UNDERSTANDING THE FUNCTIONS OF MOLECULAR CHAPERONES AND THIOL OXIDOREDUCTASES IN MHC CLASS I BIOGENESIS

by

Ursula G. Danilczyk

A thesis submitted in conformity with the requirements for the Degree of Doctor of Philosophy, Graduate Department of Immunology, University of Toronto

© Copyright by Ursula G. Danilczyk (2001)
The author has granted a non-exclusive licence allowing the National Library of Canada to reproduce, loan, distribute or sell copies of this thesis in microform, paper or electronic formats.

The author retains ownership of the copyright in this thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without the author’s permission.

L’auteur a accordé une licence non exclusive permettant à la Bibliothèque nationale du Canada de reproduire, prêter, distribuer ou vendre des copies de cette thèse sous la forme de microfiche/film, de reproduction sur papier ou sur format électronique.

L’auteur conserve la propriété du droit d’auteur qui protège cette thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

0-612-63682-8
ABSTRACT

UNDERSTANDING THE FUNCTIONS OF MOLECULAR CHAPERONES AND THIOL OXIDOREDUCTASES IN MHC CLASS I BIOGENESIS

Ursula G. Danilczyk,
Doctor of Philosophy (2001)
Graduate Department of Immunology,
University of Toronto

Calnexin (CNX) is a membrane protein of the ER that functions as a molecular chaperone and as a component of the ER quality control machinery. Calreticulin (CRT), a soluble analog of CNX, is thought to possess similar functions but these have never been directly demonstrated. Using a Drosophila expression system and the mouse class I histocompatibility molecule as a model glycoprotein, we found that CRT does possess apparent chaperone and quality control functions, enhancing class I folding and subunit assembly, stabilizing subunits, and impeding export of assembly intermediates from the ER. Indeed, the functions of CNX and CRT were largely interchangeable. We also determined that a soluble form of CNX (residues 1-387) can functionally replace its membrane-bound counterpart. However, the different topological environments of CNX and CRT are important in determining their distinct substrate specificities.

CNX and CRT bind to glycoprotein substrates in part through a lectin site specific for monoglucosylated (Glc_1,Man_7,9,GlcNAc_2) oligosaccharides. In addition to this lectin-oligosaccharide interaction, in vitro studies have demonstrated that CNX and CRT can also bind to polypeptide segments of both glycosylated and non-glycosylated proteins. We examined whether polypeptide-based interactions occur between CNX and its substrates in cultured cells. We demonstrate that CNX is capable of associating in vivo with a substrate that completely lacks Asn-linked oligosaccharides. We conclude that both lectin-oligosaccharide and polypeptide-based interactions occur between CNX and
diverse proteins in such a system and that the strength of the latter interaction varies substantially between protein substrates.

ERp57 is a thiol oxidoreductase that binds to CNX and CRT and, at least in vitro, can promote disulfide bond formation in glycoproteins that interact with CNX or CRT. As yet, there is no direct information concerning how ERp57 functions in class I biogenesis. By using a heterologous Drosophila expression system, we showed that ERp57 does not affect the formation of the disulfide bond in cells expressing free H chain. Furthermore, the rate of class I molecules folding and assembly with β2m remains unchanged in the presence of ERp57. These are important findings that challenge the current view of ERp57 function in class I biogenesis.
To

My Husband

The Person Who Changed Everything
ACKNOWLEDGMENTS

In the course of working on this thesis, I have met extraordinary people who have made my graduate career a highly enjoyable experience. To begin, I gratefully acknowledge the contribution of my supervisor, Dr. D. Williams. David has provided a constructively critical environment that promotes development of the many personal and professional skills necessary for the transition from student to self-directed, self-reliant scientist. He has guided me through many problems of clarity and consistency in my graduate work and taught me the importance of communication skills in science.

I thank Dr. T. Watts and Dr. D. Isenman for graciously agreeing to be on my supervisory committee and for their help and encouragement during this lengthy endeavor. Special thanks to Dr. D. Isenman for providing me with the opportunity to work as a teaching assistant in the Biochemistry course 471. It was a highly stressful but very rewarding experience.

Every member of the Williams’ laboratory, both present and past, deserves my thanks for contributing to a remarkably friendly working environment. In particular, I would like to thank Dr. Davinder Chawla for sharing the lab bench when I first arrived at the lab, Dr. Kathy Vasilakos for our frequent scientific discussions, Dr. Nancy Berg for her infectious love of running and consequently my first 10 km race, Dr. Yoichiro Noguchi for sharing his sweets with me, Dr. Yoshito Ihara and Dr. Yoshiro Saito for their sense of humor and our evening and weekend discussions about life in Japan and Canada, Marie-Eve Paquet for my winter visit to Quebec City and for skiing with me (it was the biggest compliment), Vicki Stronge for running with me during the hot days of last summer and Michael Leach for introducing me to the Robert Adams Book Review Series.

Special thanks to Myrna Cohen-Doyle, for patiently answering innumerable laboratory questions, for her technical help, for all the after work activities (my first baseball game, aqua fit, swimming lessons, summer cottage parties, the ‘play hooky’ days at the cottage) and much, much more….

I am grateful to my friends Irena, Jurgen, Kasia, Bart, Ewa and Maciek who have responded to my difficult moods with amusement and understanding.

And finally, I am deeply indebted to my mother and my sister, who have approached my graduate studies with an enthusiasm that far exceeded my own.

Thank you all.
TABLE OF CONTENTS

ABSTRACT ................................................................................................................................. ii

ACKNOWLEDGMENTS .............................................................................................................. iv

TABLE OF CONTENTS .............................................................................................................. vi

LIST OF FIGURES AND TABLES .............................................................................................. ix

LIST OF ABBREVIATIONS ......................................................................................................... xi

1. CHAPTER I: INTRODUCTION ............................................................................................. 1

1.1 MHC Class I Molecules and the Immune Response to Viruses and Tumors ................. 2
1.2 The Structure and Assembly of MHC Class I Molecules .................................................. 4
  1.2.1 The Structure of MHC Class I Molecules .................................................................. 4
  1.2.2 The Assembly of Class I Complexes in the Endoplasmic Reticulum ....................... 8
  1.2.3 Accessory Molecules Involved in Class I Biogenesis ............................................... 13
    1.2.3.1 The Proteasome and Peptide Generation ......................................................... 13
    1.2.3.2 TAP, Tapasin and Peptide Transport .................................................................... 15
    1.2.3.3 Molecular Chaperones Involved in Class I Biogenesis .................................... 18
1.3 The Protein Folding and Molecular Chaperones in the Cytoplasm and ER ................ 20
  1.3.1 Common Aspects of Intracellular Protein Folding .................................................... 20
  1.3.2 Molecular Chaperones in the Cytoplasm ..................................................................... 24
    1.3.2.1 Hsp70 Family ....................................................................................................... 24
    1.3.2.2 Hsp60 Family ....................................................................................................... 27
  1.3.3 The Protein Folding in the ER .................................................................................... 30
    1.3.3.1 Signal Sequence and Translocation ...................................................................... 31
    1.3.3.2 Glycosylation ....................................................................................................... 32
    1.3.3.3 Basis of Oxidative Protein Folding ....................................................................... 36
1.3.4 Molecular Chaperones in the ER ................................................................................ 37
  1.3.4.1 Classical ER Molecular Chaperones ................................................................. 40
    1.3.4.1.1 BiP .................................................................................................................... 40
    1.3.4.1.2 GRP94 ............................................................................................................. 43
  1.3.4.2 Protein Disulfide Isomerase (PDI). ......................................................................... 47
1.3.4.3 CNX and CRT: Protein Structure and Function ......................................... 52
  1.3.4.3.1 CNX ................................................................. 52
  1.3.4.3.2 CRT ................................................................. 61
1.4 Models of CNX and CRT Functions ................................................................ 66
  1.4.1 Lectin Only Model ................................................................................ 66
  1.4.2 Dual Binding Model .............................................................................. 70
1.5 Lectins ........................................................................................................... 71
1.6 Experimental Objectives ............................................................................... 75

2. CHAPTER II: FUNCTIONAL RELATIONSHIP BETWEEN CALRETICULIN, CALNEXIN AND THE ENDOPLASMIC RETICULUM LUMINAL DOMAIN OF CALNEXIN ................................................................................................ 77

  2.1 Introduction .................................................................................................... 78
  2.2 Materials and Methods ................................................................................ 79
    2.2.1 Cell Lines and Antibodies .................................................................... 79
    2.2.2 Construction of Calnexin Mutants and Expression in Drosophila Cells .... 80
    2.2.3 Construction of HA-tagged CNX and Ck1' mutants and Expression in L Cells ........................................................................................................... 84
    2.2.4 Metabolic Labeling, Immunoprecipitation, and Gel Electrophoresis ..... 85
  2.3 Results ........................................................................................................... 86
    2.3.1 CRT Can Substitute for CNX in Murine Class I Biogenesis .................... 86
    2.3.2 Association of Truncated Forms of CNX with H Chain Substrates ....... 96
    2.3.3 Quality Control Function of Truncated Forms of CNX ....................... 98
    2.3.4 Molecular Chaperone Functions of Truncated Forms of CNX ............. 101
    2.3.5 Substrate Specificities of CNX, CNX Truncation Mutants, and CRT .... 103
  2.4 Discussion ...................................................................................................... 107

3. CHAPTER III: THE LECTIN-CHAPERONE CALNEXIN UTILIZES POLYPEPTIDE-BASED INTERACTIONS TO ASSOCIATE WITH MANY OF ITS SUBSTRATES IN VIVO ........................................................................................................ 113

  3.1 Introduction .................................................................................................... 114
  3.2 Material and Methods .................................................................................. 116
    3.2.1 Cell Lines and Antibodies .................................................................... 116
    3.2.2 Construction of N-Glycosylation Mutants and Expression in L Cells .... 117
    3.2.3 Metabolic Radiolabeling and Immunoprecipitation ......................... 118
    3.2.4 Glycerol Density Gradient Centrifugation .......................................... 119
    3.2.5 Immunoblotting ..................................................................................... 119
3.3 Results
3.3.1 CNX Associates with Many Substrate Proteins when the Formation of Monoglucosylated Oligosaccharides is Blocked.................................................................120
3.3.2 Differential Effects of Castanospermine on defined CNX-Glycoprotein Complexes......................................................................................................................124
3.3.3 Calnexin Associates with Nonglycosylated Class I Molecules in vivo................131
3.4 Discussion............................................................................................................135

4. CHAPTER IV: THE ROLE OF ERp57 IN DISULFIDE BOND FORMATION DURING FOLDING AND ASSEMBLY OF THE MAJOR HISTOCOMPATIBILITY COMPLEX CLASS I IN DROSOPHILA EXPRESSION SYSTEM........................................................................................................141

4.1 Introduction..........................................................................................................142
4.2 Material and Methods........................................................................................144
  4.2.1 Cell Lines and Antibodies.............................................................................144
  4.2.2 Construction of ERp57 and its Expression in Drosophila Cells......................145
  4.2.3 Metabolic Labeling, Immunoprecipitation, and Gel Electrophoresis............146
4.3 Results..................................................................................................................147
  4.3.1 Expression of ERp57 in Drosophila Cells.....................................................147
  4.3.2 Co-transfected ERp57 Does Not Affect Disulfide Bond Formation of Class I Molecules in Drosophila Cells.................................................................151
  4.3.3 ERp57 Does Not Modulate the Molecular Chaperone Function of CNX.................................155
4.4 Discussion............................................................................................................160

5. CHAPTER V: DISCUSSION......................................................................................165

5.1 Summary of Results............................................................................................166
5.2 Molecular Chaperones in Class I Biogenesis.....................................................167
5.3 Transport of MHC Class I Molecules from the ER to the Cell Surface..............175
5.4 Degradation of MHC Class I Molecules ..........................................................176
5.5 Quality Control Based on Glycan Trimming.....................................................177
5.6 CNX/CRT Dual Binding Model ......................................................................179
5.7 Final Remarks....................................................................................................184

6. CHAPTER VI: REFERENCES...................................................................................186
LIST OF FIGURES AND TABLES

Figure 1. Schematic diagram of a class I MHC molecule ........................................... 5
Figure 2. A schematic view of the assembly of class I molecules in the ER .................. 10
Figure 3. The domain structure of Hsp70 ................................................................. 26
Figure 4. Model of the GroEl-GroES reaction cycle .................................................. 29
Figure 5. Structure of the oligosaccharide added to proteins in the ER ....................... 34
Figure 6. A proposed structural model for the ER Hsp90 family member, GRP94 ............ 45
Figure 7. Structure of the ERp57 ............................................................................... 49
Figure 8. Functional sites in CNX and CRT .............................................................. 60
Figure 9. A. Lectin-only model, B. Dual binding model ............................................. 69
Figure 10. CNX and CRT mutants ........................................................................... 83
Figure 11. Effects of CNX and CRT on the assembly of K^b H chains with B2m in Drosophila cells .......................................................... 90
Figure 12. Folding of D^b heavy chains in the presence of CNX and CRT .................. 94
Figure 13. CRT and CNX stabilize partially folded class I H chains ......................... 95
Figure 14. Expression of CNX and its truncated forms in Drosophila cells ................. 97
Figure 15. Association of truncated forms of CNX with class I H chains ................. 99
Figure 16. Intracellular transport of K^b-B2m heterodimers in the presence of truncated forms of CNX .......................................................... 100
Figure 17. Effects of truncated forms of CNX on K^b assembly with B2m ................. 102
Figure 18. Effect of truncated forms of CNX on the stabilization of class I H chains ........ 104
Figure 19. Effect of transmembrane and cytoplasmic segments on the substrate specificities of CNX and CRT ............................................. 106
Figure 20. Association of newly synthesized proteins with CNX in the absence of glucosidase activity ................................................................. 122
Figure 21. Castanospermine inhibits CNX interaction with K\textsuperscript{b} H chains in EL4 cells but not in Drosophila cells .......................................... 126
Figure 22. HLA-B27 and invariant chain remain associated with CNX in CAS-treated C1R-B27 cells ................................................................. 130
Figure 23. Association of K\textsuperscript{b} H chain glycosylation mutants with CNX .................. 134
Figure 24. Expression of ERp57 in Drosophila cells ............................................. 148
Figure 25. Co-immunoprecipitation of H chains with CNX in the presence of ERp57 ............................................................................................................. 150
Figure 26. Unassembled class I H chains in Drosophila cells contain 1 or 2 intramolecular disulfide bonds ................................................................. 152
Figure 27. Class I K\textsuperscript{b} H chains - \textsuperscript{b2}m heterodimers in Drosophila cells contain 2 intramolecular disulfide bonds .............................................. 154
Figure 28. Effect of ERp57 on the formation of disulfide bonds in the D\textsuperscript{b} H chains .............................................................................................................. 156
Figure 29. Folding of D\textsuperscript{b} heavy chains in the presence of CNX and ERp57 .......... 157
Figure 30. Effects of ERp57 on K\textsuperscript{b} assembly with \textsuperscript{b2}m .................................. 159

Table I. Summary of chaperone diversity and function ........................................... 22
Table II. Molecular chaperones and folding enzymes in the ER .............................. 38
Table III. Table of CNX and CRT \textit{in vivo} substrates ..................................................... 53
LIST OF ABBREVIATIONS

\( \beta_{2m} \)  \( \beta_2 \)-microglobulin
CAS  castanospermine
CNX  calnexin
CRT  calreticulin
CFTR  cystic fibrosis transmembrane conductance regulator
CS  citrate synthase
CTL  cytotoxic T lymphocyte
endoH  endoglycosidase H
ER  endoplasmic reticulum
Glc  glucose
GlcNAc  N-acetylg glucosamine
Grp  glucose regulate protein
HA  influenza hemagglutinin
H chain  heavy chain of class I histocompatibility molecule
Hsp  heat shock protein
INF-\( \gamma \)  interferon-\( \gamma \)
Man  mannose
MAb  monoclonal antibody
MDH  malate dehydrogenase
MHC  major histocompatibility complex
NK  natural killer cell
NP  nucleoprotein
PAGE  polyacrylamide gel electrophoresis
PDI  protein disulfide isomerase
SBA  soybean agglutinin
TCR  T cell receptor
TAP  transporter associated with antigen processing
Tm   transmembrane
UGGT UDP-glucose: glycoprotein glucosyltransferase
VSV  vesicular stomatitis virus
 CHAPTER I
Introduction

There is no form of prose more difficult to understand and more tedious to read than the average scientific paper

-Francis Crick-
1.1 MHC Class I Molecules and the Immune Response to Viruses and Tumors

In the immune response, cytotoxic T lymphocytes (CTL) and natural killer (NK) cells are responsible for recognition and clearance of cells infected with intracellular pathogens or cells undergoing malignant transformation. To allow CTL and NK cells to continuously survey the interior of cells for the presence of pathogens and infectious agents, an elegant system has evolved. The final products of this process are the class I molecules of the major histocompatibility complex (MHC). The class I molecules are found on nearly every cell in higher eukaryotes (Klein, 1975). They were discovered as proteins involved in rejection of foreign or so-called non-self tissues. In humans, class I molecules are encoded by three loci, HLA-A, B, C. Similarly, mice have three class I loci, named H-2K, D, and L (Hansen et al., 1993). The MHC alleles are designated by numbers in humans and by lower case letters (e.g. b, d, k) in mice. At the cell surface class I molecules are recognized by two types of receptors, the T-cell receptor (TCR) on CTL and NK receptors (NKR) on NK cells. CTL use their clonally expressed and somatically rearranged TCR to monitor the cell surface for intracellular peptide fragments in complex with class I molecules (Zinkemagel and Doherty, 1974). In contrast, the specificity of some NK receptors is predominantly influenced by MHC-class I itself and is unaffected by the particular peptide bound to the class I molecule.

CTL and NK cells play complementary roles in the cytolytic immune response. Activation of the CTL response requires recognition of the class I molecules presenting viral or tumor antigens on affected cell surfaces whereas NK cell action is inhibited by interaction of their receptors with MHC class I (Parham, 1996). Hence, the proper expression of MHC class I molecules is critical in the host response to viral infection or in the in vivo control of cancer. However, a number of studies show that tumor or virally infected cells reduce the amounts of class I molecules they display by interfering with
intracellular processes required for the formation of the functional class I complex (reviewed in Alcami and Koszinowski, 2000).

At the cell surface, the class I molecule is comprised of a highly polymorphic heavy (H) chain non-covalently associated with β2-microglobulin (β2m), and the antigenic peptide, generally 8 or 9 residues long. The expression of class I molecules on the cell surface is the outcome of a very complex and highly regulated process that takes place in the compartment termed the endoplasmic reticulum (ER). Within this compartment the class I H chains are synthesized and they undergo structural changes that allow them to bind β2m and peptides. Peptides, bound by class I molecules, are generated from endogenously synthesized antigens by the proteasome in a process termed antigen processing. These cytosolic peptides are delivered to the ER by the transporter associated with antigen processing (TAP). The fully assembled trimeric complexes are transported to the cell surface where they display the bound peptides for recognition by CTLs, in a process named antigen presentation.

Several recent studies suggest that tumor cells or viruses have developed strategies to interfere with almost every major step in the biogenesis of class I molecules: protein synthesis, peptide production and delivery to the ER, subunit assembly and transport to the cell surface (Hengel et al., 1997; Lehner et al., 1997; Ziegler et al., 2000). Understanding the molecular mechanisms involved in the generation of stable class I complexes would not only provide insight into a basic immunological process, but would be essential for developing new approaches to immunotherapies that exploit CTL and NK cells. Furthermore, understanding how class I molecules are assembled would contribute to our knowledge about complex biochemical processes involved in the formation of multi-subunit proteins. It was from this perspective that the involvement of different molecular chaperones in class I folding and assembly was explored. The mode of interaction of two of these molecular chaperones, termed calnexin (CNX) and calreticulin (CRT), was investigated. To place this work into the proper context, I will describe in this chapter the basic features of protein folding and how this information applies to the
biogenesis of class I molecules. But first, let me begin with reviewing the structure and biogenesis of class I molecules.

1.2 The Structure and Assembly of MHC Class I Molecules

1.2.1 The Structure of MHC Class I Molecules

The class I H chain of 44 kDa in humans or about 47 kDa in mouse consists of three extracellular domains, designated α1(N terminal), α2 and α3, a transmembrane region and a cytoplasmic tail (Fig.1A) (Nathenson et al., 1981). The three extracellular domains each contain about 90 amino acids and can be cleaved from the surface with the proteolytic enzyme papain. The homologous α1 and α2 domains are polymorphic and are involved in the formation of the single peptide binding site. The α3 domain and β2m are relatively conserved and show amino acid sequence similarity to immunoglobulin constant domains.

The H chain extends from the α3 segment into a short connecting region and then into a stretch of 25 predominantly hydrophobic amino acids (Fig.1A). This region traverses the lipid bilayer, most probably in an α helical conformation, and anchors the MHC molecule in the membrane. As with all transmembrane proteins, the hydrophobic sequence is immediately terminated at its carboxy terminal end with a cluster of basic amino acid residues that are thought to interact with the phospholipid head group of the inner leaflet of the membrane bilayer. Some class I H chains contain a Cys residue within the hydrophobic sequence that may be modified by esterification with myristic acid, however the significance of such a covalent modification is unclear.

The cytoplasmic domain is 30-40 residues long. Several specific features are highly conserved among different MHC class I molecules. All class I H chains contain amino acid residues that form consensus phosphorylation sites for cyclic adenosine monophosphate (cAMP)-dependent protein kinase A and for Src tyrosine kinase. All class I H chains contain in their carboxy termini a glutamine residue that is a suitable
Figure 1. A. Schematic diagram of a MHC class I molecule. NH$_2$ and COOH refer to amino and carboxy termini of the two polypeptide chains, class I H chain ($\alpha$1, $\alpha$2 and $\alpha$3) and $\beta_2$-microglobulin ($\beta_2$m), respectively, S-S to intrachain disulfide bonds; CHO to carbohydrate. B. A ribbon diagram of the soluble extracellular fragment of H-2K$^b$ is shown. The three domains of the H chain ($\alpha$1, $\alpha$2 and $\alpha$3) in orange make a contact with $\beta_2$m (green). The $\alpha$1 and $\alpha$2 domain form the peptide binding groove. The peptide is depicted in gray. C. Side view of the peptide binding groove of H-2K$^b$. Asn at positions 86 and 176 are depicted in green, Cys at positions 101 and 164 are red.
substrate for transpeptidation by the enzyme transglutaminase. The functional
significance of these structural features is unknown, but they may play a role in regulating
the interactions of class I molecules with other membrane proteins or with cytoskeletal
elements.

The β2m polypeptide of 12kDa is a non-polymorphic protein in man, but is
dimorphic in mice (a single amino acid change at position 85). This molecule also
associates with a number of other class I-related structures, for example the products of
the CD1 genes and the Fc receptor that mediates the uptake of IgG from milk in intestinal
cells of neonatal rat.

The three-dimensional structures of the extracellular portions (α1, α2, α3
and β2m) of several class I molecules have been elucidated by X-ray diffraction including
the human molecules HLA-A2 (Bjorkman et al., 1987; Collins et al., 1994; Saper et al.,
1991), Aw68 (Garrett et al., 1989; Guo et al., 1992; Madden et al., 1991; Silver et al.,
1992), B27 (Madden et al., 1991; Madden et al., 1992), B35 (Smith et al., 1996b), B53
(Smith et al., 1996a) and the mouse molecules K^b (Fremont et al., 1992; Fremont et al.,
1995; Zhang et al., 1992), D^b (Young et al., 1994), L^d (Balendiran et al., 1997; Speir et al.,
1998) and D^d (Achour et al., 1998). From these studies, a detailed picture of the structure
of class I molecules has emerged. Since the murine H-2K^b molecule was used as a model
to study the function of ER molecular chaperones in this thesis, I will illustrate the
general principles of class I structure by referring specifically to this complex.

The H-2K^b H chain has three structural elements α1, α2, and α3 (Fig.1B). The α1
and α2 domains fold into a single, membrane-distal peptide binding domain which
consists of a platform of eight anti-parallel β strands on top of which run two anti-
parallel α helices. A groove approximately 25-30 Å long, 12 Å wide, and 11 Å deep,
separates the α helices of the α1 and α2 domains (Bjorkman et al., 1987). The ends of
this peptide binding groove are closed and the N- and C-termini of bound peptides are
buried at the ends of the groove through hydrogen-bond interactions with residues
conserved among all class I allelic forms. The majority of polymorphic residues are positioned pointing into the groove, for peptide contact, or out from the helices, to contact the T cell receptor (Zhang et al., 1992). The groove has a number of subsites formed from ridges and pockets with which amino acid side chains could interact. Amino acid variations within the groove vary the positions of the pockets, providing a structural basis for differences in peptide binding affinity to different class I alleles. Pooled sequencing of peptides bound to the H-2K\textsuperscript{b} H chain reveals a strong preference for octamers with Tyr or Phe at position 5, Leu or Met at position 8, and to a lesser extent a Tyr at position 3 (Falk et al., 1991). Peptides bind within the \(\alpha 1/\alpha 2\) groove in an extended conformation, with a defined N- and C-terminal orientation.

The H-2K\textsuperscript{b} H chain has an N-linked complex oligosaccharides attached at Asn 86 and Asn 176. Asn 86 is in a loop connecting \(\alpha 1\) to \(\alpha 2\). Residue 176 is located in \(\alpha 2\) at a position structurally homologous to Asn 86 in \(\alpha 1\). The presence of carbohydrate at position 176 increases the approximate symmetric nature of the \(\alpha 1-\alpha 2\) `intramolecular dimer' (Fig. 1C). Residue 256 which has attached glycans only in some murine class I molecules is located on the loop between strands 5 and 6 in \(\alpha 3\) and is accessible to solvent in the H-2K\textsuperscript{b} H chain. The disulfide bond (Cys 101-Cys 164) in the \(\alpha 2\) domain connects the N-terminal \(\beta\) strand to the \(\alpha\) helix of the \(\alpha 2\) domain (Fig. 1C).

The \(\alpha 3\) and \(\beta_{2m}\) domains are both \(\beta\)-sandwich structures composed of two antiparallel \(\beta\)-pleated sheets, one with four \(\beta\)-strands and one with three \(\beta\)-strands, connected by a disulfide bond. The non-covalently bound \(\beta_{2m}\) interacts both with the \(\alpha 3\) domain and sits under the \(\alpha 1\) and \(\alpha 2\) domains. The position and interaction of the \(\beta_{2m}\) subunit relative to the H chain varies between different class I molecules, as does the position of the \(\alpha 3\) domain, which can even be removed proteolytically without causing any major changes to MHC conformation (Collins et al., 1995).
1.2.2 The Assembly of Class I Complexes in the Endoplasmic Reticulum

The assembly of MHC class I molecules has much in common with the folding and assembly of other multisubunit glycoproteins in the ER (for details see section 1.2). Class I MHC assembly is now understood to involve the interplay of multiple intra- and intermolecular events in a defined chronological order which ensure continual flow of class I molecules to the cell surface. Many different proteins that reside in the ER, such as calnexin (CNX), calreticulin (CRT), ERp57, tapasin and TAP have been found to be integral to class I assembly (Fig. 2).

During protein synthesis N-linked oligosaccharides are added to H chains as they are co-translationally imported into the ER. Human HLA-A,-B,-C molecules have an N-linked, complex oligosaccharide attached to Asn at position 86. This residue is conserved among all class I H chains so far sequenced (primates, mouse, rat, cat, rabbit, cow, horse, cheetah, squirrel, pig, horse, frog, and salmon) (Lian et al., 1998). Murine H-2 molecules have an additional glycosylation site at Asn 176 and H-2Ld, H-2Db and H-2Kd molecules have a third at Asn 256 (Lian et al., 1998). Since the glycosylation site at residue 176 was not found among any of the nonmurine MHC class I molecules, this indicates that site 176 likely arose after the divergence of mice from other species and it became fixed in all murine MHC class I genes.

The roles N-linked glycans play in class I functions are still not well understood. However, carbohydrates on residue 86 are thought to be important for the binding of class I H chains to the chaperones CNX and CRT (Parham, 1996; Zhang and Salter, 1998). Degen et al (1991) have shown that the oligomeric assembly of MHC class I proteins occurs while the H chain is bound to CNX. Also, N-glycan chain length seems to be an important determinant for cell surface expression of newly synthesized MHC class I proteins (Bennett and Kearse, 1999).

During or shortly after synthesis, the class I H chain forms one disulfide bond in the α2 domain and a second in the α3 domain. The α2 disulfide bond forms a loop of
Figure 2. A. A schematic view of the assembly of murine class I molecules in the ER. Newly synthesized class I H chain binds rapidly, possibly co-translationally to CNX and remains associated with CNX during assembly with β2m. The H chain-β2m-CN X complex then interacts with tapasin and the TAP transporter. Upon peptide binding, TAP and tapasin dissociate from the trimeric complex, however CNX remains associated with the now fully assembled class I complex. What regulates the release of CNX is not known. ERp57 is detected in association with class I H chain at all stages of assembly.

B. Schematic of the assembly of human class I molecules in the ER. The H chain associates with CNX. Upon β2m binding CNX is replaced by CRT. The CRT associated H chain-β2m heterodimer interacts now with TAP via tapasin. Peptides are loaded onto TAP-associated class I molecules and the mature class I trimer is rapidly transported out of the ER.
about 63 amino acid residues between Cys101 and Cys164. In the α3 domain the loop is 56 amino acids and is formed between Cys residues at positions 203 and 259.

Disulfide bond formation in the H chain is initiated before β2m and peptide binding \textit{in vivo}. Several lines of evidence support this view. When Cys at position 101 was substituted by serine by site directed mutagenesis in the L^d molecule, thereby blocking the formation of the α2 domain disulfide bond, this mutant accumulated in the ER and did not bind β2m (Smith et al., 1995). Furthermore, only a low number of mutated L^d H chains were able to bind peptides added at very high concentrations to cell lysates. Also, mutant HLA-A2 molecules in which the residues Cys101 or Cys164 were mutated to Ser or to Ala, respectively, were poorly loaded with peptides in the ER, poorly transported from the ER and were not expressed at high levels on the plasma membrane (Warburton et al., 1994). When infected by influenza virus, cells expressing the HLA-A2 mutants were either poorly recognized (Cys164 mutant) or not at all recognized (Cys 101) by an HLA-A2 restricted cytotoxic T cell line (Warburton et al., 1994).

The importance of disulfide bond formation to class I assembly raises the question of whether it is a regulated process. A candidate protein that might influence this process is the thiol oxidoreductase ERp57. ERp57 has recently been shown to bind to class I molecules (Hughes and Cresswell, 1998; Lindquist et al., 1998).

H chain associates with β2m within 4 min after translation (Neefjes et al., 1993). β2m is required for normal class I assembly and surface expression. The role of β2m in class I surface expression was supported by the discovery of β2m deficient cell lines (Allen et al., 1986; Bix and Raulet, 1992). Class I molecules have not been found on the surface of β2m deficient human Daudi cells (Bix and Raulet, 1992). Likewise, on murine β2m deficient cells, H chains are undetectable by flow cytometry, with the exception of low levels of D^b and L^d (Allen et al., 1986). These phenotypes were corrected by re-expression of β2m in these cells.
Several lines of evidence support the involvement of \( \beta_2m \) in correct folding and assembly of class I molecules. First, on the murine \( \beta_2m \) deficient cell line R1E/D\( ^b \), the few D\( ^b \) molecules that reached the surface were misfolded (Solheim, 1999). Second, when a modified version of \( \beta_2m \), which was ER-retained due to a C-terminal KDEL sequence, was expressed in R1E/D\( ^b \) cells, D\( ^b \) H chain was well expressed on the surface although no \( \beta_2m \) was present at the cell surface of these transfectants (Solheim et al., 1995). In addition, the D\( ^b \) molecules were properly conformed. Thus \( \beta_2m \) is not required for post-assembly stability, nor for transport, but is crucial at the point of folding in the ER. Third, class I molecules synthesized in the absence of \( \beta_2m \) in a cell-free translation system fail to form one of two internal disulfide bridges (Ribaudo and Margulies, 1992). Based on the reactivity of the \textit{in vitro} synthesized MHC class I molecules with domain-specific monoclonal antibodies it was established that in the absence of \( \beta_2m \) only the disulfide bond in the \( \alpha3 \) domain is formed. Fourth, H-2L\( ^d \) molecules from \( \beta_2m \) expressing cells have both the \( \alpha2 \) and \( \alpha3 \) intrachain disulfide bonds, whereas about 50\% of the H-2L\( ^d \) molecules from \( \beta_2m \) deficient cells have only one of these bonds (Solheim, 1999). All of the free H chains from \( \beta_2m \) expressing cells can bind exogenous peptide and \( \beta_2m \) stably \textit{in vitro}. H chains from \( \beta_2m \) deficient cells, which have both intrachain disulfide bonds, do not bind \( \beta_2m \) and peptide stably \textit{in vitro} (Solheim, 1999).

H chain-\( \beta_2m \) heterodimers can be recovered within 4 min after translation and the peptide joins the complex about 2 min later (Neefjes et al., 1993). Thus in the cell, H chain-\( \beta_2m \) heterodimers are formed first and await peptide for final assembly. The first evidence for class I binding peptides came from experiments showing that peptides from a viral protein added to uninfected cells will stimulate virus-specific class I restricted cell lysis (Zinkernagel and Doherty, 1974). X-ray crystallographic studies of class I molecules confirmed their function as peptide receptors. Finally, class I binding peptides have been identified by their elution from class I molecules (Falk et al., 1991) and by exposing class
I molecules to synthetic peptides (Townsend et al., 1990; Townsend et al., 1989). Class I associated peptides are derived from proteins degraded mainly in the cytoplasm and represent samples of the whole cellular protein content. The majority of peptides that associate with MHC class I are 8-10 residues long, with allotype-specific binding motifs containing up to three anchor positions (Falk et al., 1991). Peptide length is critical for optimal binding and antigenic activity. Peptide binding by H chain-β2m heterodimers is essential for formation of stable MHC class I complexes. Lack of peptides results in poor cell surface expression (Townsend et al., 1989).

1.2.3 Accessory Molecules Involved in Class I Biogenesis

1.2.3.1 The Proteasome and Peptide Generation

The observation that in the ER H chain-β2m heterodimers bind peptides originating mainly from cytosolic proteins implied that a protein degradation system in the cytoplasm generates class I restricted peptides. Proteasomes are complex multisubunit proteases in the cytosol and nucleus responsible for the degradation of most cellular proteins (Baumeister et al., 1998). They are abundantly expressed in eukaryotic cells and exist also in archaeabacteria and some eubacteria. All types of proteasome complexes share the proteolytically active core, the 20S proteasome, which is composed of 7 different α and 7 different β subunits, arranged in four rings surrounding a central channel in which proteolysis takes place. The two outer rings, which play mainly a structural role, contain the α-type subunits. The two inner rings consist of the β-type subunits with active sites displayed at the inner wall of the central chamber. These subunits give rise to the three major peptidase activities. One site cleaves preferentially on the carboxylic side of large hydrophobic residues (chymotrypsin-like), one after basic ones (trypsin-like) and another after acidic amino acids (post acidic or caspase-like sites). The principal function of the complex is to degrade ubiquitin-conjugated proteins progressively to oligopeptides (Baumeister et al., 1998; Coux et al., 1996).
The 20S complex associates with two 19S regulatory cap complexes. The 19S structures are each composed of approximately 20 subunits and contain polyubiquitin binding sites that allow recognition of the ubiquitin marked proteins (Baumeister et al., 1998; Coux et al., 1996). The polypeptide ubiquitin is first conjugated to the ε-amino group on proteins and the polyubiquitin chain that forms on the protein serves as a molecular tag that targets the protein for degradation. The ends of 20S core particle can also associate with PA28, an 11S subunit (200kDa) which is induced by interferon γ (INF-γ). PA28, like the 19S particle, can markedly enhance the catalytic rate of the proteasome’s peptidase sites.

Several lines of evidence support the involvement of the proteasome in antigen processing. First, membrane permeable peptide aldehydes, inhibitors of the proteasome, were shown to block the class I presentation of peptides from microinjected proteins (Harding et al., 1995; Rock et al., 1994; Sijts et al., 1996). In the presence of proteasome inhibitors, there is a profound reduction in cell surface expression of many class I alleles (Benham et al., 1998; Benham and Neefjes, 1997; Craiu et al., 1997; Glas et al., 1998; Rock et al., 1994). Second, genetic mutants in the ubiquitination pathway showed that defects in targeting of proteins for degradation resulted in decreased antigen presentation (Tanaka and Kasahara, 1998). Furthermore, enhancing the proteasome mediated degradation of proteins through ubiquitination results in more efficient antigen presentation (Townsend et al., 1988). Third, in vertebrates with an adaptive immune system, the proteolytically active subunits of the proteasome have homologs inducible by INF-γ, a cytokine which coordinately regulates the expression of other antigen presentation molecules. These inducible subunits are termed low molecular mass peptide (LMP)2, multicatalytic endopeptidase complex-like (MEL)1 and LMP7. LMP2 and LMP7 are encoded in the major histocompatibility complex (Tanaka and Kasahara, 1998). LMP7-deficient mice show reduced surface expression of MHC class I (presumably because of inefficient or inappropriate peptide generation) (Schmidtke et al., 1998; Stohwasser et al., 1996) and reduced presentation of at least one antigen (Fehling et al.,
Similarly, mice lacking LMP2 are less efficient in presenting specific peptides such as the viral influenza nucleoprotein \textit{in vitro} and mount a poor CTL response to this antigen \textit{in vivo} (Van Kaer et al., 1994). The IFN-γ induced subunits are not essential for MHC class I antigen presentation, but they do enhance this process.

Although the proteasome is clearly the site for generating most class I presented peptides, other proteases may also generate antigenic peptides that will be presented on MHC class I molecules. Evidence for a proteasome-independent generation of MHC class I peptides comes from an analysis of the presentation of peptides that are derived from leader sequences. Leader sequences are cleaved from proteins as they are cotranslationally transported into the ER by the ER-resident signal peptidase, which releases a short peptide from the protein. There are several instances of peptides from leader sequences being presented by HLA-A2 and the non-classical MHC class I HLA-E and Qa (Henderson et al., 1992). Furthermore, \textit{in vitro} assays using microsomes revealed that peptides are trimmed in the ER (Roelse et al., 1994). Finally, Elliott et al. (1992) have demonstrated that in cells long influenza NP fragments targeted to the ER were trimmed 40 amino acids from the N-terminus or 120 residues from the C-terminus to generate peptides capable of binding to class I.

\subsection*{1.2.3.2 TAP, Tapasin and Peptide Transport}

Since antigenic peptides produced in the cytosol have to meet with H chain-β2m heterodimers in the ER an essential protein for MHC class I assembly is a peptide transporter. The transporter associated with antigen processing (TAP) belongs to a large ABC (ATP binding cassette) family of transporters found in species ranging from bacteria to humans (Higgins, 1992). A role for TAP in the transport of class I binding peptides was originally supported by the discovery of two proteins, encoded by the MHC-linked genes. TAP1 and TAP2, homologous to other transport proteins (Deverson et al., 1990; Monaco et al., 1990; Powis et al., 1992; Spies et al., 1990; Trowsdale et al., 1990).
Functional data supporting a role for TAP in class I assembly came from mutant cell lines, lacking either of the TAP proteins (Attaya et al., 1992; Kelly et al., 1992; Spies et al., 1992). These cells have decreased cell surface expression of class I molecules and unstable H chain-β2m dimers accumulate intracellularly. This phenotype was corrected by re-expression of TAP protein in these cells (Attaya et al., 1992). Furthermore, TAP1 knockout mice have a defective peptide supply which is manifested by reduced class I expression, diminished number of CD8+ cells and inability to present cytosolic peptides (Van Kaer et al., 1992). Finally, a physical association between MHC class I molecules and TAP was observed in human and murine systems (Ortmann et al., 1994; Suh et al., 1994). This association was maintained after solubilization in the detergent digitonin and was dependent on the formation of class I H chain-β2m dimers. Proximity of the class I molecules to TAP has been suggested to enhance peptide binding by minimizing the distance peptide would have to diffuse in the ER before encountering a class I molecule (Suh et al., 1996).

