunt ends with T4 DNA polymerase. These cDNAs were then subcloned into the *Sma* I site of plasmid pREP5 (Arkinstall *et al.*, 1995) to create pFP $\alpha$ i or pFP $\alpha$ z (see Table 7). pREP1 based plasmid constructs pCNX560, pCNX524 and pCNX484 have been previously described and are summarized in Table 7 (Parlati *et al.*, 1995a).

#### Yeast strains and transformations

All yeast strains were S.pombe. Wild type S.pombe  $cnx1^+$  strain Q 360 ( $h^+$  leu1-32 ura4-D18 ade6-M216) was transformed with the LEU2 based vectors pREP1, pCNX560, pCNX524 or pCNX484 (Maundrell, 1993; Parlati *et al.*, 1995a). S.pombe  $\Delta cnx1$ pCNX560, S.pombe  $\Delta cnx1$  pCNX524, and S.pombe  $\Delta cnx1$  pCNX484 no longer harboring the genomic copy of  $cnx1^+$ , were rescued from death with a plasmid copy of  $cnx1^+$  (see Table 7). These strains were subsequently transformed with plasmid pREP5, pFP $\alpha$ i or pFP $\alpha$ z. Strains were grown in Edinburgh minimal media (EMM) plus supplements, and standard LiCl transformation procedures were used (Moreno *et al.*, 1991).

#### Immunodetection of secreted $\alpha_1$ -antitrypsin, BiP and Cnx1p

For detection of secreted  $\alpha_1$ -antitrypsin, cells were grown to an OD<sub>600</sub> of 1. Cells were harvested, centrifuged and resuspended in 0.1 volumes of EMM. 3 µl of cell suspension was then applied to an EMM agar plate and overlaid with nitrocellulose. After an overnight incubation, the nitrocellulose was washed and immunoblotted with  $\alpha_1$ -antitrypsin antiserum and the immunoreaction was visualized by chemiluminescence (Amersham) as described previously (Parlati *et al.*, 1995b).

For detection of secreted BiP and Cnx1p, cells were grown overnight in EMM, washed and diluted to an  $OD_{600}$  of O.5 in 200 ml of EMM, and then grown to an  $OD_{600}$  of 1. Cells were removed from media by centrifugation. The pH of the cell free media was adjusted to pH 6.2 and protease inhibitors were added to final concentration of 1 mM

PMSF, 1 µg/mł pepstatin, 1 µg/ml E-64, 1 µg/ml leupeptin and 1 µg/ml aprotinin). Media was concentrated using an Amicon Filtration Unit through a YM30 filter (Amicon). The filter was subsequently washed twice with 20 ml of MBS (10 mM MES 150 mM NaCl, pH 6.2) and proteins were finally resuspended into 350 µl of MBS. 40 µl of concentrated media proteins were electrophoresed and immunoblotted as described previously with Cnx1p (1:10,000 dilution) and BiP (1:10,000 dilution) antiserum (Pidoux and Armstrong, 1993; Parlati *et al.*, 1995a).

#### Detection of intracellular $\alpha_1$ -antitrypsin and Cnx1p

Strains were grown in EMM to an  $OD_{600}$  of 1. Cells (5 OD<sub>600</sub>) were collected by centrifugation, washed and resuspended in 0.5 ml of fresh EMM and labeled with 150 µCi of <sup>35</sup>S methionine (Express, NEN) at 30°C. After 30 minutes, cells were harvested, washed, lysed and immunoprecipitated with  $\alpha_1$ -antitrypsin antiserum (1:1000 dilution) or Cnx1p (1:2000 dilution), essentially as previously described (Franzusoff *et al.*, 1991). Briefly, after <sup>35</sup>S radiolabeling, cells were washed in TBS (10 mM Tris 150 mM NaCl, pH 7.5) and lysed in lysis buffer (TBS plus 1% Triton, 0.1% SDS, 0.1 % BSA + inhibitors) using acid washed glass beads (Sigma). Cell lysates were incubated with  $\alpha_1$ -antitrypsin or Cnx1p antiserum followed by incubation with Protein A Sepharose (Bio-Rad). Immunoprecipitates were finally washed 3 times with washing buffer (TBS + 0.5 % Triton X-100, 0.05 % SDS) and washed once with TBS. Samples were boiled in Laemmli buffer, electrophoresed on SDS-PAGE and subjected to fluorography with Amplify (Amersham).

Intracellular Cnx lp levels were determined by immunoblotting. Total membrane fractions were isolated as previously described (Parlati *et al.*, 1995a). 40  $\mu$ g of protein was electrophoresed on SDS-PAGE, transferred to nitrocellulose and immunoblotted with anti-Cnx lp antibodies also as previously described (Parlati *et al.*, 1995a).

Plasmid	Parent Plasmid	Insert	Amino Acids	
pREP1	pREP1			
pCNX 560	pREP1	cnx1+	1-560	
pCNX 524	pREP1	cnx1+	1-524	
pCNX 484	pREP1	cnx1+	1-484	
pREP5	pREP5			
pFPαi	pREP5	α <sub>1</sub> -antitrypsin Pl M (Val213)		
pFPaz	pREP5	α <sub>1</sub> -antitrypsin PI Z variant		

Table 7: Description of constructs used in chapter 4

•

### 4.4 Results

#### Deletion of Cnx1p cytosolic tail leads to PI Z secretion

We first confirmed that the *S.pombe* ER quality control apparatus can distinguish between the wild type  $\alpha_1$ -antitrypsin PI M1 (Val 213) variant and the PI Z variant (Brantly et al., 1988). Wild type human  $\alpha_1$ -antitrypsin PI M1(Val 213) variant (encoded by plasmid pFP $\alpha$ i) was efficiently secreted from both wild type S.pombe cnx1+ (Fig. 25A, lane 1) and from a S. pombe  $\Delta cnxI$  deletion strain expressing a plasmid copy of full-length CnxIp (see Table 7), termed S.pombe  $\Delta cnx1$  pCNX560 (Fig 25A, lane 2). The PI Z variant (encoded by pFP $\alpha z$ ) was poorly secreted from both of these strains (Fig. 25A, lanes 1, 2), confirming that the *S. pombe* quality control apparatus recognizes and retains this mutant protein. S. pombe strains which no longer express the Cnx1p cytosolic tail (termed S. pombe  $\Delta cnxI$  pCNX524) or the Cnx1p cytosolic tail and adjacent transmembrane domain (termed S.pombe Acnx1 pCNX484) are viable (Parlati et al., 1995a). However, we reasoned that presentation of misfolded proteins to the ER quality control apparatus may be impeded in these strains. This was confirmed in strains which expressed either Cnx1p truncated mutant in the absence of full length Cnx1p (Fig. 25A, lanes 3, 4). Levels of the secreted PI Z variant were equivalent to levels of the secreted PI M1 (Val 213) variant in S.pombe  $\Delta cnx1$  pCNX524 or S.pombe  $\Delta cnx1$  pCNX484 (Fig. 25A, lanes 3, 4).

We next determined if secretion of the PI Z variant was a consequence of the Cnx1p mutants associating with the PI Z variant and 'dragging' the PI Z variant through the secretory pathway. This possibility was tested by measuring the secretion of both the PI M1 (Val 213) and PI Z variants in strains coexpressing full length and either Cnx1p truncated mutants. Expression of a plasmid copy of full length Cnx1p in wild type *S.pombe* did not result in secretion rescue of the PI Z variant (Fig. 25B, lane 1). Likewise, truncated Cnx1p mutants were not able to rescue the secretion of the PI Z variant when full length Cnx1p was also expressed in cells (Fig. 25B, lanes 2, 3). We therefore consider it unlikely that secretion rescue of the PI Z variant is a direct consequence of its association with truncated Cnx1p, thereby resulting in its co-secretion with Cnx1p. Rather, secretion of the PI Z variant is likely a consequence of a defect in the quality control apparatus when the cytosolic tail of Cnx1p is deleted.

# Figure 25: Rescued secretion of the PI Z variant expressed in S. pombe cells.

 $\alpha_1$ -antitrypsin secretion levels were determined using an overlay assay for secreted proteins. Panel A: Secretion levels of wild type human  $\alpha_1$ -antitrypsin PI M1 (Val 213) (pFP $\alpha$ i) or the PI Z variant (pFP $\alpha$ z) were evaluated in wild type *S.pombe cnx1*+ transformed with control plasmid pREP1 (lane 1), *S.pombe \Delta cnx1* pCNX560 (lane 2), *S.pombe \Delta cnx1* pCNX524 (lanes 3) or *S.pombe \Delta cnx1* pCNX484 (lanes 4). Strains were also transformed with vector alone (pREP5) in order to monitor non-specific background. Panel B: Wild type *S.pombe cnx1*+ cells were transformed with pCNX560 (lane 1), pCNX524 (lane 2), or pCNX484 (lane 3). Secretion of  $\alpha_1$ -antitrypsin was assessed after transformation with pREP5, pFP $\alpha$ i or pFP $\alpha$ z. Identical results were found for two independent sets of transformants.







# Deletion of the Cnx1p cytosolic tail and not intracellular levels of $\alpha_1$ -antitrypsin or Cnx1p accounts for PI Z secretion.

We confirmed that the observed secreted levels of the PIZ variant were specifically due to the absence of the Cnx1p cytosolic tail (Fig. 25) rather than differences in  $\alpha_1$ antitrypsin expression levels (Fig. 26A). Intracellular levels of the PI M1 (Val 213) and PI Z variants were monitored in all four  $cnxl^+$  backgrounds by metabolically radiolabeling cells followed by immunoprecipitating  $\alpha_1$ -antitrypsin from cell lysates. We detected slightly less PI Z variant as compared to PI M1(Val 213) variant in wild type S.pombe cnx1<sup>+</sup> (Fig. 26A, lanes 1, 2). The 51 kDa polypeptide was the precursor to the other Nlinked glycosylated forms (see legend to Fig. 26A). As in wild type S. pombe  $cnxl^+$  (Fig. 26A, lanes 1, 2), slightly less PI Z variant was expressed compared with the PI MI (Val 213) normal variant in S.pombe Δcnx1 pCNX524 and S.pombe Δcnx1 pCNX484 (Fig. 26A, lanes 5-8). Although the PIZ variant was efficiently expressed intracellularly in all three strains (Fig. 26A, lanes 2, 6, 8), secretion of the PI Z variant was only occurred in S.pombe expressing truncated Cnx1p mutants (compare Fig. 25A, lanes 3 (or 4) to lane 1). Furthermore, differences in mRNA levels did not account for differences observed for  $\alpha_1$ antitrypsin PI Z levels (results not shown). The data remain consistent with a correlation between the presence of the cytosolic tail of Cnxlp and intracellular retention of the PIZ variant. Finally, in S. pombe Acnx1 pCNX560, the expression levels of wild type PI M1 (Val 213)  $\alpha_1$ -antitrypsin were slightly reduced (Fig 26A, compare lanes 1 to 3), whereas virtually no PIZ variant was detected intracellularly in this strain (Fig 26A, lane 4). The lower intracellular amounts of the PIZ variant were likely due to more efficient targeting of the PI Z variant to the ER protein degradation apparatus when full length Cnx1p was overexpressed (see Figs. 26B and C, lane 2).

