THE MOLECULAR CHAPERONES, CALNEXIN AND CALRETICULIN: 
STUDIES ON FUNCTION AND MECHANISM OF ACTION

By

Aikaterini Vassilakos

A thesis submitted in conformity with the requirements for the degree of Doctor of Philosophy
Graduate Department of Biochemistry
University of Toronto

©Copyright by A. Vassilakos 1998
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.

0-612-35353-2
CLASS I HISTOCOMPATIBILITY MOLECULES ARE THE CELL SURFACE MOLECULES RESPONSIBLE FOR PRESENTATION OF ENDOGENOUSLY SYNTHESIZED FOREIGN ANTIGENS TO CYTOTOXIC T CELLS. CALNEXIN, A CALCIUM BINDING, TYPE I INTEGRAL MEMBRANE PROTEIN, ASSOCIATES TERNARILY WITH INCOMPLETELY ASSEMBLED CLASS I HISTOCOMPATIBILITY MOLECULES AND ALSO A HOST OF OTHER GLYCOPROTEINS WITHIN THE ENDOPLASMIC RETICULUM (ER). THE FOCUS OF THIS RESEARCH WAS TO ASCERTAIN THE NATURE OF THE ASSOCIATION BETWEEN CALNEXIN AND ITS SUBSTRATES AND TO CLARIFY THE FUNCTIONAL CONSEQUENCES OF ITS ASSOCIATION WITH CLASS I MOLECULES.

The results presented in Chapter IV suggest that calnexin functions as a bona fide molecular chaperone which increases the efficiency of class I folding and assembly. Calnexin prevents aggregation and degradation of class I heavy chains, thereby increasing productive folding and assembly. The results in Chapters II and III demonstrate that calnexin and calreticulin, a luminal ER protein homologous to calnexin, act as lectins which recognize Glc,Man,GlcNAc oligosaccharides that are present on newly synthesized glycoproteins. In addition to lectin-mediated interactions, evidence is presented that supports the hypothesis that calnexin interacts with polypeptide portions of newly synthesized polypeptides. A detailed examination of the lectin properties of these proteins was undertaken. The results demonstrate that the Glcα1-3Man structure is essential for the interaction of calnexin and calreticulin with oligosaccharide but secondary interactions occur along the length of the glucose-containing
branch. There is also a requirement for at least one mannos residue on the polymannose branch of the oligosaccharide. Binding competition experiments revealed that calnexin and calreticulin bind oligosaccharide with similar relative affinities, suggesting similar lectin sites in the two proteins. EGTA treatment of calnexin and calreticulin resulted in a loss of oligosaccharide binding, indicating that calcium is required for lectin function. In contrast, ATP did not affect oligosaccharide binding in a specific manner. Reduction of disulfide bonds in calnexin did not affect oligosaccharide binding. In contrast, reduction of disulfide bonds in calreticulin resulted in a loss of lectin activity, suggesting that the disulphide bond is important for the integrity of its lectin site. Finally, the lectin sites of calnexin and calreticulin were mapped to a region dominated by two sequence motifs that are tandemly repeated. Similar sequences have not been identified in other lectin families and as a result this segment constitutes a novel lectin site.
ACKNOWLEDGEMENTS

Five years ago David Williams invited me to become part of his research group and it has been a wonderful experience. David was a great supervisor. My continuing interest in science is in large part due to the enthusiasm and confidence he maintained over the course of my graduate work.

When starting graduate school one hears horror stories - they don’t show up, they don’t read the report you stayed up to 4am writing, they fall asleep while you’re talking or worse; they laugh uncontrollably - This was the exact opposite of my supervisory committee. Dr. Brian Barber and Dr. Reinhart Reithmeier were always interested, insightful, and challenging. For all of their help I am grateful.

Myrna taught me all I know about tissue culture and cell labelling, and without a doubt she is the best technical assistant around, anywhere! But that is only a small part of what Myrna brings to the lab. On difficult days, i.e. exams, seminars, she gives hugs that chase even the worst fears away. I am convinced that all those kisses I got when I was pregnant with Nathaniel contributed to his wonderful personality. I will be forever grateful for Myrna's friendship.

The Williams' lab has seen many wonderful people pass by its benches. Whether its science, family, politics or passing along the latest jokes, these people have enriched my graduate years. Thank you to Steve Doyle, Steve Pind, Erin Mitchell, Thuy Nguyen, Nav Ahluwalia, Kyung Suh, Ursula Danilczuk, and Mike Leach.

I would like to thank my parents, Panagiota Vassilakos, Chris Vassilakos and Linda Vassilakos for their support over the years. I am incredibly grateful for the support of my parents-in-law Jim and Mary MacPherson. My Uncle Peter deserves special thanks for being the first to recognize in me the desire to learn and to encourage me. I would also like to thank my sister Helen. She has always had faith in my abilities and willingly gave up many afternoons to watch Nathaniel while I made revisions to my thesis.

The research presented in this thesis would not have been possible without the excellent work done in the laboratories of our collaborators: Dr. Per Peterson, Dr. Mark Lehrman and Dr. Marek Michalak.

I am very grateful to the many people and organizations that have provided reagents for this work. These people are acknowledged in the Materials and Methods. I am also grateful for the financial support I received through an Ontario Graduate Scholarship and an NCIC Steve Fonyo Studentship.

The use of figures, published previously by others, has been with the written permission of the journal and where necessary the author. Figures not generated by myself are cited in the figure legends and copies of the permission forms are attached to the thesis.
It is not always easy to share your life with a graduate student. You have had to live through the ups and downs of my research (not to mention the late nights, lost weekends, and shortened vacations). You once said jokingly that the ultimate display of undying love is not poetry but typing references for a thesis.

Thank You.

For my children, Erin and Nathaniel

and

in memory of my

Aunt Meropi
### TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Page</th>
<th>Section</th>
</tr>
</thead>
</table>
| 1    | CHAPTER I  
INTRODUCTION |
| 1    | 1.0 Preface |
| 2    | 1.1 The Protein Folding Problem and the Discovery of Molecular Chaperones. |
| 7    | 1.2 Hsp 70 Family |
| 7    | 1.2.0 General Features |
| 10   | 1.2.1 Functions of Hsp 70/Hsp 40/Hsp (DnaK/DnaJ/GrpE) |
| 14   | 1.2.2 Structure |
| 19   | 1.2.3 Reaction Cycle |
| 22   | 1.3 Hsp 60 Family and the Chaperonins |
| 24   | 1.3.0 General features |
| 27   | 1.3.1 Functions of chaperonins |
| 32   | 1.3.2 Structure |
| 34   | 1.3.3 Reaction cycle of Gro EL/ES |
| 38   | 1.4 Hsp 90 Family |
| 38   | 1.5 Other Chaperones |
| 39   | 1.6 Cooperativity Between Chaperone Machines |
| 53   | 1.7 Protein Biogenesis in the Eukaryotic Secretory Pathway |
| 53   | 1.7.0 Secretory pathway and glycoprotein biogenesis |
| 57   | 1.7.1 Glycoproteins and lectins in the secretory pathway - Emerging functions |
| 60   | 1.7.2 ER chaperones |
| 70   | 1.7.3 Unfolded-protein response and the regulation of ER chaperone expression |
| 73   | 1.8 Calnexin and Calreticulin |
| 73   | 1.8.0 Calnexin |
| 82   | 1.8.1 Calreticulin |
| 86   | 1.8.2 The ER chaperone machine - Emerging model |
| 87   | 1.9 Major Histo Compatibility Complex Molecules |
| 87   | 1.9.0 General properties and function |
| 92   | 1.9.1 Class I molecules - General overview of biogenesis |
| 93   | 1.9.2 Generation of class I peptide ligands |
| 98   | 1.9.3 Transport of peptides via TAP1/2 and interaction of class I dimers with TAP1/2 |
| 102  | 1.9.4 Accessory molecules involved in class I biogenesis. |
| 107  | 1.9.5 Model of class I biogenesis |
| 110  | 1.10 Issues Addressed in this Thesis |
CHAPTER II
THE MOLECULAR CHAPERONE CALNEXIN BINDS Glc,Man,GlcNAc
OLIGOSACCHARIDE AS AN INITIAL STEP IN RECOGNIZING
UNFOLDED GLYCOPROTEINS

2.0 Abstract
2.1 Introduction
2.2 Materials and Methods
2.2.0 Production and analysis of dolichol-linked oligosaccharides
2.2.1 Preparation of soluble, hexahistidine-tagged H-2K and
calanexin proteins
2.2.2 Incubation of radiolabeled Glc,Man,GlcNAc oligosaccharides with immobilized proteins
2.2.3 Preparation of [H]Glc,Man,GlcNAc oligosaccharides and incubation with calnexin-agarose
2.2.4 Isolation and analysis of radiolabeled calnexin-class I heavy chain complexes
2.2.5 Isolation and Analysis of Calnexin-a,-Antitrypsin Complexes
2.3 Results
2.3.0 Calnexin selectively binds Glc,Man,GlcNAc from an oligosaccharide mixture
2.3.1 Glc,Man,GlcNAc oligosaccharides are relatively poor ligands for calnexin
2.3.2 Incubation of calnexin-glycoprotein complexes with α-methylglycosides
2.3.3 N-linked oligosaccharides are not involved in maintaining stable complexes between newly synthesized glycoproteins and calnexin
2.4 Discussion

CHAPTER III
OLIGOSACCHARIDE BINDING CHARACTERISTICS OF THE
MOLECULAR CHAPERONES, CALNEXIN AND CALRETIJULIN
3.0 Abstract
3.1 Introduction
3.2 Materials and Methods
3.2.0 Preparation of homogeneous [H]Glc,Man,GlcNAc oligosaccharides
3.2.1 Preparation of Soluble Proteins
3.2.2 Incubation of Radiolabeled Oligosaccharides with Immobilized Proteins
3.2.3 EGTA and ATP treatment of Calnexin - Assessment of Protease Sensitivity and Oligomerization State
3.3 Results
3.3.0 Oligosaccharide binding specificity of calnexin
3.3.1 Comparison of calnexin and calreticulin binding to the Glc,Man,GlcNAc
TABLE OF ILLUSTRATIONS

<table>
<thead>
<tr>
<th>Page</th>
<th>Table/Figure</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>Table I. Components of the HSP70 and chaperonin systems in bacteria and eukaryotic cells</td>
</tr>
<tr>
<td>13</td>
<td>Figure 1. Role of the mitochondrial Hsp70 chaperone system in translocation of polypeptides across the mitochondrial membrane</td>
</tr>
<tr>
<td>16</td>
<td>Figure 2. A. Schematic representations of Hsp70 and DnaJ B. Model for interaction of Hsp70 and DnaJ-like proteins (Hsp40)</td>
</tr>
<tr>
<td>17</td>
<td>Figure 3. A. Crystal structure of the polypeptide binding region of DnaK. B. ATPase domain of DnaK C. Ribbon representation of the structure of the DnaJ domain</td>
</tr>
<tr>
<td>20</td>
<td>Figure 4. A. Reaction cycle of bacterial Hsp70 (DnaK/DnaJ/GrpE). B. Reaction cycle of eukaryotic Hsp70 (Hsp70/Hsp40/Hip/Hop)</td>
</tr>
<tr>
<td>28</td>
<td>Figure 5. A. Schematic representation of the GroEL double doughnut B. A space-filling representation of the GroEL/ES molecule</td>
</tr>
<tr>
<td>33</td>
<td>Figure 6. Reaction cycle of the bacterial GroEL and GroES chaperonin system</td>
</tr>
<tr>
<td>40</td>
<td>Table II. Summarizes work on a number of chaperone molecules that are not part of the Hsp70, 60 or 90 families</td>
</tr>
<tr>
<td>46</td>
<td>Figure 7. Protein folding in the mitochondrial matrix</td>
</tr>
<tr>
<td>48</td>
<td>Figure 8. A. ER translocation pathways in mammals and yeast B. Schematic of the eukaryotic translocon complex in SRP mediated targeting and translocation</td>
</tr>
<tr>
<td>50</td>
<td>Figure 9. Cooperative interactions between nascent polypeptide chains, nascent chain binding proteins and chaperones in the cytosol of bacteria and eukaryotes</td>
</tr>
<tr>
<td>56</td>
<td>Figure 10. A. Synthesis and translocation of the dolichol-linked core oligosaccharide into the ER B. Pathway for transfer and processing of N-linked oligosaccharide</td>
</tr>
<tr>
<td>72</td>
<td>Figure 11. The unfolded protein response pathway in yeast (UPR)</td>
</tr>
<tr>
<td>75</td>
<td>Figure 12. Linear representations of calnexin and calreticulin</td>
</tr>
<tr>
<td>76</td>
<td>Table III. Table of calnexin and calreticulin substrates</td>
</tr>
<tr>
<td>81</td>
<td>Figure 13. The consensus motif 1 and 2 sequences of calnexin and calreticulin are shown along with an alignment of sequences from various species</td>
</tr>
<tr>
<td>89</td>
<td>Figure 14. A. Genomic organization of the human and mouse MHC B. Exon-intron structure of class I and II genes</td>
</tr>
<tr>
<td>91</td>
<td>Figure 15. A. Schematic of class I molecule B. Ribbon diagram of the soluble extracellular fragment of a class I molecule with bound peptide C. Ribbon diagram of the peptide binding groove of class I viewed from the top</td>
</tr>
</tbody>
</table>
CHAPTER IV

198  Figure 1. Drosophila expression system
206  Figure 2. Effects of calnexin on Kβ assembly with β2m in Drosophila cells
209  Figure 3. Effects of temperature and length of induction on Kβ assembly in the presence or absence of calnexin in Drosophila cells
210  Figure 4. β2m expression compared to heavy chain expression in the presence or absence of calnexin
211  Figure 5. Effects of calnexin on Dβ assembly with β2m in Drosophila cells
214  Figure 6. Calnexin enhances folding of Dβ heavy chains expressed in Drosophila cells
217  Figure 7. Calnexin inhibits aggregation of Kβ heavy chains in Drosophila cells
221  Figure 8. Effect of castanospermine on heavy chain-β2m assembly in mouse cells
225  Figure 9. Effect of castanospermine on heavy chain-β2m assembly in MDA.Y.D2 cells
227  Figure 10. Castanospermine treatment impairs folding of Dβ heavy chains in mouse cells
229  Figure 11. Effect of castanospermine treatment on the aggregation state of Dβ molecules
230  Figure 12. Aggregation state of class I molecules in EL4 cells
232  Figure 13. Castanospermine increases turnover of class I molecules in murine cell lines
234  Figure 14. Effect of castanospermine on cell surface expression of class I molecules
236  Figure 15. Effect of castanospermine on the folding and assembly of the human HLA-B27 molecule

CHAPTER V

248  Figure 1. Models of calnexin and calreticulin interactions with newly synthesized polypeptides
### ABBREVIATIONS

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Term</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>AAT</td>
<td>α₁-antitrypsin</td>
<td>human leukocyte antigen</td>
</tr>
<tr>
<td>Ab</td>
<td>antibody</td>
<td>high performance liquid chromatography</td>
</tr>
<tr>
<td>β₂m</td>
<td>β₂-microglobulin</td>
<td>heat shock protein</td>
</tr>
<tr>
<td>BiP</td>
<td>binding protein</td>
<td>immunoglobulin heavy chain</td>
</tr>
<tr>
<td>BSA</td>
<td>bovine serum albumin</td>
<td>immunoglobulin light chain</td>
</tr>
<tr>
<td>CCT</td>
<td>chaperonin containing TCP-I</td>
<td>low molecular mass protein</td>
</tr>
<tr>
<td>CFTR</td>
<td>cystic fibrosis</td>
<td>major histocompatibility complex</td>
</tr>
<tr>
<td>CHAPS</td>
<td>3-(3-cholamidopropyl)dimethylammonio-1-propanesulfonic acid</td>
<td>nascent chain associated complex</td>
</tr>
<tr>
<td>CNX</td>
<td>calnexin</td>
<td>N-ethylmaleimide</td>
</tr>
<tr>
<td>CNX-His</td>
<td>histidine tagged calnexin</td>
<td>NEM</td>
</tr>
<tr>
<td>cpn</td>
<td>chaperonin</td>
<td>NMR</td>
</tr>
<tr>
<td>CPY</td>
<td>carboxypeptidase</td>
<td>NP-40</td>
</tr>
<tr>
<td>CRT</td>
<td>calreticulin</td>
<td>PBS</td>
</tr>
<tr>
<td>CRT-His</td>
<td>histidine tagged calreticulin</td>
<td>PCR</td>
</tr>
<tr>
<td>CTL</td>
<td>cytotoxic T lymphocyte</td>
<td>PDI</td>
</tr>
<tr>
<td>DTT</td>
<td>dithiothreitol</td>
<td>protein disulfide isomerase</td>
</tr>
<tr>
<td>EGTA</td>
<td>ethylene glycol-bis(β-aminoethyl ether)</td>
<td>sodium dodecyl sulphate</td>
</tr>
<tr>
<td>endo H</td>
<td>endoglycosidase H</td>
<td>sodium dodecyl sulphate-polyacrylamide gel</td>
</tr>
<tr>
<td>LiK</td>
<td>endoplasmic reticulum</td>
<td>electrophoresis</td>
</tr>
<tr>
<td>fBS</td>
<td>fetal bovine serum</td>
<td>signal recognition particle</td>
</tr>
<tr>
<td>FITC</td>
<td>fluorescein isothiocyanate</td>
<td>signal recognition particle receptor</td>
</tr>
<tr>
<td>Gal</td>
<td>galactose</td>
<td>transporter associated with antigen presentation</td>
</tr>
<tr>
<td>Glc</td>
<td>glucose</td>
<td>tailless complex</td>
</tr>
<tr>
<td>GlcNAc</td>
<td>N-acetylglucosamine</td>
<td>polypeptide-I</td>
</tr>
<tr>
<td>Grp</td>
<td>glucose regulated protein</td>
<td>TCR</td>
</tr>
<tr>
<td>GST</td>
<td>glutathione-S-transferase</td>
<td>TRIC</td>
</tr>
<tr>
<td>GST-CN X</td>
<td>glutathione-S-transferase/calnexin fusion</td>
<td>TCP-1</td>
</tr>
<tr>
<td>GST-CRT</td>
<td>glutathione-S-transferase/calreticulin</td>
<td>UPRA</td>
</tr>
<tr>
<td>HA</td>
<td>influenza hemagglutinin</td>
<td>unfolded protein response element</td>
</tr>
<tr>
<td>HC</td>
<td>heavy chain</td>
<td>VSV</td>
</tr>
<tr>
<td>Hepes</td>
<td>4-(2-hydroxy ethyl)-1-piperazineethanesulfonic acid</td>
<td>vesicular stomatitis virus</td>
</tr>
<tr>
<td>HLA</td>
<td>human leukocyte antigen</td>
<td></td>
</tr>
<tr>
<td>HPLC</td>
<td>high performance liquid chromatography</td>
<td></td>
</tr>
<tr>
<td>Hsp</td>
<td>heat shock protein</td>
<td></td>
</tr>
<tr>
<td>IAA</td>
<td>iodoacetamide</td>
<td></td>
</tr>
<tr>
<td>IgG</td>
<td>immunoglobulin</td>
<td></td>
</tr>
<tr>
<td>IgL</td>
<td>immunoglobulin</td>
<td></td>
</tr>
<tr>
<td>kDa</td>
<td>kilodalton(s)</td>
<td></td>
</tr>
<tr>
<td>LMP</td>
<td>low molecular mass protein</td>
<td></td>
</tr>
<tr>
<td>mAb</td>
<td>monoclonal antibody</td>
<td></td>
</tr>
<tr>
<td>Man</td>
<td>mannose</td>
<td></td>
</tr>
<tr>
<td>Me</td>
<td>methyl</td>
<td></td>
</tr>
<tr>
<td>met</td>
<td>methionine</td>
<td></td>
</tr>
<tr>
<td>MHC</td>
<td>major histocompatibility complex</td>
<td></td>
</tr>
<tr>
<td>NAC</td>
<td>nascent chain associated complex</td>
<td></td>
</tr>
<tr>
<td>NEM</td>
<td>N-ethylmaleimide</td>
<td></td>
</tr>
<tr>
<td>NMR</td>
<td>nuclear magnetic resonance</td>
<td></td>
</tr>
<tr>
<td>NP-40</td>
<td>Nonidet P-40</td>
<td></td>
</tr>
<tr>
<td>PBS</td>
<td>phosphate buffered saline</td>
<td></td>
</tr>
<tr>
<td>PDI</td>
<td>protein disulfide isomerase</td>
<td></td>
</tr>
<tr>
<td>SDS-PAGE</td>
<td>sodium dodecyl sulphate-polyacrylamide gel</td>
<td></td>
</tr>
<tr>
<td>SRP</td>
<td>signal recognition particle</td>
<td></td>
</tr>
<tr>
<td>SRPR</td>
<td>signal recognition particle receptor</td>
<td></td>
</tr>
<tr>
<td>TAP</td>
<td>transporter associated with antigen</td>
<td></td>
</tr>
<tr>
<td>TCP-1</td>
<td>tailless complex</td>
<td></td>
</tr>
<tr>
<td>TRIC</td>
<td>TCP-1 ring complex</td>
<td></td>
</tr>
<tr>
<td>VSV</td>
<td>vesicular stomatitis virus</td>
<td></td>
</tr>
</tbody>
</table>
CHAPTER I
INTRODUCTION

1.0 Preface

Calnexin and calreticulin are resident proteins of the endoplasmic reticulum (ER) that constitute part of a large and diverse family of molecules termed molecular chaperones. The work presented here addresses several aspects of calnexin and calreticulin action. The chaperone functions of calnexin were examined using class I histocompatibility molecules as model substrates. The effect of calnexin on folding, assembly, aggregation and stability of class I molecules was examined in both homologous and heterologous expression systems. In addition, the molecular interactions which govern calnexin association with newly synthesized proteins were examined. Initial interactions of calnexin with newly synthesized proteins were shown to involve a lectin domain in calnexin that recognizes an early oligosaccharide processing intermediate on the substrate glycoprotein. Work by others has indicated that calreticulin binds newly synthesized glycoproteins in a similar fashion. These interactions were further characterized in detail using an *in vitro* system consisting of purified soluble recombinant calnexin and calreticulin and radiolabeled oligosaccharides.

The following introduction deals with much of the background to the work summarized above. Molecular chaperones will be discussed in terms of their physiological significance, discovery, structure and function. An overview of the eukaryotic secretory pathway will follow with emphasis on N-linked glycoprotein biogenesis and ER resident molecular chaperones. This will be followed by a detailed description of calnexin and calreticulin. Finally, the function, biogenesis, and structure of class I histocompatibility molecules will be covered.
Any discussion of protein folding and assembly must inevitably begin with the observations of Anfinsen about 4 decades ago (Anfinsen, 1973). The demonstration that under favourable conditions in vitro purified proteins such as ribonuclease can fold into a native structure without the aid of any other protein has led to an entire body of literature which deals with the thermodynamics and kinetics of protein folding in isolation at high dilution (reviewed in Baldwin, 1989). This research has been incredibly useful in determining the forces that drive a linear polymer, consisting of amino acids, into a compact folded structure and ultimately may provide enough information to predict how the information in this linear array will fold into a functional molecule without prior knowledge of its structure. Briefly, urea and heat denaturation of polypeptides followed by spontaneous renaturation have provided a kinetic model for how polypeptides fold. In the time scale of micro- to milliseconds, secondary structure forms and compact folding intermediates appear (reviewed in Friere, 1995, Fink, 1995, and Englander, 1993). These intermediates, also known as “molten globule” structures, retain significant surface exposed hydrophobicity and exhibit unstable tertiary structure. It remains controversial whether these “molten globule” structures are part of the productive folding pathway of a polypeptide or whether these constitute digressive pathways which unfold prior to productive folding (Creighton, 1997). In the seconds to minutes time scale, folding to a stable tertiary structure occurs. The rate-limiting step at this stage appears to be cis-trans isomerization of prolyl peptide bonds. This rate-limiting step is overcome in vivo by catalysis of isomerization by a family of folding enzymes called peptidyl-prolyl cis-trans isomerases (reviewed in Lorimer 1992). Under oxidizing conditions in subcellular compartments such as the ER, the formation of
intermolecular and intramolecular disulfide bonds may occur and stabilize the final structure of some proteins (reviewed in Lorimer, 1992, Freedman et al., 1994, Bardwell and Beckwith, 1993 and Ruddon et al., 1996). In vivo, formation and re-arrangement of disulfide bonds during later stages of folding is catalysed by protein disulfide isomerase (PDI) in the ER of eukaryotes and by dsb gene products in the E. coli periplasm.

That a polypeptide contains necessary and sufficient information to determine its final structure is not in question but, in vitro, the conditions required for folding, the resulting kinetics of folding, and the efficiency of folding are not consistent with those observed in vivo. In vitro, the concentrations of unfolded proteins typically used to achieve successful refolding are considerably lower than those observed in vivo (0.01-0.02 mg/ml compared to several mg/ml) (reviewed in Ruddeon and Bedows, 1997). At physiological concentrations of polypeptide, the in vitro folding reaction does not produce correctly folded molecules, but instead aggregation occurs (reviewed in Ruddeon and Bedows, 1997). Furthermore, the time scale of folding in vitro is on the order of minutes to hours, which is not consistent with considerably faster folding rates observed in vivo. While some unfolding and refolding of proteins occurs in the cell under stress conditions, the majority of folding occurs during biosynthesis. Herein lies another difference from the in vitro folding reaction. In the in vitro folding reaction, thermal or chemical denaturation followed by refolding approximates refolding of proteins after stress, and not folding which occurs during protein biogenesis under normal conditions. During translation the N-terminal residues emerge from the ribosome first but often cannot fold properly until C-terminal residues emerge that interact in the final structure. This presents a problem in that nascent polypeptides must avoid aggregation under conditions of high
temperature and protein concentration (reviewed in Gething and Sambrook, 1992).

The disparate observations in vitro and in vivo were reconciled with the discovery of a class of molecules whose synthesis is dramatically upregulated during exposure to stress, i.e., heat shock, oxidative stress, glucose starvation, and osmotic shock (reviewed in Georgopolous and Welch, 1993, and Craig et al., 1993). Exposure to mild heat treatment prior to exposure to lethal temperatures was found to protect cells from the subsequent lethal temperature. This phenomenon is termed thermotolerance and further research led to the discovery of several conserved families of molecules, termed heat shock proteins, whose expression is upregulated following stress (Table 1). The ability of these molecules to protect cells from stress was linked to their ability to prevent and/or reverse denaturation and/or aggregation of proteins. In addition, these molecules were associated with degradation of damaged proteins (reviewed in Hayes and Dice, 1996). The observed expression of inducible stress proteins under non-stress conditions and the subsequent identification of constitutive isoforms of the inducible heat shock proteins resulted in the understanding that these molecules are more than just “heat shock proteins” and led to the use of the term “molecular chaperone” (reviewed in Gething and Sambrook, 1992, Hendrick and Hartl, 1993, Hartl et al., 1994, Hartl and Martin, 1995, Hartl, 1996, and Johnson and Craig, 1997). The importance of these molecules in the normal functioning of the cell is highlighted by the demonstration that deletion or mutation of the genes encoding molecular chaperones often leads to lethal or growth retarded phenotypes (Table 1). By definition, molecular chaperones are molecules that interact with newly synthesized, unfolded, or misfolded proteins, assist in the folding and/or assembly of these molecules but are not components of the final structure.
Table 1

Components of the Hsp70 and chaperonin systems in bacteria and eukaryotic cells (Table taken from Hartl, 1996).

<table>
<thead>
<tr>
<th>Component System</th>
<th>Bacteria</th>
<th>Eukaryotic homologues</th>
<th>Yeast</th>
<th>Higher eukaryotes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hsp70 system</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hsp70</td>
<td></td>
<td>Octosin</td>
<td></td>
<td>Hsp70; stress-inducible</td>
</tr>
<tr>
<td></td>
<td>nucleotide exchanger for Hsp70</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hsp40</td>
<td></td>
<td>Octosin</td>
<td></td>
<td>Hsp40; binds nascent polypeptides</td>
</tr>
<tr>
<td></td>
<td>nucleotide-exchange factor for Hsp70</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GroE</td>
<td></td>
<td>Octosin</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>GroE family (GroES)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Two rings of 7-80K subunits</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>ATPase activity, binds protein-folding intermediates, promotes folding together with Hsp10 and DnaK</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>GroES</td>
<td>Mesochochrome</td>
<td>Hsp60/GroE; required for folding of imported proteins</td>
<td>Hsp60/GroE10</td>
</tr>
<tr>
<td></td>
<td>TRC family</td>
<td>Mesochochrome</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Heterodimer, 2 rings of 6</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>subunits, ATPase activity, binds protein-folding intermediates, promotes folding in eukaryotes and Archaea</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chaperonin system</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Mesochochrome</td>
<td>Hsp60/GroE; required for folding of imported proteins</td>
<td>Hsp60/GroE10</td>
<td></td>
</tr>
<tr>
<td></td>
<td>GroES</td>
<td>Mesochochrome</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>TRC family</td>
<td>Mesochochrome</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Heterodimer, 2 rings of 6</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>subunits, ATPase activity, binds protein-folding intermediates, promotes folding in eukaryotes and Archaea</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* GroEL (Sto) is also known as GroESL-substrate-exchange system.
† The subunits of GroEL's oligomeric factor GroES consist of two fused GroEL homologous domains.
‡ TRC is also known as GroESL (chaperonin containing eubacterial polypeptides).
A wide range of molecules are now classified as molecular chaperones and are found in prokaryotes and in most organelles of eukaryotes (mitochondria, chloroplasts, cytosol, nucleus and ER). The best studied of these are the Hsp60 and 70 families. Much of our understanding of these molecules comes from studies of bacterial and eukaryotic organisms such as yeast, in which expression can be easily manipulated.

Molecular chaperones appear to be involved in most processes requiring protein folding and/or assembly including translation, translocation across organelle and plasma membranes, DNA replication, steroid hormone action, oncogene targeting, protection and recovery from stress, and finally, protein degradation (reviewed in Gething and Sambrook, 1992, Hendrick and Hartl, 1993, Hartl et al., 1994, Hartl and Martin, 1995, Hartl, 1996, and Johnson and Craig, 1997). In general, chaperones appear to function by binding to incompletely folded proteins and preventing inappropriate inter- and intramolecular interactions, thereby allowing the protein to fold into its native structure under conditions which would otherwise lead to aggregation and misfolding (reviewed in Ruddon and Bedows, 1997). In addition, chaperones can promote renaturation of misfolded substrates and disaggregate damaged proteins. This is effected by binding with higher affinity to the free unfolded state and driving the equilibrium towards the unfolded state, thereby allowing for additional attempts to fold along the correct pathway. The mechanisms by which chaperones bind to and facilitate protein folding are not well understood. For the Hsp60 (bacterial GroEL/ES) and Hsp70 (bacterial Dna K) families a better understanding has come from elucidation of their X-ray crystal, NMR, and electron microscopic
structures. This, in combination with elegant \textit{in vitro} experiments using purified recombinant proteins, has led to detailed mechanistic models of the reaction cycles involved (see subsequent sections). What has become clear from these studies and others is that chaperones, while essential for rapid and efficient folding of proteins \textit{in vivo}, do not alter the productive folding pathway to native structure but assist folding by preventing non-productive folding pathways and rescuing proteins from these pathways. By unfolding non-productive conformers of a folding protein and providing an “Anfinsen Cage” (pseudo infinite dilution) chaperones such as GroEL/ES provide an opportunity for proteins to fold properly (reviewed in Hartl, 1996). Therefore, the forces that drive correct protein folding in the cell are not different than those observed \textit{in vitro}, just complicated by the incorrect pathways which predominate in the non-ideal protein folding environment within cells.

The following sections will discuss structural and mechanistic aspects of the major chaperone families. This will provide the background necessary for discussing the experimental data presented in chapters II, III, and IV in terms of chaperone structure and function.

\subsection*{1.2 Hsp70 Family}

\subsubsection*{1.2.0 General Features}

Members of the Hsp70 family of molecular chaperones are found in diverse species as well as in various organelles (Table I). Hsp70 homologues are found in the cytosol, mitochondria and ER, all sites of protein folding in the eukaryotic cell. With 60-80\% amino acid identity between eukaryotic Hsp70s and 40-60\% identity between the bacterial and eukaryotic proteins, they are one of the most highly conserved gene families (Lindquist, 1986). The
bacterial homologue, DnaK, was originally identified by complementation of a series of mutant strains that were deficient in bacteriophage λ growth and E. coli DNA replication (reviewed in Freidman et al., 1984). DnaK was subsequently found to allow initiation of DNA replication by disassembling the helicase complex at the origin of replication (Alfano et al., 1989, Zylicz et al., 1989). Many of these early studies suggested that DnaK may somehow unfold proteins (Alfano and McMacken, 1989). Given that many Hsp70's are dramatically upregulated in response to heat treatment, much of the early work in this area also focused on DnaK's role in the heat shock response. That eukaryotic Hsp70's and bacterial DnaK are essential under normal conditions was demonstrated by deleting these genes in bacteria and yeast with resulting growth defect phenotypes. The reasons for these observations became clear as the various functions of Hsp70's were characterized in more detail (see below).

Hsp70's bind ATP and ADP and are weak ATPases (reviewed in Georgopoulos and Welch, 1993). In addition to ATPase activity, Hsp70's were found to bind peptides. In vitro binding assays demonstrated that peptide binding results in stimulation of the ATPase activity (Flynn et al., 1989 and Flynn et al., 1991, reviewed in Georgopoulos and Welch, 1993). This will be discussed further below.

Not long after the discovery of these molecules was the discovery that both eukaryotic and prokaryotic Hsp 70's (both constitutive and induced isoforms) require accessory molecules for optimal function (Liberek et al., 1991, Langer et al., 1992, Cyr et al., 1992, and reviewed in Frydman and Höhfeld, 1997). Although DnaK was sufficient to rescue and promote re-folding of heat inactivated RNA polymerase in the presence of ATP in vitro, it did so only at a large molar excess of DnaK (Skowyra et al., 1990). Further work established that in addition to DnaK, DnaJ
and GrpE were essential for proper function (Ziemienowicz et al., 1993). In eukaryotes, homologues of DnaJ, termed Hsp40, have also been found to be required for efficient Hsp70 function (Table 1). Just as there are Hsp 70 homologues unique to each organelle, Hsp40 homologues have been characterized in cytosol, mitochondria and ER (reviewed in Cyr et al., 1995 and Silver and Way., 1993). As outlined below, the sequences of these homologues diverge considerably outside of the region that binds Hsp70 (J domain) (reviewed in Silver and Way., 1993 and Cyr et al., 1994). Interestingly, one of the the DnaJ homologues in the ER (Sec 63p) is membrane anchored (Sadler et al., 1989, Feldheim et al., 1992, and reviewed in Silver and Way., 1993 and Cyr et al., 1994). DnaJ, like DnaK, interacts with peptides. In addition, it has been demonstrated that J homologues interact specifically with the corresponding Hsp70 in a reaction cycle that is outlined below. Schlenstedt et al. demonstrated that two non-ER DnaJ J domains cannot be swapped into the ER DnaJ homologues, Sec63p or Sec1p (Schlenstedt et al., 1993). Similarly, early work on the functional differences between Hsp70's in different organelles demonstrated that cytosolic Hsp70 could not functionally replace BiP, the ER homologue, and vice versa (Brodsky et al., 1993). It is interesting to speculate that this phenomenon may be in part the result of there not being the correct DnaJ homologue present. The last components of the Hsp70 machinery are GrpE, a 22 kDa protein in the prokaryotic cytosol and in mitochondria, Hip, a 48 kDa protein in the eukaryotic cytosol, and Hop a recently discovered protein in the eukaryotic cytosol (reviewed in Frydman and Höhfeld, 1997). In these last components the eukaryotic and prokaryotic Hsp70 systems and resulting mechanisms diverge (see below). GrpE is a nucleotide exchange factor that interacts with DnaK and stimulates release of ADP and the concomitant binding of ATP. Since Hsp70 in the ATP form binds unfolded substrates weakly
this results in rapid release of substrate. In contrast Hip, interacts with Hsp 70 and prevents the release of ADP thereby increasing the lifetime of the Hsp 70/Hsp 40/peptide complex. Hop, like GrpE, is a nucleotide exchange factor that stimulates release of ADP, binding of ATP and thus dissociation of the Hsp70/substrate/Hip complex.

1.2.1 Functions of Hsp 70/Hsp 40/Hip (DnaK/DnaJ/GrpE)

In cooperation with DnaJ (Hsp40) and GrpE or Hip/Hop, Hsp70 family members function in diverse processes such as protection of cells from stress, translation of RNA, translocation of polypeptides across membranes, DNA replication and in protein folding.

While there are essential roles for Hsp70 under normal conditions, the inducible forms of Hsp70 can be upregulated to prevent and/or correct damage induced by a number of stresses. The ability of DnaK, DnaJ and GrpE to co-ordinately prevent and repair heat induced damage has been demonstrated in vivo and in vitro (Schröder et al., 1993, Ziemienowicz et al., 1993). It has been determined that DnaK reversibly binds to unfolded polypeptides that occur during heat denaturation, prevents their aggregation and even disaggregates preformed aggregates (Hwang et al., 1990, Skowyrta et al., 1990 and Wickner et al., 1991). For example, after denaturation of firefly luciferase by heat shock, activity is restored in wild type bacterial strains but in dnak or dnaj mutant strains the activity is lost and the enzyme aggregates (Schröder et al., 1993). Several in vitro studies with substrates like RNA polymerase, rhodanese and luciferase confirmed that DnaK, in conjunction with its accessory molecules, could protect and even rescue heat denatured proteins (Skowyrta et al., 1990, Langer et al., 1992, Schröder et al., 1993, Ziemienowicz et al., 1993). Collectively, these results suggest that Hsp70 acts as a molecular
chaperone. In some cases, as with rhodanese and UmuC, efficient refolding requires GroEL in addition to DnaK (Langer et al., 1992 and Petit et al., 1994).

As alluded to in the introduction, there is a need for chaperone activity at the site of protein synthesis in order to maintain the newly synthesized protein in a folding competent state until emergence of the entire molecule. Hsp70 has been identified as one of the components of the nascent chain associated complex (NAC), a complex that co-sediments with ribosomes actively translating message in rabbit reticulocyte lysate (Beckman et al., 1990, Nelson et al., 1992, Frydman et al., 1994, Hansen et al., 1994, and reviewed in Bukau et al., 1996). The involvement of Hsp70 in translation is further supported by the finding that Sis1p (DnaJ homologue in S. cerevisiae) is required for translational initiation (Zhong and Arndt, 1993). In E. coli, DnaK interaction with nascent chains appears to be regulated by DnaJ and Grp E, in a reaction cycle similar to that of binding to unfolded proteins (see section 1.2.3) (Kudliki et al., 1996). As will be discussed in a subsequent section, molecular chaperones are co-ordinately involved in protein synthesis in both eukaryotes and prokaryotes. Hsp70's, chaperonins (see section 1.6) and nascent chain associated molecules, like NAC and Trigger Factor can be found associated with polypeptides as they emerge from the ribosome. Interaction of Hsp70 with the nascent polypeptide appears to be downstream of binding by Trigger Factor/NAC's (nascent chain associated proteins) and upstream of Gro EL/TRIC (chaperonins) interactions, although there is evidence that these components may be associated at different points along the same elongating chain (reviewed in Bukau et al., 1996).

The involvement of Hsp70 and its co-chaperones in facilitating translocation of polypeptides across membranes during protein targeting to organelles such as the ER,
mitochondria, peroxisome and more recently the nucleus has been clearly demonstrated

(Deshaies et al., 1988, Nguyen et al., 1991, Brodsky and Schekman, 1993, Craven et al., 1996, Vogel et al., 1990, Kang et al., 1990, Komiya et al., 1996, Ungermann et al., 1996, Schluja et al., 1996, Yang and DeFranco, 1994, Lyman and Schekman, 1997, reviewed in Pfanner et al., 1994, Stuart et al., 1994, Wickner, 1994, Brodsky, 1996 and Bukau et al., 1996). It would appear that cytosolic Hsp70's are required to maintain polypeptides, synthesized in the cytosol, in a relatively unfolded conformation until targeting to the translocation machinery of the appropriate organelar membrane occurs. Members of the Hsp70 family and corresponding Hsp40 homologues are also found on the lumenal side of the organelar membrane at the site of the translocation machinery, e.g., ER: BiP or Lhs1p (Hsp70 homologues)/Sec63 or Sec1 (Hsp40 homologues), mitochondria: mHsp70/MDJ1 (Hsp40 homologue). Not only do these function to prevent the polypeptide from aggregating prior to translocation of the entire molecule but, in mitochondria (mtHsp70) and ER (BiP) have been implicated directly in the translocation process. In mitochondria, mtHsp70 binds to Tim 44 (a membrane bound protein) on the inner membrane surface of the mitochondrial membrane and is thought to act as a molecular “ratchet” to pull the polypeptide in the correct direction and prevent “back-sliding” to the cytosol (Figure 1). The involvement of BiP in translocation will be discussed in a subsequent section.