Competition studies (Heemels et al., 1993) and direct translocation assays (Momburg et al., 1994) have shown that TAP prefers peptides of approximately the size usually found associated with class I molecules. The efficiency of translocation drops significantly for substrates shorter than 8 and longer than 14 amino acid residues (Schumacher et al., 1994). Within the optimal size range, which is nine amino acids, peptides differ in their ability to be transported by the TAP (Heemels et al., 1993; Momburg et al., 1994; Schumacher et al., 1994).

The TAP transporter is a heterodimer consisting of two subunits, TAP1 and TAP2 (Androlewicz et al., 1993; Spies et al., 1992), that is localized to the ER and cis-Golgi (Kleijmeer et al., 1992). Each subunit is composed of an N-terminal hydrophobic region containing multiple transmembrane segments and a C-terminal hydrophilic region which is cytosolic. By using radiolabeled peptides and photoactivable crosslinkers, several groups have demonstrated that both TAP1 and TAP2 contribute to the peptide binding site (Androlewicz et al., 1993; Gileadi and Higgins, 1997; Nijenhuis et al., 1996).
The interaction between TAP and class I molecules is mediated by an accessory molecule termed tapasin (TAP-associated glycoprotein). Compelling evidence for the importance of the TAP-class I association, and the critical role played by tapasin, came from the discovery of a human cell line, called .220, which lacked tapasin (Sadasivan et al., 1996). In this cell line, MHC class I molecules do not associate with TAP, resulting in inefficient peptide loading and impaired surface expression of MHC class I molecules. Re-expression of tapasin cDNA corrects the phenotype of .220.B8, restoring the association of class I molecules with TAP, cell surface class I expression, and recognition of virus-infected .220.B8 cells by class I-restricted CTL (Ortmann et al., 1997). Analysis of the β2m negative mutant cell line, Daudi, showed that tapasin can associate with TAP in the absence of assembled class I molecules (Bangia et al., 1999; Solheim et al., 1997). Also, in a TAP-negative cell line, .174, complexes exist containing H chain-β2m dimers and tapasin. This evidence suggested that tapasin has independent binding sites for H chain-β2m dimers and TAP and that it forms a bridge between them (Sadasivan et al., 1996). Quantitative analysis of the components of the TAP-tapasin-H chain-β2m complex, purified from human B-cell lines using an anti-TAP 1 affinity column, suggested that it contains approximately four tapasin molecules per TAP dimer and that each tapasin molecule associates with a single H chain-β2m dimer. Furthermore, class I-tapasin complexes, whether or not they were associated with TAP, were found to contain CRT and later ERp57 (Hughes and Cresswell, 1998).

Tapasin is a type I transmembrane glycoprotein which, like MHC molecules, is a member of the immunoglobulin superfamily. The mature protein has 428 amino acids with a single N-linked glycosylation site at position 233. The C-terminal cytosolic sequence has lysine residues at positions -3, -4, and -5 consistent with known cytoplasmic ER retrieval signals that have lysine residues at -3, and -4 or -3 and -5. Of the 392 lumenal residues of the mature protein, 50 are prolines (Ortmann et al., 1997).
Recent evidence suggest that tapasin has other functions in addition to bridging class I and TAP (Sadasivan et al., 1996; Solheim et al., 1997). Tapasin is regarded as a class I specific chaperone which retains class I molecules in the ER, protects them from degradation and prolongs their peptide binding capacity (Schoenhals et al., 1999). Furthermore, tapasin is thought to influence the population of peptides that bind in the class I groove (Suh et al., 1999; Yu et al., 1999) and to increase the level of TAP in the cell (Lehner et al., 1998). Finally, experiments with recombinant soluble tapasin, that associates with class I molecules but no longer binds TAP or increases TAP levels, showed that tapasin is sufficient to facilitate loading and assembly of class I molecules (Lehner et al., 1998). This has called into question a requirement for physical association of class I with TAP for efficient presentation.

1.2.3.3 Molecular Chaperones Involved in Class I Biogenesis

As with other proteins that are co-translationally translocated into the ER the assembly of a functional trimeric class I complex is now recognized to be a highly regulated process. This process termed ‘quality control’ involves a number of chaperones and accessory proteins and is designed to ensure that only correctly folded H chain-β2m dimers are loaded with peptide ligands for transport to the cell surface (Fig.2).

The ER chaperone BiP, an Hsp 70 homolog, is believed to interact with proteins during their translation to ensure the unidirectional movement of the protein through the ER membrane. BiP has been shown to bind to free class I H chains, which is consistent with a potential role in the translocation of class I molecules. BiP exhibits species specificity; it interacts with human but not mouse class I molecules (Degen et al., 1992; Nossner and Parham, 1995).

The dominant chaperone associated with newly synthesized class I H chains is CNX (Degen et al., 1992; Jackson et al., 1994). CNX is a lectin, with specificity for monoglucosylated N-linked glycans, that facilitates the folding and assembly of class I H chains with β2m. CNX also retains class I assembly intermediates in the ER. CNX is
believed to interact with its substrates in a two step process, an initial weak interaction with monoglucosylated sugars brings the substrate into proximity with CNX and allows binding to protein determinants. Studies in the mouse indicate that the MHC class I-CNX interactions persist after β2m association (Degen et al., 1992). However, with the exception of HLA-B27 (Carreno et al., 1995), only free human H chains and not H chain-β2m dimers have been shown to bind to CNX (Fig. 2). In human cells, the dominant chaperone associated with H chain-β2m dimers is CRT (Sadasivan et al., 1996). CRT is a soluble homologue of CNX.

The interactions of H chain with CNX and H chain-β2m dimers with CNX or CRT have not been kinetically separated from their interaction with an additional housekeeping molecule called ERp57 (Fig. 2). ERp57 is a member of the thioredoxin family of enzymes, a subset of which resides in the ER. Hammerling and co-workers showed an ERp57 interaction with class I H chain-CN X complexes that are generated early during the MHC class I assembly pathway (Lindquist et al., 1998). In addition, ERp57 is part of the late assembly complexes consisting of class I H chain-β2m dimers, CNX, CRT tapasin and TAP (Lindquist et al., 1998, Cresswell et al., 1999). It is thought that disulfide bond formation is initiated when class I H chains are associated with CNX and completed after β2m binding and CRT association, with ERp57 being the thiol oxidoreductase facilitating their formation. Salter and co-workers (Tector et al., 1997) have shown that disulfide bond formation is only partially completed in CNX-associated human MHC class I H chains.

The chaperone assisted formation of H chain-β2m heterodimers leads to the next step in class I assembly, the binding of peptide ligand. This process is facilitated by association of MHC class I molecules with TAP and tapasin (Cresswell et al., 1999). H chain-β2m heterodimers associate with TAP in both mouse and human cells forming large complexes that contain also the bridging molecule tapasin, the molecular chaperones CNX and/or CRT and the oxidoreductase ERp57 (Lindquist et al., 1998). In the human
system CRT and not CNX was found in association with complexes containing TAP (Sadasivan et al., 1996). The binding of peptides to human class I results in the dissociation of class I complexes from TAP/tapasin and CRT (van Leeuwen and Kearse, 1996b), although in the murine system CNX remains associated with the now fully assembled trimeric complex (Suh et al., 1996). CRT is responsive to class I peptide binding whereas CNX is found associated with the folded, peptide occupied form of class I as well as the unfolded form (Carreno et al., 1995). The role of CNX, CRT and ERp57 in the class I-TAP complex remains unclear.

The fully assembled class I molecules are now transported to the cell surface. There are probably a number of other ER and Golgi proteins which interact with the assembled class I molecules and assist them in the transport to the cell surface, however, such proteins have yet to be discovered.

1.3 Protein Folding and the Functions of Molecular Chaperones in the Cytoplasm and ER

1.3.1 Common Aspects of Intracellular Protein Folding

Protein folding is a process by which the linear information contained in the amino acid sequence of a polypeptide gives rise to the well-defined three-dimensional conformation of the functional protein. The characteristic fold is determined by each amino acid sequence, a concept known as Anfinsen's dogma. Anfinsen's dogma was derived from the spontaneous refolding of pure protein samples after denaturation, which revealed that the necessary information for folding is embedded within the amino acid sequence itself (Anfinsen, 1973).

In the cell, however, the correct folding of many proteins depends on the function of a pre-existing protein machinery, the molecular chaperones. Molecular chaperones are defined as a group of unrelated classes of proteins that mediate the correct folding and assembly of other proteins, but are not themselves components of the final functional
structures (Ellis, 1987). A wide range of molecules are now classified as molecular chaperones, many of them stress proteins. Synthesis of stress proteins is upregulated during exposure to stress such as heat shock, oxidative stress, glucose starvation and osmotic shock. They are found ubiquitously in prokaryotes and in most organelles of eukaryotes such as the cytoplasm, ER, mitochondria, and nucleus (Table 1). The best studied of these are the heat shock protein (Hsp) 60 and Hsp70 families.

Molecular chaperones bind to and stabilize otherwise unstable conformers of another protein and by controlled binding and release, facilitate its correct fate in vivo. be it folding, oligomeric assembly, transport to a particular sub-cellular compartment, or disposal by degradation (Hendrick and Hartl, 1993). They do not contain steric information specifying correct folding, but they prevent incorrect interactions within and between non-native polypeptides. Since molecular chaperones increase the yield but not the rate of folding reaction this distinguishes them from folding catalysts such as protein disulfide isomerases (PDI) and peptidyl-prolyl isomerases (PPI) (Lorimer, 1996). These enzymes accelerate intrinsically slow steps in the folding of some proteins, namely the formation and rearrangement of disulfide bonds in secretory proteins and the cis-trans isomeration of peptide bonds preceding proline residues, respectively.

Folding in vivo depends on molecular chaperones that prevent protein aggregation. This requirement is dictated by the conditions in which newly synthesized proteins undergo folding, i.e., the cell cytosol or the ER is a highly crowded macromolecular environment and folding in the cell must be accomplished in the context of the vectorial synthesis of polypeptide chains on ribosomes (Netzer and Hartl, 1997). The concentration of nascent (ribosome-bound) chains in the cytoplasm of E.coli is as high as ~35 μM, assuming a uniform distribution (Hartl, 1996). Stable folding requires the presence of at least a complete protein domain (usually ~100-200 amino acid residues in length) (Netzer and Hartl, 1997). As the C-terminal ~ 30 amino acids of a
Table 1. Summary of chaperone diversity and function (reviewed in Agashe and Hartl, 2000)

<table>
<thead>
<tr>
<th>Chaperone Family</th>
<th>Prokaryotes</th>
<th>Eukaryotes</th>
<th>Yeast/other eukaryotes</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Name</td>
<td>Function</td>
<td></td>
</tr>
<tr>
<td>HSP70</td>
<td>DnaK</td>
<td>Stress response regulation. Binds to and stabilizes hydrophobic segments in unfolded proteins (in cooperation with DnaJ and GrpE)</td>
<td>Ssαl-4/Hsp72</td>
</tr>
<tr>
<td></td>
<td>Hsc66</td>
<td>Cooperates with a DnaJ-like protein Hsc20</td>
<td>Ssbl-2/Hsc73</td>
</tr>
<tr>
<td>HSP40</td>
<td>DnaJ</td>
<td>A co-factor of DnaK for protein folding, binds unfolded substrates and stimulates DnaK ATPase</td>
<td>Ydj 1/Hsp40</td>
</tr>
<tr>
<td>GrpE</td>
<td>GrpE</td>
<td>ATP/ADP nucleotide exchange factor for DnaK</td>
<td>Mge 1</td>
</tr>
<tr>
<td>Chaperonins</td>
<td>GroEL</td>
<td>Group I chaperonin, cooperates with groES and interacts with ~10-15% of newly synthesized proteins</td>
<td>TRiC or CCT/ TriC or CCT</td>
</tr>
<tr>
<td></td>
<td>GroES</td>
<td>A co-factor for GroEL that forms a heptameric ring complex, functions by capping the GroEL cylinder during a protein folding cycle</td>
<td></td>
</tr>
</tbody>
</table>
Table I. Summary of chaperone diversity and function cont.

<table>
<thead>
<tr>
<th>Small Hsp</th>
<th>lbpA or B</th>
<th>Inclusion body proteins (lbp) that are stress inducible and found associated with inclusion bodies</th>
<th>Hsp 12- Hsp42, Generally stress-inducible, bind and stabilize unfolded proteins for refolding by the Hsp70/chaperonin system</th>
</tr>
</thead>
<tbody>
<tr>
<td>Trigger factor</td>
<td>TF</td>
<td>A 48-kDa protein, associates with the ribosome and can be cross-linked to nascent chains, chaperones nascent chains and can also catalyze prolyl isomerization in non-native proteins with high efficiency</td>
<td></td>
</tr>
<tr>
<td>HSP90</td>
<td>HtpG</td>
<td>Stress-inducible, Hsp90 homologue</td>
<td>Hsp82/Hsp90</td>
</tr>
<tr>
<td>Hsp100/Clp</td>
<td>ClpA/X</td>
<td>Involved in proteolytic degradation with ClpP, which has protease activity and forms double ring structures capped at both ends by rings of either ClpA or ClpX that possess protein unfolding activity. The whole complex can, thus, unfold and degrade misfolded proteins</td>
<td>Hsp104</td>
</tr>
</tbody>
</table>
newly synthesized polypeptide is topologically restricted by the ribosome, a nascent chain remains unfolded until an entire domain has emerged. This large population of aggregation-sensitive nascent chains must be maintained in a folding-competent conformation. This is achieved in the cytoplasm by the co-translational binding of the nascent chain to molecular chaperones such as Hsp70 and DnaJ or in the ER to BiP.

1.3.2 Molecular Chaperones in the Cytoplasm

1.3.2.1 Hsp70 Family

The Hsp70s appear to be tailored to prevent misfolding and aggregation at an early stage when a nascent polypeptide chain is still extended, i.e. either ribosome bound or in a state of transit across a membrane and thus not available for folding. The members of this chaperone family are highly conserved ATPases of relative molecular mass about 70,000. These proteins are found in all eubacteria and eukaryotes, but only in a subset of the Archaea. In eukaryotes, the Hsp70 chaperones are found in the cytosol, nucleus, mitochondria, chloroplasts, and ER.

In both eukaryotes and prokaryotes the Hsp70 chaperones have been shown to bind newly synthesized polypeptide chains (Beckmann et al., 1990; Pfund et al., 1998; Teter et al., 1999; Thulasiraman et al., 1999). This observation suggests that Hsp70s play an essential role in de novo protein folding and membrane translocation. Furthermore, by using peptide libraries and phage display methods, it was established that the various Hsp70s all have binding specificity for hydrophobic peptides of 7 amino acids in length (Blond-Elguindi et al., 1993; Flynn et al., 1991; Rudiger et al., 1997). Hsp70s interact with a wide range of unfolded polypeptides in an extended conformation and ignore the native state of the proteins (Beckmann et al., 1990). More than one Hsp70 may interact simultaneously with an unfolded polypeptide (Feldman and Frydman, 2000). Thus, the
key element of substrate recognition by Hsp70 is an affinity for redundant, hydrophobic structural features that are generally exposed by non-native proteins.

Both constitutive and induced isoforms of Hsp70s require accessory molecules for optimal function (Frydman and Hohfeld, 1997). Although DnaK, the main E.coli Hsp 70, was sufficient to rescue and promote re-folding of heat inactivated RNA polymerase in the presence of ATP in vitro, it did so at a large molar excess (Skowyra et al., 1990). Further work established that DnaJ/Hsp40 and GrpE were essential for proper DnaK/Hsp70 function (Ziemienowicz et al., 1993).

Structurally, Hsp70s have two functional domains (Fig.3A). The C-terminal domain of the protein (~25kDa) contains the peptide binding site and the N-terminal domain ~45kDa contains the ATPase site. Conformational changes in the ATPase domain that are induced by ATP binding and hydrolysis are transmitted to the C-terminal peptide-binding domain causing changes in peptide binding affinity. The peptide binding domain consists of two four-stranded antiparallel β-sheets with a peptide binding cleft and a single α-helical extension that acts as a lid, covering the peptide binding domain.

Mutational, crystallographic and NMR analysis of Hsp70 and its co-factors has provided a detailed picture of the structure-function relationship in these molecules. The Hsp70 reaction cycle in protein folding is best understood for the bacterial Hsp70 homologue DnaK and its cofactors DnaJ and GrpE (McCarty et al., 1995; Szabo et al., 1994). This cycle has been analyzed using denaturant-unfolded firefly luciferase and rhodenase as substrate proteins. Hsp70s achieve their versatile functions by binding and releasing hydrophobic segments of an unfolded polypeptide chain in an ATP-hydrolytic reaction cycle (Fig.3B). In this cycle, binding results in the stabilization of the unfolded state, and controlled release allows progression along folding pathways. Initially, the unfolded polypeptide interacts with the co-chaperone DnaJ. This association is necessary for targeting the substrate protein to DnaK. DnaK binds to the polypeptide in its ATP-bound state, but the interaction with DnaJ stimulates the hydrolysis of ATP by DnaK.
Figure 3. A. The domain structure of Hsp70. N-terminal and C-terminal signal occur in forms located in the mitochondria or ER. The ATPase and peptide binding domains are well conserved among all Hsp70s. The EEVD motif is involved in regulating interactions with the ATPase domain and with co-chaperones. B. Model of the Hsp70 reaction cycle for bacterial DnaK, DnaJ and GrpE (Taken from Hartl, 1996).
Upon ATP hydrolysis a stable ternary complex of unfolded substrate polypeptide and DnaK in its ADP-bound state is formed. The rate-limiting step in this cycle is the release of ADP from DnaK by the nucleotide-exchange factor GrpE. Substrate dissociates from DnaK upon subsequent ATP binding to DnaK. It then has the option of either folding, rebinding to DnaJ and DnaK, or being transferred to another chaperone system such as the chaperonin GroEL, for final folding.

1.3.2.2 Hsp 60 Family

Following their interaction as nascent chains with Hsp70, newly synthesized polypeptide chains must continue to fold in the crowded cytosol. A second group of chaperones identified in the cytosol to assist folding of proteins released from the ribosome, was termed chaperonins or Hsp60. There are two groups of chaperonins; members of the GroEL (or Hsp60) family, and the chaperonins of the TRiC family. The GroEL/Hsp60 proteins are expressed in eubacteria, mitochondria and chloroplasts. They cooperate with a smaller cofactor, GroES (Hsp10). The TRiC (TCP-1 ring complex) is expressed in archaebacteria and in the eukaryotic cytosol.

Pioneering experiments by Horwich et al., (1993), implicated chaperonins as critical players in *in vivo* folding of proteins. Shift of a conditional groEL mutant to a nonpermissive temperature resulted in the disappearance of a number of newly synthesized cytosolic proteins from the soluble cell fraction and increased aggregation of at least nine proteins. Experiments directly analyzing the flux of newly synthesized proteins through GroEL indicated that it transiently associates with approximately 12% of all newly synthesized proteins in *E. coli*; this figure increases 2-3 fold under stress (Hesterkamp and Bukau, 1998). The majority of these substrates range between 10 and 55 kDa and are enriched for a specific subset of approximately 300 polypeptides (Houry et al., 1999). Interestingly, structural analysis of over 50 natural GroEL substrates
revealed a significant preference for proteins composed of multiple α/β domains (Houry et al., 1999). As β sheets are assembled from discontinuous regions of the polypeptide, the binding of these hydrophobic surfaces to GroEL might facilitate the correct packing of strands within the β sheet. The GroEL family of proteins is known to bind to many unfolded polypeptides, but not to their corresponding folded forms. The interaction with the substrate seems to be primarily hydrophobic in nature.

GroEL is a cylindrical complex of ~150 Å height and ~140 Å width that encloses a central cavity ~50 Å wide. It consists of two stacked seven-membered rings made up of 57 kDa subunits. The central cavity represents the polypeptide binding site (Chen et al., 1994; Fenton et al., 1994). Each subunit is composed of the apical domains that forms the flexible opening of the cylinder, and the equatorial domain that contains the ATP-binding site.

GroES is a heptameric ring of 10 kDa subunits, possessing a donut-like shape. Each GroES subunit contains a functionally critical loop region that is involved in the docking of GroES onto GroEL.

A series of mutational studies coupled with the X-ray crystallographic structure of GroEL has led to our understanding of the regions and residues involved in the binding of substrate polypeptides (Fenton et al., 1994). A detailed reaction mechanism has been proposed for GroEL/GroES based on the in vitro folding of rhodanese and ornithine transcarbamoylase (Bukau et al., 2000). GroEL can bind up to two polypeptides, one per ring (Fig.4). The central hole of GroEL is a place where polypeptides are sequestered and protected from possible inadvertent inter-polypeptide aggregation. The binding of GroES to GroEL is asymmetrical and induces dramatic conformational changes in the interacting GroEL ring. Since the hydrophobic binding region of GroEL for GroES overlaps with that for polypeptide binding, GroES binding leads to the displacement of the bound substrate from its binding site and forces the substrate into the hydrophilic cage which is permissive for folding. GroES prevents the substrate protein from diffusing out of the
Figure 4. Model of the GroEl-GroES reaction cycle (Hartl, 1996). Single round of GroEL/GroES-mediated protein folding is initiated by ATP and GroES binding to the \textit{cis}-ring. This reaction causes a massive conformational change in the position of the apical GroEL domain resulting in the formation of a large folding cage. Binding of ATP to the \textit{trans}-ring triggers the release of the folded protein and is rapidly followed by the binding of GroEs to form a new folding-active \textit{cis} assembly. Both rings may alternate in protein folding.
cage. The binding and release of GroES is regulated by the ATPase activity of GroEL in a ‘two stroke’ mechanism. In a single round of GroEL/GroES-mediated protein folding (Fig. 4), initially one ring of GroEL binds ATP and GroES. ATP hydrolysis in this GroEL subunit leads to the formation of GroEL-ADP-GroES complex. Note that substrate is discharged into the cavity here. A functional cycle requires the presence of both rings. ATP binding in the opposite ring causes dissociation of the ADP and GroES. Upon release of GroES and ADP from GroEL, the ATP associated GroEL ring becomes an acceptor site for unfolded polypeptide (Martin et al., 1993a; Martin et al., 1993b; Martin et al., 1993c). The folded polypeptide leaves, whereas incompletely folded polypeptide rebinds. Chaperonin-dependent model substrates such as rhodanese, typically require multiple rounds of GroEL/GroES assisted folding to reach their native stage.

Apart from the prominent chaperone families Hsp70 and Hsp60 discussed above, more than 20 different families of chaperones are currently known. Despite the accumulation of a wealth of biochemical information, the cellular roles of some of these molecular chaperones remain unclear. For the sake of brevity an overview of the important cytoplasmic and mitochondrial chaperones is given in Table 1.

Since all of the processes I investigated in this thesis are taking place in the ER it is important to provide some basic background to the reader concerning the complex environment of this particular cell compartment. In the next part of my Introduction, I will attempt to give you insight into the unique aspects of protein folding in the ER.

1.3.3 Protein Folding in the ER

The ER represents a specialized subcellular folding compartment for secreted, plasma membrane and various organellar proteins. While some of the aspects of protein folding are common for cytosol and ER compartments such as molecular crowding, ionic conditions and the need for ATP for many of the reactions that are directly or indirectly
involved in folding proteins, several conditions make the ER unique as a folding compartment. They arise from the unique physiological functions of the ER that are defined not only by the different nature of the proteins that fold in the ER but also by the nature of the machinery that segregates them across a membrane from cytosolic proteins. The ER differs from other folding compartments (e.g., the cytosol, mitochondria) in the presence of carbohydrates and glycosylation machinery, its high oxidizing potential (Hwang et al., 1992), and its high Ca\(^{2+}\) concentration (Montero et al., 1995). These conditions provide the evolutionary pressure for the presence of unique molecular chaperones and enzymes to facilitate the folding and assembly of newly synthesized polypeptides in the ER.

1.3.3.1 Signal Sequence and Translocation

Membrane or secreted proteins are synthesized on membrane-associated ribosomes and are inserted co-translationally into the ER through a specialized structure known as the translocon. Most proteins that fold in the ER are targeted there by an N-terminal signal sequence. The presence of the signal sequence influences the folding of the nascent chain. As soon as a typical 20 amino acid signal sequence of a secretory protein protrudes from the large subunit of a ribosome it is bound by the Signal Recognition Particle (Stevens and Argon, 1999), which prevents premature folding and targets the ribosome to the ER membrane. At this stage, the newly synthesized protein is translocated into the lumen of the ER. Processing of the signal sequence can occur either co- or post-translationally, but even for those proteins that are processed co-translationally, cleavage does not occur until the nascent polypeptide reaches a minimal size of \(~80\%\) of the final length (Josefsson and Randall, 1981a; Josefsson and Randall, 1981b). As long as the signal sequence is not cleaved, the N-terminus of the mature protein is constrained and in many cases this causes a delay in folding. For example, a signal sequence that is not cleaved immediately prolongs the association of HIV-1 gp120
with the ER chaperone, CNX (Li et al., 1996). Thus the presence of the signal sequence can profoundly affect folding of the mature portion of the polypeptide.

It is also important to note that there is a vast discrepancy between the rate of folding and translation/translocation across the ER membrane. Translation/translocation proceeds at 4 +/- 2 residues per second, while initial folding steps like hydrophobic collapse can occur within nanoseconds. Thus, a polypeptide chain starts folding in the lumen of the ER well before its synthesis is completed (Bergman and Kuehl, 1978; Bergmann and Lodish, 1979). The folding opportunities for \( \alpha \) helices and \( \beta \) sheets are intensely different during translocation into the ER. Helix forming amino acids are translocated sequentially and therefore all the interactions that stabilize an \( \alpha \) helix can be formed without delay. However, the formation of a \( \beta \) sheet is severely delayed by the translocation process. Each peptide that assumes an extended strand conformation must wait until the other strands are synthesized to form hydrogen bonds. Thus, a mechanism for delaying folding in the ER is necessary. The extended \( \beta \) strand peptides are substrates for the peptide-binding chaperones of the Hsp70 family (Landry et al., 1992) which prevent inappropriate folding until partner amino acids become available in the lumen.

A similar folding constraint exists for Cys residues involved in disulfide bond formation. In fact, many proteins have disulfide bonds separated by many amino acids. Since disulfide bond formation in the ER occurs for a number of proteins such as hCG\( \beta \) or the hemagglutinin of influenza virus without a large population of incorrect S-S bonds, that would suggest that binding of molecular chaperones is used to limit possible disulfide errors.

1.3.3.2 Glycosylation

Addition of N-linked carbohydrates to newly synthesized proteins occurs specifically in the ER. N-linked glycans are added to appropriate Asn residues (i.e., those incorporated into the sequence Asn-X-Ser/Thr, where X is any amino acid other than
proline) during the translocation process. In this process a pre-formed oligosaccharide (Fig. 5) composed of N-acetylglucosamine, mannose, and glucose and containing a total of 14 sugar residues (Glc₃Man₉GlcNAc₂) is transferred en bloc to the side-chain NH₂ group of Asn (Hubbard and Ivatt, 1981). The transfer is catalyzed by a membrane-bound enzyme, oligosaccharyl transferase, which has its active site exposed on the luminal side of the ER membrane (Ballou et al., 1986). The precursor oligosaccharide is held in the ER membrane by a special lipid molecule called dolichol, and it is transferred to the target Asn in a single enzymatic step immediately after that amino acid emerges in the ER lumen during protein translocation (Hubbard and Ivatt, 1981).

All of the diversity of the N-linked oligosaccharide structures on mature glycoproteins results from post ER modifications of the original Glc₃Man₉GlcNAc₂ structure. However, while still in the ER, the terminal three glucose residues are removed quickly by the action of two enzymes; the first glucose by the membrane bound α-glucosidase I (Kornfeld and Kornfeld, 1985), and the other two by the soluble α-glucosidase II (D'Alessio et al., 1999; Trombetta et al., 1996). These reactions likely occur cotranslationally on the nascent polypeptide and can be blocked by castanospermine (CAS) or N-methyl deoxynojirimycin which inhibit α-glucosidase I and α-glucosidase II.

The vast majority of proteins passing through the secretory pathway are N-glycosylated. The generality of N-glycan addition in the ER is reflected by a recent survey of the Swiss-Prot database which showed that, out of 1823 complete animal protein entries with reported extracellular features, 1671(91.7%) were described as 'glycoproteins' in the keyword field. The remaining 8.3%, representing 152 potentially non-glycosylated plasma membrane proteins, contained 116 proteins with multiple transmembrane regions, 15 proteins that are known to associate with glycosylated subunits in a complex such as CD3 chains and seven that contained 5-38 potential glycosylation sites, yet were not marked as glycoproteins. This leaves only 14 sequences (0.7%) that are candidates for non-glycosylated, non-complex plasma membrane proteins.
Figure 5. Structure of the oligosaccharide added to proteins in the ER.

A Glc$_3$Man$_9$GlcNAc$_2$ oligosaccharide is transferred to Asn residues in nascent polypeptide chains. Glc$_1$Man$_9$GlcNAc$_2$, the specific oligosaccharide recognized by the ER chaperones CNX and CRT, is obtained by oligosaccharide processing reactions catalyzed by two enzymes, glucosidase I and glucosidase II, both localized to the ER. Glucosidase I removes the external Glc unit, whereas glucosidase II excises both remaining Glc units.
Less frequently, oligosaccharides are linked to the hydroxyl group on the side chain of serine, threonine, or hydroxylysine. These O-linked oligosaccharides are primarily formed in the Golgi apparatus.

N-linked glycosylation serves several roles in protein folding: 1) to increase the solubility of the glycoprotein as carbohydrates are hydrophilic in nature, 2) to make the process of translocation across the ER membrane less reversible by increasing the energy barrier to back-translocation and 3) to mark the surface of the folding polypeptide for binding with CNX or CRT.

One function of N-linked glycans is to facilitate protein folding and conformational maturation. When N-linked chains are eliminated by site-directed mutagenesis of the Asn residues or by treatment of cells with agents that block addition of N-linked glycans, many glycoproteins misfold, aggregate, and get degraded within the ER (Helenius, 1994). Detailed studies with the viral glycoproteins, vesicular stomatitis virus G protein and influenza hemagglutinin, whose Asn-linked glycosylation sites were systematically mutated, showed that in many cases glycosylation is needed for proper folding; under-glycosylated proteins formed intracellular aggregates and were retained in the ER (Gallagher et al., 1992; Machamer et al., 1985). Interestingly, however, no glycan at a specific site was necessary for folding of such viral glycoproteins (Gallagher et al., 1992; Machamer et al., 1985). Furthermore, expression and secretion of functional single-chain Fv molecules that are normally not glycosylated was enhanced by the addition of a glycan (Jost et al., 1994). In contrast, there are other glycoproteins that seem to fold efficiently without their N-linked glycans (Cresswell et al., 1987). The only rule that seems to emerge here is that glycoproteins, often large ones, that fold slowly in the first place have even more trouble folding if their N-linked glycans are missing (Ruddon et al., 1996).
1.3.3.3 Basis of Oxidative Protein Folding

The presence of disulfide bonds is regarded as a general property that distinguishes proteins that fold in the ER from cytosolic proteins. To support efficient disulfide formation, cells actively promote oxidation in the ER lumen. The high oxidative redox potential is maintained in the ER by the major redox buffer, the Cys-containing tripeptide glutathione (Hwang et al., 1992). The ratio of reduced to oxidized glutathione in the ER ranges from 1:1 to 3:1, whereas the overall cellular ratio ranges from 30:1 to 100:1. The mechanism by which such an oxidative environment is maintained in the ER is not clear. Recently, data from yeast show that a second mechanism, a protein-based chain of oxidation/reduction reactions may contribute to the oxidative environment in the ER. Ero1p is a newly characterized membrane bound glycoprotein required for an oxidative environment in the ER lumen. Genetic studies in Saccharomyces cerevisiae showed that the loss of Ero1p results in the accumulation of reduced PDI and the cessation of disulfide bond formation. Furthermore, this effect is glutathione-independent (Frand and Kaiser, 1998). Over the past few decades, a number of factors have been suggested to contribute to disulfide formation in the ER, including secretion of reduced thiols, uptake of oxidized thiols, and a variety of redox enzymes and small molecule oxidants (Frand and Kaiser, 1998; Hwang et al., 1992). The physiological importance of any of these to disulfide formation has not been established.

Disulfide bonds play a very important role in dictating protein folding pathways by forming covalent folding intermediates and severely restricting available conformations. Formation of disulfide bonds may begin co-translationally, as described for the immunoglobulin polypeptide (Bergmann and Lodish, 1979), or much later in the folding sequence as in the case of the hemagglutinin of influenza virus (Segal et al., 1992). Interestingly, proteins can undergo post-translational oxidation and achieve the same native structure as with co-translational oxidation. Addition of reducing agents to the medium of live cells prevented disulfide bond formation in newly synthesized class I
molecules (Solheim, 1999). When the reductant was washed out, the reduced proteins rapidly oxidized, correctly folded and assembled (Tector et al., 1997).

Several ER lumenal enzymes are capable of catalyzing the formation of disulfide bonds, reducing disulfide bonds, and correcting aberrant disulfide bonds. These include the most abundant ER disulfide isomerase, the 60 kDa protein disulfide isomerase (PDI) as well as two structurally-related proteins, ERp72 and ERp57. Each contains two or more thioredoxin-like domains which can catalyze the thiol exchange reactions \textit{in vitro} (Darby and Creighton, 1995; Füllekrug et al., 1994).

1.3.4 Molecular Chaperones in the ER
Given the unique demands of folding secreted and membrane proteins, it is not surprising that the ER contains a number of specific molecular chaperones and folding enzymes (Table II). The most abundant and well characterized of them include: the immunoglobulin heavy chain binding protein (BiP, also called glucose regulated protein 78 or GRP78) which is a member of the Hsp70 family and interacts with hydrophobic stretches of amino acids in a wide range of proteins; the glucose regulated protein 94 (GRP94, also called endoplasmin, which is a member of the heat shock protein 90 (Hsp90) family of proteins: protein disulfide-isomerase (PDI, also called ERp59) which acts as molecular chaperone and is involved in the formation and rearrangement of disulfide bonds to form protein secondary structure; CNX, the only membrane-spanning chaperone identified so far (also called p88 and IP90) and its soluble homologue CRT which are both lectins that react with monoglycosylated N-linked oligosaccharides on glycoproteins. All of these proteins are selectively retained in the ER either by a specific KDEL retention/retrieval signal at the carboxy terminus, or as integral membrane proteins. The ER seems to lack two types of molecular chaperones present in the cytoplasm: small molecular weight proteins of the HSP27 family and chaperonins of the HSP60 family.
Table II Molecular chaperones and folding enzymes in the ER.

<table>
<thead>
<tr>
<th>PROTEIN CLASS</th>
<th>REFERENCES</th>
</tr>
</thead>
<tbody>
<tr>
<td>Disulfide isomerases</td>
<td></td>
</tr>
<tr>
<td>PDI</td>
<td>(Darby and Creighton, 1995)</td>
</tr>
<tr>
<td>ERp72</td>
<td>(Rupp et al., 1994)</td>
</tr>
<tr>
<td>ERp57</td>
<td>(Elliott et al., 1997)</td>
</tr>
<tr>
<td>Ero1p</td>
<td>(Frands and Kaiser, 1998)</td>
</tr>
<tr>
<td>Prolyl isomerases</td>
<td></td>
</tr>
<tr>
<td>FKBP13</td>
<td>(Johnson and Toft, 1994)</td>
</tr>
<tr>
<td>FKBP65</td>
<td></td>
</tr>
<tr>
<td>s-Cyclophilin</td>
<td></td>
</tr>
<tr>
<td>CCYLP</td>
<td></td>
</tr>
<tr>
<td>Cyclophilin B</td>
<td></td>
</tr>
<tr>
<td>Hsp70</td>
<td></td>
</tr>
<tr>
<td>BiP</td>
<td>(Freiden et al., 1992)</td>
</tr>
<tr>
<td>Lhs1p</td>
<td>(Dierks et al., 1996)</td>
</tr>
<tr>
<td>GRP170</td>
<td>(Lin et al., 1993)</td>
</tr>
<tr>
<td>Dna-J-like</td>
<td></td>
</tr>
<tr>
<td>Sec63p</td>
<td>(Feldheim et al., 1992)</td>
</tr>
<tr>
<td>Scj1p</td>
<td>(Schlenstedt et al., 1995)</td>
</tr>
<tr>
<td>Hsp90</td>
<td></td>
</tr>
<tr>
<td>GRP94</td>
<td>(Argon and Simen, 1999)</td>
</tr>
<tr>
<td>Lectins</td>
<td></td>
</tr>
<tr>
<td>CNX</td>
<td>(Degen and Williams, 1991; Ihara et al., 1999; Wada et al., 1991)</td>
</tr>
<tr>
<td>CRT</td>
<td>(Michalak et al., 1999; Peterson et al., 1995; Saito et al., 1999)</td>
</tr>
</tbody>
</table>
Molecular chaperones and folding enzymes in the ER are thought to co-associate and act concurrently on proteins to facilitate their folding. This view is supported by the identification of complexes which contain folding intermediates associated with different molecular chaperones. For example, complexes of substrates with BiP and GRP94 or BiP and CNX have been described (Melnick et al., 1994) and may occur within a matrix of luminal proteins. Furthermore, proline isomerases make disulfide isomerases more efficient (Rupp et al., 1994) and disulfide isomerase interacts with CRT in a Ca\(^{2+}\) regulated fashion (Corbett et al., 1999). The concerted action of molecular chaperones may function either to hold a polypeptide in a specific conformation to facilitate appropriate modifications, or to prevent incorrect interactions of exposed protein domains during polypeptide folding. The notion of chaperone cooperativity in the ER is also appealing because of the complementary specificity of the chaperones, hydrophobic polypeptides for BiP or glycans for CNX and CRT. It is important to note that, so far, none of the known ER chaperones provide shielding from the environment in the same way that the chaperonin cavity does. It is possible that such a micro-environment is less critical in the ER because many of the proteins are glycosylated and constrained by the signal sequence.

The folding process in the ER is not very efficient and some polypeptides do not achieve correct folding, particularly when the cell is under conditions of stress or when mutations have been introduced into the polypeptide. The efficiency of folding in the ER can vary from almost 100% for molecules like MHC class I or viral glycoproteins to only \(\sim 15\%\) for proteins like the cystic fibrosis transporter. To deal with this problem, an elegant process termed ‘quality control” has evolved (Hammond and Helenius, 1995; Kopito, 1997). A characteristic of quality control is that it continually monitors for misfolding, actively retaining and degrading incompletely folded or aggregated intermediates, misfolded polypeptides and unassembled subunits. ‘Quality control’ is an important process as it is essential to avoid the export of structurally/functionally altered
proteins and/or protein aggregates, from the ER. The molecular chaperones BiP, Grp94, PDI, CNX and CRT have all been implicated as part of the ‘quality control’ process in the ER.