In order to determine if intracellular stability and eventual secretion of the PI Z variant was due to Cnx1p levels, Cnx1p biosynthetic and steady state levels were assessed. The amounts of Cnx1p were measured by pulse labeling cells and immunoprecipitating Cnx1p from cell lysates. The levels of all the plasmid encoded forms of Cnx1p in *S.pombe*  $\Delta cnx1$  were greater than those of endogenous Cnx1p in wild type *S.pombe* (Fig. 26B, compare all lanes). Cnx1p steady state intracellular levels corresponded to Cnx1p biosynthetic levels except in *S.pombe*  $\Delta cnx1$  pCNX484 (Fig. 26C). In this case, lower steady state levels of this Cnx1p truncated mutant were observed (compare Figs. 26B and C, lane 4) due to its secretion into the extracellular medium (Fig. 27B, lane 4). In *S.pombe*  $\Delta cnx1$  pCNX524, Cnx1p biosynthetic and steady state levels were similar to

Cnx1p levels in *S.pombe*  $\Delta cnx1$  pCNX560 (Figs. 26B and C, lanes 2, 3). In the former but not the latter strain, secretion rescue of the PIZ variant was effected (Fig 25A, compare lane 3 to lane 2). Hence, Cnx1p biosynthetic or steady state levels can not simply account for the secretion rescue of the PIZ variant in strains which express truncated Cnx1p mutants.

Regardless of its intracellular levels, expression of full length Cnx1p (Figs. 26B and C, compare lanes 1 and 2) did not result in the secretion of the PI Z variant (Fig. 25A, lanes 1, 2). However the increased expression of full length Cnx1p correlated with enhanced instability of the PI Z variant, presumably by increasing its access to the ER associated protein degradation machinery (Fig. 26A, lane 4).

Figure 26: Intracellular expression of the wild type  $\alpha_1$ -antitrypsin PI M1 (Val 213) variant, the PI Z variant and Cnx1p in all four  $cnx1^+$  backgrounds.

Panel A: Intracellular expression levels of the PI M1 (Val 213) variant, abbreviated  $\alpha i$  (lanes 1, 3, 5, 7) and the PI Z variant, abbreviated  $\alpha z$  (lanes 2, 4, 6, 8) were determined in wild type *S.pombe cnx1*<sup>+</sup> transformed with pREP 1 (lanes 1, 2), *S.pombe \Delta cnx1* pCNX560 (lane 3, 4), *S.pombe \Delta cnx1* pCNX524 (lane 5, 6) or *S.pombe \Delta cnx1* pCNX484 (lanes 7, 8). Polypeptides with mobilities which correspond to 51 kDa, 53 kDa, 56 kDa and 59 kDa were detected. The arrow indicates the 51 kDa form of  $\alpha_1$ -antitrypsin. The results of pulse-chase studies indicated that the 51 kDa protein was a precursor of the other forms which are N-linked glycosylation intermediates. These higher molecular weight forms are absent in cells which had been treated with the N-linked glycosylation inhibitor tunicamycin (results not shown).

Panel B: Biosynthetic levels of Cnx1p in wild type *S.pombe cnx1*<sup>+</sup> transformed with pREP1 (lane 1) or *S.pombe \Delta cnx1* pCNX560 (lane 2), *S.pombe \Delta cnx1* pCNX524 (lane 3), *S.pombe \Delta cnx1* pCNX484 (lane 4) were evaluated. The size of the radiolabeled protein corresponds to that expected of the full length (91 kDa, lanes 1, 2) or the truncated Cnx1p mutants (85 kDa, lanes 3; 80 kDa lane 4). By densitometry (arbitrary units), the ratio of expressed Cnx1p was determined as a function of that expressed in wild type cells (lane 1).

Panel C: Steady state levels of Cnx1p were determined in wild type *S.pombe*  $cnx1^+$  transformed with pREP1 (lane 1), *S.pombe*  $\Delta cnx1$  pCNX560 (lane 2), *S.pombe*  $\Delta cnx1$  pCNX524 (lane 3) and *S.pombe*  $\Delta cnx1$  pCNX484 (lane 4). As in Panel B, Cnx1p levels are indicated in arbitrary units (bottom row).





# Deletion of the Cnx1p cytosolic tail does not lead to perturbations of ER lumenal contents.

We determined if a non-specific mechanism, such as the bulk secretion of ER lumenal contents, was responsible for PI Z variant secretion in the Cnx1p C-terminal deletion mutant strains. Mislocalized Cnx1p truncation mutants might conceivably result in the perturbation of the ER lumenal matrix, thereby causing a non specific secretion of ER lumenal resident proteins (Sambrook, 1990). Proteolytic fragments of mislocalized Cnx1p were detected in culture media from *S.pombe*  $\Delta cnx1$  pCNX484 (Fig. 27B, lane 4). In *S.pombe*  $\Delta cnx1$  pCNX524, low levels of proteolytic fragments of Cnx1p were also found in the media, which were likely due to mislocalization of the truncated form of Cnx1p (Fig 27B, lanes 3).

We determined if BiP (Pidoux and Armstrong, 1992) was secreted from *S.pombe*  $\Delta cnx1$  pCNX524 and *S.pombe*  $\Delta cnx1$  pCNX484 strains. Although calnexin was mislocalized in these strains (Fig 27B, lanes 3, 4), non-specific secretion of ER lumenal contents was not detected. No extracellular BiP was detectable in culture media from wild type *S.pombe*  $cnx1^+$  (Fig. 27A, lane 1) or from *S.pombe*  $\Delta cnx1$  pCNX524 and  $\Delta cnx1$ pCNX484 (Fig. 3A, lanes 3, 4). Therefore, this situation does lead to PI Z variant secretion (Fig. 25A, lanes 3, 4) without causing non-specific secretion of ER lumenal proteins (Fig. 27A, lanes 3, 4).

We detected BiP secretion from *S.pombe*  $\Delta cnx1$  pCNX560, which overexpresses full length Cnx1p (Fig. 27A, lane 2). However, intracellular levels of BiP in this strain were not affected (data not shown). Despite the loss of BiP from these cells, secretion rescue of the PI Z variant was not observed (Fig. 25A, lane 2).

# s ----- Si

### Figure 27: Evaluation of resident ER lumenal protein BiP secretion.

BiP secretion (Panel A) or Cnx1p secretion (Panel B) levels were determined in *S.pombe*  $cnx1^+$  (lane 1), *S.pombe*  $\Delta cnx1^-$  pCNX560 (lane 2), *S.pombe*  $\Delta cnx1^-$  pCNX524 (lane 3) or *S.pombe*  $\Delta cnx1^-$  pCNX484 (lane 4).



### 4.5 Discussion

A mutation (D342 --> K) in the PI Z  $\alpha_1$ -antitrypsin variant prevents formation of a salt bridge which is predicted to constrain the mobility of the reactive center loop (Loebermann *et al.*, 1984; Stein and Carrell, 1995). Loss of this structural constraint is thought to initiate the 'loop-sheet aggregation' of the PI Z variant (Lomas *et al.*, 1992; Sifers, 1992a). Although this may function as the actual mechanism resulting in intrahepatic accumulation of the PI Z variant, the mechanism of the initial retention of this variant is unresolved (Sifers, 1995). An arrest in the folding of the PI Z variant has been observed and its association with calnexin has been shown. This interaction with calnexin may be responsible for the ER retention and storage and subsequent degradation of the PI Z variant (Le *et al.*, 1990; Le *et al.*, 1992; Ou *et al.*, 1993; Le *et al.*, 1994). Here, we have shown that truncations in calnexin result in efficient secretion of this otherwise intracellularly retained mutant protein.

We have shown that the ER quality control apparatus in *S.pombe* can distinguish between the mutant PI Z variant and wild type PI M1 (Val 213) human  $\alpha_1$ -antitrypsin (Fig. 28A). Using a COS cell expression system, ER retention of unassembled and normally ER retained human CD3e chain was effected in the presence of human calnexin (Rajagopalan et al., 1994). Deletion of the cytosolic tail of human calnexin in this system caused cell surface expression of the human CD3e chain. In a second study using a Drosophila expression system, heterologously expressed MHC I heavy chains were ER retained only when mammalian calnexin was also coexpressed (Jackson et al., 1994). Hence, the endogenous calnexin in either COS cells or *Drosophila* cells could not retain unassembled protein subunits in the ER. In most mammalian systems, the endogenous calnexin will ER retain unfolded or unassembled proteins (Bergeron et al., 1994; Hammond and Helenius, 1995). Unlike endogenous COS or *Drosophila* calnexin, *S. pombe* calnexin acts as a *bona fide* constituent of the ER quality control apparatus. Thus caluexin's role in the secretory defect associated with a 'protein trafficking disease' was directly affected by genetic manipulation of endogenous S.pombe calnexin. Deletion of the cytosolic tail of S.pombe calnexin was sufficient to rescue the secretion defective PI Z variant without causing a general perturbation of ER lumenal contents (Fig 28C). However, secretion rescue of the PI Z variant was not observed when full length and C-terminal truncated Cnx1p were coexpressed (Fig 28D). We have also observed the same behavior when the PI Z variant and truncated forms of calnexin were transfected into human 293A cells (Hemming et al., in preparation). Hence, there may be different criteria for the retention of misfolded protein
such as the PI Z variant as compared to unassembled subunits of multimeric complexes (Rajagopalan *et al.*, 1994; Jackson *et al.*, 1994). Nevertheless, our data is consistent with a model in which the cytosolic tail mediated localization of Cnx1p is responsible for presentation of the PI Z variant to the ER quality control apparatus (Fig. 28).