In addition to the above functions there is evidence from several studies that implicate HSP 70's in the degradation of aberrant proteins which are damaged or terminally misfolded (reviewed in Hayes and Dice, 1996). Evidence for this comes from the observation that aberrant polypeptides are “hypodegraded” in GroEL or DnaK mutants (Keller et al., 1988, Strauss et al., 1988, and Sherman and Goldberg, 1992). Furthermore these chaperones appear to associate both
Figure 1
Role of the mitochondrial Hsp70 chaperone system in translocation of polypeptides across the mitochondrial membrane. mHsp70 binds to translocating proteins tightly as ATP is hydrolysed. GrpE mediated exchange of ADP for ATP allows for release and re-binding of mHsp70.
(Taken from Hartl, 1996)
with unstable proteins as well as proteases.

The preponderance of work in the area of Hsp 70 function focuses on its role in the folding of newly synthesized proteins and of denatured proteins following stress conditions. In addition, there are examples of equally essential, but quite specialized functions of Hsp 70. In these cases Hsp 70 appears to interact with hydrophobic segments exposed on functional molecules. For example, a role for the DnaK/DnaJ/GrpE reaction cycle in regulation of σH activity has been reported (Liberek et al., 1992, Gamer et al., 1992, Liberek and Georgopolous, 1993, and Gamer et al., 1996). The σ subunit of RNA polymerase is required for specific initiation of transcription at promoter sites. σH is a specialized sigma subunit required for transcription of genes under the control of the heat shock promoter. Binding of DnaK/DnaJ to σH appears to inhibit its activity and this is partially reversed by GrpE activity (Gamer et al., 1996). Furthermore, Hsp 70 is required for some phage, plasmid and chromosomal DNA replication in *E. coli* (Wickner et al., 1991, Wickner et al., 1992, Zylicz et al., 1989, and reviewed in Georgopolous and Welch, 1993). Mammalian cytosolic Hsp 70 has also been implicated in the ATP dependent uncoating of clathrin coated vesicles. It appears to interact with vesicles via interactions with auxilin which contains a J domain (Chappell et al., 1986, Gao et al., 1991, reviewed in Rothman and Schmid, 1986, and Hartl., 1996).

1.2.2 Structure

Mutational, crystallographic and NMR analyses of Hsp 70/DnaK, Hsp 40/DnaJ and GrpE have provided a detailed picture of the structure/function relationship in these molecules (reviewed in Hartl., 1996). The most highly conserved region of Hsp70 molecules is the 44 kDa
N-terminal domain that contains the ATPase activity (Figure 2). The isolated 44 kDa fragment displays constitutive ATPase activity, whereas in the full length molecule the ATPase activity is activated by the binding of Hsp40/DnaJ and peptide (Chappell et al., 1987). In Figure 3B the crystal structure of the 44 kDa fragment is shown (Flaherty et al., 1990). The structure is also similar to that of actin, prokaryotic cell cycle proteins and hexokinase, an ATP-binding protein that undergoes substrate-mediated conformational changes in its ATP binding domain (Flaherty et al., 1991 and Bork et al., 1992). Therefore it was proposed that Hsp70 could undergo conformational changes in the 44 kDa N terminal domain. In a subsequent study, the crystal structures of the ATP and ADP bound forms of the 44 kDa fragment were solved (Flaherty et al., 1994). Consistent with its structural similarity to hexokinase, the 44 kDa fragment exists in two different conformations depending on which nucleotide is bound.

The peptide binding domain of Hsp70/DnaK has been localized to an 18 kDa C-terminal region that is less conserved among Hsp70s (Figure 2) (Wang et al., 1993). The crystal structure of the 18 kDa fragment has been solved (Figure 3A) (Zhu et al., 1996). The peptide is bound in a hydrophobic cleft formed by two, four stranded β sheets and the peptide binding domain is stabilized by an α helix. This structure was not expected given that earlier theoretical models predicted a structure similar to the peptide binding domain of class I histocompatibility molecules (reviewed in Hartl, 1996). Circular dichroism studies on the full length Hsp70 molecules demonstrated that the molecule can undergo conformational changes, in both the 44 kDa and 18 kDa fragment regions, upon peptide or ATP binding (Park et al., 1993). The most striking feature of the Hsp70 structure is that it can undergo such conformational changes in response to different interactions. The nucleotide bound at the N terminal ATPase domain can
Figure 2

A. Schematic representations of Hsp70 and DnaJ showing the peptide binding regions, the ATPase domain of Hsp70, and other conserved regions. GrpE binds to DnaK (bacterial Hsp70) via the ATPase domain (Taken from Hartl, 1996).

B. Model for interaction of Hsp70 and DnaJ-like proteins (Hsp40) in a ternary complex with polypeptide (not to scale). The filled ovals are the zinc finger-like regions of the polypeptide binding domain of Hsp40 and PBD is the polypeptide binding domain of Hsp70 (Taken from Cyr et al., 1994).
Figure 3

A. Crystal structure of the polypeptide binding region of DnaK. The polypeptide binding domain consists of two four-stranded $\beta$ sheets and one $\alpha$ helix (left panel). The $\beta$ sandwich contacts the peptide (top view of bound peptide in right panel), whereas the $\alpha$ helix stabilizes the polypeptide binding domain. The arrows point out the bound peptide. This does not resemble the peptide binding domain of class I molecules in spite of the fact that they both bind peptides in an extended conformation (Taken from Zhu et al., 1996).

B. ATPase domain of DnaK (Flaherty et al., 1990).

C. Ribbon representation of the structure of the DnaJ domain (Szyperski et al., 1994).
alter the peptide binding properties at the C terminal substrate binding domain and vice versa. Therefore it is somewhat surprising that the two domains can function when separated.

Interestingly, the coupling of ATPase activity and peptide binding is dependent on an absolutely conserved amino acid, E171 (Buchberger et al., 1994). E171 mutants lose peptide stimulated ATPase activity and there is no ATP mediated substrate dissociation. This in spite of the fact that the mutant can bind peptide and retains ATPase activity. Therefore this residue is critical for the coupling of the two domains (Buchberger et al., 1994).

Structure/function studies of DnaJ have determined that the N-terminal J region (highly conserved motif) and the adjacent Gly/Phe rich region are sufficient to bind to DnaK, regulate the conformation of DnaK, stimulate its ATPase activity and activate binding of DnaK to σ2 (Figure 2) (Wall et al., 1994 and Szyperski et al., 1994). The structure of the DnaJ J region (Figure 3C) has been solved by NMR and consists of four helices, two of which are amphipathic and form an antiparallel helical bundle that is stabilized by hydrophobic interactions along the interacting face of the helices (Figure 3C) (Szyperski et al., 1994). The Gly/Phe region was found to be highly disordered in the NMR structure. In addition to the J domain, full chaperone activity of the DnaK/DnaJ/GrpE chaperone machine requires additional residues in DnaJ. The cysteine-rich central domain is also required for full function. A recent structure-function analysis of DnaJ mapped the polypeptide binding region to the 90 amino acid cysteine-rich central region (Szabo et al. 1996). This region contains two zinc-finger domains that are required for peptide binding to DnaJ.

Peptide binding/ATPase assays have demonstrated that Hsp 70 can bind peptides as long as 25 amino acids but optimal ATPase stimulation occurs with peptides that are 7 amino acids in
length (Flynn et al., 1989, Flynn et al., 1991). From a number of studies it has been suggested that Hsp 70 preferentially binds hydrophobic amino acids. NMR and crystal structures confirm that peptides bind, in an extended conformation, via hydrophobic interactions with hydrophobic residues in the peptide binding pocket (Flynn et al., 1991, Landry et al., 1992, Langer et al., 1992, Blond-Elguindi et al., 1993a, Richarme and Kohiyama, 1993, Gragerov et al., 1994 and Zhu et al., 1996). While all the Hsp 70 homologues are thought to bind extended conformations of unfolded polypeptides via hydrophobic interactions, it has been suggested that their binding specificities are subtly different. Thus, it is not surprising that Hsp 70s located within different intracellular compartments are not functionally interchangeable (Brodsky et al., 1993).

1.2.3 Reaction Cycle

The Hsp 70 reaction cycle is probably the best characterized of the chaperone “machines”. What has emerged from a number of in vivo and in vitro studies is a mechanism that intimately involves all of the components in a cycle of binding and release controlled by ATP (see Figure 4A for prokaryotic and Figure 4B for eukaryotic cytosolic mechanisms) (reviewed in Hartl., 1996). While DnaK (HSP70) has received much of the attention in the past it is becoming increasingly clear that DnaJ/Hsp40 is a central component involved in ATP independent substrate binding and presentation of substrate to DnaK/Hsp70 (Szabo et al., 1994). DnaJ/Hsp40 bind first to the unfolded substrate and via interactions between DnaJ/Hep40 and DnaK/Hsp70, present the substrate to the ATP bound form of DnaK/Hsp70. Interaction of DnaJ/Hsp40 to DnaK/Hsp70 stabilizes binding of substrate to DnaK/Hsp70. Upon binding of
Figure 4
A. Reaction cycle of bacterial Hsp70 (DnaK/DnaJ/GrpE) (Taken from Hartl, 1996). B. Reaction cycle of eukaryotic Hsp70 (Hsp70/Hsp40/Hip/Hop) (Taken from Frydman and Hohfeld, 1997).
DnaJ/Hsp40 and substrate to DnaK/Hsp70 the ATPase domain is stimulated, resulting in the ADP bound form of DnaK/Hsp70. This form binds substrate more tightly than the ATP bound state. In the prokaryotic system, release of substrate is effected by the nucleotide exchange activity of GrpE, which exchanges ADP for ATP on DnaK (Figure 4A). In the eukaryotic cytosol there is no GrpE homologue, but instead the third component to the Hsp70 machinery is Hip, a molecule that stabilizes the ADP bound state and prevents premature substrate dissociation from Hsp70. Recently, another molecule, termed Hop, has been incorporated into the eukaryotic mechanism (reviewed in Frydman and Hohfeld, 1997). Hop dissociates the Hsp70/substrate/Hip complex by stimulating the exchange of ADP for ATP (Figure 4B). This difference in the reaction cycles is likely the result of eukaryotic and prokaryotic cycles having different rate limiting steps and as a result require different accessory molecules. What has become clear in the last few years is that unfolded proteins that bind to HSP 70/40 are released in an unfolded state, suggesting that folding actually occurs while released from the chaperones and not while bound (reviewed in Hartl, 1996). It is only after repeated cycles of binding and release that proteins fold correctly, being aided by the fact that interaction with Hsp 70/40 prevents aggregation and mis-folding. In some cases, such as rhodanese, the DnaK/J/GrpE cycle is insufficient to facilitate folding and only in the presence of the Hsp60 chaperone system, GroEL/ES, can folding ultimately occur (Langer et al., 1992). This will be discussed further in the Chaperone Co-operativity section. In some cases if the protein never folds it is eventually degraded.

The fact that proteins fold once released from DnaK is in contrast to the Hsp60 chaperonin family members, that bind to substrate and release it into an internal cavity. This
provides an "Anfinsen cage" in which the substrate can fold, at least partially, prior to release (see below). While the chaperonin system also functions through a cycle of ATP-regulated binding and release, folding of unfolded proteins (and unfolding of mis-folded proteins) does occur while the molecule is within the complex (reviewed in Hartl., 1996). This difference does not seem to be important in the vast majority of folding reactions since there are a number of substrates that are equally well chaperoned by either the Hsp 70 or Hsp 60 machinery in vitro. In vivo it is more likely that, under normal conditions, there is a co-ordinated and sequential action by chaperone machines that provide optimal folding. This co-ordinated action is sometimes dictated by topology such as in the mitochondrial import and folding pathway where Hsp 70 molecules on either side of the membrane must act first to promote translocation which is then followed by Hsp 60/10 mediated folding and assembly. In other cases, such as the synthesis of cytosolic proteins, there is evidence that interactions with the nascent polypeptide are ordered (see below).

1.3 Hsp 60 Family and the Chaperonins

1.3.0 General features

The first Hsp60 member to be identified was the bacterial GroEL molecule along with its co-chaperone, GroES, also termed Hsp10. Mutations in the E. coli groE locus resulted in the loss of bacteriophage λ growth. Further examination revealed that this locus contained two genes that coded for the GroEL (58 kDa large subunit) and GroES (10 kDa small subunit) polypeptides (reviewed in Friedman et al., 1984). Although these proteins were originally identified as essential for bacteriophage head assembly, it was soon determined that they were required for
synthesis of *E. coli* proteins as well (reviewed in Zeilstra-Ryalls et al., 1991). The importance of these gene products under both normal and stress-induced conditions was emphasized by the lethal phenotype observed in mutants that have deletions in this region (Fayet et al., 1989).

GroEL had all the hallmarks of a molecular chaperone. It bound transiently to greater than 50% of all newly synthesized *E. coli* proteins and was found to bind exclusively to unfolded or unassembled proteins (Bochareva et al., 1988 and Viitanen et al., 1992).

Since the identification of GroEL/ES(Hsp60) there have been numerous additions to the Hsp60 family of proteins and also the larger family of molecules called chaperonins (based on their characteristic structural organization - see below). Bacterial Hsp60 homologues have been identified in thermophilic bacteria and in mycobacteria. There are also homologues in a variety of eukaryotic cells including yeast, mammals and plants that are located within mitochondria and chloroplasts. In addition to Hsp60 homologues, eukaryotic cells contain a cytosolic chaperonin, TCP1 (TRiC), which is an oligomeric chaperone bearing no sequence similarity to GroEL (reviewed in Hartl., 1996).

Hsp 60's are similar to the Hsp70 family in that they are ATP binding proteins that exhibit weak ATPase activity. As the name implies, most Hsp 60 members, except the chloroplast homologue, are induced by heat shock and other stresses but, like the Hsp 70's, are also constitutively expressed. Hsp 60 molecules differ substantially from Hsp70 in that they exist as multimeric structures. The functional structure of bacterial GroEL/ES consists of 14 identical GroEL subunits arranged in two stacked rings. The GroES subunits form similar ring structures that can bind to either end of the GroEL "double doughnut". This similar oligomeric structure is found in the TCP1 complex (TRiC), the eukaryotic cytosolic chaperonin, with the exception that
each double ring contains 8 or 9 members of a conserved family of 52-65 kDa proteins unrelated to GroEL (reviewed in Kim et al., 1994 and Nelson and Craig, 1992). TRiC, although similar in organization to the Hsp 60 family, is not inducible and, unlike the organellar Hsp60 chaperonins that associate with a 10 kDa GroES-like co-chaperone (i.e., Hsp 10), TRiC does not have such a co-chaperone. In lies of an Hsp10 homologue in the eukaryotic cytosol, a number of co-chaperones have been identified that are termed co-factors A, B, C, D and E (Gao et al., 1993, Gao et al., 1994, and Tian et al., 1996). These function in conjunction with cytosolic chaperonin in the folding of α and β tubulin. Co-factor A, a 13 kDa molecule that exists as dimer, is not related to Hsp10 in either sequence or structure and it modulates the ATPase activity of cytosolic chaperonin during α and β tubulin folding. In contrast, the folding of actin and γ tubulin can be facilitated by chaperonin in the absence of these co-factors (Gao et al., 1992, Melki et al., 1993).

1.3.1 Functions of chaperonins

GroEL/ES

By far the most intensively studied of the chaperonins is the GroEL/ES complex. As mentioned previously, under normal conditions GroEL/ES is involved in the biogenesis of many E. coli proteins and, under stress conditions, GroEL/ES functions to protect and even rescue proteins from unfolding and aggregation (reviewed in Georgopoulos and Welch, 1993). Temperature-sensitive mutants of GroEL/ES have been isolated that produce a lethal phenotype at the non-permissive temperature. Upon shift to the non-permissive temperature there is little effect on translation in general, but there is a complete ablation of folding and assembly of citrate synthase, α-ketoglutarate dehydrogenase and polynucleotide phosphorylase (Horwich et
al., 1993). Since GroEL binds transiently to more than half of the newly synthesized proteins in E. coli, it is likely that there are other proteins that require this chaperonin system to fold correctly. These studies demonstrate the importance, in vivo, of GroEL/ES expression. Interestingly, even though there are examples of GroEL/ES interacting with nascent chains, this does not appear to be necessary for general translation to proceed normally (see Chaperone Cooperativity below).

What followed next was a long series of in vitro studies that examined the molecular and mechanistic basis for the in vivo effects of GroEL/ES on protein folding. Substrates in these studies included citrate synthase, DHFR (dihydrofolate reductase) (Ostermann et al., 1989 and Viitanen, 1991), rhodanese (Mendoza et al., 1991, Martin et al., 1993, and Weissman et al., 1994), luciferase (Flynn et al., 1993), ornithine transcarbamylase (Zhang et al., 1993 and Weissman et al., 1994), phage 22 tailspike protein (Brunschi et al., 1993), pre-β-lactamase (Bochkaereva and Girshovich, 1992), malate dehydrogenase (Hartman et al., 1993), glutamate synthase (Fisher, 1993 and Fisher, 1994), barnase (Gray et al., 1993, Gray and Fersht, 1993 and Corales and Fersht, 1995), cyclophilin (Zahn et al., 1994), RNA polymerase (Ziemienowicz, 1993) and more. It appears that GroEL/ES functions by binding to and stabilizing partially folded intermediates and releasing them for repeated attempts to fold productively. It is thought that GroEL binds with high affinity to unfolded or partially folded protein conformations thereby shifting the equilibrium away from aggregation (Walter et al., 1996). In addition to stabilizing unfolded and partially folded conformations, GroEL/ES can promote unfolding of mis-folded proteins. Instead of directly catalyzing unfolding, GroEL appears to function by binding to unfolded conformations thereby shifting the equilibrium from mis-folded to unfolded
conformations (Corrales and Fersht, 1996, Todd et al., 1996, Walter et al., 1996 and reviewed in Hartl., 1996). This allows a polypeptide additional chances to fold. In contrast to Hsp70, chaperonins bind substrate and release it into a central cavity where folding can proceed in the absence of other polypeptides, thus providing what is termed an “Anfinsen cage.” A detailed mechanism for GroEL/ES mediated folding is presented below as well as a discussion of the nature of the polypeptides that bind to the chaperonin.

Finally, there is evidence that in E. coli, GroEL/ES is involved in the degradation of proteins (Kandor et al., 1994, Kandor et al., 1995 and reviewed in Hayes and Dice, 1996). This process is ATP dependent and involves other chaperones as well as the Clp P or Lon proteases (reviewed in Hayes and Dice, 1996). The presence of both GroEL and GroES is necessary for efficient degradation and there has been one report of Trigger Factor being involved in the GroEL/ES mediated degradation of unfolded or mis-folded proteins (Kandor et al., 1995).

**Mitochondrial Hsp60/Hsp10 (cpn60/cpn10)**

Mitochondrial Hsp60/Hsp10 (cpn60/cpn10) is an essential mitochondrial chaperonin. In temperature sensitive strains of S. cerevisiae that harbour mutations in the Hsp60 gene, a shift to the non-permissive temperature arrests cell growth and a number of mitochondrial matrix proteins, including Hsp60, fail to fold and assemble (Hallberg et al., 1993). These proteins are correctly targeted to the mitochondrial matrix but, instead of folding, they aggregate as precursor forms that retain their mitochondrial import sequences. This phenomenon was not observed for a number of proteins targeted to the intermembrane space (Hallberg et al., 1993). These latter proteins are correctly targeted, processed and assembled at the non-permissive temperature. The
function of cpn60 is dependent on the expression of its co-chaperonin, cpn10 (Hohfeld and Hartl, 1994). Interestingly, the yeast cpn10 cannot be functionally replaced by GroES, in spite of the fact that cpn10 can replace GroES in E. coli (Rospert et al., 1993).

**Eukaryotic chaperonins (TRiC/TCP-1 complex/c-cpn)**

The functions of eukaryotic chaperonins (TRiC/TCP-1 complex/c-cpn) are not well understood. Whereas isolated complexes can bind, prevent aggregation, and promote ATP dependent renaturation of luciferase and tubulin in vitro, it is not yet clear whether TRiC acts as a general chaperone in vivo (Frydman et al., 1992). Only a small number of substrates have been identified for TRiC and its related complexes. Mammalian TRiC interacts with and assists in the folding of actin, α,β,γ tubulin and Hepatitis B virus capsid proteins (Lingappa et al., 1994, Melki et al., 1993, Tian et al., 1996, Gao et al., 1992, 1993, 1994). TRiC does not interact with an Hsp10 homologue but in some cases requires the presence of several unrelated co-chaperones (see section 1.3.1). In plants, a TCP1-related complex appears to be involved in the biogenesis of phytochrome and although no other substrates have been reported, its abundance suggests a broader role in the cell (Mummert et al., 1993). In yeast, the TCP20 and TCP1 subunits of the TRiC chaperonin, are required for viability suggesting that although TRiC may not be a general chaperone, its function is indispensable (Li et al., 1994).

**1.3.2 Structure**

The crystal structures of GroEL and GroES have been determined and refined to 2.8Å resolution (Figure 5A and B) (Briërg et al., 1994). GroEL is a cylinder-shaped molecule 150Å in
Figure 5
A. Schematic representation of the GroEL double doughnut viewed from the a. top, b. side and c. side including the GroES subunit. 1, 2 and 3 represent the apical, intermediate and equatorial domains respectively. Mutations affecting ATP binding and hydrolysis, GroES and peptide binding are indicated by the circle and ovals (Taken from Hartl, 1994).

B. A space-filling representation of the GroEL/ES molecule based on the crystal structure. Top and side view are shown along with their dimensions. The coloured areas represent the apical (red and purple), intermediate (orange and yellow) and equatorial domains (blue and green) of adjacent subunits of the GroEL multimer (Taken from Hartl, 1996).
height and 140Å wide. The central cavity is approximately 50Å wide and contains the polypeptide binding site. GroEL can be subdivided into 3 domains. The apical domain forms the flexible opening to the cylinder and provides the contact sites for both GroES and peptide. This is the least conserved domain in the GroEL structure. The intermediate domain, formed by a series of antiparallel β-strands, is a small region that appears to form a hinge that allows for movement of the apical and equatorial domains with respect to each other. The equatorial domain, consisting of tightly packed α-helices, is the site of interaction of the two heptameric rings of GroEL and also contains the ATP binding site. This is the most highly conserved region between Hsp60 homologues and through mutagenesis it was determined that the equatorial domain contains the ATPase activity (Kim et al., 1994). The structure also contains some openings at the side of the cylinder that may serve to allow nucleotide movement in and out of the cavity when GroES is bound. GroES is a dome-shaped molecule that binds to the apical domain via a flexible loop region that, upon docking to GroEL, becomes structured (Landry et al., 1992; Hunt et al., 1996). Critical mutations in this loop region result in loss of binding to GroEL and impaired GroEL function (Landry et al., 1992). There are large changes in GroEL conformation that accompany the binding of GroES (Chen et al., 1994a). Binding of GroES causes a conformational change in the GroEL ring that produces an enlarged 65X80Å cavity. The nature of the nucleotide bound to the equatorial domains can also affect the conformation of the apical domain dramatically. In cryo-electron microscopy studies, ATP binding at the equatorial domain results in large rotations of the apical domain (Roseman et al., 1996).

Recently, a crystal structure of the GroEL-GroES-(ADP) complex was published (Xu et al., 1997). Large movement of the cis ring’s apical and intermediate domains allow for GroES to
stabilize an enlarged “folding chamber” with ADP bound to the cis ring (shown in Figure 6, complex 5). In this complex the cavity is doubled in volume and the residues near the opening of the ring that interact with hydrophobic peptides (see next paragraph and Fenton et al., 1994) become buried either in the GroES interacting surface or else between GroEL subunits. As a result the polypeptide is displaced into the cavity. In addition, the walls of the cavity are lined with hydrophilic residues, favouring binding of non-native polypeptides and promoting commitment to folding to the native state. Concomitant conformational changes in the trans ring prevent binding of GroES to the trans ring until the cis complex is dissociated by ATP binding to the trans ring. Interestingly, the T4 bacteriophage encodes Gp31, a specialized co-chaperonin that can functionally replace GroES. Although this protein bears little sequence identity to GroES (14%), the crystal structure shows similar tertiary and quaternary structure and a similar effect on the size of the cavity within GroEL has been demonstrated (Hunt et al., 1997). Furthermore, this co-chaperonin can also alter the hydrophilicity of the cavity. Gp31 interaction with GroEL, but not GroES, promotes folding and assembly the T4 capsid protein, Gp23.

A series of mutational studies coupled with the structure of GroEL has led to a good understanding of the region and residues involved in binding to exposed polypeptides on substrate molecules (Fenton et al., 1994). In the absence of GroES two polypeptides can bind, one in each of the two cavities. It appears as though cooperative binding of polypeptide to more than one member of the heptamer stabilizes the interaction. The interacting surface on the apical domain of GroEL contains several essential hydrophobic residues that, when mutated to charged or polar residues, abolish peptide binding (Fenton et al., 1994). These mutations had no effect on the binding and hydrolysis of ATP but did abolish GroES binding, leading to the suggestion that
GroES and substrate bind on the same or overlapping surface of GroEL. It is hypothesized that binding of GroES to the apical domain is responsible for displacing the unfolded polypeptide and sending it into the central cavity where it can fold. The crystal structure of Xu et al., supports this hypothesis (see previous paragraph). The observed hydrophobic nature of the peptide binding domain, coupled with studies where the binding of individual amino acids to GroEL correlated with their hydrophobicity, leads to the conclusion that the nature of polypeptide binding is hydrophobic, much like Hsp70 (Richarme and Kohiyamma, 1994). In addition to binding hydrophobic amino acids, GroEL bound some polar amino acids found most often in α-helices. This would allow for GroEL to bind and stabilize amphipathic structures that exist in intermediate molten globule-like states.

In contrast to Hsp70s, that associate very early in protein folding (even co-translationally prior to the formation of significant secondary structure), the polypeptides bound by GroEL are, in general, more compact in structure. Substrates bound to GroEL contain significant α-helical content and exist in low energy molten globule-like conformations (Martin et al., 1991, Landry et al., 1992, Flynn et al., 1993, Hayer-Hartl et al., 1994, Robinson et al., 1994 and Tian et al., 1995). Landry et al. examined the structure of peptides bound to DnaK and GroEL by transferred NOE NMR methods. Consistent with DnaK acting at an earlier stage in folding, peptides bound to DnaK were in an extended conformation whereas peptides bound to GroEL were helical (Landry et al., 1992). In another study, α-lactalbumin folding intermediates were used to determine the conformers bound by GroEL. GroEL binding to a series of intermediates ranging from extended to native-like molten globules was examined (Hayer-Hartl et al., 1994). Interestingly, unfolding of α-lactalbumin to an extended conformation resulted in decreased
binding to GroEL as did complete folding. In addition, only some partially folded intermediates interacted with GroEL. Finally, hydrogen exchange kinetics, measured by electron spray mass spectrometry, of bound α-lactalbumin determined that it was bound to GroEL in a disordered but partially folded conformation similar to what is observed in early folding in vitro (Robinson et al., 1994). These results suggest that GroEL/ES function downstream of Hsp70 after initial collapse of the polypeptide to a compact unfolded state.

1.3.3 Reaction cycle of GroEL/ES

A detailed reaction mechanism has been proposed for GroEL/ES-mediated protein folding based on in vitro folding reactions using purified wild type and mutant GroEL molecules and stop-flow kinetic analysis of folding of OTC and rhodanese (Weissman et al., 1995, reviewed in Hartl, 1996). This model is shown in Figure 6. Briefly, binding of GroES to one end of the GroEL cylinder produces an asymmetrical ADP bound complex that can efficiently bind substrate in the opposite ring to form the “trans” complex (Figure 6, step 1). ATP binding and hydrolysis in the polypeptide binding ring cause dissociation of ADP and GroES from the opposing ring (steps 2 and 3). ATP and GroES will then re-bind to either end of the ring with equal probability and, in the case of the polypeptide-bound ring, will displace it into the GroEL cavity where folding can take place (step 4). If GroES binds at the opposite surface polypeptide will be released in a “non-native” state. ATP binding in the GroES bound side results in increased interaction of GroES and GroEL while ATP hydrolysis at the opposite end results in dissociation of GroES (step 6). Folded polypeptide can leave the cavity while incompletely folded or mis-folded polypeptide will re-bind for another round of folding (step 6). Binding by
Figure 6
Reaction cycle of the bacterial GroEL and GroES chaperonin system (Taken from Hartl, 1996).
GroEL can also randomly unfold mis-folded or “kinetically trapped” intermediates of proteins during this cycle and thus allow for additional folding chances (Todd et al., 1996). Folding by GroEL/ES occurs in an iterative fashion where each binding cycle leads to a fraction of the polypeptide folding to completion. The number of cycles required to fold a protein varies from 1 to many, depending on the protein’s intrinsic ability to fold in the 20-25 seconds that it is held in the GroEL cavity. The number of cycles will of course reflect the distribution of compact intermediate conformations bound by GroEL and their propensity to fold or produce “kinetically trapped” intermediates. For some proteins such as DHFR, 30% of the molecules can fold completely while in the cavity, in a single cycle, whereas for others like rhodanese, only 5% are released in a form that is committed to folding (Mayhew et al., 1996). In experiments where GroES release is inhibited by non-hydrolysable ATP analogues or with single ring GroEL mutants, complete refolding of rhodanese can occur while in the cavity, suggesting that for some proteins a single cycle may be sufficient (Weissman et al., 1996). In addition, binding and release of GroES to and from GroEL in the ADP bound state appears to be sufficient for folding of rhodanese in vitro (Hayer-Hartl et al., 1996). Folding is slow under these conditions, so ATP binding and hydrolysis appear to provide conformational changes in GroEL necessary for GroES cycling at a rate that is relevant to in vivo folding.

3.4 Hsp90 Family

Although Hsp90 is, constitutively, the most abundantly expressed heat shock protein its function and structural basis of interaction are not clearly understood. Several Hsp90
homologues have been identified including Hsp90 in the cytosol and grp94 in the ER (reviewed in Georgopolous and Welch, 1993 and Johnson and Craig, 1997). Hsp90 molecules bind ATP and evidence was recently presented for peptide-stimulated ATPase activity of grp94, although the relevance of this to Hsp90 function is not clear (Li and Shrivastiva, 1993). In addition, cytosolic Hsp90 and grp94 can be phosphorylated by casein kinase II (Cal a and Jones, 1994 and Shi et al., 1994). The in vivo significance of phosphorylation is unclear but it has been suggested that phosphorylation regulates Hsp90 activity.

While there appear to be many substrates for the Hsp70 and 60 families there are but a few specific substrates for Hsp90 identified in vivo, i.e., steroid hormone receptors (Dalman et al., 1989, Picard et al., 1990, and Schowalter et al., 1991), viral oncogenes (Oppermann et al., 1981, Schuh et al., 1985 and Hutchison et al., 1992), cellular kinases (Rose et al., 1987, Miyata and Yahara, 1992, and A ligue et al., 1994) and cytoskeletal components such as actin and tubulin (Sanchez et al., 1988, and Miyata and Yahara, 1991). The two best characterized substrates are steroid hormone receptors (e.g., glucocorticoid receptor) and pp60v-src. In the case of steroid receptors, hsp90 is required for the receptor to achieve a hormone binding conformation and is required subsequently to maintain this conformational state. In the absence of steroid hormone, the receptor remains associated with the hsp90 chaperone complex and is kept in a state that is inactive for DNA binding. Binding of hormone to the receptor causes a translocation of the receptor/chaperone complex to the nucleus in a process that appears to require the presence of hsp90 (reviewed in Pratt, 1993). Once in the nucleus the hormone-associated receptor dissociates from hsp90 and is active for binding to specific DNA elements, thus inducing or repressing transcription of hormone regulated genes.
Hsp90 involvement in pp60\textsuperscript{V-Src} biogenesis follows somewhat similar principles as the steroid hormone receptor mechanism. Almost immediately after its synthesis, pp60\textsuperscript{V-Src} interacts with Hsp90 and, in this complex, the kinase remains inactive (Brugge et al., 1981 and Oppermann et al., 1981, reviewed in Brugge, 1986). Hsp90 appears to escort pp60\textsuperscript{V-Src} to the plasma membrane at which time pp60\textsuperscript{V-Src} dissociates from Hsp90, is inserted into the membrane and the kinase is activated (Brugge et al., 1983). This ensures that the kinase domain is not prematurely activated prior to reaching its site of function. Not only is Hsp90 responsible for keeping pp60\textsuperscript{V-Src} inactive but it is also required for stabilization and proper activation of pp60\textsuperscript{V-Src} kinase activity (reviewed in Brugge, 1986). The importance of Hsp90 in pp60\textsuperscript{V-Src} function was demonstrated in yeast cells where decreased expression of Hsp90 resulted in a dampening of the pp60\textsuperscript{V-Src} toxicity that is observed when wild type levels of Hsp90 are present (Xu and Lindquist, 1993).

The observation that Hsp90 binds to and prevents premature activation of its substrates, coupled with its ability to translocate with its substrate to its site of action, suggests a common mechanism for Hsp90 functions. By exploiting yeast genetics, one group was able to show that several different Hsp90 mutants could affect both glucocorticoid receptor and pp60\textsuperscript{V-Src} function suggesting that Hsp90\textsuperscript{S} mechanism of action was similar for both substrates (Nathan and Lindquist, 1995). Further support for this view comes from data implicating Hsp90 interaction with the cytoskeleton (via Hsp56) as a means to regulate both steroid receptor and pp60\textsuperscript{V-Src} targeting to their site of function (reviewed in Pratt, 1993).

The ability of Hsp90 to function as a molecular chaperone has been demonstrated for a number of substrates. For citrate synthase and antibody Fab fragments, Hsp90 could prevent
aggregation and support efficient refolding in the absence of ATP (Weich et al., 1992). For other substrates such as luciferase and β-galactosidase, hsp90 has been shown to prevent aggregation and keep substrates in a folding competent state but requires the subsequent addition of ATP, Hsp70 and its co-chaperones to reactivate the enzymes (Schumacher et al., 1994, and Freeman and Morimoto, 1996). Whether these chaperone functions are relevant in vivo is not yet clear.

From the limited number of substrates identified in vivo it is difficult to determine whether hsp90 acts as a general molecular chaperone.

In vivo, hsp90 is not usually found as a monomer but exists as a dimer when active for substrate binding. Mutational analysis of Hsp90 has identified the 49 carboxy terminal amino acids as the region responsible for dimerization (Minami et al., 1994). In yeast this region is not only required for dimer formation but is required for normal functioning of Hsp90 (Minami et al., 1994). In addition to dimer formation, Hsp90 exists as a large complex with several other cytosolic proteins. These include members of the Hsp70 chaperone machinery as well as small heat shock proteins and cyclophilins. All in all, nine different components of the Hsp90 complex have been characterized in mammalian cells (reviewed in Pratt, 1993, and Johnson and Craig, 1997). These include the Hsp90 dimer, Hsp70, Hip, p60/Hop (Hsp70/90 organizing protein), p48/Hip, Hdj1, p23 and any one of the large peptidyl prolyl cis-trans isomerases (PPI) - FKBP52/p50/Hsp56, FKBP54, or cyclophilin 40 (CyP40). Most of these proteins interact directly with specific regions of Hsp90 to form a large complex in the cytosol. This complex is pre-assembled in the cytosol and can be isolated in the absence of steroid receptor or other substrates (Hutchison et al., 1994a and 1994b). For this reason it is referred to as the "apo-receptor complex", the "super-chaperone" complex or "foldosome". In vitro folding assays determined
that some of the components of this "super-chaperone complex" can act individually as molecular chaperones by binding to and stabilizing folding intermediates, i.e., CyP40, FKBP52, and p23 can act independently to bind and stabilize folding intermediates (Bose et al., 1996). While they do not appear capable of completely refolding unfolded substrates, these molecules can maintain them in a folding competent conformation. It is interesting that the large PPI's have appeared to evolve chaperone domains capable of binding folding intermediates independent of their isomerase function. The Hsp90 complex appears to have evolved such that all or most of the components contribute to protein folding. This is supported by the observed requirement for both Hsp70 and p23 expression to achieve efficient interaction of receptors with Hsp90 in reticulocyte lysates (Hutchison et al., 1994a and 1994b and Johnson and Toft, 1994). In addition, the efficient reactivation of luciferase in reticulocyte lysate and in a purified folding assay requires ATP as well as Hsp70 and Hsp90 in a 1:2 ratio (Schumacher et al., 1994).

Finally, the Hsp90 super-complex may mediate degradation of associated polypeptides. This is suggested by a study in which the drug ansamycin can inhibit refolding and in turn promote degradation of denatured firefly luciferase. This degradation appears to be the result of preventing dissociation of luciferase from the Hsp90 complex (Schneider et al., 1996). This drug-induced shift of Hsp90 to a degradation chaperone is artificial and may not reflect an in vivo function for the Hsp90 complex (under either normal conditions or during stress). Further studies are necessary to determine what role, if any, Hsp90 complexes have in protein turnover.  

1.5 Other Chaperones

In addition to the Hsp70, 60 and 90 families discussed above there are a growing number
of proteins that also behave as molecular chaperones. Not much is understood about their structures or mechanisms of action but there is evidence that each acts as a molecular chaperone. Table II outlines some of these proteins, their subcellular localization, substrate molecules and functions. In some cases the molecules are presented as a family of chaperones by virtue of functional and sequence similarity. The definition of a molecular chaperone is quite broad and as a result this Table is by no means complete. There are a great number of proteins involved in protein biogenesis and trafficking that are not classically termed molecular chaperones but may, in the course of their action protect nascent proteins from aggregation, misfolding and degradation.

1.6 Cooperativity Between Chaperone Machines

There are many examples where chaperone machines function in concert to fold newly synthesized or newly translocated polypeptides. While redundancy has been observed \textit{in vivo} using mutants in one of the two main chaperone systems (Hsp70 and Hsp90), it is more likely that under normal conditions there is a co-ordinated and sequential action by chaperone machines that provides optimal folding. This co-ordinated action is sometimes dictated by topology such as in the mitochondrial import and folding pathway where Hsp 70 molecules on either side of the membrane must act first to promote translocation which is then followed by Hsp 60/10 activity (Figures 1 and 7). Another example of such co-ordinated action is in the targeting, import and folding of polypeptides in the ER (reviewed in Brodsky, 1996). As with nuclear-encoded mitochondrial proteins, synthesis of proteins bound for the secretory pathway is initiated and sometimes completed in the cytosol prior to translocation into the ER where post-translational modifications, folding, assembly and subsequent transport to the Golgi
### Chaperone Family

<table>
<thead>
<tr>
<th>General Features</th>
<th>Examples</th>
<th>Species</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Periplasmic Chaperones</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>PapD</td>
<td><em>E. coli</em></td>
</tr>
<tr>
<td></td>
<td>PapC</td>
<td></td>
</tr>
<tr>
<td></td>
<td>FimC</td>
<td></td>
</tr>
</tbody>
</table>

- **Periplasmic Chaperones**
  - **PapD**
    - Immunoglobulin-like molecules
    - Soluble in periplasm
    - C-terminal domain is CD4-like
    - Crystal structure solved: 2 Ig-like domains shaped like a boomerang with the substrate binding site formed by the cleft between the two domains
    - PapD involved in the assembly of P pili
  - **PapC**
    - Membrane protein of the periplasm
    - Called a molecular usher because it shuttles plus subunits (Pap G, K, F, E) from chaperone complexes to the assembling pilus.
  - **FimC**
    - Homologous to PapD
    - Involved in the assembly of type 1 pili

**Table II**

This table summarizes work on a number of chaperone molecules that are not part of the Hsp70, 60 or 90 families. The species distribution, subcellular localization, preferred substrates and functions are included.
<table>
<thead>
<tr>
<th>Clp/Hsp100</th>
<th>E. Coli</th>
<th></th>
</tr>
</thead>
</table>
| Clp A     | • 44 kDa enzyme, forms dimers and then hexamers of dimers  
            • binds ATP and associates with the ClpP protease  
            • heat shock inducible expression  
            • targets proteins for proteolysis  
            • PI DNA replication activated by targeting of RepA for degradation  |
| Clp X     | • heat shock inducible expression  
            • binds ATP and associates with ClpP  
            • protects HO protein from heat induced aggregation and can disaggregate aggregates once formed  
            • protection from aggregation requires only ATP(ADP)6 binding but disaggregation requires ATP hydrolysis  |
| Clp B     |  |
| Clp Y/HslU|  |

| Yeast     |  |
|-----------|  |
| Hsp104    | • homologous to ClpA but does not appear to target proteins for degradation  
            • reactivates insoluble aggregates following heat damage and can reanimate proteins once disaggregated (e.g., firefly luciferase)  
            • low levels can provide thymine rescue  
            • cytosolic localization  |
| Hsp78     | • soluble mitochondrial matrix protein, nuclear encoded gene  
            • homologous to Hsp104 and ClpB  
            • 91 kDa basic protein  
            • non-lethal phenotype  
            • can partially substitute for mHsp70 loss  
            • prevents aggregation of proteins in the mitochondria but unlike mHsp70 cannot target misfolded proteins to PIM 1 ATP protease.  |

| Plant     | Clp C   |  |

41
## Cyclophilins
(cyclosporin A binding proteins)

### General features:
- members of the immunophilin family of proteins
- more than 20 members in a number of species
- considered folding enzymes, have peptidyl-prolyl cis-trans isomerase activity (PPIase)
- PPI catalysis increases cis-trans isomerization 300X in vitro, 10X in vivo
- folding of carbonic anhydrase by cyclophilins can occur via a PPIase dependent and independent pathway suggesting multifunctional chaperone functions.