1.3.4.1 Classical ER Molecular Chaperones

1.3.4.1.1 BiP

BiP was originally identified as a 78 kDa glucose regulated protein or Grp78 that accumulates in the ER in response to a block in glycoprotein synthesis or glucose deprivation (Pouyssegur et al., 1977). Later it was also discovered as the immunoglobulin heavy chain binding protein and termed BiP. The first indication that BiP functions as a broad specificity molecular chaperone was provided by the discovery that BiP is the ER-located member of the family of Hsp70 molecular chaperones (Munro and Pelham, 1986). Functional data supporting a molecular chaperone role for BiP came from experiments demonstrating that BiP binds transiently to newly synthesized proteins in the ER and more permanently to misfolded, underglycosylated or unassembled proteins whose transport from the ER is blocked (Gething and Sambrook, 1992). Also, BiP does not interact with native polypeptides. Finally, BiP was shown to bind with low affinity (in the range of 1-100 mM) to 7-8 residue peptides whose sequence contains hydrophobic amino acids (Blond-Elguindi et al., 1993). The most direct evidence for BiP’s molecular chaperone function was obtained from studies in yeast. The *Saccharomyces cerevisiae* Kar2 (BiP) protein is essential for protein folding in the ER lumen; also Kar2 mutants that lacked a functional ATPase domain bind more permanently to misfolded substrates (Simons et al., 1995).

Experiments performed in yeast provide evidence for additional roles for BiP in the ER. BiP is essential for protein translocation across the ER membrane (Lyman and Schekman, 1995; McClellan et al., 1998) and for retrograde protein transport of aberrant polypeptides destined for degradation by the proteasome (Brodsky et al., 1999; Plemper
et al., 1997). Finally, in *Saccharomyces cerevisiae* Kar2 is required during a/α cell mating for the stage of karyogamy that involves fusion of the nuclear membranes, the outer of which is contiguous with the ER membrane (Latterich and Schekman, 1994; Ng et al., 1996; Ng and Walter, 1996). In mammalian cells, BiP is implicated in maintaining the permeability barrier of the ER membrane by sealing the lumenal end of the translocon pore before and early in translocation (Hamman et al., 1998).

BiP, similarly to other Hsp70 proteins, consist of two domains, an N-terminal domain with an ATPase catalytic site and a C-terminal domain that binds substrate (Fig.3A). The duration and affinity of polypeptide binding is regulated by the nucleotide binding site in the N-terminal domain (Flynn et al., 1989). When chaperoned by BiP an unfolded polypeptide undergo cycles of binding and release regulated by the rate of ATP hydrolysis and ADP to ATP exchange as described for Hsp70 (Fig.3B)(see section on Hsp70).

BiP binding sites may be located in regions of the polypeptide chain that take up their final conformation late in the folding pathway (Gething, 1996) or may be present within sequences that fold rapidly, precluding their interaction with the chaperone (Hellman et al., 1999). It thus appears that the rate and stability of the folding of a protein determines whether or not a particular site is recognized, with BiP preferentially binding to proteins that fold slowly or somewhat unstably (Hellman et al., 1999).

BiP, like all other Hsp70 proteins, performs its chaperoning functions with the help of J-domain containing Hsp40 proteins. In *S. cerevisiae* Kar2p (BiP) is thought to interact with three ER-located Hsp40 proteins, Sec63p (Feldheim et al., 1992), Scj1p (Schlenstedt et al., 1995) and Jem1p (Nishikawa and Endo, 1997). A possible mammalian counterpart of Sec63p has been identified, Mtj1 (Brightman et al., 1995), and the physical interaction between Mtj and BiP has been recently demonstrated in the mammalian ER (Chevalier et al., 2000).
In vivo, mammalian BiP exists in interconvertible monomeric and oligomeric forms (Blond-Elguindi et al., 1993; Freiden et al., 1992) and can be post-translationally modified by phosphorylation (Gaut, 1997) and by ADP ribosylation (Ledford and Leno, 1994). These modifications may be important in regulating the synthesis and polypeptide binding activity of the molecule (Freiden et al., 1992). It is well established that conditions that increase the levels of unfolded proteins in the ER lumen increase also the proportion of BiP’s monomeric species and cause a decrease in the extent of modification of BiP (Laitusis et al., 1999; Leno and Ledford, 1990; Leustek et al., 1991). Only unmodified, monomeric BiP molecules are found in complexes with unfolded or unassembled polypeptides (Freiden et al., 1992; Hendershot et al., 1988).

In the ER, BiP cooperates with a number of other molecular chaperones, including the Hsp90 family member Grp94 and CNX. BiP and Grp94 can be isolated in a ternary complex with newly synthesized immunoglobulin chains (Melnick et al., 1994). BiP binds an early intermediate of immunoglobulin light chain and dissociates within a few minutes whereas Grp94 binding is prolonged with a half-time of 50 min and it interacts with disulfide bonded species (Melnick et al., 1994). Similarly, during the folding of vesicular stomatitis virus G protein BiP binds to early folding intermediates but in this case is replaced after a short lag period by CNX (Hammond and Helenius, 1994). Chemical crosslinking experiments revealed that newly synthesized thyroglobulin may interact with multiple ER proteins such as BiP, Grp94, ERp72 and grp170. Similarly, newly synthesized influenza hemagglutinin was found to be part of a complex that contains BiP, Grp94, CNX and CRT (Kuznetsov et al., 1997). These findings support the view that a variety of different molecular chaperones act either simultaneously or sequentially during the conformational maturation of secretory or plasma membrane proteins, with the nature and order of involvement of the various chaperones depending on particular features displayed by the folding polypeptide at different stages of their maturation.
Another representative of the Hsp70 family in the ER is GRP170, which is associated with immunoglobulin in B cells, but whose role is not yet defined (Lin et al., 1993).

### 1.3.4.1.2 GRP94

GRP94 is the member of the Hsp90 family with an intriguing pattern of evolutionary expression. It is present in vertebrates, some invertebrates like *C.elegans*, and in plants. GRP94 is absent from bacteria and yeast and from *Drosophila* (Argon and Simen, 1999). GRP94 was first described as one of the three proteins induced by glucose starvation by the same group that identified Grp78(BiP) (Shiu et al., 1977). It was also isolated as the most abundant calcium binding protein of the ER lumen and termed endoplasmin (Koch et al., 1986) or CaBP4 (Van et al., 1989). When detected in differentiating B lymphocytes as one of the three major induced proteins GRP94 was named ERp99 (Lewis et al., 1985). Finally, GRP94 was also identified as gp96, a glycoprotein which plays a role as a major tumor rejection antigen (Srivastava et al., 1986). GRP94 is one of the most abundant proteins of the ER, accounting for 5-10% of the ER lumenal contents with an estimated concentration of 10 mg/ml.

GRP94 is currently thought to have two major functions; as a molecular chaperone during folding of secretory and membrane proteins and as a peptide carrier for T cell immunization. In this introduction I will only examine the current data dealing with GRP94 as a molecular chaperone.

The GRP94 protein consists of 782 residues that can be divided into 6 distinct structural domains (Fig.6) whose functions are only beginning to be elucidated. The N-terminal nucleotide binding domain is homologous to that of Hsp90, and it has been shown to bind ATP (Schulte et al., 1999). GRP94 has two acidic domains, the 7-repeat imperfect leucine zipper-like sequence implicated to serve as a protein-protein interaction module, the KEKE domain with another putative protein interaction motif (Realini et al.,
1994) and the dimerization domain. Similarly to Hsp90, GRP94 exists in the ER as a homodimer and a C-terminal 23 kDa fragment was identified to be both necessary and sufficient for dimerization. Residence in the ER is maintained by the C-terminal tetrapeptide KDEL. GRP94 is a glycoprotein with six potential glycosylation sites, but only one. Asn 196 is glycosylated. Glycosylation is, however, unrelated to substrate binding since de-glycosylated GRP94 can still associate with newly synthesized peptides (Argon and Simen, 1999). GRP94 is phosphorylated on Ser and Thr, but not on Tyr residues. Phosphorylation may regulate the activity of GRP94, as phosphorylated GRP94 does not associate with immunoglobulin light chain in the cell (Argon and Simen, 1999). Like the other abundant lumenal proteins, GRP94 is a low-affinity, high-capacity calcium binding protein. It is thought to have 15 calcium-binding sites, 4 with moderate affinity (Kd~2 μM) and 11 with low affinity (Kd~600 μM) (Van et al., 1989). A fundamental difference between GRP94 and chaperones like BiP or CNX is that it interacts with a restricted set of protein substrates. While BiP binds to a large number of proteins that fold in the ER by recognizing a broad range of peptides, and while CNX recognizes a glycan moiety common on many glycoproteins, GRP94 has been shown to associate only with few proteins. GRP94 interacts with immunoglobulin chains (Melnick et al., 1994). MHC class II (Schaif et al., 1992), thyroglobulin (Kuznetsov et al., 1997), erbB2 (Chavany et al., 1996), a herpes virus glycoprotein (Ramakrishnan et al., 1995), apolipoprotein B (Linnik and Herscovitz, 1998), collagen (Ferreira et al., 1994), protein C (Katsumi et al., 1996) and pancreatic bile salt-dependent lipase (Bruneau et al., 1998). GRP94 has not been detected in association with class I molecules or with many viral glycoproteins.

The emerging picture of GRP94 function is of a chaperone specific for advanced intermediates in the biosynthesis of proteins, which works downstream of or in conjunction with other chaperones. Functional data supporting this view came from in vitro experiments showing Hsp90 preference for late folding intermediates during in vitro
Figure 6. A proposed structural model for the ER Hsp90 family member, GRP94.

The protein is depicted as a dimer. The N-terminal signal sequence and the C-terminal KDEL sequence target and localize GRP94 to the ER, respectively. An Asn-linked oligosaccharide (CHO) is attached at position 196.
refolding reactions (Freeman and Morimoto, 1996; Jakob et al., 1995; Wiech et al., 1992). Further evidence that GRP94 associates with advanced folding intermediates of its substrate proteins, or with incompletely assembled oligomers, but not early folding intermediates was obtained by demonstrating that \textit{in vivo} GRP94 associates with a late fully disulfide bonded intermediate of immunoglobulin light chain and this association persists for the majority of the residence time of the light chain in the ER. Also, during the biosynthesis of MHC class II molecules, GRP94 binds to incompletely assembled intermediates, lacking the invariant chain (Schaiff et al., 1992). Another feature of GRP94 interaction with its substrate proteins is that each of these substrates has also been demonstrated to interact with either BiP or CNX (Schaiff et al., 1992; Ferreira et al., 1994; Kuznetsov et al., 1997; Melnick et al., 1994). It is not known whether GRP94 binding to proteins depends on other molecular chaperones.

By analogy to Hsp90, it is likely that the functions of GRP94 are enhanced by interactions with co-chaperones. However, there is currently no direct evidence for the existence of ER homologs of HSP90-associated proteins like p23, HOP, FKBPS or cyp40 (Chen and Smith, 1998; Johnson and Toft, 1994).

An unresolved issue in the understanding of GRP94 function, as with other HSP90 family members, is the relationship between substrate binding and adenine nucleotide binding and hydrolysis. The peptide binding chaperones whose action are understood in great detail are all ATPases. GRP94 has been reported to bind immobilized ATP (Clairmont et al., 1992; Dierks et al., 1996) and to cleave ATP (Li and Srivastava, 1993). However, the inherent ATPase activity of GRP94 is much slower than the activity of BiP and the ATP binding activity is at least 20-fold weaker than that of BiP (Dierks et al., 1996). Furthermore, \textit{in vitro} peptide binding by GRP94 is ATP-independent (Wearsch and Nicchitta, 1997). So, the role of adenine nucleotides in regulating GRP94 substrate binding remains unclear.
1.3.4.2 Protein Disulfide Isomerase (PDI)

In the preceding section I have described two classical molecular chaperones that function in the ER. In addition, some thiol oxidoreductases of the ER have been implicated as molecular chaperones. The most studied member of the family is PDI. PDI is a multifunctional enzyme with diverse activities including oxidation, reduction and isomerization of disulfides, chaperone and anti-chaperone activity, and transglutaminase activity (reviewed in Noiva, 2000). PDI also acts as a subunit of prolyl 4-hydroxylases (Koivu et al., 1987) and the microsomal triglyceride transfer protein complexes (Wetterau et al., 1991). In addition to its ER function, PDI catalyses redox reactions with proteins at the cell surface and in endosomes. PDI also interacts with cell surface integrins (Mou et al., 1998), actin microfilaments (Farwell et al., 1990), and peptides transported into the ER by TAP for interaction with MHC class I (Lammert et al., 1997; Spee and Neefjes, 1997). I will limit my discussion only to the functions of PDI which relate to their molecular chaperone and oxidoreductase functions in the ER and discuss them in the context of one particular isoform, ERp57.

The thiol oxidoreductase, ERp57, was first discovered in a temperature-sensitive fibroblast cell line as one of the three proteins that accumulated in the ER upon glucose starvation, and as such was termed glucose regulated protein, GRP58 (Shiu et al., 1977). Over the years ERp57 has appeared in the literature under a host of pseudonyms as murine ER protein, ERp61 (Lewis et al., 1986; Mazzarella et al., 1994), estrogen induced protein, HIP-70 (Mobbs et al., 1990), a new rat isozyme of thiol:protein-disulfide oxidoreductase Q-2 (Srivastava et al., 1991), and as a new component of the class I H chain complex. ER-60 (Lindquist et al., 1998).

ERp57 is an ER-resident protein and a member of the PDI family. It was initially thought to be an isozyme of phosphoinositide-specific phospholipase C (Bennett et al., 1988), however, the activity was never demonstrated (Srivastava et al., 1993). Several
other activities have since been attributed to ERp57, including that of a cysteine protease (Urade et al., 1992) and carnitine palmitoyl transferase (Murthy and Pande, 1994). However, ERp57 is now generally regarded as a thiol oxidoreductase. The first indication of ERp57 function was provided by the analysis of the predicted amino acid sequence from a number of cDNA clones (Bennett et al., 1988). ERp57 was noted to share considerable sequence identity with protein disulfide isomerase, PDI. Several subsequent studies demonstrated that ERp57 has a thiol-dependent oxidoreductase activity (Hirano et al., 1995; Srivastava et al., 1991; Srivastava et al., 1993; Zapun et al., 1998). The direct evidence of ERp57 thiol-dependent reductase activity was provided by in vitro studies analyzing the refolding of a monoglucosylated substrate (Zapun et al., 1998). Also recently, a role for ERp57 in the specific catalysis of native disulfide bond formation in glycoproteins has been demonstrated in cells. The formation of a mixed disulfide species formed between ERp57 and a specific viral glycoprotein substrate was detected in a pulse-chase experiment coupled with 2-dimensional electrophoresis (Molinari et al., 1999).

ERp57 belongs to the superfamily of protein-thiol oxidoreductase enzymes with sequence and structural similarity to thioredoxin. ERp57 is composed of 481 amino acids forming four domains with thioredoxin folds (a-b-b'-a'), followed by a c region that probably does not have secondary or tertiary structure. There is extensive internal sequence similarity between the a and a' domains as well as between the b and b' domains (Fig.7A). The two protein-thiol oxidoreductase active sites of ERp57 (-Cys-Gly-His-Cys-) are located in the a and a' domains. Recent NMR studies of the a domain of ERp57 (Kemmink et al., 1996) indicate that it has a thioredoxin fold with two grooves that are probable sites of enzyme-substrate interaction. Similarly, the b domain has a thioredoxin fold (Darby and Creighton, 1995; Darby 1996) despite the absence of any Cys-Xaa-Xaa-Cys-redox active sites or sequence similarity to thioredoxin. The proposed structure of ERp57 is a total of four domains with thioredoxin-like folds, two with redox active
Figure 7. A. Structure of ERp57 based on a model proposed by Darby et al. ERp57 is a modular protein comprised of domains a, b, b' a' and c. The a and a' domains contain the sequence Cys-Gly-His-Cys (CGHC), which represents the active catalytic sites for isomerase activity. The C-terminal–KDEL motif is involved in the ER localization of ERp57. B. Enzymatic reaction catalyzed by the protein-thiol oxidoreductases. The reaction catalyzed by ERp57 can be either a reduction of an existing disulfide, an oxidation of two thiols or isomerization of disulfides.
sites, and two with no redox activity. The c domain has been defined as the region from amino acid residue 465 to the C-terminus of the ERp57 polypeptide. The c domain is of considerable interest because of the presence of the C-terminal KDEL ER retention signal and lack of similarity with classical PDI.

The family of thioredoxin-like enzymes include, apart from the ERp57, other eukaryotic enzymes such as PDI, ERp72, CABP1, prokaryotic enzymes such as DsbA, as well as the ubiquitous enzymes thioredoxin and glutaredoxin. The catalytic mechanism used by protein thiol-disulfide oxidoreductases for the formation or reduction of disulfides has been well characterized (Gilbert, 1990). The protein thiol-disulfide oxidoreductases catalyze disulfide exchange between the enzyme and the substrate (Fig. 7B). When an oxidoreductase catalyzes the oxidation of two substrate cysteiny1 sulphydryls, it becomes reduced. Likewise, when the oxidoreductase catalyses the reduction of the disulfide bridge, it becomes oxidized. The redox state of the oxidoreductase can be regenerated by an enzyme, such as thioredoxin reductase or a small redox buffer such as glutathione. The direction of the reaction catalyzed by the oxidoreductase is determined by its substrate and product concentrations, redox potential, and the redox conditions of the cellular milieu.

In addition to its activity as a protein-thiol oxidoreductase, ERp57 is thought to function as a molecular chaperone. A role for ERp57 during glycoprotein maturation was first indicated when ERp57 was found to interact with CNX during the biosynthesis of a glycosylated membrane protein, glut 1 glucose transporter (Oliver et al., 1996). Subsequent experiments showed that ERp57 interacts specifically with a number of newly synthesized secretory and membrane proteins, and that this interaction is transient (Elliott et al., 1997; Oliver et al., 1997). The most direct evidence that ERp57 may not function solely as the thiol-dependent reductase, but may have a more general chaperone activity, comes from experiments in which ERp57 was shown to be crosslinked to glycoproteins that do not contain any cysteine residues and, therefore, lack disulfide bonds (Oliver et al., 1997). The observation that ERp57 is a stress-inducible protein
(grp58), which accumulates under similar conditions as BiP/Grp78 and GRP94 (Shiu et al., 1977) further supports the view that it may act as a generic molecular chaperone.

ERp57 appears to be adapted to function together with CNX and CRT in promoting folding, assembly and disulfide bond formation of glycoproteins in the ER (Elliott et al., 1997; Oliver et al., 1997; Zapun et al., 1998). A model has been proposed in which CNX or CRT binds to a monoglucosylated N-linked glycan of a newly synthesized glycoprotein, and ERp57 associates with these chaperones. A crucial discovery that promotes this view was the finding that ERp57, which does not have any intrinsic lectin-like properties, interacts specifically with N-glycosylated polypeptides. Both binding of ERp57 to newly synthesized glycoproteins, and its subsequent release, required the trimming of glucose residues from the N-linked glycan (High et al., 2000).

Using a variety of approaches, ERp57 was shown to form distinct complexes with both CNX and CRT. These discrete complexes were formed both within the lumen of the ER and when the proteins were mixed in solution (Oliver et al., 1999). Furthermore, the binding of ERp57 to CNX and CRT is direct and does not require the presence of the glucose- trimmed glycoprotein (M. Leach manuscript in preparation). Finally, the catalysis of disulfide bond formation by ERp57 was greatly enhanced in the presence of CNX or CRT, when the formation of disulfide bonds in ribonulease B was monitored in vitro (Zapun et al., 1997).

Several groups have demonstrated the presence of ERp57 during the assembly of the MHC class I complex. ERp57 was found to interact with both class I-H chain-CNX and H chain-β2m-CRT complexes (Lindquist et al., 1998; Hughes and Cresswell, 1998). The most interesting aspect of the ERp57 interaction with class I is that it is only in this context that the CRT-ERp57 interaction can be detected without chemical crosslinking (Elliott et al., 1997; Hughes and Cresswell, 1998).

This introduction has described a number of different molecular chaperones that contribute to the production of functional proteins in the cell. Indeed, a wealth of
information is now available regarding the folding and assembly of a number of proteins and the mechanisms which are responsible for these functions. However, the molecular chaperone composition of the ER, at least as currently known, differs from that of the other folding compartments in eukaryotes and prokaryotes, as it contains a unique type of chaperone-lectin that binds glycoproteins. In this section, I will discuss two such chaperones: CRT, which is a lumenal protein, and CNX, the only membrane spanning chaperone identified so far.

1.3.4.3 CNX and CRT: Protein Structure and Function.

1.3.4.3.1 CNX

CNX was termed p88 when it was initially described as a novel 88 kDa protein that transiently recognizes murine MHC class I molecules (Degen and Williams, 1991). In human cells, CNX was detected as a 90 kDa protein (IP90) that associates with incompletely assembled T cell receptor, immunoglobulin and MHC class I subunits (Hochstenbach et al., 1992). Finally, when identified as a 90 kDa phosphoprotein of the ER lumen it was termed CNX based on the protein's homology with CRT and its ability to bind calcium (Wada et al., 1991). Co-immunoprecipition experiments with CNX specific antibody, demonstrated that this protein interacts with over 50 newly synthesized proteins in the ER (Hochstenbach et al., 1992). In recent years, a large number of these proteins have been identified (Table III).

Several lines of evidence indicate that CNX may function as a molecular chaperone and a component of the ER quality control system. First, CNX binds transiently to a diverse array of both membrane bound and soluble proteins shortly following their translocation into the ER (Hochstenbach et al., 1992). Second, CNX associates with folding intermediates but not fully folded proteins. For the major secretory glycoprotein gp80 of the MDCK cells, CNX dissociation correlates with gp80 precursor folding an
Table III. Table of CNX and CRT in vivo substrates.

<table>
<thead>
<tr>
<th>ASSOCIATE WITH CALNEXIN</th>
<th>REFERENCES</th>
</tr>
</thead>
<tbody>
<tr>
<td>α₁-antitrypsin</td>
<td>(Le et al., 1994)</td>
</tr>
<tr>
<td>Acid phosphatase</td>
<td>(Jannatipour et al., 1998)</td>
</tr>
<tr>
<td>AMPA receptor</td>
<td>(Rubio and Wenthold, 1999)</td>
</tr>
<tr>
<td>ATPase H⁺ vacuolar</td>
<td>(Li et al., 1998)</td>
</tr>
<tr>
<td>ATPase Na,K (α,β subunit)</td>
<td>(Beggah and Geering, 1997)</td>
</tr>
<tr>
<td>CD3ε (TCR receptor)</td>
<td>(Rajagopalan et al., 1994)</td>
</tr>
<tr>
<td>CFTR</td>
<td>(Pind et al., 1994)</td>
</tr>
<tr>
<td>Chicken hepatic lectin</td>
<td>(Feng et al., 1995)</td>
</tr>
<tr>
<td>Chorionic gonadotropin β subunit (human)</td>
<td>(Wada et al., 1994)</td>
</tr>
<tr>
<td>gp80 (MDCK glycoprotein)</td>
<td>(Rozell et al., 1998)</td>
</tr>
<tr>
<td>Gonadotropin receptor</td>
<td>(Chevet et al., 1999)</td>
</tr>
<tr>
<td>Growth factor 2 (fibroblast)</td>
<td>(Le et al., 1994)</td>
</tr>
<tr>
<td>Immunoglobulin</td>
<td>(Grupp et al., 1995)</td>
</tr>
<tr>
<td>INFγ receptor</td>
<td>(Joseph et al., 1999)</td>
</tr>
<tr>
<td>Inositol triphosphate receptor</td>
<td>(Barr et al., 1999)</td>
</tr>
<tr>
<td>Leptin receptor</td>
<td>(Hahn et al., 1997)</td>
</tr>
<tr>
<td>Ltk tyrosine kinase</td>
<td>(Snijders et al., 1997)</td>
</tr>
<tr>
<td>Meprin alpha (human)</td>
<td>(Hahn et al., 1997)</td>
</tr>
<tr>
<td>MHC class II (α,β,βIi)</td>
<td>(Arunachalam and Cresswell, 1995)</td>
</tr>
<tr>
<td>Nicotinic acetylcholine receptor</td>
<td>(Chang et al., 1997)</td>
</tr>
<tr>
<td>P-glycoprotein</td>
<td>(Loo and Clarke, 1994)</td>
</tr>
<tr>
<td>Retinol binding protein</td>
<td>(Bellovino et al., 1996)</td>
</tr>
<tr>
<td>Shaker K⁺ channel protein</td>
<td>(Nagaya et al., 1999)</td>
</tr>
<tr>
<td>Thyroglobulin</td>
<td>(Kim and Arvan, 1995)</td>
</tr>
<tr>
<td>Tyrosinase</td>
<td>(Branza-Nichita et al., 1999)</td>
</tr>
<tr>
<td>Viral proteins</td>
<td></td>
</tr>
<tr>
<td>NSP4 enterotoxin (rotavirus)</td>
<td>(Mirazimi et al., 1998)</td>
</tr>
<tr>
<td>P8 and p16 membrane protein</td>
<td>(Salmons et al., 1997)</td>
</tr>
<tr>
<td>(vaccinia virus)</td>
<td></td>
</tr>
<tr>
<td>NH protein (Newcastle disease virus)</td>
<td>(McGinnes and Morrison, 1998)</td>
</tr>
</tbody>
</table>

ASSOCIATE WITH CALNEXIN AND CALRETICULIN

<table>
<thead>
<tr>
<th>associating molecule</th>
<th>references</th>
</tr>
</thead>
<tbody>
<tr>
<td>Apo B-100</td>
<td>(Wang and White, 2000)</td>
</tr>
<tr>
<td>CD11b</td>
<td>(Huttinger et al., 1999; Sugita et al., 1997)</td>
</tr>
<tr>
<td>Coagulation factor VIII and V</td>
<td>(Kaufman et al., 1997; Pipe et al., 1998)</td>
</tr>
</tbody>
</table>
Fibrilllin (Ashworth et al., 1999)
Fibrinogen (Roy et al., 1996)
Glut 1 glucose transporter (Oliver et al., 1996)
H protein (human factor H) (Ault et al., 1997)
Insulin receptor (Bass et al., 1998)
Integrins (Lenter and Vestweber, 1994; Coppolino and Dedhar, 1999))
MHC class I H chain (Balow et al., 1995; Carreno et al., 1995; Degen et al., 1992; Degen and Williams, 1991; Harris et al., 1998; Rajagopan and Brenner, 1994; Sadasivan et al., 1996)
Myeloperoxidase (Nauseef et al., 1998)
Mucin MUC2 (McCoo1 et al., 1999)
Plasminogen activator (Allen and Bulleid, 1997)
T cell receptor (α,β subunits) (Hochstenbach, 1992)
Transferrin (Ou et al., 1993; Wada et al., 1995; Wada et al., 1997)
Thyroperoxidase (Fayadat et al., 2000)
Von Willebrand factor propeptide (Allen et al., 2000)

Viral proteins
Influenza hemmaglutinin (Hebert et al., 1996)
Gp160, gp120 (from HIV) (Ottoe1 and Moss, 1996)
Hepatitis C glycoprotein (Choukhi et al., 1998)
Envelope glycoprotein F and NH (sendai virus) (Tomita et al., 1999)
G-protein (from VSV) (Hammond and Helenius, 1994; Mathieu et al., 1996)
G1 and G2 glycoprotein (ukunieni virus) (Veijola and Pettersson, 1999)
Glycoprotein (Rabies virus) (Gaudin, 1997)

ASSOCIATE WITH CALRETICULIN

Cyclophilin (Reddy and Atreya, 1999)
Cocaine-sensitive serotonin transporter (Tate et al., 1999)
Dynein (Kraemer et al., 1999)
Perforin (Fraser et al., 1998)
PDI (Gerner et al., 1999)

Viral proteins
Herpes Simplex virus type 1 glycoprotein B (Laquerre et al., 1998)
indicated by the differential susceptibility to proteinase K of CNX-bound versus unbound molecules (Wada et al., 1991). CNX binds to incompletely oxidized transferrin and influenza HA and dissociates at the time that fully disulfide-bonded molecules are formed (Hammond and Helenius, 1994; Ou et al., 1993). Third, a number of protein complexes assemble while in association with CNX. Class I H chain binds β2m when in complex with CNX (Degan et al., 1991). Furthermore, formation of the complete ternary complex is required for efficient dissociation of CNX from class I molecules; incomplete complexes lacking β2m or peptide exhibit prolonged binding to CNX (Degen et al., 1992; Degen and Williams, 1991; Suh et al., 1996). Also class II histocompatibility molecules assemble into a large complex consisting of three invariant chains and two αβ dimers while associated with CNX. Addition of the final αβ dimer correlates with CNX dissociation (Anderson and Cresswell, 1994). Fourth, CNX also interacts in prolonged fashion with misfolded proteins or subunits that are unable to assemble. A number of incompletely assembled proteins such as class I H chain (Degan et al., 1992; Rajagopalan et al., 1994), subunits of the T cell receptor (David et al., 1993) and integrins (Lenter and Vestweber, 1994), remain associated with CNX and are retained in the ER. Misfolded mutant proteins such as the ts045 mutant of VSV G protein (Hammond and Helenius, 1994), a truncated variant of α1-antitrypsin (Le et al., 1994), and the ΔF508 mutant of CFTR (Pind et al., 1994) are all retained in the ER as complexes with CNX.

The first direct evidence that CNX regulates the intracellular transport of its substrates and, hence, functions as a component of the ER quality control system came from experiments in Drosophila cells co-expressing mammalian CNX along with free class I H chains or H chain-β2m heterodimers (Jackson et al., 1994). Co-expression of canine CNX with the mouse H chain and β2m in these Drosophila cells drastically reduced intracellular transport rates of peptide-deficient heterodimers and resembled the very slow kinetics seen in TAP-deficient mouse cells. Furthermore, export of free H chains from the ER was virtually eliminated (Jackson et al., 1994). CNX's ER quality control
function was also demonstrated by expressing either full length CNX or a truncated variant that lacks an ER localization signal in cells expressing the T cell receptor ε subunit or the class I H chain. These subunits were retained in the ER in association with intact CNX, whereas their association with truncated CNX resulted in their redistribution to the Golgi complex or cell surface (Rajagopalan and Brenner, 1994; Rajagopalan et al., 1994).

Crucial evidence regarding the function of CNX as a molecular chaperone was also obtained from studies in cells (Chang et al., 1997; Hammond and Helenius, 1994; Vassilakos et al., 1996). In the absence of CNX in Drosophila cells, newly synthesized class I H chains are either degraded rapidly or they exhibit a substantial degree of aggregation (depending on the isotype examined). However, upon co-expression of CNX, H chains are more stable, and their folding and assembly with β2m are enhanced (Vassilakos et al., 1996). CNX also enhanced the assembly of the nicotinic acetylcholine receptor and its cell surface expression (Chang et al., 1997).

The long awaited evidence that CNX functions as a bona fide molecular chaperone was provided recently by in vitro studies. A soluble form of CNX was shown to exhibit many of the functional characteristics of a classical molecular chaperone. It prevented the aggregation of unfolded proteins when added in stoichiometric amounts, protected a protein against thermal denaturation and enhanced the efficiency of protein folding by maintaining an unfolded substrate in a folding-competent state (Ihara et al., 1999). Similar properties have been described for the Hsp90, Hsp70/40, Hsp60 and small heat shock proteins (Buchner, 1999; Jakob et al., 1995).

The issue of whether CNX discriminates between unfolded and native states of substrate proteins remains controversial. CNX was capable of discriminating between unfolded and native conformational states of some substrate proteins, such as citrate synthase (CS) and malate dehydrogenase (MDH) (Ihara et al., 1999), however, it failed to discriminate between reduced and native conformational states of monoglycosylated RNase B in vitro and in microsomes (Rodan et al., 1996; Zapun et al., 1997).
Furthermore, CNX does not enhance the oxidative folding of RNase B (Zapun et al., 1997). Such controversial results may, however, be explained by CNX's substrate selectivity, a phenomenon well documented for the cytosolic HSP90 and TRiC chaperones (Gething, 1997).

Two properties make CNX unique as a molecular chaperone, its preference for Asn-linked glycoproteins and its membrane disposition. The lectin binding property of CNX was first observed in human hepatoma cells (Ou et al., 1993). In these cells CNX's transient association with a series of glycoproteins was prevented by tunicamycin, an inhibitor of protein N-glycosylation. Moreover, the main non-glycosylated protein secreted by these cells, albumin, did not interact with CNX (Ou et al., 1993). Elegant experiments by Hammond and co-workers established that not only tunicamycin, but also the oligosaccharide processing inhibitors, castanospermine and 1-deoxynojirimycin, prevented the interaction of glycoproteins with CNX (Hammond et al., 1994). Castanospermine and 1-deoxynojirimycin inhibit glucosidases I and II, enzymes that remove glucose residues from the Glc3Man9GlcNAc2 oligosaccharide that is attached to nascent polypeptide chains (Fig.5). In this experiment only the monoglucosylated and not the diglucosylated G protein of the ts045 VSV mutant was found to associate with CNX. Further work confirmed this conclusion and in a large number of cases the addition of glucosidase I and II inhibitors prevented CNX-glycoprotein association in vivo (Vassilakos et al., 1996; Balow et al., 1995; Chang et al., 1997; Kearse et al., 1994; van Leeuwen and Kearse, 1996a; Zhang et al., 1995). Furthermore, in vitro modification or removal of the RNase B oligosaccharide resulted in its dissociation from CNX (Rodan et al., 1996; Zapun et al., 1997).

Direct evidence for CNX's specificity for monoglucosylated oligosaccharides came from experiments with purified immobilized CNX. Recombinant soluble CNX, consisting of the entire ER luminal domain, was shown to interact only with monoglucosylated oligosaccharides when challenged with a mixture of tri-, di-, mono- and un-glucosylated
glycans (Ware et al., 1995). The strongest interaction was observed with
Glc₃Man₉GlcNAc₂, and removal of mannose units from this species reduced the binding
(Vassilakos et al., 1998). No difference in glycan binding was observed between CNX
and CRT (Vassilakos et al., 1998; Spiro et al., 1998).

In addition to interacting with oligosaccharide, there is substantial evidence that
CNX also recognizes the polypeptide portion of incompletely folded glycoproteins.
First, complexes between CNX and several membrane-bound or soluble glycoproteins
cannot be disrupted by treatment with endoglucosidase H which removes the
oligosaccharide from the glycoprotein (Vassilakos et al., 1996). Second, CNX binds to
some proteins that lack Glc₃Man₉GlcNAc₂ oligosaccharides either naturally (Rajagopalan
and Brenner, 1994), have lost them through mutagenesis (Loo and Clarke, 1994) or by
inhibition of glycosylation or oligosaccharide processing (Arunchalam and Cresswell,
1995; Kearse et al., 1994). Third, soluble CNX was showed to interact in vitro with the
polypeptide portion of unfolded proteins, as evidenced by its ability to prevent the
aggregation of glycoforms of soybean agglutinin that do not possess monoglucosylated
oligosaccharides and also the aggregation of the unglycosylated proteins, CS and MDH.
Stable complexes of CNX and unglycosylated CS and MDH could be recovered by size
exclusion chromatography (Ihara et al., 1999).

CNX homologues have been identified and cloned from a wide range of organisms
such as dog (Wada et al., 1991), human (David et al., 1993), worm (Hawn et al., 1993),
plant (Huang et al., 1993), mouse (Schreiber et al., 1994), rat (Tjoelker et al., 1994),
S. pombe (Jannatipour and Rokeach, 1995), S. cerevisiae (de Virgilio et al., 1993; Parlati et
al., 1995), and Drosophila melanogaster (Christodoulou et al., 1997). The mammalian
homologues share 93-98% amino acid identity (Tjoelker et al., 1994). The S. pombe
homologue, Cnp1p, is 40% identical in amino acid sequence, to mammalian CNX’s
(Jannatipour and Rokeach, 1995; Parlati et al., 1995) whereas the S. cerevisiae homologue,
Cne 1p, is 24% identical. Cne 1p unlike other CNX’s is N-glycosylated, does not bind
calcium and lacks an ER retention motif (Parlati et al., 1995). Deletion of this gene in
*S. cerevisiae* results in lack of detectable phenotype suggesting that its expression is not
essential (Parlati et al., 1995). Similarly, the CNX deficient mammalian cell line,
CEM.NKR, grows well, assembles class I molecules efficiently and is apparently
unimpeded by the absence of CNX (Sadasivan et al., 1995). In contrast, deletion of the
*S. pombe* Cnx 1p gene leads to a lethal phenotype (Jannatipour and Rokeach, 1995; Parlati
et al., 1995). Two CNX homologs have also been described in testis as calmegin (CMG)
or CNX-t (Watanabe et al., 1994). Both showed lectin properties similar to CNX.

A few years ago, a protease-resistant core of CNX, that retains biological activity
as shown by its binding to monoglucosylated oligosaccharides and its functional
interaction with ERp57, was identified and crystallized (Hahn et al., 1998). However, the
obtained crystals diffracted only to medium resolution, with the best found occasionally
diffracting to 2.8Å resolution. Since the structure is still unavailable, I will described what
is known about CNX molecules based on cDNA and biochemical analysis.

Canine CNX is predicted to be a 573 amino acid, type I integral membrane protein
with a molecular weight of 65.4 kDa. The protein consists of a large 463 amino acid
lumenal domain, a single transmembrane domain and a 90 amino acid cytoplasmic tail
(Fig. 8) (Wada et al., 1991). Mammalian CNX is remarkably acidic with a predicted pI of
~4.5. CNX is targeted to the ER via an N-terminal signal sequence that is cleaved upon
translocation. The lumenal domain shares 29% sequence identity with CRT, the major
soluble calcium binding protein in the lumen of the ER (discussed in the next section) and
64% sequence identity with calmegin, a calcium-binding protein that is expressed in
spermatogenesis (Watanabe et al., 1994).

A central domain in the luminal portion of CNX contains two tandemly repeated
motifs (Fig 8). The number of these repeats vary between different members of the CNX
Figure 8. Functional sites in CNX and CRT. Segments of sequence homology between the two proteins are indicated by rectangles. The sequence motifs 1 and 2 repeated 4 times in CNX and 3 times in CRT are located within the homologous regions and are depicted by the numbers 1 and 2. The ERp57 binding site and the high affinity site for Ca$^{2+}$ binding are localized to the repeat region. The substrate selection site was mapped to the transmembrane domain of CNX. The polypeptide substrate binding site is mapped to the membrane proximal region.
family. CNX and calmegin contains four, and CRT three of these proline-rich elements (Wada et al., 1991; Smith and Koch, 1989; Watanabe et al., 1994). The consensus sequence of the motifs are: motif 1, I[XX][DPXAKPDW][XD/E] and motif 2, G[XX][WXX][P][IX]NP. The region encompassing the motif 1 repeats contains a high affinity calcium binding site. Calcium binding was shown to be essential for the lectin properties of CNX since EGTA treatment abrogated oligosaccharide binding (Vassilakos et al., 1998). The luminal domain of CNX has at least one disulfide bond, that is not essential for its lectin properties (Vassilakos et al., 1998). ATP has previously been shown to bind to CNX although the functional effect remains unclear. Of all CNX homologous, only that from *S. cerevisiae* possesses a consensus sequence for Asn-linked glycosylation within the luminal domain.