When full length Cnx1p was overexpressed, the intracellular instability of the PI Z variant was increased. This supports the hypothesis that the Cnx1p (via its tail) presents misfolded proteins to the ER quality control apparatus (Fig 28B). Interestingly, BiP secretion is also observed in this strain. Since calnexin is a high affinity calcium-binding protein (Wada *et al.*, 1991; Tjoelker *et al.*, 1994; Parlati *et al.*, 1995a) elevated levels of calnexin in the ER may perturb the free calcium levels leading to the release of BiP from the ER lumen and its secretion from the cell (Sambrook, 1990; Wada *et al.*, 1991). Alternatively, overexpression of full length Cnx1p may lead to an upregulation of BiP causing its secretion from cells. Nevertheless, overexpression of full length Cnx1p does not lead to secretion rescue of the PI Z variant (Fig. 28B)

We have previously shown that the absence of calnexin in *S.cerevisiae* leads to a two to three fold increase in cell surface expression of  $\alpha_1$ -antitrypsin (Parlati *et al.*, 1995b). Recent results from McCracken and Brodsky show that Cnelp is in part responsible for efficient degradation of a mutant secretory protein in yeast (McCracken and Brodsky, 1996). In this respect, our results in *S.pombe* are consistent with observations made in S.cerevisiae. However, there are three significant differences between observations made in this study using *S.pombe* and previous studies using *S.cerevisiae*. First, secretion of the wild type PI M1 (Val 213) variant is not significantly affected by manipulation of S.pombe cnx1<sup>+</sup>. Since the PI M1 (Val 213) variant is normally folded and secreted in mammalian cells, it should not be a substrate for the quality control apparatus and therefore unaffected by manipulating S.pombe cnx1+. Whereas in S.cerevisiae, secretion of the PI M1 (Val 213) and PI Z  $\alpha_1$ -antitrypsin variants were upregulated to the same extent when CNE1 was deleted. Second, PI Z variant secretion was restored to wild type levels when the cytosolic tail of S.pombe Cnx1p was deleted. Deletion of the entire CNEI gene in S.cerevisiae did not lead to PI Z variant secretion levels equivalent to levels observed for the PI M1 (Val 213) variant. Third, deletion of the cytosolic tail alone was sufficient to effect secretion rescue of the PIZ variant in S. pombe. This phenotype would be predicted since the cytosolic tail of mammalian calnexin mediates calnexin ER retention. Therefore calnexin ER retention would be required for its interaction with other components of the ER quality control apparatus. In S. cerevisiae, deletion of the entire CNE1 gene resulted only in a small increase in  $\alpha_1$ -antitrypsin secretion. Hence, calnexin mediated quality control in *S.pombe* is more prominent than in *S.cerevisiae*.

**Figure 28:** Quality control regulation by Cnx1p. Panel A: Presentation of the PI Z variant to the ER protein degradation machinery by full length Cnx1p. Panel B: Overexpression of full length Cnx1p and efficient presentation of the PI Z variant to the ER protein degradation machinery. Overexpression of full length Cnx1p leads to BiP secretion. Panel C: Expression of truncated Cnx1p mutants in *S.pombe*  $\Delta cnx1$  leads to the sequestration of the PI Z variant from the ER protein degradation machinery and its subsequent secretion. Panel D: Full length Cnx1p is dominant over Cnx1p truncation mutants. Coexpression of both full length Cnx1p and either Cnx1p truncation mutant leads to presentation of the PI Z variant to the ER protein degradation machinery.











# **CHAPTER 5**

Conclusions

The study of mammalian calnexin and calreticulin has been dynamic and the function of mammalian calnexin has become clearer during the progress of this work. At the start of this work, little was known about the role of calnexin in the ER except for its physical properties as an ER integral membrane calcium binding phosphoprotein which was found in a complex with three other glycoproteins, pgp 35, gp25H (at one time considered subunits of a putative signal sequence receptor) and gp25L (Wada *et al.*, 1991). Yeasts are genetically manipulable and the function of a gene can be directly determined by gene disruption and the expression of mutants. The *S.cerevisiae* and *S.pombe* homologues of calnexin were identified and cloned in order to shed light on the role of calnexin. Some observations made during this study with yeast calnexins are not consistent with observations of mammalian calnexin. In this chapter, I will discuss observations made in this study and their broader relevance to the calnexin/calreticulin chaperone field. Many questions remain unanswered as to the function of yeast calnexin in quality control and I will suggest experiments in order to extend the observations made during this study.

## 5.1 Identity of S.cerevisiae CNE1

5

Whether *CNE1* encoded a *bona fide* calnexin homologue remained a contentious issue for some time. Although Cne1p was an ER membrane protein and shares a significant level of overall amino acid identity with mammalian calnexin (including the conserved cysteine residues), many other amino acid motifs were not present. Unlike its mammalian counterpart, only one of four putative lumenal calcium binding repeats was present in Cne1p and Cne1p did not bind calcium *in vitro* using a <sup>45</sup>Ca<sup>2+</sup> overlay assay. We can not rule out that Cne1p does in fact bind calcium *in vivo* at ER calcium concentration levels [i.e. millimolar concentrations] (Sambrook, 1990). However, unlike its mammalian and *S.pombe* counterparts, Cne1p is not a strong calcium binding protein since it does not bind calcium at micromolar concentrations.

Cnelp does not have a cytoplasmic tail, and thus does not have the cytosolic low affinity calcium binding site, the cytosolic phosphorylation sites or an ER retention/retrieval sequence like its mammalian counterparts. How Cnelp remains located to the ER is enigmatic. Many type I ER membrane proteins have the sequence motif KKXX at their Cterminus, which confers an ER localization (Pelham *et al.*, 1995). Cnelp does not have a cytosolic tail, and thus does not contain any classical retrieval motifs. Deletions of the transmembrane domain of Cnelp did not produce a secreted form of Cnelp, as observed for *S.pombe* and mammalian calnexin (my unpublished results). In fact, truncated Cne1p remained membrane-associated. It is possible that Cne1p is complexed with another ER membrane protein which thereby confers ER localization to Cne1p.

Species often encode a number of closely related genes, which have related and sometimes overlapping functions in the cells. For example, there are at least ten HSP 70 homologues (Craig, 1989; Rassow *et al.*, 1995) and four PDI homologues in *S.cerevisiae* (Bardwell and Beckwith, 1993; Tachikawa *et al.*, 1995), some of which have overlapping functions. In some cases, deletion of one homologue does not lead to an obvious phenotype since related proteins can functionally replace the missing protein. Calnexin is related to calreticulin in mammals, thus it is reasonable to expect that *S.cerevisiae* codes for a calreticulin homologue (Wada *et al.*, 1991). Furthermore, since *CNE1* is distantly related to calnexin and deletion of *CNE1* resulted in a weak phenotype, we expected that *S.cerevisiae* might have a gene with closer homology to mammalian calnexin. Recently, the entire genome of *S.cerevisiae* has been sequenced but no other calnexin or calreticulin homologue has been found.

Thus, it is conclusive that Cnelp is the only calnexin or calreticulin homologue in *S.cerevisiae*. The divergence between calnexin in *S.cerevisiae* and mammalian cells is possibly indicative of different physical properties and cellular functions. The 38% amino acid identity between the central domain of both calnexins (see Chapter 2), which is the most conserved domain amongst all calnexins and calreticulins, implies that some functions are likely to be shared.

## 5.2 Identity of S. pombe cnx1+

Sequencing of the *cnx1*<sup>+</sup> gene revealed that it coded for a protein which was indeed a *bona fide* calnexin homologue. It also shared a higher degree of amino acid identity with mammalian calnexin (34%) than *S.cerevisiae* Cne1p (24%). It shared mammalian calnexin's type I integral membrane topology and possesses most of the characteristic calnexin motifs, including four lumenal calcium binding repeats, four conserved cysteine residues and a cytosolic tail containing casein kinase II phosphorylation sites. The Cnx1 protein is indeed a calcium binding membrane protein and its cytosolic tail can be phosphorylated *in vitro* (my unpublished observations). The main difference between mammalian and *S.pombe* calnexin is the absence of the low affinity, high capacity calcium binding sequences (aspartic/glutamic acid stretch) in the cytosolic tail of *S.pombe* calnexin. This low capacity calcium binding domain is also absent from plant calnexins and calreticulins from *O.volvulus* and Schistosomes. Hence, this domain may have been acquired in some species to fulfill a specialized function.

## 5.3 Phenotype of S.cerevisiae Cne1p

In *S. cerevisiae*, deletion or overexpression of *CNE1* revealed no detectable phenotype. Following the identification of *CNE1*, mammalian calnexin was reported to transiently associate with several glycoproteins. In mammalian cells, immunoprecipitation studies of pulse-labeled cells showed that calnexin-glycoprotein associations were short-lived (Ou *et al.*, 1993). Equivalent experiments in *S. cerevisiae* using anti-Cne1p antibodies, could detect no transient association with cellular proteins (my unpublished observations). To further investigate a possible role for Cne1p in protein secretion, the amount of secreted or cell surface expressed proteins was evaluated in various *CNE1* backgrounds. No differences in the secretion or cell surface expression of the proteins acid phosphatase,  $\alpha$ -factor and Ste2p were observed in cells overexpressing or deleted for *CNE1* as compared with wild type cells. However, a function in ER quality control was found for Cne1p when misfolded and intracellularly retained proteins in *S. cerevisiae* were examined. Increased cell surface expression of intracellularly retained Ste2-3p at non permissive temperatures and increased secretion of heterologously expressed  $\alpha_1$ -antitrypsin was observed (see Chapter 2) in *Δcne1* cells as compared with wild type cells.

The role for Cne1p in quality control was further corroborated by data from experiments showing *CNE1* involvement in unglycosylated pro- $\alpha$ -factor degradation. Using an *in vitro* import assay, nonglycosylated pro- $\alpha$ -factor was imported into yeast microsomes and observed to be proteolytically degraded with a high life of 7.5 minutes, in an ATP and a cytosol dependent manner. Using the same assay with microsomes prepared from a  $\Delta cne1$  strain, the half life of unglycosylated pro- $\alpha$ -factor proteolytic degradation increased to 20 minutes (McCracken and Brodsky, 1996). Although Cne1p is not likely to be a protease, a  $\Delta cne1$  strain shows a defect in the degradation of pro- $\alpha$ -factor. My results, together with results from McCracken and Brodsky, are the first evidence that calnexin promotes intracellular retention and proteolytic degradation of secretory proteins.

## 5.4 Phenotype of S. pombe cnx1+

In stark contrast to *S.cerevisiae CNE1*, *S.pombe cnx1+* is essential for viability. Mammalian calnexin, mammalian calreticulin, *S.cerevisiae CNE1*, and various C-terminal truncations of cnx1+ were tested for rescue of the cnx1::ura4+ lethal phenotype. Mammalian calnexin and calreticulin as well as *CNE1* were unable to rescue the cnx1::ura4+ lethal phenotype, although they were efficiently synthesized in diploid cnx1+/cnx1::ura4+ cells (Chapter 4, and my unpublished results). Likewise, mammalian BiP, which has sequence and functional homology with its yeast counterpart Kar2p, is also unable to complement the lethal phenotype associated with a *kar2* null allele (Normington *et al.*, 1989). However, mammalian BiP was able to complement partially the *kar2-1* mutation (Normington *et al.*, 1989). Slight variations in function may explain the failure of a mammalian gene to complement a disruption of a homologous yeast gene. Truncated Cnx1p mutants deleted for either the i) the cytosolic tail, ii) the cytosolic tail and the transmembrane domain or iii) the cytosolic tail, the transmembrane domain and the ten most C-terminal lumenal amino acids were able to rescue the lethal *cnx1::ura4+* phenotype.