<table>
<thead>
<tr>
<th>Sub-Family</th>
<th>CyP</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>A</td>
<td>ubiquitous, 18 kDa cytosolic protein, found in all cell types</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- 40kDa cyclophilin associated with Hsp90/steroid receptor complexes</td>
</tr>
<tr>
<td></td>
<td>B</td>
<td></td>
</tr>
<tr>
<td></td>
<td>C</td>
<td></td>
</tr>
<tr>
<td></td>
<td>D</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td></td>
<td>vertebrate enzymes, 18 kDa</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- CyP-B, C have ER and Golgi distribution, some secreted</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- CyP-C is predominantly a kidney enzyme</td>
</tr>
<tr>
<td></td>
<td></td>
<td>yeast, ER and Golgi localized</td>
</tr>
<tr>
<td>3</td>
<td>R</td>
<td>N. crassa, mitochondrial protein involved in preprotein translocation and folding in cooperation with mHsp70 and Hsp60</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td></td>
<td>has both signal sequence and C-terminal hydrophobic sequence</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- anchored in the ER membrane</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- involved in photoreceptor biogenesis, increases rhodopsin synthesis by 10X, required for transport of rhodopsin through secretory pathway - Quality Control?</td>
</tr>
<tr>
<td></td>
<td></td>
<td>topology like nina A</td>
</tr>
<tr>
<td>5</td>
<td></td>
<td>nuclease involved in apoptotic DNA degradation in lymphocytes</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- 18kDa protein</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CyP-A, B, C also found to have nuclease activity like NUC 18</td>
</tr>
</tbody>
</table>

### NUC 18

- nuclease involved in apoptotic DNA degradation in lymphocytes
- 18kDa protein
- CyP-A, B, C also found to have nuclease activity like NUC 18
**FKBP's**  
(FK506/rapamycin binding proteins)

**General features:**  
• members of the immunophilin superfamily  
• display prolyl-isomerase activity (i.e. are PPIases)

<table>
<thead>
<tr>
<th>Mammalian</th>
<th>Yeast</th>
</tr>
</thead>
<tbody>
<tr>
<td>FKBP52/59 (p56, Hsp56) and FKBP54</td>
<td><strong>FKR2 (FKBP13)</strong></td>
</tr>
</tbody>
</table>
| • associate with Hsp90 in the inactive steroid receptor complex  
• cytosolic localization | • *S. cerevisiae homolog* of FKBP-13  
• localized to the ER lumen, expression up-regulated at the transcriptional level by heat shock and accumulation of unfolded proteins in the ER (tunicamycin treatment, sec32-8 mutant)  
• **FKR2** contains the 21bp UPR element |
| FKBP-12 | |
| • binds to the Ca\(^{2+}\) release channel in skeletal muscle  
• ER localization | |
| s-cytochrome c | |
| • co-localizes with calreticulin  
• ER localization | |
<table>
<thead>
<tr>
<th>Crystallins/ small Hsp's</th>
<th>Plant</th>
<th>Mammalian</th>
</tr>
</thead>
<tbody>
<tr>
<td>General Features:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>• 16-30kDa in size, can be phosphorylated in response in stress and mitogens</td>
<td>P. Sativum (pea)</td>
<td>Hsp 27</td>
</tr>
<tr>
<td>• upregulated during stress</td>
<td>G. max (soya)</td>
<td>Hsp 25</td>
</tr>
<tr>
<td>• 70 amino acid C-terminal domain conserved among all family members, possible function for oligomerization, N-terminal region not conserved</td>
<td></td>
<td></td>
</tr>
<tr>
<td>• exist in two populations of high molecular weight aggregates 1-200-500 kDa</td>
<td></td>
<td></td>
</tr>
<tr>
<td>• thought to maintain cell structure during stress</td>
<td></td>
<td></td>
</tr>
<tr>
<td>• no enzymatic activity? Do not bind ATP</td>
<td></td>
<td></td>
</tr>
<tr>
<td>• found in many eukaryotes, yeast (1. bus fci), human, mouse, etc.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>• plants have a large number of small Hps that can be grouped into 4 gene families by virtue of N-terminal similarity and sub-cellular localization</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chaperone Functions:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>• in vitro bovine α crystallin, helHsp27, multihsp25 can prevent heat induced aggregation and stabilize α-glucosidase and citrate synthase. Promotes refolding of denatured molecules as well. ATP independent.</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>&gt;76% amino acid identity to G. max</td>
<td></td>
</tr>
<tr>
<td></td>
<td>signal sequence, ER retention motif</td>
<td></td>
</tr>
<tr>
<td></td>
<td>hamster Hsp27 expressed in NIH3T3 cells produces resistance to lethal heat shock</td>
<td></td>
</tr>
<tr>
<td></td>
<td>thermal tolerance requires high levels at time of heat shock not via induced pool</td>
<td></td>
</tr>
<tr>
<td></td>
<td>partial protection from actin depolymerizing effects of cytochalasin D suggests role in stabilization of actin fibers</td>
<td></td>
</tr>
<tr>
<td></td>
<td>high levels of Hsp27 found at leading edge of polarized fibroblasts (lamellopodia and ruffles), suggests role in actin polymerization</td>
<td></td>
</tr>
<tr>
<td></td>
<td>overexpression of a human non-phosphorylatable mutant results in decreased F-actin at the cortex and decreased proapoptosis, opposite effect with wild type hsp27</td>
<td></td>
</tr>
<tr>
<td></td>
<td>if heat cells with agonists that induce phosphorylation are shift from aggregate to 70 kDa molecules, stress induces the same shift - significance?</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Hsp25 is phosphorylated by MAPKAP kinase 2, not essential for thermotolerance activity of Hsp25. Non-phosphorylated mutants same level of thermotolerance</td>
<td></td>
</tr>
<tr>
<td></td>
<td>phosphorylation of Hsp25 abolishes actin polymerization activity of Hsp 25 monomers and oligomers in vitro</td>
<td></td>
</tr>
<tr>
<td></td>
<td>chaperone activity in eye lens, prevents heat-induced aggregation and scattering of α and β crystallins, prevents cataracts by preventing β crystallins cleavage</td>
<td></td>
</tr>
<tr>
<td></td>
<td>mediate intermediate filament (IF) assembly, α crystallins co-isolated with IF</td>
<td></td>
</tr>
<tr>
<td></td>
<td>in vitro α crystallins inhibit assembly of GFAP and vimentin by binding along the length of GFAP in regular spaced arrays. ATP independent.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>α crystallins can autophosphorylate serine residues.</td>
<td></td>
</tr>
<tr>
<td>Misc. Chaperones</td>
<td>General Features</td>
<td></td>
</tr>
<tr>
<td>-----------------</td>
<td>-----------------</td>
<td></td>
</tr>
</tbody>
</table>
| YTA10-12 Complex (E. Coli) | • inner mitochondrial membrane complex, 850 kDa complex made up of Yta10p and Yta12p  
• AAA protease family member ("AAA" ATPases associated with a variety of cellular activities, ATP dependent metalloproteases)  
• Yta10p related to D. cervinia Yme1p, chloroplast PfH, K. Cde FtsH and Lso protease.  
• Yta12p does not display protease activity  
• YTA10-12 complex in mitochondria required for assembly of ATP synthase and degradation of unassembled subunits of mitochondrial complexes that reside in the inner membrane. |
| UreD (K. aerogenes) | • UreD required for assembly of the nickel metallocentre of K. aerogenes urease complex (Ure A,B and C)  
• UreD binds to urease monoprotein and facilitates nickel binding at which time it dissociates. |
| PuIS (K. oxytoca) | • periplasmic protein of gram-negative bacteria that appears to be necessary for association of PuID (needed for pullulanase secretion) with the outer membrane.  
• PuID and PuIS are both part of the GSP (general secretion pathway) complex in these bacteria. |
| Lim (limA gene) (E. coli) | • Lim required for lipase (LipA/lip) secretion in a variety of bacterial expression systems.  
• Lim not involved in transcription or translation but is present required for expression of active lipase.  
• In vitro denaturation and subsequent renaturation of lipase requires Lim suggesting that it is a 33 kDa protein acts as a chaperone, supported by observation that Lim directly interacts with lipase |
| SycE, SycH, SycD (P. mirabilis) | • putative cytosolic molecular chaperones involved in Yop secretion (Sec-independent secretion pathway) |
| MSP (mammalian) | • cytosolic chaperone capable of unfolding, disaggregating and targeting of mitochondrial precursors to the mitochondrial import pathway  
• ATPase activity of MSP required for activity |
| Hop1p, APG-1/Hop70p (hamster/mouse/human) | • genes implicated in spermatogenesis  
• appear to be expressed both constitutively and in response to a "low temperature heat shock"  
• maximal induction - 32-37°C shift, not the traditional 37-42°C shift |
| SecB (E. coli) | • E. coli cytosol, involved in binding to and targeting of newly synthesized proteins for secretion  
• X. laevis nucleus, binds histones, involved in nucleosome assembly  
• ER lumen, involved in X ecoli receptor assembly |

45
Figure 7
Protein folding in the mitochondrial matrix. The mt-Hsp70 chaperone machine (consisting of Hsp70, MDJ (DnaJ homologue) and MGE (GrpE homologue)) interact to promote translocation of the polypeptide after which the substrate is released for either an additional round of Hsp70 binding (not shown) or transfer to Hsp60 (cpn60/cpn10). R is the receptor for the preprotein, GIP is the general insertion pore, MIM is the mitochondrial inner membrane import machinery, and MPP is the processing peptidase that removes the mitochondrial targeting signal (Taken from Pfanner et al., 1994).
occurs. Both cytosolic and ER chaperones are involved in translocation of newly synthesized polypeptides into the ER (Figure 8), an aspect of ER chaperone function that will be discussed in more detail in a subsequent section. In addition to facilitating translocation, there are a growing number of examples where ER chaperones, such as BiP (ER Hsp70), calnexin, calreticulin, grp94 (ER Hsp90), Erp72 and Hsp170, act concurrently and in succession to fold and assemble nascent and fully translocated proteins, i.e., BiP, calnexin and calreticulin are involved in the biogenesis of human class I histocompatibility molecules (reviewed in Watts and Williams, 1995 and Suh et al., 1996), BiP and grp94 in immunoglobulin biogenesis (Melnick et al., 1992, Melnick et al., 1994, and reviewed in Melnick and Argon, 1995), calnexin and calreticulin in influenza haemagglutinin biogenesis (Hebert et al., 1996), BiP and calnexin in VSV G glycoprotein biogenesis (Hammond and Helenius, 1994), BiP, grp94 and PDI in thyroglobulin biogenesis (Kuzmatsov et al., 1994), BiP, grp94, Erp72 and Grp170 in oligomeric thyroglobulin and thrombospondin biogenesis (Kuzmatsov et al., 1994, Kuzmatsov et al., 1997). It is interesting to note that these molecules may also be involved in the degradation of mis-folded or mis-assembled molecules (see BiP section below).

One of the best studied examples of chaperone cooperativity is in the folding of newly synthesized polypeptides in the cytosol of prokaryotes and eukaryotes (Figure 9). (reviewed in Bukau et al., 1996 and Hartl, 1996) There is evidence that during the synthesis of cytosolic proteins, eukaryotic NAC (nascent chain associated complex) and prokaryotic trigger factor (prolyl isomerase) bind to the nascent polypeptide first. Subsequent chaperone binding occurs in an ordered and sequential fashion beginning with the binding of the Hsp70/40 chaperone machine which is then followed by interaction with cytosolic chaperonins (Frydman et al.,
Figure 8
A. ER translocation pathways in mammals (left) and yeast (right). In yeast, post-translational translocation requires cytosolic Hsp70/Hsp40 (Ssa1p and Ydj1p) for targeting and translocation. The secretory protein initially interacts with the Sec67p/Sec71p/Sec72p complex after which it is transferred to the Sec61p complex which forms the translocation pore. For translocation to proceed the polypeptide must interact with the ER lumenal Hsp70, BiP. Sec63p is a putative DnaJ homologue that is also required for translocation. In the mammalian system (left), most secretory proteins are translocated co-translationally. This process is initiated by interactions between an N-terminal signal peptide on the nascent polypeptide and signal recognition particle (SRP). This complex is then targeted to the ER membrane via interaction of SRP and the signal recognition particle receptor (SRα and SRβ). Subsequently, SRP dissociates and the ribosome/nascent chain complex is transferred to the translocon (Sec61p/TRAM), described in greater detail in part B of this Figure (Taken from Corsi and Schekman, 1996).

B. Schematic of the eukaryotic translocon complex in SRP mediated targeting (left) and translocation (right). SRP receptor (SR), signal peptidase (SP), ribosome receptor (RR) and oligosaccharyl transferase (OST) are shown. Sec61 with two additional low molecular weight proteins are thought to form the translocation pore. TRAM is a protein that is required for efficient translocation and is found to make contacts with the translocating polypeptide. BiP binding is required for translocation to proceed. The exact architecture of the translocon is not known (Taken from Gilmore, 1993).
Figure 9
Cooperative interactions between nascent polypeptide chains, nascent chain binding proteins [NAC, Trigger Factor (TF)] and chaperones in the cytosol of bacteria (left) and eukaryotes (right) (Taken from Hartl, 1996). TriC is the chaperonin identified in the eukaryotic cytosol.
A portion of the mammalian Hsp70 pool co-purifies with active ribosomes. Hsp73 interacts in a puromycin sensitive fashion, whereas Hsp72 binding is unaffected by puromycin. This suggests that Hsp72 may be involved in maintaining the ribosomal structure, while Hsp73 interacts directly with the nascent chain (Beck and Maio., 1994). During the translation of firefly luciferase in a reticulocyte lysate, TRiC, Hsp 70 and Hsp 40 are found in association with the active ribosome and Hsp70 appears to be required at the time of synthesis for proper folding to occur. Luciferase translation was not affected by immunodepletion of cytosolic Hsp70, but folding was abrogated, and so it appears as though interaction with chaperones is only required for the efficient folding of nascent polypeptides. The N-terminal portion of a nascent polypeptide can be folding within the TRiC cavity while Hsp70/40 retain the emerging residues in a folding competent state. Additional evidence for sequential addition of chaperones to nascent chains comes from the observation that Hsp70 and Hsp40 can associate with both short and long puromycyl fragments (prematurely released polypeptide fragments) while TRiC only associates with longer fragments (Frydman et al., 1994).

Although there is evidence for this ordered addition of molecules to the growing nascent chain these data must be interpreted with caution. In the studies where association of nascent chains with chaperones was observed, the nascent chain was stalled artificially on the ribosome (reviewed in Bukau et al., 1996). It was only under these conditions that chaperones were required for folding of model substrates in E. coli (Kudlik et al., 1994). In cases where translation was ongoing, no direct interaction with nascent chains was observed and model proteins folded in the absence of E. coli chaperones. The stalling of nascent chains may contribute to misfolding and as a result may artificially recruit chaperones to the site of
synthesis. This may also be the case for the results using puromycyl fragments to measure direct interactions. These fragments are incomplete polypeptides and as a result are more likely to be unfolded than full length molecules (reviewed in Bukau et al., 1996). Under these conditions unfolded molecules accumulate and are present for extended periods. It would therefore not be surprising to find chaperones recruited to these fragments to prevent aggregation. The differences in length preference observed between Hsp70/40 and TRiC may be the result of Hsp 70/40 having a higher affinity for short fragments and TRiC having a requirement for longer fragments to promote cooperative binding to each subunit of its ring. Although Hsp70 and Hsp40 molecules have been found to be associated with active ribosomes there is some evidence that this interaction is directed to the ribosome and not necessarily to the nascent chain. Furthermore, yeast and bacterial mutants lacking Hsp70 and Hsp40 do not display a dramatic translational or folding defect at normal temperatures. It is only when all major chaperones are missing that a partial defect in folding of newly synthesized proteins is observed. One of the most significant arguments against chaperonin association with translating proteins is the amount of either GroEL/ES or TRiC complexes in the cytosol. These complexes do not appear to be abundant enough to interact with all nascent chains during protein synthesis (Lorimer, 1996). It is more likely that there are a limited number of substrates that require the chaperonin machinery at the time of synthesis. This is supported by the finding that TRiC complexes are almost exclusively associated with actin and tubulin (Lewis et al., 1996). For many proteins the presence of NAC/trigger factor may be sufficient for folding during synthesis and a requirement for chaperonin may occur only after synthesis is complete.
1.7 Protein Biogenesis in the Eukaryotic Secretory Pathway

As mentioned previously, the focus of this thesis is the function of calnexin and calreticulin, putative molecular chaperones of the ER. To place this work in proper context, the process of protein translocation into the ER, post-translational modifications that occur in this organelle, packaging of proteins for export to the Golgi, and the function of various ER chaperones will now be discussed.

1.7.0 Secretory pathway and glycoprotein biogenesis

In eukaryotic cells the vast majority of proteins destined for cell surface expression, lysosomal targeting and secretion must transit through the endoplasmic reticulum (ER) (reviewed in Gilmore, 1993, Schatz and Dobberstein, 1996, Cori and Schekman, 1996 and Andrews and Johnson, 1996). Translation of mRNA encoding secretory proteins begins in the cytosol on free ribosomes but, in most cases, synthesis is retarded by the binding of SRP (signal recognition particle) to the N-terminal ER targeting motif known as the signal sequence. The SRP is a GTPase made up of 6 polypeptides (9,14,19,54,68, and 72 kDa) on a 7S RNA scaffold. The SRP 54 subunit is responsible for recognition and binding to the hydrophobic signal sequence, a region which is usually poorly recognized by nascent chain associated proteins (NAC's). The other subunits are responsible for retarding translation (SRP9, 14), targeting, and GTP-mediated interaction with the SRP receptor (SR). The nascent chain is released from SRP once it binds to the SRP receptor on the ER membrane. SRPR subsequently transfers the nascent chain complex to the translocation machinery (see BiP discussion below for more details), a large protein complex collectively known as the translocon (Figure 8A and 8B). Binding of the
ribosome to the cytosolic face of the translocon appears to act as a signal to open a proteinaceous aqueous pore in the membrane (gated channel) through which the nascent chain threads. This co-translational process allows for the unidirectional translocation of the polypeptide without the loss of integrity of the ER compartment. There are examples of alternative targeting and translocation pathways to the ER, particularly in yeast, that include post-translational targeting involving cytosolic Hsp70 (Figure 8A, yeast). It is not exactly clear at this point how transmembrane sequences are integrated into the ER membrane as they are encountered in the translocon but it seems that a membrane exposed segment opens laterally to allow transfer (reviewed in Andrews and Johnson, 1996). Another important aspect of ER translocation is the presence of chaperone molecules that are required for efficient translocation (see ER Chaperone section).

In addition to the molecules that interact with the signal sequence and those involved in forming the translocation pore, the translocon is associated with several enzymatic activities. The signal peptidase complex removes cleavable N-terminal signal sequences co-translationally while uncleaved signal anchor sequences act as transmembrane segments in the final protein structure. Most relevant to the work presented here is the association of the multi-subunit oligosaccharyl transferase with the translocon (Figure 8B). The vast majority of proteins that are translocated into the ER contain the sequence Asn-X-Ser/Thr (where X is not Pro) that acts as a signal for the addition of a core oligosaccharide to the Asn residue via an N-glycosidic bond (reviewed in Kornfeld and Kornfeld, 1985 and Hirschberg and Schneider, 1987). The number of N-glycans added to a polypeptide can vary from one to many. In cell-free translation/translocation studies it was found that glycosylation occurs co-translationally and the
position of the Asn-X-Ser/Thr site relative to the membrane is important for efficient transfer of oligosaccharide (Nilsson and von Heijne, 1993). The pre-assembly of the oligosaccharide on a dolichol carrier and its translocation into the ER for additional modification are shown in Figure 10A and 10B, along with the processing events that occur after the addition of the oligosaccharide to the polypeptide. Processing occurs in a sequential "assembly-line" fashion where the product of the previous step is the substrate for the next and availability of processing enzymes in the Golgi determines which modifications occur. Unlike the template-driven process of DNA, RNA and polypeptide synthesis, the addition of oligosaccharide to polypeptide does not result in homogeneous oligosaccharide structures once processing is complete. Even within a single polypeptide the oligosaccharides at different sites may differ in structure and a structure present at a given site may vary between copies of the polypeptide chain. This results in thousands of possible structures whose differences may or may not be physiologically relevant. What is clear though is that N-glycosylation provides the potential for increased numbers and types of recognition domains and interacting surfaces on a protein.

Glycoproteins and glycolipids are involved in a great number of cellular processes ranging from protein folding, protein targeting, mucosal surface protection, cell-cell interactions, growth receptor-ligand interactions and cell signalling, cell adhesion, removal of damaged or old proteins from the circulation and even apoptosis. Many of these functions require the presence of lectin molecules (molecules that non-covalently associate with sugar residues on glycoproteins and glycolipids) on cell surfaces and/or intracellular compartments (reviewed in Rademacher et al., 1988, Gahmberg and Tolvanen, 1996, and Lasky, 1995). Lectin biology and structure will not be discussed further but have been extensively reviewed
Figure 10

A. Synthesis and translocation of the dolichol-linked core oligosaccharide into the ER. The sugar donors in this reaction are nucleotide sugars and the reactions catalysed by glycosyltransferases. After assembly of Man₅GlcNAc₂-PP-Dolichol in the cytosol (top), the oligosaccharide is translocated into the ER lumen where additional sugar residues are added to produce Glc₅Man₅GlcNAc₅-PP-Dolichol. This core oligosaccharide is transferred, en bloc, to newly synthesized polypeptides (Modified from Darnell et al., 1986).

B. Pathway for transfer and processing of N-linked oligosaccharide. The ovals represent the dolichol carrier, squares - N-acetylglucosamine, open circles - mannose, triangles - glucose, closed circles - galactose or sialic acid. Successive processing of the oligosaccharide occurs as the glycoprotein moves through the ER and Golgi. In addition, the localization of the major processing glycosidases and glycosyltransferases is shown. (Taken from Austen and Westwood, 1991)
elsewhere (Drickamer and Taylor, 1993, Rini, 1995 and Weis and Drickamer, 1996). However, the involvement of lectins in protein biogenesis within the secretory pathway is particularly relevant to the focus of this thesis, the function of calnexin and calreticulin, and hence will be discussed below.

1.7.1 Glycoproteins and lectins in the secretory pathway - Emerging functions

The role of glycosylation in efficient protein folding and assembly has been dramatically demonstrated by the deleterious effects of glycosylation inhibitors such as tunicamycin and of processing inhibitors such as castanospermine and deoxynojirimycin (reviewed in Elbein, 1987, Paulson, 1989, and Helenius et al., 1992). Treatment of cells with tunicamycin results in the accumulation of unfolded proteins in the ER and eventually leads to a complete loss of translation/translocation of secretory proteins (reviewed in Elbein, 1987). Drugs such as castanospermine have been shown to cause increased turnover of some proteins such as class I histocompatibility molecules due to the misfolding and retention of these molecules in the ER (Moore and Spiro, 1993). It has been postulated that the addition of a hydrophilic carbohydrate chain to newly synthesized polypeptides plays a major role in the protein folding pathway by preventing aggregation of molecules via exposed hydrophobic regions in the unfolded molecule. From these early studies it became clear that glycosylation of many newly synthesized polypeptides is an important component in their biogenesis and transit through the ER.

Until a few years ago the only direct example of a lectin being involved in the trafficking of glycoproteins in the secretory pathway was the mannose-6-phosphate receptor. This receptor recognizes phosphorylated mannose residues on lysosomal proteins as they transit the Golgi and
directs them to the lysosome (reviewed in Kornfeld and Mellman, 1989). Defects in the processing of oligosaccharide to produce the mannose-6-phosphate moiety results in mis-targeting of a subset of lysosomal enzymes to external secretions, resulting in defects in lysosomal function (e.g., I-cell disease).

There is increasing evidence that lectins may also function in the vesicular transport of proteins between different compartments in the secretory pathway (reviewed in Fiedler and Simons, 1995). In order to maintain the integrity of each compartment (i.e., ER, Golgi, plasma membrane), transport of material through the secretory pathway during protein biogenesis occurs via membrane vesicles that bud off of the source membrane and fuse with the target membrane in a highly ordered and specific fashion (reviewed in Barinaga, 1993, Ferro-Novick and Jahn, 1994, Benderek et al., 1996, Aridor and Balch, 1996, and Rothman and Wieland, 1996). During the process of vesicle formation there is an elaborate sorting process that occurs to allow for appropriate loading of cargo into each vesicle. For example, proteins exiting the ER for the Golgi are sorted such that ER resident proteins are, for the most part, excluded from the vesicles. There is also evidence for the additional active sorting of cargo protein into vesicles based on an as yet poorly understood basis (reviewed in Aridor and Balch, 1996, and Rothman and Wieland, 1996). One problem with the latter possibility is that it is difficult to envision the multitude of receptors that would be required to incorporate diverse cargo proteins into vesicles. The discovery that ERGIC53, a protein that shuttles between the ER, intermediate compartment (vesicular structure between the ER and Golgi; reviewed in Hauri and Schweizer, 1992) and cis-Golgi, is a lectin that specifically recognizes mannose residues, has provided a possible solution to this problem (Arar et al., 1995, and Itin et al., 1995). Although a sorting receptor function has
not yet been demonstrated for ERGIC53, in vivo, it is interesting to speculate on its function.

After a glycoprotein folds correctly in the ER, it exits this organelle bearing terminal mannose residues on its oligosaccharide(s) (Figure 10B). Binding of this structure by ERGIC53 may be an efficient way of sorting any folded glycoprotein into transport vesicles. VIP16 is another putative sorting receptor with homology to ERGIC53 and legume lectins (Fielder and Simons, 1994). VIP16 is a membrane protein that cycles between the Golgi and the plasma membrane. It remains to be seen whether this molecule actually functions as a lectin and, if so, whether the specificity matches the oligosaccharide structures expected in this part of the secretory pathway.

These are interesting findings because they suggest that oligosaccharides may have more than one function in the lifetime of a protein and that this function may change as the oligosaccharide itself matures. At very early stages of biogenesis, the oligosaccharide may be involved in folding and assembly and, at later stages, may participate in sorting the protein into transport vesicles. Finally, when the protein is mature, oligosaccharides can function in diverse processes such as cell adhesion, receptor recognition, etc. This suggestion is borne out in part by the results presented in this thesis, in combination with the work of other researchers, in which it is demonstrated that the ER proteins calnexin and calreticulin recognize a correctly processed oligosaccharide moiety as a means to effect their functions as molecular chaperones (reviewed in Helenius, 1994, Hammond and Helenius, 1995 and Williams, 1995).
1.7.2 ER chaperones

BiP

The Hsp70 homologue of the ER lumen was originally identified as a 78 kDa polypeptide that accumulated in the ER upon glucose starvation, thus it was called glucose regulated protein of 78 kDa or Grp78 (Pouységur et al., 1977). This same protein was later called BiP, for binding protein, due to its association with immunoglobulin heavy chain molecules (Haas and Wabl., 1983). At about the same time, cloning of rat p72 revealed that this was an Hsp70 related protein and further work established that it was identical to Grp78 and BiP (Munro and Pelham, 1986). The importance of BiP expression is highlighted by the finding that, in yeast, deletion of the BiP gene is lethal (Nicholson et al., 1990). Consistent with its function as an Hsp70 homologue, BiP was found to have weak ATPase activity; ATP dissociates bound substrate (peptide) and ADP prolongs this association (Kassenbrock and Kelly, 1989, Munro and Pelham, 1986, Dorner et al., 1990, and Toledo et al., 1993). A peptide binding motif has been identified for BiP through the screening of bacteriophage peptide display libraries for BiP binding (Blond-Elguindi et al., 1993a). These studies determined that BiP binds preferentially to peptides whose sequence contains hydrophobic residues at every other position, with some amino acid preferences in the hydrophobic positions (i.e., alternating hydrophobic/non-hydrophobic residues).

BiP has been found to exist in three forms in the ER, as a monomer, dimer (or oligomer) and complexed with other proteins (Frieden et al., 1992, Staddon et al., 1992 and Blond-Elguini et al., 1993b). Free monomeric BiP appears to be the form that is ATPase active whereas the dimeric (or aggregated) form is inactive (Blond-Elguini et al., 1993b). Incubation of dimeric BiP
with ATP results in the conversion from dimers to monomers suggesting a role for ATP in BiP activity (Blond-Elguini et al., 1993b). In addition, there is evidence that interconversion between these forms is regulated by two post-translational modifications, ADP ribosylation and serine/threonine phosphorylation (potentially autophosphorylation) of BiP (Frieden et al., 1992, Staddon et al., 1992, and Gaut and Hendershot, 1993b). These modifications are only found in inactive, oligomeric forms of BiP and are lost upon conversion to the active monomer. The exact roles of ATP, ADP ribosylation and phosphorylation in the activation of BiP in vivo have not yet been established.

Much of the early work on BiP function centred around its putative role in ER “quality control”, the mechanism by which newly synthesized proteins are screened for proper folding and assembly prior to transport out of the ER. This was suggested by the observation that BiP associates with a number of misfolded and/or mis-assembled proteins in the ER for prolonged periods of time, e.g., mutant forms of influenza haemagglutinin (HA) (Gething et al., 1986, Hurtly et al., 1989), VSV G protein (Machamer et al., 1990), insulin receptor (Accili et al., 1992), and immunoglobulin heavy chains (IgH) expressed in the absence of light chains (Haas and Wabl, 1983). Several subsequent studies led to the suggestion that BiP was directly involved in the biogenesis of newly synthesized polypeptides and not just in recognition and retention of misfolded polypeptides (Bole et al., 1986, and Hendershot, 1990). Like many chaperones examined to date, BiP was found to associate with a wide range of newly synthesized proteins. These interactions were transient in nature and appeared to involve immature conformations of these substrates, consistent with a role as a molecular chaperone. Direct evidence that BiP functions as a molecular chaperone comes from a variety of studies. The first
evidence comes from studies examining mutant IgH molecules (lacking the C\textsubscript{H1} domain) that could not associate with BiP. These molecules were found to be poorly assembled into mature IgG suggesting interaction with BiP is required for assembly (Hendershot et al., 1987).

Furthermore, unassembled mutant, but not wild type, heavy chains were secreted suggesting BiP functions in ER “quality control”. BiP was found to associate with nascent thyroglobulin and that this association was responsible for dissolution of aggregates and also in prevention of aggregation (Kim et al., 1992). The most direct evidence for BiP acting as a molecular chaperone comes from studies using BiP mutants in both yeast and mammalian cell systems. In the yeast system, maturation of carboxypeptidase Y (CPY) was examined at the permissive and non-permissive temperatures in \textit{kur2} (yeast BiP is called Kar2p) temperature sensitive mutant strains (Simons et al., 1995). At the non-permissive temperature CPY aggregated, did not form disulfide bonds and was retained in the ER. Unlike wild type BiP, these mutants lacked functional ATPase domains and as a result there was also a prolonged association of mis-folded CPY with the mutants. In a more recent study, immunoglobulin light chain (IgL) folding was examined in COS cells transfected with several BiP ATPase mutants (Hendershot et al., 1996). While both wild type and mutant BiP molecules bound to unfolded or partially folded IgL, the mutants did not release the IgL chains. As a result, light chains could not complete disulfide bond formation and were retained in the ER. That binding to BiP interferes with the formation of disulfide bonds is consistent with \textit{in vitro} studies where BiP was found to have “anti-chaperone” activity during protein disulfide isomerase (PDI) catalysed refolding of lysozyme (Puig and Gilbert, 1994). Finally, mutant BiP molecules that could not hydrolyse ATP were also found to bind to but not release IgH in an ATP dependent fashion (Gaut and Hendershot, 1993b).
same mutants when expressed in COS cells led to a dramatic disruption of the ER as examined by EM suggesting that the release of polypeptide from BiP is essential, not only for BiP chaperone function, but for maintenance of the ER environment (Hendershot et al., 1995).

In addition to quality control and general chaperone functions for BiP there is also growing evidence that BiP is involved in the degradation of mis-folded polypeptides that are retained in the ER. Immunoglobulin light chain molecules, in the absence of heavy chains, degrade in a pre-Golgi compartment, presumably the ER. Although secreted light chains only transiently interact with BiP, non-secreted forms interact quantitatively, for extended times, in partially oxidized form (Knittler et al., 1995). This study also demonstrated that two different non-secreted light chains degrade with half lives that coincide with their dissociation from BiP. These results suggest that interaction with BiP may regulate degradation, possibly by maintaining non-secreted molecules in a partially unfolded state that is accessible to proteases or, the putative retro-translocation machinery (see below). DRβ molecules, constituents of HLA class II molecules, are rapidly degraded in a pre-Golgi degradative pathway if expressed in the absence of DRA and invariant chains. DRβ molecules appear to form large aggregates with BiP as intermediates in degradation (Cotner and Pious, 1995). Unlike monomeric proteins, that aggregate and remain in the ER for extended periods of time, aggregation of subunits of multimeric proteins appears to be only an intermediate in the degradation pathway. Although BiP is found in these aggregates its role in the degradation pathway is not yet clear. Some recent studies examining degradation of CFTR, class I histocompatibility molecules and CPY, suggest that degradation of ER proteins occurs in the cytosol after translocation of the protein out of the ER (Ward et al., 1995, Jenson et al., 1995, Weitz et al., 1996a and 1996b, Hughes et al., 1997
and Hiller et al., 1996). Recent evidence suggests that transport of molecules out of the ER for degradation occurs via Sec61p, a component of the translocon (Weirtz et al., 1996b, Pilon et al., 1997 and reviewed in Kopito, 1997). BiP may act to partially unfold and subsequently stabilize proteins in a translocation competent state for interaction with this putative "retro"-translocation machinery.

These studies suggest that BiP functions in a manner similar to cytosolic Hsp70's. Indeed, like the cytosolic chaperones, BiP requires the presence of a DnaJ homologue to function in the ER. To date there are two yeast DnaJ homologues, Sec63p and Sec11p, in the ER that interact and function with BiP (Kar2p) (Feldheim et al., 1992, Scidmore et al., 1993 and Blumberg and Silver, 1991).

While much of the work on BiP function has involved its chaperone activity, BiP (and more recently a novel ER Hsp70 termed Lah1), have also been implicated in the translocation of polypeptides into the ER in S. cerevisiae (Kang et al., 1990, Vogel et al., 1990, Scherer et al., 1990, Nguyen et al., 1991, Brodsky et al., 1993 and Craven et al., 1996). For BiP, this function has been demonstrated both in intact cells and in reconstituted proteoliposomes. In mutant strains of S. cerevisiae where BiP (Kar2p) expression can be regulated, reduction of BiP to 15% of wild type levels results in a loss of secretion of α factor and invertase. Accumulation of unglycosylated and signal-uncleaved forms of α factor indicated that the block in secretion was at the level of translocation (Nguyen et al., 1991). In another study, proteoliposomes reconstituted from yeast microsomes could support translocation of model proteins only when BiP and cytosolic Hsc70 were present (Brodsky et al., 1993). Hsc70 could not substitute for BiP and vice-versa, suggesting that the two Hsp70 homologues have topologically distinct roles in
translocation of proteins across the ER membrane (Brodsky et al., 1993). In these experiments BiP function was reliant on the expression of the J homologue, Sec63p (Brodsky et al., 1993). Mutations in the J domain of Sec63p that prevent BiP association result in impaired translocation suggesting BiP must interact with Sec63p (Lyman and Schekman, 1995). These studies suggest that BiP plays an important role in binding to hydrophobic regions of translocating polypeptides as they emerge from the translocon, thereby preventing aggregation and mis-folding. In this way BiP may function like mitochondrial Hsp70. In a recent study, Sec63p and BiP were also found to be required for the ATP-mediated release of nascent polypeptides after their binding to a Sec62p, Sec71p and Sec72p translocon sub-complex. The Sec62p, Sec71p and Sec72p translocon sub-complex appears to be required for the initial delivery of nascent polypeptides bearing signal sequences to the site of translocation. The release from this complex, subsequent insertion into the translocation pore and exit of the growing polypeptide chain into the ER constitute the three steps involved in translocation into the ER (Lyman and Schekman, 1997). It is therefore a significant finding that BiP/Sec63p are involved in an early and critical step in the translocation process. Further studies will hopefully address how the interaction of BiP/Sec63p with either the nascent chain and/or sub-complex results in release of the nascent chain and, more importantly, how this activity differs from its chaperone activity.

**grp94**

Grp94 is an abundant ER glycoprotein that is a member of the Hsp90 family of molecular chaperones. Grp94, as the name implies, was originally identified as a 94 kDa protein whose
synthesis increased during glucose starvation. Although grp94 appears to be the most abundant of the ER resident proteins there is very little known about its chaperone function. The synthesis of grp94, as expected for a molecular chaperone, is regulated by accumulation of unfolded polypeptides in the ER (see section 1.7.3) and it associates with a number of unfolded or partially folded polypeptide intermediates (Lee, 1987, Kozutsumi et al., 1988, Lenny and Green, 1991, Dorn et al., 1992, and Cala and Jones, 1994). Its substrates include immunoglobulin heavy chains, class II histocompatibility molecules, thyroglobulin, collagen and HSV-1 glycoprotein B (Melnick et al., 1992, Schaift et al., 1992, Navarro et al., 1991, Kuznetsov et al., 1994 and Ferreira et al., 1994). In many instances grp94 interacts with newly synthesized proteins in a cooperative fashion with other ER chaperones (see section 1.6). Interestingly, it binds at a later stage of protein folding, after BiP, suggesting that it interacts with relatively mature proteins (Melnick et al., 1994. This is consistent with the finding that Hsp90 in the cytosol recognizes substrates such as src and steroid hormone receptors that undergo relatively limited conformational changes upon release from Hsp90. Several lines of evidence suggest that grp94 can interact with peptides. Firstly, it can bind to immobilized denatured proteins in an ATP-regulated fashion (Nigam et al., 1994). Secondly, peptides associated with grp94 have been identified as tumour rejection antigens in murine sarcomas (see section 1.9.4). The observation that peptides associated with grp94 can be used to immunize mice and elicit cytotoxic T lymphocyte (CTL) responses confirms that grp94 can indeed bind peptides (see section 1.9.4). How relevant this is to antigen presentation in general or even its chaperone activity has not yet been established. Structurally, grp94 appears to be similar to other Hsp90 family members. Grp94 is N-glycosylated and contains an ER localization sequence, KDEL, at
its C-terminus (Kang and Welch, 1991). Grp94 can bind ATP and ATPase activity has been demonstrated \textit{in vitro} (Clairemont et al., 1992, and Li and Srivastava, 1993). In addition, grp94 binds Ca\textsuperscript{2+} and can be phosphorylated by casein kinase II, but phosphorylation does not appear to be relevant \textit{in vivo} (Cala and Jones, 1994). A recent study has used purified grp94 and a variety of \textit{in vitro} methods to determine the domain structure of grp94 (Wearsch and Nicchitta, 1996). The hydrodynamic properties of native grp94 suggest it exists as a rod-like dimer in solution. This dimerization can be reversed by incubation with detergent, suggesting that dimerization occurs via hydrophobic interactions. These results are consistent with the structure and oligomeric state of cytosolic Hsp90. Using a cell free reticulocyte lysate expression system and a series of grp94 mutants, this study also determined that the dimerization domain, like in other Hsp90 molecules, is found within the 200 amino acid C-terminal domain. By fusing smaller segments of the C terminal region to maltose binding protein (MBP) and examining dimerization of MBP it was determined that the dimerization sequence was in the hydrophobic patch of amino acids between amino acids 676 and 719. Electron microscopic analysis of wild type grp94 dimers, monomers and mutant grp94 suggested that the dimers form a rod like structure where the central portion is made up of two C-terminal domains interacting in opposite orientation to each other. The electron microscopic analysis could only distinguish low resolution shapes of the entire molecule and does not provide a detailed atomic structure. As a result the C-terminal/C-terminal interaction has not been established unequivocally. X-ray crystallographic structure determination would provide the necessary detail to confirm or refute this suggested structure.

It is important to reiterate at this point that a chaperone function for grp94 has only been
inferred from its binding behaviour in the ER. To unequivocally establish it as a molecular chaperone there has to be more work on the mechanism of interaction with newly synthesized polypeptides and a detailed analysis of how deletion or mutation of grp94 affects the folding and assembly of its substrate molecules.

Other chaperones and folding enzymes

In addition to BiP, Grp 94, calnexin and calreticulin (see below) the ER also contains a number of other putative molecular chaperones. Some of these like Grp170 (an Hsp70/Hsp10-like-molecule) interact with a number of newly synthesized proteins, such as unfolded immunoglobulin, thyroglobulin and thrombospondin, in co-operation with BiP and grp94 (Chen et al., 1996, Lin et al., 1993, Kuznetsov et al., 1994, Kuznetsov et al., 1997). Grp170 has also been implicated in protein transport into the ER (Dierks et al., 1996). Other putative chaperones, such as Hsp47 (reviewed in Nagata, 1996), RAP (Bu et al., 1995; and Willnow et al., 1996), and invariant chain (reviewed in Williams and Watts, 1995) have been identified as substrate-specific chaperones for collagen, LDL receptor-related protein, and class II histocompatibility molecules, respectively. All three of the substrate-specific chaperones bind to their substrates early in biogenesis. Hsp47 appears to be involved in collagen processing and secretion and, under stress conditions, provides a "quality control" function by retaining improperly folded procollagen in the ER (reviewed in Nagata, 1996). Both RAP and invariant chain have similar functions in that they prevent the premature binding of ligand to their respective substrates. RAP binds to LDL receptors early in the ER and prevents binding of such ligands as apolipoprotein B, thus preventing premature aggregation of the receptor (Bu et al., 1995, and Willnow et al.,...
1996). Invariant chain binds to the peptide binding groove of class II molecules, thereby preventing the binding of ER peptides. It also targets the class II dimers to a specialized endocytic compartment where they acquire peptides derived from endocytosed proteins (reviewed in Williams and Watts, 1995). This ensures that class II molecules receive the appropriate peptides prior to cell surface expression. Further research may reveal a more general role for molecules such as RAP and Hsp47 in ER protein folding or may establish these as truly substrate-specific molecular chaperones. In all cases these molecules interact with immature substrate and appear to dissociate at some stage in the secretory pathway before expression at the cell surface.