The C-terminal cytosolic domain of CNX contains four defined regions; a lysine-rich juxtamembrane domain, a glutamic acid-rich acidic domain, a phosphorylation domain and an ER localization motif (RKPRRE). The phosphorylation domain contains well conserved potential serine phosphorylation sites, three in CNX and four in calmegin. Mass spectometry analysis revealed *in vivo* CNX phosphorylation sites, two of them were within protein kinase CK2 sites while the third was within a PKC/PDK site. It is thought that phosphorylation of the CNX cytosolic domain can regulate the ribosome-binding properties of CNX and its localization to the translocon (Chevet et al., 1999; Wong et al., 1998).

### 1.3.4.3.2 CRT

CRT (CRT) was first identified as a Ca\(^{2+}\) binding protein of the muscle sarcoplasmic reticulum (SR) (Ostwald et al., 1974). Only more than a decade later did it become clear that this protein is highly abundant in nonmuscle tissue and that it is one of the major Ca\(^{2+}\) binding proteins of the ER (Michalak et al., 1992). CRT is an ubiquitous
protein synthesized by all mammalian cells tested except erythrocytes. Surprisingly there is still a considerable controversy in the literature concerning the subcellular localization of CRT. CRT was originally identified as an ER/SR membrane protein (Ostwald et al., 1974). Numerous studies confirmed the ER localization of CRT in many diverse species, including plants (Crofts et al., 1998; Michalak et al., 1999; Michalak et al., 1992). The protein has also been localized to the cytoplasmic granules of cytotoxic T cells (Michalak et al., 1999), sperm acrosomes (Nakamura et al., 1993), tick saliva (Jaworski et al., 1996), the cell surface (Arosa et al., 1999; Gray et al., 1995) and secreted in the bloodstream (Sueyoshi et al., 1991). Subsequent to its broad localization in the cell, CRT has been implicated to participate in many cellular function within and outside of the ER, such as lectin-like chaperoning, Ca\(^{2+}\) storage and signalling, regulation of gene expression, cell adhesion and autoimmunity.

The initial indication that CRT, the soluble homologue of CNX, may function as a molecular chaperone came from the elucidation of CNX functions. Like CNX, anti-CRT antibodies could co-isolate a large number of newly synthesized proteins (Peterson et al., 1995). CRT, similarly to CNX, was found to associate transiently with these proteins in the ER (Peterson et al., 1995). Soon after, several CRT substrates were identified and the population of proteins that bind to CRT was found to partially overlap with those that bind CNX (Table III) (Allen and Bulleid, 1997; Bass et al., 1998; Hammond and Helenius, 1995; Hebert et al., 1996; Hebert et al., 1995; Krause and Michalak, 1997). Furthermore, it was established that CRT interacts primarily with incompletely folded or assembled intermediates and conformationally trapped forms (Peterson et al., 1995). Ultimately, the molecular chaperone function of CRT was directly demonstrated by using purified components \textit{in vitro} (Saito et al., 1999). CRT was shown to suppress the thermal aggregation of a number of different substrates such as IgY, CS and MDH. In addition, CRT also conferred protection against thermal inactivation and maintains substrates in a folding-competent state (Saito et al., 1999). Furthermore, CRT was shown to form stable
complexes with unfolded substrates but not with native proteins (Svaerke and Houen, 1998). Detailed analysis of the substrates shared by CNX and CRT revealed that these chaperones interact with different folding or assembly intermediates of the substrate protein and form a link of lectin-like chaperones handing over the glycoprotein from one to the other to ensure proper folding and assembly. Such a collaboration is best illustrated with the assembly of the human class I complex where β2m binding to H chain displaces CNX and the resulting H chain-β2m heterodimer binds CRT. In addition, CRT forms structural and functional complexes with other ER molecular chaperones such as BiP, GRP94, PDI, and ERP57 (Hebert et al., 1995; Kim and Arvan, 1995; Kuznetsov et al., 1997; Linnik and Herscovitz, 1998; Oliver et al., 1999). Complexes of ER chaperones including CRT are associated with maturation of thyroglobulin (Kim and Arvan, 1995; Kuznetsov et al., 1997) and apolipoprotein B (Linnik and Herscovitz, 1998). Binding of CRT to most substrates is oligosaccharide dependent and requires the trimming of glucose residues from Asn-linked core glycans by glucosidase I and II (Peterson et al., 1995; Otteken and Moss, 1996)

CRT's lectin binding properties were first demonstrated by the co-purification of CRT with endomannosidase on a Glc-Man-containing matrix (Spiro et al., 1996) Direct evidence that CRT is a lectin with Glc1Man9GlcNAc2 specificity was provided by studies carried out with immobilized CRT and a series of radiolabeled oligosaccharides derived from N-linked carbohydrate units. Although optimal binding occurred with Glc1Man9GlcNAc2, substantial interaction with CRT was retained after sequential trimming of the polymannose portion down to the Glc1Man5GlcNAc2 stage (Spiro et al., 1996). Furthermore, binding of a monoglycosylated chicken IgG and the monoglycosylated oligosaccharide; Glc1Man9GlcNAc2 with CRT was recently measured using surface plasmon resonance and the K_a value was found to be 4.87x10^5 M^-1 at 25°C (Patil et al., 2000). Using deletion mutants of CRT, expressed in E. coli as glutathion S-
transferase fusion proteins, the lectin site in CRT was localized to the N- and C-domain, with 60% of lectin function mapped to this region (M. Leach, personal communication).

It should be emphasized, that like CNX, monoglycosylated high-mannose carbohydrates may not be the only requirement for substrate binding to CRT. Several lines of evidence indicate that the functions of CRT may involve protein-protein interactions. CRT binds directly to several non-glycosylated proteins including PDI (Baksh et al., 1995), ERp57 (Corbett et al., 1999), perforin (Andrin et al., 1998), cytoplasmic domain of integrin (Coppolino et al., 1995) and the DNA-binding domain of steroid receptors (Burns et al., 1994). Furthermore, glucosidase I and II inhibitors do not affect association between CRT and factor VIII (Pipe et al., 1998) or between CRT and mucin (McCool et al., 1999). Association of CRT with non-glycosylated peptides has also been demonstrated both in vitro and in vivo (Basu and Srivastava, 1999; Nair et al., 1999). CRT has been shown to discriminate in its binding between native and non-native conformations of non-glycosylated proteins in vitro (Svaerke and Houen, 1998). Finally, CRT was demonstrated to function as a molecular chaperone for non-glycosylated proteins discriminating between protein conformational states through polypeptide based interactions (Saito et al., 1999). Thus, like CNX, CRT possesses both lectin and molecular chaperone functions.

CRT is a 46 kDa protein with an N-terminal cleavable signal sequence which is responsible for targeting CRT to the ER lumen. Depending on the species, CRT may have one or more potential N-linked glycosylation sites. The glycosylation pattern appears not to be conserved. The glycosylation of CRT is more common in plants than animal cells (Navazio et al., 1996). Interestingly, heat shock may trigger glycosylation of CRT (Jethmalani and Henle, 1998); however, the functional consequence of this stress-induced glycosylation are presently not clear. CRT has three cysteine residues (Cys 88, Cys 120 and Cys 146) all located in the N-terminal part of the protein. Importantly, the location of these Cys residues is conserved in all cloned CRT from higher plants to humans.
A disulfide bridge is formed between Cys120 and Cys146 (Matsuoka et al., 1994) and plays a significant role in maintaining a functional lectin site.

Structural analysis of CRT suggests that CRT is an asymmetric elongated molecule likely to be flexible in solution (Bouvier and Stafford, 2000). Furthermore, CRT is monodisperse and monomeric in solution (Bouvier and Stafford, 2000). CRT has at least three domains termed N-, P-, and C-domain (Fig. 8). The amino acid sequence of the N-domain is extremely conserved in all CRT proteins (Michalak et al., 1999). The N-domain binds Zn$^{2+}$ (Baksh et al., 1995b; Heilmann et al., 1993) and Zn$^{2+}$ binding involves four of the histidine residues found in this domain (Baksh et al., 1995a). The N-domain interacts with the DNA-binding domain of the glucocorticoid receptor in vitro (Burns et al., 1994), with rubella virus RNA (Singh et al., 1994), $\alpha$-integrin (Rojiani et al., 1991), protein disulfide isomerase (PDI) and ERp57 (Baksh et al., 1995a; Corbett et al., 1999). Interaction of the N-domain with PDI inhibits PDI chaperone activity (Baksh et al., 1995a), but enhances ERp57 function (Zapun et al., 1998). These protein-protein interactions are regulated by Ca$^{2+}$ binding to the C-domain of CRT (Corbett et al., 1999).

The P-domain of CRT is comprised of a proline-rich sequence with three repeats of motif 1, followed by three repeats of motif 2. The NMR structure of the rat CRT P-domain (residues 189-288) shows a hairpin fold that involves the entire polypeptide chain, and has the two chain ends in close spatial proximity (Ellgaard et al., 2001). This region of the protein binds Ca$^{2+}$ with high affinity, but low capacity ($K_d \approx 1\mu M$; 1mol Ca$^{2+}$/mol of protein) and these repeats are essential for the high affinity Ca$^{2+}$ binding (Baksh et al., 1995b; Tjoelker et al., 1994). The P-domain interacts also with PDI (Baksh et al., 1995b; Corbett et al., 1999) and perforin, a component of the cytotoxic T-cell granules (Andrin et al., 1998). The P-domain is the most interesting region of CRT because it is the major site of ERp57 binding (Leach et al., 2001) and has high sequence similarity to CNX, calmegin and CALNUC, a Golgi Ca$^{2+}$ binding protein (Lin et al., 1998).
The C-domain is highly acidic and terminates with the ER retrieval sequence, KDEL (Fliegel et al., 1989; Smith and Koch, 1989). This domain binds over 25 mol of Ca\(^{2+}\)/mol of protein with a relatively low affinity of \(K_r\approx 2\) mM (Baksh and Michalak, 1991). Ca\(^{2+}\) binding to this domain is thought to regulate CRT interaction with PDI, ERp57 and perhaps other chaperones (Corbett et al., 1999).

To date, cDNA and genes encoding CRT have been isolated from several vertebrates (human, mouse, rat, rabbit, bovine, \(X\ laevis\)), invertebrates (\(D.\ melanogaster\). \(C.\ elegans\). \(A.\ californica\) (sea slug) and plants (barley, maize) (Michalak et al., 1999). There is no CRT gene in yeast and prokaryotes. The CRT gene knockout is lethal for the mouse embryo (12-14 days old) and death is due to circulatory problems (Mesaeli et al., 1999). It is not clear which of CRT’s many functions are responsible for this phenotype but it may be the role of CRT in modulation of Ca\(^{2+}\) homeostasis or chaperoning.

1.4 Models of CNX and CRT functions

In spite of this abundance of \textit{in vivo} and \textit{in vitro} data, the mechanism whereby CNX or CRT enhances glycoprotein folding is poorly understood. There are two models, depicted in Fig. 9, that summarize the current thinking on how CNX or CRT interacts with glycoprotein folding intermediates, the ‘lectin-only’ model initially proposed by A. Helenius and co-workers (Hammond et al., 1994) and the ‘dual binding’ model favored by our laboratory (Williams, 1995).

1.4.1 Lectin Only Model

When oligosaccharides are added to a newly synthesized glycoprotein from a dolichol phosphate-bound precursor (Hubbard and Ivatt, 1981), one of the branched chains of the N-linked glycan terminates in three glucose residues (Fig.5). These are successively
removed in the ER by the action of two glucosidases, the membrane bound glucosidase I and the ER lumenal glucosidase II (Fig.5). CNX and CRT bind to the N-linked glycan intermediate bearing a single glucose residue, Glc$_1$Man$_9$GlcNAc$_2$. Precisely how CNX/CRT facilitate folding is not understood, but they are postulated to be involved in a binding and release cycle together with a specific glucosyl transferase, the ER resident enzyme UDP-glucose: glycoprotein glucosyltransferase (UGGT) (Fig.9A). This enzyme can discriminate between native and non-native forms of glycoproteins (Sousa and Parodi, 1995), which is actually a characteristic property of chaperones. In this scheme, following interaction with CNX/CRT, the glucose residue on the monoglycosylated N-linked glycan is removed by glucosidase II. If after release from CNX/CRT the protein is correctly folded it will leave the ER. However, if the glycoprotein has not achieved its native conformation, it is recognized by UGGT, and is reglucosylated. Reglucosylation leads to the second round of interaction with CNX/CRT and another attempt at folding. In this cycle, UGGT plays the role of a folding sensor and only when a native conformation is finally achieved the N-linked glycan is not reglucosylated and the glycoprotein is released from further CNX/CRT interactions. In this model, CNX/CRT function as "holders" of substrate proteins and the observed molecular chaperone function is achieved by recruitment of other chaperones and folding enzymes. Indeed, CNX/CRT bind to the thiol oxidoreductase ERp57, an interaction that promotes disulfide bond formation (High et al., 2000). This model is supported by the finding that glucosidases I and II inhibitors block CNX/CRT substrate interactions and, if added after complexes are formed, can block substrate dissociation (Zhang et al., 1997). Furthermore, CNX/CRT were shown to dissociate from complexes with RNase B in in vitro studies following digestion with glucosidase II (Rodan et al., 1996). Finally, in vitro experiments with RNase B demonstrated that CNX/CRT do not discriminate between reduced and native forms of the enzyme (Zapun et al., 1997).
Figure 9. **Lectin-only model.** Glucose residues (G) in the N-linked oligosaccharides are trimmed by glucosidases I and II. When trimmed to the monoglucosylated form, the oligosaccharides mediate the binding of the glycoprotein to CNX or CRT. CNX and CRT function solely as lectin. Release of the substrate glycoprotein depends on hydrolysis of the remaining glucose residue by glucosidase II. If the protein is not completely folded, the glycans are selectively re-glucosylated by the glucosyltransferase, allowing the oligosaccharide to rebind to CNX. Once the glycoprotein reaches its mature conformation, it is no longer recognized by the glucosyltransferase.

**B. Dual binding model.** Following removal of two glucose residues, newly synthesized glycoproteins contact CNX or CRT via their monoglucosylated oligosaccharide chains. Carbohydrate-mediated association and stronger protein-protein interactions occur between the chaperone and peptide elements exposed on the surface of the unfolded substrate protein. Release of the substrate depends on conformational changes that eliminate the protein-protein interactions. Following the release, the last glucose is removed by glucosidase II and the reglucosylation cycle begins.
1.4.2 Dual Binding Model

In the dual binding model, CNX/CRT function as both a lectin and as a molecular chaperone (Fig. 9B). According to this model (Williams, 1995), the specificity of these molecular chaperones is guided by their oligosaccharide binding site which brings the protein substrate in proximity to a polypeptide binding site that is able to recognize unfolded glycoproteins. This model incorporates the features of the lectin-only model with UGGT representing one of the protein conformational sensors. However, the dual model requires that the dissociation of CNX/CRT complexes occurs through conformational change in the polypeptide binding site followed by the action of glucosidase II. In this model, CNX/CRT play roles as second sensors for substrate folding. They function as classical chaperones to prevent aggregation of the substrate proteins. Cycles of binding and release continue until the glycoprotein acquires a conformation no longer recognized by UGGT and CNX/CRT.

Support for this model comes from the finding that the association of CNX and a variety of glycoproteins persist after enzymatic removal of the oligosaccharide (Zhang et al., 1995; Vassilakoss, 1998). Furthermore, CNX/CRT were found to interact with several non-glycosylated proteins such as PDI, ERp57, T cell receptor ε subunit, unglycosylated H-2Lδ molecules (Baksh et al., 1995; Leach et al., 2000; Rajagopalan and Brenner, 1994; Carreno et al., 1995) and also to non-glycosylated peptides in vivo and in vitro (Basu and Srivastava, 1999; Svaerke and Houen, 1998). Finally, CNX/CRT have been reported to behave as classical chaperones in cell-free folding assays. Both molecular chaperones were shown to inhibit the aggregation of not only denatured glycoproteins devoid of the specific monoglucosylated oligosaccharides species, but also non-glycosylated proteins (Ihara et al., 1999; Saito et al., 1999).
Discussion about the involvement of CNX/CRT in class I biogenesis cannot be complete without understanding the nature of lectin interactions. In the following section I will introduce a few characteristics of lectins relevant to the topic of CNX/CRT functions and especially to their substrate specificity.

1.5 Lectins

Lectins are a structurally diverse group of proteins defined as non-enzymatic, sugar binding proteins. Comparison of various lectin structures indicates that the ability to bind carbohydrates has evolved independently on several separate occasions in the context of various protein structural frameworks. This conclusion is consistent with the lack of sequence similarity among various groups of lectins (Drickamer, 1995). Although a large number of lectins from plant, animal, bacteria and viruses have been characterized, our understanding of lectin functions is widely variable.

The function of many plant lectins that have been known for some time remains somehow enigmatic. In contrast, in the animal kingdom, a number of important biological phenomena depend on carbohydrate-protein interactions. Lectins such as hemagglutinins of influenza and other viruses are utilized by pathogens as a means of attachment to eukaryotic cell surfaces (Sauter et al., 1992; Weis et al., 1988). Several animal lectins serve as part of the innate host immune system by selectively binding to the surfaces of bacterial and viral pathogens and initiating steps toward their neutralization. Thus, selective recognition of sugars can provide a means of distinguishing self from non-self. Selectins, a group of lectins residing in cell membranes, are known to mediate the initial recognition of immunologically important events such as lymphocyte routing and neutrophil and monocyte recruitment to injury sites (Lasky, 1992). Several studies have shown that 'rolling' of neutrophils and monocytes along capillary endothelia results from reversible interactions between selectins presented on endothelial cells and leukocytes (Alon et al., 1997; Somers et al., 2000). Other animal lectins such as CNX/CRT are involved in ER quality control (Trombetta and Helenius, 1998; Williams, 1995). Sorting
of some newly synthesized glycoproteins within the luminal compartment of the ER is mediated by the lectin ERGIC 53 (Nichols et al., 1998) and delivery of newly synthesized lysosomal enzymes from the Golgi apparatus to lysosomes requires the recognition of mannose-6-phosphate on these enzymes by a specific lectin receptor (Kornfeld, 1987). Although the number of animal lectins continues to increase, a recent classification based on structural considerations indicates that most fall into one of four major groups: C-type or Ca\(^{2+}\)-dependent lectins, galectins, thiol dependent β-galactoside-binding lectins (S-type lectins), and P-type lectins.

In spite of the enormous diversity, lectin binding sites that have been analyzed so far share a few key features, which are presumably related to the fact that the interaction of lectins with individual sugars is of modest affinity. Most of the binding sites are relatively shallow indentations on the protein surface. The sugar-binding activity can be ascribed to a limited portion of most lectin molecules, typically a globular carbohydrate-recognition domain which is ~135 amino acids for C-type and ~130 amino acids for galectins (Drickamer, 1988). This arrangement differs from those in which monosaccharides are completely surrounded by the protein, as in the bacterial periplasmic sugar-binding proteins and hexokinase (Fletterick et al., 1975; Quiocco, 1993), and the active sites of many other enzymes that utilize sugars or sugar conjugates as substrates, which are often deep clefts in the protein (Janakiraman et al., 1994).

An essential part of lectin-carbohydrate interactions is hydrogen bonding to sugar hydroxyl groups, and an important aspect of differential binding of sugars is selective binding to epimeric hydroxyls. Two common features of hydrogen bonding between carbohydrates and lectins are evident. First, charged or polar planar groups form the majority of hydrogen-bond donors and acceptors with sugar hydroxyl groups. The oxygen atom of the hydroxyl can act as an acceptor of two hydrogen bonds and as a donor of a single hydrogen bond. Generally, one acidic side chain is used as a hydrogen bond acceptor from one or two sugar hydroxyl groups. Hydrogen-bond donors come primarily from main-chain amide groups and the side-chain amide group of asparagine.
and, less frequently, glutamine. Second, the sugar-binding sites are preformed, in the sense that few changes occur upon saccharide binding. In no case, have global changes in protein structure been observed; instead small movements are restricted to the immediate vicinity of the sugar. High-resolution structures of unliganded lectins reveal that discrete water molecules form hydrogen bonds with those polar atoms that form hydrogen bonds with the sugar in the lectin-sugar complexes. These water molecules serve as mimics of the hydrogen bonding pattern of sugar hydroxyls (Weis and Drickamer, 1996). Other features common to lectin binding sites are also evident. The van der Waals contacts between sugar and protein often include packing interactions with aromatic amino acid side chains. Such packing is particularly common in the Gal-specific sites, but is also observed in legume lectins that bind Man and Glc. In all lectin-Gal complex structures, the apolar patch of the Gal face packs against the face of tryptophan or phenylalanine. The coordination bonds that link sugar to the Ca\(^{2+}\) in C-type animal lectins contribute substantially to the affinity of this binding site. It will be interesting to see how many of the features of the lectin-carbohydrate interactions are shared by CNX and CRT as analyzed by crystallography or NMR.

Lectins often bind to natural polysaccharides with high affinities, reflected in dissociation constants in the nanomolar range, yet their interactions with simple monosaccharide is far weaker. These carbohydrates bind to the lectins at a primary binding site with dissociation constants in the 0.1-1.0 mM range. Increase in affinities for oligosaccharides derive in many instances from extended binding sites where more than just terminal sugar residues of an oligosaccharide make contact, either directly or through bridging water molecules, with the surface of the lectin. However, the affinities resulting from these additional contacts are still in the micromolar range at best (Rini, 1995). Dissociation constants in the nanomolar range are described only for lectins with more than one lectin subunit either contained on the same polypeptide chain or from separate chains of an oligomer which interact with a multivalent ligand. Many lectins achieve much
higher affinity by clustering several similar or identical binding domains, often by formation of polypeptide oligomers.

For a number of lectins the division of lectin into carbohydrate binding (CRD) and effector domains reflects the coming together of an effector function with sugar-binding as a means to target a desired activity with greater precision. Some lectins are known to have both carbohydrate-binding and peptide-peptide interaction sites. For the lymphocyte low affinity IgE Fc receptor, CD23, binding of IgE is mediated by the CRD-like portion of the receptor and requires Ca$^{2+}$, although deglycosylation of IgE does not affect binding to the receptor (Bettler et al., 1989). Binding is inhibited by specific peptides from the Fc region, but not by sugar (Richards and Katz, 1990). Peptide-peptide interaction is also important for the binding of the C-type lectin, Ly49A, with its ligand, murine class I molecules. Natural killer cell function is controlled by interaction of Ly49A, a NK receptor with MHC class I molecules expressed on target cells. The peptide-peptide interaction was supported by studies showing that Ly49 binds independently of carbohydrate and Ca$^{2+}$ and shows specificity for MHC class I but not bound peptide (Natarajan et al., 1999). Furthermore, the MHC specificity of Ly49 A was clearly tied to protein allotype. Finally, the crystals obtained for the dimers of the carbohydrate recognition domain of Ly49A plus the entire extracellular part of the class I (H-2D$^d$) were obtained in the absence of N-liked glycans (Tormo et al., 1999). However, the analysis of the crystal structure of the sugar-free complex point almost directly at the involvement of the carbohydrate in the binding of Ly49A (Natarajan et al., 1999). Furthermore, the C-type lectin-like interaction of Ly49A was established by showing that high molecular weight sulfated carbohydrates inhibit ligand binding (H-2D$^d$) to the receptor in cell adhesion assay (Kane, 1994; Chang et al., 1996). Also, direct binding experiments established that Ly49A has specificity for the complex carbohydrate fucoidan (Daniels et al., 1994). When the role of carbohydrate in the Ly49A/MHC class I interaction was examined in vivo, the carbohydrate at position 176 of the murine H chain contributed to both adhesion and functional protection from NK cytolysis (Lian et al., 1998).
Also, when P-selectins were analyzed in their interaction with PSGL-1, a mucin-like homodimeric glycoprotein expressed by virtually all subset of leucocytes, two conformations were detected: a low affinity form that supports sialyl Lewis^x-like glycan (SLEx) binding and a high affinity form that is capable of making extended contacts with PSGL-1. High affinity P-selectin binding requires both a SLEx-containing O-glycan and one or more tyrosine sulfate residues within the anionic N terminus of the PSGL-1 polypeptide (Pouyani and Seed, 1995; Sako et al., 1995; Wilkins et al., 1996). Both tyrosine sulfation and SLEx modification of this region of PSGL-1 are required for P-selectin function under the influence of shear flow (Goetz et al., 1997; Ramachandran et al., 1999).

1.6 Experimental Objectives

The experiments described in this thesis were designed to investigate the role of two ER resident proteins, CRT and ERp57 in the biogenesis of MHC class I molecules. Furthermore, as our understanding of the processes involved in protein folding and assembly expanded, and the mechanism of some of the molecular chaperones became clear the possibility arose of addressing some of the controversies regarding the mode of action of the ER molecular chaperones CNX and CRT. Within this framework, I designed experiments to investigate 1) the role of CRT in the biogenesis of murine class I molecules, 2) the contribution of transmembrane and cytoplasmic domains of CNX in its molecular chaperone functions, 3) the contribution of transmembrane and cytoplasmic domains of CNX to its substrate specificity, 4) the effect of N-glycans on the interaction between substrate protein and CNX/CRT and 5) the involvement of ERp57 in the early steps of H chain folding.

In Chapter 2, I describe the studies on the molecular chaperone functions of CRT and soluble CNX. By using the Drosophila expression system, CRT's ability to function as a molecular chaperone for MHC class I molecules was directly demonstrated. Also, soluble variants of CNX were shown to retain chaperone functions. When truncated
forms of CNX were utilized to map sites involved in substrate selection. CNX expressed as a soluble protein associated with an array of glycoproteins that resemble those that bind to CRT. Conversely, membrane-anchored CRT bound to similar set of glycoproteins as membrane-bound CNX. Thus, the distinct topologies of CNX and CRT contribute to their distinct substrate selectivities.

Chapter 3 describes the contribution of N-glycans to the interaction of substrate proteins with CNX/CRT in vivo. While there is general agreement that the lectin function versus polypeptide-based association of CNX/CRT is important for glycoprotein folding, the question of whether these two proteins function solely as lectins or both as lectins and as a classic chaperones remains to be answered. Recent in vitro studies with purified components lend support to the notion that CNX/CRT bind to polypeptide segments of unfolded or partially folded proteins. In this chapter, the requirement for monoglucosylated oligosaccharides for interaction with CNX/CRT in vivo was re-examined. Evidence is presented which demonstrates that a large number of proteins retain the ability to bind to CNX, after the formation of monoglucosylated glycans is blocked, and that an unglycosylated substrate associates with CNX in vivo in a physiologically relevant way.

In Chapter 4 the involvement of ERp57 in class I biogenesis is explored using the Drosophila expression system. The objective was to monitor the formation of disulfide bonds within the class I H chain in the presence or absence of ERp57. We showed that in insect cells, contrary to what was reported in vitro, neither the thiol oxidoreductase nor the molecular chaperone function of ERp57 could be detected.

Finally, Chapter 5 provides an overview summary and discussion of the thesis projects.
CHAPTER II
Functional Relationship Between CRT, CNX
and the Endoplasmic Reticulum Luminal Domain of CNX

---

1 Most of the results presented in this chapter have been published in Ursula G. Danilczyk, Myrna F. Cohen-Doyle, and David B. Williams (2000) The Journal of Biological Chemistry vol.275, pp. 13089-13097. Reprinted with the permission from Ursula G. Danilczyk, Myrna F. Cohen-Doyle, David B. Williams and The American Society for Biochemistry and Molecular Biology, Inc.
2.1 Introduction

CNX and CRT are lectins that bind transiently to many newly synthesized membrane and soluble proteins as they pass through the ER (Williams, 1995; Helenius et al., 1997). Both CNX and CRT exhibit a marked preference for binding to Glc$_1$Man$_9$GlcNAc$_2$ oligosaccharides on Asn-linked glycoproteins (Spiro et al., 1996; Ware et al., 1995). Indeed, treatment of cells with tunicamycin or with castanospermine, an inhibitor which prevents the formation of the Glc$_1$Man$_9$GlcNAc$_2$ oligosaccharide, abrogates the association of CNX and CRT with many glycoproteins (Ou et al., 1993; Hammond et al., 1994; Hammond and Helenius, 1994; Kearse et al., 1994; Peterson et al., 1995).

CNX is thought to function as a molecular chaperone since its expression enhances the folding and assembly of class I histocompatibility molecules (Vassilakos et al., 1996), the nicotinic acetylcholine receptor (Chang et al., 1997), and the vesicular stomatitis G glycoprotein (Hammond and Helenius, 1994). It also participates in quality control, retarding the export of incompletely assembled protein subunits from the ER (Jackson et al., 1994; Rajagopalan et al., 1994; Rajagopalan and Brenner, 1994). CRT is believed to function in a similar manner since the simultaneous inhibition of CNX and CRT binding by castanospermine treatment is accompanied by impaired folding and subunit assembly, more rapid degradation, and premature release of glycoproteins from the ER in a variety of model systems (Moore and Spiro, 1993; Tector and Salter, 1995; Hebert et al., 1996; Vassilakos et al., 1996; Chang et al., 1997; Zhang et al., 1997; Bass et al., 1998; Toyofuku et al., 1999). However, CRT's individual role in these processes has never been examined.

Given the identical lectin specificities of CNX and CRT, it is not surprising that there is overlap in the glycoproteins that they bind and that they can, in some instances, associate simultaneously with the same glycoprotein (Helenius et al., 1997). However, it is clear from an examination of the overall spectrum of glycoproteins co-isolated with CNX or CRT that there are distinct differences in binding specificity (Peterson et al., 1995; Wada
et al., 1995; van Leeuwen and Kearse, 1996b). Furthermore, it has been demonstrated that
the vesicular stomatitis virus G glycoprotein binds to CNX but not to CRT (Peterson et
al., 1995), that CRT dissociates more rapidly than CNX as the folding/assembly of the T
cell receptor (van Leeuwen and Kearse, 1996b) and the influenza virus hemagglutinin
(Hebert et al., 1997) proceeds, and that CNX and CRT act at different stages in the
assembly of class I histocompatibility molecules (Noessner and Parham, 1995; Sadasivan
et al., 1996). These observations raise the question of what the functional relationship is
between CNX and CRT. Do they possess distinct functions that are utilized at different
stages in glycoprotein biogenesis? Alternatively, are they functionally interchangeable but
bind differentially to certain glycoproteins by virtue of their distinct membrane versus
soluble dispositions or through differences in polypeptide binding specificity?

To address these questions, we first asked whether CRT alone is capable of enhancing
protein folding and participating in quality control processes using the well characterized
mouse class I histocompatibility molecule as a model glycoprotein. We then compared the
results with those previously obtained for CNX to determine the extent to which the
functions of these two proteins are interchangeable. Furthermore, we examined the
influence of the different topological environments of CNX and CRT by removing the
cytoplasmic and transmembrane segments of CNX and assessing the impact on its
chaperone/quality control functions and its substrate binding specificity. We found that
CRT does indeed function as an apparent chaperone and component of the ER quality
control machinery and that these functions are largely interchangeable with those of CNX.
Furthermore, CNX retains its functions when expressed as a soluble molecule but its
substrate specificity is altered to resemble that of CRT.

2.2 Materials and Methods

2.2.1 Cell Lines and Antibodies

*Drosophila melanogaster* Schneider cells were maintained in Schneider’s insect
medium (Sigma) with 10% FBS and antibiotics. Stably transfected derivatives were cultured in the same medium supplemented with 500 μg/ml Geneticin (GIBCO-BRL). Mouse L cells were grown in Dulbecco’s modified Eagle’s minimum essential medium supplemented with 10% FBS and antibiotics.

The following mAb were used for the isolation of class I molecules: mAb 20-8-4S which reacts with H-2Kb heavy (H) chains associated with β2-microglobulin (β2m) (Ozato and Sachs, 1981), mAb B22-249.R1 which recognizes H-2Db H chains associated with β2m (Lemke and Hammerling, 1979), and mAb 28-14-8S which recognizes a conformational epitope in the α3 domain of free or β2m-associated Db H chains (Ozato et al., 1980). A rabbit antiserum (anti-8) directed against the C-terminus of the H-2Kb H chain which reacts with all conformational states of Kb was provided by Dr. Brian Barber, University of Toronto (Smith et al., 1986). Unassembled mouse class I heavy chains were isolated using a rabbit antiserum (anti-HC) provided by Dr. Hidde Ploegh, Harvard University (Machold et al., 1995). A rabbit antiserum raised against the C-terminal 14 amino acids of CNX was used to isolate full-length CNX (Jackson et al., 1994) whereas CNX mutants lacking the C-terminus were detected with a rabbit antiserum (αpp90) directed against the N-terminal 268 residues of CNX (provided by Dr. Ikuo Wada, Sapporo Medical University). MAb 111CA5 was used to detect influenza hemagglutinin (HA)-tagged CNX and CRT mutants and was provided by Dr. Paul Hamel, (University of Toronto).

2.2.2 Construction of CNX Mutants and Expression in Drosophila Cells

Figure 10 summarizes the C-terminal truncation mutants of canine CNX that were generated and expressed in Drosophila melanogaster cells. In the Δcyt mutant, CNX’s cytoplasmic tail was replaced with that of the much shorter adenovirus E3/19K glycoprotein resulting in a protein consisting of CNX residues 1-486 fused to the sequence –KYKSRRSFIDEKKMP. The E3/19K cytoplasmic tail contains a functional ER localization sequence (Jackson et al., 1990). Two PCR
primers, one upstream from the BspM II restriction site 5'CCCGAAGATACCAAATCCGG3' and the other downstream from the Tm segment 5'CCGCGGATCCCTATGGCATCTTTTTCGTCTATAAGGCTCCTCTACTTTT ATACTTTCCAGAGCAGCAGAAGAGG3' were used to generate a 564 bp PCR fragment flanked by BspM II and Bam HI sites that encodes the new 15 residue cytoplasmic tail. The PCR fragment was digested with BspM II and Bam HI and ligated into the compatible sites of CNX cDNA subcloned into the Drosophila pRMHa3 expression vector (CNX-pRMHa3).

A soluble form of CNX corresponding to its complete ER luminal domain (residues 1-463; designated Δcyt.Tm) was generated by cleaving CNX-pRMHa3 at the unique Dsa I and Kpn I sites (Kpn I being at the 3' end of the cDNA in the pRMHa3 multiple cloning site). A double stranded oligonucleotide cassette 5'CGTGGGAAGGACGAGCTGTAAGGTACCG3'/5'GATCCGGTACCTTACAGCTGTCCTTC3' containing an ER localization signal (KDEL) and a stop codon flanked by Dsa I and Kpn I restriction sites was inserted into DsaI/KpnI-digested CNX cDNA.

A truncated form of the ER luminal domain corresponding to CNX residues 1-387 (designated Δcyt,Tm,388-462) was generated by inserting an oligocassette 5'CCGGATAAGGACGAGCTGTAAG GTACCG3'/5'GATCCGGTACCTTACAGCTGTCCTTC3' containing the KDEL ER localization signal and a stop codon flanked by BspM II and Kpn I sites into CNX-pRMHa3 cleaved at the unique BspM II and Kpn I sites.

Rabbit CRT cDNA in the pBluescript vector (pB-CR-2) was obtained from Dr. M. Michalak (University of Alberta). For expression in Drosophila cells, the KpnI/Ecl136 II restriction fragment of pB-CR-2, containing full length CRT cDNA, was subcloned into the KpnI and Hinc II sites of the pRMHa3 expression vector.

In the Drosophila expression vector pRMHa3, cDNAs are under the control of the metallothionein promoter (Bunch et al., 1988). Stably transected Drosophila melanogaster Schneider cell lines were established by co-transfecting a plasmid containing the neomycin resistance gene (phshsneo) plus multiple pRMHa3 plasmids encoding CNX or CNX mutants.
Figure 10. CNX and CRT mutants. (A) CNX deletion mutants expressed in *Drosophila* cells. Lightly shaded, hatched and stippled boxes represent the N-terminal signal sequence, the tandemly repeated sequence motifs, and the transmembrane (Tm) segment, respectively, of CNX. The region containing the repeat motifs binds ERp 57 and is also the site of high affinity calcium binding. The Cnx:Acyt mutant has 87 residues of the 89 residue cytoplasmic tail deleted and replaced with the 15 residue tail of the adenovirus E3/19K glycoprotein which contains the ER localization signal -DEKKMP (underlined). Cnx:Acyt,Tm corresponds to the entire ER luminal portion of CNX fused to the ER localization motif -KDEL (underlined). Cnx:Acyt,Tm,388-462 is a truncated form of the ER luminal segment which retains the repeated sequence motifs and hence the ability to bind ERp 57 and calcium (Vassilakos et al., 1998). It is fused to the -KDEL ER localization motif. Restriction sites used in construction of the mutants are indicated. (B) CNX mutants expressed in L cells. The truncation mutants depicted in (A) were modified by the insertion of an influenza hemagglutinin epitope tag, -YPYDVPDYA-, as indicated in bold type. In addition, the Cnx:19KTm,Δcyt(HA) mutant consists of the entire ER luminal portion of CNX plus two Tm residues (W463L464) fused to the Tm segment and cytoplasmic tail (with HA epitope inserted) of the adenovirus E3/19K glycoprotein. The new 22 residue Tm segment (-WLFCSTALLITALVCTLLYL-) is the same length as CNX’s Tm segment (-WLWVYVLTVTLPVFVLVIFCC-). (C) CRT mutants expressed in L cells. CRT containing the HA epitope has been described previously (Wada et al., 1995). Crt:CnxTm/cyt(HA) has also been described previously (Wada et al., 1995) and is a membrane anchored form of CRT fused at residue 340 to the Tm and HA-tagged cytoplasmic segments of CNX (residues 459-573).
along with H-2K\textsuperscript{b} or D\textsuperscript{b} H chains in the presence or absence of mouse β\textsubscript{2}m (Jackson et al., 1994). Following induction with 1 mM CuSO\textsubscript{4} for 16 h, expression of class I molecules was analyzed by metabolic radiolabeling followed by immunoprecipitation with K\textsuperscript{b} or D\textsuperscript{b} specific antibodies. The expression of CNX mutants was detected by Western blotting. To ensure that all three proteins, CNX or CNX mutant, H chain and β\textsubscript{2}m were expressed within one cell, the cells were cloned by the soft agar technique and screened for expression (Hapel et al., 1981). Six clones were selected for each transfected cell line, and only clones expressing comparable levels of these proteins were used in all subsequent experiments.