I next tested whether, like mammalian calnexin, Cnx1p transiently associated with secretory proteins. As with *S. cerevisiae* Cnelp, no transient association between calnexin and other cellular proteins was detected directly. To assess whether Cnx1p like S.cerevisiae Cnelp plays a role in quality control, both the wild type PI M1 (Val 213) and the PI Z variant  $\alpha_1$ -antitrypsin were expressed in *S. pombe*. In *S. pombe* cells expressing full length Cnx1p, wild type  $\alpha_1$ -antitrypsin was efficiently secreted whereas the PI Z variant was not. The fates of wild type PI M1 (Val 213) and PI Z variant  $\alpha_1$ -antitrypsin were examined in S.pombe cnx1::ura4+ strains which were complemented with truncated Cnx1p mutants. Wild type and PI Z variant  $\alpha_1$ -antitrypsin were expressed in S.pombe strains  $\Delta cnx1$  pCNX524 (cytosolic tail deleted Cnx1p mutant) or S. pombe  $\Delta cnx1$ pCNX484 (both cytosolic tail and transmembrane domain deleted). In both strains, efficient secretion of both wild type and the PI Z variant  $\alpha_1$ -antitrypsin occurred. However, efficient secretion of PI Z variant was not observed in cells coexpressing both full length and either C-terminal truncation mutants of Cnx1p (see Chapter 4). These experiments suggest that the cytosolic tail is responsible for mediating PI Z intracellular retention. Although the Cnx1p cytosolic tail is not essential for viability, the cytosolic tail is probably important for Cnx1p localization, and perhaps maintains the integrity of the ER quality control apparatus.

Recent studies suggest that the cytoplasm is the site of ER protein degradation (Ward *et al.*, 1995; Jensen *et al.*, 1995; McCracken *et al.*, 1996). Therefore calnexin may play a role in protein export to the cytoplasm, which may explain the phenotype observed with the C-terminal truncated Cnx1p mutants. The Cnx1p C-terminal domain may play a role in anchoring the protein in a specific region of the ER, which permits protein export to the cytosol. For Cne1p, an additional protein may be required for Cne1p's ER localization as well as association with a postulated ER protein export apparatus.

## 5.5 S.cerevisiae and S.pombe calnexin compared

At first glance, the calnexin homologues of *S.cerevisiae* and *S.pombe* appear to be quite divergent. They share only 22% amino acid identity, *S.pombe* cnx1+ encodes a cytosolic tail and is clearly a calcium binding protein. Most significantly, *S.pombe* cnx1+ is essential for cell viability whereas *S.cerevisiae* CNE1 is not an essential gene. Despite these differences, both genes play roles in the quality control of secretory proteins, the only phenotype determined thus far for both genes.

An increase in the secretion of intracellularly retained proteins is likely to be the effect of a decreased degradation of misfolded proteins in *S.cerevisiae*  $\Delta cne1$  cells. Deletion of the *S.pombe* Cnx1p cytosolic tail also leads to enhanced secretion of intracellularly retained misfolded proteins. In this system, the effect is more pronounced. Moreover, overexpression of full length Cnx1p leads to increased instability of a mutant protein. This observation is analogous to the situation in *S.cerevisiae*, where mutant proteins are more unstable in cells expressing wild type levels of Cne1p compared to  $\Delta cne1$  cells. Therefore in *S.pombe*, deletion of the Cnx1p cytosolic tail leads to an equivalent phenotype to a *CNE1* deletion in *S.cerevisiae*.

Since deletion or 'knock out' of calnexin in *S.pombe* and *S.cerevisiae* leads to different phenotypes, there must be fundamental differences in function between these two proteins in their respective host cells. In *S.pombe*, the quality control apparatus can be bypassed without causing cell death. Thus, the essential phenotype of *S.pombe cnx1* + may not be the same as its role in quality control. The other major difference between Cne1p and Cnx1p is their calcium binding abilities. Calcium binding is important in calnexin structure as well as in the ability of mammalian calnexin to bind to secretory proteins (Ou *et al.*, 1993; Ou *et al.*, 1995). However, Cne1p (at best a weak calcium binding protein) is still able to function as a component of the quality control apparatus.

One function of calnexin's high affinity calcium binding ability may be in regulating ER calcium stores. *S.pombe* Cnx1p as well as mammalian calnexin and calreticulin have a highly conserved central domain containing a high affinity (in calreticulin  $K_d$  1.6 µm), low capacity (for calreticulin 1 mol/mol) calcium binding site (Baksh and Michalak, 1991). For calreticulin, it has been postulated that this domain serves as a calcium sensor, by prolonging IP3 induced rise in cytosolic calcium when ER stores of calcium are full and the high capacity calcium binding site is occupied (Camacho and Lechleiter, 1995). When stores are depleted, the high capacity calcium binding site is unoccupied thereby causing the flux of calcium to shift in order to fill ER stores. It is likely that S.pombe calnexin as well as mammalian calnexin have a similar role. If S.pombe calnexin, like S.cerevisiae calnexin is the only calnexin/calreticulin homologue, then deletion of Cnx1p in S.pombe would cause cells to be unresponsive to IP3 mediated ER calcium release, and presumably cause cell death. This hypothesis assumes that inositol trisphosphate signaling system is present and essential in S.pombe, but either not present or not essential in S.cerevisiae. Thus far, there is little evidence of IP<sub>3</sub> mediated release of ER calcium stores in either species (Carpenter and Cantley, 1996).

## 5.6 Yeast calnexin mode of action

From the data presented, the calnexin homologues in the yeasts *S.cerevisiae* and *S.pombe* play a role in protein quality control in the ER. Both *CNE1* and *cnx1+* have been shown to interact genetically with mutant proteins and affect their proteolytic degradation. Thus far, no transient association of yeast calnexin with folding glycoproteins has been observed. Technical problems that account for lack of observed calnexin-glycoprotein physical interaction cannot be ruled out. However, there is recent evidence which may explain the observed lack of transient association between calnexin and glycoproteins in yeast.

In the mammalian calnexin-glycoprotein binding model, calnexin associates with proteins containing the N-linked GlcNAc<sub>2</sub>Man<sub>9</sub>Glc<sub>1</sub> moiety. This moiety is generated either by glucosidase I/II trimming or by reglucosylation of N-linked unfolded GlcNAc<sub>2</sub>Man<sub>7-9</sub> proteins by UGGT (Figure 1). However, some of these components are either not present in yeast or not essential for cell viability. Glucosidase II is present in both yeasts (Herscovics and Orlean, 1993; Ziegler *et al.*, 1994; Trimble and Verostek, 1995) but is not essential for viability in *S.cerevisiae* (unpublished observations). The only UGGT sequence related homologue is Kre5p in *S.cerevisiae*, important in cell wall

synthesis (Meaden *et al.*, 1990). Since the UGGT activity is not present in *S.cerevisiae*, it is unlikely that Kre5p has UGGT activity (Fernandez *et al.*, 1994). In fact, Kre5p is required for (1-6)- $\beta$ -D-glucan synthesis (Meaden *et al.*, 1990). The UGGT activity is present in *S.pombe* and a gene coding for the UGGT, *gpt1*+, has been isolated (Fernandez *et al.*, 1996). Cells deleted for *gpt1*+ are viable and deficient in UGGT activity. Hence, in both species, reglucosylation of secretory glycoproteins and subsequent calnexin interaction is not necessary for cell viability.

Two models may explain the role of glucose trimming, calnexin and UGGT in yeast ER quality control:

#### Model 1

In the first model, calnexin-glycoprotein associations occur in yeasts, as in mammalian cells. In wild type *S.cerevisiae* cells or  $\Delta gpt1$  *S.pombe* cells, substrate glycoproteins for calnexin can only be generated by glucosidase I/II trimming. Glycoproteins containing the GlcNAc<sub>2</sub>Man<sub>9</sub>Glc<sub>1</sub> moiety associate with calnexin to produce a completely folded protein. After calnexin release, the glycoprotein has a GlcNAc<sub>2</sub>Man<sub>9</sub>Glc<sub>0</sub> oligosaccharide and is unable to reassociate with calnexin. In this scenario, the requirement for UGGT is bypassed and cells can generate mature glycoproteins in a one step process.

This model may explain the observed lack of calnexin-glycoprotein association in yeasts. A glycoprotein may have a single chance to bind calnexin, which leads to a completely folded protein. In this model, yeast calnexin-glycoprotein associations may be more transient, thus difficult to detect. In some cases, mammalian calnexin-glycoprotein associations have a half life of up to 240 minutes, thus lending support to this model (Degen and Williams, 1991; Ou *et al.*, 1993; Anderson and Cresswell, 1994; Lenter and Vestweber, 1994). In nutrient rich media, wild type yeasts divide every two to three hours compared to 18-24 hours for many cultured mammalian cell types. It is likely that proteins in mammalian cells have more time to undergo maturation in the ER, as compared with proteins in yeasts. Mammalian calnexin mediated folding involves repetitive cycles of protein binding to calnexin, dissociation from calnexin due to glucose removal and reassociation to calnexin due to reglucosylation (Fig. 1). This folding pathway may not occur in yeast since this process may be inherently slow and inefficient.

This model can be tested using an *in vitro* calnexin-glycoprotein binding assay. Zapun and colleagues have recently shown that recombinant mammalian calnexin can bind ribonuclease with the GlcNAc<sub>2</sub>Man<sub>9</sub>Glc<sub>1</sub> moiety *in vitro* (manuscript in preparation). Likewise, the ability of recombinant Cne1p and Cnx1p to bind ribonuclease in this assay can be assessed. If yeast calnexin has this activity, then closer inspection of *in vivo* calnexin interactions are warranted.

#### Model 2

In this second model, normal folding intermediates of wild type proteins do not associate with yeast calnexin. In yeasts, calnexin may act as a folding sensor for misfolded proteins and is responsible for 'presenting' these proteins to the ER proteolytic degradation machinery.

In *S.pombe*, this process may involve reglucosylation of high mannose oligosaccharides by UGGT. This role for UGGT and Cnx1p is consistent with the fact that UGGT transcription is highly induced by agents which cause protein misfolding (Fernandez *et al.*, 1996). Under stress conditions, elevated levels of UGGT would facilitate reglucosylation of aggregation prone high mannose N-linked proteins. In this scenario, the presence of the GlcNAc<sub>2</sub>Man<sub>9</sub>Glc<sub>1</sub> intermediate on misfolded proteins is sensed by Cnx1p, which mediates ER-retention and subsequent proteolytic degradation of misfolded proteins. However, *S.cerevisiae* Cne1p must have some other mechanism for sensing misfolded proteins since UGGT activity is not present in this species. In fact, experiments by McCracken and Brodsky demonstrate that nonglucosylated  $\alpha$ -factor is degraded in a calnexin dependent fashion (McCracken and Brodsky, 1996). Thus, the glycosylation state of unfolded proteins is likely not important for Cne1p recognition.