In addition to what would be defined as molecular chaperones, the ER contains a number of folding enzymes that are required for the maturation of newly synthesized proteins (reviewed in Helenius et al., 1992). The ER milieu is highly oxidizing compared to the cytosol and, as a result, cysteine residues can form intramolecular and/or intermolecular disulfide bonds (Hwang et al., 1992). The formation and re-arrangement of these bonds is catalysed by protein disulfide isomerase (PDI, ERp72 and Erp61). In addition to catalyzing disulphide bond formation there are examples where PDI can act as a chaperone in the absence of disulfide bond catalysis. In addition to disulfide bond formation, the ER contains a number of other enzymes that modify specific amino acids in proteins. Prolyl hydroxylase, lysyl hydroxylase and peptidyl-prolyl cis-trans isomerase are examples of such "folding" or "maturation" enzymes. While discussion of these enzymes is outside the scope of this introduction it is essential to stress how important these reactions are in the maturation of proteins. In many cases the reactions catalysed by these enzymes are the rate-limiting step in the maturation and folding of newly synthesized proteins.
As a result, any analysis of molecular chaperones and protein folding in vivo must include the action of such enzymes in the overall folding mechanism.

1.7.3 Unfolded-protein response and the regulation of ER chaperone expression

The first results to demonstrate an ER to nucleus signalling pathway were from experiments in which mutant forms of influenza haemagglutinin protein were found to induce the expression of several ER chaperones including BiP, grp94, and PDI. Through the use of a variety of stresses and chemical perturbants it was subsequently determined that the common stimulus for induction of this pathway was the accumulation of unfolded proteins in the ER, hence the unfolded-protein response or UPR (reviewed in Pahl and Baeuerle, 1997). This pathway could be demonstrated in both yeast and mammalian cells but, due to ease of manipulation, the UPR has been most intensively studied in yeast cells.

Analysis of the _kur2_ (yeast BiP) gene identified a 22-base pair sequence that is conserved in the 5' regions of a number of ER chaperone genes and is responsible for the UPR (Mori et al., 1992). During stress this UPR element (UPRE) binds a novel transcription factor, Hac1p (Cox and Walter, 1996). Deletion of the gene for Hac1p abolishes the UPR in yeast cells. Hac1p appears to be expressed in an unstable form constitutively, but stress (accumulation of unfolded proteins) induces alternative splicing of the mRNA to produce a stable form of Hac1p that binds to the UPRE and increases expression of ER chaperones (Figure 11) (Cox and Walter, 1996, and Sidrauski et al., 1996). A ser/thr kinase (Ire1p) found in the ER membrane is an essential component of the UPR and functions upstream of Hac1p (Cox et al., 1993 and Mori et al.,...
Consistent with a structure similar to mammalian transmembrane kinase receptors, Ire1p is activated upon dimerization (Shamu and Walter, 1996). Dimerization is induced by stress and it has been proposed, but not experimentally demonstrated that free BiP, present in excess under normal conditions, binds to the Ire1p monomer and prevents dimerization. As the concentration of unfolded proteins increases, BiP becomes limiting and Ire1p dimerization occurs (reviewed in Pahl and Baeuerle, 1997). This signals kinase activation from the lumenal domain to the cytosolic kinase domain. This is only a hypothesis, especially since any of the ER chaperones could fulfill the same function as BiP. Activation of the kinase domain somehow increases alternative splicing of the Hac1p mRNA. Sequence similarity of Ire1p to other endonucleases suggests that activated Ire1p may be directly involved in splicing of Hac1p mRNA to produce the more stable product (Sidrauski et al., 1996).

The mechanisms involved in the UPR in mammalian cells are not well understood. A 28 base pair GRP core element has been identified in mammalian GRP genes. Although this core is 50% identical to the yeast UPRE, its deletion does not abrogate the UPR in mammalian cells (Chang et al., 1989, Wooden et al., 1991). It would appear that in mammalian cells no one element is responsible for the UPR but instead there are multiple redundant elements that contribute to the overall effect. The inability to identify an Ire1p homologue or other kinase in mammalian cells that is responsible for the transmembrane signalling also underscores the divergent nature of the UPR in yeast and mammals (reviewed in Pahl and Baeuerle, 1997). In contrast to yeast, there may be more than one mammalian kinase involved in the UPR signalling mechanism.
Figure 11
The unfolded protein response pathway in yeast (UPR). The pathway is described in detail in the text (Taken from Pahl and Baeuerle, 1997).
1.8 Calnexin and Calreticulin

The bulk of this thesis deals with the question of the functions of calnexin and calreticulin and their mode of interaction with newly synthesized proteins within the ER. Given that the background and discussion of these issues is presented in detail in subsequent chapters and in the general discussion, the following will focus on the discovery, general properties and structures of calnexin and calreticulin.

1.8.0 Calnexin

Calnexin was discovered at approximately the same time by three independent research groups and, as a result, appears in the literature as p88, calnexin and Ip90. Calnexin was first described as a novel 88 kDa (p88) protein that can be cross linked to murine class I histocompatibility molecules early in their biogenesis (Degen and Williams, 1991). Class I molecules are trimeric molecules consisting of a transmembrane heavy chain, a soluble light chain termed β2-microglobulin (β2m) and a short 8-10 residue peptide ligand. p88 interacts transiently, but quantitatively, with newly synthesized class I heavy chains and dissociates shortly before they acquire resistance to digestion with endoglycosidase H. This suggested that p88 interacts with class I molecules in the ER where oligosaccharides on the heavy chain are sensitive to endoglycosidase H. In cell lines deficient in one of the subunits of the class I heterotrimer, p88 remains associated with unassembled heavy chains for prolonged periods suggesting a role for calnexin in ER quality control (Degen et al., 1992). At about the same time, another group identified a 90 kDa phosphoprotein from dog pancreatic microsomes which they called calnexin (Wada et al., 1991). The name calnexin was based on the protein’s homology with
calreticulin and its ability to bind calcium. From the cDNA encoding this protein, calnexin was predicted to be a 573 amino acid, 65.4 kDa type I integral membrane protein with a 90 amino acid cytoplasmic tail, a single transmembrane domain and a large 463 amino acid lumenal domain (Figure 12). This topology has been confirmed by proteolytic digestion of intact microsomal membranes. Calnexin is targeted to the ER via an N-terminal signal sequence that is cleaved upon translocation. The 90 amino acid cytoplasmic domain contains an ER localization motif (RKPRRE) as well as sites for phosphorylation by casein kinase II (Wada et al., 1991, Ou et al., 1993a). The role of phosphorylation in calnexin function has not been elucidated at this time. The human homologue of calnexin was also reported at this time. It was identified as a 90 kDa protein (IP90) that co-isolates with incompletely assembled T cell receptor, immunoglobulin and MHC class I subunits (Hochstenbach et al., 1992). The monoclonal antibody AF8 was generated to this protein and used to localize IP90 expression to the ER. Co-immunoprecipitation experiments with this antibody demonstrated that IP90 interacts transiently with over 50 newly synthesized proteins in the ER (Hochstenbach et al., 1992). The sequence of IP90 was compared to the canine sequence and exhibited 95% amino acid identity (David et al., 1993). Furthermore, comparative peptide mapping combined with antibody cross-reactivity demonstrated that p88 and calnexin are the same protein (Ahuwalia et al., 1992).

It was not long after these first reports that the number of substrates recognized by calnexin began to grow rapidly (Table III). Table III is likely to keep growing given the vast number of as yet unidentified proteins that co-precipitate with calnexin-specific antibodies. The transient nature of its interactions coupled with its observed binding to incompletely folded
Figure 12

Linear representations of calnexin and calreticulin. The homologous regions are shown by the rectangles and the positions of the two tandemly repeated sequence motifs are indicated by the numbers 1 and 2 (Reproduced with permission of D. B. Williams). The single high affinity site for Ca\(^{2+}\) binding is localized to the repeat 1 region in both calnexin and calreticulin. The ER localization sequences, RKPRRE or KDEL, are indicated at the C-terminus of each protein. The N, P and C domains of calreticulin are also shown.
<table>
<thead>
<tr>
<th>CLASS OF SUBSTRATE</th>
<th>CALNEXIN</th>
<th>CALNEXIN/CALRETICULIN</th>
<th>CALRETICULIN</th>
</tr>
</thead>
<tbody>
<tr>
<td>SOLUBLE</td>
<td>$\alpha_1$-antitrypsin, $\alpha_1$-chymotrypsin, C3, Apo B-100, gp80 (MDCK glycoprotein), thyroglobulin, $\alpha$ fetoprotein</td>
<td>transferrin, HIV gp160(120)</td>
<td>myeloperoxidase</td>
</tr>
<tr>
<td>MEMBRANE ASSOCIATED/</td>
<td>B cell receptor (membrane bound immunoglobulin), MHC class II ($\alpha$, $\beta$, $\kappa$), VSV-G, nicotinic acetylcholine receptor $\beta$ subunit, IFN$\gamma$ receptor, chicken hepatic lectin</td>
<td>human and mouse MHC class I heavy chain, influenza hemagglutinin, Glut 1 glucose transporter, T cell antigen receptor ($\alpha$, $\beta$ subunits), integrins</td>
<td></td>
</tr>
<tr>
<td>MEMBRANE ANCHORED</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>POLYTOPIC MEMBRANE</td>
<td>CFTR, P glycoprotein</td>
<td></td>
<td></td>
</tr>
<tr>
<td>NON-GLYCOSYLATED</td>
<td>TCR CD3c (natural), P glycoprotein (mutant), MHC class I heavy chain (I$\alpha$, I$\beta$ mutant)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MISC. INTERACTIONS</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(interactions with mature proteins)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table III
Table of calnexin and calreticulin substrates sub-divided into soluble, membrane associated, polytopic, non-glycosylated and misc. categories. The common substrates are listed in the third column.
polypeptides suggested a chaperone-like function for calnexin. Evidence that calnexin functions as a molecular chaperone is indirect. It binds transiently to a diverse array of both membrane and soluble glycoproteins shortly following their translocation into the ER, and its dissociation can usually be correlated with some stage in polypeptide folding or subunit assembly. In the case of transferrin and influenza HA, calnexin binds to incompletely oxidized folding intermediates and dissociates at about the time fully disulfide-bonded molecules are formed (Ou et al., 1993b, and Hammond et al., 1994). For the major secretory glycoprotein of MDCK cells, gp80, calnexin dissociation correlates with gp80 precursor folding as judged by the differential susceptibility to proteinase K of calnexin-bound versus released molecules (Wada et al., 1994). Furthermore, the dissociation of calnexin from the gp80 precursor can be blocked by modulating disulfide bond formation with either DTT or diamide. In addition to folding, subunit assembly can apparently occur while polypeptides are associated with calnexin. Assembly of the heavy chain, β2-microglobulin (β2m), and peptide ligand of mouse class I histocompatibility molecules takes place on calnexin. Formation of the complete ternary complex is required for efficient dissociation of calnexin since incomplete complexes lacking β2m or peptide exhibit prolonged binding to calnexin (Degen and Williams, 1991 and Degen et al., 1992, Suh et al., 1996). By contrast, in human cells, assembly of class I heavy chain-β2m heterodimers appears to be sufficient to trigger calnexin dissociation (Sugita and Brenner, 1994, and Nolten and Parham, 1995). Class II histocompatibility molecules assemble into a large complex consisting of three invariant chains and two αβ dimers while associated with calnexin. Addition of the final αβ dimer correlates with calnexin dissociation (Anderson and Cresswell, 1994). As stated above, the consistent observation that calnexin interacts with incompletely folded or assembled
proteins, but is absent from native (or nearly native) structures, suggests a molecular chaperone function for calnexin. Furthermore, calnexin interacts in prolonged fashion with misfolded glycoproteins or with subunits that are unable to assemble. Incompletely assembled forms of class I molecules (Dejen et al., 1992, and Rajagopalan and Brenner, 1994), the T cell receptor (David et al., 1993), and integrins (Lenter and Vestweber, 1994) remain stably associated with calnexin and are not transported out of the ER. Misfolded mutant proteins such as those produced by the metabolic incorporation of amino acid analogs (Ou et al., 1993b), the ts045 mutant of VSV G protein (Hammond et al., 1994), and a truncated variant of α1-antitrypsin (Le et al., 1994) are also retained as complexes with calnexin. A direct demonstration of calnexin's capacity to retain incompletely assembled proteins was provided by co-expressing calnexin along with free class I heavy chains or heavy chain-β2m heterodimers in Drosophila cells (Jackson et al., 1994). The aberrant transport of these assembly intermediates out of the ER that occurs in Drosophila cells was impeded when calnexin was co-expressed. In a separate study, retention by calnexin was demonstrated by expressing either full length calnexin or a truncated variant that lacks an ER localization signal in cells that also express the T cell receptor ε subunit or class I heavy chain subunit. Whereas these subunits were retained in the ER in association with intact calnexin, their association with truncated calnexin resulted in redistribution to the Golgi complex or cell surface (Rajagopalan et al., 1994, and Rajagopalan and Brenner, 1994). These studies demonstrate that calnexin functions as a component in ER “quality control”, the process by which proteins are retained intracellularly until completion of folding and assembly.

Two properties set calnexin apart from known molecular chaperones, its membrane disposition and its marked preference for Asn-linked glycoproteins. The latter property
originally became apparent when human hepatoma cells were treated with tunicamycin and the level of complexes between calnexin and newly synthesized secretory glycoproteins was dramatically reduced (Ou et al., 1993b). Subsequent studies revealed that treatment of cells with the oligosaccharide processing inhibitors castanospermine and 1-deoxynojirimycin also prevented calnexin binding (Hammond et al., 1994; Hammond and Helenius, 1994). These compounds inhibit glucosidases I and II, enzymes that remove glucose residues from the Glc3Man9GlcNAc2 oligosaccharide that is co-translationally attached to nascent polypeptide chains as they enter the ER. Based on this and additional data it was suggested that glucose removal is required for calnexin binding and that calnexin specifically recognizes the Glc1Man9GlcNAc2 structure (Hammond et al., 1994). In addition to potentially interacting with oligosaccharide, there is substantial evidence that calnexin also recognizes the polypeptide portion of incompletely folded glycoproteins. There are many examples of calnexin binding to proteins that lack Glc1Man9GlcNAc2 oligosaccharides either naturally (Rajagopalan et al., 1994), by inhibition of glycosylation or oligosaccharide processing (Kearse et al., 1994; Arunachalam and Cresswell, 1995), or by mutagenesis (Leo and Clarke, 1994; Carreno et al., 1995a).

At this time nothing is known about the three dimensional structure of calnexin but the primary sequence contains some very interesting regions in terms of topology and homology with calreticulin (Figure 12). The lumenal domain of human calnexin is 33% identical to calreticulin, a soluble protein in the ER (see below) (reviewed in Williams, 1995). The region of highest identity between the two proteins is dominated by two repeated motifs, known as motif 1 and motif 2. These motifs are tandemly repeated 4 times in calnexin and 3 times in calreticulin.
In calnexin the motif 1 repeats contain a single high affinity calcium binding site (Tjoelker et al., 1994). Ca$^{2+}$ binding appears to be required for stable association of calnexin and newly synthesized proteins and also for stabilization of calnexin conformation (reviewed in Bergeron et al., 1994). The region encompassing the two motifs in both calnexin and calreticulin is also the most conserved segment across species (Figure 13). The N- and C-terminal regions of the ER lumenal domain of calnexin are acidic in nature and account for calnexin's aberrant mobility on SDS-PAGE (runs 23-25 kDa higher than the actual molecular weight). These regions also contain low affinity calcium binding sites (Tjoelker et al., 1994). To date there has been only scant evidence that calnexin binds ATP and no evidence for ATPase activity has been reported (Ou et al., 1994). Whereas radiolabeled ATP was found to interact with calnexin in overlay blots, evidence for its binding in vivo has not been demonstrated (Ou et al., 1994).

Calnexin homologues have been identified and cloned from a wide range of organisms including dog (Wada et al., 1991), mouse (Schreiber et al., 1994), human (David et al., 1993), rat (Tjoelker et al., 1994), worm (Hawt et al., 1993), plant (Huang et al., 1993), frog (Yamamoto et al., 1996), fruit fly (Chrisodoulou et al., 1997) and yeast (Jannatipour and Rokeach, 1995, Parlati et al., 1995a, de Virgilio et al., 1993, and Parlati et al., 1995b). The mammalian homologues share 93-98% amino acid identity (Tjoelker et al., 1994). The S. pombe homologue, Cnx1p, is 40% identical, in amino acid sequence, to mammalian calnexins (Jannatipour and Rokeach, 1995, and Parlati et al., 1995b) whereas the S. cerevisiae homologue, Cre1p, is only 24% identical (Parlati et al., 1995b). Cre1p differs from other homologues in its overall length and, unlike other calnexins, it is N-glycosylated, does not bind calcium and does not contain an ER retention motif (Parlati et al., 1995b). Although it retains significant identity in the motif 1
Alignment of motif 1 and 2 repeats in calnexin and calreticulin

The consensus (CON) motif 1 and 2 sequences of calnexin and calreticulin are shown along with an alignment of sequences from various species. For a residue to be included in the consensus sequence it must be present in at least 66% of the sequences from various species. (Reproduced with permission of D. B. Williams)
and 2 repeats (Figure 13) it remains to be determined if, in fact, it functions as a calnexin homologue. Deletion of this gene does not lead to any detectable phenotype in *S. cerevisiae* suggesting its expression is not essential (Parlati et al., 1995b). In contrast, deletion of the *S. pombe* Cnx1p gene led to a lethal phenotype (Jannatipour and Rokeach, 1995, and Parlati et al., 1995b). Although the *S. pombe* homologue is up-regulated in response to stress, to date there is no evidence for the up-regulation of mammalian calnexin in response to stress (i.e., no UPR) (Parlati et al., 1995b).

1.8.1 Calreticulin

Calreticulin is a 46 kDa soluble protein that was originally identified as a calcium storage protein of the ER and sarcoplasmic reticulum (reviewed in Michalak et al., 1992, and Krause and Michalak, 1997). It contains a cleavable N-terminal signal sequence as well as a KDEL ER localization motif at its C-terminus (Figure 12). The N- and C-terminal domains of calreticulin are negatively charged and contain low affinity, high capacity calcium binding sites. A single high affinity calcium binding site is located in the central P domain which contains the motif 1 and motif 2 repeats described above for calnexin (Figures 12 and 13) (Baksh and Michalak, 1991). Calreticulin is constitutively expressed in most cells but expression can also be induced by heat shock via a putative heat shock element located in the 5' region of the human calreticulin gene (Conway et al., 1995).

Calreticulin is often referred to as a “multi-functional” protein because there are reports of its localization to the ER, nucleus, cytosol and cell surface (reviewed in Michalak et al., 1992, and Krause and Michalak, 1997). There is evidence for its role in diverse cellular processes such
as calcium homeostasis, steroid hormone receptor action, integrin function, modulation of IP3-mediated Ca\textsuperscript{2+} signalling (Camacho and Lechleiter, 1995), modulation of intracellular Ca\textsuperscript{2+} store-operated Ca\textsuperscript{2+} influx from the extracellular space (Mery et al., 1996), binding to rubella virus RNA (Singh et al., 1994), initiation of spreading of laminin-adherent cells, Ro52/SS-A autoantigen binding (Singh et al., 1994, and Cheng et al., 1996) and, finally, in ER folding as a chaperone (reviewed in Michalak et al., 1992, and Krause and Michalak, 1997). Evidence supporting the expression of calreticulin in the cytosol or nucleus is not definitive and appears to occur under conditions of high expression of calreticulin when mis-targeting to the cytosol can occur at low frequency. In addition, this mis-targeted pool could translocate into the nucleus via a cryptic nuclear localization signal (NLS) in calreticulin. Under normal conditions it appears as though calreticulin is strictly an ER lumenal protein (reviewed in Michalak et al., 1992, and Krause and Michalak, 1997).

Evidence that calreticulin interacts directly with glucocorticoid receptors was demonstrated in vitro and appears to be via a conserved KXFF(K/R)R motif found in the DNA binding domain of the receptor (Luisi et al., 1991, and Burns et al., 1994). This sequence is almost identical to the integrin α subunit sequence, KXGFFKR, which is a highly conserved segment within the cytoplasmic domain. This sequence was found to be necessary for the interaction of integrins with calreticulin and a synthetic peptide corresponding to this sequence could bind calreticulin in vitro (Rojiani et al., 1991, Leung-Hagesteijn et al., 1994). For both steroid receptors and integrins there is evidence that calreticulin is involved in modulating their activity. Interaction between integrins at the cell surface and calreticulin is inducible and it appears as though binding of calreticulin is required for stabilizing the high affinity state of the
receptor (Coppolino et al., 1995). In a recent study, calreticulin double knock-out embryonic stem (ES) cells were constructed as well as fibroblasts isolated from mouse embryos lacking calreticulin and both examined for integrin-mediated adhesion (Coppolino et al., 1997). Although integrin expression was unaltered, integrin-mediated adhesion to extracellular matrix was impaired. Transfection of calreticulin in the ES cells rescued integrin-mediated adhesion. Interestingly, the amount of calcium in endoplasmic stores was unaltered in these cells, suggesting calreticulin is not essential for calcium storage in the ER. Whether the in vivo interaction between calreticulin and the integrin receptor is actually via the cytoplasmic peptide shown to interact in vitro is not clear. More work needs to be done to map the in vivo binding site. If it is the same as the in vitro site then it will be necessary to determine how an ER protein is redirected to the cytoplasmic face of the plasma membrane.

As stated above, calreticulin interacts with the DNA binding domain of the glucocorticoid receptor and interaction can inhibit binding of the receptor to the glucocorticoid response element in vitro (Burns et al., 1994). Over-expression of calreticulin in mouse fibroblasts resulted in decreased transcriptional activation of a glucocorticoid-sensitive reporter gene and an endogenous target gene suggesting that calreticulin modulates the activity of the glucocorticoid receptor (Burns et al., 1994). A direct interaction between calreticulin and glucocorticoid receptors has not been demonstrated in vivo. The recent demonstration that the in vivo effect is dependent on over-expression of ER-localized calreticulin and not cytosolic calreticulin suggests that the modulation of steroid responses occurs via an indirect route, possibly by the effects of calreticulin on calcium signalling from the ER (Michalak et al., 1996).

Two independent lines of evidence led to the suggestion that calreticulin acts as a lectin.
Firstly, a cell surface form of calreticulin was identified as the molecule responsible for initiating spreading of adherent cells onto laminin in culture (White et al., 1995). This function was found to be the result of a direct interaction between calreticulin and oligomannosides covalently linked to laminin (Chandrasekaran et al., 1996a and 1996b, McDonnell et al., 1996). These oligosaccharides are unique in that they resemble oligosaccharide structures present on newly synthesized glycoproteins in the ER and it was shown that calreticulin could bind mannan, Man5, Man9 and laminin but not Man3 structures (Chandrasekaran et al., 1996a and 1996b). Interestingly, calreticulin binding to laminin was dependant on Ca++ since EDTA abolished the interaction (McDonnell et al., 1996). The aberrant localization of calreticulin at the cell surface may be the result of co-secretion with plasma membrane proteins, although this has not been tested experimentally. Secondly, a lectin activity for calreticulin was postulated because of its homology to calnexin and the growing evidence, presented below, that it too might bind oligosaccharide.

The earliest suggestion that calreticulin could bind newly synthesized or unfolded proteins came from a study in which ER lysates were passed over a number of denatured protein columns (Nigam et al., 1994). Several ER proteins selectively bound to these columns and did so in an ATP- dissociable fashion. BiP, grp94, PDI, Erp72, Erp50 and calreticulin were isolated. The fact that all of the other proteins were folding factors suggested a similar role for calreticulin. The caveat to this study is that calreticulin may have indirectly associated with the column through interaction with any of the other proteins (see PDI example in Chapter III). Soon after this report there were several reports of calreticulin binding to newly synthesized glycoproteins and its subsequent dissociation concomitant with folding (see Table III for
Like calnexin, anti-calreticulin antibodies could co-isolate a large number of newly synthesized proteins (Peterson et al., 1995). The interaction is transient in nature and can be inhibited by glycosylation processing inhibitors (Peterson et al., 1995). These interactions were reminiscent of calnexin-mediated interactions and prompted one group to examine whether differences in the pattern of polypeptides that bind to calreticulin compared to calnexin could be accounted for by different ER localization (Wada et al., 1995). When calreticulin was expressed as a membrane-anchored form there was a shift in the pattern of proteins that co-isolated with calreticulin-specific antibodies (Wada et al., 1995). Interestingly, this pattern was almost identical to the pattern observed in calnexin co-precipitates, suggesting that membrane anchoring may play a role in substrate binding (for more discussion see Chapter V).

For both calnexin and calreticulin, studies have suggested that these molecules act as molecular chaperones and lectins. These functions will be elaborated in subsequent chapters.

1.8.2 The ER chaperone machine - Emerging model

It is clear from a number of studies that ER chaperones function in concert with each other and with ER folding enzymes to promote the efficient folding of newly synthesized proteins and degradation of terminally mis-folded proteins. These factors bind both concurrently and in succession and calnexin and calreticulin are no exception. There are a growing number of examples of glycoproteins that interact with both proteins, either in a ternary complex (gp160, influenza haemagglutinin, Glut 1 transporter) or in succession (MHC class I molecules)(see section 1.6). These folding pathways may also include any of the other ER chaperones and enzymes providing an ER chaperone/folding machine. The efficiency with which this machine
functions would be improved by the association of the different factors in super-complexes similar to cytosolic Hsp90 and evidence for this possibility is growing. ER folding factors have been isolated in weakly associated complexes that are proposed to form a folding matrix in the ER (Tatu and Helenius, 1997). Interactions such as those between calreticulin and PDI, BiP and PDI, and calreticulin/calnexin and Erp57 (a thiol dependant reductase) all support this model of an ER folding matrix. Future work will undoubtedly identify the specific interactions involved and the overall architecture of the matrix.

1.9.0 Major Histocompatibility Complex Molecules

Dr. Williams' laboratory has been studying the biogenesis of mouse class I histocompatibility molecules as a convenient system to characterize calnexin's interactions (Degen and Williams, 1991; Degen et al., 1992; Margolese et al., 1993, Jackson et al., 1992). These molecules are ideal model substrates for a number of reasons. The genes for many class I molecules have been cloned and the crystall structures determined for several human and mouse class I allele products. Furthermore, there are a large number of conformation sensitive monoclonal antibody reagents available to assess folding and assembly events during class I biogenesis. To understand the general properties of these model substrates, the following sections will review class I function, structure, biogenesis and interaction with ER molecules.

1.9.0 General properties and function

Major Histocompatibility Complex (MHC) molecules were first identified functionally as the factors which were responsible for graft rejection in mouse models of skin transplantation.
(reviewed in Abbas et al., 1991). By the generation of inbred and congenic (identical in all but one genetic locus) mouse strains investigators were able to deduce that there was a single genetic region which was responsible for graft rejection. This region was linked to a blood group antigen originally termed Antigen II (later termed H-2). The occasional recombination event observed during interbreeding of congenic mouse strains resulted in the conclusion that more than one gene was encoded by the H-2 region and that graft rejection was determined by several different genes clustered in this region. Subsequent work identified these genes and mapped the genes within the H-2 region. Further work using alloantisera (antibodies raised to white blood cells of donors) from transplant patients, multiparous women and immunized volunteers, established that a similar region is present in humans that contains 6 genetic loci responsible for non-self recognition (HLA- human leukocyte antigens). Further work on the nature of MHC molecules led to their classification as either class I or class II molecules. Class I genes are termed K, D and L in the mouse and A, B and C in the human genome whereas class II genes are I-A and I-E in the mouse and DQ, DR and DP in the human genome (Figure 14A). The organization of the mouse and human MHC as well as the structure of class I and II genes are shown in Figure 14. To date there are up to 20 mouse alleles and greater than 40 human alleles at each of the class I and II loci. This makes the MHC one of the most polymorphic regions in the mouse and human genomes.

Class I molecules present fragments of endogenously synthesized foreign antigens to CD8+ T cells (i.e., peptides derived from viral proteins synthesized during infection) whereas class II present fragments of exogenously synthesized foreign antigens to CD4+ T cells (i.e., peptides derived from antigens endocytosed and processed via the endocytic pathway). CD8+
Figure 14
A. Genomic organization of the human and mouse MHC, localized on mouse chromosome 17 and human chromosome 6.
B. Exon-intron structure of class I and II genes.

( Abbas et al., 1991)
cells are cytotoxic T cells responsible for killing infected or aberrant cells whereas CD4+ T cells are the T helper cells responsible for eliciting help from other components of the immune system to clear an infection. In the case of class I molecules, the cytotoxic T cell surveys class I molecules expressed on the surface of cells. If the class I molecule contains a peptide fragment from a foreign protein, as in the case of viral infection, or from proteins unique to transformed or malignant cells, the infected or malignant cell is destroyed. If the class I molecules contain only peptide fragments from normal cellular proteins, as is the case under most circumstances, the cell is ignored by the cytotoxic T cell (reviewed in Germain and Margulies, 1993, Neefjes and Momburg, 1993, and York and Rock, 1996).

While it was clear that class I and II molecules are a necessary component in the recognition of foreign proteins (as peptide fragments) by T cells, the exact nature of how this is accomplished was not clearly understood until the crystal structures of these molecules were solved. With the solution of these structures came the understanding not only of how these molecules interacted with a great number of peptides but also gave the structural basis for observed differences in T cell recognition of class I and II molecules. The remainder of this discussion will focus on class I molecules.

In Figure 15, schematic representations and ribbon structures of class I molecules are shown. The class I heavy chain contains a large extracellular segment consisting of a peptide binding region (α1 and α2 domains) and an immunoglobulin-like region (α3 domain) (Figure 15A and 15B). The polymorphic heavy chain non-covalently associates with the non-polymorphic, non-MHC encoded β2-microglobulin subunit (β2m) (Figure 15B). The heavy chain contains a single transmembrane domain and a cytoplasmic domain that can vary in both length...
Figure 15

A. Schematic of class I molecule. The various regions are defined in the text (Abbas et al., 1991).

B. Ribbon diagram of the soluble extracellular fragment of a class I molecule with bound peptide. \(a_1, a_2, a_3\) and \(\beta_m\) are defined in the text.

C. Ribbon diagram of the peptide binding groove of class I viewed from the top. The \(\beta\) sheet makes up the floor of the groove and the two \(\alpha\) helices define the sides of the groove (Abbas et al., 1991).

D. Side view of the peptide binding groove of class I (ribbon) and bound peptide (space filling).
and sequence. The most polymorphic region, the peptide binding domain, is defined by a β sheet floor with α helical sides (Figure 15C). This peptide binding clefts can accommodate a large number of peptides with stable binding kinetics due to unique binding characteristics. Each class I molecule displays a preference for binding peptides with specific motifs usually made up of one or two essential anchor residues whose side chains make specific contacts in the groove (e.g., peptide residues 5 and 9 in Figure 15D). Other peptide residues are completely variable. For class I molecules, the N- and C-termini of the 8-10 amino acid peptide make necessary contacts with the ends of the binding groove to provide stable binding of a wide range of peptides (reviewed in Rammensee et al., 1993, and Barber and Parham, 1993). The ends of the peptide binding cleft are closed in class I molecules and as a result peptides longer than 8-10 residues cannot bind. There are rare examples where larger peptides bind by bulging in the central region or where one of the two termini protrudes out of the binding groove (Chen et al., 1994b, and Urban et al., 1994). These unique binding characteristics ensure that each of the 3 to 6 class I allotypes present in any individual is capable of binding to a very large and diverse array of peptides. Consequently, at least one peptide from any foreign antigen will be presented to cytotoxic T cells by class I molecules.

Given that class I molecules are the model substrates we chose to use in our examination of calnexin chaperone functions, the remainder of this discussion will focus on class I biogenesis in the ER.

1.9.1 Class I molecules - General overview of biogenesis

Class I molecules are assembled in the endoplasmic reticulum and, as described above,
consist of a trimolecular complex of the membrane anchored heavy chain, β₂m and peptide ligand. Both the heavy chain and β₂m are co-translationally translocated into the ER. The β₂m subunit is essential for maintaining the heavy chain in a functional peptide binding conformation (reviewed in Jackson and Peterson, 1993, and York and Rock, 1996). β₂m is also important in facilitating the surface expression of functional class I molecules as emphasized by studies in β₂m-deficient cell lines, such as the human Daudi cell line and mouse R1E cells. In these cells, heavy chain molecules are synthesized but are retained intracellularly (reviewed in Jackson and Peterson, 1993). Peptides presented by class I molecules are generated primarily in the cytoplasm and are transported into the ER, via a specialized transporter termed TAP, in order to associate with newly synthesized heavy chain/β₂m dimers. Binding of peptide stabilizes the complex and results in a functional trimolecular complex that can be recognized by CD8+ cytotoxic T cells (see below) (reviewed in Heemels and Ploegh, 1995, and Howard, 1995). Peptide not only acts as the antigenic component of the class I complex but, like β₂m, it facilitates its transport to the cell surface. In cells unable to transport peptides into the ER, due to mutations in the TAP peptide transporter, heavy chain/β₂m dimers are transported very slowly to the cell surface and are thermally unstable (Ljunggren et al., 1990, Townsend et al., 1989, and reviewed in Parham, 1990). The finding that only completely assembled class I molecules are efficiently exported from the ER suggests that there exists a quality control mechanism in the ER that prevents the transport of assembly intermediates to the cell surface.

1.9.2 Generation of class I peptide ligands

Analysis of the peptides bound to class I molecules has demonstrated that the vast
majority of peptides are derived from proteins initially synthesized in the cytosol (reviewed in Heemels and Ploegh, 1995, Suh et al., 1996, and York and Rock, 1996). For this reason much of the focus has been on the 20S and 26S multicatalytic proteasome, the machinery responsible for the bulk of cytosolic protein degradation (reviewed in Rivett, 1993, Ciechanover, 1994, Coux et al., 1996, and Hilf and Wolf, 1996). The 20S proteasome is the catalytic core of the 26S proteasome, and is a 700 kDa complex of 12-15 different subunits. The multicatalytic protease activities reside in a central cavity of the 20S particle. This particle alone requires substrates to be unfolded in vitro. Addition of regulatory subunits to the 20S particle results in the 26S particle that contains ATPase activity, responsible for unfolding substrates, and ubiquitin-binding sites. As a result this particle can degrade ubiquitinated proteins in vitro. This is important given that ubiquitination of proteins in the cytosol targets them for degradation in vivo (reviewed in Ciechanover, 1994).

Several lines of evidence lead to the conclusion that the proteasome generates the bulk of class I-binding peptides. Firstly, the proteasome produces peptides that range in length from 3-15 residues, a range that allows for the generation of the 8-10 amino acid peptides bound to class I molecules (Wenzel et al., 1994). Second, the discovery of two interferon-γ (IFN-γ) inducible proteasome subunits encoded in the MHC led to the proposal that these low-molecular-mass polypeptides (LMP2 and LMP7) are co-ordinately regulated with other IFN-γ regulated MHC genes (such as class I heavy chains and β2m) to increase antigen presentation at the cell surface. When expressed, LMP2 and LMP7 replace resident proteasome subunits and there is some evidence that they change the specificity of the proteasome to favour peptides that can be transported into the ER and bind to class I molecules (reviewed in Suh et al., 1996, and York and Rock, 1996).
LMP 7 but not LMP 2 knock-out mice displayed a 55-75% reduction in the levels of stable class I molecules at the cell surface (Fehling et al., 1994, and van Kaer et al., 1994). This defect could be reversed by exogenous addition of peptide, suggesting a defect in peptide supply. Thirdly, indirect evidence for the proteasome being involved in peptide generation came from several studies using inhibitors of the proteasome and assessing their effects on class I antigen presentation. A good correlation was made between the potency of the inhibitor and the extent of inhibition of antigen presentation to T cells (Rock et al., 1994). In addition there is evidence using mutants in the ubiquitination pathway to show that decreases in ubiquitination, and hence defects in targeting of proteins for degradation, resulted in poorer antigen presentation (Michalek et al., 1993). In the converse experiment, when proteins were mutated to make them more susceptible to ubiquitination there was an increase in presentation of peptides within these proteins and this was inhibited by proteasome and ubiquitination inhibitors (Grant et al., 1995).

Within any given polypeptide there may exist several potentially antigenic peptides that could bind equally well to a given class I allele (e.g., residues 257-264 and 55-62 of ovalbumin), but often what is observed in vivo is that one of the peptides is presented most often (e.g., Ova 257-264) while the other is not (e.g., Ova 55-62). This is the result of the proteasome generating one epitope intact while often destroying the other during degradation. This pattern can be reversed by relocating the peptides to different sites within the protein (Niedermann et al., 1995). This is just one example of how the specificity of the proteasome may alter significantly the spectrum of peptides available for binding to class I.

Cells that are deficient in the TAP transporter, which is responsible for peptide transport into the ER, display a profound decrease in cell surface expression of stable class I molecules.
suggesting that peptide supply is severely limited. While it is clear from these studies that the cytosol is the main source of peptides for binding to class I, there is some evidence for alternative processing pathways that can lead to the generation of peptides suitable for binding to class I molecules (Figure 16, steps 3 and 4). The early evidence supporting ER peptide generation was in transporter-deficient cells (where cytosolic peptides cannot gain access to the ER), or in systems where normally cytosolic peptides were redirected to the ER via fusion to signal sequences or by insertion of epitopes into secreted or plasma membrane proteins (Bacik et al., 1993, Godelaine et al., 1993). For example, HLA-A2 molecules isolated from transporter-deficient cell lines contained peptides derived from signal sequences cleaved off proteins translocated into the ER, provided clear evidence for ER generated peptides (Henderson et al., 1992, Wei and Cresswell, 1992). In contrast, there are examples where epitopes within the signal sequence of proteins are not presented in TAP-deficient cells suggesting that signal cleavage may not be sufficient for production of these epitopes (Aldrich et al., 1994, Hombach et al., 1995, Isobe et al., 1995). A possible explanation for the above differences comes from studies on the proteolytic activities present in the ER (Link et al., 1994, Elliot et al., 1995). While these studies demonstrated that, in addition to signal peptidase, the ER contains an aminopeptidase activity and possibly an endopeptidase or carboxypeptidase activity, generation of specific epitopes was limited by both position in the polypeptide as well as length of the polypeptide translocated into the ER. Therefore, the observed differences in presentation of ER derived epitopes may be antigen specific and may be limited by the protease repertoire of the ER. While a certain degree of processing may occur in the ER generating some epitopes, there
Peptide processing and transport pathways. Peptides generated in the cytosol enter the ER via TAP (1). If the peptide is the correct length and sequence it will bind immediately to the TAP-associated class I heterodimer (2). If the peptide is too long or not the correct sequence it will enter ER peptide pool (3, TAP at top of figure). Trimming of peptides by ER processes may produce peptides that can bind to class I molecules (4). Alternatively, peptides may be transported out of the ER in an ATP-dependent fashion (5). Once in the cytosol peptides may be trimmed (6) and transported back into the ER, via TAP, for a second chance at binding to class I.

Alternative pathways for peptide entry into the ER pool include signal peptidase and generation of peptides by ER degradation of translocated polypeptides (7, ribosomes at bottom of figure) (Taken from Suh et al., 1996).
may be an additional requirement for processing in the cytosol for other epitopes (Hammond et al., 1993, and Hammond et al., 1995). ATP-dependent efflux of peptides out of the ER has been demonstrated and suggests that some peptide fragments of ER directed proteins may get to the cytosol for further processing prior to TAP-dependent transport and binding to class I molecules (Figure 16, step 5).

Given that the majority of TAP transported peptides are a suitable length for binding to class I molecules there is some debate as to how significant ER trimming is to the generation of class I binding peptides. Perhaps ER proteases function primarily in the trimming of longer peptides (transported via TAP) to optimal binding lengths. TAP transporters transport peptides as long as 13 amino acids, and in some cases even longer, suggesting that there may be some trimming which occurs if these peptides are to bind to class I molecules. There has been both in vivo and in vitro evidence that trimming of longer peptides can produce peptides of a length suitable for presentation by class I (Eisenlohr et al., 1992, Ojcius et al.,1993, Roesle et al., 1994, Limmk-Schneider et al., 1994, and Elliot et al., 1995).