2.2.3 Construction of HA-tagged CNX and CRT Mutants and Expression in L Cells

For expression in mouse L cells, CNX truncation mutants were modified by adding a hemagglutinin (HA) tag adjacent to the ER localization signal as depicted in Figure 10B. Full length CNX tagged with HA (pCN (HA)) and HA-tagged CRT (pCR (HA)) were generous gifts of Dr. Ikuo Wada (Sapporo Medical University School of Medicine; Wada et al., 1995). The Δcyt mutant of CNX tagged with HA (Δcyt (HA)) was generated by subcloning dog CNX cDNA into the Kpn I and Xba I sites of the pcDNA3 vector (Invitrogen; CNX-pcDNA3). Two PCR primers, one upstream from the unique BspMII restriction site

5'CCCGAAGATACCAATCCGG3' and the other downstream from the Tm segment

5'CCGCGGATCCTTATGGCATTTCATTTTCTGTCAGCATAATCTGGAACATCATATGGATATCCAGACGG3' were used to generate a PCR fragment that encodes the HA tag inserted adjacent to the adenovirus E3/19K ER localization sequence (Fig. 19B). The PCR product digested with BspMII and BamHI was subcloned into the compatible sites of CNX-pcDNA3. The ER luminal domain of CNX tagged with HA (Δcyt, Tm (HA)) was obtained by inserting a double-stranded oligonucleotide cassette

5'CGTTGATATCCAGATGTTCCAGATTATCTAAGGACGAGCTGTAAG3'/
5'GATCCTTACAGCTCCTTAGCATAATCTGGAACATCATATGGATAC3' encoding
the HA tag followed by the KDEL localization signal into DsAI/BamHI digested CNX-pRMHa3. The mutated CNX insert was isolated by digestion with BspMII and BamHI and subcloned into compatible sites of CNX-pcDNA3. The truncated ER luminal domain of CNX tagged with HA (Δcyt, Tm, 388-462 (HA)) was obtained by inserting a double-stranded oligonucleotide cassette

5'CCGGATTATCCATATGATGTTCCAGATTATG CTAAGGACGAGCT GTAAG3' / 5'GATCCTTACAGCTCCTTAGCATAATCTCGGACAT CATATGGATAAT3' encoding the HA tag followed by the KDEL localization signal into BspM II/BamHI digested CNX-pcDNA3. To replace CNX’s transmembrane domain with that of the adenovirus E3/19K glycoprotein, the Δcyt (HA) construct was digested with Kpn I and Bam HI and subcloned into the KpnI and BamHI sites of the pRMHa3 vector. The vector was then digested with DsAI and EcoRV and a double-stranded oligonucleotide cassette

5'CGTGGCTCTTTTGGTTCCACCGCTCTGCTTATTACAGCGCTTTGGTATGTA CCTTACTTTATCTCAAATACAAAT 3'/5'ATTGATATTGAGATA AAGTAAGGTAC ATACCAAAAGCAAGCGCTGTAATAAGCAGAGCGGTGGAACAAAAGAGC encoding the E3/19 transmembrane domain was inserted. The 1588bp KpnI/BamHI fragment was then subcloned into KpnI/BamHI digested pcDNA3 and termed Cn:19KTm, Δcyt (HA).

cDNA encoding CRT residues -17-340 fused to CNX’s transmembrane and cytoplasmic segments (residues 459-573) was obtained from Dr. Ikuo Wada, Sapporo Medical University and designated Crt:CnxTm/cyt(HA) (Wada et al., 1995). In all cases, the recombinant plasmids were introduced into L cells using Superfect (Qiagen). The cells were analyzed two days after transfection.

2.2.4 Metabolic Labeling, Immunoprecipitation, and Gel Electrophoresis

Labeling of Drosophila cells with [35S]Met, lysis, and immunoprecipitation were carried out as described previously (Vassilakos et al., 1996). Briefly, following induction with 1mM CuSO4 for 16 h, transfected Drosophila cells were incubated for 30 min in Met-free Schneider’s medium. Cells were then radiolabeled with [35S]Met for 5 min or as
indicated in figure legends, chased for various times and lysed in a buffer containing 1% digitonin, PBS, pH 7.4, 10mM iodoacetamide, 1% aprotinin and 10μg/ml each of chymostatin, leupeptin, antipain and pepstatin. Lysates were incubated for 2 hours at 4°C with anti-class I or anti-CNX antibodies followed by a 1 hour incubation with protein A-agarose beads. Immune complexes were analyzed by SDS-PAGE using 10% gels (Laemmli, 1970) and radioactive proteins were visualized by fluorography. For quantitation of bands, fluorograms were scanned using an EPSON 100C scanner and analyzed using NIH Image software.

L cells at a density of 5x10^5 cells/60mm dish were radiolabeled for 30 min with 200μCi/ml [35S]Met, lysed for 30 min at 4°C in 1 ml lysis buffer and incubated with anti-HA antibodies for 2 hours. Immune complexes were collected on protein A-agarose and analyzed by SDS-PAGE as described previously (Vassilakos et al., 1996).

For detection of CNX and CNX mutants by immunoblotting, proteins resolved by SDS-PAGE were transferred to Immobilon-P membrane (Millipore) (Suh et al., 1994). The membrane was incubated with rabbit anti-CNX antiserum (α-pp90) at 1:5,000 dilution, followed by a donkey anti-rabbit IgG horseradish peroxidase conjugate at 1:10,000 dilution (Jackson Laboratories). Immune complexes were visualized using an enhanced chemiluminescence system (Amersham Life Science).

### 2.3 Results

#### 2.3.1 CRT Can Substitute for CNX in Murine Class I Biogenesis.

To investigate the function of CRT in class I biogenesis we selected the *Drosophila melanogaster* expression system based on the following characteristics: First, as is reflected by the absence of MHC molecules and their accessory proteins, *Drosophila melanogaster*, like all other invertebrate cells, lack an adaptive immune system. Hence, *Drosophila melanogaster* Schneider cells provide a valuable system for reconstitution of the MHC class I pathway. Secondly, since the capacity for diverse yet regulated
oligosaccharide expression characteristic of vertebrate species is apparent in *Drosophila*
this system is also ideal to investigate the function of mammalian lectins. Finally,
*Drosophila* cells are used routinely to produce functional mammalian proteins and can
easily be transfected with multiple cDNAs.

We previously expressed mouse class I H chains and B2m in *Drosophila*
*melanogaster* Schneider cells and found that *Drosophila* homologs of CNX and CRT
could not be detected in association with murine class I molecules as assessed either by
chemical cross-linking (Jackson et al., 1994) or by coimmunoprecipitation with anti-H
chain mAbs (Vassilakos et al., 1996). Consistent with this observation, these cells were
unable to support the efficient folding or assembly of class I molecules and they lacked the
quality control capacity to retard the export of incompletely assembled class I molecules
from the ER. Since *Drosophila* cells possess genes encoding CNX and CRT and also a
deglucosylation-reglucosylation system (Smith, 1992; Parker et al., 1995; Christodoulou et
al., 1997), it is unclear why *Drosophila* CNX and CRT do not bind detectably to mouse
class I H chains. Nevertheless, these cells provide a useful system to assess the functions
of mammalian CNX and CRT in class I biogenesis. Indeed, co-expression of mammalian
CNX along with class I H chain and B2m in *Drosophila* cells revealed that CNX promotes
efficient H chain folding and assembly with B2m, stabilizes H chain conformation, and
functions as a component of the quality control machinery to retain assembly intermediates
within the ER (Jackson et al., 1994; Vassilakos et al., 1996). It is important to note,
however, that class I molecules do not acquire peptides in *Drosophila* cells and hence their
assembly cannot be studied beyond the formation of H chain-B2m heterodimers (Jackson et

Unlike CNX, which binds rapidly to newly synthesized free H chains and is present
throughout the whole process of murine class I assembly, CRT has only been detected with
the products of some class I alleles and only after assembly of H chains with B2m (Solheim
et al., 1997; Harris et al., 1998). Hence it's role in facilitating H chain folding or assembly
with β2m, if any, remains unclear. In fact, CRT’s ability to function as a molecular chaperone has never been clearly demonstrated nor has its functional relationship with CNX, apart from its identical lectin specificity, been assessed.

To examine the functions of CRT in mouse class I biogenesis and to compare it to those of CNX, Drosophila cells were co-transfected with cDNAs encoding rabbit CRT or dog CNX, mouse K\textsuperscript{b} or D\textsuperscript{b} H chains, and mouse β2m. Initially, the abilities of CRT and CNX to augment the assembly of K\textsuperscript{b} H chains with β2m were examined. Cells expressing K\textsuperscript{b} H chains and β2m in the absence or presence of co-expressed CRT or CNX were subjected to pulse-chase radiolabeling and the levels of newly synthesized K\textsuperscript{b} H chains were detected using three different antibodies: a rabbit antiserum (anti-8) that recognizes total (assembled and unassembled) K\textsuperscript{b} H chains, a β2m dependent mAb (20-8-4S) that only recognizes assembled K\textsuperscript{b}-β2m heterodimers, and a rabbit antiserum (anti-HC) that recognizes unassembled K\textsuperscript{b} H chains.

Figure 11A demonstrates that immunoprecipitation with each antibody was complete with more than 97% of the particular H chain population being recovered in the first round of immunoprecipitation (Fig. 11A). Also, there was no cross-reactivity between the population-specific antibodies since the amount of unassembled H chain that could be recovered with anti-HC Ab after prior clearance with the β2m dependent mAb (20-8-4S) was comparable to that recovered in a direct precipitation with anti-HC and vice versa (Fig. 11A, compare lanes 5 and 6 and lanes 3 and 8). In the absence of co-expressed CRT or CNX less than 50% of total H chains assembled with β2m. Furthermore, the kinetics of H chain association with β2m were slow with less than 25% of H chains associating during the 5 min pulse period (Fig. 11B and 11C). By contrast, and consistent with our previous observations (Vassilakos et al., 1996), co-expression of CNX resulted in more rapid assembly (40% of H chains associated with β2m during the pulse period) and up to 90% of H chains assembled with β2m during the 80 min chase period. Treatment with 2 mM castanospermine abolished this CNX-enhanced assembly but had no effect on the ~40-50%
Figure 11. Effects of CNX and CRT on the assembly of K\textsuperscript{b} H chains with \(\beta_2m\) in *Drosophila* cells. (A) *Drosophila* cells expressing K\textsuperscript{b} H chains, \(\beta_2m\), and CNX were radiolabeled with [\textsuperscript{35}S]Met for 5 min. K\textsuperscript{b} molecules were isolated with antibodies recognizing total (\(\alpha\)-8), \(\beta_2m\)-associated(20-8-4S), or unassembled (\(\alpha\)-HC) H chain in the first round of immunoprecipitation (1\textsuperscript{st} i.p.). The supernatants were then subjected to a second round of immunoprecipitation (2\textsuperscript{nd} i.p.) either with the same antibodies (to assess the efficiency of antibody binding) or with the complementary antibody (to assess the degree of cross-reactivity) and analyzed by reducing SDS-PAGE and fluorography. The fluorogram was slightly overexposed to detect the trace levels of antigen recovered in the 2\textsuperscript{nd} i.p. with the same antibody. (B) *Drosophila* cells expressing K\textsuperscript{b} H chains and \(\beta_2m\) in the absence or presence of CNX or CRT were radiolabeled with [\textsuperscript{35}S]Met for 5 min and then chased in the presence of excess unlabeled Met for the indicated times. K\textsuperscript{b} molecules were isolated with antibodies recognizing total, \(\beta_2m\)-associated, or unassembled H chains and analyzed by reducing SDS-PAGE. The mobilities of mature and immature H chains are indicated. (C) Results from five independent experiments for cells expressing K\textsuperscript{b} and \(\beta_2m\) alone or in the presence of CNX and from two independent experiments for cells co-expressing CRT were quantified by densitometry. The averaged amounts of unassembled and \(\beta_2m\)-associated H chains at each time point were expressed as a percentage of the total H chains present in the pulse sample.
assembly observed in cells lacking exogenous CNX (data not shown). This suggested that the assembly observed in cells expressing only H chain and β2m occurred without the involvement of Drosophila CNX or CRT. CRT was just as effective as CNX in enhancing the kinetics and efficiency of heterodimer assembly. During the pulse, 55% of H chains assembled and assembly was nearly quantitative after 80 min of chase (Fig. 11B and 11C).

As reported previously, CNX plays an important role in ER quality control (Jackson et al., 1994; Rajagopalan et al., 1994; Rajagopalan and Brenner, 1994). Co-expression of CNX with murine class I molecules in Drosophila cells dramatically slowed export from the ER of the free H chain and peptide-deficient H chain-β2m assembly intermediates (Jackson et al., 1994). To determine whether CRT can also retain assembly intermediates in the ER, we assessed the rates at which newly synthesized H chain-β2m heterodimers are converted to mature forms that possess Golgi-processed N-linked oligosaccharides. In Drosophila cells, mature heavy chains that have been transported through the Golgi apparatus migrate more rapidly than their immature precursors due to processing of their N-linked oligosaccharides to smaller, complex forms. As shown in Fig. 11B (β2m-associated panels), Kβ heterodimers were converted to the smaller, mature form with a t1/2 of ~20 min in cells lacking mammalian CNX or CRT. This is indicative of poor quality control for class I molecules in Drosophila cells since the same peptide-deficient heterodimers are exported very slowly in mouse cells with a t1/2 of 100 min (Degen et al., 1992). Consistent with previous findings (Jackson et al., 1994), co-expression of CNX resulted in a dramatic slowing of intracellular transport with export of heterodimers to the Golgi occurring with a t1/2 well in excess of 80 min, Fig. 11B. CRT proved to be just as effective as CNX in retarding the transport of Kβ heterodimers to the Golgi since mature heterodimers were formed at a comparably slow rate.

CNX has also been shown to increase the yield of folded class I H chains (Vassilakos et al., 1996). However, since CRT has only been detected in association with H chain-β2m heterodimers, its capacity to interact with free H chains and influence folding are
unknown. To address this issue, we examined H chain folding in control and CRT- or CNX-transfected cell lines by monitoring the formation of a conformational epitope in the α3 domain of the D^b H chain defined by mAb 28-14-8S. Note that in these experiments D^b folding was measured in the absence of β2m using Drosophila transfectants expressing only free H chains. As shown in Fig. 12A, essentially all D^b H chains (total or 28-14-8s reactive) could be recovered in a single round of immunosolation. Fig. 12B depicts a pulse-chase radiolabeling experiment followed by immunosolation of total or 28-14-8S reactive H chains. Similar to results obtained previously (Vassilakos et al., 1996), CNX enhanced H chain folding by ~2-fold; all D^b H chains folded into a 28-14-8S-reactive conformation in the presence of CNX, whereas only 50-60% of H chains acquired the epitope in its absence. CRT also enhanced folding of H chains although the effect was slower and somewhat less efficient than observed in the presence of CNX. Densitometric analysis revealed that 80% of H chains acquired the 28-14-8S epitope after a 5 min pulse in the presence of CNX but this level was reached only after a 20 min chase in the presence of CRT (Fig. 12C).

In addition to promoting H chain folding, CNX has been shown to stabilize free H chains against unfolding and/or degradation (Moore and Spiro, 1993; Jackson et al., 1994). Consequently, as a final assessment of the functional relationship between CNX and CRT, we compared the abilities of CNX and CRT to stabilize folded free D^b H chains. Cells expressing D^b H chains in the presence and absence of CNX or CRT were subjected to pulse-chase radiolabeling followed by immunosolation of 28-14-8S reactive H chains (Fig. 13). In control cells expressing only D^b, the half-life of 28-14-8S reactive H chains was 80 min. By contrast, co-expression of CNX increased the half-life of the partially folded D^b H chains 4-fold (t_{1/2}=320 min). Surprisingly, CRT, which has never been detected in association with free H chains, was just as effective as CNX in stabilizing folded free D^b H chains (t_{1/2}=320 min). These effects of CNX and CRT occurred primarily through stabilization of the 28-14-8S reactive conformation rather than through
Figure 12. Folding of Db heavy chains in the presence of CNX and CRT. (A) Drosophila expressing Db H chain and CRT were radiolabeled with [35S]Met for 5 min. H chains with a folded α3 domain were isolated with mAB 28-14-8S and total Db H chains were isolated with a combination of mAB 28-14-8S plus HC serum(1st). Supernatants from the first round of immunoprecipitation were then subjected to a second immunoprecipitation with the same antibodies (2nd). Immune complexes were analyzed by SDS-PAGE. (B) Drosophila cells expressing Db H chain alone or in the presence of CNX, CRT, or a truncated form of the CNX luminal domain (Δcyt, Tm, 388-462) were subjected to pulse-chase radiolabeling with [35S]Met. Total Db H chains and H chains with a folded α3 domain were isolated with the antibodies described in panel (A) and immune complexes were analyzed by SDS-PAGE. (C) The results from three independent experiments were quantified by densitometry and the averaged amounts of 28-14-8S reactive H chains at each time point were expressed as a percentage of the total heavy chain present in the pulse sample.
Figure 13. CRT and CNX stabilize partially folded class I H chains. Drosophila cells expressing $D^b$ H chain in the absence or presence of CNX or CRT were incubated for 5 min with $[^{35}S]$Met and then chased for the times indicated. (A) $D^b$ H chains were immunoisolated with mAb 28-14-8S and analyzed by reducing SDS-PAGE. (B) Fluorograms from two independent experiments were quantified by densitometry and the averaged amount of H chain recovered at each time point were expressed as a percentage of the amount present in the pulse sample.
prevention of H chain degradation since immunoisolation of H chains with a conformation-insensitive Ab revealed minimal differences in degradation rates during a 160 min chase period (data not shown).

Overall these findings indicate that CRT can largely replace CNX in enhancing H chain folding and assembly with β<sub>2</sub>m, in retaining H-chain-β<sub>2</sub>m heterodimers in the ER, and in stabilizing partially folded H chains.

### 2.3.2 Association of Truncated Forms of CNX with H chain Substrates

Since CRT, a soluble analog of CNX, appears to function in a manner that is largely interchangeable with CNX, the question arises as to whether CNX can also function as a soluble molecule or if its cytoplasmic and transmembrane segments are essential to its overall functions as a molecular chaperone and component of the ER quality control machinery. To address this issue, the truncated forms of CNX depicted in Figure 10 were constructed. To study the functions of CNX's cytoplasmic tail, it was replaced with the much shorter cytoplasmic tail of the adenovirus E3/19K glycoprotein (Δcyt). In addition, the combined functions of CNX's transmembrane and cytoplasmic segments were examined by generating a soluble form of CNX corresponding to its entire ER luminal segment (Δcyt,Tm). To study the functions of residues 388-462, a region that has been proposed to interact with polypeptides (Wada et al., 1995), the ER luminal segment was further truncated to produce a 387 residue fragment (Δcyt, Tm, 388-462) that retains about half the lectin function of the entire ER luminal segment (Vassilakos et al., 1998). For each mutant, an ER localization signal was added at the C-terminus (Fig. 10A).

Immunoblots of lysates from stably transfected *Drosophila melanogaster* cells revealed that the various CNX deletion mutants were expressed in amounts comparable to that observed for full length CNX. Furthermore, the mutants were detected in cell lysates but not in the culture medium, indicative of their intracellular retention (Fig. 14).

To examine how individual segments of CNX affect its association with class I molecules, transfectants expressing CNX deletion mutants, D<sup>b</sup> H chain, and β<sub>2</sub>m were radiolabeled with
Figure 14. Expression of CNX and its truncated forms in *Drosophila* cells. Aliquots (5x10^5 cell equivalents/sample) of cell lysates and culture media from *Drosophila* cells expressing full length CNX or various CNX deletion mutants were subjected to SDS-PAGE analysis and then proteins were transferred to Immobilon-P membrane. The membrane was probed with rabbit anti-CNX antibody (anti-pp90) and bands were visualized using a chemiluminescence-based detection system.
[35S] Met. and then class I-CNX complexes were immunoisolated with class I specific mAb. As shown in Fig. 15, full length CNX and CNX lacking either its cytoplasmic tail (Δcvt) or both its cytoplasmic tail and transmembrane segment (Δcvt, Tm) formed stable complexes with Db H chains. However, a reduction in the intensity of the soluble Δcvt, Tm mutant relative to either full length CNX or the Δcvt mutant suggests that the strength of the interaction decreases upon loss of the transmembrane segment. Furthermore, removal of residues 388-462 from the ER luminal segment of CNX resulted in a very weak association with class I molecules that was detected only when the immune complexes were washed under mild conditions (data not shown). This finding suggests that residues 388-462 may also contribute, either directly or indirectly, to the interaction of CNX with class I H chains.

2.3.3 Quality Control Function of Truncated Forms of CNX

To test whether transmembrane and cytoplasmic segments contribute to the quality control function of CNX, we examined the ER to Golgi transport rates of peptide-deficient Kb-β2m heterodimers in the presence of full length CNX or CNX truncation mutants. It is well established that the rate at which a class I molecule acquires resistance to digestion with endoglycosidase H (endo H) provides a measure of its rate of transport from the ER to the medial Golgi cisterna. As shown in Fig. 16, the transport rates of Kb heterodimers were significantly slowed from a half-time of ~20 min in CNX-deficient cells to well over 80 min in cells expressing CNX with a truncated cytoplasmic tail (Δcvt) or the soluble ER luminal segment of CNX (Δcvt, Tm). Interestingly, these truncated forms of CNX slowed the transport of peptide-deficient heterodimers more than full length CNX (t1/2 ~80 min). CNX's quality control function was slightly impaired by further truncation of the ER luminal segment (Δcvt, Tm, 388-462) with Kb heterodimers being transported out of the ER with a t1/2 of ~60 min. Comparable trends were observed when the CNX mutants were co-expressed with peptide-deficient Db heterodimers (data not shown).

Collectively, these data indicate that the cytoplasmic tail and the transmembrane segment are
Figure 15. Association of truncated forms of CNX with class I H chains. *Drosophila* cells expressing *D*\(^b\) H chain and *β*\(_2\)m in the presence of full length or truncated forms of CNX were radiolabeled with \[^{35}S\]Met for 10 min. *D*\(^b\)-*β*\(_2\)m heterodimers were isolated with mAb B22-249RI. The mobilities of co-isolated full length CNX, CNX with truncated cytoplasmic tail (Δcyt), and the soluble ER luminal segment of CNX (Δcyt,Tm) are indicated by asterisks.
Figure 16. Intracellular transport of $K^b\beta_2m$ heterodimers in the presence of truncated forms of CNX. *Drosophila* cells expressing peptide-deficient $K^b\beta_2m$ heterodimers in the absence or presence of CNX and various CNX deletion mutants were incubated with [$^{35}$S]Met for 5 min and then in the presence of excess unlabeled Met for the indicated times. $K^b\beta_2m$ heterodimers were immunoisolated with mAb 20-8-4S, digested with endo H, and analyzed by reducing SDS-PAGE. The mobilities of the endo H-resistant (r) and endo H-sensitive (s) forms of the $K^b$ H chain are indicated.
not essential for the quality control function of CNX. However, the presence of the segment corresponding to CNX residues 388-462 slightly enhances the efficiency of quality control.

2.3.4 Molecular Chaperone Functions of Truncated Forms of CNX

To assess the contribution of CNX’s transmembrane and cytoplasmic segments to its molecular chaperone functions, we compared the ability of full length or truncated forms of CNX to enhance H chain assembly with B2m, to promote H chain folding, and to stabilize partially folded H chains. H chain-B2m assembly was analyzed as described in Fig. 1 using specific antibodies to detect either total or B2m-associated K\(^b\) H chains. The results obtained for control cells lacking CNX and for cells expressing either full length CNX or the mutant with the largest truncation (Δcyt, Tm, 388-462) are depicted in Fig. 17. The mutant, which consists only of residues 1-387 of CNX’s ER luminal segment, was essentially as effective as full length CNX in enhancing the kinetics and efficiency of K\(^b\)-B2m assembly.

The folding of free D\(^b\) H chains in the presence of CNX or CNX truncation mutants was assessed by monitoring the formation of the conformational epitope defined by mAb 28-14-8S and comparing it to the total amount of D\(^b\) H chains. Similar to the results described for H chain assembly with B2m, the truncated ER luminal segment of CNX (Δcyt, Tm, 388-462) resembled full length CNX in its ability to enhance H chain folding at an early stage (Fig. 12B). This was a very rapid process happening largely within the 5 min pulse labeling period. Thus, CNX’s cytoplasmic tail, its transmembrane segment and residues 388 to 462 are not required for its functions in enhancing H chain folding or assembly with B2m.

We also examined the ability of CNX mutants possessing less extensive truncations to promote H chain folding and assembly with B2m. Surprisingly, CNX with a truncated cytoplasmic tail (Δcyt) and the soluble ER luminal segment of CNX (Δcyt, Tm) exhibited little or no enhancement of folding or assembly relative to the CNX-deficient control (data
Figure 17. Effects of truncated forms of CNX on $K^b_2m$ assembly with $\beta_2m$. (A) Drosophila cells expressing $K^b$ H chains and $\beta_2m$ in the absence or presence of full length or truncated forms of CNX were incubated with $[^{35}S]$Met for 5 min and then with excess unlabeled Met for the times indicated. $K^b$ molecules were isolated with antibodies recognizing total (anti-8 antiserum) or $\beta_2m$-associated (mAb 20-8-4S) H chains. Following digestion with endo H, proteins were analyzed by reducing SDS-PAGE. The mobilities of the endo H-resistant (r) and endo H-sensitive (s) H chains are indicated. (B) Results from three independent experiments were quantified by densitometry and the averaged amounts of $\beta_2m$-associated H chains were expressed as a percentage of the total H chains present in the pulse sample.
not shown). Clearly, this is not due to some essential role for the cytoplasmic or transmembrane segment in CNX's chaperone function since the Δcyt, Tm, 388-462 mutant also lacks these segments and promotes H chain folding and assembly with β2m almost as well as full length CNX. The inability of the Δcyt and Δcyt, Tm mutants to enhance folding/assembly may be related to the fact that both mutants retain class I molecules in the ER more efficiently than full length CNX (Fig. 16) and hence may be impaired in their abilities to release class I H chains to undergo productive folding reactions.

To assess the role of CNX’s cytoplasmic and transmembrane segments in stabilizing partially folded H chains, cells expressing free Dβ H chains in the absence or presence of CNX or CNX truncation mutants were subjected to pulse-chase radiolabeling followed by immunoisolation of 28-14-8S reactive H chains (Fig. 18). In the absence of CNX, partially folded Dβ H chains disappeared with a t1/2 of 80 min. Upon coexpression of full length CNX, loss of the folded epitope was markedly slowed such that only 20% was lost after 160 min and 40% after 320 min. All CNX mutants provided extensive protection with only 10-20% of H chains being lost after 160 min. These results suggest that the cytoplasmic tail, the transmembrane segment, and residues 388-462 of CNX’s ER luminal segment are not essential for stabilizing partially folded H chains.

2.3.5 Substrate Specificities of CNX, CNX Truncation Mutants, and CRT

The preceding results have shown that CNX’s cytoplasmic tail, it’s transmembrane region, and residues 388 to 462 of its ER luminal segment are not essential for its molecular chaperone or quality control functions. Consequently, we asked whether these segments might influence the spectrum of proteins with which CNX interacts. To approach this question, various CNX truncation mutants were expressed transiently in mouse L cells and compared to similarly expressed CNX and CRT. To distinguish the transfected protein products from endogenous CNX and CRT, a hemagglutinin (HA) epitope tag was inserted
Figure 18. Effect of truncated forms of CNX on the stabilization of class I H chains.

*Drosophila* cells expressing free Db H chains alone or in the presence of CNX or various CNX truncation mutants were radiolabeled for 5 min with [35S]Met and then chased for the times indicated. Db H chains were immunoisolated with mAb 28-14-8S and analyzed by reducing SDS-PAGE. Fluorograms were quantified by densitometry and the amounts of H chain recovered at each time point were expressed as a percentage of the amount present in the pulse sample.
near the carboxy terminus of each construct (Fig. 10B,C). The L cell transfectants were radiolabeled with [35S]Met and then cell lysates were incubated with anti-HA mAb. The patterns of proteins co-immunisolated with HA-tagged CNX, truncated CNX mutants, and CRT are shown in Fig. 19. Numerous radiolabeled proteins were found to form complexes with CNX and CRT, but the patterns of these proteins differed between the two chaperones (Fig. 19; compare Cnx(HA) and Crt(HA), lanes 2 and 5). The majority of co-isolated proteins were N-glycosylated since treatment of cells with castanospermine prior to radiolabeling resulted in a marked reduction or complete loss of proteins associating with either CNX or CRT (data not shown).

Replacement of CNX’s cytoplasmic tail with the HA tag and the adenovirus E3/19K ER localization signal had little effect on the spectrum of proteins co-isolated with CNX (Fig. 19; compare Cnx(HA) and Cnx:Acyt(HA), lanes 2 and 6). However, upon removal of both the transmembrane and cytoplasmic segments of CNX, a substantial change in the pattern of co-isolated proteins was observed such that pattern closely resembled that observed for CRT (Fig. 19; compare Cnx(HA), Cnx:Acyt,Tm(HA), and Crt(HA), lanes 2, 3, and 5). Further truncation of residues 388-462 from the ER luminal portion of CNX almost completely abolished stable interactions between CNX and the diverse proteins with which it interacts (Fig. 19; Cnx:Acyt,Tm,388-462(HA), lane 4). This was unlikely to be due to misfolding of the mutant protein because a non-HA tagged version of this truncated form of CNX also lacked stable interactions with class I H chains (Fig. 15) yet was fully capable of functioning as a molecular chaperone for class I molecules (Figs. 12, 17, and 18). This construct also served to demonstrate that the various proteins co-isolated with CNX and CRT were specifically associated since these proteins could not be detected when the Cnx:Acyt,Tm,388-462(HA) mutant was immunoisolated under identical conditions.

The finding that the soluble ER luminal segment of CNX (Cnx:Acyt,Tm(HA)) associated with a similar spectrum of proteins as CRT suggested that CNX’s Tm segment somehow influences substrate specificity. To determine if this is due to some unique
Figure 19. Effect of transmembrane and cytoplasmic segments on the substrate specificities of CNX and CRT. HA-tagged CNX (Cnx(HA), lane 2), the soluble ER luminal portion of CNX (Cnx:Δcyt,Tm(HA), lane 3), a truncated ER luminal segment of CNX (Cnx:Δcyt,Tm,388-462(HA), lane 4), CRT (Crt(HA), lane 5), CNX with a truncated cytoplasmic tail (Cnx:Δcyt(HA), lane 6), CNX with a truncated cytoplasmic tail but containing the Tm segment of the adenovirus E3/19K glycoprotein (Cnx:19KTm, Δcyt(HA), lane 7), or membrane bound CRT possessing CNX's transmembrane and cytoplasmic segments (Crt:CnxTm/cyt(HA), lane 8) were transiently expressed in L cells. The cells were radiolabeled for 45 min with [35S]Met and then lysed with buffer containing 1% digitonin. The HA-tagged molecules and associated proteins were immunoprecipitated with anti-HA mAb 12CA5. Dots indicate the mobilities of CNX and CRT constructs. Lane 1 represents an anti-HA immunoprecipitate of lysate from untransfected L cells.
property of CNX's Tm segment, such as a site of protein interaction, or if it is simply a consequence of anchoring CNX within the ER membrane, we replaced CNX's Tm segment with the Tm segment of the adenovirus E3/19K glycoprotein (Cnx:19KTm,Δcyt (HA); see Fig. 19B). As shown in Fig. 19, the spectrum of proteins associating with CNX anchored by the E3/19K Tm segment (lane 7) closely resembled that observed for CNX anchored by its own Tm segment (Cnx(HA) and Cnx:Δcyt(HA), lanes 2 and 6). This finding suggests that the primary basis for the difference in proteins associating with CNX and CRT is their different topological environments rather than some specific property of CNX's Tm segment.

To confirm this finding, we examined a membrane-anchored form of CRT to determine if its pattern of associated proteins would be similar to that of CNX. This construct consisted of CRT residues 1-340 fused to the Tm and cytoplasmic segments of CNX (designated Crt:CnxTrm/cyt(HA)). As shown in Fig. 19 (lane 8), this chimera lacked the distinctive pattern of associated proteins observed with soluble CRT (Fig. 19, lane 5) but rather it closely resembled the pattern observed with CNX or the Cnx:Δcyt(HA) mutant (Fig. 19, lanes 2 and 6).

2.4 Discussion

To date, the concept that CRT functions to facilitate the folding of newly synthesized glycoproteins and to retain incompletely folded or misfolded glycoproteins within the ER has been based on indirect and correlative evidence. For example, CRT has been shown to bind transiently to folding intermediates but not to native forms of glycoproteins in a variety of in vivo studies (Nauseef et al., 1995; Peterson et al., 1996; Sadasivan et al., 1996; Wada et al., 1997; Harris et al., 1998). CRT can also discriminate between denatured and native states of both non-glycosylated and glycosylated proteins in vitro (Svaerke and Houne, 1998). Furthermore, its primary sequence similarity and identical lectin specificity with CNX, which clearly does participate in glycoprotein folding and quality control, has
suggested similar functions for CRT. In the present study we show for the first time that CRT is indeed capable of enhancing glycoprotein folding in vivo and that it can retain incompletely folded/assembled glycoproteins within the ER. By heterologous expression of mouse class I subunits in Drosophila melanogaster cells in the absence or presence of co-expressed CRT, we demonstrate that CRT enhances the folding of class I heavy chains as well as their subsequent assembly with β2m. CRT also stabilizes free H chains and impedes the export of incompletely assembled H chain-β2m heterodimers from the ER.

The finding that CRT stabilizes free H chains and promotes free H chain folding is surprising given that CRT has not been detected in association with either mouse or human H chains prior to assembly with β2m (Sadasivan et al., 1996; van Leeuwen and Kearse, 1996a). Only CNX has been shown to associate stably with free H chains in mouse or human cells (Degen et al., 1992). Remarkably, we have also been unable to detect a CRT-free H chain complex in Drosophila cells by co-immunoprecipitation although it presumably exists based on the functional data. This raises the question of whether CRT plays a more significant role in the earliest stages of class I biogenesis in mouse and human cells than was previously thought; its involvement being unappreciated given its weak association with free H chains. Alternatively, the participation of CRT in free H chain folding and stabilization that we observe in Drosophila cells may be a consequence of providing H chains with only a single chaperone. In mouse or human cells, where both CNX and CRT are present, CNX may be utilized preferentially due to its membrane disposition or perhaps its proximity to the sec 61 translocon that translocates nascent H chains into the ER (Chevet et al., 1999). The ability of CRT to substitute for CNX under conditions of CNX depletion provides a likely explanation for why a human CNX-deficient cell exhibits minimal alterations in the assembly or intracellular transport of class I molecules (Sadasivan et al., 1995; Scott and Dawson, 1995).

Overall, our experiments indicate that CRT's chaperone and quality control functions are largely interchangeable with those of CNX. The two molecules are virtually
indistinguishable in their abilities to promote the assembly of K\textsuperscript{b} H chains with β\textsubscript{2}m, to retard the export of peptide-deficient H chain-β\textsubscript{2}m heterodimers from the ER, and to prevent the rapid degradation of free H chains. Only in promoting free H chain folding does CNX function more efficiently than CRT. Given this interchangeability of soluble CRT with membrane-bound CNX, we questioned whether CNX’s cytoplasmic and transmembrane segments contribute to either its chaperone or quality control functions. Our results indicate that neither segment is required since a truncated form of CNX’s ER luminal domain consisting of residues 1-387 functions essentially as well as full length CNX. This finding is consistent with the observation that expression of CNX’s ER luminal domain complements the lethal phenotype accompanying the disruption of the CNX gene in *Schizosaccharomyces pombe* (Parlati et al., 1995). However, as assessed by co-immunoprecipitation experiments, we noted that the stability of CNX-H chain interactions is significantly reduced upon the removal of CNX’s transmembrane segment, an observation also made by Ho et al. (1999). Complementary results have been obtained in studies that have attempted to map sites of CNX interaction on the class I H chain. Both Margolese et al. (1993) and Carreno et al. (1995) have shown that CNX associates with the H chain at least in part through interactions in the vicinity of its Tm segment. Similarly, CNX was found to associate less efficiently with a soluble form of influenza hemagglutinin when compared with the normal, membrane-bound form (Hebert et al., 1997). We also observed a dramatic reduction in stability of CNX-H chain interactions upon truncation of its ER luminal domain from 462 to 387 residues. This truncated ER luminal construct failed to form stable complexes with a variety of other glycoproteins expressed in L cells as well. This may be due to the loss of a short stretch of hydrophobic amino acids (Phe\textsuperscript{400}-Val\textsuperscript{421}) that has previously been suggested to play a role in the binding of CNX to polypeptide segments of non-native glycoproteins (Wada et al., 1995). However, we cannot exclude the possibility that the reduced binding of this truncated form of CNX to diverse glycoproteins is due to weaker lectin-oligosaccharide interactions since this mutant
retains approximately half of the lectin activity observed with the intact ER luminal domain (Vassilakos et al., 1998). Although our findings suggest that the avidity of CNX’s interactions with diverse glycoproteins may be enhanced by polypeptide-mediated interactions through its transmembrane segment and possibly through ER luminal residues 388-462, it is clear that such interactions are not essential (at least for class I biogenesis) since residues 1-387 of CNX retain both chaperone and quality control functions.

Although CRT and CNX possess identical lectin binding specificities (Vassilakos et al., 1998) and our current findings suggest that they are essentially redundant in their chaperone and quality control functions, it is clear that they divide the labour of chaperoning the synthesis of newly synthesized glycoproteins. Numerous studies have documented overlapping but distinct substrate binding specificities for the two chaperones (Helenius et al., 1997). This raises the important question of what determines their selectivities for different glycoproteins. Most attempts to address this issue have focused on the structural characteristics of the glycoprotein substrates, particularly the number and location of N-linked oligosaccharide chains. Extensive mutagenesis studies of the 7 glycosylation sites in influenza hemagglutinin revealed that CRT binds preferentially to N-glycans located at the top (membrane distal) domain of the molecule whereas CNX binds equally well to the top and membrane-proximal stem domains (Hebert et al., 1997). Consistent with this observation, Harris et al. (1998) showed that of the three glycosylation sites in the mouse Ld H chain, removal of the site in the membrane-distal α1 domain (residue 86) ablates CRT binding whereas CNX binding is unaffected. These findings suggest that the ER luminal versus membrane-bound locations of CRT and CNX may influence their glycoprotein binding preferences. In a separate study, Zhang and Salter (1998) investigated the basis for the apparently exclusive binding of CRT to human class I H chain-β2m heterodimers in contrast to the predominant binding of CNX to mouse heterodimers. Since human H chains possess a single glycan at residue 86 and mouse H chains possess a second glycan at residue 176, they engineered a second glycosylation site
at residue 176 of the human H chain and found that CNX binding to heterodimers was increased whereas CRT binding decreased. It is difficult to rationalize these results in terms of chaperone topology since residues 86 and 176 are located roughly the same distance from the membrane. Alterations in polypeptide determinants also appear to influence CNX and CRT binding. Point mutations at residue 134 in the human class I HLA-A2.1 molecule (Lewis and Elliott, 1998) and at residue 227 in the mouse Ld molecule (Harris et al., 1998) were accompanied by a loss of careticulin binding but there was no effect on CNX interactions. The glycosylation state of these molecules was not altered by the mutations nor were they misfolded as evidenced by their capacity to bind both β2m and several conformation-sensitive mAbs. These findings are most readily explained by a polypeptide component to CNX and CRT binding in addition to the lectin-oligosaccharide interaction. There is considerable evidence to support such a polypeptide binding component (Arunachalam and Cresswell, 1995; Ware et al., 1995; Zhang et al., 1995; Harris et al., 1998; Jannitipour et al., 1998; Svaerke and Houen, 1998; Basu and Srivastava, 1999) and these studies have recently been bolstered by our finding that both CNX and CRT are able to form discrete complexes with non-glycosylated proteins in vitro and prevent their thermally induced aggregation in a manner analogous to other chaperone families (Ihara et al., 1999). A strong candidate for a polypeptide binding site in CNX is the stretch of hydrophobic residues at residues Phe400-Val421 that upon deletion results in a dramatic loss in stable binding to diverse glycoproteins. A similar hydrophobic patch is also present in CRT at residues Phe294-Ile315 (Wada et al., 1995).