This model would explain the observed lack of calnexin-glycoprotein physical association in *S.cerevisiae* and *S.pombe*. It would also explain the lack of any effect on secretion of wild type proteins in a *S.cerevisiae*  $\Delta cne1$  strain. This model is also consistent with observations made with mutant protein secretion and degradation in calnexinmanipulated *S.cerevisiae* and *S.pombe* strains. Whether physical association of yeast calnexin with misfolded proteins is required their 'presentation' to an ER proteolytic degradation machinery remains to be tested.

## 5.7 Future experiments

#### S.cerevisiae

In experiments performed in the course of this work, it was observed that the phenotype which associated *S.cerevisiae* Cnelp with quality control was not strong. It is possible that Cnelp either plays a minor role in quality control or that in *S.cerevisiae* there is an alternative pathway. Hence, other quality control components were sought using a synthetic lethal screen. This approach was unsuccessful in identifying genes which interact genetically with *CNE1* (my unpublished results).

An alternative approach to identify gene products that interact with Cne1p is the two-hybrid screen (Allen *et al.*, 1995). This genetic screen is used to identify proteins which may interact with a known protein, in this case Cne1p. Cne1p association with ER membranes is likely due to association with other ER membrane proteins. In addition, Cne1p may also physically interact with other components of the ER quality control apparatus. Therefore, the two-hybrid screen provides a feasible approach to identify such proteins. Once identified, the gene encoding these proteins can be deleted, and their effects on cell viability, ER quality control and Cne1p localization can be determined. Subsequently, one can use two-hybrid and synthetic lethal screens to identify proteins which interact with Cne1p interacting genes. Alternatively, if the interacting genes are essential for viability, temperature sensitive alleles can be sought and genes complementing the temperature sensitive phenotype can be obtained.

Proteins expected to interact with Cne1p would include ER proteins implicated in quality control, in addition to transmembrane proteins which would keep Cne1p associated with ER membranes. ER proteins involved in quality control would include BiP, Scj1p, Sec63p, Ire1p, glucosidase I, glucosidase II, PDI homologues and cyclophilin homologues (Shamu *et al.*, 1994; Freedman, 1995; Brodsky, 1996). If Cne1p associates with proteins cotranslocationally, Cne1p may in fact associate with components of the translocation apparatus, such as components of the Sec61p complex and/or the Sec62p/Sec63p complex (High and Stirling, 1993). Association with the translocation machinery may enable Cne1p to remain associated with ER membranes.

#### S.pombe

From results presented in this thesis, it is likely that the cytosolic tail of Cnx1p interacts with cytosolic factors which help mediate protein degradation as well as ER retention. In addition, the role of kinases which phosphorylate Cnx1p can be determined. Thus, proteins which interact with the cytosolic (or lumenal) domain of Cnx1p can also be sought using the two-hybrid screen.

Since  $cnxl^+$  is an essential gene, temperature sensitive alleles should be isolated and characterized. If the  $cnxl^+$  essential function does involve protein folding, shifting cnxl ts cells to non-permissive temperature should affect protein processing, secretion, degradation or a combination of these events. Alternatively, temperature sensitive  $cnxl^+$  mutant alleles may give insight into other possible Cnx1p functions. In addition, temperature sensitive alleles of  $cnxl^+$  will allow identification of other genes which genetically interact with  $cnxl^+$ . Complementation of the temperature sensitive  $cnxl^+$  alleles with a high copy *S.pombe* library will identify other genes which function in the same pathway as Cnx1p. Characterization of these proteins and their interaction with Cnx1p will bring about a better understanding of the cellular pathway in which calnexin's function plays a role.

## **Concluding remarks**

The evidence presented in this thesis points to a function for yeast calnexin in ER protein quality control. Therefore, the ER quality control apparatus can be dissected by identifying other factors implicated in this pathway.

## REFERENCES

Abeijon, C., Orlean, P., Robbins, P. W. and Hirschberg, C. B. (1989). Topography of glycosylation in yeast: characterization of GDP mannose transport and lumenal guanosinediphosphatase activities in Golgi-like vesicles. Proc. Natl. Acad. Sci. USA. *86*, 6935-6939.

Ahluwalia, N., Bergeron, J. J., Wada, I., Degen, E., and Williams, D. B. (1992). The p88 molecular chaperone is identical to the endoplasmic reticulum membrane protein, calnexin. J. Biol. Chem. 267, 105 4-10918.

Allen, J. B., Walberg, M. W., Edwards, M. C., and Elledge, S. J. (1995). Finding prospectives partners in the library: the two hybrid system and phage display find a match. Trends Biochem. Sci. 20, 511-516.

Amara, J., Cheng, S. H., and Smith, A. E. (1992). Intracellular protein trafficking defects in human disease. Trends. Cell Biol. 2, 145-149.

Amin, J., Ananthan, J., and Voellmy, R. (1988). Key features of heat shock regulatory elements. Mol. Cell Biol. 8, 3761-3769.

Anfinsen C. B. (1973). Principles that govern the folding of protein chains. Science 181, 223-230.

Anderson, K. S., and Cresswell, P. (1994). A role for calnexin (IP90) in the assembly of class II MHC molecules. EMBO J. *13*, 675-682.

Argos, P., Rao, J. K., and Hargrave, P. A. (1982). Structural predictions of membranebound proteins. Eur. J. Biochem. 128, 565-575.

Arkinstall, S., Payton, M., and Maundrell, K. (1995). Activation of phospholipase Cγ in *Schizosaccharomyces pombe* by coexpression of receptor or nonreceptor tyrosine kinase. Mol. Cell Biol. *15*, 1431-1438.

Baksh, S., and Michalak, M. (1991). Expression of calreticulin in *Escherichia coli* and identification of its Ca2+ binding domains. J. Biol. Chem. 266, 21458-21465.

Bandlow, W. (1972). Membrane separation and biogenesis of the outermembrane of yeast mitochondria. Biochem. Biophys. Acta 282, 105-122.

Bardwell, J. C. A., and Beckwith, J. (1993). The bonds that tie: catalyzed disulfide bond formation. Cell 74, 769-771.

Benedetti, C. E. and Turner, J. G. (1995). Nucleotide Sequence of an Arabidopsis thaliana cDNA. Plant Physiol. *109*, 338.

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

Bonifacino, J. S., and Lippincott-Schwartz, J. (1991). Degradation of proteins with the endoplasmic reticulum. Curr. Opin. Cell Biol. *3*, 592-600.

Bonnerot, C., Marks, M. S., Cosson, P., Robertson, E. J., Bikoff, E. K., Germain, R. N., and Bonifacino, J. S. (1994). Association with BiP and aggregation of class II MHC molecules synthesized in the absence of invariant chain. EMBO J. *13*, 934-44.

Bordier, C. (1981). Phase separation of integral membrane proteins in Triton X-114 solution. J. Biol. Chem. 256, 1604-1606.

Bowman, B. J., and Slayman, C. W. (1979). The effects of vanadate on the plasma membrane ATPase of *Neurospora crassa*. J. Biol. Chem. 254, 2928-2934.

Boyce, J. M., Coates, D., Fricker, M. D., and Evans, D. E. (1994). Genomic sequence of a calnexin homolog from *Arabidopsis thaliana*. Plant Physiol. *106*, 1691.

Brantly, M., Nukiwa, Y. and Crystal, R. G. (1988). Molecular basis of alpha-1-antitrypsin deficiency. Am. J. Med. 84, 13-31.

Brodsky, J. L., and Schekman, R. (1993). A Sec63p-BiP complex from yeast is required for protein translocation in a reconstituted proteoliposome. J. Cell Biol. *123*, 1355-1363.

Brodsky, J. L. (1996). Post-translational protein translocation: not all hsc70s are created equal. Trends Biochem Sci. 21, 122-129.

Cala, S. E., Ulbright, C., Kelley, J. S., and Jones, L. R. (1993). Purification of a 90-kDa protein (Band VII) from cardiac sarcoplasmic reticulum. J. Biol. Chem. 268, 2969-2975.

Camacho, P., and Lechleiter, J. D. (1995). Calreticulin inhibits repetitive calcium waves. Cell 82, 765-771.

Cannon, K. S., Hebert, D. N., and Helenius, A. (1996). Glycan dependent and independent association of vesicular stomatitis virus G with calnexin. J. Biol. Chem. (in press).

Carpenter, C. L., and Cantley, L. C. (1996). Phosphoinositide kinases. Curr. Opin. Cell Biol. *8*, 153-158.

Carreno, B. M., Schreiber, K. L., Mckean, D. J., Stroynowski, I., and Hansen, T. H. (1995). Aglycosylated and phosphatidylinositol-anchored MHC class 1 molecules are associated with calnexin: evidence implicating the class 1-connecting peptide segment in calnexin association. J. Immunol. *154*, 5173-5180.

Chen, F., Hayes, P. M., Mulronney, D. M., and Pan, A. (1994). Identification and characterization of cDNA clones encoding plant calreticulin in barley. Plant Cell 6, 835-843.

Chen, W., Helenius, J., Braakman, I., and Helenius, A. (1995). Cotranslational folding and calnexin binding during glycoprotein synthesis. Proc. Natl. Acad. Sci. USA 92, 6229-6233.

Cheng, S. H., Gregory, R. J., Marshall, J., Paul, S., Souza, D. W., White, G. A., O'Riordan, C. R., and Smith, A. E. (1990). Defective intracellular transport and processing of CFTR is the molecular basis of most cystic fibrosis. Cell 63, 827-34. Clark, A. R. (1996). Molecular chaperones in protein folding and translocation. Curr. Opin. Struct. Biol. *6*, 43-50.

Conway, E. M., Liu, L. L., Nowakowski, B., Steinermosonyi, M., Ribeiro, S. P., and Michalak, M. (1995). Heat shock-sensitive expression of calreticulin - *in vitro* and *in vivo* up-regulation. J. Biol. Chem. 270, 17011-17016.

Costigan, C., Gething, S. and Snyder, M. (1992). A synthetic lethal screen identifies *Slk1*, a novel protein kinase homolog implicated in yeast cell morphogenesis and cell biology. Mol. Cell Biol. *12*, 1162-1178.

Cox, D. W. (1989). Alpha<sub>1</sub>-antitrypsin deficiency. The Metab. Basis of Inher. Dis., 6th edition, Vol. 2, 2409-2438.

Craig, E. (1989). Essential roles of 70 kDa heat shock inducible proteins. Bioessays 11, 48-52.

Cross, G.A. M. (1990). Glycoprotein anchoring of plasma membrane proteins. Annu. Rev. Cell Biol. *6*, 1-39.