In addition to the above routes of peptide generation there are a growing number of examples where class I molecules present peptides derived from exogenous proteins phagocytosed by macrophages or dendritic cells (reviewed in Harding, 1995). There are only a limited number of examples of this pathway and hence it may not reflect a significant contribution in vivo. As a result this area will not be discussed further.

1.9.3 Transport of peptides via TAP1/2 and interaction of class I dimers with TAP1/2

Given that generation of peptides for binding to class I molecules occurs predominantly
in the cytosol, there must exist a mechanism for transporting peptides into the ER, where class I molecules are assembled. As mentioned briefly above, the transporter associated with antigen processing (TAP) serves this function. The transporter is actually a heterodimer consisting of TAP 1 and TAP 2 subunits. These subunits were originally identified as proteins encoded in the MHC adjacent to LMP2 and LMP7. TAP 1 and TAP 2 are members of the ATP binding cassette family of transporters and each contains 6-8 transmembrane segments in its N-terminal portion and a C-terminal cytosolically exposed ATP binding domain (reviewed in Higgins, 1992). Immunolocalization studies have detected TAP proteins in the ER as well as the cis-Golgi (reviewed in Howard, 1995).

The first evidence that TAP functions in class I assembly came from mutant cell lines lacking either or both of the TAP subunits. Class I surface expression was reduced in these cells and empty (peptide deficient) heterodimers accumulated in the ER (reviewed in Momburg et al., 1994). Those heterodimers that are inefficiently transported to the cell surface in these cells are unstable suggesting that they lack peptide ligand. Addition of peptide exogenously to cells stabilizes the cell surface class I molecules confirming that the defect is in peptide supply.

Using microsomal or permeabilized cell systems it was subsequently shown that TAP is necessary for ATP-dependent peptide transport into the ER lumen. In this system, peptides that do not contain a glycosylation site are rapidly transported out of the ER to the cytosol in an ATP-dependent, TAP-independent mechanism giving rise to the possibility of peptide recycling (Figure 16, see pathway 1-3-5-6). As a result of this retro-transport, in the peptide transport assay it was necessary to use peptides with glycosylation sites in order to observe an accumulation of peptides in the microsomes following transport via TAP. The optimal peptide length for
transport appears to be 8-15 residues but peptides as long as 33 amino acids have been reported to be transported, albeit inefficiently. The length of peptide transported is similar to the length of peptides which bind optimally to class I molecules. TAP transporters also display C-terminal residue preferences that appear to be species specific. In the mouse system the preferred residues at the C-terminus are hydrophobic whereas in the human system any residue is efficiently transported. This C-terminal preference correlates well with C-terminal preferences observed for peptide binding to class I molecules. Peptide binding by TAP is independent of ATP and requires peptide contact with both TAP 1 and TAP 2 subunits. In contrast, transport of bound peptide into the ER does require ATP. The first three residues and the ninth residue of a peptide appear to determine how well the peptide is transported by TAP. Interestingly, a positive correlation has been made between class I binding peptides and the efficiency of transport via TAP. This correlation does not always hold true as some optimal class I-binding peptides require additional amino acids to be transported efficiently (reviewed in Suh et al., 1996). Such peptides are presumably trimmed to optimal length by ER proteases (see above).

The discovery that class I heavy chain/β2m dimers, but not free heavy chains, interact with TAP has led to the suggestion that this interaction serves to efficiently load peptide onto class I molecules thereby completing assembly in the ER (Suh et al., 1994, and Ortmann et al., 1994). This is supported by the observation that class I molecules rapidly dissociate from TAP following peptide loading into the class I binding site (Suh et al., 1994). The interaction with TAP ensures that the peptide binding site is in proximity to newly translocated peptides and therefore allows class I binding to compete effectively with other processes such as dilution of peptide in the ER lumen, chaperone binding and export from the ER. Although it is clear that
peptides can bind to class I molecules independently of TAP, as in the case of signal sequence-derived peptides, there is evidence that the bulk of peptide loading is dramatically enhanced by the interaction of class I and TAP. For example, a mutant class I heavy chain that assembles well with β2m but cannot interact with TAP presents antigenic peptide 4-fold less efficiently than the wild type molecule (Suh et al., 1996). The best evidence thus far for the importance of TAP in antigen presentation is the discovery of a mutant human cell line (0.220 cell line) that lacks a gene encoded in the MHC that is necessary for the efficient association of class I heterodimers and TAP. These cells display a marked decrease in cell surface expression of class I molecules and an accumulation of “empty” heterodimers in the ER. TAP peptide transport function is normal in these cells but there is a lack of class I-TAP interaction (Grandea et al., 1995). These results suggest that the physical interaction of class I molecules with TAP constitutes an important step in peptide loading. The gene that is lacking in the 0.220 cell line has been characterized and is called tapasin. Tapasin is a 48 kDa membrane glycoprotein resident in the ER that has been found in complexes with TAP, class I heterodimers and with both TAP and class I heterodimers. This molecule is necessary for interaction of class I with TAP and so it was suggested that it acts as a bridging molecule (Sadasivan et al., 1996). Whether it truly is the sole bridge between TAP and class I or whether it acts by stabilizing the direct interaction between class I and TAP until peptide binds remains to be determined. It will be interesting to see how these molecules, as well as calnexin and calreticulin (see below), are spatially oriented in the TAP complex and how peptide binding alters the conformation of the complex such that dissociation occurs. It is noteworthy that a mouse homologue of tapasin has not been identified and raises the question of whether the human and murine systems differ in the requirement for a
bridging molecule.

### 1.9.4 Accessory molecules involved in class I biogenesis.

**Calnexin**

As outlined above, several observations in cell lines defective in class I assembly, suggested that the ER contains a molecule or molecules which are involved in maintaining ER “Quality Control” during class I biogenesis. The most likely candidate for this function is calnexin, which associates with murine and human class I heavy chains very early after synthesis.

Immediately following synthesis, murine class I heavy chains bind quantitatively to calnexin and assembly with β2m occurs shortly thereafter (Figure 17A) (Degen and Williams, 1991). The heavy chain-β2m heterodimers then associate with the TAP peptide transporter (Ortmann et al., 1994; Suh et al., 1994). Class I heterodimers still remain associated with calnexin at this stage (Suh et al., 1996). Subsequently, peptide binding, release from TAP, and dissociation from calnexin all appear to occur at about the same time (Degen et al., 1992; Suh et al., 1994).

For murine class I molecules the ultimate regulator of transport out of the ER appears to be dissociation from calnexin. A recent study demonstrated that calnexin remains associated with murine class I heterodimers even after interaction with TAP (Suh et al., 1996). Using detergent permeabilization of tissue culture cells to deliver class I binding peptides to the cytosol, this study found that murine class I molecules remain associated with calnexin under conditions where peptide binding and dissociation from TAP has occurred (Suh et al., 1996). These results
demonstrate that calnexin dissociates downstream of TAP, and therefore is more likely to regulate class I transport out of the ER.

While the demonstration of calnexin interactions with murine class I heavy chain-β₂m heterodimers is well established, it was not until recently that calnexin interactions with human heterodimers was demonstrated (Carreno et al., 1995b). These interactions appear to be weak, probably due to human heavy chains having a single N-glycan as opposed to the 2 or 3 on murine heavy chains (see Chapter II). Consistent with this is the observation that addition of a second N-glycan to human class I molecules results in easier isolation of calnexin-heterodimer complexes. It has been suggested that for human molecules the main regulator of export from the ER is dissociation from TAP but, given the difficulty in isolating calnexin complexes with human dimers, this suggestion needs to be examined more closely. Furthermore, human molecules also interact with calreticulin after assembly with β₂m and, as a result, calreticulin may regulate export instead of calnexin (see below).

Although the evidence thus far suggests that calnexin acts as a chaperone in class I biogenesis, there are several pieces of data which suggest the involvement of other chaperones. The observation that, in the absence of calnexin interactions, 20-30% folding and assembly does occur suggests that other chaperones can substitute for calnexin (see Chapter IV). In addition, the recent characterization of a calnexin deficient cell line which was normal for class I biogenesis suggests that in the chronic absence of calnexin other chaperones can functionally replace calnexin (Scott and Dawson, 1995).
Calreticulin

There are at least three reports of calreticulin being associated with both human and murine class I heavy chain/β2m dimers although as observed for calnexin there appear to be differences in the extent to which interaction occurs (Sadasivan et al., 1996, Van Leeuwen and Kearse, 1996c, and A. Vassilakos, data not shown). Under normal conditions the predominant interactions early in both human and murine class I biosynthesis are with calnexin with little or no calreticulin associated with free heavy chains. Upon association with β2m, there is little detectable association of calnexin with human dimers (Figure 17B) (Sadasivan et al., 1996). Instead, it appears as though calreticulin is the predominant molecule associated with dimers. In contrast, a large proportion of total murine dimer molecules remain associated with calnexin throughout their biogenesis (Suh et al., 1996). Calreticulin can be found associated with dimers in the murine system but it appears as though this occurs at a low level (A. Vassilakos, data not shown). Contrary to this finding, Kearse’s group reported that calreticulin is the main accessory molecule interacting with murine dimers that are associated with TAP transporters, suggesting that the murine system behaves like the human system (Van Leeuwen et al., 1996c). This study was limited in that the isolation methods used were not optimal and as a result there were few quantitative data to assess the levels of calnexin versus calreticulin associated with class I molecules present at any given time in the experiments (discussed in greater detail in chapter IV). The exact roles for calnexin and calreticulin in class I biogenesis remain to be carefully dissected but from preliminary studies it would appear as though they act at different stages of class I folding and assembly. The possibility remains for at least some overlap and potential functional redundancy in the ER (see Chapter V for more discussion on this topic).
BIP

Mouse and human class I molecules also differ in their interactions with another chaperone in the ER, BIP. Mouse heavy chain molecules that are not associated with Blm are found exclusively in association with calnexin whereas human free heavy chains associate with both calnexin and BIP (reviewed in Williams and Watts, 1995). As described previously, BIP is a member of the Hsp70 family of molecular chaperones that resides in the ER and associates with newly synthesized and misfolded polypeptides. While it has been suggested that either BIP or calnexin may facilitate the folding and assembly of human class I molecules, in the acute loss of calnexin interactions BIP is not capable of completely functionally replacing calnexin suggesting their roles are not completely redundant (see Chapter IV, Vassilakos et al., 1996).

Grp94

Proteins of 95-105 kDa have been found in immunoprecipitates of mouse class I molecules from TAP-deficient cells and from human cells expressing mouse heavy chains and the adenovirus E3/19K protein (Feurbach and Berger, 1993, Townsend et al., 1990). These proteins appear to be related to or identical to grp94, an Hsp90 family member (also known as gp96). Given that grp94 binds peptides and it can be released from class I molecules by the addition of specific class I binding peptides, it has been suggested that this chaperone may be involved in peptide loading of class I molecules (Feurbach and Berger, 1993, Townsend et al., 1990). This hypothesis is supported by the observation that expression of grp94 can be up-regulated by IFN-γ and IFN-α in a similar manner to the TAP genes (Anderson et al., 1994). Evidence for grp94 being a source of peptides for class I molecules comes from several studies.
When mice are immunized with grp94 (or more precisely the peptides bound to grp94) isolated from certain sarcomas, specific immunity is elicited against a subsequent challenge with the source tumor (Shrivastava et al., 1994). These mice produce cytotoxic T lymphocytes (CTL) against proteins expressed in the grp94 source cells. Furthermore, grp94 isolated from cells expressing VSV nucleocapsid protein can be fed to phagocytic cells, resulting in the latter cells being able to stimulate VSV-specific CTL. This same effect can be obtained by loading isolated grp94 with viral peptides in vitro before incubation with phagocytic cells (Suto and Shrivastava, 1995). The above experiments involve the transfer of grp94 bound peptides to class I molecules via some undefined phagocytic route. Direct evidence for grp94 being involved in peptide transfer in the ER where it is expressed normally has not been demonstrated. Why there would be a requirement for grp94 in addition to TAP for peptide loading is not clear, but may involve loading of different pools of peptides. A recent study in which the grp94 pools in the ER were dramatically reduced using an antisense approach resulted in normal class I surface expression and presentation of defined β-galactosidase epitopes (Lammert et al., 1996). The results of this study would suggest that grp94 is not a major peptide source for class I molecules. However the possibility remains that this route may present a suitable alternative under conditions where cytosolic degradation and TAP transporter function are compromised during viral infection.

1.9.5 Model of class I biogenesis

Efficient class I biogenesis in the ER is the result of a series of interactions with ER resident proteins that serve to enhance either folding, assembly with β2m or peptide loading of class I molecules. Figure 17 includes all of the interactions known to date for murine (Figure
and human (Figure 17B) class I molecules and attempts to show when they occur during class I passage through the ER.
Figure 17

A. Schematic of murine class I biogenesis including the interactions important for folding, assembly and peptide loading in the ER. Class I heavy chains and \( \beta_2m \) are co-translationally translocated into the ER. The newly synthesized class I heavy chain binds to calnexin very rapidly. Within 3-5 minutes the heavy chain folds and binds to \( \beta_2m \). After assembly of the heterodimer the calnexin/heavy chain/\( \beta_2m \) complex interacts with the TAP transporter. Peptide binding, dissociation from TAP and, finally, dissociation from calnexin occurs in a time scale of 20-60 minutes depending on the class I allele. The mature class I trimer is then rapidly transported out of the ER to the cell surface, via the Golgi (Reproduced with permission from Suh et al., 1996).

B. Schematic of human class I biogenesis including the interactions important for folding, assembly and peptide loading in the ER. Class I heavy chains and \( \beta_2m \) are co-translationally translocated into the ER. The newly synthesized class I heavy chain binds to calnexin and BiP very rapidly. Within minutes the heavy chain folds and binds to \( \beta_2m \). After assembly calnexin dissociates and calreticulin associates with the heterodimer. The calreticulin/heavy chain/\( \beta_2m \) complex interacts with tapasin and the TAP transporter. Once peptide binds, the mature class I trimer is rapidly transported out of the ER to the cell surface, via the Golgi (reproduced with permission of D.B. Williams).
A  Mouse

B  Human
1.10 Issues Addressed in this Thesis

This thesis focuses on several issues relating to the functions of calnexin and calreticulin.

1. Mode of Interaction of Calnexin and Calreticulin

At the time I embarked on these studies, the lectin function of these proteins had been suggested based only on studies with inhibitors of glycosylation or oligosaccharide processing inhibitors. Therefore the putative lectin function of calnexin and calreticulin remained to be tested directly. The experiments in Chapters II and III address the issue of whether calnexin and calreticulin are indeed lectins. Chapter II also addresses the question of whether calnexin binds to the polypeptide portion of glycoproteins, similar to chaperones of the Hsp60 and Hsp70 families. In Chapter III the lectin properties of calnexin and calreticulin are examined in detail. These results provide a basis for comparing the mechanisms of action of these homologous proteins and allow speculation on how their functions are related in vivo.

2. Functions of Calnexin

Prior to the commencement of this study the evidence for a chaperone function for calnexin was indirect. Calnexin was thought to be a chaperone by virtue of its association with various folding intermediates and mis-folded proteins and not by direct assessment of its effect on folding and assembly. The experiments in Chapter IV were designed to address the question of whether calnexin acts as a bona fide molecular chaperone. The effects of calnexin on the efficiency of class I folding and assembly is examined by two different approaches in an attempt to unequivocally answer this question.
CHAPTER II
THE MOLECULAR CHAPERONE CALNEXIN BINDS Glc,Man,GlcNAc OLIGOSACCHARIDE AS AN INITIAL STEP IN RECOGNIZING UNFOLDED GLYCOPROTEINS
Abstract

Calnexin is a molecular chaperone that resides in the membrane of the endoplasmic reticulum. Most proteins that calnexin binds are N-glycosylated and treatment of cells with tunicamycin or inhibitors of initial glucose trimming steps interferes with calnexin binding. To test if calnexin is a lectin that binds early oligosaccharide processing intermediates, a recombinant, soluble calnexin was created. Incubation of soluble calnexin with a mixture of Glc,Man,GlcnAc oligosaccharides resulted in specific binding of Glc,Man,GlcnAc species. Furthermore, Glc,Man,GlcnAc oligosaccharides bound relatively poorly suggesting that, in addition to a requirement for the single terminal glucose residue, at least one of the terminal mannose residues was important for binding. To assess the involvement of oligosaccharide-protein interactions in complexes of calnexin and newly synthesized glycoproteins, α1-antitrypsin or the heavy chain of the class I histocompatibility molecule were purified as complexes with calnexin and digested with endoglycosidase H. All oligosaccharides on either glycoprotein were accessible to this probe and could be removed without disrupting the association with calnexin. Furthermore, addition of 1M α-methyl-glucoside or α-methylmannoside had no effect on complex stability. These findings suggest that once complexes between calnexin and glycoproteins are formed, oligosaccharide binding does not contribute significantly to the overall interaction. However, it is likely that the binding of Glc,Man,GlcnAc oligosaccharides is a crucial event during the initial recognition of newly synthesized glycoproteins by calnexin.
2.1 Introduction

One of the most conspicuous characteristics of calnexin is its apparent specificity for glycoproteins that possess Asn-linked oligosaccharides. Ou et al. originally showed that pretreatment of human hepatoma cells with tunicamycin prevents the formation of complexes between calnexin and many newly synthesized secretory glycoproteins (Ou et al., 1993). This finding was subsequently reproduced with the integral membrane glycoproteins, influenza HA and VSV-G. Furthermore, pre-treatment with the α-glucosidase I and II inhibitors, castanospermine and 1-deoxynojirimycin, blocks the binding of calnexin to HA or G proteins (Hammond et al., 1994). In contrast, the α-mannosidase inhibitor 1-deoxy-mannojirinimycin has no effect. The data obtained with the oligosaccharide processing inhibitors suggest that glucose trimming of newly synthesized glycoproteins is a requirement for calnexin binding. Additional studies showed that the HA glycoprotein isolated from a complex with calnexin likely contains oligosaccharides with one or two terminal glucose residues and that the VSV G ts045 mutant, which possesses monoglycosylated oligosaccharides for extended periods in the ER, exhibits prolonged association with calnexin. All of these observations led to the proposal that for calnexin binding to occur, a glycoprotein must possess oligosaccharides that have undergone partial trimming from the initial Glc3Man9GlcNAc2 structure to one containing either two or, more likely, a single glucose residue (Hammond et al., 1994 and Helenius, 1994).

One interpretation of these results is that calnexin is a lectin with specificity for monoglycosylated oligosaccharides. However, as discussed recently (Bergeron et al., 1994), other interpretations are equally plausible. In fact some observations are difficult to reconcile with recognition of monoglycosylated oligosaccharide being a prerequisite for calnexin binding. For
example, both the T cell receptor $\alpha$ subunit that lacks Asn-linked oligosaccharides and a recombinant form of the multi-drug resistance P glycoprotein in which N-glycosylation sites are absent, form stable and long-lived complexes with calnexin (Rajagopalan et al., 1994 and Loo and Clarke, 1994). Conversely, removal of the transmembrane and cytoplasmic domains from the T cell receptor $\alpha$ subunit almost completely eliminates calnexin binding but this truncated subunit still possesses its full complement of N-linked oligosaccharides (Margolese et al., 1993). Finally, cross-linking experiments have indicated that Asn-linked oligosaccharides are unlikely to be the sole mode of association between calnexin and class I heavy chains (Margolese et al., 1993). This latter study identified a region encompassing the transmembrane domain and three flanking amino acids of the heavy chain as a site of interaction with calnexin.

In an effort to clarify the involvement of Asn-linked oligosaccharides in the binding of calnexin to newly synthesized glycoproteins, the ability of calnexin to function as a lectin by assaying its ability to bind to a series of oligosaccharide processing intermediates was tested. Additionally, the relative contribution of protein-carbohydrate and protein-protein interactions in maintaining the association between calnexin and newly synthesized soluble or transmembrane glycoproteins was assessed. These findings indicate that calnexin is indeed a lectin with specificity for the Glc,Man,GlcNAc$_2$ oligosaccharide. They also lead to the proposal of a model in which binding to this oligosaccharide is a critical event that occurs during initial recognition of newly synthesized glycoproteins by calnexin. However, once calnexin-glycoprotein complexes are formed, protein-protein interactions become predominant and oligosaccharide binding contributes minimally, if at all, to the overall association.
2.2 Materials and Methods

2.2.0 Production and Analysis of Dolichol-linked Oligosaccharides

[1-3H]dolichol-linked oligosaccharides were produced in M. Lehrman's laboratory by high-efficiency labeling of Chinese hamster ovary cells with [2-3H]mannose, cleaved from dolichol pyrophosphate, and reduced with NaBH₄, essentially as described earlier (Zhu et al., 1992). Approximately 5 x 10⁶ cpm of free, reduced oligosaccharides were obtained from each 100 mm dish of cells. Oligosaccharides were fractionated with an ISCO dual-pump HPLC system, employing Chemresearch software (ISCO Corp.), with a 150 mm Absorbsphere HS 3µ silica column (Alltech) as described (Sykulev et al., 1994). Radioactive material eluting from the column was identified with a Radiomatic A-140 detector with either a 1.0 or 2.5 ml detector cell: the former was chosen for experiments requiring optimal peak resolution and the latter for those needing optimal sensitivity.

2.2.1 Preparation of Soluble, Hexahistidine-tagged H-2Kb and Calnexin Proteins

Soluble class I H-2Kb heavy chain (with His, at the C-terminus) in association with β₂m was purified from culture medium of transfected Drosophila melanogaster SC2 cells by Ni-NTA-agarose (Qiagen) and MonoQ (Pharmacia) anion exchange chromatography as described previously (Sykulev et al., 1994). To prepare His₆-tagged soluble calnexin, canine calnexin cDNA under control of the metallothionein promoter in plasmid pRMHa3 (Bunch et al., 1988) was digested with Dsa I and Sal I to remove the segment corresponding to the transmembrane and cytosolic domains. The deleted segment was replaced with an oligonucleotide cassette such
that the final construct encoded the complete ER luminal domain of calnexin terminating at Pro$^{162}$ followed by the sequence: SRRSWGSHHHHHH. The plasmid was co-transfected with a neomycin resistance plasmid, phleo-neo, into Drosophila SC2 cells and G418-resistant stable cell lines were obtained (Jackson et al., 1992). Cells grown in Insect-Xpress medium (Whittaker) at 23°C were treated with 1 mM CuSO$_4$ for 3 days to induce expression of soluble calnexin. Culture supernatants (2 litres) were concentrated 10-fold by ultrafiltration through an Amicon YM30 membrane and were dialyzed against phosphate-buffered saline, pH 7.4, (PBS) prior to loading onto a 5 ml column of Ni-NTA agarose. After recirculating the sample overnight, the column was washed extensively with PBS followed by 20 mM imidazole in PBS. Calnexin was eluted with 100 mM imidazole in PBS and then dialyzed against 50 mM Tris-Cl, pH 8.5. The sample was applied to a MonoQ 10/10 anion exchange column (Pharmacia) which was subjected to a linear 0-600 mM NaCl gradient in 50 mM Tris, pH 8.5. Calnexin eluted at ~400 mM NaCl and was concentrated using a Centricon 30 concentrator (Amicon). Both soluble calnexin and H-2K$^k$ were essentially homogeneous as judged by SDS-PAGE analysis and silver staining. About 0.5 - 1 mg of each protein was recovered per liter of culture supernatant.

2.2.2 Incubation of Radiolabeled Glc$_3$Man,GlcNAc$_2$ Oligosaccharides with Immobilized Proteins

Each assay employed 6 - 7.5 μg samples of hexahistidine-tagged calnexin or class I H-2K$^k$ molecule immobilized on 12 - 15 μl of nickel-agarose in binding buffer (10 mM Hepes- Na, pH 7.5, containing 0.15 M NaCl and 10 mM CaCl$_2$). Unless indicated otherwise, all procedures were performed at 23°C. The agarose was washed briefly with binding buffer just prior to
incubation with oligosaccharides, suspended in 100 μl binding buffer containing approximately 20,000 cpm of a mixture of Glc₆,Man₆,GlcNAC₉ oligosaccharides, and incubated for 1 hour with agitation on an orbital shaker at 200 rpm. The samples were then centrifuged for 5 minutes at 2,600 x g and the supernatant collected. The agarose beads were rinsed briefly with 100μl binding buffer, centrifuged as above, and the supernatant collected. The agarose pellet was saved for later analysis (below). The two supernatants were pooled, 0.1 ml (approximate packed volume) water-washed Dowex 50W X8 (H' form) was added, and the samples were agitated at 300 rpm for 5 minutes. The supernatants were recovered, approximately 20 Amberlite MB-3 beads were added, and the samples were mixed at 350 rpm for 15 minutes. If all of the dye on the MB-3 beads changed color, indicating that salts remained in the sample, the process was repeated with additional beads. The supernatants were evaporated to dryness, dissolved in 200 μl of 80% acetonitrile, 20% water, 0.2% 1,4-diaminobutane, divided into two equal portions, and each was analyzed separately by HPLC as described above with a 1.0 ml detector cell. In all experiments reported, similar results were obtained with each of the two samples.

To analyze the oligosaccharides that bound to the immobilized proteins, the nickel-agarose beads were subjected to a series of sequential washes. The first wash was performed briefly (1-2 min) at 4°C with 100 μl binding buffer followed by centrifugation for 5 minutes at 2,600 x g and recovery of supernatant. Four subsequent washes were then performed, each for 1 h at 23°C with agitation at 200 rpm, consisting of 100 μl of binding buffer alone, followed by binding buffer supplemented with (in order) 0.1 M α-methyl-D-galactopyranoside, 0.1 M α-methyl-D-mannopyranoside, or 0.1 M α-methyl-D-glucopyranoside (Aldrich). The sixth and final wash was performed with agitation overnight at 23°C with 100 μl 0.1 M α-methyl-D-glucopyranoside
An aliquot of each supernatant (10%) was analyzed by liquid scintillation counting, and the remaining six wash supernatants for each immobilized protein were pooled and treated with Dowex and Amberlite beads as described above. The entire eluate sample was then analyzed by HPLC with a 2.5 ml detector cell which gives greater detector sensitivity relative to a 1.0 ml detector cell, but also increases peak widths and decreases peak resolution. To detect oligosaccharides which remained bound to the proteins after the elution procedure, the nickel-agarose beads were boiled for 15 minutes in 100 µl of binding buffer. No tritium was detected in the resulting supernatants for either H-2K<sup>a</sup>-agarose or calnexin-agarose.

### 2.2.3 Preparation of [H]<sub>2</sub>Glc<sub>6</sub>Man<sub>4</sub>GlcNAc<sub>2</sub> oligosaccharides and incubation with calnexin-agarose

Pure [H]<sub>2</sub>Glc<sub>6</sub>Man<sub>4</sub>GlcNAc<sub>2</sub> was isolated from a mixture of dolichol-linked [H]<sub>2</sub>Glc<sub>6</sub>Man<sub>4</sub>GlcNAc<sub>2</sub> oligosaccharides by preparative HPLC. [H]<sub>2</sub>Glc<sub>6</sub>Man<sub>4</sub>GlcNAc<sub>2</sub> oligosaccharides were then generated by partial digestion with Jack bean α-mannosidase (Boehringer-Mannheim), and their monosaccharide compositions were assigned by comparison of their elution times with those of related dolichol-linked oligosaccharides (Zhu et al., 1992 and Zeng and Lehrman, 1991). Two isomeric configurations are possible for Glc<sub>6</sub>Man<sub>4</sub>GlcNAc<sub>2</sub> and three are possible for Glc<sub>6</sub>Man<sub>4</sub>GlcNAc<sub>2</sub> but these were not characterized. [H]<sub>2</sub>Glc<sub>6</sub>Man<sub>4</sub>GlcNAc<sub>2</sub> was detectable only after brief digestion times and the quantities generated were insufficient for further analysis. A mixture of [H]<sub>2</sub>Glc<sub>6</sub>Man<sub>4</sub>GlcNAc<sub>2</sub> oligosaccharides was prepared, and 2,000 cpm were analyzed directly by HPLC (see Figure 3, starting fraction). Another portion (approximately 4,000 cpm) was dissolved in 100 µl binding
buffer and incubated with calnexin-agarose as described above. After collection of the supernatant (approximately 3,400 cpm) by centrifugation, the pellet was eluted with two aliquots of 100 µl binding buffer containing 0.1 M α-methylglucoside at 23°C for 1 hr and 16 hr, respectively. Approximately 1000 cpm were recovered in the pooled eluates. Both the supernatant and eluate fractions were desalted and characterized by HPLC as described above, except that individual 0.5 ml fractions were collected, mixed with 4.0 ml scintillation cocktail, and analyzed with a liquid scintillation counter.

2.2.4 Isolation and Analysis of Radiolabeled Calnexin-Class I Heavy Chain Complexes

Transfected Drosophila melanogaster Schneider cells expressing canine calnexin and either H-2Kb or D* class I heavy chains were maintained in Schneider's insect medium (Sigma) supplemented with 10% fetal bovine serum, antibiotics, and 0.5 mg/ml Geneticin (Gibco-BRL). Synthesis of calnexin and class I heavy chains was induced by treatment with 1 mM CuSO4 for 24 h. Following induction, cells (1-3 x 10^7) were incubated for 15 minutes at 23°C in methionine-free Grace's insect medium (Sigma) and then were resuspended at 5x10^7 cells/ml in Met-free Grace's medium supplemented with 0.5 mCi/ml [35S]Met (>800 Ci/mmol; Amersham). Radiolabeling was carried out for 10 min at 23°C and then the cells were washed twice with phosphate-buffered saline, pH 7.4, (PBS) followed by lysis at 0.2-0.3 x 10^7 cells/ml in PBS containing 0.5 % digitonin (Sigma), 10 mM iodoacetamide, 1% aprotinin, 0.25 mM 4-(2-aminoethyl)benzenesulfonyl fluoride, and 10 µg/ml each of chymostatin, leupeptin, antipain, and pepstatin. After 30 minutes on ice, lysates were divided into 1 ml aliquots and centrifuged at 16,000 x g for 10 minutes. To isolate calnexin-class I heavy chain complexes, lysates were
incubated for 2 hours at 4°C with 20 µl of a rabbit antiserum raised against a peptide corresponding to the C-terminal 14 amino acids of canine calnexin (Jackson et al., 1994).

Immune complexes were recovered by shaking with 15 µl of protein A-agarose (Gibco-BRL) for 1 h at 4°C. The agarose beads were washed 3 times with 0.2% digitonin in PBS and resuspended in 50 µl wash buffer. To this bead-bound complex of calnexin and class I heavy chain was added either 5-20 mU (IUB) endo H (New England Biolabs), 1 M α-methyl-D-mannopyranoside, or 1 M α-methyl-D-glucopyranoside and the samples were incubated at room temperature for 1-4 hours. Control incubations with no additions were included for each experiment. At the end of the incubation the beads were centrifuged, the supernatant removed, and the beads were washed once with 50 µl wash buffer which was pooled with the supernatant. The supernatants and washed beads were boiled in SDS-PAGE sample buffer containing 100 mM DTT and proteins were separated by SDS-PAGE using 10% gels and visualized by fluorography (1M sodium salicylate). In addition to the non-denaturing conditions used for the endo H digestions outlined above, digestions performed under denaturing conditions (0.1% SDS) were included in each experiment to provide size standards for completely deglycosylated class I heavy chains (Jackson et al., 1994).

2.2.5 Isolation and Analysis of Calnexin-α1-Antitrypsin Complexes

Human HepG2 hepatoma cells were grown in α-MEM supplemented with 10% fetal bovine serum and antibiotics. Radiolabeling was performed as described for Drosophila cells except that the pre-incubation (30 min) and labeling (10 min) were performed at 37°C in Met-free α-MEM supplemented with 9 mM Heps and 2 mM glutamine. Cell lysis, isolation of anti-
calnexin immune complexes, and digestion with endo H were also performed as described above except that 2% sodium cholate (Sigma) replaced digitonin in the lysis buffer and protein A-agarose beads were washed with 0.5% sodium cholate in PBS. Unlike the situation in *Drosophila* cells, calnexin binds to many radiolabeled proteins in HepG2 cells. In order to detect α1-antitrypsin in the supernatant and bead fractions following endo H digestion of anti-calnexin immune complexes, it was necessary to re-isolate with anti-α1-antitrypsin antibody.

Supernatant and bead fractions were boiled in 0.1 ml PBS containing 0.2% SDS for 10 minutes and then 1.2 ml PBS, pH 8, containing 1% Nonidet P40 was added. The samples were incubated with 10 μl rabbit anti-α1-antitrypsin antibody (Calbiochem) for 2 hours on ice and immune complexes were recovered with protein A-agarose beads. The beads were washed 4 times with PBS containing 0.5% Nonidet P40 and proteins were analyzed by SDS-PAGE as described above. Under these conditions there was no residual endo H activity during the immune isolation of α1-antitrypsin.

### 2.3 Results

#### 2.3.0 Calnexin selectively binds Glc,Man,GlcNAc from an oligosaccharide mixture

To obtain radiolabeled oligosaccharides that could be used in a direct test of binding by calnexin, we took advantage of the fact that the dolichol-linked oligosaccharide pool in CHO cells can be efficiently labeled with [3H]mannose (Zhu et al., 1992). This pool consists of the full length dolichol-linked oligosaccharide,Glc,Man,GlcNAc, (Figure 1), as well as less...
Figure 1. Structure of the full length dolichol-linked oligosaccharide. This oligosaccharide is transferred to Asn-X-Ser(Thr) sequences during translocation of nascent polypeptides into the ER lumen. It is subsequently processed through the action of α-glucosidases I and II within the ER and by α-mannosidases in both the ER and Golgi apparatus. The sites of cleavage by the glucosidases are indicated. The Glc α1-3Man, Glcα1-3Man oligosaccharide that is bound by calnexin is enclosed by the box. Mannose residues removed by digestion with Jack bean α-mannosidase (see Fig. 3) are shown in italics.
abundant precursors possessing shorter oligosaccharide chains. After release from dolichol pyrophosphate with mild HCl treatment and reduction with NaBH₄, the oligosaccharides can be resolved with an HPLC system that separates neutral oligosaccharides on the basis of increasing size (Zeng and Lehrman, 1991). Although the predominant radiolabeled dolichol-linked oligosaccharide is Glc₃Man₃GlcNAc₂ (Zhu et al., 1992 and references therein), HPLC analysis revealed appreciable quantities of precursors eluting as Glc₃Man₃GlcNAc₂ (Figure 2, Panel A). These assignments are based on previous studies in which dolichol-linked oligosaccharides from normal and mutant CHO lines labeled in vivo and in vitro were characterized (Zeng and Lehrman, 1991, Zhu et al., 1992 and Camp et al., 1993). In addition, we found that only species expected to contain glucose residues were labeled after incubation of cells with [³H]galactose, which is converted intracellularly into UDP-[³H]glucose. The relative incorporation of label reflected the anticipated glucose contents of the oligosaccharides (data not shown). Thus, in cells incubated with [³H]galactose, Glc₃Man₃GlcNAc₂ was by far the most efficiently labeled dolichol-linked oligosaccharide, with Glc₃Man₃GlcNAc₂ labeled in lesser amounts.

Solutions of [³H]mannose-labeled Glc₃Man₃GlcNAc₂ were incubated with a soluble His₆-tagged form of calnexin, or, as a control, a soluble His₆-tagged form of the class I H-2Kb histocompatibility protein, each immobilized on nickel-agarose. After 1 h, the agarose was removed by centrifugation and the supernatants were analyzed by HPLC. An incubation without agarose beads was also included. As shown in Figure 2, panel A, four distinct oligosaccharide species (Glc₃Man₃GlcNAc₂) were present in the sample incubated without agarose. After incubation with H-2Kb-agarose, the amount of each oligosaccharide remaining in solution was not significantly altered (data not shown). In contrast, the Glc₃Man₃GlcNAc₂ oligosaccharide
Figure 2. Selective binding of Glc,Man,GlcNAc, oligosaccharide by immobilized calnexin.

A mixture of Glc,Man,GlcNAc, oligosaccharides was dissolved in binding buffer (see Materials and Methods), divided into three equal portions, and incubated for 1 h at 23°C either alone, with H-2Kb-agarose, or with calnexin-agarose. Panel A: HPLC analysis of the oligosaccharide mixture incubated in the absence of immobilized protein. The arrowheads indicate the elution times of the various oligosaccharides in the mixture. Panel B: HPLC analysis of oligosaccharides eluted from calnexin-agarose (closed circles). The elution profile of a mixture of the following oligosaccharide standards is also included (open circles):

Man,GlcNAc, Man,GlcNAc, Glc,Man,GlcNAc, Glc,Man,GlcNAc, and Glc,Man,GlcNAc.
was selectively depleted after incubation with calnexin-agarose. In several experiments, the amount of Glc,Man,GlcnAc_{3} oligosaccharide recovered was reduced by 20 - 45% whereas no depletion of other oligosaccharides from the mixture by calnexin could be detected (data not shown).

To examine material which specifically bound to calnexin, a six step elution procedure (Table I) was employed with the H-2K^{b}-and calnexin-agarose samples recovered after the experiment. Each successive step involved a condition expected to be more effective for elution of bound oligosaccharide from calnexin. In total, about 80 cpm was recovered from H-2K^{b}-agarose, whereas approximately 560 cpm was recovered from calnexin-agarose. Surprisingly, a large fraction of the radioactive material eluted from calnexin-agarose in the initial steps which employed buffer alone or buffer plus α-methyl-D-galactoside, a compound not expected to inhibit calnexin. These data suggested that the oligosaccharide was not tightly bound to calnexin. After the sequential elution procedure, no additional radioactivity was recovered by boiling the agarose. The fractions eluted from calnexin-agarose were then pooled and analyzed by HPLC (Figure 2B). A single peak was observed which co-eluted with Glc,Man,GlcnAc_{3}. In other experiments the small amount of radioactivity eluting from H-2K^{b}-agarose was analyzed by HPLC, but no discrete peaks were observed. In contrast, radioactivity recovered from calnexin-agarose always eluted as Glc,Man,GlcnAc_{3}.

2.3.1 Glc,Man,GlcnAc_{3} oligosaccharides are relatively poor ligands for calnexin

The preceding experiments demonstrated that calnexin exhibits a strong preference for the Glc,Man,LeNac_{3} oligosaccharide over oligosaccharides that possess ninemannoses and
Table I

Elution of bound radioactive material from immobilized proteins

The agarose samples recovered from the experiment shown in Figure 2 were sequentially eluted in a six step procedure as described under Experimental Procedures. The tritium content of 10% of each eluate was determined by liquid scintillation counting. Values listed in the table reflect the total estimated radioactivity recovered for each step.

<table>
<thead>
<tr>
<th>Step</th>
<th>Glycoside</th>
<th>Time (h)</th>
<th>Per Step (c.p.m.)</th>
<th>Total (c.p.m.)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>H-3K&lt;sup&gt;o&lt;/sup&gt;</td>
<td>Calnexin</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>H-3K&lt;sup&gt;o&lt;/sup&gt;</td>
<td>Calnexin</td>
</tr>
<tr>
<td>1</td>
<td>0.02</td>
<td>50</td>
<td>110</td>
<td>50</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>20</td>
<td>100</td>
<td>70</td>
</tr>
<tr>
<td>3</td>
<td>0.1 M αMe-Gal</td>
<td>1</td>
<td>10</td>
<td>120</td>
</tr>
<tr>
<td>4</td>
<td>0.1 M αMe-Man</td>
<td>1</td>
<td>0</td>
<td>80</td>
</tr>
<tr>
<td>5</td>
<td>0.1 M αMe-Glc</td>
<td>1</td>
<td>0</td>
<td>40</td>
</tr>
<tr>
<td>6</td>
<td>0.1 M αMe-Glc</td>
<td>16</td>
<td>0</td>
<td>110</td>
</tr>
</tbody>
</table>
either 0, 2, or 3 glucose residues. The Glcα1,3 determinant is present at the non-reducing terminus of one of three mannose-containing branches present in Glc,Man,GlcNAc2 (Figure 1).

To examine whether either of the two remaining branches on Glc,Man,GlcNAc2 is important for binding to calnexin, purified [3H]Glc,Man,GlcNAc2 was digested with Jack bean α-mannosidase to generate glucosylated molecules with 5, 6, or 7 mannose residues. These were added to a sample of the undigested material and the resulting mixture of [3H]Glc,Man,GlcNAc2 oligosaccharides was incubated with calnexin-agarose. The starting mixture, unbound radioactive material in the supernatant, and bound material eluted from the pellet were analyzed by HPLC (Figure 3). Comparison of the relative amounts of each oligosaccharide in the bound fraction to the relative amounts in the starting material indicated a 3-fold enrichment of the Glc,Man,GlcNAc2 species bound to calnexin-agarose. The other oligosaccharides were not significantly enriched. These data suggest that at least one of the mannosidase-sensitive residues on Glc,Man,GlcNAc2 may be important for binding to calnexin. It is not clear whether the Glc,Man,GlcNAc2 oligosaccharides detected in the bound fraction were specifically bound to calnexin. These species were not significantly depleted from the bound fraction by repeating the experiment with a brief wash with binding buffer prior to the elution steps (data not shown).

Thus, it is possible that calnexin binds Glc,Man,GlcNAc2, although less avidly than Glc,Man,GlcNAc2.