In the present study, we approached the issue of glycoprotein binding specificity from the chaperone side of the interaction. We found that although CNX’s transmembrane segment is not required for its chaperone or quality control functions it clearly influences substrate specificity. When CNX was expressed as a soluble molecule corresponding to its entire ER luminal domain, the spectrum of associated glycoproteins shifted to a pattern similar to that observed for CRT. CNX’s transmembrane segment was not unique in
conferring altered binding specificity since its replacement with the transmembrane segment from the adenovirus E3/19 K glycoprotein resulted in binding to a similar set of glycoproteins. Therefore, a luminal versus membrane topology appears to be a major factor influencing the differential binding specificities of CRT versus CNX. This suggestion was further supported by anchoring CRT to the ER membrane via CNX’s transmembrane segment and observing an accompanying shift in the pattern of associated glycoproteins to one similar to that of CNX. The latter experiment has independently been performed by Wada and co-workers with comparable results (Wada et al., 1995).

It is not difficult to imagine that the different topological orientations of CNX and CRT place constraints on the glycoproteins they bind. Presumably CRT binds preferentially to glycan and polypeptide segments that are most luminaly exposed and less readily to those that are less accessible. This is consistent with the observed preferential binding of CRT to glycans on the membrane distal domains of influenza hemagglutinin (Hebert et al., 1997) and the class I H chain (Harris et al., 1998) versus those that are located closer to the membrane. CNX, on the other hand, has its binding sites anchored to the membrane and at least a portion of CNX molecules appears to be further constrained to the environment of the translocon (Chevet et al., 1999). CNX’s lectin site is indeed quite close to the ER membrane since it has been shown to bind to an N-linked glycan located 12-13 residues from the membrane (Andersson et al., 1996). This disposition must necessarily place constraints on the sites that CNX can bind on a newly synthesized glycoprotein. Given these constraints, it would suggest that CRT and CNX divide the labor of chaperoning different glycoproteins or different stages in the maturation of a single glycoprotein in a controlled manner. However, in a CNX or CRT deficient situation, they may substitute for one another if there are productive binding sites available on a particular glycoprotein, such as we have documented in the case of the folding and stabilization of free class I H chains.
CHAPTER III

The lectin-chaperone calnexin utilizes polypeptide-based interactions to associate with many of its substrates in vivo\textsuperscript{1}

---

Please Note:

In Fig. 23 Myrna F. Cohen-Doyle assisted in preparation of the gradients.

---

\textsuperscript{1}The results presented in this chapter have been published in Ursula G. Danilczyk and David B. Williams (2001) \textit{The Journal of Biological Chemistry} vol 276, pp. 25532-25540. Reprinted with the permission from Ursula G. Danilczyk, David B. Williams and The American Society for Biochemistry and Molecular Biology, Inc.
3.1 Introduction

Glycoprotein folding within the ER is facilitated in part by the membrane-bound chaperone CNX and its soluble homolog CRT (reviewed in Parodi, 2000). These proteins are unique among molecular chaperones in that they utilize a lectin site as a means to associate with unfolded glycoproteins (Hammond et al., 1994; Spiro et al., 1996; Ware et al., 1995). It is a widely held view that CNX and CRT associate with glycoproteins solely through lectin-oligosaccharide interactions. The concept is based primarily on experiments wherein cultured cells were treated with tunicamycin to block Asn-linked oligosaccharide addition or with the glucosidase I and II inhibitors castanospermine or deoxynojirimycin to prevent the conversion of the Glc₃Man₉GlcNAc₂ precursor to the monoglucosylated Glc₁Man₉GlcNAc₂ species. Subsequent immunoprecipitation with anti-CNX or anti-CRT antibodies frequently revealed a dramatic reduction in the amounts of various glycoproteins co-isolating as complexes with these chaperones (Hammond et al., 1994; Nauseef et al., 1995; Otteken and Moss, 1996; Ou et al., 1993; Peterson et al., 1995; Toyofuku et al., 1999; Vassilakos et al., 1996). Similar results were obtained with mutant cell lines that lack the glucosidases involved in producing the monoglucosylated oligosaccharide (Balow et al., 1995; Ora and Helenius, 1995).

In contrast with the preceding results, many lines of evidence have suggested that CNX and CRT can also associate with non-native proteins via protein-protein interactions as described in detail in chapter 1. Briefly, several studies have shown that complexes between CNX and either membrane-bound or soluble glycoproteins cannot be dissociated by enzymatic removal of oligosaccharides. There are several examples of CNX interacting with proteins that either lack Asn-linked oligosaccharides naturally or have lost them through mutagenesis or underglycosylation (Rajagopalan et al., 1994; Carreno et al., 1995; Kim and Arvan, 1995; Loo and Clarke, 1995). Furthermore, following treatment of cells with castanospermine to block the formation of monoglucosylated oligosaccharides, CNX or CRT have been detected in association with
specific glycoproteins. Both CNX and CRT have been shown to bind specifically to unglycosylated peptides both in vitro and in vivo (Basu and Srivastava, 1999; Jorgensen et al., 2000; Nair et al., 1999; Spee et al., 1999). Finally, using purified components in vitro, it has been demonstrated that CNX and CRT are capable of functioning as molecular chaperones to suppress the aggregation and enhance the folding not only of glycoproteins bearing monoglucosylated oligosaccharides but of non-glycosylated proteins as well (Ihara et al., 1999; Saito et al., 1999).

Despite this accumulated information, the concept that CNX and CRT are capable of associating in vivo with unfolded proteins via polypeptide-based interactions in addition to lectin-oligosaccharide binding has been largely discounted. It has been speculated that the lack of dissociation of CNX-substrate complexes following complete deglycosylation may be due to the trapping of the two species within the same detergent micelle (Parodi, 2000; Rodan et al., 1996) or that the substrate, being non-native, might become insoluble upon dissociation (Parodi, 2000; Rodan et al., 1996; Zapun et al., 1997). Similarly, it has been suggested that the association of CNX with non-glycosylated proteins in vivo may arise through non-specific inclusion of CNX within misfolded protein aggregates (Cannon et al., 1996; Parodi, 2000; Zapun et al., 1997). However, apart from a single instance in which CNX was detected in association with aggregates of unglycosylated VSV G protein (Cannon et al., 1996), there has been no direct evidence to support such speculations. Finally, the in vitro studies demonstrating direct binding of non-glycosylated peptides to CRT or the molecular chaperone functions of CNX and CRT with non-glycoproteins have been questioned in terms of their relevance to the in vivo situation (Parodi, 2000).

In an effort to address the question of the existence of polypeptide-based interactions between CNX and its diverse substrates in vivo, we chose to utilize the same methodology used most commonly in previous studies to demonstrate the apparent exclusivity of lectin-oligosaccharide interactions, i.e., block the formation of
monoglucosylated oligosaccharides and assess by co-immunoisolation if complexes between diverse substrates and CNX can be detected. We reasoned that upon loss of the lectin-oligosaccharide interaction any remaining polypeptide-based association might be too weak to survive rigorous immunoisolation conditions. Consequently, care was taken to employ mild, yet highly specific isolation procedures. Using either pharmacologic or genetic methods to block the formation of the Glc$_1$Man$_9$GlcNAc$_2$ oligosaccharide in diverse cell types, we show that although many complexes were lost, a large number of CNX-substrate complexes remained readily detectable. Complementary results were also obtained using a substrate that lacked oligosaccharides through mutation of its Asn-X-Ser(Thr) sequons. Interactions with CNX (and CRT) were maintained in the absence of any detectable aggregation. We conclude that in addition to the well established lectin-oligosaccharide interaction, polypeptide-based association does indeed exist \textit{in vivo} between CNX or CRT and a diverse array of protein substrates.

3.2 Material and Methods

3.2.1 Cell Lines and Antibodies

Murine BW5147 thymoma cells, its glucosidase II deficient variant Pha$^8$2.7 (Reitman et al., 1982) (both provided by Dr. R. Kornfeld, Washington University), and L cells were grown in Dulbecco’s modified Eagle’s minimum essential medium. Murine EL-4 thymoma cells and the human C1R cell line that stably expresses the HLA-B27 molecule (Anderson and Cresswell, 1994) (provided by Dr. P. Cresswell, Yale University) were cultured in RPMI 1640. CHO-K1 and its glucosidase I deficient variant CHO-Lec23 (Ray et al., 1991) were obtained from Dr. A. Helenius, Swiss Federal Insitute of Technology, and were grown in $\alpha$-MEM. Stably transfected \textit{Drosophila melanogaster} Schneider cells that express CNX along with H-2K$^b$ H chains in the presence of mouse $\beta_2$m (Danilczyk et al., 2000) were maintained in Schneider’s insect medium (Sigma). All media were supplemented with 10% FBS and antibiotics.
A rabbit antiserum (anti-8) directed against the C-terminus of the H-2K^b H chain which reacts with all conformational states of K^b was provided by Dr. Brian Barber, University of Toronto. Antiserum UCSF#2 reacts with the cytoplasmic tail of class I HLA H chains and was provided by Dr. Frances Brodsky, Stanford University. MAb PIN1.1, which reacts with invariant chain, was obtained from Dr. Tania Watts, University of Toronto. mAb 28-14-8S was used to detect D^b H chains (Ozato et al., 1980). Two rabbit antisera were used to isolate CNX. One is directed against the C-terminal 14 amino acids (anti-C-CNX) and the second was raised against the entire 462 residue ER luminal domain (anti-N-CNX). MAb 12CA5 that reacts with the influenza hemagglutinin (HA) epitope tag on CNX(HA) and the CNX 1-387(HA) mutant was provided by Dr. Paul Hamel, University of Toronto.

3.2.2 Construction of N-glycosylation Mutants and Expression in L Cells

The N-glycosylation mutants of the H-2 K^b H chain were generated by mutating the consensus glycosylation sequence, Asn-X-Ser/Thr, using the QuikChange™ Site-Directed Mutagenesis Kit (Stratagene) and full length H-2K^b cDNA in pcDNA3 (Invitrogen) as template. The QuikChange™ method requires 2 complementary mutagenic oligonucleotides of which only the coding strand is shown in the examples below. To remove the glycosylation site at residue 176, Asn 176 was changed to Lys using the mutagenic oligonucleotides (mutated base in lower case): 5' GC AGA TAT CTG AAC GGG AAG 3'. The glycosylation site at residue 86 was removed by substituting Asn 86 with Lys using 5' C CTG CTC GGC TAC TAC AAG CAG AGC AAG GGC GGC GGC GGC GGC GGC GGC GGC GGC GGC GGC 3' as the mutagenic oligonucleotide. To add a glycosylation site at position 256, Tyr 256 was changed to Asn with the mutagenic oligonucleotide 5' GGG AAG GAG CAG aAT TAC ACA TGC CAT GTG TAC C 3'. Recombinant plasmids were introduced into L cells using the SuperFect™ transfection reagent (Qiagen) and stably transfected cells were established by G418 selection.
Transient expression of HA-tagged calnexin (CNX(HA)) and a truncated ER luminal segment of calnexin (CNX 1-387(HA)) in L cells was conducted as described previously (Danilczyk et al., 2000).

3.2.3 Metabolic Radiolabeling and Immunoisolation

BW5147, Pha2.7, CHO-K1, CHO-Lec23 or L cells at a density of 1x10^7 cells/100mm dish were incubated in Met-free medium for 60 min at 23°C to deplete intracellular Met pools. They were then radiolabeled at 23°C for the times indicated in the various figure legends by the addition of 400μCi/ml [35S]Met. Castanospermine (CAS), when added, was present throughout the prelabelling and labelling periods. Cells were lysed for 30 min at 4°C in 1 ml of lysis buffer containing either 1% digitonin or 1% CHAPS in PBS, pH 7.4, 10mM iodoacetamide, 60 μg/ml Pefabloc® (Boehringer Mannheim) and 10 μg/ml each of leupeptin, antipain and pepstatin. For isolation of CNX and associated molecules, lysates were incubated with preimmune or anti-CNX antibodies for 2 h. Immune complexes were collected for 1 h using protein A-agarose beads and analyzed by SDS-PAGE as described previously (Danilczyk et al., 2000). For detection of calnexin-associated Kb, Db, HLA-B27 and invariant chain, digitonin lysates were subjected to sequential immunoprecipitation (Suh et al., 1999). Briefly, CNX-substrate complexes were recovered with anti-CNX antiserum, dissociated by heating at 42°C for 1 h in 0.2% SDS, adjusted to 2% Nonidet P-40, 5% skim milk and incubated with anti-class I H chain or anti-invariant chain antibodies. Immune complexes were collected and analyzed as above.

Radiolabeling of transfected Drosophila cells with [35S]Met, lysis, and immunoisolation were carried out as described previously (Vasilakos et al., 1996). Briefly, following induction of the metallothionein promoter with 1mM CuSO4 for 16 h, Drosophila cells were incubated for 1h in Met-free Schneider's medium in the presence or absence of 500 μg/ml of CAS. Cells were then radiolabeled with 0.5 mCi/ml [35S]Met
for 5 min in the presence or absence of CAS and lysed in digitonin lysis buffer. Lysates were incubated with anti-class I H chain or anti-C-CN X antibodies and immune complexes were collected on protein A-agarose followed by SDS-PAGE analysis using 10% gels (Laemmli, 1970). Radioactive proteins were visualized by fluorography.

3.2.4 Glycerol Density Gradient Centrifugation

L cells (1x10^7) expressing wild type or non-glycosylated H-2K^b were radiolabeled with [35S]Met for 30 min, lysed in 1 ml 1% digitonin buffer, and centrifuged briefly at top speed in an Eppendorf microfuge to remove insoluble material. A 0.5 ml aliquot of lysate was loaded onto a 12 ml, 10-40% (w/v) linear glycerol gradient prepared in digitonin buffer. The gradients were centrifuged at 4°C for 15 hr at 35,000 rpm using a Beckman SW 41 rotor. Fractions (0.75 ml) were collected from the top of the gradients and K^b H chains were immunoisolated from each fraction using anti-8 antiserum. As a control for the total amount of K^b molecules loaded onto the gradient an additional 0.5 ml sample of lysate was also immunoisolated with anti-8 antiserum.

3.2.5 Immunoblotting

For detection of H-2K^b-CN X complexes by immunoblotting, 1x10^7 L cells expressing wild type K^b or various K^b glycosylation mutants were lysed in digitonin lysis buffer and immunoisolated with anti-8 antiserum and protein A-agarose. Following SDS-PAGE analysis, proteins were transferred to nitrocellulose membrane (Suh et al., 1994) and the membrane was incubated with rabbit anti-N-CN X antiserum at 1:5,000 dilution followed by donkey anti-rabbit IgG horseradish peroxidase conjugate at 1:10,000 dilution (Jackson Laboratories). Immune complexes were visualized using enhanced chemiluminescence (Amersham).
3.3 Results

3.3.1 Calnexin Associates with Many Substrate Proteins when the Formation of Monoglucosylated Oligosaccharides is Blocked

To establish whether CNX associates with its substrates only via its lectin site or whether protein–protein interactions also contribute to this association, we examined the formation of CNX-substrate complexes in glucosidase I- or glucosidase II-deficient cell lines and in the presence of the glucosidase inhibitor, castanospermine (CAS). In an effort to minimize protein aggregation and to preserve potentially weak protein-protein interactions, metabolic radiolabeling of cells was conducted at 25°C and lysis was performed using the mild detergents digitonin or CHAPS. Furthermore, immune complexes were washed for the minimum number of times (typically three) required to preserve CNX-substrate interactions while minimizing recovery of non-specifically associated proteins.

Initially, the BW5147 mouse lymphoma cell line and its glucosidase II-deficient mutant, Pha².7, were radiolabeled with [³⁵S]Met and digitonin lysates were subjected to immunoisolation with two separate anti-CNX antisera. The anti-C-CNX antibody recognizes the last 14 residues of the cytoplasmic tail of CNX and the anti-N-CNX antibody is directed against the entire ER luminal domain (residues 1-462). As reported previously (Balow et al., 1995; Ora and Helenius, 1995), in addition to CNX which appeared as a major band of 90 kDa, a large number of newly synthesized proteins co-isolated as complexes with CNX from the parental BW5147 cells (Fig. 20A). A similar pattern of proteins was observed with the two independent anti-CNX antisera (Fig. 20A, lanes 2 and 3). A substantial number of these proteins were lost or reduced in intensity in the glucosidase II-deficient Pha².7 cells, reflecting their apparent requirement for monoglucosylated oligosaccharides for stable association with CNX. However, it is noteworthy that many other proteins remained firmly associated with CNX (Fig 20A, lanes 5 and 6). A similar result was obtained when parental BW5147 cells were treated
Figure 20. Association of newly synthesized proteins with CNX in the absence of glucosidase activity. A, BW5147 and Pha^{b}2.7 cells were incubated for 1 h in the absence (lanes 1-6) or presence (lane 7) of 1 mM CAS in Met-free medium, labeled for 10 min with [^{35}S]Met and lysed in 1% digitonin lysis buffer. Lysates were incubated with preimmune serum (PI) or with two polyclonal anti-CNX antisera, anti-N CNX and anti-C CNX, as indicated. Immune complexes were analyzed by reducing SDS-PAGE. B, CHO-K1 and CHO-Lec23 cells were radiolabeled and lysed as in A and then immunoprecipitated either with preimmune serum or with anti-N-CNX or anti-C-CNX antisera as indicated. C, HA-tagged CNX (CNX(HA)) and a truncated ER luminal segment of CNX (CNX 1-387(HA)) were transiently expressed in L cells. The cells were radiolabeled as in A and then lysed with buffer containing 1% CHAPS. The HA-tagged molecules and associated proteins were immunoprecipitated with mAb 12CA5. The mobilities of CNX in each panel are indicated.
with CAS to block glucosidase activity (Fig. 20A, compare lanes 2 and 3 with lane 7). Indeed, the patterns of CNX-associated proteins were remarkably similar in the PhaR2.7 and CAS-treated BW5147 cells. These proteins were not the result of non-specific interactions with the precipitating antibodies since a similar spectrum of proteins was obtained with the two independent anti-CNX antibodies and they were absent from control isolations performed with preimmune serum (Fig. 20A, lanes 1 and 4). To confirm these findings, we also compared glucosidase I-deficient Lec23 cells to their parental CHO cell line (Fig. 20B). Lec23 cells have been shown to possess little or no glucosidase I activity and no monoglucosylated oligosaccharides could be detected on glycoproteins (Ray et al., 1991). Remarkably, despite the block in formation of monoglucosylated oligosaccharides, there was no obvious reduction in the number of CNX-associated proteins recovered with each antiserum although some differences in the patterns of recovered proteins were apparent (Fig. 20B, compare lane 2 with lane 5 and lane 3 with lane 6). Again, these proteins were absent in control isolations performed with preimmune antiserum.

To further exclude the possibility that the proteins remaining associated with CNX in glucosidase-deficient cells or after CAS treatment were due to non-specific associations with immune complexes, HA-epitope tagged CNX (CNX(HA)) and a soluble variant (CNX 1-387(HA)) were prepared and transfected into mouse L cells. We showed previously that the CNX 1-387(HA) variant fails to form complexes with newly synthesized proteins (Danilczyk et al., 2000). The transfectants were radiolabeled with [35S]Met, lysed in 1% CHAPS, and subjected to immunosolation with anti-HA mAb. As shown in Fig. 20C, lane 1, a large number of newly synthesized proteins were recovered with the anti-HA mAb from cells expressing full length CNX(HA). Consistent with the experiments presented in Fig. 20, panels A and B, many proteins were also recovered in association with CNX(HA) following treatment of cells with CAS (Fig. 20C, lane 3). In contrast, only trace levels of proteins were recovered in association with the binding-impaired CNX 1-387(HA) variant (Fig. 20C, lane 2). This was also the case when the
CNX 1-387(HA) variant was isolated from CAS-treated cells (data not shown). Since the CNX 1-387(HA) variant was recovered under identical conditions of immune isolation as the CNX(HA) construct, this establishes that the CNX-associated proteins remaining after CAS treatment are indeed *bona fide* complexes and not merely proteins non-specifically adsorbed to anti-HA immune precipitates.

### 3.3.2 Differential Effects of Castanospermine on Defined CNX-Glycoprotein Complexes

To obtain further insight into the nature of proteins that remain associated with CNX under conditions that block the formation of monogluinosylated oligosaccharides, we examined the interactions of CNX with several defined glycoproteins. Initially, complexes between CNX and the mouse class I histocompatibility molecules, H-2Kβ and H-2Dβ, were studied using the EL4 thymoma cell line. Murine class I molecules consist of three subunits that assemble within the ER, a ~46 kDa transmembrane heavy (H) chain that possesses 2 or 3 Asn-linked glycans, the soluble 12 kDa β2m subunit, and an 8-10 residue peptide ligand. EL-4 cells were incubated with or without CAS and screened for proteins associated with CNX by immunoisolation with anti-CNX antibodies. As expected, many proteins co-isolated as complexes with CNX (Fig. 21A, lane 3). After CAS treatment, many of these complexes were lost but others were preserved (Fig. 21A, lane 4). To confirm that CAS was effective in inhibiting glucosidase activity in this and preceding experiments, the H-2Kβ and H-2Dβ proteins were isolated directly and their H chain subunits were shown to exhibit the characteristic decrease in mobility following CAS treatment that reflects unprocessed N-linked oligosaccharides (Fig. 21A, compare lanes 5 and 6 and lanes 7 and 8). The interactions of H-2Kβ and H-2Dβ H chains with CNX were completely blocked following CAS treatment. This was evidenced by the loss of the intense 46 kDa species in the anti-CNX immunoprecipitate from CAS-treated cells (Fig. 21, compare lanes 3 and 4) and also by the absence of H chains following their immune isolation from solubilized anti-CNX immunoprecipitates (Fig. 21A, compare
Figure 21. Castanospermine inhibits CNX interaction with K\textsuperscript{b} H chains in EL4 cells but not in Drosophila cells. A, EL-4 cells were incubated for 1 h in the absence or presence of 1 mM CAS as indicated and then radiolabeled as in Fig. 1. Digitonin lysates were treated either with preimmune serum (lanes 1, 2), anti-N CNX antiserum (lanes 2, 3), anti-8 antiserum to isolate K\textsuperscript{b} H chains (lanes 4, 5) or mAb 28-8-14S to isolate D\textsuperscript{b} H chains (lanes 7, 8). To recover K\textsuperscript{b} or D\textsuperscript{b} H chains that were associated with CNX, CNX and associated proteins were first isolated with anti-N CNX antiserum. Immune complexes were then dissociated in SDS and subjected to a second round of immune isolation with anti-8 or 28-14-8S antibodies (lanes 9-12). Isolated proteins were analyzed by SDS-PAGE. B, Drosophila cells expressing K\textsuperscript{b} H chains, \(\beta_2m\) and CNX were incubated for 1 h in the absence or presence of 1mM CAS, radiolabeled with \([^{35}\text{S}]\text{Met}\) for 10 min, and lysed in digitonin lysis buffer. Lysates were treated with anti-C-CNX or anti-8 antiserum as indicated and immune complexes were either analyzed directly or following digestion with endo H to remove immature N-linked glycans. C, Drosophila cells expressing D\textsuperscript{b} H chains, \(\beta_2m\) and CNX were incubated for 1 h in the absence or presence of 1mM CAS, radiolabeled with \([^{35}\text{S}]\text{Met}\) for 10 min, and lysed in digitonin lysis buffer. Lysates were treated with 28-14-8S mAb and immune complexes were heated in sample buffer at 55°C prior to analysis by SDS-PAGE under non-reducing conditions.
lanes 9 and 10 and lanes 11 and 12). Thus, murine class I molecules in EL4 cells appear to depend extensively on lectin-oligosaccharide interactions for their stable association with CNX.

A different situation was observed when we expressed the Kb H chain, β2m, and CNX in Drosophila melanogaster Schneider cells. These cells possess the glucosidases and glucosyltransferase required for a functional deglucosylation-reglucosylation cycle (Christodoulou et al., 1997; Parker et al., 1995; Smith, 1992) but they lack the specialized TAP transporter that transports peptide ligands for class I molecules from the cytosol to the lumen of the ER. We showed previously that under these conditions the H chain assembles with β2m but remains peptide deficient and the heterodimer is retained in the ER in association with CNX (Jackson et al., 1994; Vassilakos et al., 1996). In transfected Drosophila cells, the Kb molecule is the major substrate associated with CNX and it can be identified directly in anti-CNX immunosolutes. When these cells were incubated in the absence or presence of CAS, the Kb H chain mobility was reduced in response to CAS treatment reflecting a block in glucose trimming (Fig. 21B, lanes 5, 6). However, when CNX-Kb complexes were recovered with anti-CM Ab, a significant fraction of Kb molecules remained associated with CNX after CAS treatment (Fig. 21B, lanes 1, 2), and these molecules possessed the slower electrophoretic mobility indicative of unprocessed oligosaccharides. To confirm that the change in mobility of the CNX-associated Kb H chain following CAS treatment was due to carbohydrate modification, the immune complexes were digested with Endo H to remove Asn-linked oligosaccharides after which they possessed the same mobility with or without CAS treatment (Fig. 21B, lanes 3, 4). It is noteworthy that CNX-Kb complexes could also be detected in the absence or presence of CAS by immunosolating Kb molecules and visualizing the associated CNX band (Fig. 21B, lanes 5, 6). Similar results were obtained with the Dp H chain (Fig. 21C). It is conceivable that the altered glycans resulting from CAS-treatment promoted aggregation of Kb and Dp molecules and the non-specific inclusion of CNX in such aggregates. We showed previously that a mix of disulfide-linked and non-disulfide-linked
aggregates of K\textsuperscript{b} or D\textsuperscript{b} H chains can readily be detected under conditions of non-reducing SDS-PAGE when samples are heated only to 55°C (Vassilakos et al., 1996). Under these conditions, no aggregated D\textsuperscript{b} molecules could be detected at the top of the separating gel (Fig. 21C). Thus, in Drosophila cells, incompletely assembled class I H chains are capable of associating with CNX in the apparent absence of monoglucosylated oligosaccharides on the H chain. The difference relative to EL4 cells is probably related to detection sensitivity since H chains are the major substrates associated with CNX in Drosophila cells.

Since the murine K\textsuperscript{b} H chain possesses two glycans at positions 86 and 176, there was the formal possibility that CAS treatment may not fully block the formation of monoglucosylated oligosaccharides at both sites thereby permitting some degree of interaction with CNX. Consequently, we examined CNX association with the human class I molecule, HLA-B27, which has only a single glycan at position 86. C1R cells stably expressing HLA-B27 were treated for 1 h with 1 mM CAS, lysed, and CNX-associated proteins were isolated with anti-CN Ab. As shown in Fig. 22A, CAS treatment resulted in the loss of many, but not all, complexes of newly synthesized proteins with CNX. To assess the fate of CNX-HLA-B27 complexes, immune complexes containing CNX and associated proteins were dissociated and the released proteins were subjected to a second round of immune isolation with anti-HLA antibody. As shown in Fig. 22B, left panel, equivalent amounts of HLA-B27 H chains were recovered from CNX complexes in the absence or presence of 0.25 or 1 mM CAS. Furthermore, the CNX-associated H chains from CAS-treated cells exhibited the reduced mobility indicative of blocked glucose trimming. This was most apparent at 1 mM CAS; at the lower concentration of 0.25 mM the mobility shift was less pronounced suggesting an incomplete block in glucose trimming (Fig. 22B, left and center panels). We also assessed the tendency for HLA-B27 to form insoluble aggregates following CAS treatment. However, no H chains could be sedimented following centrifugation at 100,000 x g (Fig. 22B, right panel). Therefore, for a singly glycosylated glycoprotein which clearly lacked
Figure 22. HLA-B27 and invariant chain remain associated with CNX in CAS-treated C1R-B27 cells. A, C1R-B27 cells were incubated in the absence or presence of 1 mM CAS and then were radiolabeled with [³⁵S]Met for 5 min and lysed in digitonin lysis buffer. CNX and associated proteins were immunoisolated either with anti-N-CN X (lanes 2, 5) or anti-C-CN X (lanes 3, 6) antisera. Lanes 1 and 4 depict lysates treated with preimmune (PI) serum. The mobilities of CNX and invariant chain are indicated. B, C1R-B27 cells were incubated with various amounts of CAS as indicated prior to radiolabeling and lysis as in panel A. To isolate CNX-associated HLA-B27 molecules, CNX-containing complexes were first recovered with anti-N-CN X serum and, following dissociation in SDS, HLA-B27 H chains were isolated with UCSF#2 anti-serum (left panel). In addition, 1/10 the amount of each lysate was treated directly with UCSF#2 antiserum (center panel). To test for the presence of insoluble H chain aggregates, radiolabeled cell lysates were centrifuged at 100,000 x g for 30 min and HLA-B27 H chains were immunoisolated from supernatant (S) and solubilized pellet (P) fractions (right panel). C, Radiolabeled lysates of C1R-B27 cells were prepared as in panel B and CNX-associated invariant chains were recovered using anti-N-CN X antibody followed by a second round of immune isolation with the anti-invariant chain mAb PIN1.1 (left panel). Invariant chains were also isolated directly from 1/10 the amount of each lysate using mAb PIN1.1 (center panel). Radiolabeled cell lysates were also centrifuged at 100,000 x g as described in panel B to detect any invariant chain aggregates (right panel).
monoglucosylated oligosaccharide, interactions with CNX were fully maintained.

We extended these experiments to include an endogenous glycoprotein of C1R cells, the invariant chain. Invariant chain is CNX's major substrate in these cells and it can be readily observed as an intense ~35 kDa band in anti-CNX immune isolates (Fig. 22A, lanes 2 and 3). Invariant chain also appeared to be present at reduced intensity in anti-CNX immune complexes following 1 mM CAS treatment (Fig. 22A, lanes 5 and 6). To confirm this, immune complexes containing CNX and associated proteins were isolated from cells treated with 0-1 mM CAS, dissociated, and subjected to re-immunoinosolation with anti-invariant chain mAb. In this experiment, the bulk of invariant chain molecules remained associated with CNX after CAS treatment (Fig. 22C, left panel) and they exhibited the reduced electrophoretic mobility that accompanies a block in glucose trimming (Fig. 22C, left and center panels). Invariant chain also did not form insoluble aggregates following CAS treatment (Fig. 22C, right panel).

We conclude that for three different glycoproteins, the presence of monoglucosylated oligosaccharides is not required for association with CNX. Furthermore, the associations detected in CAS-treated cells are unlikely to be a consequence of the non-specific inclusion of CNX in large glycoprotein aggregates.

3.3.3 Calnexin Associates with Nonglycosylated Class I Molecules in vivo

Although CNX fails to bind to purified Glc$_3$Man$_9$GlcNAc or Glc$_2$Man$_9$GlcNAc oligosaccharides in vitro (Ware et al., 1995), the possibility remained that the CNX-substrate complexes observed in CAS-treated or glucosidase deficient cells could be mediated through lectin interactions with di- or tri-glucosylated oligosaccharides. To address this issue, we tested the association of CNX with K$^b$ H chains mutated to lack Asn-linked oligosaccharides. In addition, we examined the consequence of varying the number of N-linked glycans on CNX binding. The various K$^b$ glycosylation mutants containing 0, 1, 2, or 3 glycans are depicted in Fig. 23A. The wild type K$^b$ H chain and glycosylation mutants were stably expressed in murine L cells (which contain $\beta_2$m and all
other components required for normal assembly of class I molecules). Fig. 23B shows that the various forms of the K\(^b\) H chain were synthesized in L cells and possessed the electrophoretic mobilities expected based on their different numbers of oligosaccharide chains. We also established that the differentially glycosylated H chains associated normally with \(\beta_2\)m and that the singly, doubly, and triply glycosylated proteins were all transported from the ER to the Golgi apparatus at comparable rates (data not shown). In the case of the nonglycosylated molecule, its stability was similar to wild type K\(^b\) and it could be detected at the cell surface by flow cytometry albeit at lower levels than observed for the wild type protein (data not shown). Collectively, these findings suggest that changes in the number of N-glycans do not cause major misfolding of the K\(^b\) molecule.

The various L cells transfectants were radiolabeled, lysed in digitonin buffer, and CNX-associated proteins were isolated with anti-CN X Ab. K\(^b\) H chains were then recovered from the anti-CN X immunoprecipitate in a second round of immune isolation with anti-H chain antiserum. As demonstrated in Fig. 23C, left panel, all of the K\(^b\) glycosylation mutants associated with CNX including the non-glycosylated mutant. The same result was obtained when K\(^b\) H chains were isolated from non-radiolabeled cell lysates and the presence of associated CNX was detected by immunoblotting (Fig. 23C, right panel). However, a comparison of the relative amounts of K\(^b\) mutants synthesized with the relative amounts recovered in association with CNX indicated that the nonglycosylated mutant formed complexes with CNX somewhat less efficiently that the glycosylated forms (compare Fig. 23B with 23C, left panel). This suggests that the presence of one or more N-linked glycans increases the stability of the CNX-K\(^b\) interaction.

Because aggregation is a common fate for glycoproteins that have been treated with tunicamycin to block N-glycosylation or that have been mutated to lack Asn-linked glycans, it was essential to determine whether the co-isolation of the nonglycosylated K\(^b\)
Figure 23. Association of K\textsuperscript{b} H chain glycosylation mutants with CNX. A. The locations of glycosylation sites in the various K\textsuperscript{a} H chain glycosylation mutants are indicated by residue number (CHO* represents the nonglycosylated protein). The K\textsuperscript{b} H chain is depicted in black and the β\textsubscript{2}m subunit is shown in gray. B, C. Untransfected L cells (lane 1) or L cells expressing wild type K\textsuperscript{b} with two glycosylation sites (lane 2), K\textsuperscript{b} with a single glycan (lane 3), nonglycosylated K\textsuperscript{b} (lane 4), or K\textsuperscript{b} with three glycans (lane 5) were radiolabeled with \textsuperscript{35}S]Met for 10 min and lysed in digitonin lysis buffer. K\textsuperscript{b} H chains were immunoisolated with anti-8 antiserum (panel B) and CNX-associated K\textsuperscript{b} H chains were recovered by immunoisolating first with anti-N-CNX antiserum followed by complex dissociation and isolation of K\textsuperscript{b} H chains with anti-8 antiserum (panel C, left). In panel C, right, CNX-associated K\textsuperscript{b} H chains were also detected by isolating the various K\textsuperscript{b} glycosylation mutants from lysates of transfected cells using anti-8 antiserum.

Following separation of proteins by reducing SDS-PAGE, proteins were transferred to nitrocellulose membrane, probed with anti-N-CNX antiserum, and visualized with a chemiluminescence-based detection system. Lane 1, untransfected cells; lane 2, wild type K\textsuperscript{b}; lane 3, unglycosylated K\textsuperscript{b}; lane 4, K\textsuperscript{b} with a single glycan; lane 5, K\textsuperscript{b} with three glycans. Lane 6 represents 1/100 of the cell lysate used for immunoprecipitation and serves as a positive control for CNX detection. D, L cells expressing either wild type or unglycosylated K\textsuperscript{b} H chains were radiolabeled, lysed in digitonin-containing buffer and the lysates applied to the top of a 10-40% glycerol density gradient. Following centrifugation for 15 hours at 34,000 rpm in an SW41 rotor, the gradients were fractionated and K\textsuperscript{b} H chains were isolated from each fraction with anti-8 antiserum. Isolated proteins were analyzed by reducing SDS-PAGE.
A  
\[ K^b \text{ CHO-} \]  \[ K^b \text{ 86} \]  \[ K^b \text{ 86,176} \]  \[ K^b \text{ 86,176,256} \] (wild type)

B  
anti-8

C  
1st IP  N-CNX  2nd IP  anti-8
IP: anti-8  Blot: N-CNX

D  
direction of sedimentation

\[ K^b \text{ H chain-WT} \]  \[ K^b \text{ H chain-unglycosylated} \]

Fraction #  1  2  3  4  5  6  7  8  9  10  11  12  13
protein with CNX occurred because of a specific protein-protein interaction or was due merely to the non-specific inclusion of CNX and possibly many other glycoproteins in a large aggregate of unglycosylated H chains. To test this possibility, transfectants expressing the wild type K\textsuperscript{b} H chain or its nonglycosylated mutant were radiolabeled, lysed in digitonin lysis buffer, and the lysates subjected to sedimentation through glycerol density gradients. Fractions were collected and K\textsuperscript{b} molecules were isolated from each fraction. As shown in Fig. 23D, the wild type K\textsuperscript{b} H chain was detected mainly in fractions 5 and 6. A similar distribution was observed for the nonglycosylated K\textsuperscript{b} H chain; there was no evidence of large aggregates that sediment near the bottom of the gradient. When CNX-H chain complexes were isolated from fractions with anti-CN X Ab, the nonglycosylated H chain was again found to sediment primarily in fractions 5 and 6 indicating that CNX-unglycosylated H chain complexes are not found in large aggregates (data not shown). Therefore, we conclude that the association between CNX and nonglycosylated K\textsuperscript{b} H chains is a specific polypeptide-based interaction that can occur independently of lectin-oligosaccharide binding.

3.4 Discussion

Our findings indicate that when care is taken to preserve weak interactions during metabolic radiolabeling and immune isolation, complexes between CNX and a diverse array of proteins lacking monoglucosylated oligosaccharides can be detected. This was observed when the formation of Glc\textsubscript{3}Man\textsubscript{7,9}GlcNAc\textsubscript{2} oligosaccharides was blocked with the glucosidase inhibitor CAS or through the use of cells lacking either glucosidase I or glucosidase II. Complexes were not likely due to weak interactions between di- and triglucosylated oligosaccharides on the glycoprotein substrate and CNX's lectin site because interactions were maintained even when a non-glycosylated substrate was tested. Also, CNX does not bind detectably to these oligosaccharides \textit{in vitro} (Ware et al., 1995). Furthermore, the association of CNX with proteins lacking monoglucosylated oligosaccharides did not arise from the non-specific inclusion of CNX into protein
aggregates nor was it due to non-specific adsorption of proteins onto anti-CNX immune precipitates. Consequently, we conclude that CNX is capable of associating with diverse proteins \textit{in vivo} through polypeptide-based interactions in addition to its well-characterized lectin-oligosaccharide interaction.

It is noteworthy that the interactions between various proteins and CNX were affected to markedly different extents when the formation of monoglucosylated oligosaccharides was blocked. Examination of the patterns of proteins recovered in anti-CNX immune complexes (Figs 20, 21A and 22A) revealed that some proteins were completely lost, others appeared to be present at reduced levels, and others seemed unaffected. This situation was also observed when complexes between CNX and specific glycoproteins were examined. The murine class I molecules, H-2K\textsuperscript{b} and H-2D\textsuperscript{b}, could not be recovered in association with CNX following CAS treatment of EL4 cells although some interaction could be detected when overexpressed in \textit{Drosophila} cells. However, unglycosylated H-2K\textsuperscript{b} remained firmly associated with CNX. This likely reflects different conformational states between the glycosylated and unglycosylated proteins with the accompanying presentation of different polypeptide determinants to CNX. In contrast to the glycosylated murine class I molecules, the human class I molecule, HLA-B27, and the human MHC class II invariant chain were largely unaffected in their interactions with CNX following CAS treatment. Collectively, these findings suggest that different glycoproteins exhibit different dependencies on lectin-oligosaccharide interactions for their stable association with CNX. In other words, in the absence of lectin-oligosaccharide binding, different proteins exhibit differences in the strength of their polypeptide-based interactions with CNX.