David, V., Hochstenbach, F., Rajagopalan, S., and Brenner, M. B. (1993). Interaction with newly synthesized and retained proteins in the endoplasmic reticulum suggests a chaperone function for human integral membrane protein IP90 (calnexin). J. Biol. Chem. 268, 9585-92.

de Duve, C. (1975). Exploring cells with a centrifuge. Science 189, 186-194.

de Silva, A. M., Balch, W. E., and Helenius, A. (1990). Quality control in the endoplasmic reticulum: folding and misfolding of vesicular stomatitis virus G protein in cells and *in vitro*. J. Cell Biol. 111, 857-866.

de Virgilio, C., Burckert, N., Neuhaus, J. M., Boller, T., and Wiemken, A. (1993). *CNE1*, a *Saccharomyces cerevisiae* homologue of the genes encoding mammalian calnexin and calreticulin. Yeast 9, 185-8.

Degen, E., Cohen-Doyle, M. F., and Williams, D. B. (1992). Efficient dissociation of the p88 chaperone from major histocompatibility complex class I molecules requires both  $\beta_2$ -microglobulin and peptide. J. Exp. Med. *175*, 1653-1661.

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

Denecke, J., Carlsson, L. E., Vidal, S., Hoglund, A. S., Ek, B., van Zeijl, M. J., Sinjorgo, K. M., and Palva, E. T. (1995). The tobacco homolog of mammalian calreticulin is present in protein complexes *in vivo*. Plant Cell 7, 391-406.

Dorner, A. J., Bole, D. G., and Kaufman, R. J. (1987). The relationship of N-linked glycosylation and heavy chain-binding protein association with the secretion of glycoproteins. J. Cell Biol. *105*, 2665-2674.

Feldheim, D., Rothblatt, J. and Schekman, R. (1992). Topology and functional domains of Sec63p, an endoplasmic reticulum membrane protein required for secretory protein translocation. Mol. Cell Biol. *12*, 3288-3296.

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

Fernandez, F. S., Jannatipour, M., Hellman, U., Rokeach, L. A., and Parodi, A. J. (1996). A new stress protein: synthesis of *Schizosaccharomyces pombe* UDP-glucose: glycoprotein glucosy!transferase mRNA is induced by stress conditions but the enzyme is not essential for cell viability. EMBO J. *15*, 705-713.

Flynn, G. C., Pohl, J., Flocco, M. T., and Rothman, J. E. (1991). Peptide-binding specificity of the molecular chaperone BiP. Nature *353*, 726-30.

Forsayeth, J. R., Gu, Y., and Hall, Z. W. (1992). BiP forms stable complexes with unassembled subunits of the acetylcholine receptor in transfected COS cells and in C2 muscle cells. J. Cell Biol. 117, 841-7.

Franzusoff, A., Rothblatt, J., and Schekman, R. (1991). Analysis of polypeptide transit through yeast secretory pathway. Methods Enzymol. *194*, 662-74.

Freedman, R. B., Hirst, T. R., and Tuite M. F. (1994). Protein disulfide isomerase: building bridges in protein folding. Trends Biochem. Sci. 19, 331-336.

Freedman, R. B. (1995). The formation of protein disulphide bonds. Curr. Opin. Struct. Biol. 5, 85-91.

Frigerio, G., and Pelham, H. R. (1993). A *Saccharomyces cerevisiae* cyclophilin resident in the endoplasmic reticulum. J. Mol. Biol. 233, 183-8.

Galvin, K., Krishna, S., Ponchel, F., Frohlich, M., Cummings, D. E., Carlson, R., Wands, J. R., Isselbacher, K. J., Pillai, S., and Ozturk, M. (1992). The major histocompatibility class 1 antigen-binding protein p88 is the product of the calnexin gene. Proc. Natl. Acad. Sci. 89, 8452-8456.

Gelman, M. S., Chang, W. S., Thomas, D. Y., Bergeron, J. J. M., and Prives, J. M. (1995). Role of the endoplasmic reticulum chaperone calnexin in subunit folding and assembly of nicotinic acetylcholine receptors. J. Biol. Chem. 270, 15085-15092.

Gething, M. J., and Sambrook, J. (1992). Protein folding in the cell. Nature 355, 33-45.

Goode, J. H., Settlage, S. B., Wilson, R. F., and Dewey, R. E.(1995). Isolation of a calnexin homolog (GenBank U20502) from developing scybean seeds. Plant Physiol. 108, 1341.

Graham, K. S., Le, A., and Sifers, R. N. (1990). Accumulation of the insoluble PI Z variant of human alpha 1-antitypsin within the hepatic endoplasmic reticulum does not elevate the steady- state level of grp78/BiP. J. Biol. Chem. 265, 20463-20468.

Grimm, C., Kohli, J., Murray, J. and Maundrell, K. (1988). Genetic engineering of *Schizosaccharomyces pombe*: a system for gene disruption and replacement using the *ura4*+ gene as a selectable marker. Mol. Gen. Genet. 215, 81-86.

Hammond, C., Braakman, I., and Helenius, A. (1994). Role of N-linked oligosaccharide recognition, glucose trimming, and calnexin in glycoprotein folding and quality control. Proc. Natl. Acad. Sci. USA *91*, 913-917.

Hammond, C., and Helenius, A. (1994a). Folding of VSV G protein: sequential interaction with BiP and calnexin. Science 267, 456-458.

Hammond, C., and Helenius, A. (1994b). Quality control in the secretory pathway: retention of a misfolded viral membrane glycoprotein involves cycling between the ER, intermediate compartment, and the Golgi apparatus. J. Cell Biol. *126*, 41-52.

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

Hann, B. C., Poritz, M. A., and Walter, P. (1989). *Saccharomyces cerevisiae* and *Schizosaccharomyces pombe* contain a homologue to the 54 kDa subunit of the signal recognition particle that in *S.cerevisiae* is essential for growth. J. Cell Biol. *109*, 3223-3230.

Hardwick, K. G., Lewis, M. J., Semenza, J., Dean, N., and Pelham, H. R. (1990). *ERD1*, a yeast gene required for the retention of lumenal endoplasmic reticulum proteins, affects glycoprotein processing in the Golgi apparatus. EMBO J. 9, 623-30.

Hartl, F. U., and Martin, J. (1995). Molecular chaperones in cellular protein folding. Curr. Opin. Struct. Biol. 5, 92-102.

Hawn, T. R., Tom, T. D., and Strand, M. (1993). Molecular cloning and expression of SmIrV1, a *Schistosoma mansoni* antigen with similarity to calnexin, calreticulin, and OvRal1. J. Biol. Chem. 268, 7692-7698.

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

Hendershot, L., Bole, D., Kohler, G., and Kearney, J. F. (1987). Assembly and secretion of heavy chains that do not associate post-translationally with immunoglobulin heavy chain-binding protein. J. Cell Biol. *104*, 761-767.

e

Hendershot, L. M. (1990). Immunoglobulin heavy chain and binding protein complexes are dissociated *in vivo* by light chain addition. J. Cell Biol. 111, 829-837.

Hendrick, J. P., and Hartl, F.-U. (1993). Molecular chaperones functions of heat shock proteins. Annu. Rev. Biochem. 62, 349-384.

Herscovics, A. and Orlean, P. (1993). Glycoprotein synthesis in yeast. FASEB J. 7, 540-550.

High, S., and Stirling, C. J. (1993). Protein translocation across membranes: common themes in divergent organisms. Trends Cell Biol. *3*, 335-339.

Hochstenbach, F., David, V., Watkins, S., and Brenner, M. (1992). Endoplasmic reticulum resident protein of 90 kilodaltons associates with the T- and B-cell antigen receptors and major histocompatibility complex antigens during their assembly. Proc. Natl. Acad. Sci. USA 89, 4734-4738.

Hoheisel, J. D., Maier, E., Mott, R., McCarthy, L., Grigoriev, A. V., Schalkwyk, L. C., Nizetic, D., Francis, F., and Lehrach H. (1993). High resolution cosmid and P1 maps spanning the 14 Mb genome of the fission yeast *S.pombe*. Cell *73*, 109-120.

Huang, L., Franklin, A. E., and Hof?man, N. E. (1993). Primary structure and characterization of an *Arabidopsis thaliana* calnexin-like protein. J. Biol. Chem. 268, 6560-6566.

Huggins, M. C., Gibbs, J., and Moloney, N. A. (1995). Cloning of a *Schistosoma japonicum* gene encoding an antigen with homology to calreticulin. Mol. Biochem. Parasitol. 71, 81-87.

Hugueney, P., Bouvier, F., Badillo, A., d'Harlingue, A., Kuntz, M. and Camara, B. (1995). Identification of a plastid protein involved in vesicle fusion and/or membrane protein translocation. Proc. Natl. Acad. Sci. U.S.A. 92, 5630-5634.

Hurtley, S. M., Bole, D. G., Hoover-Litty, H., Helenius, A., and Copeland, C. S. (1989). Interactions of misfolded influenza virus hemagglutinin with binding protein (BiP). J. Cell Biol. *108*, 2117-2126.

Huttner, W. B., and Tooze, S. A. (1989). Biosynthetic protein transport in the secretory pathway. Curr. Opin. Cell Biol. *1*, 648-654.

Ito, H., Fukuda, Y., Murata, K. and Kimura, A. (1983). Transformation of intact yeast cells treated with alkali cations. J. Bacteriol. *153*, 163-168.

Jackson, M. R., Cohen-Doyle, M. F., Peterson, P. A., and Williams, D. B. (1994). Regulation of MHC class I transport by the molecular chaperone, calnexin (p88, IP90). Science *263*, 384-387.

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

Jaworski, D. C., Simmen, F. A., Lamoreaux, W., Coons, L. B., Muller, M. T., and Needham, G. R. (1995). A secreted calreticulin protein in ixodid tick (*Amblyomma americanum*) saliva. J. Insect Physiol. *41*, 369-375.

Jensen, T. J., Loo, M. A., Pind, S., Williams, D. B., Goldberg, A. L., and Riordan, J. R. (1995). Multiple proteolytic systems, including the proteasome, contribute to CFTR processing. Cell 83, 129-135.

Jones, J. S., and Prakash, L. (1990). Yeast *Saccharomyces cerevisiae* selectable markers in pUC18 polylinkers. Yeast *6*, 363-366.

Kahn-Perlès, B., Salamero, J., and Jouans, O. (1994). Biogenesis of MHC class I antigens: involvement of multiple chaperone molecules. Eur. J. Cell Biol. *64*, 176-185.

Kearse, K. P., Williams, D. B., and Singer, A. (1994). 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.

Kennedy, T. E., Kuhl, D., Barzilai, A., Sweatt, J. D., and Kandel, E. R. (1992). Longterm sensitization training in Aplysia leads to an increase in calreticulin, a major presynaptic calcium-binding protein. Neuron 9, 1013-1024.