2.3.2 Incubation of calnexin-glycoprotein complexes with α-methyl-glycosides

The finding that calnexin binds the Glc,Man,GlcNAc2 oligosaccharide raises the question of whether this is the sole means of interaction between calnexin and newly synthesized
Starting Material

Figure 3. Binding of calnexin to mannosidase-treated glucosylated oligosaccharides.
A mixture of $[^3H]$Glc$_n$Man$_x$GlcNAc$_y$ oligosaccharides was incubated with calnexin-agarose. Unbound material was collected in the supernatant fraction following centrifugation, and bound material was collected by eluting the pellet with 0.1 M α-methylglucoside. The starting, unbound, and bound fractions were analyzed by HPLC. In each case, the flow rate was 1 ml/min and fractions were collected every 0.5 min. Fraction 20 corresponds to 35 min in the solvent program. The elution positions of the various oligosaccharides are indicated by the arrowheads.
glycoproteins or if protein-protein interactions are also involved. Because α-linked glucose and mannose residues participate in the binding of the Glc,Man,GlcNAc oligosaccharide to calnexin (Figs. 2 and 3), immunoprecipitated calnexin-glycoprotein complexes were incubated with 1 M α-methyl-glucoside or α-methyl-mannoside in an effort to inhibit the oligosaccharide component of the interaction. In these experiments, *Drosophila* cells expressing canine calnexin and the heavy chain of either the H-2K^b^ or D^a^ mouse class I histocompatibility molecule were used as a convenient source of defined calnexin-glycoprotein complexes. Class I heavy chains are Type 1 transmembrane polypeptides that possess 2-3 N-linked oligosaccharides and we have shown previously that in the absence of their partner subunits, β^2^m and peptide ligand, heavy chains form long-lived complexes with calnexin in mouse cells or in transfected *Drosophila* cells (Degn et al., 1992 and Jackson et al., 1994). In both mouse and transfected *Drosophila* cells the calnexin-heavy chain interaction appears to depend on N-linked oligosaccharides since it is dramatically reduced by treatment of cells with the oligosaccharide processing inhibitor castanospermine (Chapter IV and Figure 4). The advantage of using the *Drosophila* transfectants in the present experiments is that the calnexin-heavy chain complex is the predominant species recovered with anti-calnexin antibodies.

Calnexin-heavy chain complexes were immunoprecipitated from digitonin lysates of metabolically radiolabeled *Drosophila* transfectants using an antibody raised against the carboxy-terminal 14 amino acids of calnexin followed by collection on protein A-agarose. The agarose-bound calnexin-K^b^ or D^a^ complexes were incubated with either 1 M α-methyl-glucoside or α-methyl-mannoside and then separated into supernatant (S) and agarose bead (B) fractions (Figure 5). Following this treatment, nearly all heavy chains remained in the bead fraction.
Figure 4. Castanospermine inhibits calnexin interactions with D^b and α1-antitrypsin molecules.

*Drosophila* cells (left panel) and HepG2 cells (right panel) were pre-incubated with castanospermine (1 mM) for the indicated times prior to a 10 minute radiolabeling. Castanospermine appears to be effective in preventing glucosidase activity in both *Drosophila* and HepG2 cells as evidenced by the shift to slower mobility of D^b and AAT (indicative of larger glucosylated oligosaccharide chains). By 120 minutes of pre-incubation there is a loss of almost all the D^b molecules from calnexin complexes. In HepG2 cells the reduction is not as dramatic, with a loss of about 70% of calnexin-AAT complexes by 60 minutes. It is clear from these results that castanospermine treatment results in an inhibition of calnexin interactions.
Figure 5. Incubation of calnexin-glycoprotein complexes with α-methyl-glycosides.

*Drosophila* cells expressing H-2K^b^ or D^b^ heavy chains and calnexin were radiolabeled with [35S]Met for 10 min. Cells were lysed in buffer containing 0.5% digitonin and calnexin-heavy chain complexes were isolated with anti-calnexin antibody followed by protein A-agarose. The agarose beads were incubated with either 1M α-methyl-glucoside or α-methyl-mannoside and then were centrifuged to form bead-bound (B) and supernatant (S) fractions. Radiolabeled proteins in these fractions were analyzed by SDS-PAGE. Panel A. Effect of α-methyl-glycosides on calnexin-K^b^ heavy chain complexes. Panel B. Effect of α-methyl-glycosides on calnexin-D^b^ heavy chain complexes. The mobilities of calnexin (cnx), the K^b^ and D^b^ heavy chains, as well as molecular weight standards are indicated.
associated with calnexin. Small amounts of heavy chain and calnexin were detected in the supernatant fraction at approximately the same ratio as observed in the agarose bead fraction. The amounts of calnexin and heavy chain in the supernatant were variable from experiment to experiment and likely reflect the difficulty in sedimenting agarose beads efficiently through viscous solutions of 1 M glycoside. Similar results were obtained when the experiment was repeated using Nonidet P40 rather than digitonin for cell lysis and immune isolation (data not shown). The inability to dissociate calnexin-heavy chain complexes under conditions known to be effective in dissociating other lectin-glycoprotein complexes (Osawa and Tsuji, 1987) suggests that oligosaccharide-protein interactions may not contribute substantially to the calnexin-heavy chain association once complexes are formed. Alternatively, α-methyl-glucoside and α-methyl-mannoside may be inefficient competitors of the binding of calnexin to the Glc,Man,GlcNAc, oligosaccharide.

2.3.3 N-linked oligosaccharides are not involved in maintaining stable complexes between newly synthesized glycoproteins and calnexin

As an additional approach to assess the relative involvement of protein-oligosaccharide versus protein-protein interactions in complexes between calnexin and newly synthesized glycoproteins, endoglycosidase H (endo H) was used as a probe to monitor accessibility of N-linked oligosaccharides in calnexin-glycoprotein complexes. In the first set of experiments, radiolabeled calnexin-class I heavy chain complexes were isolated from digitonin lysates of Drosophila transfectants using anti-calnexin antibody followed by collection on protein A-agarose. Figure 6 depicts the results of digesting calnexin-κ heavy chain (panel A) and
calnexin-D* heavy chain (panel B) complexes with endo H and analyzing proteins in supernatant (S) and agarose bead (B) fractions. Remarkably, the two oligosaccharides on the K* heavy chain and the three oligosaccharides on the D* heavy chain were completely accessible to endo H resulting in total deglycosylation of the heavy chains (compare the endo H-digested agarose bead fraction, lane 6, with the deglycosylated heavy chain standard, lane 2). Furthermore, removal of heavy chain oligosaccharides by endo H did not result in dissociation of heavy chains from calnexin. All of the deglycosylated heavy chains remained associated with calnexin in the bead-bound complex; none were present in the supernatant fraction (compare the endo H-digested S and B lanes). The same results were obtained when complexes of calnexin and the H-2* heavy chain were digested with endo H and also when the experiments were repeated using NP-40, CHAPS, or sodium cholate detergents in place of digitonin for cell lysis (data not shown).

In Figure 7, a similar experiment as in Figure 6 is shown with the exception that the endo H was added to the cell lysate prior to immune isolation of calnexin/heavy chain complexes. This addresses the possibility that the complexes in Figure 6 were stable after endo H treatment due to low detergent concentrations or alternatively from being in complexes with antibody and protein A-agarose. After the indicated incubation, a 10-fold excess of detergent lysis buffer containing yeast mannan, to inhibit further endo H activity, was added and then complexes were immune isolated as above. The results are consistent with Figure 6 in that endo H treatment did not prevent subsequent isolation of calnexin/class I heavy chain complexes. The results suggest that complexes of calnexin with membrane proteins survive after both oligosaccharide removal and addition of excess detergent. The experiment was also attempted under detergent-free conditions by delivering endo H via transient permeabilization at high pH into microsomes.
Figure 6. Digestion of calnexin-glycoprotein complexes with endo H.
Panels A and B. Radiolabeled calnexin-heavy chain complexes were isolated with anti-calnexin antibody followed by protein A-agarose as described in the legend to Figure 5. The agarose beads were incubated in the absence or presence of endo H and then were centrifuged to form bead-bound (B) and supernatant (S) fractions (lanes 3-6). In addition, separate samples of calnexin-heavy chain complexes were dissociated in SDS and incubated in the absence or presence of endo H to provide mobility standards of glycosylated and deglycosylated heavy chains, respectively (lanes 1 and 2). To confirm that the heavy chain standards were completely deglycosylated, limited endo H digests were performed to visualize all partially deglycosylated species (data not shown). The major heavy chain bands in lane 2, panels A and B, do indeed represent the K heavy chain lacking its two oligosaccharides and the D heavy chain lacking its three oligosaccharides, respectively. Panel C. HepG2 cells were radiolabeled for 10 min with \[^{15}S\]Met. Cells were lysed in buffer containing 2% sodium cholate and calnexin-glycoprotein complexes were isolated, digested with endo H, and separated into supernatant and bead-bound fractions as in Panels A and B. Subsequently, the fractions were boiled in PBS containing 0.2% SDS, adjusted to 1% NP-40, and subjected to a second round of immunoprecipitation using anti-\(\alpha\)-antitrypsin antibody (lanes 3-6). Lanes 1 and 2 contain standards of glycosylated and deglycosylated \(\alpha\)-antitrypsin, respectively. Complete deglycosylation of the \(\alpha\)-antitrypsin standard was confirmed as described for the class I heavy chains.
Figure 7. Digestion of calnexin-glycoprotein complexes with endo H in cell lysates

Radiolabeled cell lysates (0.5% digitonin) from Drosophila cells expressing D³ and calnexin were incubated in the absence or presence of endo H for 24 hours at 4°C (2 additions of endo H at t=0 and t=12 hours, 50μl total) (lanes 1 and 2). After the incubation, a 10-fold excess of detergent lysis buffer was added (volume diluted 10-fold) containing yeast mannan to inhibit further endo H activity. Radiolabeled calnexin-heavy chain complexes were isolated with anti-calnexin antibody followed by protein A-agarose. In addition, separate samples of calnexin-heavy chain complexes were dissociated in SDS and incubated in the absence or presence of endo H to provide mobility standards of glycosylated and deglycosylated heavy chains, respectively (lanes 1 and 2). The complete deglycosylation of D³ molecules did not reduce recovery heavy chains with anti-calnexin antibodies (compare lanes 3 and 4). Although the overall signal is reduced in the endoH-treated samples, the calnexin-heavy chain ratio is similar.
containing radiolabeled calnexin-heavy chain complexes (Nicchitta and Blobel, 1993).

Subsequent immunosolubilization of complexes, from solubilized microsomes, with anti-calnexin antibodies revealed that, although the efficiency of endo H digestion was quite low, a small proportion of completely deglycosylated heavy chains could be recovered in association with calnexin. This only occurs under conditions where the microsomes were permeabilized to allow access to the lumen and at 24°C where endo H is enzymatically active (data not shown).

Thus, by three different approaches, complete removal of oligosaccharide by endo H treatment did not result in loss of calnexin interactions. These results suggest that oligosaccharide-mediated interactions are not the sole mode of interaction between calnexin and its substrates.

Calnexin binds not only to membrane-associated proteins such as class I molecules but also to a number of soluble, secretory glycoproteins (Ou et al., 1993). Since secretory glycoproteins lack transmembrane segments that have been implicated in the binding of membrane proteins to calnexin (Margolese et al., 1993), they may rely more heavily on oligosaccharide-protein interactions to maintain stable associations with calnexin. Furthermore, experiments involving endo H digestion of complexes containing calnexin and secretory glycoproteins avoid the potential complication of the deglycosylated glycoprotein being unable to dissociate from calnexin due to co-localization within the same detergent micelle. Thus, to evaluate the role of N-linked oligosaccharides in the interaction between calnexin and a secretory glycoprotein, anti-calnexin immune complexes from sodium cholate lysates of metabolically radiolabeled HepG2 cells were treated with endo H. Previous studies have shown that such immunoprecipitates contain complexes between calnexin and many secretory
glycoproteins including α₁-antitrypsin, α₁-antichymotrypsin, α-fetoprotein, transferrin, complement component C3, and apoB-100 (Ou et al., 1993). Following separation into supernatant and bead fractions, the samples were boiled in SDS to disrupt calnexin-glycoprotein complexes (and to inactivate endo H) and then were subjected to a second round of immunoprecipitation with antibodies against α₁-antitrypsin (Figure 6, Panel C). At least 50% of α₁-antitrypsin molecules present in complexes with calnexin could be completely deglycosylated by endo H (compare the endo H-digested agarose bead fraction, lane 6, with the deglycosylated α₁-antitrypsin standard, lane 2). Furthermore, the deglycosylated molecules remained associated with calnexin in the bead fraction (compare the endo H-digested S and B lanes). These results were consistent with those obtained using calnexin-class I heavy chain complexes and suggest that once calnexin-glycoprotein complexes are formed N-linked oligosaccharides are dispensable in maintaining an association with calnexin. The less efficient deglycosylation observed with α₁-antitrypsin relative to class I heavy chains may reflect the fact that a large number of substrates that compete for endo H are present in calnexin immunoprecipitates from HepG2 cells whereas calnexin-heavy chain complexes constitute the major species immunoprecipitated from Drosophila transfectants.

2.4 Discussion

These findings indicate that the ER luminal domain of calnexin has the capacity to bind to the early N-linked oligosaccharide processing intermediate, Glc₃ManGlcNAc₂. The binding is specific because calnexin selected this oligosaccharide from a mixture containing Glc₃ManGlcNAc₂ species. Calnexin appears to show some preference for this oligosaccharide
when presented with a mixture containing Glc,Man,GlcNAc. Thus, in addition to recognizing a single terminal glucose, calnexin may recognize at least one terminal mannose residue in Glc,Man,GlcNAc (Figure 1). This issue is addressed in greater detail in Chapter III.

A detailed analysis of the binding kinetics was precluded by the limited quantities of purified, soluble calnexin available.

In none of the experiments was hydrolysis of any oligosaccharide detected, suggesting that calnexin is a lectin rather than an enzyme that modifies oligosaccharide structure, e.g., a glycosidase. However, unlike most lectins that bind large oligosaccharides with micromolar dissociation constants (Osawa and Tsuji, 1987), calnexin binds the Glc,Man,GlcNAc oligosaccharide weakly; extensive dissociation of the complex is detectable following short incubations in buffers lacking sugar hapten. Furthermore, primary sequence comparisons revealed that calnexin lacks the carbohydrate-recognition domains characteristic of each of the three major groups of animal lectins. These include the C- (Ca$^{2+}$-dependent) type, the galectins (S-type), and the P- (mannose-6-phosphate) type lectins (Drickamer and Taylor, 1993). Although some lectins, such as concanavalin A, are capable of binding to processing intermediates of Asn-linked oligosaccharides, none exhibit the binding specificity associated with calnexin (Drickamer and Taylor, 1993 and Kornfeld, 1992). Consequently, it can be concluded that calnexin constitutes a new type of lectin with unique specificity for the Glc,Man,GlcNAc$_2$ oligosaccharide.

Treatment of cultured cells with either tunicamycin or the glucosidase inhibitors, castanospermine or 1-deoxynojirimycin, results in a dramatic decrease in the formation of complexes between calnexin and a large array of newly synthesized glycoproteins (Figure 4, Ou
et al., 1993, Hammond et al., 1994). Consistent with previous speculation (Hammond et al., 1994 and Helenius, 1994), the demonstration that calnexin binds selectively to the Glc,Man,GlcNAc₂ processing intermediate provides a clear molecular explanation for the action of these drugs. The remarkable efficacy of tunicamycin and the glucosidase inhibitors also underscores the crucial role that oligosaccharide binding must play in the formation and/or maintenance of calnexin-glycoprotein complexes.

Two observations lead to the suggestion that the binding of Glc,Man,GlcNAc₂ oligosaccharide is important in the initial formation of complexes with calnexin but it cannot be responsible for maintaining complexes once they are formed. First, the apparent affinity of calnexin for the oligosaccharide is low and it is unlikely that complexes maintained through this interaction alone would survive the prolonged immune isolation procedures used for their purification. It is possible that calnexin could possess more than one carbohydrate binding site or could exist as a homo-oligomer (features that might increase the avidity of the interaction) but such properties would not be effective for the many glycoproteins having a single oligosaccharide chain. Second, in purified complexes of calnexin with either class I heavy chains or α₁-antitrypsin, all oligosaccharides are accessible to endo H and their removal is not accompanied by complex dissociation. Consequently, a two-step mechanism for the interaction of newly synthesized glycoproteins with calnexin is envisioned (Figure 8).

In this model, the initial interaction between nascent glycoproteins and calnexin occurs through binding of the Glc,Man,GlcNAc₂ oligosaccharide. This intermediate first appears on nascent chains (Hubbard and Robbins, 1979) and, given the rapid association observed between calnexin and some glycoproteins (Degen and Williams, 1991, Anderson and Cresswell, 1994,
Hammond et al., 1994, and Wada et al., 1994), it is possible that the initial binding of calnexin could occur co-translationally. Having been brought in proximity to calnexin through this initial interaction, the unfolded glycoprotein then binds to calnexin through segments of its polypeptide chain. The physical features recognized by calnexin in this second interaction are largely unknown but they may be hydrophobic segments or patches as is thought to be the case for chaperones of the Hsp 60 and Hsp 70 families (Hendrick and Hartl, 1993). Due to the reversibility of carbohydrate binding, dissociation of the oligosaccharide likely occurs as the polypeptide chain and calnexin interact, perhaps aided by new steric constraints placed on the oligosaccharide and calnexin. Subsequently, the polypeptide folds in association with calnexin and in conjunction with folding enzymes until sites for calnexin binding are buried in the folded molecule. Whether folding (and assembly) occurs while the polypeptide is tethered to calnexin or during cycles of calnexin binding and release is unknown. Although calnexin does not possess consensus sequences for nucleotide binding, it has recently been demonstrated that the ER luminal domain of calnexin binds ATP in vitro (Ou et al., 1994). This raises the possibility that calnexin could undergo cyclic interactions with unfolded glycoproteins in a manner analogous to other chaperones. Conceptually, the two step binding model is reminiscent of the mechanism that regulates leukocyte localization in the vasculature. Circulating leukocytes are brought into proximity with endothelial cells via transient selectin-carbohydrate interactions followed by tight adhesion mediated by integrins and Ig superfamily adhesion receptors (Springer, 1994).

The existence of monoglucosylated oligosaccharides on newly synthesized glycoproteins is prolonged by a cycle of deglucosylation by glucosidase II and re-glucosylation via an ER enzyme known as UDP-glucose:glycoprotein glucosyltransferase (Gallán et al 1991 and Trombeta and
Two-step model for calnexin binding to newly-synthesized glycoproteins

Figure 8. Two-step model for binding of calnexin to unfolded glycoproteins. Following removal of two glucose residues, newly synthesized glycoproteins initially contact calnexin via their monoglucosylated oligosaccharide chains (Step 1). Having been placed in proximity to calnexin by this first interaction, the unfolded polypeptide associates directly with additional sites on calnexin (Step 2). The oligosaccharide chains are accessible to exogenous probes at this stage. Reglucosylation by UDP-glucose:glycoprotein glucosyltransferase may play an important role in recovering proteins that have lost all three glucoses prior to any contact with calnexin or in facilitating re-binding to calnexin during cycles of folding. Asterisks indicate steps blocked by castanospermine and 1-deoxynojirimycin.
Parodi, 1992). The latter enzyme re-glucosylates only non-native glycoproteins and it has been suggested by Helenius that the purpose of the cycle is to ensure that non-native glycoproteins continually oscillate between calnexin-bound and unbound states. Once a glycoprotein folds it is no longer a substrate for reglucosylation and it dissociates from calnexin (Helenius, 1994). Although attractive, this model requires that oligosaccharide structure is the main regulator not only of calnexin binding but of release as well. This is inconsistent with the observation that oligosaccharides do not participate in maintaining calnexin-glycoprotein complexes. Rather, the data presented here suggest that the function of reglucosylation may be to provide newly synthesized glycoproteins with additional opportunities to bind calnexin on those occasions when all three glucoses are removed before an initial interaction with calnexin can take place (Figure 8, recovery pathway). This may explain the increased level of association with calnexin that occurs for some glycoproteins during the 5-10 min period post synthesis (Oc et al., 1993 and Hammond et al., 1994). Reglucosylation followed by two-step binding may also be the means whereby a folding glycoprotein can re-bind to calnexin if dissociation from calnexin occurs before folding is complete (Figure 8).

Why has calnexin evolved to utilize oligosaccharide binding for its initial interaction with unfolded glycoproteins? Unlike soluble chaperones of the ER, calnexin is constrained within the ER membrane and, at least in some cell types, it may be associated with components of the translocation apparatus (Wada et al., 1991). Given this disposition, polypeptide binding sites on calnexin may have limited access to nascent glycoproteins. Oligosaccharide addition, being among the first covalent modifications that occur on nascent chains, ensures that a conserved and well-exposed site for calnexin binding is present at an early stage in the folding of nascent...
chains. Calnexin could even be associated with glucosidase I, poised to capture the monoglcosylated oligosaccharide as soon as it is formed. For some proteins, however, this initial stage of oligosaccharide capture can be by-passed since they bind to calnexin even though they are unglycosylated. This is the case for the T cell receptor α subunit (Rajagopalan et al., 1994) and also for variants of the multi-drug resistance P glycoprotein (Loo and Clarke, 1994) and the class I H-2 L3 heavy chain (Capps and Zuilige, 1993) that lack consensus sequences for N-glycosylation. Presumably these proteins have accessible sites on their polypeptide chains for interaction directly with calnexin.

As suggested previously (Helenius, 1994), the finding that calnexin binds to the Gc,Man,GlcNAc, processing intermediate may explain why eukaryotes from yeasts to humans initiate N-glycosylation with a common, pre-assembled oligosaccharide (Figure 1). If the purpose of early attachment of oligosaccharide was solely to ensure that segments of a folding polypeptide remained exposed to solvent, then oligosaccharides of diverse size and composition would likely suffice. Calnexin is an abundant protein in virtually all eukaryotic cell types examined including yeast (in which it has an essential function) (Parlati et al., 1995 and Jannatipour et al., 1995), plants, worms, and mammals (see Bergeron et al., 1994 for references). Its intimate association with N-linked oligosaccharides as part of its quality control and chaperone functions may be the major factor responsible for preserving the "en bloc" mode of glycosylation that originated in early eukaryotes.
CHAPTER III
OLIGOSACCHARIDE BINDING CHARACTERISTICS OF THE MOLECULAR CHAPERONES, CALNEXIN AND CALRETIULIN

Please Note: All the data were produced by myself with the following exceptions:
For Figure 5 my contribution was in the preparation of immobilized calreticulin for oligosaccharide binding analysis performed in Dr. M. Lehrman's laboratory.

The results in this chapter are in press in the journal Biochemistry:
Aikaterini Vassilakou, Marik Michelak, Mark A. Lehrman, and David B. Williams (1998)
3.0 Abstract

Calnexin and calreticulin are homologous proteins that are thought to function as molecular chaperones of the endoplasmic reticulum. Their binding to newly synthesized glycoproteins is mediated, at least in part, by a lectin site that recognizes the early N-linked oligosaccharide processing intermediate, Glc,Man,GlcNAc. In an effort to determine the basis for reported differences in their association with various glycoproteins, the oligosaccharide binding specificities of calnexin and calreticulin were examined. Using mono-, di-, and oligosaccharides to inhibit the binding of Glc,Man,GlcNAc to calreticulin and to a truncated, soluble form of calnexin, we show that the entire Glcα1-3Manα1-2Manα1-2Man structure, extending from the α1-3 branch point of the oligosaccharide core, is recognized by both protein. Furthermore, analysis of the binding of monoglycosylated oligosaccharides containing progressively fewer mannose residues suggests that the α1-6 mannose branch point of the oligosaccharide core is also essential for recognition for both proteins. Consistent with their similar substrate specificities, calnexin and calreticulin exhibited the same relative affinities when competing for binding to the Glc,Man,GlcNAc oligosaccharide.

Thus, differential glycoprotein binding cannot be attributed to differences in the lectin specificities or binding affinities of calnexin and calreticulin. The effects of ATP, calcium, and disulfide reduction on the lectin properties of calnexin and calreticulin were also examined. Whereas oligosaccharide binding was only slightly enhanced for both proteins in the presence of high concentrations of a number of adenine nucleotides, removal of bound calcium abrogated oligosaccharide binding, an effect that was largely reversible upon re-addition of calcium. Disulfide reduction had no effect on oligosaccharide binding by calnexin but binding by
calreticulin was inhibited by 70%. Finally, deletion mutagenesis of calnexin and calreticulin mapped their oligosaccharide binding sites to a central proline-rich segment characterized by two tandem repeat motifs. This segment bears no sequence homology to the carbohydrate recognition domains of other lectins and thus constitutes a novel lectin binding site.

3.1 Introduction

Calreticulin and calnexin are homologous calcium binding proteins that reside within the endoplasmic reticulum (Williams, 1995). Whereas calreticulin is a soluble, luminal protein, calnexin is a type I membrane protein with most of its mass luminaly disposed (Williams, 1995 and Michalak et al., 1992). Segments of these proteins share amino acid sequence identity ranging from 42% to 78% (Wada et al., 1991). The most striking and highly conserved segment contains two tandem repeat motifs, repeated four times each in calnexin and three times each in calreticulin (Michalak et al., 1992, Wada et al., 1991 and David et al., 1993). This region, which in calreticulin is known as the proline-rich P domain, also contains a site for high affinity calcium binding (Bakhsh and Michalak, 1991 and Tjoelker et al., 1994).

As outlined in the Introduction, both calreticulin and calnexin bind transiently to diverse membrane or secretory glycoproteins that are translocated into the ER. They also exhibit prolonged interaction with mutant glycoproteins that fail to fold or assemble correctly (Williams, 1995, Helenius, 1994, Nauseef et al., 1995, Peterson et al., 1995 and Otteken and Moss, 1996). During these interactions, calnexin functions as part of the ER quality control system that prevents the export of misfolded or incompletely assembled glycoproteins along the
A number of recent studies have provided evidence that calnexin also acts as a molecular chaperone that facilitates the folding and assembly of glycoproteins (see next chapter for details). Specifically, in the absence of calnexin binding, the formation of fully oxidized forms of the vesicular stomatitis virus G glycoprotein or the heavy chain of the human class I histocompatibility molecule is impaired or delayed, respectively (Hammond and Helenius, 1994 and Tector and Salter, 1995). Likewise, murine class I histocompatibility molecules synthesized in the absence of calnexin assemble inefficiently due to heavy chain misfolding and aggregation (Vassilakos et al., 1996). Calnexin also protects many nascent glycoproteins from rapid intracellular degradation (Jackson et al., 1993, Kearse et al., 1994 and Moore and Spiro, 1993). Although calreticulin-deficient mouse embryos have been produced (giving rise to an embryonic lethal phenotype), the effect of calreticulin depletion on the folding of nascent glycoproteins has not yet been studied (Coppolino et al., 1997). However, it is likely that calreticulin possesses chaperone properties. When the interactions of calreticulin and calnexin with influenza hemagglutinin are simultaneously prevented with castanospermine treatment the overall efficiency of hemagglutinin maturation decreases (Hebert et al., 1996).

One of the most remarkable aspects of calnexin and calreticulin function is their striking preference for binding to Asn-linked glycoproteins. The basis for this selectivity arises from the fact that both proteins are lectins that recognize the oligosaccharide processing intermediate Glc,Man,GlcNAc₂ (Ware et al., 1995, Spiro et al., 1996, Hammond et al., 1994 and Hebert et al., 1995). The results of the Ware et al. study are presented in Chapter II. In an earlier study from Spiro’s group, calreticulin was identified as a 60 kDa polypeptide that was co-isolated with a rat
liver Golgi endomannosidase preparation (Hiraizumi et al., 1994). The preparation required chromatography on a Glcα1-3Man containing matrix, suggesting that calreticulin acts as a lectin. A subsequent study, using immobilized calreticulin and a series of radiolabelled oligosaccharides, identified optimal binding to the monoglucosylated Glc,Man,GlcNAc, oligosaccharide (Spiro et al., 1996). Trimming of the polymannose branch demonstrated that binding could occur as far down as Glc,Man,GlcNAc, but the α1-6Man branch point was essential for binding to calreticulin. Furthermore, substrate specificity studies have identified the single terminal glucose residue as a critical determinant recognized by both chaperones since oligosaccharides containing 0, 2, or 3 glucose residues fail to bind (Ware et al., 1995 and Spiro et al., 1996).

Oligosaccharide binding is clearly crucial for the formation of complexes between glycoproteins and calnexin or calreticulin. If formation of the Glc,Man,GlcNAc, oligosaccharide is blocked with tunicamycin or if production of the Glc,Man,GlcNAc, species is prevented by treatment with the glucosidase inhibitors, castanospermine or deoxynojirimycin, the binding of calnexin/calreticulin to the vast majority of proteins is inhibited (Nauseef et al., 1995, Peterson et al., 1995, Hammond et al., 1994 and Ou et al., 1993). It has been suggested that these chaperones interact with nascent glycoproteins in cycles of binding and release, regulated by glucosidase II which removes the single glucose residue from the Glc,Man,GlcNAc, oligosaccharide and by UDP-glucose: glycoprotein glucosyltransferase which re-attaches it (Hebert et al., 1995). However, it is not apparent how such a lectin-only mode of interaction can give rise to the observed molecular chaperone functions of these molecules, e.g., the suppression of aggregates. There is abundant evidence suggesting that calnexin and calreticulin recognize the
polypeptide segments of newly synthesized glycoproteins as well (discussed extensively in Chapter 11). Consequently, an alternative model that incorporates both oligosaccharide and polypeptide interactions into the above cycle was proposed in the previous chapter (Chapter 11, Figure 9). Since there are in vitro binding experiments that support either a lectin-only (Rodan et al., 1996 and Zapun et al., 1996, discussed in Chapter V) or a lectin/polypeptide dual mode of binding (Ware et al., 1995, Carrero et al., 1995a, Zang et al., 1995, Arunachalam and Cresswell, 1995, Margolese et al., 1993, Rojiani et al., 1991, van Leeuwen and Kearse, 1996a, Baksh et al., 1995 and Nigam et al., 1994, discussed in Chapter 11), it is possible that different modes of interaction may be used for individual glycoproteins. Clearly, further experiments are required to resolve this issue.

A number of studies have examined cofactors that may influence the structure and function of calnexin and calreticulin. Mg-ATP has been shown to bind to calnexin resulting in oligomerization and increased sensitivity to protease digestion (Ou et al., 1995). In vivo studies have suggested a role for ATP in maintaining the association of calnexin with a soluble glycoprotein in Madin-Darby canine kidney cells but it is unclear if this effect is due to a direct interaction of nucleotide with calnexin (Wada et al., 1995). Indirect evidence suggests that ATP may bind to calreticulin as well, but in contrast to its effect on calnexin in vivo, ATP triggered the dissociation of calreticulin from a variety of denatured proteins in vitro (Nigam et al., 1994). Both calnexin and calreticulin bind Ca²⁺ (Baksh and Michalak, 1991, Tjoelker et al., 1994). In the case of calnexin, Ca²⁺ appears to have the opposite effect of Mg-ATP since a monomeric, protease-resistant conformation accompanies Ca²⁺ binding (Ou et al., 1995). Ca²⁺ is also required for the binding of calnexin to glycoproteins as assessed either in vivo (Capps et al., 1994) or in
The effects of Ca$^{2+}$ on the structure of calreticulin have not been established but it does not appear to be required for the binding of calreticulin to denatured proteins. Indeed, the addition of Ca$^{2+}$ enhanced the ATP-induced dissociation of calreticulin from denatured proteins. By contrast, Ca$^{2+}$ has been shown to be required for the binding of calreticulin to laminin. Although preliminary and sometimes conflicting, these studies nevertheless implicate ATP and Ca$^{2+}$ as cofactors that affect the structure and function of calnexin and calreticulin. Whether or not these cofactors exert their effects by modulating the lectin functions of these chaperones remains to be addressed.

It is unclear why the ER of most eukaryotic species contains two homologous chaperones with lectin properties. This may reflect redundancy in the ER quality control and protein folding systems or the two chaperones could conceivably function in a coordinated fashion. Although redundant action is suggested by several studies in which calreticulin and calnexin were shown to bind to the same glycoprotein substrates, other work has demonstrated differences either in the spectrum of glycoproteins bound or in the stage of glycoprotein maturation recognized by each chaperone. The study presented in this chapter provides a detailed examination of the oligosaccharide binding sites in calnexin and calreticulin in an effort to better understand the nature of their interaction with glycoproteins and to discern any differences that might account for their function.
differential binding to various glycoproteins. The data indicate that binding interactions between calnexin or calreticulin and the GlcManGlcNAc oligosaccharide are much more extensive than previously thought (Ware et al., 1995 and Spiro et al., 1995) and include the entire glycosylated branch of the oligosaccharide. Furthermore, the two proteins bind the GlcManGlcNAc oligosaccharide with similar affinity and are indistinguishable in terms of their oligosaccharide binding specificities. The effect of Ca\(^{2+}\) and Mg-ATP on oligosaccharide binding was examined and it was found that bound Ca\(^{2+}\) is essential to the lectin functions of both proteins. Experiments presented here were unable to reproduce previously reported conformational changes associated with Mg-ATP binding to calnexin and, consistent with this finding, observed only minimal enhancements of oligosaccharide binding to calnexin or to calreticulin in the presence of nucleotide. Finally, the oligosaccharide binding sites in calnexin and calreticulin were localized and found that they both reside with a central homologous segment distinguished by two tandem repeat motifs. This segment constitutes a novel lectin binding site. Given that the lectin sites of these proteins are similar in terms of primary sequence, specificity, binding affinity, and the effect of cofactors, it is unlikely that they are responsible for differences observed in the binding of calnexin and calreticulin to nascent glycoproteins. Rather, such differences may be due to differential recognition of the polypeptide chains of glycoproteins or to the distinct topological environments wherein the two chaperones reside.
3.2 Materials and Methods

3.2.0 Preparation of homogeneous [\(^{\text{3}}\)H]Glc,Man,GlcNAc oligosaccharides

Dolichol-linked [\(^{\text{3}}\)H]oligosaccharides from the \textit{S. cerevisiae \textit{alg}8} strain (Runge and Robbins, 1986), prepared in Dr. M. Lehrman's laboratory by metabolic labeling with [\(^{\text{3}}\)H]mannose (Ware et al., 1995), were used as starting material. Oligosaccharides were cleaved from dolichol pyrophosphate with mild acid, reduced with sodium borohydride, and desalted (Ware et al., 1995). The material from \textit{alg}8 cells was mostly Glc,Man,GlcNAc, (which yielded acetolysis fragments migrating as Man, Man, and Glc,Man,GlcNAc by HPLC) but also included smaller amounts of oligosaccharides tentatively identified as being Man, GlcNAc, which were readily separated from Glc,Man,GlcNAc by preparative HPLC (Ware et al., 1995). For other specific oligosaccharide fragments the Glc,Man,GlcNAc, was treated as indicated, and the fragments isolated by preparative HPLC. The structures of these fragments are indicated in Figure 1.

Glc,Man,GlcNAc, lacking the terminal mannose of the center branch was prepared by digestion with purified recombinant \textit{S. cerevisiae} ER \(\alpha,1,2\) mannosidase (the kind gift of A. Herscovics, McGill University). Glc,Man,GlcNAc, was obtained by acetolysis, a treatment which selectively cleaves \(\alpha,1,6\)-linked mannose residues. Incomplete cleavage of \(\alpha,1,6\) mannosyl linkages also yielded appreciable quantities of Glc,Man,GlcNAc, A distinct isomer of Glc,Man,GlcNAc, was obtained by digestion with purified, recombinant Golgi mannosidase I (kindly provided by K. Moreman, U. of Georgia), which removes all exposed \(\alpha,1,2\)-linked mannose residues. In all cases the purified fragments had the expected elution times based on
Figure 1. Preparation of Glc$_3$Man$_{1-3}$GlcNAc$_2$ oligosaccharides to calnexin.

Structure of the Glc$_3$Man$_{1-3}$GlcNAc$_2$ oligosaccharide recognized by calnexin. The indicated treatments were used to remove specific mannose residues (A-E) for the production of Glc$_3$Man$_{1-3}$GlcNAc$_2$ oligosaccharides.
the known properties of the HPLC column, which separates oligosaccharides on the basis of increasing sugar content. To verify that Glc,Man,GlcNAc₂ was, in fact, glucosylated, it was shown to be fully resistant to jack bean α-mannosidase treatment. By comparison, Man,GlcNAc₂ was highly susceptible whereas Man,GlcNAc₂ was resistant as expected.

[¹H] Glc,Man,GlcNAc₂ mixtures were isolated from [¹H]-mannose labeled CHO cells as described (Ware et al., 1995). Either this mixture or the predominantly Glc,Man,GlcNAc₂ preparation from αgal yeast cells was used in various experiments. The oligosaccharide preparation used for each experiment (CHO or αgal) is indicated in the figure legends.

3.2.1 Preparation of Soluble Proteins

His-tagged Proteins – The soluble class I histocompatibility H-2Kb heavy chain (with His₉ at the C-terminus) in association with β₂-microglobulin was purified by Erin Mitchell of our laboratory from the culture medium of transfected Drosophila melanogaster SC2 cells using Ni-NTA-agarose (Qiagen) and Mono Q (Pharmacia) anion exchange chromatography as described previously (Sykulev et al., 1994). Expression of soluble calnexin in baculovirus-infected Sf9 cells was accomplished using the MAXBAC baculovirus expression system from Invitrogen. The soluble calnexin construct (CNX-His) encoded the N-terminal signal sequence and ER luminal domain of calnexin followed by the sequence SRRSWGSHHHHHH (Ware et al., 1995). This was cloned into the pVI.1393 vector and was kindly provided by T. Jensen and J. R. Riordan, Mayo Clinic, Scottsdale. Erin Mitchell performed the purification as follows: Sf9 cells were maintained at 27°C in Grace's insect cell
medium (Life Technologies) supplemented with 3.33 g/L lactalbumin hydrolysate, 3.33 g/L yeastolate, 2 mM glutamine and 10% fetal bovine serum. For infection with recombinant baculovirus for protein production, the cells were grown at 24°C in roller bottles containing 150 ml of a 50:50 mixture of supplemented Grace's medium and Insect Xpress medium (Biowhittaker). The recombinant virus was added to a confluent culture of SF9 cells (15 ml was added to each roller bottle) and incubated with slow rotation for 72 hours. At 60 h and 72 h, the protease inhibitor AEBSF ([4-(2-aminoethyl)-benzenesulfonyl fluoride], Calbiochem) was added to the medium at a concentration of 0.2 mM. Culture supernatant (typically 1 liter from 7-8 roller bottles) was collected and cellular debris was removed by centrifugation at 2,500 x g for 20 minutes at 4°C. All of the following procedures were performed at 4°C unless otherwise specified. The culture supernatant was concentrated 5-fold by ultrafiltration through an Amicon YM30 membrane and dialyzed against three changes of HSC buffer [10 mM Hapes (pH 7.5), 150 mM NaCl, 10 mM CaCl₂]. This was loaded onto a 5 ml Ni-NTA-agarose column (Qiagen) equilibrated with HSC buffer and a flow rate of 0.5-0.75 ml/min was established. After sample loading, the column was washed with two column volumes of HSC buffer and then with 20 mM imidazole in HSC buffer. The 20 mM imidazole wash was continued until the OD₂₈₀ of the effluent dropped to <0.01 (approximately 10 column volumes). The soluble CNX-His was eluted with 250 mM imidazole in HSC buffer and 1 ml fractions were collected. Proteins in the fractions were resolved by SDS-PAGE and visualized by both Coomassie blue staining and immunoblotting with anti-Npp90 to identify calnexin-containing fractions. Anti-Npp90 is a rabbit antiserum raised against amino acids 8-268 of canine calnexin and was the kind gift of J. M. Bergeron, McGill University (Wada et al., 1991). Fractions containing CNX-His were
pooled and dialyzed against 20 mM Tris (pH 8.5), 10 mM CaCl₂ prior to separation on a fast protein liquid chromatography Mono Q column (10/10) (Pharmacia) equilibrated with the same buffer at room temperature. Proteins were eluted with a linear gradient of 200-700 mM NaCl over 40 min at 2 ml/min. Fractions (1-1.5 ml) were collected and samples were subjected to SDS-PAGE and visualized by Coomassie blue staining and immunoblotting with anti-Npp90. Fractions containing homogeneous CNX-His were pooled and dialyzed against 10 mM Tris (pH 7.5), 10 mM CaCl₂.