Consistent with the above \textit{in vivo} findings, substrate-specific differences have also been observed when the interactions of CNX or CRT with unglycosylated protein or peptide substrates were studied \textit{in vitro}. For example, when the ability of CNX to suppress the aggregation of various unfolded proteins was assessed, equimolar amounts of CNX effectively suppressed the aggregation of the naturally unglycosylated proteins
citrate synthase and malate dehydrogenase (Ihara et al., 1999). However, more than a 3-fold molar excess was required to fully suppress the aggregation of enzymatically deglycosylated soybean agglutinin (Ihara et al., 1999) and a 30-fold molar excess was only partially effective in preventing the aggregation of enzymatically deglycosylated α-mannosidase (V.S. Stronge and D. B. Williams, unpublished observations). CNX did not bind to any of these proteins without their prior denaturation. An analysis of the binding of 39 different peptides to CRT also revealed its marked preference for certain peptides. In general, hydrophobic peptides lacking acidic residues were favored and there also appeared to be a minimum length requirement (Jorgensen et al., 2000). A preference for binding to hydrophobic peptides is a common characteristic among molecular chaperone families and, for some chaperones such as the cytosolic Hsp90 and Tric proteins, there are clear preferences for certain protein substrates over others (Gething, 1997).

Differences in the strength of polypeptide-based interactions between CNX or CRT and various glycoprotein substrates coupled with variations in the stringency of isolation conditions can account for many of the conflicting results reported in the literature concerning the association of these chaperones with specific glycoproteins following treatment with glucosidase or glycosylation inhibitors. For example, if glycosylation is prevented or if glucosidases are inhibited, complexes are not detected between CNX or CRT with the α and β subunits of the T cell receptor (Kearse et al., 1994), influenza hemagglutinin (Hammond et al., 1994), VSV G glycoprotein (Hammond and Helenius, 1994), ribonuclease B (Rodan et al., 1996), myeloperoxidase (Nauseef et al., 1995), cruzipain (Labriola et al., 1999), and tyrosinase (Toyofuku et al., 1999). However, complexes can readily be detected at normal or reduced levels under conditions of deglycosylation or glucosidase inhibition with the ε and δ subunits of the T cell receptor (Rajagopalan et al., 1994; van Leeuwen and Kearse, 1996), P glycoprotein (Loo and Clarke, 1995), erythrocyte AE1 (Popov and Reithmeier, 1999), acid phosphatase (Jannatipour et al., 1998), MHC class II α and β chains (Arunachalam and Cresswell, 1995), MHC class II invariant chain (Zhang and Salter, 1998), MHC class I H chain (this
study), and HIV gp160 (Otteken and Moss, 1996). Other interesting examples of variability include the finding that CAS treatment almost completely prevented the formation of complexes between CNX and coagulation factors V and VIII but only partially inhibited the formation of complexes with CRT (Pipe et al., 1998). Furthermore, CAS prevented the formation of complexes between CNX and the α subunit of the acetylcholine receptor in one study (Chang et al., 1997) but had little effect on complex formation in another (Keller et al., 1998). The main difference appeared to be the use of Triton X-100 for cell lysis in the former study as opposed to the milder CHAPS detergent in the latter. In addition to these studies that focused on complexes of specific glycoproteins with either CNX or CRT, variable results have also been reported when the entire spectrum of CNX or CRT associated proteins were examined. In agreement with our current study, Kearse et al. observed strong association of many proteins with CNX in the glucosidase II-deficient Pha^82.7 cell line and in CAS-treated wild-type cells, even though associations with TCRα and TCRβ were virtually eliminated (Kearse et al., 1994). In contrast, Helenius and co-workers observed an almost complete elimination of CNX- or CRT-associated proteins in Pha^82.7 cells and in CAS-treated wild type cells (Ora and Helenius, 1995; Peterson et al., 1995). Likewise, in glucosidase I deficient Lec23 cells, we observed very few changes in the number of CNX-associated proteins relative to wild type CHO whereas Ora and Helenius documented a substantial loss of these complexes (Ora and Helenius, 1995). It should be noted however that the two studies varied greatly in that we detected a large array of CNX-associated proteins whereas Ora and Helenius recovered only four distinct complexes.

The preceding survey illustrates how conflicting conclusions can arise when generalizations are made on the basis of studying the association of a single glycoprotein with either CNX or CRT or when only a single type of isolation condition is employed. It is exceedingly unlikely that all of the examples of complexes between CNX or CRT with unglycosylated or non-glucose trimmed proteins can be dismissed on the basis of their non-specific inclusion in aggregates, trapping within detergent micelles, and insolubility
of folding intermediates as has frequently been claimed (Cannon et al., 1996; Parodi, 2000; Rodan et al., 1996; Zapun et al., 1997). Indeed we have eliminated most of these possibilities in the present study. Rather, a view more consistent with the data is that CNX and CRT possess both a lectin site and a polypeptide binding site and that the latter binds with varying affinities to polypeptide segments of different glycoproteins. Such a dual binding model (Ware et al., 1995) has the advantages that it accommodates the in vitro demonstrations of polypeptide interactions between diverse substrates and either CNX or CRT (Ihara et al., 1999; Jorgensen et al., 2000; Saito et al., 1999) and it evokes the possibility of an enhanced avidity of chaperone-glycoprotein interactions via contacts through two binding sites.

In fact, a lectin-only type of interaction is rather difficult to rationalize in light of the stable complexes that occur between CNX or CRT and various monoglycosylated glycoproteins. Recent studies have demonstrated that the affinity of CRT for IgG carrying a single Glc,Man,GlcNAc, oligosaccharide is 1-2 μM (Patil et al., 2000). Glycans with dissociation constants greater than about 1 μM are typically retarded on immobilized lectin columns rather than binding tightly (Osawa, 1987). Such chromatographic behavior has been documented for the interaction of monoglycosylated oligosaccharides with immobilized CRT (Spiro et al., 1996) and is consistent with our observation that Glc,Man,GlcNAc, bound to immobilized CNX is readily released upon washing (Ware et al., 1995). An increased affinity could result if CNX and CRT are oligomeric and capable of binding to multiple oligosaccharides on a glycoprotein substrate. Indeed there have been several reports of enhanced CNX interactions when a glycoprotein is converted from a singly to a doubly glycosylated form (Rodan et al., 1996) (Cannon et al., 1996; Popov and Reithmeier, 1999). This could be due either to oligomeric CNX or to a bivalent CNX interaction that is induced by the precipitating anti-CNX antibody. Although we and others have provided gel filtration evidence suggestive of CNX or CRT oligomers (Ihara et al., 1999; Saito et al., 1999; Zapun et al., 1998) this appears not to be the case. First, no increase in apparent affinity was observed
when CRT was incubated *in vitro* with IgG possessing two monoglucosylated oligosaccharides (Patil et al., 2000). Second, recent biophysical studies have demonstrated that CRT is monomeric and possesses a highly asymmetric structure that accounts for its anomalous gel filtration behaviour (Bouvier and Stafford, 2000). Third, CNX behaves as a monomer as evidenced by sedimentation through a sucrose density gradient (Tatu and Helenius, 1997) and it is highly asymmetric with a single lectin site as revealed by its recently solved x-ray crystallographic structure (M. Cygler, personal communication). Finally, using epitope-tagged variants of CNX or CRT in transfected cells we have consistently been unable to co-isolate these proteins as mixed oligomers with the endogenous chaperones (U. Danilczyk and D. Williams, unpublished observations). Despite the weak lectin-oligosaccharide binding affinity and the apparent lack of CNX or CRT oligomers, complexes of these chaperones even with singly glycosylated glycoproteins (such as the human class I molecule, α-fetoprotein, and the nicotinic acetylcholine receptor α-subunit) are stable to detergent solubilization and immune isolation. Such stable associations can most readily be explained by a dual binding mechanism that encompasses both lectin-oligosaccharide and polypeptide-based interactions. Further proof of this model must await the identification of the polypeptide binding sites on CNX and CRT and the elucidation of their fine binding specificities.
CHAPTER IV
The Role of ERp57 in Disulfide Bond Formation During Folding and Assembly of Major Histocompatibility Complex Class I as Assessed in a *Drosophila* Expression System

Please Note: The cDNA ERp57 construct used in this study was prepared by D. Tessier (see section Material and Methods)
4.1 Introduction

On the cell surface class I molecules are composed of a polymorphic transmembrane H chain, a soluble, noncovalently associated polypeptide termed \( \beta_2m \), and a peptide ligand consisting of \( \sim 8-10 \) amino acids (Nathenson et al., 1981). The assembly of class I molecules take place in the ER where, prior to stable interaction with \( \beta_2m \), the H chains undergo glycosylation and two highly conserved disulfide bonds are formed (Tector and Salter, 1995). The intra-chain disulfide bonds, within the \( \alpha_2 \) and \( \alpha_3 \) domains, are probably formed co-translationally or shortly after synthesis. The formation of these bonds is required for correct folding and assembly and only the fully disulfide bonded form is found in association with \( \beta_2m \) and peptides (Smith et al., 1995; Warburton et al., 1994).

CNX is the dominant molecular chaperone that associates with newly synthesized class I H chains early during H chain translocation into the ER (Degen et al., 1992; Jackson et al., 1994). CNX was shown to associate with H chains which have none, one or two disulfide bonds formed, suggesting that this protein functions to protect the unfolded protein from nonspecific association prior to the formation of disulfide bonds (Tector and Salter, 1995).

The interaction of MHC class I with CNX has not been kinetically separated from its interactions with an additional house-keeping molecule called ERp57. ERp57 is a member of the thioredoxin family of enzymes, a subset of which resides in the ER. The prototypical ER resident member of the family is protein disulfide isomerase (PDI). These enzymes catalyze the formation and isomerization of the disulfide bonds in proteins undergoing folding and assembly in the ER, ensuring that the correct interchain and intrachain disulfide bonding patterns of proteins or assemblies of proteins are achieved (Gilbert, 1990). Apart from being a thiol oxidoreductase, ERp57 was also implicated to function as a cysteine-dependent protease (Urade et al., 1992) and as a carnitine palmitoyl transferase (Murthy and Pande, 1994).
Recently, ERp57 was reported to interact specifically with glycosylated versions of membrane and secretory proteins (Oliver et al., 1997; Elliott et al., 1997). Although CNX and CRT have been identified as lectins, there has been no suggestion that ERp57 is a lectin. ERp57 associates transiently with newly synthesized glycoproteins apparently in an indirect fashion by binding to CNX and CRT (van der Wal et al., 1998). Consequently, it was suggested that ERp57 may function together with CNX and CRT in promoting folding and assembly of glycoproteins in the ER (Oliver et al., 1997; Elliott et al., 1997; Zapun et al., 1998). In the case of class I molecules, ERp57 has been shown to interact with class I H chain-CN X complexes in addition to H chain-β2m-CRT complexes (Lindquist et al., 1998; Hughes and Cresswell, 1998). The thiol-dependent oxidoreductase activity of ERp57 is believed to be important for the formation of H chain disulfide bonds (Cresswell et al., 1999; Lindquist et al., 1998).

The ERp57 thiol oxidoreductase function was recently demonstrated in vitro. ERp57 promoted disulfide bond formation in monoglucosylated RNase B and its activity was greatly enhanced when ERp57 and RNase B were incubated in the presence of CNX or CRT (Zapun et al., 1998). Based on these in vitro studies a model has been proposed in which CNX or CRT binds to a monoglucosylated N-linked glycan of a newly synthesized glycoprotein, and ERp57 associates with these chaperones. The association of ERp57 with CNX and CRT is direct and does not require the presence of the complex's substrate. In this model, by virtue of its stable association with CNX and CRT, ERp57 is brought into contact with the newly synthesized glycoproteins and is able to modulate their folding. The glycoprotein substrate is released from CNX and CRT, and thereby its specific association with ERp57, after the removal of the last glucose residue from the glycan by glucosidase II (as described in details for the "lectin only" model in chapter I). A combination of the binding and release mechanism incorporating glucosidase II and UGGT as well as formation and/or isomerization of disulfide bonds
mediated by ERp57 generates a properly folded glycoprotein with correct arrangements of disulfide bonds (High et al., 2000).

As yet there is no direct information concerning how ERp57 functions in class I biogenesis. Since ERp57 is found in class I H chain complexes with CNX, and these complexes are generated early during the class I assembly pathway, an obvious possibility is that ERp57 may promote intrachain disulfide bond formation in the H chain and additionally affect H chain structure to render it competent to form complexes with $\beta_2$m. ERp57 was also detected at the later stage of assembly, namely in complexes consisting of class I heterodimers and the TAP transporter. It is thought that ERp57 may be recruited to this peptide-loading complex to rearrange disulfide bonds and, by doing so, regulate the peptide binding ability of class I molecules (Cresswell et al., 1999)

The purpose of this study is to assess the involvement of ERp57 in the early stages of class I biogenesis in vivo. Since ERp57, at least in vitro, can promote disulfide bond formation in glycoproteins that interact with CNX, we examined the contribution of ERp57 to the formation of disulfide bonds within the class I H chain. In addition, we assessed the effect of ERp57 on the efficiency of H chain folding and on the rate of H chain assembly with $\beta_2$m. In these studies, we used the heterologous Drosophila expression system previously developed to investigate the molecular chaperone functions of CNX and CRT. Since Drosophila cells do not express class I molecules, $\beta_2$m or the TAP transporter, they provide an excellent system to investigate early events leading to class I heterodimer formation. Furthermore, based on a search of the Drosophila genome these cells do not possess the gene encoding ERp57.

4.2 Material and Methods

4.2.1 Cell Lines and Antibodies

Drosophila melanogaster Schneider cells were maintained in Schneider's insect medium (Sigma) with 10% FBS and antibiotics. Stable transfected derivatives were
cultured in the same medium supplemented with 500 μg/ml Geneticin and 300 μg/ml hygromycin (GIBCO-BRL).

The following mAb were used for the isolation of class I molecules: mAb 20-8-4S which reacts with H-2K\textsuperscript{b} H chains associated with β\textsubscript{2}m (Ozato and Sachs, 1981), and mAb 28-14-8S which recognizes a conformational epitope in the α3 domain of free or β\textsubscript{2}m-associated D\textsuperscript{b} H chains (Ozato et al., 1980). A rabbit antiserum (anti-8) directed against the C-terminus of the H-2K\textsuperscript{b} H chain which reacts with all conformational states of K\textsuperscript{b} was provided by Dr. Brian Barber, University of Toronto (Smith et al., 1986).

Unassembled mouse class I H chains were isolated using a rabbit antiserum (anti-HC) provided by Dr. Hidde Ploegh, Harvard University (Machold et al., 1995). A rabbit antiserum raised against the C-terminal 14 amino acids of CNX was used to isolate full-length CNX (Jackson et al., 1994). Polyclonal antibody to detect ERp57 was provided by Dr. David Thomas, Biotechnology Research Institute, Montreal.

4.2.2 Construction of ERp57 cDNA and its Expression in Drosophila Cells

Human ERp57 cDNA in the pRMHa3 vector was obtained from Dr. Daniel Tessier (Biotechnology Research Institute, Montreal). The coding portion of ERp57 (without its signal peptide) was fused to the honeybee melittin signal peptide in the pFastBac-1 vector. For expression in Drosophila cells, the 1,532 bp BamHI-SacI restriction fragment of pFastBac-1 containing full length Mel-ERp57 cDNA, was subcloned into the BarnHI and Hinc II sites of the pRMHa3 expression vector. The sequence of the Mel-ERp57 construct at the 5’ end is as follows:

```
S’GGATCCCGGTCCGCACAAGATGAAATTCTTAGTCAACGTTGCA
CTAGTSSSTATGGTGTACATCTTACATCDWXG/GCCTCCGACGT
GCTAGAA...
```

The underlined sequence represents the BamHI site, the sequence between start codon ATG and codon representing the ERp57 mature protein in bold, encodes the Melittin Signal Peptide (21 amino acids).
In the *Drosophila* expression vector pRMHa3, cDNAs are under the control of the metallothionein promoter (Bunch et al., 1988). Stably transfected *Drosophila melanogaster* Schneider cell lines containing either D^b^ H chain or K^b^ H chain-β_2m^ and CNX were established by co-transfecting a plasmid containing the neomycin resistance gene (phshsneo) plus multiple pRMHa3 plasmids encoding CNX along with H-2K^b^ or D^b^ H chains in the presence or absence of mouse β_2m^ (Danilczyk et al., 2000). These stable transfectants were subsequently co-transfected with the Mel-ERp57-pRMHa3 construct and a plasmid containing the hygromycin resistance gene (pCoHYGRO)(Invitrogen). Following induction with 1 mM CuSO_4^ for 16 h, expression of class I molecules was analyzed by metabolic radiolabeling followed by immunoprecipitation with K^b^ or D^b^ specific antibodies. The expression of ERp57 was detected by Western blotting. To ensure that all four proteins ERp57, CNX, H chain and β_2m^ were expressed within one cell, the cells were cloned by the soft agar technique and screened for expression (Hapel et al., 1981). Two clones were selected for each transfected cell line, and only clones expressing comparable levels of these proteins were used in all subsequent experiments.

### 4.2.3 Metabolic Labeling, Immunoprecipitation, and Gel Electrophoresis

Labeling of *Drosophila* cells with [35S]Met, lysis, and immunoprecipitation were carried out as described previously (Vassilakos et al., 1996). Briefly, following induction with 1 mM CuSO_4^ for 16 h, transfected *Drosophila* cells were incubated for 30 min in Met-free Schneider's medium. Cells were then radiolabeled with [35S]Met for 5 min or as indicated in figure legends, chased for various times and lysed in a buffer containing 1% digitonin or 1% NP-40 as described in chapter 2. Lysates were incubated for 2 hours at 4°C with anti-class I or anti-CNAX antibodies followed by a 1 hour incubation with protein A-agarose beads. Immune complexes were analyzed by reducing or non-reducing SDS-PAGE using 10% gels (Laemmli, 1970) and radioactive proteins were visualized by
fluorography. For quantitation of bands, fluorograms were scanned using an EPSON 1000C scanner and analyzed using NIH Image software.

For detection of ERp57 by immunoblotting, proteins resolved by SDS-PAGE were transferred to Nitrocellulose (Millipore) (Danilczyk et al., 2000). The membrane was incubated with rabbit anti-ERp57 antiserum at 1:2,500 dilution, followed by a donkey anti-rabbit IgG horseradish peroxidase conjugate at 1:10,000 dilution (Jackson Laboratories). Immune complexes were visualized using an enhanced chemiluminescence system (Amersham Life Science).

4.3 Results

4.3.1 Expression of ERp57 in Drosophila Cells

As described in chapter 2, when the D^b H chain was expressed in Drosophila cells, a folded conformer could be formed as judged by immunoprecipitation with the conformation-specific antibody 28-14-8S. Furthermore, the association of H chains with β_2m was observed in these cells, when both cDNAs were cotransfected. However, the folding and the assembly of these heterodimers was inefficient and the heterodimers were transported rapidly to the cell surface. In these cells the endogenous molecular chaperones, in particular the homologs of CNX and CRT, could not be detected with the co-transfected H chain (Jackson et al., 1994; Vassilakos et al., 1996). Consistent with this observation, co-transfection of mammalian CNX or CRT not only increased the efficiency of folding and heterodimer assembly but these partially assembled molecules were now retained in the ER. Upon co-transfection with CNX or CRT a phenotype was obtained that was similar to TAP-deficient mammalian cells.

Since the Drosophila expression system permits an examination of H chain folding and assembly with β_2m, separately, we selected the Drosophila expression
Figure 24. Expression of ERp57 in Drosophila cells. (A) Drosophila cells previously shown to express D\textsuperscript{b} H chains or K\textsuperscript{b} H chains - β\textsubscript{2}m heterodimers in the presence of CNX were co-transfected with ERp57. This system allows us to monitor H chain folding and assembly with β\textsubscript{2}m in the presence of selected ER resident proteins. However only early events in class I biogenesis may be investigated in these cells, since they are deficient in TAP. (B) Detergent lysates from Drosophila cells expressing D\textsuperscript{b} H chains (D\textsuperscript{b}) or K\textsuperscript{b} H chains - β\textsubscript{2}m heterodimers (K\textsuperscript{b}), CNX (CNX) and ERp57 were analyzed by Western blotting with rabbit anti-ERp57 antibody. HC, H chain.
system to investigate the contribution of ERp57 to the formation of H chain disulfide bonds at each stage. For this purpose, *Drosophila* cells expressing mouse Dp H chains or Kp H-chain-β2m heterodimers with CNX were transfected with the human ERp57 cDNA (Fig. 24A). The presence of CNX was essential since it is required to recruit ERp57 to the class I H chain (High et al., 2000).

Our initial experiments were designed to demonstrate by Western blotting that the co-transfected ERp57 was expressed in the *Drosophila* cells. As shown in Fig. 24B, ERp57 was not present in the untransfected cells, but was easily detected in cells expressing ERp57 only or cells co-expressing H-chain and CNX. These experiments show that stably transfected *Drosophila* cells expressing two or three transgenes can be additionally co-transfected with ERp57 cDNA. However, the expression levels of ERp57 were reduced with 3 transgenes and further reduced when 4 heterologous cDNAs were expressed in these cells.

In these initial experiments, we also investigated whether ERp57 forms stable complexes with CNX and class I molecules in *Drosophila* cells. Originally ERp57 was detected in association with substrate proteins only when cross-linked. However, Hughes and Cresswell reported that, in human cell lines, ERp57 interaction with class I heterodimers can be detected without chemical crosslinking (Hughes and Cresswell 1998). These results were further confirmed by Lindquist et al. (1998). They also showed that free H chain which associates with CNX can be detected in association with ERp57.

To address the question of whether such a stable association exists between free H chains and ERp57 in *Drosophila* cells, cells expressing Dp H chains and CNX or Kp H chain-β2m heterodimers and CNX in the presence or absence of ERp57 were analyzed. The complexes formed between H chains and the co-transfected proteins were immunosolated with antibodies recognizing unfolded H chain (anti-HC) or partially folded H chain (28-14-8S mAb). These proteins were analyzed by non-reducing SDS-PAGE. In these experiments, we specifically isolated free H chains, since the formation
Figure 25. Co-immunoisolation of H chains with CNX in the presence of ERp57. (A) 
*Drosophila* cells expressing CNX and D^b^ H chains (A) or CNX and K^b^ H chains and B^2m^ (B) in the presence or absence of ERp57 were radiolabeled with [35S]Met for 5 min and subsequently incubated in media containing excess unlabeled Met for 10 min. Cells were lysed in 1% digitonin lysis buffer and D^b^ molecules were isolated with antibodies recognizing unfolded (anti-HC) H chains or partially folded (28-14-8S) H chains. K^b^ molecules were isolated with antibodies recognizing unassembled (α-HC) H chain. Proteins were analyzed by non-reducing SDS-PAGE. The mobilities of co-isolated CNX and D^b^ or K^b^ H chains are indicated.
of disulfide bonds is initiated at this stage, whereas \( \beta_2m \) associated H chains have mainly fully formed disulfide bonds. As shown in Fig.25 CNX was detected in association with the unfolded \( \text{D}^b \) or \( \text{K}^b \) H chain (anti-HC) and the partially folded \( \text{D}^b \) H chain (28-14-8S). However, ERp57 was not detected in any of these complexes. These results may indicate that, in \textit{Drosophila} cells, the co-transfected ERp57 does not associate with the free H chain. Alternatively, one might also conclude that the population of ERp57-CNIX-H chains complexes represent a very small fraction of the CNX-H chain complexes and as such may be difficult to detect within the whole population of ER retained H chains. Finally, the interaction between free H chain and ERp57 may not be not strong enough to withstand the process of co-immunoprecipitation. The lack of ERp57 in CNX-H chain complexes, in \textit{Drosophila} cells, would suggest that ERp57 does not exist as a subunit of a discrete complex formed with CNX (ERp57-CNIX).

### 4.3.2 ERp57 Does Not Affect Disulfide Bond Formation in Class I Molecules Expressed in \textit{Drosophila} Cells.

Since the inability to detect ERp57-class I complexes may reflect either a small steady state pool of such complexes or weak interaction between these two molecules, we next investigated ERp57 thiol oxidoreductase function and its involvement in H chain disulfide bond formation. Mature class I molecules have two disulfide bonds present; one in the \( \alpha2 \) domain between Cys residues at positions 101 and 164 and the second in the \( \alpha3 \) domain between Cys residues at positions 203 and 259. Several studies have demonstrated that in mammalian cells lacking \( \beta_2m \) only a fraction of H chains have two disulfide bonds formed (Ribaudo and Margulies, 1992; Tector et al., 1997; Solheim, 1999). To determine whether disulfide bond formation in \textit{Drosophila} cells mimics mammalian cells, \textit{Drosophila} cells expressing \( \text{D}^b \) H chains in the presence of CNX were radiolabeled with \(^{35}\text{S}\)Met and the \( \text{D}^b \) molecules were isolated with antibodies recognizing unfolded H chains or partially folded H chains. The H chains were separated
Figure 26. Unassembled class I H chains in *Drosophila* cells contain 1 or 2 intramolecular disulfide bonds. *Drosophila* cells expressing D$_b^b$ H chain and CNX were radiolabeled with $[^{35}\text{S}]$Met for 5 min and subsequently incubated in media containing excess unlabeled Met for 10 min. Cells were lysed in 1% NP-40 lysis buffer and D$_b^b$ molecules were isolated with antibodies recognizing unfolded (anti-HC) H chains or partially folded (28-14-8S) H chains. Proteins were separated by reducing (+DTT) or non-reducing (-DTT) SDS-PAGE. H chains containing 0, 1 or 2 disulfide bonds are indicated.
by reducing or non-reducing SDS-PAGE. As shown in Fig. 26, upon reduction of disulfide bonds only one band was detected in unfolded or partially folded H chain. This band represents the population of H chains with fully reduced disulfides (0 disulfide bonds) and serves as a mobility marker for reduced H chains. In contrast, under non-reducing conditions two populations of H chains could easily be detected. The slower migrating H chains contained one disulfide bond, whereas the faster migrating H chains contained two disulfide bonds. Interestingly, H chains with two disulfide bond were detected in the unfolded H chains and in the partially folded H chains defined by reactivity with mAb 28-14-8S. These results compare well with the published results for β_{2m} deficient Daudi cells (Tector and Salter, 1995). Disulfide bond formation is initiated in both mammalian cells and Drosophila cells in the absence of β_{2m}, and therefore presumably takes place before β_{2m} binding.

When the same approach was used to analyze the status of disulfide bonds in cells expressing K^b H chain and β_{2m}, we established that H chains containing 1 or 2 disulfide bonds could be detected in Drosophila cells (Fig. 27). However, only H chains with two disulfide bonds were co-isolated with β_{2m} (20-8-4S population, right panel). Furthermore, H chains with two disulfide bonds were already detected in the pulse sample, suggesting that in the presence of β_{2m} the process of disulfide bond formation is very fast. Interestingly, the population of free H chains detected with anti-HC antibody contained predominantly one disulfide bond. These results demonstrate that the Drosophila system mimics mammalian cells and may be used to evaluate formation of disulfide bonds.

To investigate the capacity of ERp57 to influence formation of disulfide bonds within free H chains we carried out similar pulse-chase radiolabeling experiments on cells expressing H chain and CNX in the presence or absence of ERp57. We asked if levels of newly synthesized class I molecules with none, one or two disulfide bonds are altered by the presence of ERp57. Radiolabeled cells were lysed and unfolded H chains
Figure 27. Class I K\textsuperscript{b} H chains-\textbeta\textsubscript{2}m heterodimers in *Drosophila* cells contain 2 intramolecular disulfide bonds. *Drosophila* cells expressing K\textsuperscript{b} H chains, \textbeta\textsubscript{2}m, and CNX were radiolabeled with [\textsuperscript{35}S]Met for 10 min and subsequently incubated in media containing excess unlabeled Met for 20 min. Cells were lysed in 1% digitonin lysis buffer and K\textsuperscript{b} molecules were isolated with antibodies recognizing K\textsuperscript{b} H chains -\textbeta\textsubscript{2}m heterodimers (20-8-4S) or unassembled H chains (anti-HC). Proteins were separated by reducing (+DTT) or non-reducing (- DTT) SDS-PAGE. H chains containing 0, 1 or 2 disulfide bonds are indicated.
were immunoprecipitated with anti-HC antibodies. Upon analysis by non-reducing SDS-PAGE the acquisition of one disulfide bond in class I H chains was assessed. As shown in Fig. 28 the amount of H chain with one disulfide bond when expressed as percentage of total H chain did not change upon co-expression of ERp57. These results demonstrate that in the absence of β2m we are unable to increase the efficiency of disulfide bond formation by co-expression of ERp57. The process of disulfide bond formation in the presence or absence of ERp57 is very fast and maximal formation of disulfide bonds occurred within the pulse time i.e., there was no change in the pattern of disulfide bonds when the pulse and 10 min chase samples were compared. However, this experimental system does not rule out the possibility that in the presence of ERp57 the disulfide bonds are formed more efficiently but the fully oxidized H chains, in the absence of β2m, are so unstable that we could not detect them.

4.3.3 ERp57 Does Not Modulate the Molecular Chaperone Function of CNX

Although we could not detect effects of ERp57 on disulfide bond formation, it is possible that ERp57 can still exert some influence on H chain folding and assembly, beyond that which we demonstrated previously for CNX. Consequently, we examined the effect of ERp57 on the efficiency of H chain folding and assembly with β2m in the presence of CNX. In these experiments we applied the same approach that was previously used to demonstrate the molecular chaperone function of CNX and CRT (Vassilakos et al., 1996; Danilczyk et al., 2000) i.e., the conformation-dependent mAb that is specific for a folded epitope in the α3 domain of the Dβ H chain was used to monitor the kinetics of folding and the β2m specific antibody was used to monitor Kβ H chain assembly with β2m in pulse-chase radiolabeling experiments.

Since we previously showed that CNX enhanced H chain folding by ~2-fold and that 80% of H chains acquired the folded epitope after a 5 min pulse in the presence of CNX we decided to monitor the effect of ERp57 at shorter pulse times. Drosophila cells
Figure 28. Effect of ERp57 on the formation of disulfide bonds in the $\text{D}^b$ H chains. (A)

*Drosophila* cells expressing $\text{D}^b$ H chain and CNX were radiolabeled with $[35\text{S}]\text{Met}$ for 5 min and subsequently incubated in media containing excess unlabeled Met for 10 min. Cells were lysed in 1% digitonin lysis buffer and $\text{D}^b$ molecules were isolated with antibodies recognizing unfolded (anti-HC) H chains. Proteins were separated by non-reducing (- DTT) SDS-PAGE. H chains containing 1 or 2 disulfide bonds are indicated. (B) Weakly exposed fluorograms were quantified by densitometric analysis. The amount of H chains with one disulfide bond (HC-1) was expressed as percentage of total H chains.
Figure 29. Folding of Db H chains in the presence of CNX and ERp57. Drosophila cells expressing CNX and Db H chain in the presence or absence of ERp57 were radiolabeled with $[^35]$S]Met for 2 min and chased for 2 and 4 min. H chains with a folded $\alpha_3$ domain were isolated with mAb 28-14-8S and total Db H chains were isolated with a combination of mAb 28-14-8S plus HC serum. The immune complexes were analyzed by SDS-PAGE and were quantified by densitometry. The amount of 28-14-8S reactive H chains at each time point were expressed as a percentage of the total H chain present in the pulse sample.
expressing D\textsuperscript{b} H chain and CNX in the presence or absence of ERp57 were radiolabeled for 2 min and chased for 2 or 4 min. As depicted in Fig. 29, following pulse-chase radiolabeling, the total or 28-14-8S reactive H chains were immunoisolated and the amount of partially folded H chain was determined as a percentage of total H chain. Similar to results obtained previously (Vassilakos et al., 1996; Danilczyk et al., 2000), ~50% of D\textsuperscript{b} H chains folded into a 28-14-8S-reactive conformation in the presence of CNX at the 0 chase time. However, co-expression of ERp57 did not change the efficiency of H chain folding and no increase was observed over the 4 min chase period. These results demonstrate that ERp57 does not enhance the molecular chaperone function of CNX at early stage of class I biogenesis.

To examine the involvement of ERp57 in a later stage of class I assembly we monitored the effect of ERp57 on the efficiency of H chain assembly with β\textsubscript{2}m. As reported previously, co-expression of CNX with murine class I molecules in Drosophila cells augments the assembly of K\textsuperscript{b} H chains with β\textsubscript{2}m from about 20% to about 50% during a 5 min pulse period (Vassilakos et al., 1996; Danilczyk et al., 2000).

To determine if ERp57 influences the formation of class I heterodimers beyond that observed with CNX, cells expressing K\textsuperscript{b} H chains, β\textsubscript{2}m and CNX in the absence or presence of co-expressed ERp57 were subjected to pulse-chase radiolabeling (5 min pulse and 10 min chase) and the levels of newly synthesized K\textsuperscript{b} H chains were detected using three different antibodies: a rabbit antiserum (anti-8) that recognizes total (assembled and unassembled) K\textsuperscript{b} H chains, a β\textsubscript{2}m dependent mAb (20-8-4S) that only recognizes assembled K\textsuperscript{b}-β\textsubscript{2}m heterodimers, and a rabbit antiserum (anti-HC) that recognizes unassembled K\textsuperscript{b} H chains.

Figure 30 demonstrates that ERp57 did not enhance heterodimer assembly beyond that observed in the presence of CNX alone. Regardless of the presence of ERp57, ~55% of H chains assembled with β\textsubscript{2}m during the pulse, and assembly was nearly completed by
Figure 30. Effects of ERp57 on Kb H chain assembly with β2m. Drosophila cells expressing CNX and Kb H chains and β2m in the absence or presence of ERp57 were incubated with [35S]Met for 5 min and then with excess unlabeled Met for 10 min. Kb molecules were isolated with antibodies recognizing unassembled (anti-HC), β2m-associated (mAb 20-8-4S) or the total population (α8) of H chains. Proteins were analyzed by reducing SDS-PAGE and were quantified by densitometry. The amounts of β2m-associated or unassembled H chains were expressed as a percentage of the total H chains present in the pulse sample.
10 min of chase (Fig. 30). Furthermore, the increase in H chain assembly with β2m was reflected by a decrease in the percentage of unassembled H chains, a drop from 40% to ~20% was observed between 0 and 10 min chase. These results demonstrate that ERp57 does not enhance the ability of CNX to enhance the association of H chain with β2m.

4.4 Discussion

The objective of this study was to assess the involvement of ERp57 in the formation of disulfide bonds in class I H chains and to clarify the role ERp57 plays in class I biogenesis. Our findings indicate that ERp57 does not cooperate with CNX in enhancing H chain folding and assembly with β2m. This was observed when the formation of a partially folded epitope or the association of H chain with β2m was monitored with conformation specific antibodies. Furthermore, we were unable to demonstrate the thiol oxidoreductase function of ERp57 as measured by enhanced formation of disulfide bonds in partially folded H chains in vivo. The levels of newly synthesized class I molecules with one or two disulfide bonds were unaffected by the co-expression of ERp57. Finally, in our experiments, ERp57 and MHC class I molecules do not co-immunoprecipitate by anti-class I antibodies in Drosophila cells even though interactions between class I H chains and CNX are easily detected in this system.

Interestingly, ERp57 was detected in association with class I molecules in other systems. ERp57 was found in complexes with class I molecules and CNX that are generated early in MHC class I assembly in murine EL4 cells (Lindquist et al., 1998). In addition, ERp57 was shown to be a part of the late assembly complex consisting of H chain, tapasin, TAP, CRT and CNX in both EL4 and human 45.1 cell lines (Lindquist et al., 1998; Hughes and Cresswell, 1998). The relative abundance of MHC class I H chains co-precipitating with ERp57 with respect to CNX or CRT favor ERp57-H chain
complexes as the predominant species observed in the anti-ERp57 immunoprecipitates in EL4 cells (Lindquist et al., 1998).

Presently it is difficult to reconcile our studies in Drosophila cells with studies published on mammalian cells. A possible explanation for the absence of the ERp57 in the anti-H chain immunoprecipitates is that these H chains represent the entire pool of class I molecules, while ERp57 interacts with only a small portion of these ER localized H chains. Hence, the ERp57-H chain complex would consist only of a small percentage of the total pool of class I. In part, differences in the ability to immunoprecipitate ERp57 in complexes with class I molecules, observed between Drosophila cells and mammalian cells, may also lie in the ability of class I molecules to interact with additional components in the mammalian cells that stabilize ERp57-H chain-CN X interactions. Hughes and Cresswell have shown that in Daudi cells, which lack β2m, the association between class I, CRT and ERp57 can not be detected. Also, antibodies to CRT fail to co-precipitate ERp57 in these cells. Likewise, in .220-B8 cells which lack tapasin, the CRT-ERp57-class I interaction is undetectable. In .174 cells, which lack TAP but can form a subcomplex containing H chain-β2m dimers, tapasin, CRT and ERp57 the interaction is easily seen. The probable implication of these observations is that CRT and ERp57 bind independently to the H chain-β2m-tapasin complex and to each other with relatively low affinity and that the interactions act cooperatively to stabilize the complex (Hughes and Cresswell, 1998).

There are however, several studies that contradict this conclusion. Most notable, Lindquist et al., demonstrated that the association of ERp57 with H chain was independent of β2m in the β2m negative mouse line S3. In this cell line, ERp57-CN X-H chain complexes are shown to bind to TAP. Also, the human tapasin-negative cell line 721.220 was used to demonstrate the association of ERp57 with class I H chains. Interestingly, in this study, the detection of class I-ERp57 was possible when anti-H chain immunoprecipitations were performed from digitonin extracts. Tapasin negative
cells do not express MHC class I on the cell surface and the entire pool of class I molecules reside within the ER (Lindquist et al., 1998).

The *Drosophila* expression system resembles the tapasin and TAP deficient cells, since they lack genes encoding these proteins, and as such association between class I-ERp57 should reflect these mammalian mutants. However, our studies on murine class I further complicate the issue of class I association with ERp57. Even though we used in our experiments the murine class I molecule, our result supports the observation for human H chains, i.e., ERp57 does not form stable complex with class I in the absence of B_{2}m or tapasin.

Although the role of B_{2}m and tapasin in the stabilization of class I-ERp57 complexes can be disputed, it is clear that ERp57 is found associated with class I molecules in mammalian cells. This raises the question: what role does ERp57 play in class I biogenesis? Recently a semipermeabilized cell system was established that reproduces the folding and assembly of a class I complex as it would occur in intact cells. By using this system the authors demonstrated that the thiol-dependent oxidoreductase ERp57 initially interacts with non-disulfide bonded H chains. Furthermore, these H chains become rapidly disulfide-bonded during their interaction with ERp57. Consequently, they concluded that H chains folding occurs during its interaction with ERp57 (Farnery et al., 2000).

Although recruitment of a thiol-dependent oxidoreductase to unassembled H chains would suggest that at this stage ERp57 functions as a thiol oxidoreductase it is clear that the function of ERp57 cannot be concluded based only on the fact that ERp57 associates with both the native and oxidized H chain. However, the function of ERp57 as a thiol oxidoreductase was demonstrated in vitro (Zapun et al., 1998). The authors found that the thiol oxidoreductase activity of ERp57 on the refolding of monoglucosylated ribonuclease B is greatly enhanced when this glycoprotein was associated with CNX or CRT. In this system CNX does not modulate glycoprotein
folding and serves probably to recruit the ERp57 in the proximity of the substrate protein. It is the combination of ERp57 association with CNX that specifically modulates glycoprotein folding, in this case by promoting the formation of native disulfide bonds. This is in contrast to our findings. First, in the Drosophila expression system CNX promotes efficient H chain folding and assembly with β₂m during class I biogenesis. Secondly, we did not see any increase in disulfide bond formation when ERp57 and CNX were coexpressed in Drosophila cells compared to CNX alone.