Kim, P. S., and Arvan, P. (1995). Calnexin and BiP act as sequential molecular chaperones during thyroglobulin folding in the endoplasmic reticulum. J. Cell Biol. *128*, 29-38.

Kim, P. S., Bole, D., and Arvan, P. (1992). Transient aggregation of nascent thyroglobulin in the endoplasmic reticulum: relationship to the molecular chaperone, BiP. J. Cell Biol. *118*, 541-549.

Knittler, M. R., and Haas, I. G. (1992). Interaction of BiP with newly synthesized immunoglobulin light chain molecules: cycles of sequential binding and release. EMBO J. 11, 1573-81.

Kohno, K., Normington, K., Sambrook, J., Gething, M. J., and Mori, K. (1993). The promoter region of the yeast KAR2 (BiP) gene contains a regulatory domain that responds to the presence of unfolded proteins in the endoplasmic reticulum. Mol. Cell Biol. *13*, 877-90.

Kubota, S., Yoshida, Y., and Kumota. (1977). Studies on the microsomal electrontransport system of anaerobically grown yeast. Purification and characterization of NADPH-cytochrome c reductase. J. Biochem. 81, 187-195.

Kwiatkowski, B. A., Zielinskakwiatkowska, A. G., Migdalski, A., Kleczkowski, L. A., and Wasilewska, L. D. (1995). Cloning of two cDNAs encoding calnexin-like and calreticulin-like proteins from maize (*Zea mays*) leaves: Identification of potential calciumbinding domains. Gene *165*, 219-222.

Kyte, J., and Doolittle, R. F. (1982). A simple method for displaying the hydropathic character of a protein. J. Mol. Biol. 157, 105-132.

Labriola, C., Cazzulo J. J., and Parodi, A. J. (1995). Retention of glucose units added by the UDP-glucose: glycoprotein glucosyltransferase delays exit of glycoproteins from the endoplasmic reticulum. J. Cell Biol. *130*, 771-779.

Le, A., Ferrell, G. A., Dishon, D. S., Le, Q. Q., and Sifers, R. N. (1992). Soluble aggregates of the human PiZ alpha 1-antitrypsin variant are degraded within the endoplasmic reticulum by a mechanism sensitive to inhibitors of protein synthesis. J. Biol. Chem. 267, 1072-80.

Le, A., Graham, K. S., and Sifers, R. N. (1990). Intracellular degradation of the transport-impaired human PiZ alpha 1- antitrypsin variant. Biochemical mapping of the degradative event among compartments of the secretory pathway. J. Biol. Chem. 265, 14001-7.

Le, A., Steiner, J. L., Ferrell, G. A., Shaker, J. C., and Sifers, R. N. (1994). Association between calnexin and a secretion-incompetent variant of human alpha 1-antitrypsin. J. Biol. Chem. 269, 7514-7519.

Lenter, M., and Vestweber, D. (1994). The integrin chains beta 1 and alpha 6 associate with the chaperone calnexin prior to integrin assembly. J. Biol. Chem. 269, 12263-12268.

Li, Z., and Srivastava, P. K. (1993). Tumor rejection antigen gp94/grp94 is an ATPase: implications for protein folding and antigen presentation. EMBO J. *12*, 3143-3251.

Liu, N., Fine, R. E., and Johnson, R. J. (1993). Comparison of cDNAs from bovine brain coding for two isoforms of calreticulin. Biochim. Biophys. Acta *1202*, 70-76.

Loebermann, H., Tokuoka, R., Deisenhofer, J., and Huber, R. (1984). Human alpha 1proteinase inhibitor. Crystal structure analysis of two crystal modifications, molecular model and preliminary analysis of the implications for function. J. Mol. Biol. *177*, 531-557.

Lomas, D. A., Evans, D. L., Finch, J. T., and Carrell, R. W. (1992). The mechanism of Z alpha 1-antitrypsin accumulation in the liver. Nature *357*, 605-607.

Loo, T. W., and Clarke, D. M. (1994). Prolonged association of temperature-sensitive mutants of human P-glycoproteins with calnexin during biogenesis. J. Biol. Chem. 269, 28683-28689.

Loo, T. W., and Clarke, D. M. (1995). P-glycoprotein, J. Biol. Chem. 270, 21838-21844.

Maniatis, T., Fritsch, E. F. and Sambrook, J. (1982) *Molecular Cloning. A Laboratory Manual.* (Cold Spring Harbor Lab., Cold Spring Harbor, NY).

Margelese, L., Waneck, G. L., Suzuki, C. K., Degen, E., Flavell, R. A., and Williams, D. B. (1993). Identification of the region on the class I histocompatibility molecule that interacts with the molecular chaperone, p88 (calnexin, IP90). J. Biol. Chem. 268, 17959-17966.

Marquardt, T., and Helenius, A. (1992). Misfolding and aggregation of newly synthesized proteins in the endoplasmic reticulur. J. Cell Biol. *117*, 505-13.

Matsuoka, K., Seta, K., Yamakawa, Y., Okuyama, T., Shinoda, T., and Isobe, T. (1994). Covalent structure of bovine brain calreticulin. Biochem. J. 298, 435-42.

Maundrell, K. (1993). Thiamine-repressible expression vectors pREP and pRIP for fission yeast. Gene *123*, 127-130.

McCracken, A. A., and Kruse, K. B. (1993). Selective protein degradation in the yeast exocytic pathway. Mol. Biol. Cell 4, 729-36.

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

Meaden, P., Hill, K., Wagner, J., Slipetz, D., Sommer, S. S., and Bussey, H. (1990). The yeast *KRE3* gene encodes a probable endoplasmic reticulum protein required for (1-6)- $\beta$ -D-glucan synthesis and normal cell growth. Mol. Cell Biol. *10*, 3013-3019.

Melnick, J., Aviel, S., and Argon, Y. (1992). The endoplasmic reticulum stress protein GRP94, in addition to BiP, associates with unassembled immunoglobulin chains. J. Biol. Chem. 267, 21303-6.

Melnick, J., Dul, J. L., and Argon, Y. (1994). Sequential interaction of the chaperones BiP and GRP94 with immunoglobulin chains in the endoplasmic reticulum. Nature *370*, 373-375.

Michalak, M. (1996). Calreticulin (R. G. Landes, Austin).

Michalak, M., Milner, R. E., Burns, K., and Opas, M. (1992). Calreticulin. Biochem. J. 285, 681-92.

Mori, K., Sant, A., Kohno, K., Normington, K., Gething, M. J., and Sambrook, J. F. (1992). 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, 2583-93.

Moreno, S., Klar, A., and Nurse P. (1991). Molecular genetic analysis of fission yeast *Schizosaccharomyces pombe*. Methods Enzymol. *194*, 795-823.

Nauseef, W. M., McCormick, S. J., and Clark, R. A. (1995). Calreticulin functions as a molecular chaperone in the biosynthesis of mycloperoxidase. J. Biol. Chem. 270, 4741-4747.

Nigam, S. K., Goldberg, A. L., Ho, S., Rohde, M. F., Bush, K. T., and Sherman, M. (1994). A set of endoplasmic reticulum proteins possessing properties of molecular chaperones includes Ca(2+)-binding proteins and members of the thioredoxin superfamily. J. Biol. Chem. 269, 1744-1749.

Nijenhuis, M., and Neefjes, J. (1994). Early events in the assembly of major histocompatibility complex class II heterotrimers from their free subunits. Eur. J. Immunol. 24, 247-256.

Nobner, E., and Parham, P. (1995). Species-specific differences in chaperone interaction of human and mouse major histocompatibility complex class I molecules. J. Exp. Med. *181*, 327-337.

Normington, K., Kohno, K., Kozutsumi, Y., Gething, M. J., and Sambrook, J. (1989). *S.cerevisiae* encodes an essential protein homologous in sequence and function to mammalian BiP. Cell *57*, 1223-36.

Nuoffer, C., Horvath, A., And Reizman, H. (1993). Analysis of the sequence requirements for glycosylphosphatidylinositol anchoring of *Saccharomyces cerevisiae* Gas1 protein, J. Biol. Chem. 268, 10558-10563.

Ohsako, S., Hayashi, Y., and Bunick, D. (1994). Molecular cloning and sequencing of calnexin-t. An abundant male germ cell-specific calcium-binding protein of the endoplasmic reticulum. J. Biol. Chem. 269, 14140-8.

Ortmann, B., Androlewicz, M. J., and Cresswell, P. (1994). MHC class 1/beta 2microglobulin complexes associate with TAP transporters before peptide binding. Nature *368*, 864-867.

Otteken, A., and Moss, B. (19 96). Calreticulin interacts with newly synthesized human immunodeficiency virus type 1 envelope glycoprotein, suggesting a chaperone function similar to that of calnexin. J. Biol. Chem. 271, 97-103.

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

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

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

Palade, G. (1975). Intracellular aspects of the process of protein synthesis. Science 189, 347-58.

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

Parlati, F., Dominguez, M., Bergeron, J. J. M., and Thomas, D. Y. (1995b). *Saccharomyces cerevisiae CNE1* encodes an endoplasmic reticulum (ER) memorane 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.

Pelham, H. B., and Bienz, M. (1982). Expression of a *Drosophila* heat-shock protein in Xenopus oocytes: conserved and divergent regulatory signals. EMBO J. *1*, 1473-1477.

Pelham, H. R. (1990). The retention signal for soluble proteins of the endoplasmic reticulum. Trends Biochem Sci. 15, 483-6.

Pelham, H.R. (1995). Sorting and retrieval between the endoplasmic reticulum and the Golgi apparatus. Curr. Opin. Cell Biol. 7, 530-535.

Peter, F., Van, P. N., Soling, H. D. (1992). Different sorting of Lys-Asp-Glu-Leu proteins in the rat liver. J. Biol. Chem. 267, 10631-10637.

Peterson, J. R., Ora, A., Van, P.N., and Helenius, A. (1995). Transient, lectin-like association of calreticulin with folding intermediates of cellular and viral glycoproteins. Mol. Biol. Cell 6, 1173-1184.

Pidoux, A. L., and Armstrong, J. (1992). Analysis of the BiP gene and identification of an ER retention signal in *Schizosaccharomyces pombe*. EMBO J. 11, 1583-91.

Pidoux, A. L., and Armstrong, J. (1993). The BiP protein and the endoplasmic reticulum of *Schizosaccharomyces pombe*: fate of the nuclear envelope during cell division. J. Cell Sci. *105*, 1115-20.

Pind, S., Riordan, J. R., and Williams, D. B. (1994). Participation of the endoplasmic reticulum chaperone calnexin (p88, IP90) in the biogenesis of the cystic fibrosis transmembrane conductance regulator. J. Biol. Chem. 269, 12784-12788.

Pringle, J. R., Adams, A. E. M., Drubin, D. G., and Harrer, B. K. (1991). Immunofluorescence methods for yeast. Methods Enzymol. 194, 565-602.