Production of His-tagged calreticulin was carried out in Dr. M. Michalak’s laboratory (University of Alberta, Edmonton) as follows: PCR was used to synthesize cDNA encoding full length mature calreticulin. PCR products were purified by polyacrylamide gel electrophoresis and ligated into pRSETB (Invitrogen) to generate a carboxy-terminal fusion of calreticulin with His. Transformed E. coli K38 cells were grown overnight at 30°C and then were collected by centrifugation for 15 min at 3,000 x g in a GSA rotor. The pellet was suspended in PBS containing 0.1% Triton X-100 followed by sonication for 3 x 45 seconds. The suspension was then centrifuged for 12 min at 8,000 x g. The supernatant fraction was filtered and loaded onto a Zn⁺⁺ chelating Sepharose column (Pharmacia) equilibrated with a buffer containing 50 mM Na₂HPO₄, pH 7.6, and 100 mM NaCl. Protein fractions were eluted with a 0 to 1 M imidazole gradient, analyzed by SDS-PAGE, pooled, and then concentrated using an Amicon YM30 membrane. CRT-His was further purified using an FPLC Resource Q column. The sample was loaded at a flow rate of 1 ml/min onto a 6 ml column equilibrated with buffer containing 20 mM Na₂HPO₄, pH 7.6, and 100 mM NaCl. CRT-His was eluted with a 100 to 750 mM NaCl gradient containing 50 mM Na₂HPO₄, pH 7.6. Protein fractions were analyzed by SDS-PAGE, pooled, and
Glutathione S-transferase (GST)-fusion Proteins – GST-rabbit calreticulin fusions

(full-length, or encompassing calreticulin residues 1-182, 1-273, 139-273, 139-401, and 270-401) were expressed and purified in M. Michalak’s laboratory from *E. coli* as previously described (Baksh and Michalak, 1991). GST-CRT (139-220) was generated by A. Vassilakos by digestion of the pGEX vector (Pharmacia) containing CRT fragment 139-401 with Bsi WI and Eco RI to remove nucleotides corresponding to CRT residues 321-401. The digest was treated with mung bean nuclease to generate blunt ends and the plasmid was re-circularized using T4 DNA ligase. The ligation reaction was transformed into DH5α cells.

GST-calnexin constructs (GST-CNX) were prepared by A. Vassilakos as follows (see Figure 2 for details). Canine calnexin cDNA in the Blaescript vector was modified by first removing a Dsa I site in the vector followed by insertion of an Eco RI 8mer linker into the Dsa I restriction site positioned within the calnexin coding sequence at the luminal side of the transmembrane domain. A DNA fragment containing the sequence for the entire luminal domain of calnexin was generated by Bsp HI digestion, incubation with Klenow to generate a blunt 5' end followed by Eco RI digestion. The calnexin fragment (encoding amino acids 1-461) was subcloned into the Sma I and Eco RI sites of pGEX 3X (Pharmacia). This construct, GST-CNX, was used to generate a C-terminal deletion (a.a. 1-391) by digesting with Bgl II (partial digest) and Eco RI. The digests were incubated with Klenow fragment to generate blunt ends and then the vector was gel purified, re-circularized using T4 ligase, and used to transform DH5α cells. The deletion mutant 204-391 was generated by Bgl II/Bam HI digestion of mutant 1-391, incubation with
Figure 2. Flowchart of calnexin-GST fusion construction.

All constructs were sequenced (HSC/Pharmacia Biotechnology Service Centre, Toronto), and tested for expression in E. coli.
Klenow, and re-circularization of the plasmid as above.

Fusion proteins were expressed in *E. coli* and isolated by single step purification using glutathione-agarose (Sigma). A 1 ml aliquot of overnight cultures was diluted to 10 ml and grown for 90 min. Isopropylthio-β-D-galactoside (0.1 mM) was added to induce synthesis of fusion protein and cultures were incubated for an additional 3 h. Bacteria were isolated by centrifugation for 10 minutes at 4,000 x g and resuspended in 0.5 ml solubilization buffer [10 mM Hepes, pH 7.5, containing 0.15 M NaCl, 10 mM CaCl₂, and 1% Nonidet P40 (NP-40)].

Cells were disrupted by sonication and cell debris was removed by centrifugation for 5 min at 12,000 x g. Fusion proteins were isolated by incubation for 30 min at 4°C with 50-100 ml glutathione-agarose. The beads were washed twice with 1 ml solubilization buffer and three times with binding buffer (10 mM Hepes, pH 7.5, containing 0.15 M NaCl and 10 mM CaCl₂).

Purified proteins were left bound to the agarose beads for use in oligosaccharide binding assays. They were typically >80% pure as assessed by Coomassie blue staining of SDS-PAGE gels.

### 3.2.2 Incubation of Radiolabeled Oligosaccharides with Immobilized Proteins

Each assay employed 5 μg samples of CNX-His, CRT-His, GST-CRT, or GST-CNX molecules immobilized on 5-10 μl of Ni-agarose or glutathione-agarose in binding buffer (10 mM Hepes, pH 7.5, containing 0.15 M NaCl, and 10 mM CaCl₂). Unless indicated otherwise, all procedures were performed at 21-23°C. Agarose-beads were washed briefly with binding buffer just prior to each experiment and then were suspended in 100 μl binding buffer which contained either no additions, competitive inhibitors (protein or mono-, di-, and oligosaccharides),
nucleotides, dithiothreitol (DTT) or ethylene glycol-bis-(β-aminopropyl ether)
N,N',N'-tetraacetic acid (EGTA). Incubations involving DTT or EGTA included a 30 minute
pre-treatment at 30°C prior to addition of oligosaccharide. Unless otherwise indicated,
approximately 4,000 cpm of oligosaccharides isolated from the dolichol-oligosaccharide fraction
of Chinese hamster ovary cells (mixture of Glc₃Man,GlcNAc₃) or 2000 cpm of
oligosaccharides isolated from αlβ yeast cells (predominantly Glc₃Man,GlcNAc₃) were added.
The specific oligosaccharide preparation used in each experiment is indicated in the figure
legends. The samples were incubated for 1 hour with agitation on an orbital shaker at 200 rpm,
centrifuged for 5 seconds at 10,000 x g, and the supernatant fraction removed. The agarose
beads were rinsed briefly with 200 µl of binding buffer, centrifuged as above, and the
supernatant fraction removed using a Hamilton syringe. The agarose beads were boiled for 5 min
in 100 µl of binding buffer and bead-associated radioactivity was analyzed by liquid scintillation
counting. Results are reported as specific binding (radioactivity in each sample minus
background radioactivity associated with control incubations using either GST or K3-His(s).

3.2.3 EGTA and ATP treatment of Calnexin - Assessment of Protease Sensitivity and
Oligomerization State

To test the effects of EGTA and ATP on the protease sensitivity of calnexin, 5 µg samples of
CNX-His or GST were immobilized on 5-10 µl of nickel-agarose or glutathione-agarose,
respectively, in binding buffer. For ATP-treated samples, beads were suspended in 20 µl of
binding buffer containing 30 µg/ml proteinase K, either 1 mM or 10 mM ATP, and 5 mM MgCl₂
For samples treated with EGTA, the agarose beads were washed three times with 1 ml of calcium-free binding buffer and then suspended in 20 μl of calcium-free binding buffer containing 30 μg/ml protease K and 10 mM EGTA. pH 7.5 was maintained in all of the incubations. After incubation at 30 °C for 30 min, the samples were boiled in sample buffer and analyzed by SDS-PAGE. Resolved proteins were visualized by Coomassie Blue staining.

For determination of oligomerization state, 5 μg samples of CNX-His or GST (not immobilized) were incubated for 30 min at 30°C in 100 μl binding buffer containing 10 mM ATP and 5 mM MgCl₂ or in calcium-free binding buffer containing 10 mM EGTA as indicated. After incubation, samples were analyzed on 7% non-denaturing polyacrylamide gels (Laemmli gels omitting the SDS) and resolved proteins were visualized by Coomassie Blue staining.

3.3 Results

To study the oligosaccharide binding properties of calnexin and calreticulin, the soluble ER luminal domain of calnexin was expressed with a His, sequence at its C-terminus (CNX-His) and calreticulin was expressed either with a His, sequence at its C-terminus (CRT-His) or as a fusion protein with glutathione S-transferase (GST-CRT). The recombinant proteins were immobilized either on Ni-agarose or glutathione-agarose. In preliminary experiments it was determined that CNX-His, CRT-His, and GST-CRT bound [3H]Glc,Man,GlcNAc,oligosaccharide 20- to 40-fold above background binding to the corresponding control proteins, H-2K²-His and GST, and that binding was saturable (Figure 3). All subsequent experiments were performed with
Figure 3

A. Relative oligosaccharide binding by CNX-His, GST-CRT, GST and Kb-HIS
Equal amounts of each protein immobilized on Ni-agarose (Calnexin-HIS and Kb-HIS) or glutathione agarose (GST-CRT and GST) were incubated with radiolabeled oligosaccharide and bound radioactivity was determined. The radioactivity determined for a buffer sample is subtracted from the value for each protein. The results are presented relative to either GST-CRT (striped boxes) or CNX-His (solid boxes).

B. Calnexin binding to radiolabeled oligosaccharide is saturable.
Equal amounts of radiolabeled CNX-His (5 µg) were incubated with increasing amounts of oligosaccharide as indicated. Specific binding was determined by subtraction of radioactivity associated with GST incubated with equivalent amounts of oligosaccharide.

(CHO oligosaccharide)
sub-saturating levels of oligosaccharide and the results were expressed in terms of specific binding [radioactivity bound to CNX-His, CRT-His, or GST-CRT minus radioactivity bound to the corresponding control protein (typically 10-20 cpm)].

3.3.0 Oligosaccharide binding specificity of calnexin

We showed previously that the ER luminal domain of calnexin, when incubated with a mixture of Glc, Man, GlcNAc$_2$ oligosaccharides, bound selectively to the Glc, Man, GlcNAc$_2$ species (Ware et al., 1995 and Chapter II). Thus, the single glucose residue is an important determinant for recognition by calnexin (see Figure 1 for oligosaccharide structure). To identify additional components of the oligosaccharide that are recognized by calnexin, a variety of mono-, di- and oligosaccharides were tested for their abilities to inhibit the binding of [H]$^3$Glc, Man, GlcNAc$_2$ to immobilized CNX-His (Table 1). The efficacy of each compound was expressed as the concentration that inhibits oligosaccharide binding to 1 $\mu$M calnexin by 50%.

No inhibition was observed with the monosaccharides, a-methylglucoside or a-pNO$_2$-glucoside or mannoside. By contrast, disaccharides containing glucose $\alpha$1-3 linked to either glucose or mannose inhibited oligosaccharide binding. These findings indicate that a penultimate sugar residue is essential for inhibition and that calnexin does not distinguish between $C_2$ epimers of the penultimate residue. Preference for the $1$-$3$ linkage was evidenced by the lack of inhibition with Glc $\alpha$1-6 Glc and a 3-fold poorer inhibition with Glc $\alpha$1-4 Glc. Consistent with our previous findings, the terminal glucose residue is an important determinant for recognition by calnexin since Man $\alpha$1-3 Man was not an inhibitor.
Table 1

A. Inhibition of Glcα1-3ManGlcNAc2 binding to calnexin.

Equal samples of CNX-His immobilized on Ni-agarose (1 μM final concentration) were incubated with [3H]Glcα1-3ManGlcNAc2 oligosaccharide and various non-radioactive mono-, di-, and oligosaccharides as indicated. Radioactivity specifically bound to calnexin was determined and the results are presented as the concentration of inhibitor that produced 50% inhibition of binding.

<table>
<thead>
<tr>
<th>Inhibitor</th>
<th>Concentration providing 50% inhibition (mM)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Glc α pNO₂β, -OCH₃</td>
<td>&gt;&gt;2.5¹</td>
</tr>
<tr>
<td>Man α pNO₂β</td>
<td>&gt;&gt;2.5</td>
</tr>
<tr>
<td>Glc α1-3 Man</td>
<td>0.6</td>
</tr>
<tr>
<td>Glc α1-3 Glc</td>
<td>0.7</td>
</tr>
<tr>
<td>Glc α1-4 Glc</td>
<td>1.7</td>
</tr>
<tr>
<td>Glc α1-6 Glc</td>
<td>&gt;&gt;2.5</td>
</tr>
<tr>
<td>Man α1-3 Man</td>
<td>&gt;&gt;2.5</td>
</tr>
<tr>
<td>Glc α1-3 Man α1-2 Man</td>
<td>0.08</td>
</tr>
<tr>
<td>Man α1-3 Man α1-6 Man</td>
<td>&gt;&gt;2.5</td>
</tr>
<tr>
<td>Glc α1-3 Man α1-2 Man α1-2 Man</td>
<td>0.0045</td>
</tr>
</tbody>
</table>

* The disaccharides, lactose and sucrose, provided no detectable inhibition at 2.5 mM. Polymeric yeast mannan inhibited binding by 25% at a concentration of 0.75 mg/ml.

¹ No detectable inhibition at a concentration of 2.5 mM.
Glc α1-3 Glc and Glc α1-3 Man are present at the respective non-reducing termini of the
Glc-Man,GlNAC, and Glc-Man,GlNAC, oligosaccharide processing intermediates occurring
early in glycoprotein biosynthesis. However, only the Glc,Man,GlNAC, intermediate is
specifically recognized by calnexin (Ware et al., 1995 and Chapter II). Presumably, additional
interactions occur that provide for the observed specificity of calnexin. To test this possibility,
tri- and tetra-saccharides that reflect the structure of the glucosylated arm of the
Glc,Man,GlNAC, oligosaccharide (see Figure 1) were tested as inhibitors. Glc α1-3 Man α1-2
Man and Glc α1-3 Man α1-2 Man α1-2 Man were ~8-fold and ~130-fold more potent inhibitors
than the disaccharides, respectively (Table I). These findings indicate that calnexin recognizes
all four sugar residues on the glucosylated α3 branch of the Glc,Man,GlNAC, oligosaccharide
(Figure 1, bold type).

Interestingly, yeast mannan (a heterogeneous polymannose structure containing an α1-6
backbone and branched α1-2 and α1-3 linkages) also inhibited oligosaccharide binding,
suggesting that additional mannose residues may be involved in interactions with calnexin
(Table I). This possibility was tested by generating a series of radiolabeled oligosaccharides
containing 4-9 mannose residues and assaying for their binding to immobilized CNX-His. The
method of generation of these oligosaccharides and their structures are shown in Figure 1. In
Figure 4, specific binding of each oligosaccharide is plotted as a function of input radioactivity
to normalize for any differences in specific activities of the oligosaccharides. With the exception
of Glc,Man,GlNAC, all of the oligosaccharides bound to calnexin. The slopes of the curves for
each of the oligosaccharides that bound are within the range of variability observed for
Figure 4. Binding of Glc,Man,GlcNAc, oligosaccharides to calnexin.

Identical samples of immobilized CNX-His were incubated with increasing amounts of each of the Glc,Man,GlcNAc, oligosaccharides described in Figure 1. Bound radioactivity was determined and the results are presented as specific binding vs. input radioactivity. G1M9, Glc,Man,GlcNAc; G1M8, Glc,Man,GlcNAc; G1M7 (man), Glc,Man,GlcNAc, prepared by Golgi α-mannosidase treatment; G1M7 (acet), Glc,Man,GlcNAc, prepared by acetolysis; G1M4, Glc,Man,GlcNAc, (alg8 oligosaccharide)
Glc,Man,GlcNAc (compare the two Glc,Man,GlcNAc curves in Figure 4), suggesting that
calnexin binds to each with similar affinity. Clearly, calnexin does not require the two outermost
mannose residues ("A" and "B" in Figure 1) nor the penultimate mannose residue "D" for
binding. Given that there was no binding to Glc,Man,GlcNAc, it appears that an internal
mannose residue on the non-glucosylated arm of the Glc,Man,GlcNAc, oligosaccharide is also
required for recognition by calnexin.

3.3.1 Comparison of calnexin and calreticulin binding to the Glc,Man,GlcNAc,
oligosaccharide

Spiro and co-workers recently reported that calreticulin, like calnexin, binds the
Glc,Man,GlcNAc, oligosaccharide (Spiro et al., 1996). We confirmed this finding by incubating
CRT-His, immobilized on Ni-agarose, with a mixture of radiolabeled Glc,Man,GlcNAc, oligosaccharides, eluting bound oligosaccharide, and analyzing by HPLC (Figure 5). Shown are
the starting oligosaccharide mixture (top panel), the material eluted from an immobilized control
protein (middle), and the material eluted from CRT-His. Calreticulin bound specifically to the
Glc,Man,GlcNAc, species.

To further compare the oligosaccharide binding specificities of calnexin and calreticulin, we
tested a variety of mono-, di-, and oligosaccharides for their abilities to inhibit the binding of
[3H]Glc,Man,GlcNAc, to GST-CRT immobilized on glutathione-agarose (Table 2). The results
were remarkably similar to those obtained above with CNX-His (compare Table 2 with Table 1).
The smallest inhibitory compounds were the disaccharides, Glc α1-3 Man and Glc α1-3 Glc,
with calreticulin exhibiting a distinct preference for the 1-3 linkage over the 1-4 or 1-6 linkages.
Figure 5. Selective binding of Glc₃Man₃GlcNAc₂ oligosaccharide by immobilized calreticulin.

A mixture of [³H]Glc₃Man₃GlcNAc₂ oligosaccharides isolated from Chinese hamster ovary cells (Ware et al., 1995) was dissolved in binding buffer, divided into three equal portions, and incubated for 1 h at 23°C either alone (top), or with 6.25 µg of H-2Kb-His (middle) or CRT-His (bottom) immobilized on Ni-agarose. The beads were washed briefly with binding buffer, and [³H]oligosaccharides were eluted with two sequential 1 h washes in binding buffer containing 0.1 M α-methylmannoside and 0.1 M α-methylglucoside. The two eluates were combined, desalted, and characterized by HPLC with a Radiomatic A-140 in-line isotope detector as described earlier (Ware et al., 1995). Elution profiles are shown for each incubation and the positions of Glc₃Man₃GlcNAc₂, Glc₃Man₃GlcNAc₂, Glc₅Man₃GlcNAc₂, and Glc₅Man₃GlcNAc₂ are indicated at the top.
Table 2

*Inhibition of GlcManGlcNAc binding to calreticulin.*

Equal samples of GST-CRT immobilized on glutathione-agarose (0.7 μM final concentration) were incubated with [3H]GlcManGlcNAc oligosaccharide and potential inhibitors as indicated. Specifically bound radioactivity was determined and the results are presented as the concentration of inhibitor that produced 50% inhibition of binding.

<table>
<thead>
<tr>
<th>Inhibitor</th>
<th>Concentration providing 50% inhibition (mM)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Glc α OCH₃</td>
<td>&gt;&gt;2.5†</td>
</tr>
<tr>
<td>Glc α1-3 Man</td>
<td>0.5</td>
</tr>
<tr>
<td>Glc α1-3 Glc</td>
<td>0.4</td>
</tr>
<tr>
<td>Glc α1-4 Glc</td>
<td>1.8</td>
</tr>
<tr>
<td>Glc α1-6 Glc</td>
<td>&gt;&gt;2.5</td>
</tr>
<tr>
<td>Glc α1-3 Man α1-2 Man</td>
<td>0.005</td>
</tr>
<tr>
<td>Glc α1-3 Man α1-2 Man α1-2 Man</td>
<td>0.001</td>
</tr>
</tbody>
</table>

† No detectable inhibition at a concentration of 2.5 mM.
Furthermore, the trisaccharide, Glc α1-3 Man α1-2 Man, and the tetrasaccharide, Glc α1-3 Man α1-2 Man α1-2 Man α1-2 Man, were ~100-fold and ~500-fold more potent inhibitors than the disaccharides, respectively. These findings indicate that calreticulin, like calnexin, recognizes all four sugar residues on the glucosylated α3 branch of the Glc,Man,GlcNAc,oligosaccharide.

Spiro and co-workers have also reported that the innermost α6-linked Man residue of the Glc,Man,GlcNAc,oligosaccharide is essential for recognition by calreticulin (residue "E" in Fig. 1) (Spiro et al., 1996). This is consistent with our finding that calnexin is unable to bind the Glc,Man,GlcNAc, species that lacks this residue (Figure 4). Collectively, the results suggest that calreticulin and calnexin share identical oligosaccharide binding specificities. If so, it would be expected that they exhibit the same relative binding affinities for the Glc,Man,GlcNAc, oligosaccharide. To test this possibility, GST-CRT was immobilized on glutathione agarose and incubated with radiolabeled oligosaccharide and increasing amounts of soluble CNX-His (Figure 6). CNX-His competed with GST-CRT for oligosaccharide binding with 50% inhibition occurring at 0.7 μM CNX-His, the same concentration as the immobilized GST-CRT in the assay. These findings suggest that the two lectins possess similar binding affinities for Glc,Man,GlcNAc,.

3.3.2 Effects of EGTA, ATP, and DTT on the lectin functions of calnexin and calreticulin

A previous report indicated that the ER luminal domain of calnexin binds both Ca²⁺ and Mg-ATP and that removal of Ca²⁺ or addition of Mg-ATP leads to oligomerization of calnexin and an increase in its sensitivity to exogenous protease (Ou et al., 1995). Consequently, it was of
Calnexin and calreticulin bind oligosaccharide with similar affinities. Identical samples of GST-CRT immobilized on glutathione-agarose (0.7 μM) were incubated with a sub-saturating amount of \([\text{[^{3}H]}\text{Glc,Man,GlcNAc}]\) and increasing concentrations of soluble CNX-His (as indicated). Specifically bound radioactivity was determined and plotted against the concentration of soluble calnexin in the incubation. (CHO oligosaccharide)
Figure 7. Effects of EGTA and Mg-ATP on the conformation of calnexit.
A. Identical samples of CNX-His immobilized on Ni-agarose were incubated with or without 30 μg/ml proteinase K, 10mM EGTA, or 1 mM or 10 mM ATP as indicated for 30 min at 30°C. Total proteins were analyzed by SDS-PAGE and visualized with Coomassie blue staining. As a control protein, GST bound to glutathione-agarose was subjected to identical treatments.
B. Identical samples of CNX-His (not immobilized) were incubated with or without 10 mM EGTA, or 10 mM ATP as indicated for 30 min at 30°C. Incubation mixtures were analyzed by non-denaturing PAGE and proteins were visualized with Coomassie blue. GST was included as a control.
interest to test whether the conformational alterations induced by these treatments affects oligosaccharide binding. However, we first wished to confirm that the removal of Ca²⁺ and the addition of Mg-ATP are accompanied by conformational changes in calnexin. As shown in Figure 7A, CNX-His was incubated with or without EGTA or Mg-ATP as indicated and its sensitivity to digestion with proteinase K determined. Consistent with previous work (Ou et al., 1999), treatment of CNX-His with EGTA resulted in increased sensitivity to proteinase K digestion. Surprisingly, Mg-ATP did not alter the susceptibility of CNX-His to digestion. Neither treatment affected the protease sensitivity of GST which was included as a control. Identical samples were also examined by non-denaturing PAGE to determine whether EGTA or Mg-ATP treatment influenced the oligomerization state of these molecules (Figure 7B). Treatment of CNX-His with EGTA, but not Mg-ATP, resulted in increased oligomerization of calnexin. These results confirm that removal of bound Ca²⁺ from calnexin by EGTA treatment alters the conformation of calnexin such that the molecule is more sensitive to proteolysis and results in the formation of higher order oligomers. However, the lack of effect of Mg-ATP is in direct contrast to what has been previously reported (Ou et al., 1995).

Calreticulin also binds Ca²⁺ (Baksh and Michalsk, 1991) and has been suggested to bind Mg-ATP as well (Nigam et al., 1994). We attempted to assess conformational changes associated with removal of bound Ca²⁺ or the addition of Mg-ATP using a variety of proteases as conformational probes. Unfortunately, we were unable to obtain information on conformational changes since calreticulin was extremely sensitive to proteolysis regardless of whether or not these cofactors were present (Figure 8). Nevertheless, experiments to test whether the removal of
Figure 8. Effects of EGTA and Mg-ATP on the susceptibility of calreticulin to protease digestion.

Identical samples of CRT-His immobilized on Ni-agarose were incubated with or without 15 μg proteinase K, 10 mM EGTA, or 1 mM or 10 mM ATP as indicated for 20 min at 30°C. Total proteins were analyzed by SDS-PAGE and visualized with Coomassie blue staining.
bound Ca" or the addition of Mg-ATP influences the ability of either calnexin or calreticulin to bind the Glc,Man,GlcNAc, oligosaccharide were performed.

As shown in Figure 9, incubation with EGTA abolished oligosaccharide binding by CNX-His. This effect was completely reversible when 10 mM Ca" was re-introduced after EGTA treatment. Thus, the conformational changes associated with removal of bound Ca" appear to interfere with the function of calnexin's lectin site. Similar to the results obtained with calnexin, EGTA treatment profoundly inhibited oligosaccharide binding by CRT-His. Again, this effect could be reversed upon re-addition of Ca" although not as completely as observed for calnexin.

CNX-His was also incubated with radiolabeled oligosaccharide and a variety of nucleotides (Table 3). To exclude the possibility that the recombinant CNX-His was purified with bound ATP, CNX-His was treated with apyrase and found to have no effect on oligosaccharide binding (data not shown). There was also no effect of adding 1 mM ATP on oligosaccharide binding either in the presence or absence of 5 mM Mg". At 10 mM ATP there was a 60% enhancement of oligosaccharide binding, but this effect was independent of Mg". A similar enhancement was observed with 10 mM ATP"S, ADP and AMP (Table 3). The high concentrations of nucleotide required to produce this enhancement, the lack of Mg" requirement, and the fact that ATP, ADP, AMP, and ATP"S produced the same effect, suggest that the enhancement is non-specific. Interestingly, the phenomenon was not observed with 10 mM GTP. Oligosaccharide binding by CRT-His was also assayed in the presence of 10 mM ATP. Similar to the situation with CNX-His, an enhancement of oligosaccharide binding was reproducibly observed although it was somewhat less than that detected with CNX-His, i.e., 15-30% (data not shown).
Figure 9. Treatment of calnexin and calreticulin with EGTA abolishes oligosaccharide binding.

CNX-His and CRT-His immobilized on Ni-agarose were incubated with 10 mM and 20 mM EGTA, respectively, for 30 min at 30°C. The samples were cooled to room temperature and radiolabeled Glc$_4$-GlcNAc$_2$-oligosaccharide was added either directly or the beads were first washed to remove EGTA and then buffer containing 10 mM CaCl$_2$ was added prior to addition of oligosaccharide. As a control, the immobilized proteins were treated in the same manner as in the EGTA washout-CaCl$_2$ addition experiment but with the omission of EGTA. Specific oligosaccharide binding was determined and expressed as a percentage of the control value. Control values were 380 cpm for CNX-His and 150-360 cpm (depending on the experiment) for CRT-His. (tag8 and CHO oligosaccharide)
Table 3. Effects of nucleotide on oligosaccharide binding by calnexin

<table>
<thead>
<tr>
<th>Nucleotide</th>
<th>% Binding</th>
</tr>
</thead>
<tbody>
<tr>
<td>No addition</td>
<td>100 +/- 7.3</td>
</tr>
<tr>
<td>1 mM ATP + 5mM Mg++</td>
<td>100</td>
</tr>
<tr>
<td>1 mM ATP</td>
<td>100</td>
</tr>
<tr>
<td>10 mM ATP + 5mM Mg++</td>
<td>120</td>
</tr>
<tr>
<td>10 mM ATP</td>
<td>163 +/- 28</td>
</tr>
<tr>
<td>10 mM ATPγS</td>
<td>131 +/- 17</td>
</tr>
<tr>
<td>10 mM AMP</td>
<td>141 +/- 15</td>
</tr>
<tr>
<td>10 mM ADP</td>
<td>142 +/- 29</td>
</tr>
<tr>
<td>10mM GTP</td>
<td>100</td>
</tr>
</tbody>
</table>

Identical samples of CNX-His immobilized on Ni-agarose were incubated with radiolabeled oligosaccharide. 1mM or 10mM nucleotide and 5mM Mg ++ were included in the binding reactions as indicated. Specific oligosaccharide binding was determined and results are expressed relative to binding in the absence of any nucleotide (first row). Where possible the range is given (from experiments where the reactions were done in duplicate or triplicate). (alg6 and CHO oligosaccharide)
Calnexin contains at least one disulfide bond that has been implicated in its chaperone
function (Hebert et al., 1995, Ou et al., 1995 and Wada et al., 1995). Calreticulin also contains a
disulfide bond but its location differs from that in calnexin (Matsuoka et al., 1994). To address
whether disulfide bonds are required for the lectin properties of calnexin or calreticulin, samples
of CNX-His and CRT-His immobilized on Ni-agarose were incubated with 10 mM DTT for 30
min at 30°C. The DTT was removed with several rapid washes in binding buffer prior to
incubation with radiolabeled oligosaccharide. These conditions were sufficient to reduce
disulfide bonds in CNX-His and CRT-His as evidenced by their slower mobility on SDS-PAGE
relative to unreduced samples (Figure 10). Although disulfide reduction has been previously
shown to perturb calnexin function in vivo or in microsomes (Hebert et al., 1995 and Wada et al.,
1995) and to prevent ATP binding by calnexin's ER luminal domain in vitro (Ou et al., 1995), it
did not alter the oligosaccharide binding properties of CNX-His (Figure 11). By contrast,
treatment of CRT-His with DTT inhibited oligosaccharide binding by 70%. Thus, the single
disulfide bond in calreticulin appears to play an important role in the maintenance of a
functional oligosaccharide binding site.

There is the possibility that the treatments described above result in a loss of oligosaccharide
binding due to a loss of protein bound to the Ni-agarose beads. To discount this possibility,
CNX-His and CRT-His were treated with DTT, EGTA and ATP, the beads washed and the
bound material analyzed by SDS-PAGE and Coomassie staining. As shown in Figure 12, these
treatments did not result in a loss of material bound to the Ni-agarose beads.
Figure 10. Effect of DTT on the mobility of calnexin and calreticulin resolved by SDS-PAGE. CNX-His and CRT-His immobilized on Ni-agarose were incubated in 0 mM, 10 mM, or 60 mM DTT for 30 min at 30°C. The beads were subsequently washed several times in binding buffer to remove the DTT and run on non-reducing SDS-PAGE. The DTT treated samples were treated with NEM to block excess DTT.
Figure 11. Effect of disulfide reduction on oligosaccharide binding by calnexin and calreticulin.

CNX-His and CRT-His immobilized on Ni-agarose were incubated in the absence or presence of 10 mM DTT for 30 min at 30°C. The beads were subsequently washed several times in binding buffer to remove the DTT and then were incubated with radiolabeled oligosaccharide. Specific binding was determined and expressed as a percentage of the control sample that was not exposed to reducing agent. Control values were 120 cpm for CNX-His and 180-380 cpm (depending on the experiment) for CRT-His. (alg8 oligosaccharide)
Figure 12. Effects of EGTA, ATP and DTT on CNX-His and CRT-His on binding to Ni-agarose.

Identical samples of CNX-His and CRT-His, immobilized on Ni-agarose, were incubated for 30 min. at 30°C with EGTA, ATP or DTT as indicated. Control incubations under the same conditions but omitting EGTA or DTT were included as controls. After incubation, the samples were washed once with binding buffer and the amount of calnexin bound to the agarose beads was determined by boiling the beads in sample buffer, resolving on SDS-PAGE, and staining by Coomassie staining. GST is shown as a control.
3.3.3 Mapping the lectin sites of calnexin and calreticulin

To map the lectin sites in calnexin and calreticulin, deletion mutants were constructed as fusion proteins with glutathione S-transferase. Linear representations of calnexin and calreticulin are shown in Figure 13A with homologous regions depicted by large rectangles and the positions of two characteristic tandem repeat motifs indicated by the numbers 1 and 2. The deletion constructs are depicted in Figure 13B along with the results of binding assays with radiolabeled oligosaccharide. Given the small amounts of oligosaccharide available, it was not possible to assess the relative binding affinities of each construct but only whether or not it was capable of binding.

Both GST-CRT (residues 1-401) and GST fused to the ER luminal domain of calnexin (residues 1-461) bound oligosaccharide, indicating that for both chaperones fusion to GST supports proper folding of their lectin sites (Figure 13B). For calreticulin, constructs encompassing segments of the protein other than the tandem repeat sequences failed to bind oligosaccharide. These constructs correspond to the N-terminal 45% (residues 1-187) and the C-terminal 33% (residues 270-401) of the calreticulin polypeptide chain. By contrast, a central segment (residues 139-320) that encompassed all of the repeat sequences retained the ability to bind oligosaccharide, although binding was somewhat reduced relative to the full length molecule. Further truncation of this functional segment at its C-terminus (139-272), which removed part of the last motif 2 sequence, abolished oligosaccharide binding. Similar results were obtained for calnexin. Calnexin mutant 204-391, which also spanned the repeat motifs, was capable of binding oligosaccharide. Clearly, for both proteins, their lectin sites are located within the large homologous segment and are centered on the tandem repeat sequences.
Figure 13. Location of the oligosaccharide binding site in calnexin and calreticulin.

A. Schematic representations of calnexin and calreticulin showing their topology as well as regions of high sequence similarity between the two molecules (large rectangles). Repeat motif 1 (I-DPD/EAKPEDWD/E) and motif 2 (G-W--P-I-NP-Y) are indicated as are the high affinity calcium binding sites. In calreticulin, the segment containing the repeat motifs is also known as the proline-rich P domain.

B. Mapping the lectin site of calnexin and calreticulin. Deletion mutants of calnexin and calreticulin were expressed and purified as GST fusion proteins, immobilized on glutathione-agarose, and assayed for specific oligosaccharide binding. For these experiments, 2000–4000 cpm of oligosaccharides purified either from Chinese hamster ovary cells or from alg8 yeast cells were used. Binding was compared to GST-CRT and CNX-His or GST-CN in each experiment. Shaded boxes denote the segment containing the two repeat motifs.
3.4 Discussion

The results presented in Chapter II demonstrated that the ER luminal domain of calnexin has the capacity to bind to the early N-linked oligosaccharide processing intermediate, \( \text{Glc,Man,GlcNAc} \). Recently, the same observation was made for its soluble ER homologue, calreticulin (Spiro et al., 1996). In both studies, the single terminal glucose residue was shown to be crucial for recognition by these chaperones. To explore further the relationship between calnexin and calreticulin, the experiments presented here compare their oligosaccharide binding specificities in detail. The results demonstrate that the lectin sites in both chaperones recognize the entire glucosylated \( \alpha 3 \)-linked branch of the oligosaccharide, consisting of the sequence \( \text{Glc} \alpha 1-3\text{Man} \alpha 1-2\text{Man} \alpha 1-2\text{Man} \). Furthermore, not only the terminal glucose residue but also the \( 1-3 \) linkage to the penultimate mannose residue were shown to be important recognition elements. The possibility that residues located more internally than the tetrasaccharide are recognized as well, such as the \( \beta \)-linked mannose residue of the core, cannot be excluded (see Figure 1), because potential inhibitors larger than the tetrasaccharide were not available for testing. However, the Asn-linked GlcNAc residue can be eliminated as a possible recognition element since this residue was reduced to the corresponding alcohol in these studies, and oligosaccharides lacking this residue have previously been shown to bind well to calreticulin (Spiro et al., 1996).

The similarities in binding specificity between calnexin and calreticulin also extend to the \( \alpha 6 \)-linked branch of the oligosaccharide (residues A-E in Figure 1). That there might be interactions beyond those observed with the glucosylated arm of the oligosaccharide was initially suggested by the finding that polymeric yeast mannan inhibits calnexin binding to the
Glc,Man,GlcNAc, oligosaccharide (Table I) and also binds directly to calreticulin (White et al., 1995). Subsequent binding studies using glucosylated oligosaccharides possessing progressively fewer mannose residues revealed that calnexin is unable to bind the Glc,Man,GlcNAc, oligosaccharide but binds the Glc,Man, Man,GlcNAc, species (Figure 4). Thus terminal mannoses are not essential for binding but an internal Manα2-Man linkage appears to be important for recognition by calnexin. The same results were recently obtained by Spiro and co-workers for calreticulin. In addition, they were able to identify the determinant essential for recognition by calreticulin as the most internal Manα1-6Man branch point (Spiro et al., 1996 and residue E in Figure 1). Therefore, calnexin and calreticulin appear to possess identical oligosaccharide binding specificities. This conclusion is also supported by our demonstration that the two chaperones compete for the Glc,Man,GlcNAc, oligosaccharide with the same relative binding affinities.

It is interesting to examine the binding specificities of calnexin and calreticulin in light of the recently determined NMR structure of the Glc,Man,GlcNAc, oligosaccharide (Pterescu et al., 1997). The structure is highly extended from the single terminal glucose to the GlcNAc residue involved in linkage to asparagine. Consequently, the glucosylated α3-linked branch that is recognized by both chaperones is completely accessible. By contrast, the Manα1-6Man branch point that is also important for recognition is largely inaccessible. Molecular dynamics simulations suggest that the α6-linked Man could be involved in an extensive hydrogen bond network involving water molecules and the α3-linked Man of the glucosylated branch, resulting in the glucosylated branch adopting a single conformation. Loss of the α6-linked Man could result in a more flexible glucosylated branch and hence a structure not recognizable by either
chaperone (R. Woods, M. Wormald, and R. Dwek, personal communication). The structure of the Glc,Man,GlcNAc oligosaccharide is currently being determined to test the validity of this suggestion.

Using deletion analysis, the lectin sites of calnexin and calreticulin were mapped to a homologous region of the two molecules that contains their distinctive tandem repeat motifs and also their site of high affinity calcium binding. It was possible to remove in excess of one-third of their N-terminal and 15-20% of their C-terminal (ER luminal) segments without loss of oligosaccharide binding. By contrast, partial or complete deletion of the repeat sequences abrogated binding. The co-localization of the oligosaccharide and calcium binding sites is reminiscent of the carbohydrate-recognition domains of C-type lectins. This class of lectin requires calcium for oligosaccharide binding in a process involving direct interactions between sugar hydroxyl groups and the metal ion (Drickamer, 1993). However, the lectin sites in calnexin and calreticulin lack the consensus sequence characteristic of the carbohydrate-recognition domains in C-type lectins or, for that matter, in other major lectin classes as well. Consequently, a novel type of lectin binding site has been defined which is probably defined by the motif 1 and motif 2 tandem repeats and which may bind oligosaccharide in a manner distinct from other classes of lectins. It is noteworthy that the same tandem repeat motifs have been found in calmegin, a Ca$^+$ binding protein specifically expressed during male meiotic germ cell development that is required for sperm fertility (Watanabe et al., 1994, Oshako et al., 1994 and Ikawa et al., 1997), suggesting that calmegin may also function as a lectin.

Several properties of calnexin and calreticulin have been described that have been proposed to influence their chaperone functions. They are calcium-binding proteins (Baksh and Michalak,
1991 and Tjoelker et al., 1994), contain at least one disulfide bond (Hebert et al., 1995, Ou et al., 1995 and Matsuoka et al., 1994), and appear to bind adenosine-containing nucleotides (Nigam et al., 1994 and Ou et al., 1995). These properties were examined to determine their effect on lectin function. The results indicate that calcium is essential for the lectin properties of calnexin since EGTA treatment abrogated oligosaccharide binding. Consistent with previous findings (Ou et al., 1995), EGTA treatment also increased calnexin's sensitivity to protease digestion and enhanced its assembly into oligomeric forms. Since calnexin is not a C-type lectin, the results are most easily interpreted in terms of calcium being required to maintain the overall conformation of calnexin, including its lectin site. Previous reports indicating the importance of calcium in the binding of calnexin to glycoproteins both in vitro (Le et al., 1994) and in vivo (Capps et al., 1994) are consistent with our findings, because it is well established that calnexin binding to most glycoproteins depends on a functional lectin site (Hammond et al., 1994 and Ou et al., 1993). We also determined that bound calcium is essential for oligosaccharide binding by calreticulin. It is difficult to relate this finding to glycoprotein binding since there are conflicting reports concerning the role of calcium in the binding of calreticulin to various proteins. Calcium did not appear to be required for calreticulin binding to various denatured proteins or glycoproteins (Nigam et al., 1994), yet it was essential for the binding of calreticulin to glycosylated laminin (McDonnell et al., 1996).

DTT treatment had no effect on oligosaccharide binding by calnexin, indicating that its disulfide bond(s) is not essential for its lectin properties. Interestingly, addition of DTT to intact microsomes under conditions where calnexin is reduced resulted in a loss of binding to in vitro translated/translocated influenza hemagglutinin (Hebert et al., 1995). The opposite effect was
observed when MDCK cells were treated with DTT. Interactions between calnexin and the secretory glycoprotein, gp80, were dramatically prolonged (Wada et al., 1994). Thus, although not required for calnexin's lectin function, the disulfide bond(s) in calnexin may be required for some other aspect of its function as a molecular chaperone. Possibilities include the structural maintenance of a polypeptide binding site (Ware et al., 1995, Carreno et al., 1995a, and Zhang et al., 1995) or a binding site for putative co-chaperones or regulatory factors. In contrast to calnexin, DTT treatment substantially inhibited calreticulin's ability to bind oligosaccharide. Calreticulin possesses a single disulfide bond between Cys\(^{109}\) and Cys\(^{144}\) (Matsuoka et al., 1994). Only the former residue is conserved in calnexin indicating that the two chaperones differ in the location of their disulfide bonds. Apparently, the disulfide bond in calreticulin plays a significant role in maintaining a functional lectin site. The loss of this bond may be the reason for the reduced oligosaccharide binding observed with deletion constructs lacking the N-terminal 138 residues of calreticulin (Figure 13B).