Based on our results ERp57 is not involved in the formation of disulfide bonds in free H chains. Hence, the role of ERp57 as the thiol oxidoreductase at the early stages of class I biogenesis has yet to be demonstrated. Furthermore, ERp57 does not enhance the folding and assembly of class I molecules in the presence of CNX. Since we could not detect the presence of ERp57 in complexes with CNX and free H chain, these results suggest that ERp57 probably is not involved at this stage of class I biogenesis. However, the observed lack of ERp57 functions, in Drosophila cells, may also be due to the presence of endogenous thiol-oxidoreductases, that work together with the co-transfected CNX.

An interesting aspect of the ERp57 interaction with class I molecules is the fact that, in mammalian cells, ERp57 is mainly detected after β₂m association during the assembly of H chain-β₂m-CNX/CRT-tapasin-TAP complexes (Lindquist et al., 1998; Hughes and Cresswell, 1998). Since at this stage the H chains are exclusively disulfide bonded, it raises a question as to what role CNX/CRT and ERp57 play in the class I-TAP complex. The prolonged interaction of ERp57 with H chain-β₂m heterodimers may indicate that one of the disulfide bonds in the H chain is subject to repeated reduction/oxidation reactions (Hirano et al., 1995). It is hypothesized that the bond which connects the α₂ α-helix to the floor of the binding groove (Cys 101-Cyt 164) is reversibly reduced and that this process provides a mechanism for regulating the affinity of peptide binding. If associated peptide binds with high affinity it would be difficult to
reduce the bond because of the strong non-covalent interactions within the binding groove, whereas a low affinity peptide might readily permit reduction. If ERp57 plays a role in rearranging disulfide bonds before or just after peptide binding the Drosophila expression system may be used to investigate this possibility especially since cells cotransfected with TAP and tapasin are already available.

Alternatively, a role for ERp57 has been suggested in the degradation of misfolded proteins (Otsu et al., 1995). It is possible that the population of H chains that do not assemble with peptide remain associated with ERp57 to facilitate reduction prior to unfolding and dislocation from the lumen of the ER for proteasome-mediated degradation. Since the \( \beta_m \) associated H chains are fully oxidized it is possible that their degradation may require an interaction with ERp57.

Finally, ERp57 has also been suggested to possess proteolytic activity (Urade and Kito, 1992). Purified ERp57 was capable of degrading CRT and PDI \textit{in vitro} and degradation was inhibited by cysteine protease inhibitors (Urade et al., 1997). Furthermore, a cysteine protease activity was mapped to C-terminal cysteines in the CGHC motifs. Such motifs are present in the a and a' domains of all members of the thioredoxin family. The presence of ERp57 in association with CNX/CRT-class I-tapasin-TAP complexes raises the possibility that ERp57 is involved in the trimming of peptides that bind to MHC class I but are longer than the required 8-10 amino acids.

In summary, we have developed an expression system that has allowed us to evaluate the interactions of CNX and ERp57 with unassembled and partially assembled H chains. We could not demonstrate a role for ERp57 in the early steps of class I molecules biogenesis. However, the ability to add any new protein to this system such as tapasin and TAP will give us the opportunity to determine whether ERp57 is involved in the late stages of MHC class I molecules biogenesis. Using this approach, we can also address the question of ERp57 functions in the trimming of peptides or degradation.
The greatest derangement of the mind is to believe in something because one wishes it to be so.

-Louis Pasteur-

False facts are highly injurious to the progress of science, for they often endure long.

-Charles Darwin-
5.1 Summary of Results

This thesis describes the involvement of the molecular chaperones CNX and CRT in class I biogenesis. It also addresses the broader issue of what regulates the binding specificity of CNX and CRT for diverse glycoproteins. Furthermore, since CNX and CRT function both as lectins and as classical molecular chaperones *in vitro* the question of whether polypeptide-based interactions between CNX or CRT and their diverse substrate exist *in vivo* was examined. Finally, the role of ERp57 in class I biogenesis was investigated. From these studies the following conclusions can be drawn:

CRT, the soluble homologue of CNX, can largely replace CNX in its quality control and molecular chaperone functions. CRT retains peptide-deficient class I heterodimers in the ER (quality control), promotes assembly of H chain with β2m, facilitates folding of class I H chains and stabilizes folded class I intermediates.

The transmembrane and cytoplasmic domains of CNX are not required for CNX's quality control and molecular chaperone functions. The membrane disposition of CNX, however, influences the spectrum of proteins with which it interacts. When CNX was expressed as a soluble protein, the pattern of associated glycoproteins changed to resemble that of CRT. Conversely, when CRT was engineered to be membrane anchored, it bound to a similar set of glycoproteins as CNX.

CNX forms stable, discrete complexes with substrate proteins in the absence of glucose trimming *in vivo*. Under these conditions CNX associates with HLA-B27, H-2 K\(^{b}\) H chain, H-2D\(^{b}\) H chain and invariant chain. Furthermore, unglycosylated H-2K\(^{b}\) H chain associates with CNX in the absence of any detectable aggregation, suggesting that CNX forms peptide-based interactions with its substrates *in vivo*.

The levels of fully reduced, intermediate (one disulfide) and fully oxidized class I forms are not altered by the presence of ERp57. Furthermore, co-expression of
mammalian ERp57 does not affect the folding and assembly of class I molecules. This is an important finding that challenges the current view of ERp57 function in class I biogenesis.

In the discussion that follows, I will elaborate upon my findings from a perspective of protein folding and protein traffic within the cell. In particular, I will detail the mechanisms involved in class I folding and assembly with β2m and its relevance to the basic issues of protein folding in the ER. The class I dominant molecular chaperones, CNX and CRT, are unique among molecular chaperones due to their lectin properties. However, carbohydrate-protein interactions are viewed increasingly as important mechanisms for biological information transfer between different subcellular compartments. From this perspective I will discuss the involvement of lectins in the class I secretory pathway and speculate on their function in class I egress from the ER.

The discovery of intracellular lectins offers a possible explanation of why the glycosylation machinery, especially that of N-glycosylation, is so remarkably complex. Within this framework, I will speculate whether there are some processes early in class I translocation that may involve the N-linked glycan intermediate Glc₂Man₉GlcNAc₂.

Since both CNX and CRT were shown to function as lectins and molecular chaperones in class I biogenesis the question arises as to their mechanism of action. As my findings suggest that the polypeptide component plays a role in CNX’s interaction with substrate proteins, the current models for CNX/CRT functions will be discussed with specific emphasis on peptide-peptide interactions.

5.2 Molecular Chaperones in Class I Biogenesis

The experiments described in chapter 2 that established the role of CRT in class I biogenesis are part of the growing body of evidence regarding the involvement of molecular chaperones and folding catalysts in class I biogenesis. The discovery of ERp57 in association with H chain introduced a new thiol oxidoreductase into the multisubunit functional complex facilitating class I folding and assembly (Hughes and Cresswell, 1998;
Lindquist et al., 1998). As many features of the mechanisms involved in molecular chaperone function and substrate specificity have been elucidated in recent years a more detailed picture of class I biogenesis has become apparent. In the model that follows, I will incorporate the new concepts regarding protein folding in eukaryotes and I will attempt to include all of the players so far indicated to be involved in class I biogenesis. I will also speculate on the involvement of some molecular chaperones which, based on their substrate specificity and class I structural data, are good candidates for being class I molecular chaperones.

According to Hartl, protein folding in the cell differs from folding observed in vitro. Contrary to the in vitro denaturation experiments where a full length polypeptide folds into its final structure, nascent polypeptide chains in cells emerge from the ribosome-translocon system in a linear fashion (Hartl, 1996). Consequently, a single-domain protein can only fold post-translationally after its release from the ribosome. However, polypeptides consisting of two or more consecutive domains may fold co-translationally and thus avoid incorrect interactions between domains during folding (Netzer and Hartl, 1997). This observation implies that each of the \( \alpha_1, \alpha_2 \) and \( \alpha_3 \) domain of class I H chain may fold independently, as it leaves the translocon and in the discussion that follows I will describe the requirements for molecular chaperones and other folding factors within this framework.

In the ‘individual domain’ folding model, as class I H chain emerges from the translocon, it will fold into a mature conformation in a multistep process. It is hypothesized that the folding of each domain is initiated once the whole newly synthesized domain enters the ER lumen, and such folding is independent of the presence of other domains. Since each domain is different in structure and/or in its post-translational modifications, such as glycosylation and disulfide bond formation, the requirements for molecular chaperones and folding catalysts may differ for each domain. This model suggests that protein folding may be assisted by a large number of molecular chaperones, each used at different stages of
protein folding. It also suggests that more than one molecular chaperone may associate with a given protein at the same time.

Class I biogenesis starts with the H chain being targeted to the ER by an N-terminal signal sequence. As soon as this 23 amino acid signal sequence protrudes from the large subunit of a ribosome it is bound by the Signal Recognition Particle. As the amino acids of the α1 domain enter the ER they are constrained by the N-terminal 23 amino acid long signal sequence that targets the ribosome to the ER membrane and the translocon. Following the signal sequence, the amino acids form four strands of β pleated sheets, strand 1: between amino acids (aa) 2-14, strand 2; aa 20-27 strand 3; aa 32-37, and strand 4; aa 44-48 that are followed by two α helical structures formed between aa 50-55 and aa 58-84. Since the formation of the β sheet is severely delayed by the translocation process and each peptide that assumes an extended strand conformation must wait until the other strands are synthesized to form the hydrogen bonds, the extended β strand peptides are potential substrates for the peptide-binding chaperones such as BiP. As the fiftieth amino acid of the α1 domain enters the ER, and the following amino acids of α1 domain (50-84) form α helices, the first domain of the H chain may fold spontaneously without the additional help of chaperones. Finally, the end of the α1 domain is marked by carbohydrate, as the N-linked glycan is added to Asn 86 during translocation. Following the removal of the last two glucose from the Glc3Man9GlcNAc2 oligosaccharide the H chain is ready for a stable association with the membrane bound CNX (Ware et al., 1995). Similarly to α1, the α2 domain contains four β pleated sheets (strand 1: aa 92-104, strand 2: aa 112-118, strand 3: aa 124-127, strand 4: aa 132-136) followed by four α helices (aa 138-148, aa 151-161, aa 163-173 and aa 175-180). Additionally, this domain contains two Cys residues at position 101 and 164 that are involved in the formation of a disulfide bond. This disulfide bridge brings the first β strand into proximity of the third α helix. Furthermore, in the α2 domain of the murine H chain, the N-linked glycan is added to the Asn at position 176 (6 amino acids from the end of domain 2) and this glycan
becomes a potential substrate for CNX. It is important to note that only murine H chains have a glycan at position 176 and this glycan is missing from all other mammalian class I H chains examined so far (Lian et al., 1998).

The structure of the immunoglobulin-like \( \alpha \) 3 domain with seven \( \beta \)-pleated strands connected with a disulfide bridge formed between strand 2 (Cys 203) and strand 6 (Cys 259) appears also to be a strong candidate for association with molecular chaperones. Since only some murine H chains have N-linked glycan added at position 256, chaperones other than CNX may be involved in facilitating the folding of this particular domain.

The strategic location of the N-glycans in the murine system at the end of each domain suggest that these positions may be crucial for chaperone association. In \( \alpha \) 1 and \( \alpha \) 2 domains the glycans are located at the end of an \( \alpha \) helical structure (four and six amino acids from the end of the domain, respectively) preceding stretch of amino acids involved in the formation of \( \beta \) pleated sheets, and in the \( \alpha \) 3 domain at the end of \( \beta \) pleated sheets, four amino acids in front of the connective peptide. Indeed, both murine and human free H chain were detected in association with the glycan-specific chaperone CNX.

The glycan dependance of CNX association with class I molecules was demonstrated in experiments described in chapter 3, in cells such as EL-4. Furthermore, since CNX was shown to associate with a glycan at each location; 86, 176 and 256 (Harris et al., 1998; Zhang and Salter, 1998), CNX may function as the dominant chaperone for murine class I molecules.

This observation, however, raises a question of which chaperone replaces CNX at position 176 and 256 in the process of human H chain folding. Several lines of evidence indicate that the classical molecular chaperone from the Hsp70 family, BiP, is a good candidate to facilitate the folding of the human H chain. The ER chaperone BiP is believed to interact with proteins during their translocation (Corsi and Schekman, 1997; McClellan et al., 1998). Also, it was established that BiP preferentially associates with \( \beta \) pleated stretches of newly synthesized proteins (Blond-Elguindi et al., 1993) and such structures
are present in all three domains of the H chain. Finally, BiP has been shown to bind to free human H chains but not murine H chains (Parham, 1996). The differential association of murine and human free H chains with CNX and BiP implies that these two molecular chaperones may compete for substrates at the translocon with CNX binding preferentially to glycosylated domains. It would also suggest that these two molecular chaperones are functionally interchangeable and further support the notion that CNX indeed functions as a classical molecular chaperone.

The lack of BiP association with murine H chain has important implications regarding the folding of the α1 domain. Since CNX as a lectin will associate with the H chain only after the entire α1 domain is synthesized (position 86), what protects the newly synthesized chain from aggregation at this early stage of synthesis remains unclear. One possibility may stem from the fact that α1 domain is constrained by the signal sequence and the folding may be delayed until the whole domain is synthesized. However, it is also conceivable that the use of more sensitive techniques on synthesis arrested H chains may allow the detection of BiP in association with the α1 domain. Finally, we showed in Chapter 3 for class I molecules expressed in Drosophila cells and human H chain expressed in C1R cells that CNX association with a substrate protein may not always be glycan dependent and glycans only contribute to CNX substrate selectivity. Thus, CNX may also work as molecular chaperone at early stages of α1 domain folding, before the addition of the glycan at residue 86.

Studies in the mouse indicate that CNX remains associated with H chains after β2m association (Degen et al., 1992; Degen and Williams, 1991). In contrast, in human cell lines, the association of β2m with the H chain triggers the release of CNX, and now CRT becomes the dominant molecular chaperone (Sadasivan et al., 1996). Since studies in Drosophila cells showed that CRT and CNX are largely interchangeable (Chapter 2) and CNX and CRT share their substrate specificity for monoglucosylated N-linked glycans,
this observation raises an important question of what accounts for the preferential use of one molecular chaperone versus the other in the mouse and human systems?

It is likely that substrate selection is a complex process regulated by at least three independent elements: 1) the membrane dispositions of CNX versus CRT; CRT substrate specificity is dramatically changed when expressed as membrane bound molecule, suggesting that altered topology results in altered substrate recognition (Chapter 2), 2) the different potential polypeptide binding specificities of CNX and CRT; experiments with class I point mutants indicate that CNX and CRT are distinct in their binding sites on class I, mutation at position 134 in H-2Ld H chain abrogates the association with CRT but not CNX (Carreno et al., 1995; Harris et al., 1998), and 3) the location of N-linked glycans on glycoprotein substrates; human H chains with a deleted N-glycan at position 86 do not associate with CRT as newly synthesized proteins (Zhang and Salter, 1998).

Our current model, based on CNX/CRT 'dual binding specificity', incorporates all of these elements contributing to substrate specificity and suggest that these elements may play a key role in the differences observed in the selection of molecular chaperones between human and murine H chains. In this model the glycan functions as a guide for CNX and CRT binding. The association with a glycan brings the H chain in proximity of CNX or CRT so that the peptide-peptide interaction may take place. Glycans also contribute to the avidity of the interaction between CNX/CRT and H chains. However, the presence of a glycan is not a prerequisite for the interaction to take place and only gives CNX/CRT a functional advantage over other classical molecular chaperones. According to this model, membrane bound CNX dissociates from the α1 domain, the N-linked glycan and the potential polypeptide interaction site, upon a CNX conformational change triggered by an as yet undefined factor. Upon folding, the luminal exposure of the polypeptide binding site and the glycan at position 86 in the α1 domain increases, the accessibility of this domain to membrane bound CNX becomes limited. In the murine system, however, the oligosaccharide at position 176 and/or 256 now become accessible to CNX and CNX
remains associated with these H chains. In contrast, in both human and murine cell lines, the glycan at position 86 and the α1 domain now become accessible to the ER luminal CRT.

Several conflicting studies can be reconciled by using this model. CRT-H chain-β2m complexes are easily detected in human cells (Sadasivan et al., 1996), but in murine cells, they are detected only in thymocytes and only for particular alleles of the L'd gene (Van Leeuven and Kearse, 1996; Solheim et al., 1996; Yu et al., 1999). We propose that, for the complexes to be detected by co-immunoprecipitation, strong association between CRT and H chain is necessary and the complexes are detected only when both sites, the lectin and the polypeptide binding sites, participate in the binding. However, the presence of a glycan at position 176 in murine molecules gives CNX the advantage to associate first with the substrate by both its lectin and the polypeptide sites. CNX's polypeptide interaction site is thought to prevent the peptide-peptide binding of CRT. Furthermore, depending on the allele, the peptide-peptide interference may vary, leading to variations in association between CRT and the H chain. This view is supported by the observation that introduction of the second glycan on the human H chain, which allows CNX to bind to the H chain at position 176, weakens CRT association (Zhang and Salter, 1998).

From my discussion it is apparent that CNX functions as a molecular chaperone in murine class I biogenesis. However, there is no indication so far whether CRT plays similar role. In chapter 2, we demonstrated that CRT can replace CNX in its molecular chaperone functions when expressed in Drosophila cells. However, very little is known about its function in vivo in murine cells. Since murine H chains-β2m heterodimers remain associated with CNX this raises a question concerning the functional consequence of their additional association with CRT, if any. Are murine class I molecules associated simultaneously with both CNX and CRT or are there two subpopulations of murine class I molecules, one associated predominantly with CNX, and the second with CRT? Are CNX and CRT redundant in their functions in vivo? Experiments with CNX or CRT deficient
cell lines, where one can evaluate the association of murine class I molecules with CRT (in a CNX deficient cell line) and human class I heterodimers with CNX (in a CRT deficient cell line) may address some of these issues.

The presence of two disulfide bonds in class I H chain make this molecule a perfect substrate for thiol oxidoreductases. Indeed, association of CNX and CRT with H chains cannot be kinetically separated from their association with the thiol oxidoreductase, ERp57. Since in chapter 4, we established that ERp57 is not involved in early class I biogenesis, the question remains, what catalyzes the formation of disulfide bond in H chains? Thein vitro system described by Zapun et al. (1998), with class I as substrate, in which the formation of disulfide bonds is monitored in the presence of other thiol oxidoreductases such as PDI, may clarify this issue. Since ERp57 was detected with CNX and CRT in the peptide loading complex it raises a question whether this molecule has a direct role in peptide binding? Is ERp57 involved in the opening of the peptide binding groove by interacting with the H chain disulfide bond and as such regulating the binding of peptides?.

Drosophila cells co-expressing tapasin and /or TAP along with ERp57 would be ideal to address these issues.

As demonstrated in this discussion, MHC class I molecules provide a particularly illustrative example of the complexity of the quality control process in the ER. The H chains undergo stepwise binding to the general folding factors CNX and CRT. Two class I-specific components are also involved: TAP as a peptide translocator in the ER membrane and tapasin, the bridging molecule with a possible peptide editor function. Together, these factors bind the newly synthesized protein ensuring that it is correctly folded and assembled and that it does not leave the ER without a correct peptide in its peptide-binding groove.
5.3 Transport of MHC Class I Molecules from the ER to the Cell Surface

Once fully assembled, class I molecules are no longer retained by elements of the primary quality control machinery e.g., CNX and CRT. They can now enter exit sites and leave the ER for the Golgi complex and the cell surface. The exact process by which export of these proteins from the ER occurs is not clear. MHC class I molecules have been generally known to arrive rapidly at the cell surface by a default pathway without requirements for specific signals (Jackson et al., 1994). However, recent evidence of sorting of MHC class I molecules in the trans-Golgi network (TGN) suggest that the post ER traffic may be a controlled process (Joyce, 1997). It is postulated that oligosaccharide can give the glycoprotein a positive migration signal. A few proteins have been identified that can act as intracellular lectins at this stage in the biosynthetic pathway. These include the ERGIC53 protein, which is homologous to plant lectins (Appenzeller et al., 1999).

Recent evidence indicates that the mannose selective lectin, ERGIC 53 serves as a cargo transport receptor for some glycoproteins, namely the cathepsin-Z-related glycoprotein (Appenzeller et al., 1999). Transport receptors are believed to concentrate selected proteins as cargo before Golgi arrival and by doing so accelerate transport from the ER. ERGIC53 recognizes glycoproteins that are carrying high-mannose N-glycans and are therefore no longer substrates for UGGT. Binding of the substrate is carbohydrate and calcium-ion dependent and is affected by untrimmed glucose residues. ERGIC 53 may represent, after CNX and CRT, an additional check point in the lectin-based ER quality control.

The next step in the secretory pathway is the Golgi sorting station with its putative sorting lectin VIP36. VIP36 is a type I integral membrane glycoprotein homologous to ERGIC 53. VIP36 shows specificity for high mannose type glycans of the Man₆,GlcNAc₂ structure but efficient binding requires the additional presence of an asparagine residue. The function of VIP36 is currently unknown. However, based on its location in cis-Golgi
and the ER-Golgi intermediate compartment (ERGIC) it is tempting to speculate that it may function in ‘quality control’ outside of the ER. It is possible that VIP36 recognizes glycoproteins that have escaped glycan trimming by cis-Golgi mannosidase I to the Man₅GlcNAc₂ structure and recycles them back for an additional round of trimming.

Finally, studies with polarized epithelial cells implicated lectins in glycoprotein transport from the Golgi to the cell surface. Introduction of N-glycosylation sites into non-glycosylated soluble or membrane proteins can target them to the apical cell surface indicating that N-glycans may operate as apical targeting signals (Gut et al., 1998; Scheiffele et al., 1995). It is interesting to note that N-glycans also mediate Golgi exit in non-polarized cells (Gut et al., 1998). It is conceivable that membrane lectins sort glycoproteins into transport vesicles in the trans-Golgi network, but such lectins remain to be discovered.

Whether any of these new lectins, implicated to play an important role in glycoprotein traffic, are involved in the secretory pathway of class I molecules remains to be seen. If these lectins, however, are not detected in association with class I complexes, what other molecules are involved in class I traffic from ER to Golgi? Are the putative class I transporters all lectins? Answers to these questions could provide valuable insights into the role lectins play in the transfer of biological information inside the cells and would renew interest in biology of lectins.

5.4 Degradation of MHC class I Molecules

For the most part, only fully assembled class I molecules leave the ER. Misfolded or unassembled class I molecules are selected for degradation. ER-associated degradation is mainly carried out by the 26S proteasome located in the cytosol. The process of degradation occurs in several steps. Terminally misfolded or unassembled proteins are recognized by ER chaperones such as CNX, BiP or by other factors such as specific mannose lectins. They are then retrotranslocated through the Sec 61 channel into the
cytosol, deglycosylated and polyubiquitinated before proteasomal degradation (Wiertz et al., 1996; Bebok et al., 1998). Because mannosidase inhibitors block degradation of some misfolded glycoproteins, it has been suggested that the removal of mannose units by slow-acting ER-resident α1,2-mannosidases might work as a timer for glycoprotein degradation (Ellgaard et al., 1999). More specifically, a Man₄GlcNAc₂ glycan with a mannose missing in the middle branch, generated by mannosidase I, was recently identified as a necessary signal for the degradation of misfolded carboxypeptidase Y in yeast (Ellgaard et al., 1999). Furthermore, proteasomal degradation of misfolded human α₁-antitrypsin requires a physical interaction with CNX (Liu et al., 1999) as well as mannose trimming by ER mannosidase I. UGGT-mediated reglucosylation, leading to monoglucosylated Glc₁Man₈GlcNAc₂ delays dissociation of misfolded α₁-antitrypsin from CNX and consequently slows down degradation. Whether CNX in conjunction with mannosidase I selects misfolded class I molecules for degradation by the proteasome is worthwhile to explore. If so, it would imply that the central role of UGGT as a folding sensor is important not only in the CNX/CRT cycle but also for targeting of misfolded proteins for degradation. Moreover, it would further support a role for CNX in the targeting of glycoprotein substrates for retro-translocation to the cytosol.

5.5 Quality Control Based on Glycan Trimming

The discovery of intracellular lectins offers a possible explanation why the glycosylation machinery, especially that of N-glycosylation is so remarkably complex. Modification of N-glycans may function as a key determinant of protein state and/or may guide the glycoprotein through the secretory pathway by interactions with lectins with different specificity. The early attachment of core N-linked structures to glycoproteins suggests that this part of the sugar functions early in the life of the glycoconjugate. This argument is particularly appealing in the case of the glucose component because these residues are transient and are lost once the glycoprotein reaches the Golgi compartment.
The transfer of the preformed oligosaccharide to the nascent polypeptide by oligosaccharyl transferase clearly represents a central event in the reaction in which glucose plays a role. It has been well documented that this transmembrane heterooligomeric enzyme requires the complete Glc$_2$Man$_9$GlcNAc$_2$ sequence for effective transfer. The processing of Glc$_2$Man$_9$GlcNAc$_2$ is initiated by an ER-situated integral membrane enzyme with a lumenally oriented catalytic domain, glucosidase I, that specifically cleaves the α 1-2 linked glucose residue (Fig. 5). This is followed by the action of glucosidase II, which releases both of the α 1-3 linked glucose components. As described in Chapter 1, the most internal of the three glucose residues, after being brought to a terminal position through the action of ER-situated glucosidases, interacts with the lectin-like chaperones, CNX and CRT.

The function of Glc$_2$Man$_9$GlcNAc$_2$ is well documented by its interaction with the molecular chaperones CNX and CRT. What remains to be clarified is the fact that three glucose are necessary for the transfer of N-linked glycan, however, the first functional role is described for the monoglucosylated species. It raises an important question whether there is a role for the intermediate formed after removal of the first glucose by the membrane bound glucosidase I, namely Glc$_2$Man$_9$GlcNAc$_2$. It is worthwhile to explore the possibility that CNX may play a role at this point. CNX, similarly to glucosidase II, may have the ability to interact with both forms, Glc$_2$Man$_9$GlcNAc$_2$ and Glc$_2$Man$_9$GlcNAc$_2$. Since immobilized CNX was not detected in the interaction with Glc$_2$Man$_9$GlcNAc$_2$ when challenged with a mixture of tri-, di-, mono and unlucosylated glycans, one may speculate that it has much lower affinity for the Glc$_2$Man$_9$GlcNAc$_2$ than Glu$_1$Man$_9$GlcNAc$_2$. However, such a double specificity may be considered, since either glucose when present in the mono- or di-glucosylated species can represent the terminal glucose with an α1-3 linkage and interestingly both forms serve as substrates for glucosidase II.

CNX being a membrane bound protein may form a functional complex at this stage with the oligosaccharyl transferase and the glucosidase I, both also membrane bound, and may play a role in the active transport through the translocon. CNX or another as yet
unknown lectin with specificity for Glc$_2$Man$_9$GlcNAc$_2$ may be involved in the process of translocation, helping to pull the newly synthesized polypeptide from the translocon. It could represent an additional step in quality control of glycoproteins being translocated into the ER. The lectins together with the glucosidases may regulate the movement of the proteins. For the unglycosylated proteins that are translocated into the ER, BiP is indicated in the translocation process. It is possible that these two molecular chaperones, BiP and CNX divide their tasks with CNX binding preferentially to glycoproteins. This view is supported by the recent discovery that the presence of a glycosylation consensus sequence close to the N-terminal of a protein determines the choice of chaperones with which the protein interacts (Molinari and Helenius, 2000). For example, the presence of a glycosylation consensus sequence close to the N-terminus of a protein determines that it will interact co-translationally with CNX/CRT instead of binding BiP. The class I molecule may serve as a model to investigate the function of CNX, if any, at the Glc$_2$Man$_9$GlcNAc$_2$ stage. It may also be used to address the question whether CNX or BiP is involved in the translocation process of class I molecules. These experiments may be addressed in a cell free in vitro translation system arrested at different stages of protein synthesis with microsomes as a source of ER.

It is evident from this discussion that lectins together with glucose residues attached to Man$_9$GlcNAc$_2$ oligosaccharides play an important role in quality control of glycoproteins. These sugars which are only transiently present on N-linked carbohydrates provide a revealing example of the importance of carbohydrates and the cooperation between lectins and an elaborate enzymatic machinery that has evolved to make the quality control process possible.

5.6 CNX/CRT Dual Binding Model

Since it has become clear that polypeptide mediated interactions described in chapter 3 are involved in CNX/CRT functions, the view that CNX and CRT bind only to the
carbohydrate moiety of a glycoprotein substrate has now largely been superseded. The emerging picture of CNX/CRT function is of a chaperone that recognizes the carbohydrate moiety of a glycoprotein substrate, probably as an initial recognition event, and as the central feature it maintains close association with the polypeptide chain. According to this ‘dual binding’ model CNX and CRT are multifunctional proteins with carbohydrate and polypeptide binding sites.

Carbohydrate binding contributes to the overall avidity of CNX/CRT binding and to their substrate specificity/selectivity for monoglucosylated substrates. Polypeptide binding contributes to their specificity for unfolded polypeptides. In the model depicted in Fig. 9, the release of the substrate from the molecular chaperone occurs through a conformational change in the polypeptide binding site followed by the action of glucosidase II to remove the terminal glucose residue. CNX/CRT can rebind upon readdition of the glucose residue by UGGT an enzyme that has specificity for nonnative glycoproteins (Sousa and Parodi, 1995). Cycles of binding and release continue until the glycoprotein acquires a conformation no longer recognized by UGGT and CNX/CRT. Hence, both carbohydrate and polypeptide binding promote CNX/CRT molecular chaperone function. How the carbohydrate regulates CNX/CRT functions is now well defined and the involvement of glucosidase II and UGGT in this process is well documented. However, the exact nature and importance of polypeptide mediated interactions during glycoprotein folding and quality control remain controversial.

The crucial question to ask at this point is: do CNX and CRT directly influence protein folding? CNX/CRT molecular chaperone functions began to emerge following the landmark demonstration in vitro that both molecules prevent the aggregation of unfolded proteins when added in stoichiometric amounts, protect a protein against thermal denaturation and enhance the efficiency of protein folding by maintaining an unfolded substrate in a folding-competent state. Hence, CNX and CRT exhibit many of the functional properties of classical molecular chaperones such as Hsp90, Hsp60 and Hsp70
(Ihara et al., 1999; Saito et al., 1999). These results were obtained with purified components by using a soluble form of CNX or full length CRT along with nonglycosylated proteins as model substrates. Furthermore, CNX and CRT interact directly with the polypeptide portion of substrate proteins. Stable complexes between CNX or CRT and unglycosylated proteins were recovered by size exclusion chromatography (Ihara et al., 1999). Evidence for a direct role of CNX/CRT in protein folding is further supported by the observation that CNX and CRT can discriminate between unfolded and native conformational states of a substrate protein. They formed complexes with MDH that was unfolded as evidenced by hypersensitivity to trypsin. However, CNX/CRT failed to associate with native, enzymatically active MDH (Ihara et al., 1999; Saito et al., 1999).

There are, however, several studies which contradict this conclusion. Most notably, Rodan et al. have demonstrated that CNX and CRT do not distinguish between folded and unfolded polypeptide (Rodan et al., 1996). In this experiment a mutant form of ribonuclease missing essential disulfide bonds (mutated Cys residue) that is more trypsin sensitive than the wild type ribonuclease was engineered. There was no difference in interaction between wild type (folded) or mutant (unfolded) ribonuclease and CNX. CNX/CRT was demonstrated to bind to ribonulease B as long as it was monoglucosylated. Additionally, experiments with RNase B demonstrated that CNX and CRT do not influence glycoprotein folding directly (Zapun et al., 1998). In this study reduced/unfolded RNase was incubated in a redox buffer in the presence or absence of CNX. Native gel electrophoresis demonstrated that there was no enhancement in the yield of correctly folded RNase as judged by the formation of disulfide bonds in the presence of CNX.

At present it is difficult to reconcile these results with the studies which demonstrate a peptide based molecular chaperone function of CNX/CRT. One possible explanation is that RNase B does not aggregate, when unfolded by reducing conditions, since it lacks a typical hydrophobic core. As such it is unlikely that RNase B needs to bind to molecular chaperones. However, as discussed below, the key factor in understanding the
polypeptide interaction site may lie in the ability of CNX/CRT to interact with additional components in the ER.

The crucial question to ask at this point is: do CNX and CRT recruit other ER components that have an effect on protein folding. To date, only ERp57 has been identified as contributing to the molecular chaperone function of CNX/CRT. It is hypothesized that ERp57 functions as a subunit of discrete complexes formed with CRT and CNX. Several lines of evidence support this view. First, using a variety of approaches ERp57 was shown to form distinct complexes with both CNX and CRT, both within the lumen of the ER and when the proteins are mixed in solution (Mike Leach et al., 2001; Corbert et al., 1999; Oliver et al., 1999). Second, the thiol oxidoreductase activity of ERp57 was demonstrated to be greatly enhanced by the presence of CRT or CNX when the refolding of RNase B was analyzed. Third, mixed disulfide species formed in vivo between ERp57 and a specific viral glycoprotein substrate was identified (Molinari and Helenius, 1999). These mixed disulfides represent transient intermediates and they were formed with ERp57 only when the substrates were both N-glycosylated and have glucose trimmed by the combined action of glucosidase I and II (Molinari and Helenius, 1999). In contrast, when specific substrate was analyzed in vivo conflicting results were obtained. First, in Daudi cells, which lack β₂m and show no class I-CRT association, antibodies to CRT fail to co-precipitate ERp57. Likewise, in.220.B.8 cells, which lack tapasin and in which the CRT-ERp57-class I interaction is undetectable, no CRT-ERp57 association is observed (Hughes and Cresswell, 1998). Furthermore, as described in chapter 4, ERp57 is not present in complexes formed between CNX and its substrate, class I H chains in the Drosophila expression system. Second, in the Drosophila expression system CNX promotes efficient H chain folding and assembly with β₂m in the absence of ERp57 during class I biogenesis. No enhancement in disulfide bond formation was observed when ERp57 and CNX were coexpressed in Drosophila cells compared to CNX alone. Clearly, further studies are required to decipher whether in vivo ERp57 interacts with CNX
directly or through its association with a substrate protein. Furthermore, the role of ERp57 in CNX and CRT functions in vivo has to be resolved. CNX mutants that have lost their substrate or ERp57 binding site may be used to further clarify these issues.

Throughout this discussion I have attempted to convey the controversy which surrounds the issues of CNX/CRT functions. The inability to reach firm conclusions is in part due to the lack of structural data, but it is also a general reflection of the complexity of protein folding in the ER. In recent years, the predominant view that CNX is a lectin and its function is regulated by the cycle of binding and release regulated by glucosidase II and UGGT has largely been superseded. It has become apparent that additional peptide-peptide interactions play a role and this poses a question of what regulates the release of substrates bond to CNX/CRT via peptide-peptide interactions.

If a cycle of binding and release is to be evoked there must be an event which would result in the release of the substrate from CNX/CRT. The peptide binding chaperones whose actions are understood in great detail are all ATPases. Members of both Hsp60 and Hsp70 chaperone families bind and release their substrate by conformational changes in a manner regulated by ATP-binding and hydrolysis. An unresolved issue in the understanding of CNX function, as with HSP90 family members, is the relationship between substrate binding and adenine nucleotide binding and hydrolysis. CNX has been reported to bind ATP (Ou et al., 1993; Ihara et al., 1999). Direct evidence that a change in CNX conformation is associated with ATP binding comes from an intrinsic fluorescence study (Ihara et al., 1999). However, the inherent ATPase activity of CNX is much slower than the activity of other known molecular chaperones (Ihara et al., 1999). Furthermore, in vitro polypeptide binding by CNX is ATP-independent. So, the role of adenine nucleotides in regulating CNX substrate binding remains unclear.

The binding and release cycle may also be regulated by oligosaccharide binding. Recent experiments demonstrated that monogluosylated oligosaccharide binding reduces the affinity of CNX for its unglycosylated protein substrate in vitro (Ihara et al., 1999). It
was also demonstrated that the conformation of CNX/CRT changes in response to oligosaccharide binding. However, how the lectin-oligosaccharide interaction affects polypeptide binding is not clear. The crystal structure of CNX bound to oligosaccharide may serve to highlight the potential role of carbohydrate in substrate release. Furthermore, mutations in the oligosaccharide binding site may help to establish whether oligosaccharide binding is involved in substrate release. *In vitro* binding experiments of glycosylated and unglycosylated substrate with such mutants may serve as a model for such experiments.

By analogy to HSP70,60 and Hsp90, it is likely that the function of CNX is regulated by interactions with co-chaperones. However, there is currently no direct evidence for the existence of such ER co-chaperones. Once such co-chaperones are identified, the question of ATP hydrolysis may be revisited.

Finally, what ultimately causes the release of the substrate and transport out of ER may be regulated by affinity of the substrate for other molecular chaperones or transporters with ERGIC-53 being a good candidate for such a function.

5.7 Final Remarks

Throughout this discussion, I have attempted to convey the complexity of class I biogenesis and to discuss class I folding in a framework of general protein folding in the ER. When commencing the work on this thesis, there was a comfortably simple view regarding the involvement of molecular chaperones in class I biogenesis i.e., CNX was the major molecular chaperone for murine H chains. It has since become apparent that additional proteins are involved in murine class I biogenesis. CRT was not only shown to function as molecular chaperone for class I molecules in my studies with *Drosophila* cells (Chapter 2), but it was also detected in association with murine class I molecules in thymocytes and in fibroblasts transfected with the L^d_ gene (Van Leeuven and Kearse, 1996; Solheim et al., 1996; Yu et al., 1999). Furthermore, the ER housekeeping molecule, ERp57 was detected in association with free H chain as well as with H-chain-β_2m.
heterodimers (Hughes and Cresswell, 1998; Lindquist et al., 1998). Finally, class I molecules were found in association with the TAP transporter, through another ER resident molecule, the membrane-bound tapasin. As we understand better the complexity associated with class I molecules, the picture of a carefully regulated and specialized system emerges that ensures the surface expression of class I complexes essential for T cell immunity.

Over time, I also witnessed a shift in the paradigm explaining the mechanism of CNX functions. When the work in this thesis was begun, there was a predominant view that CNX is a lectin and its function is regulated by the cycle of binding and release regulated by glucosidase II and UGGT. It has since become apparent that additional peptide-peptide interactions play a role and as such a new family of molecular chaperones that exploits reactions unique to the ER environment has emerged.

In my thesis I have attempted to explore a basic immunological process of class I biogenesis as a means to understand the functions of molecular chaperones in the ER. As our understanding of molecular chaperones continues to expand, so does our ability to control the process of antigen presentation. We have entered a time in immunology at which our extensive understanding of protein biology can be applied to antigen presentation of tumor and pathogen infected cells. This close relationship to basic research offers tremendous promise for development of new disease prevention methods and treatments.
CHAPTER VI

References


Silverstein, A.M., Grammatikakis, N., Cochran, B.H., Chinkers, M. and Pratt, W.B. (1998) p50(cdc37) binds directly to the catalytic domain of Raf as well as to a site on hsp90 that is topologically adjacent to the tetraricopeptide repeat binding site. *J Biol Chem*, 273, 20090-5.