Rabouille, C. and Nilsson, T. (1995). Redrawing compartmental boundaries in the exocytic pathway. FEBS Lett. *369*, 97-100.

Rajagopalan, S., and Brenner, M. B. (1994). Calnexin retains unassembled major histocompatibility complex class I free heavy chains in the endoplasmic reticulum. J. Exp. Med. *180*, 407-412.

Rajagopalan, S., Xu, Y., and Brenner, M. B. (1994). Retention of unassembled components of integral membrane proteins by calnexin. Science 263, 387-90.

Rassow, J., Voos, W., and Pfanner, N. (1995). Partner proteins determine multiple functions of Hsp70. Trends Cell Biol. 5, 207-212.

Riles, L., Dutchik, J. E., Baktha, A., McCauley, B. K., Thayer, E. C., Leckie, M. P., Braden, V. V., Depke, J. E., and Olson, M. V. (1993). Physical maps of the six smallest chromosomes of *Saccharomyces cerevisiae* at a resolution of 2.6 kilobase pairs. Genetics *134*, 81-150.

Rose, M. D., Misra, L. M., and Vogel, J. P. (1989). *KAR2*, a karyogamy gene, is the yeast homolog of the mammalian BiP/GRP78 gene. Cell 57, 1211-1221.

Sailbil, H. (1996). The lid that shapes the pot: structure and function of the chaperonin GroES. Curr. Biol. 4, 1-4.

Sambrook, J. F. (1990). The involvement of calcium in the transport of secretory proteins from the endoplasmic reticulum. Cell *59*, 197-199.

Sanders, S. L., Whitfield, K. M., Vogel, J. P., Rose, M. D., and Schekman, R. W. (1992). Sec61p and BiP directly facilitate polypeptide translocation into the ER. Cell 69, 353-365.

Schlenstedt, G., Harris, S., Risse, B., Lill, R., and Silver, P. A. (1995). A yeast DnaJ homologue, Scj1p, can function in the endoplasmic reticuluri with BiP/Kar2p via a conserved domain that specifies interactions with Hsp70s. J. Cell Biol. *129*, 979-988.

Schulze, E., and Kirschner, M. (1986). Microtubule dynamics in interphase cells. J. Cell Biol. *102*, 1020-1031.

Schreiber, K. L., Bell, M. P., Huntoon, C. J., Rajagopalan, S., Brenner, M. B., and McKean, D. J. (1994). Class II histocompatibility molecules associate with calnexin during assembly in the endoplasmic reticulum. Int. Immunol. 6, 101-111.

Shamu, C. E., Cox, J. S. and Walter P. (1994). The unfolded-protein-response pathway in yeast. Trends Cell Biol. 4, 56-61.

Sharp, H. L., Bridges, R. A., Krivit, W, and Frier, E. F. (1969). Cirrhosis associated with alpha1-antitrypsin deficiency: a previously unrecognized inherited disorder. J. Lab. Clin. Med. 73, 934-939.

Sherman, F., Fink, G. R. and Lawrence, C. W. (1979). Methods in Yeast Genetics (Cold Spring Harbor Lab., Cold Spring Harbor, NY).

Schmid, F. X. (1993). Prolyl isomerase: enzymatic catalysis of slow protein-folding reactions. Annu. Rev. Biophys. Biomol. Struct. 22, 123-42.

Sifers, R. N., Finegold, M. J., and Woo, S. L. (1989). Alpha-1-antitrypsin deficiency: accumulation or degradation of mutant variants within the hepatic endoplasmic reticulum. Am. J. Respir. Cell Mol. Biol. 1, 341-345.

Sifers, R. N. (1992a). Z and the insoluble answer. Nature 357, 541-542

Sifers, R. N., Finegold, M. J, and Woo, S. L. C. (1992b). Molecular biology of alphalantitrypsin deficiency. Seminars in Liver Disease 12, 301-310.

Sifers, R. N. (1995). Defective protein folding as a cause of disease. Nature Struct. Biol. 2, 355-357.

Sikorski, R. S., and Hieter, P. (1989). A system of shuttle vectors and yeast host strains designed for efficient manipulation of DNA in *Saccharomyces cerevisiae*. Genetics *122*, 19-27.

Simons, J. F., Ferro-Novick, S., Rose, M. D., and Helenius, A. (1995). BiP/Kar2p serves as a molecular chaperone during carboxypeptidase Y folding in yeast. J. Cell Biol. *130*, 41-49.

Smith, M. J. (1992a). Nucleotide sequence of a *Drosophila melanogaster* gene encoding a calreticulin homologue. Dna. Seq. 3, 247-50.

Smith, M. J. (1992b). A *C.elegans* gene encodes a protein homologous to mammalian calreticulin. Dna. Seq. 2, 235-40.

Smith, M. J., and Koch (1989). Multiple zones in the sequence of calreticulin (CRP55, calregulin, HACBP), a major calcium binding ER/SR protein. EMBO J. 8, 3581-3586.

Smith, D. B., and Johnson, K. S. (1988). Single-step purification of polypeptides expressed in *Escherichia coli* as fusions glutathione S-transferase. Gene 67, 31-40.

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

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

Spiro, R. G., Zhu, Q., Bhoyroo, V., and Solig, H.-D. (1996). Definition of the lectin like properties of the molecular chaperone, calreticulin, and demonstration of its copurification with endomannosidase from rat liver. J. Biol. Chem. 271, 11588-11594.

Sprague, G. F.(1991). Assay of yeast mating reaction. Methods Enzymol. 194, 77-94.

Stein, P. E., and Carrell, R. W. (1995). What do dysfunctional serpins tell us about molecular mobility and disease? Nature Struct. Biol. 2, 96-113.

<u>a a</u>

Suh, W., Cohen-Doyle, M., Fruh, K., Wang, K., Peterson, P., and Williams, D. (1994). Interaction of MHC class I molecules with the transporter associated with antigen presentation. Science 264, 1322-1326.

Sulston, J., Du, Z., Thomas, K., Wilson, R., Hillier, L., Staden, R., Halloran, N., Green, P., Thierry-Mieg, J., Qiu, L., Dear, S., Coulson, A., Craxton, M., Durbin, R., Berks, M., Metzstein, M., Hawkins, T., Ainscough, R., and Waterson, R. (1992). The *C.elegans* genome sequencing project: a beginning. Nature *356*, 37-41.

Tachikawa, H., Takeuchi, Y., Funahashi, W., Miura, T., Gao, X. D., Fujimoto, D., Mizunaga, T., and Onodera, K. (1995). Isolation and characterization of the yeast gene, *MPD1*, the overexpression of which suppresses inviability caused by protein disulfide isomerase depletion. FEBS Lett. *369*, 212-216.

Thomas, P. J., Qu, B.-H., and Pederson, P. L. (1995). Defective protein folding as a basis of human disease. Trends Biochem. Sci. 20, 456-459

Tjoelker, L. W., Seyfried, C. E., Eddy, R. L., Byers, M. G., Shows, T. B., Calderon, J., Schreiber, R. B., and Gray, P. W. (1994). Human, mouse, and rat calnexin cDNA cloning: identification of potential calcium binding motifs and gene localization to human chromosome 5. Biochemistry *33*, 3229-3236.

Tohe-e, A., Ueda, Y., Kakimoto, S.-I. and Oshima, Y. (1973). Isolation and characterization acid phosphatase mutants in *Saccharomyces cerevisiae*. J. Bacteriol. *113*, 727-738.

Treves, S., Zorzato, F., and Pozzan, T. (1992). Identification of calreticulin isoforms in the central nervous system. Biochem. J. 287, 579-581.

Trimble, R. B., and Verostek, M. F. (1995). Glycoprotein oligosaccharide synthesis and processing in yeast. Trends Glycosci. Glycotechnol. 7, 1-30.

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

Unnasch, T. R., Gallin, M. Y., Soboslay, P. T., Erttmann, K. D., Greene, B. M., (1988). Isolation and characterization of expression cDNA clones encoding antigens of *O. volvulus* infeced larvae. J. Clin. Invest. *82*, 262-269.

Van, P. N., Peter, F., and Soling, H. D. (1989). Four intracisternal calcium-binding glycoproteins from rat liver microsomes with high affinity for calcium. No indication forcalsequestrin-like proteins in inositol 1,4,5-trisphosphate-sensitive calcium sequestering rat liver vesicles. J. Biol. Chem. 264, 17494-501

Vernet T., Dignard, D., and Thomas, D. Y. (1987). A family of yeast expression vectors containing the phage f1 intergenic region. Gene *52*, 225-233.

von Heijne, G. (1986). A new method for predicting signal sequence cleavage sites. Nucleic Acids Res. 14, 4683-4691.

Wada, I., Ou, W.-J., Liu, M.-C., and Scheele, G. (1994). Chaperone function of calnexin for the folding intermediate of gp80, the major secretory protein in MDCK cells. J. Biol. Chem. 269, 7464-7472.

Wada, I., Rindress, D., Cameron, P. H., Ou, W. J., Doherty, J. J., Louvard, D., Bell, A. W., Dignard, D., Thomas, D. Y., and Bergeron, J. J. (1991). SSR alpha and associated calnexin are major calcium binding proteins of the endoplasmic reticulum membrane. J. Biol. Chem. 266, 19599-19610.

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

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

Ware, F. E., Vassilakos, A., Peterson, P. A., Jackson, M. R., Lehrman, M. A., and Williams, D. B. (1995). The molecular chaperone calnexin binds Glc<sub>1</sub>Man<sub>9</sub>GlcNAc<sub>2</sub> oligosaccharide as an initial step in recognizing unfolded glycoproteins. J. Biol. Chem. 270, 4697-4704.

Williams D. B. and Watts T. H. (1995). Molecular chaperones in antigen presentation. Curr. Opin. Immunol. 7, 77-84.

Wise, J. A. (1991). Freparation and analysis of low molecular weight RNAs and small ribonucleoproteins. Methods Enzymol. *194*, 405-408.

Wright, A., Maundrell, K., Hyer, W-D., Beach, D., and Nurse, P. (1986). Vectors for the construction of gene banks and the integration of cloned genes in *Schizosaccharomyces pombe* and *Saccharomyces cerevisiae*. Plasmid *15*, 156-158.

Wu, Y., Whitman, I., Molmenti, E., Moore, K., Hippenmeyer, P., and Perlmutter, D. (1994). A lag in the intracellular degradation of the mutant  $\alpha_1$ -antitrypsin correlates with the liver disease phenotype in homozygous PiZZ  $\alpha_1$ -antitrypsin deficiency. Proc. Natl. Acad. Sci. USA 91, 9014-9018.

Ziegler, F. D., Gemmill, T. R., and Trimble, R. B. (1994). Glycoprotein synthesis in yeast: early events in N-linked oligosaccharide processing in *Schizosaccharomyces pombe*. J. Biol. Chem. 269, 12527-12535.

57