The ER luminal domain of calnexin has been reported to bind Mg-ATP with an accompanying increase in its protease sensitivity and the formation of oligomers (Ou et al., 1995). Similar experiments were unable to reproduce such conformational effects in the present study. Furthermore, the present findings indicate that the presence of Mg-ATP does not significantly affect calnexin's lectin function. While the addition of Mg-ATP slightly enhanced oligosaccharide binding, the effect appeared to be non-specific. This was suggested by the high concentration of nucleotide required to enhance oligosaccharide binding, the lack of a requirement for Mg\(^{2+}\), and the ability of a range of adenosine-containing nucleotides to produce the same effect. Calreticulin has also been suggested to bind Mg-ATP based on its
ATP-dependent dissociation from various denatured proteins in vitro (Nigam et al., 1994). However, as was the case for calnexin, only a slight enhancement in oligosaccharide binding was observed in the presence of Mg-ATP. These results do not rule out the possibility that ATP binding or hydrolysis may be involved in other aspects of calnexin or calreticulin function, such as polypeptide mediated binding or release of bound glycoproteins, much like other well-characterized chaperones (Hartl, 1996).

Given that the lectin sites of calnexin and calreticulin exhibit the same oligosaccharide binding specificities and affinities, are similarly affected by cofactors, and are related in primary sequence, it is interesting to speculate on the basis for their observed differences in binding to glycoprotein substrates. While there are some common substrates for calnexin and calreticulin (Table III, Chapter I), there are a growing number of examples where binding to some substrates is exclusive to one or the other chaperone. For example, the binding of human class I histocompatibility molecules to either calnexin or calreticulin appears to be dictated by the assembly state of these molecules. Calnexin binds exclusively to the free heavy chain subunit, whereas calreticulin appears to bind only following assembly of the heavy chain with the β2-microglobulin subunit (Sadasivan et al., 1996). Murine class I molecules differ in that calnexin appears to be the main chaperone associated with both free and β2-microglobulin associated heavy chains (Vassilakos et al., 1996 and Degen and Williams, 1991), although low levels of calreticulin binding to the β2-microglobulin associated form have recently been reported (van Leeuwen and Kearse, 1996c). The results presented here would suggest that the lectin specificities of calnexin and calreticulin are not sufficient to provide the observed differences in overall substrate specificity. One possible explanation could be that different
spatial orientations of oligosaccharide chains in different glycoproteins might favor interaction either with the membrane-constrained calnexin or the luminallly oriented calreticulin. The fact that human class I heavy chains possess a single N-linked oligosaccharide versus two or three oligosaccharides in the murine homologues could potentially account for their differences in chaperone interaction. Furthermore, the transfer of human class I molecules from calnexin to calreticulin at the time of assembly with \( \beta \)-microglobulin could be a consequence of a different orientation of its oligosaccharide in the two folding states (Helenius et al., 1997). A much simpler explanation is that calnexin and calreticulin differ in their recognition of polypeptide segments of unfolded glycoproteins. That calreticulin is capable of binding to polypeptide in addition to the Glc,Man,GlcNAc, oligosaccharide is well established. For example, calreticulin has been purified from cell extracts by affinity chromatography using the peptide, KLGFFKR, as ligand (Rojiani et al., 1991). This peptide corresponds to a highly conserved region in the cytoplasmic domain of the \( \alpha \) subunit of integrins. In addition, another report demonstrated that calreticulin binds to protein disulfide isomerase via its P domain, a region that also contains the lectin site (Baksh et al., 1995). However, in contrast to its lectin properties, this interaction was inhibited by \( \text{Ca}^{2+} \). Calnexin has also been implicated in binding to polypeptide segments because its interactions with glycoproteins are maintained following complete removal of oligosaccharide chains through endoglycosidase H digestion. Of the wide array of membrane and soluble glycoproteins tested in this manner (Ware et al., 1995, Zang et al., 1995, Arunachalam and Cresswell, 1995 and van Leeuwen and Kearse, 1996), only ribonuclease B has been shown to dissociate from calnexin upon deglycosylation (Rodan et al., 1996 and Zapun et al., 1996). Assessment of the possibility that calnexin and calreticulin differ in their recognition
of peptide motifs within unfolded or unassembled glycoproteins will be greatly facilitated by the reconstitution of their binding interactions with diverse glycoproteins \textit{in vitro}. 
CHAPTER IV
CALNEXIN ACTS AS A MOLECULAR CHAPERONE TO FACILITATE
THE FOLDING AND
ASSEMBLY OF CLASS I HISTOCOMPATIBILITY MOLECULES

Please Note: All data were produced by myself with the following exceptions:
Uncloned Droshila cell transfecants were supplied by Drs. M. Jackson and P. Peterson (R. W. Johnson, La Jolla, CA). Flow
cytometric analysis of prepared samples was performed by C. Smith, the technician in the flow cytometry facility, Department of
Immunology, University of Toronto.
In Figures 1-3, experimental design and initial experiments were performed by myself. To prepare data for publication M. F.
Cohen-Doyle and I repeated the experiments.
In Figure 4, M. F. Cohen-Doyle assisted in preparation of the gradients.

Most of the results in this chapter have been published in:
Alaniz P. Vassilakos, Myna F. Cohen-Doyle, Per A. Peterson, Michael R. Jackson, and David B. Williams (1996). The EMBO
Journal vol 15 no 7 pp. 1495-1506.
4.0 Abstract

Calnexin, a membrane protein of the endoplasmic reticulum, is generally thought to function as a molecular chaperone based on indirect or correlative evidence. To examine calnexin's functions more directly, I reconstituted the assembly of class I histocompatibility molecules in the absence or presence of calnexin in *Drosophila melanogaster* cells. Calnexin enhanced the assembly of class I heavy chains with β2-microglobulin as much as 5-fold. The improved assembly appeared largely due to more efficient folding of heavy chains as evidenced by increased reactivity with a conformation-sensitive monoclonal antibody and by a reduction in the level of aggregates. Similar findings were obtained in mouse or human cells when the interaction of calnexin with class I heavy chains was prevented by treatment with the oligosaccharide processing inhibitor castanospermine. The ability of calnexin to facilitate heavy chain folding and to prevent the formation of aggregates provides compelling evidence that calnexin functions as a bona fide molecular chaperone.
4.1 Introduction

Proteins that are translocated into the endoplasmic reticulum fold and undergo subunit assembly with the assistance of a diverse group of soluble folding enzymes and molecular chaperones. These include protein disulfide isomerase, peptidylprolyl isomerase, and molecular chaperones of the hsp70 (BiP) and hsp90 (GRP94) families (reviewed in Gething and Sambrook, 1992; Helenius et al., 1992). In addition to these general folding factors the ER contains a putative molecular chaperone known as calnexin that is largely selective for Asn-linked glycoproteins. Calnexin itself is not glycosylated and it resides in the ER as a Type I membrane protein that contains a 461 residue luminal domain homologous to calreticulin (Wada et al., 1991).

Evidence that calnexin functions as a molecular chaperone is indirect. It binds transiently to a diverse array of both membrane and soluble glycoproteins shortly following their translocation into the ER, and its dissociation can usually be correlated with some stage in polypeptide folding or subunit assembly. Furthermore, calnexin interacts in prolonged fashion with misfolded glycoproteins or with subunits that are unable to assemble (reviewed in Williams, 1995; Helenius, 1994).

Two properties set calnexin apart from known molecular chaperones, its membrane disposition and its marked preference for Asn-linked glycoproteins. The latter property originally became apparent when human hepatoma cells were treated with tunicamycin and the level of complexes between calnexin and newly synthesized secretory glycoproteins was dramatically reduced (Ou et al., 1993). Subsequent studies revealed that treatment of cells with the oligosaccharide processing inhibitors castanospermine and 1-deoxynojirymycin also
prcvcntcd calncxin binding (Hammond et al, 1994; Hammond and Helenius, 1994). These compounds inhibit glucosidases I and II, enzymes that remove glucose residues from the Glc3Man9GlcNAc2 oligosaccharide that is cotranslationally attached to nascent polypeptide chains as they enter the ER. Based on this and additional data it was suggested that glucose removal is required for calnexin binding and that calnexin specifically recognizes the Glc1Man9GlcNAc2 structure (Hammond et al, 1994). In Chapters II and III data was presented that directly demonstrates the lectin function of calnexin. In addition to lectin mediated interactions there is evidence for polypeptide mediated interactions between calnexin and its substrates. The evidence for this is discussed in detail in Chapters II and III.

Given its unusual binding specificity and its involvement in the biogenesis of most glycoproteins that traverse the ER, there is intense interest in establishing whether calnexin indeed functions as a molecular chaperone and, if so, how it utilizes its apparent dual mode of binding to effect its function. Recently, Hammond and Helenius (1994) employed castanospermine to prevent the binding of calnexin to vesicular stomatitis virus G glycoprotein and found that maturation of G glycoprotein to a fully disulfide-bonded form was substantially impaired. Although this result is suggestive of a chaperone function for calnexin, caution should be exercised in interpreting the effects of this inhibitor since it may have pleiotropic effects on other ER chaperones and folding factors. For example, it was recently demonstrated that calreticulin, which shares extensive sequence identity with the luminal domain of calnexin, binds transiently to immature forms of myeloperoxidase (Nauseef et al., 1995) as well as other glycoproteins (Hebert et al., 1995). This chaperone-like binding could be prevented by tunicamycin treatment and castanospermine. Furthermore, the results in chapter III demonstrate
that calreticulin, like calnexin, also interacts with GlcMan3GlcNAc2 oligosaccharides.

Dr. Williams' laboratory has developed an alternative approach to study calnexin's functions in which mouse class I heavy chains and β2m are expressed in Drosophila melanogaster cells either alone or co-expressed with calnexin. Since these cells lack the TAP transporter they are deficient in supplying peptides to the ER (Figure 1). However, they can produce functional class I heavy chain-β2m heterodimers at the cell surface that are capable of binding exogenously added antigenic peptide and eliciting T cell responses (Jackson et al., 1992). Furthermore, the cells are functionally calnexin deficient since no endogenous calnexin homolog can be detected in association with the heterologously expressed mouse heavy chains (Jackson et al., 1994). The utility of this system in reconstituting calnexin functions was recently illustrated by demonstrating a role for calnexin in the intracellular retention of class I assembly intermediates (free heavy chains or peptide-deficient heavy chain-β2m heterodimers) and in protecting free heavy chains from rapid ER degradation (Jackson et al., 1994). An obvious advantage of this system is that it does not rely on potentially non-specific chemical agents to produce a calnexin-deficient environment.

In this chapter the effects of calnexin on class I heavy chain folding and subunit assembly are determined using the Drosophila expression system. Furthermore, these data were compared with the effects observed following treatment of mouse and human cells with castanospermine. The results were remarkably similar and showed that calnexin facilitates both the folding of class I heavy chains and their assembly with β2m. The ability to demonstrate these functions by two independent methods leads to the conclusion that calnexin does indeed function as a
Drosophila cells transfected with DNA encoding class I molecules synthesize, assemble and transport heavy chain β₂m dimers to the cell surface. These cells lack the transporters responsible for transport of peptide to the ER for binding to heterodimers and as a result the surface expressed molecules are peptide deficient. Empty heterodimers are largely retained in the ER of mammalian cells suggesting that Drosophila cells lack appropriate ER quality control mechanisms. There is no evidence for interaction of class I molecules with endogenous Drosophila calnexin as assessed by chemical cross-linking, suggesting that calnexin may be the missing component. Experimental support for this came from experiments where calnexin was co-expressed and found to retain dimers in the ER and to prevent their degradation (Jackson et al., 1993).
molecular chaperone and is an important component in the folding and assembly pathway of class I molecules.

4.2 Materials and methods

4.2.0 Drosophila Cell Transfection and Cloning

_Drosophila melanogaster_ Schneider cells transfected with cDNAs encoding murine H-2 K\^b or D\^b class I heavy chains and β2-microglobulin, either alone or with canine calnexin, have been described previously (Jackson et al., 1994). The cDNA's for K\^b, D\^b, murine β2-m and canine calnexin were cloned downstream of the metallothionein promoter in plasmid pRMHS. These were transfected into Schneider cells along with the selection plasmid phsneo as described in Jackson et al., 1992. After 4 weeks in selection medium, G418 resistant cells were treated for 24 to 48 hours with CuSO\_4 and tested for expression of class I molecules by flow cytometry and immunosorbent assays. All _Drosophila_ cell lines used in this study were cloned to produce cell lines that express all of the heterologous proteins in each cell. To clone cells, 6x10^5 viable cells were plated on culture dishes in Schneider cell media containing 0.3% noble agar, 0.5 mg/ml geneticin, 10% heat treated (65°C, 30 min.) FCS and 1 X 10^6 γ-irradiated feeder cells. The feeder cells were produced by exposure of G418 resistant Schneider cells (1x10^5 cells/4 ml Dulbeccos PBS) to 24,000 RADS of γ radiation. The plated cells were grown for 3 weeks, individual colonies were isolated, grown in suspension and screened for expression of heterologous proteins by radiolabeling and immunoprecipitation after CuSO\_4 treatment. Clones expressing comparable levels of heavy chain and β2m were used to assess the effects of calnexin co-
expression. Control plates containing only feeder cells contained less than 1 colony/plate.

4.2.1 Cell Culture and Antibodies

Cells were maintained at room temperature in Schneider's insect medium (Sigma) supplemented with 10% fetal bovine serum, antibiotics, and 0.5 mg/ml Geneticin (Life Technologies Inc.). Prior to radiolabeling, cells were incubated in medium containing 1 mM CuSO₄ for 16-24 h to induce expression of the cDNAs under control of the metallothionein promoter. Murine EL4 cells (H-2b lymphoma; Gorin, 1950) and R1E-D₃ cells (P₂m-deficient thymoma transfected with the D₃ heavy chain gene; Allen et al., 1986) were grown in RPMI 1640 medium supplemented with 10% fetal bovine serum and antibiotics in a humidified incubator at 37°C under a 5% CO₂ atmosphere. Human C1R cells (a B lymphoblastoid line with no detectable HLA-A or -B expression) transfected with the HLA-B27 gene were obtained from Dr. P. Cresswell, Yale University, and were maintained as described for murine cells with the addition of 0.5 mg/ml Geneticin (Alexander et al., 1989).

To assess the state of folding or assembly of class I heavy chains, the following conformation-dependent monoclonal antibodies (mAbs) were used: mAbs 20-8-4S (Ozato and Sachs, 1981), Y3 (Jones and Janeway, 1981), and B8-24-3 (Köhler et al., 1981) for mouse H-2 K₃ heavy chains associated with β₂m, mAbs B22-249 R1 (Lemke et al., 1979) and 27-11-15S (Ozato and Sachs, 1981) for mouse H-2 D₃ heavy chains associated with β₂m, mAb W5/32 for human HLA-A, B, or C heavy chains associated with β₂m (Barrattable et al., 1978), mAb 28-14-8S (Ozato et al., 1980) which recognizes a conformational epitope present in the α₃ domain of either free or β₂m-associated D₃ heavy chains, and mAb 5H7 (Smith et al., 1994) which
recognizes a folded determinant in the α3 domain of free or β2m-associated HLA-A, B, or C heavy chains (a gift from Dr. J. Bluestone). A rabbit polyclonal antibody (anti-8) raised against a C-terminal peptide of the K^b heavy chain was provided by Dr. B. Barber, University of Toronto, and was used to isolate all conformational states of K^b (Smith et al., 1986). For the selective isolation of unassembled mouse class I heavy chains, a rabbit polyclonal antibody raised against denatured mouse heavy chains (anti-HC) was provided by Dr. H. Ploegh and R. Machold (Machold et al., 1995). MAb HC-10, also obtained from Dr. H. Ploegh, was raised against denatured human HLA-B locus products and was used to isolate unassembled HLA-B27 heavy chains (Stam et al., 1986). Most mAbs were purified from hybridoma culture supernatants by protein A-agarose affinity chromatography using the MAPS system (Bio-Rad laboratories). For the isolation of calnexin and associated proteins, a rabbit antiserum raised against the C-terminal 14 amino acids of canine calnexin was used (Jackson et al., 1994). For the isolation of β2m a polyclonal antibody raised against mouse β2m was used and was provided by R. Machold and H. Ploegh.

4.2.2 Pulse-Chase Radiolabeling and Immune Isolation

Transfected Drosophila cells expressing mouse class I molecules with or without calnexin were incubated at room temperature for 20 min in Met-free Graces medium prior to radiolabeling. Cells (1-3 x 10^7) were resuspended at 5 x 10^7 cells/ml in the same medium containing 0.5 mCi/ml [35S]Met (>1000 Ci/mmol, Amersham) and were radiolabeled for 5 or 10 min at room temperature (Pulse). Nine volumes of Insect-XPRESS medium (BioWhittaker) supplemented with 1 mM unlabeled Met were added and cells were incubated for various periods (Chase). In
some experiments, the pre-incubation time was increased to 40 min and the pre-incubation, pulse, and chase media included 0.1 mM of the protease inhibitor LLnL (Sigma). After chase incubations, cells (2.5 x 10^6) were washed twice with cold phosphate-buffered saline (PBS), pH 7.4 and lysed on ice for 30 min in 1 ml of one of the following buffers. Digitonin and NP-40 lysis buffers contained either 1% digitonin (Sigma) or 1% Nonidet P-40 (BDH), respectively, in PBS, pH 7.4, supplemented with 10 mM iodoacetamide, 1% aprotinin, 0.25 mM 4-(2-aminoethyl)benzene-sulfonyl fluoride, and 10 μg/ml each of chymostatin, leupeptin, antipain, and pepstatin. RIPA lysis buffer contained 25 mM Hapes, pH 7.4, 150 mM NaCl, 1% NP-40, 1% deoxycholate, 0.1% SDS, 20 mM N-ethylmaleimide and protease inhibitors as above. All subsequent steps were performed at 4°C. Lysates were centrifuged at 16,000 x g for 10 min and supernatant fractions were incubated with anti-class I heavy chain or anti-calnexin antibody for 2 h. Immune complexes were recovered by shaking for 1 h with 20-40 μl protein A-agarose after which bead-bound complexes were washed 3 times with 1 ml of either 0.2% digitonin/PBS, pH 7.4, or NTSE (0.5% NP-40, 10 mM Tris, pH 7.4, 150 mM NaCl, 1 mM EDTA, and 0.02% sodium azide), or RIPA, depending on the detergent used for cell lysis. Most immune isolations recovered in excess of 90% of the respective class I heavy chain species in a single clearance.

When mAb B22-249.R1 was used, a second clearance was incorporated to isolate all of the β2m-associated D^β heavy chains. In several instances, immune-isolated proteins were digested with endoglycosidase H prior to analysis by SDS-PAGE (Jackson et al., 1994). Unless otherwise stated, immune complexes were heated in Laemmli sample buffer (with or without dithiothreitol as indicated) for 10 min at 70°C and proteins were separated by SDS-PAGE using 10% gels (Laemmli, 1970). Gels were fixed, equilibrated with 1 M sodium salicylate, and dried
prior to fluorography. A Hoefer GS 300 Transmittance/Reflectance Scanning Densitometer was used to quantitate data from fluorograms and data was processed using MacIntegrator I software (Rainin).

Murine EL4 and R1E.Db cells were radiolabeled, lysed and subjected to immune isolation as described for Drosophila cells with the following exceptions. Pre-incubation was conducted for 60 min in Met-free RPMI 1640 supplemented with 9 mM Hepes, 2 mM glutamine, with or without 1 mM castanospermine at 37°C. Cells were radiolabeled in the same medium at 37°C for 5 min in the presence or absence of castanospermine after which the cells were incubated for various chase times in medium supplemented with 1 mM unlabeled Met +/- castanospermine. Human CIR cells expressing HLA-B27 were treated in a similar manner with the exception that detergent lysates of radiolabeled CIR cells were incubated with 10% formalin treated Staphylococcus aureus cell suspension (Sigma) prior to immune isolation. This pre-clearance step was necessary to remove radiolabeled immunoglobulin synthesized endogenously by CIR cells.

4.2.3 Velocity Density Gradient Centrifugation

Drosophila cells expressing H-2 Kb and β2m with or without calnexin were radiolabeled with [35S]Met for 5 min and then incubated with excess unlabeled Met for 80 min. Cells (1 x 10^7) were lysed in 1 ml RIPA buffer and, following centrifugation to remove insoluble material, 0.25 ml of lysate was loaded onto a 12 ml, 10-40% (w/v) linear glycerol gradient prepared in RIPA buffer. The gradients were centrifuged at 4°C for 15 h at 35,000 rpm using a Beckman SW41 rotor (Pind et al., 1994). Fractions (0.75 ml) were collected manually from the top of the
and analyzed by SDS-PAGE. To detect any insoluble material sedimenting to the bottom of the gradient, 100 µl of 0.1% SDS was added to the centrifuge tube following removal of the final fraction. The sample was heated at 55°C for 10 min, diluted with 1.2 ml of NP-40 lysis buffer, and subjected to immune isolation with anti-8 antiserum. As a measure of the total amount of Kb molecules loaded on the gradient, an additional 0.25 ml sample of lysate was treated directly with anti-8 and analyzed by SDS-PAGE along with the gradient fractions. Gradients were calibrated with BioRad molecular mass standards consisting of thyroglobulin (670 kDa), IgG (158 kDa), ovalbumin (44 kDa), and myoglobin (17 kDa).

4.2.4 Flow Cytometric Analysis

EL4 cells were cultured under standard conditions for 24 h either alone or with the addition of castanospermine (0.27 mM) at 24 h and at 16 h prior to processing for flow cytometry. All subsequent steps were conducted at 4°C. Cells (1 x 10^6) were pelleted at 300 x g for 5 min and resuspended in PBS-BSA (Dulbecco's PBS containing 0.1% Fraction V BSA and 0.1% NaN3). Cells were re-pelleted, resuspended in 0.1 ml PBS-BSA, and incubated for 1 h with 1-2 µg of either mAb B22-249.R1 or Y3. PBS-BSA (2 ml) was added, cells were pelleted, washed once with PBS-BSA, and then incubated for 45 min in 0.1 ml PBS-BSA containing 1/50 dilution of FITC-conjugated goat anti-mouse IgG (Sigma). The cells were then washed in Dulbecco's PBS and resuspended at 2-5 x 10^6 cells/ml in Dulbecco's PBS containing 0.1% NaN3. Control incubations omitting either the first antibody or both antibodies were included for each experiment. Fluorescence data was collected and analyzed on a Coulter Epics Elite flow.
cytometer. FITC fluorescence profiles and mean FITC fluorescence values were generated for each sample.

4.3 Results

4.3.0 Calnexin facilitates assembly of class I molecules expressed in Drosophila cells

Initially, the Drosophila expression system was used to assess calnexin’s ability to facilitate the assembly of two different mouse class I heavy chains with β2m. Drosophila cells expressing either Kβ or Dβ heavy chains and β2m in the presence or absence of co-expressed calnexin were subjected to a pulse-chase analysis to determine the kinetics and efficiency of subunit assembly. The results obtained for Kβ are depicted in Figure 2A. Efficiency of heterodimer formation was assessed by comparing heavy chain reactivity to a β2m-dependent mAb and to an antibody that recognizes total Kβ heavy chains. In addition, reactivity to an antibody raised against denatured heavy chains was used to measure unassembled Kβ. Results from five separate experiments were quantified by densitometry and both β2m-associated and unassembled heavy chains were expressed as a percentage of total heavy chains present in the pulse sample (Figure 2B). In the absence of co-expressed calnexin only ~20% of Kβ heavy chains assembled with β2m (Figure 2, left panels). The small population of Kβ molecules that assembled was gradually transported out of the ER as assessed by the maturation of Asn-linked oligosaccharides to a form resistant to digestion with endoglycosidase H (t1/2 ~ ~80 min). By 320 min of chase, all assembled Kβ heterodimers were resistant to endo H digestion (data not shown). Unassembled Kβ heavy chains remained at a high level throughout the chase and were exported very slowly from the ER.
Figure 2. Effects of calnexin on K\(^b\) assembly with β2m in Drosophila cells.

(A) Drosophila cells expressing K\(^b\) and β2m in the absence (left panels) or presence (right panels) of mammalian calnexin were radiolabeled with \(^{35}\)S-Met for 5 min and subsequently incubated in media containing excess unlabeled Met for the indicated times. Cells were lysed in digitonin lysis buffer and K\(^b\) molecules were isolated with antibodies recognizing total (anti-8 antiserum), β2m-associated (mAb 20-8-4S), or unassembled (anti-HC antiserum) heavy chains. Following digestion with endoglycosidase H (endo H), proteins were analyzed by reducing SDS-PAGE. Only the heavy chain regions of the gels are shown. The mobilities of endo H-sensitive (endo HS) and endo H-resistant (endo HR) K\(^b\) heavy chains are indicated on the top panel only. The letter P signifies pulse-labeled samples.

(B) The results from five independent experiments were quantified by densitometry and the amounts of β2m-associated and unassembled heavy chains were expressed as a percentage of the total heavy chains present in the pulse samples. Data points represent mean values and error bars reflect the range of the replicate samples.
No further assembly with $\beta_2m$ was observed beyond 80 min of chase (data not shown). In marked contrast, co-expression of calnexin resulted in virtually quantitative assembly of $\beta_2m$ with $\beta_2m$ (Figure 2, right panels). As a consequence of increased assembly more heterodimers were transported out of the ER but, consistent with the findings of an earlier study (Jackson et al., 1994), these peptide-deficient molecules that had been associated with calnexin were transported more slowly than heterodimers assembled in the absence of calnexin (11/2 endo H resistance = ~120 min).

The preceding experiment was performed at 24°C, a temperature at which Drosophila cells remained viable and peptide-deficient class I molecules were stable. To ensure that the observed difference in assembly efficiency was not related to exposure of class I subunits to temperatures lower than the 37°C normally present in mammalian cells, additional experiments involving chase incubations at 30°C and 37°C were performed. Consistent with the findings at room temperature, calnexin co-expression resulted in a roughly 5-fold stimulation of assembly. However, since heavy chain-$\beta_2m$ heterodimers acquire little peptide ligand in Drosophila cells (Jackson et al., 1992), and peptide-deficient class I molecules are thermolabile (Ljunggren et al., 1990), the heterodimers dissociated rapidly upon transport out of the ER at 37°C (Figure 3, top panel). In this panel there is no transported material and there is a loss of assembled material relative to chase at lower temperatures (compare assembled to total heavy chain). Since the cDNAs transfected into Drosophila cells were under the control of the metal-inducible metallothionein promoter it was possible to assess whether there were calnexin-related assembly differences at varying levels of protein expression. Experiments performed after different periods of induction with CuSO4 (corresponding to a 4-fold difference in protein expression as
assessed by Western blot analysis, not shown) gave virtually identical results to those in Figure 2, suggesting that expression levels do not significantly alter the requirement for calnexin in class I heterodimer assembly (Figure 3, compare panels 3 and 4). In addition, the amount of β2m synthesized in cells with or without calnexin was comparable when examined as a function of how much heavy chain was produced, suggesting that β2m was not a limiting factor (Figure 4, compare levels of heavy chain to β2m in the presence, left panel, or absence, right panel, of calnexin). The combined results indicate that although Kb heavy chains can assemble with β2m in the absence of calnexin, calnexin clearly enhances the efficiency of assembly.

Db heavy chains share 81% sequence identity with Kb heavy chains, yet an analysis of Db assembly with β2m in Drosophila cells yielded somewhat different results (Figure 5A). Again, efficiency of heterodimer formation was assessed by comparing heavy chain reactivity to β2m-dependent mAbs and to antibodies that recognize total heavy chains. In the absence or presence of calnexin there appeared to be quantitative assembly of Db with β2m (Figure 5A, compare total versus 20 min chase lanes). However, in the absence of calnexin there was considerable degradation of Db heavy chains that was apparent even in the pulse sample. Typically, 20-40% of total heavy chains labeled in a 5 min pulse were present as a series of degradation products of faster mobility than the full length Db heavy chain (Figure 5A, left panel, total). At this time point little assembly with β2m had occurred (Figure 5A, left panel, compare first two lanes). Degradation continued during chase incubations and included Db-β2m heterodimers. As observed previously (Jackson et al., 1994), co-expression of calnexin dramatically inhibited degradation (Figure 5A, right panel). The observation that a substantial
Figure 3. Effects of Temperature and Length of Induction on K⁺ Assembly in the Presence or Absence of Calnexin in Drosophila Cells

Experiments presented here are essentially the same as in Figure 2 with the following changes. The temperature of the chase incubation was varied from 24° - 37°C as indicated in the right column and the time of induction with CuSO₄ was varied, either 24 or 48 hours, as indicated in the left hand column. Cell lysates and immunoprecipitation reactions were as in Figure 2. In the absence of endo H digestion, mature heavy chains transported to the Golgi apparatus migrate more rapidly than their immature precursors due to the processing of oligosaccharides to smaller, complex forms in Drosophila cells.
Figure 4. β2m expression compared to heavy chain expression in the presence or absence of calnexin

Drosophila cells expressing Kβ and β2m with or without calnexin were radiolabeled with [35S]-methionine for 5 minutes, lysed with digitonin lysis buffer and immunoprecipitated as described in the materials and methods. Kβ molecules were immunoprecipitated with anti-8 antibody and β2m molecules with an rabbit anti-mouse β2m polyclonal antibody. The relative amount of β2m expressed when compared to heavy chain expression was not significantly different in the presence or absence of calnexin expression. The exposures in this figure were adjusted to allow comparison of β2m signals at equivalent heavy chain signals.
Figure 5. Effects of calnexin on \( \beta_2 \text{m} \) assembly with \( \beta_2 \text{m} \) in \( Drosophila \) cells.

(A) \( Drosophila \) cells expressing \( \Delta \beta \) and \( \beta_2 \text{m} \) in the absence or presence of calnexin were radiolabeled with \( [35S] \text{Met} \) for 5 min., subsequently incubated in media containing unlabeled Met for the indicated times and cells were lysed in digitonin lysis buffer. \( \Delta \beta \) molecules were isolated with antibodies recognizing total (mAb 28-14-8S and anti-HC) or \( \beta_2 \text{m} \)-associated (mAbs \( B22-249.1 \) and \( b \text{m}-13 \)) molecules. Proteins were analyzed by reducing SDS-PAGE. The mobilities of calnexin, immature and mature \( \Delta \beta \), and degraded \( \Delta \beta \) molecules are indicated.

(B) The experiment performed in panel A was repeated except that cells were pre-incubated for 40 min, radiolabeled, and chased all in the presence of the protease inhibitor, LLAzL.

(C) The results in panel B were quantified by densitometry: The amount of \( \beta_2 \text{m} \)-associated \( \Delta \beta \) molecules at each time was expressed as a percentage of total heavy chains in the pulse sample.
proportion of heavy chains was degraded very early after synthesis, before assembly, raised the possibility that a decreased efficiency of assembly in the absence of calnexin could be masked by the rapid loss of unassembled molecules.

In an attempt to minimize the degradation of Dβ molecules, Drosophila cells were incubated with the membrane permeable protease inhibitor, LLnL. Treatment with LLnL slowed the degradation of Dβ expressed in the absence of calnexin but did not block it entirely. A loss of total heavy chains was still observed during the chase (Figure 5B, left panel). However, under these conditions, it was now apparent that only ~60-70% of the total Dβ heavy chains present in the pulse sample assembled with β2m in the absence of calnexin (Figure 5B and 5C). The value of 60-70% assembly is most likely an overestimate given that degradation was not completely blocked by LLnL treatment. This residual degradation appeared mainly restricted to heavy chains that did not assemble with β2m since at the 10 and 20 min time points all of the remaining heavy chains were β2m-associated and relatively stable (Figure 5B, compare total and β2m-associated lanes). The inefficient assembly was not due to degradation simply limiting availability of heavy chains but rather, as detailed in the next section, the inability of a population of heavy chains to fold correctly (see below). In cells expressing calnexin, assembly was more efficient with all of the Dβ heavy chains forming heterodimers (Figure 5B and 5C). As was observed for Kβ (Figure 2), the effect of calnexin on Dβ assembly was manifested rapidly; as early as the 5 min pulse sample calnexin enhanced assembly by about 3-fold. Consistent with previous observations (Jackson et al., 1994), transport of peptide-deficient Dββ2m heterodimers to the Golgi apparatus (appearance of mature form) was slower in calnexin-expressing cells due to their retention by the chaperone (Figure 5A and 5B).
These findings indicate that calnexin facilitates the assembly of both \( \text{D}^b \) and \( \text{K}^b \) heavy chains with \( \beta_2m \) when expressed in \textit{Drosophila} cells. A major difference between these two class I isotypes is the fate of unassembled heavy chains in the absence of calnexin. \( \text{D}^b \) heavy chains that do not assemble are rapidly degraded whereas unassembled \( \text{K}^b \) heavy chains are stable and are largely retained intracellularly in a conformation that is not competent for binding \( \beta_2m \).

### 4.3.1 Calnexin facilitates folding of \( \text{D}^b \) heavy chains expressed in \textit{Drosophila} cells

One possible explanation for the more efficient assembly of class I molecules in the presence of calnexin is that calnexin promotes folding of heavy chains to a conformation that is competent for interaction with \( \beta_2m \). For \( \text{D}^b \) molecules, a single mAb (28-14-8S) exists that recognizes a folded epitope in the \( \alpha_3 \) domain of the heavy chain independent of whether the heavy chain is free or \( \beta_2m \)-associated. Acquisition of this epitope was used as an indication of \( \text{D}^b \) heavy chain folding.

\textit{Drosophila} cells expressing only \( \text{D}^b \) heavy chains in the absence or presence of calnexin were subjected to pulse-chase radiolabeling followed by immunoprecipitation of total or 28-14-8S reactive heavy chains (Figure 6A and 6B). As observed in the previous experiment, significant degradation of total heavy chains occurred during the chase period in the absence of calnexin although the characteristic degradation intermediates were not detected in these cells that lacked \( \beta_2m \) (Figure 6A, left panel). However, since acquisition of the 28-14-8S epitope was very rapid, it was possible to observe the effects of calnexin on this process even in the presence of degradation. Whereas 40% of the total \( \text{D}^b \) heavy chains present in the pulse sample folded into a
Figure 6. Calnexin enhances folding of D\text{\textsuperscript{b}} heavy chains expressed in \textit{Drosophila} cells.

(A) \textit{Drosophila} cells expressing D\text{\textsuperscript{b}} heavy chain in the absence or presence of calnexin were subjected to pulse-chase radiolabeling as described in the legend to Figure 2A. Digitonin lysates were incubated with antibodies recognizing total D\text{\textsuperscript{b}} heavy chains (mAb 28-14-8S and anti-HC combined) or heavy chains with a folded \(\alpha3\) domain (mAb 28-14-8S) and isolated proteins were analyzed by reducing SDS-PAGE. The mobilities of the D\text{\textsuperscript{b}} heavy chain and co-isolated calnexin are indicated.

(B) The results in panel A were quantified by densitometry and the amount of 28-14-8S reactive heavy chains at each time point was expressed as a percentage of the total heavy chains present in the pulse sample.
mAb-reactive conformation in the absence of calnexin, 90% of heavy chains acquired the epitope when co-expressed with calnexin (Figure 6B). Again, calnexin exerted its effects at an early stage, enhancing the efficiency of heavy chain folding from 24% to 65% during the 5 min pulse radiolabeling. Similar to the situation observed for D\textsuperscript{b} assembly (Figure 5B), heavy chains that did not fold into a mAb-reactive conformation in the absence of calnexin were selectively degraded. By 20 min of chase only heavy chains possessing the 28-14-8S epitope remained (Figure 6A, left panel, compare total and 28-14-8S lanes). It is apparent that degradation did not limit the opportunity for heavy chains to fold. As evidenced by the pulse and 10 min time points in the presence of calnexin, the 28-14-8S epitope was capable of forming rapidly and nearly quantitatively. However, after 10 min in the absence of calnexin, only 40% of heavy chains had acquired the epitope even though unfolded heavy chains were still available (Figure 6A, left panel, compare 28-14-8S and total lanes). Clearly, most heavy chains synthesized in the absence of calnexin were simply incapable of correct folding. This conclusion was confirmed when cells lacking calnexin were treated with LDL to retard degradation and no increase in heavy chain folding was observed (data not shown). Thus, interaction with calnexin increases the efficiency of heavy chain folding as monitored by a mAb-defined conformational epitope.

4.3.2 Calnexin prevents aggregation of K\textsuperscript{b} heavy chains expressed in Drosophila cells

We next examined if the dramatic effect of calnexin on assembly of K\textsuperscript{b} heavy chains with P\textsubscript{2m} (Figure 7) could also be attributed in part to differences in the folding state of K\textsuperscript{b} heavy chains. Unfortunately, it was not possible to monitor acquisition of a conformational epitope as was done for D\textsuperscript{b}, because no mAbs specific for K\textsuperscript{b} exist that recognize conformational epitopes
in the absence of assembly with β2m. Instead, we took advantage of the fact that, unlike Δβ
heavy chains, Kβ heavy chains not competent for assembly with β2m were relatively stable in
the absence of co-expressed calnexin (Figure 2A, left panel). This permitted an assessment of
the aggregation state of Kβ.

Drosophila cells expressing Kβ and β2m with or without calnexin were pulse radiolabeled
and chased for 80 min, at which time assembly was complete and any remaining free heavy
chains were not competent for further assembly (Figure 2). To determine if unassembled heavy
chains existed in an aggregated state that might render them incompetent for subunit association,
isolated heavy chains were subjected to non-reducing SDS-PAGE. Using an antiserum that
recognizes all folding states of Kβ, a high molecular weight aggregate was detected that barely
entered a 7.5% resolving gel (Figure 7A, total, -DTT). The aggregate was stable to heating in
SDS at 55°C but could be partially dissociated upon heating at 70°C. Complete dissociation was
effected only by heating at 70°C under reducing conditions (Figure 7A, +DTT). The presence of
Kβ heavy chains in the aggregate was confirmed by excising the aggregate region of the gel,
eluting proteins by heating in DTT, and re-analyzing by reducing SDS-PAGE (data not shown).
Thus, the aggregate appeared to contain heavy chains associated through both non-covalent and
cystine interactions. The relative amount of aggregate was determined densitometrically as a
percentage of the total heavy chain recovered in both free and aggregated forms (Figure 7A, see
% aggregates). Importantly, there was a 3-fold increase in the percentage of aggregates in cells
lacking calnexin. As much as 43% of Kβ heavy chains could be recovered in the form of
aggregates in the absence of calnexin. That the aggregates contained heavy chains that were
incompetent for assembly with β2m was confirmed by comparing total
Figure 7. Calnexin inhibits aggregation of K\(^b\) heavy chains in Drosophila cells.

(A) Drosophila cells expressing K\(^b\) and β\(_2\)m with or without calnexin were radiolabeled with [\(^35\)S]Met for 10 min, subsequently chased for 80 min, and then lysed in RIPA lysis buffer. RIPA buffer proved to be more effective in solubilizing aggregated heavy chains than digitonin lysis buffer. K\(^b\) molecules were isolated with antibodies recognizing either total heavy chains (anti-8 antiserum and 20-8.4S mAb combined) or K\(^b\)-β\(_2\)m heterodimers (mAb 70.8-4S). Immune complexes were incubated in SDS-PAGE sample buffer with or without DTT at 55°C or 70°C for 15 minutes and proteins were analyzed by non-reducing SDS-PAGE (7.5% gel). The mobilities of the K\(^b\) heavy chain and heavy chain aggregates are indicated. The percentage of total heavy chains present as aggregates was determined densitometrically and is included at the bottom of the gel lanes.

(B) Drosophila cells were radiolabeled and lysed as in panel A. Cell lysates were applied to the top of a 10-40% glycerol gradient and the gradients were centrifuged for 15 hours at 34,000 rpm in an SW-41 rotor. Gradients were fractionated, total K\(^b\) heavy chains were immunoprecipitated from each fraction with anti-8 antiserum, and isolated proteins were analyzed by reducing SDS-PAGE. P fractions represent proteins that pelleted to the bottom of the gradient. T samples represent the total pool of K\(^b\) heavy chains obtained by immune precipitation of an equivalent amount of lysate to that loaded on the gradient.

(C) The data from panel B was quantified by densitometry and the relative amount of heavy chain recovered in each fraction was expressed as percentage of the total heavy chain applied to the gradient.
heavy chains with assembled Kb-β2m heterodimers recovered selectively with mAb 20-8-4S. Heterodimers recovered with this mAb were found exclusively in a non-aggregated state (Figure 7A, β2m-assoc.).

To confirm that calnexin influences the extent of aggregation of Kb heavy chains, cell lysates (radiolabeled under the same conditions as above) were fractionated on 10-40% glycerol gradients prior to immunosolation of total Kb molecules and analysis by reducing SDS-PAGE (Figure 7B and 7C). In cells expressing calnexin, 97% of the Kb molecules that could be recovered from the gradient were present in fractions 1-3. Only 3% of Kb molecules were recovered in denser fractions (fractions 4-16). This was the same distribution as that observed for Kb-β2m heterodimers when mAb 20-8-4S was used for the isolation (data not shown). This reflects the fact that, in cells expressing calnexin, Kb heavy chains were present almost exclusively as heterodimers by 80 min of chase (Figure 2). In the absence of calnexin, 78% of heavy chains that could be recovered from the gradient were present in fractions 1-3 and there was a substantial proportion (22%) that was found in the denser (4-16) fractions (Figure 7B and 7C).

Thus, by two independent techniques, it could be demonstrated that in the absence of calnexin interactions there was a substantial increase in the degree of Kb heavy chain aggregation that precluded assembly with β2m. The extent of aggregation assessed by these techniques (22-43%) was somewhat less than that expected based on the observed low efficiency of subunit assembly in cells lacking calnexin (Figure 2). It is possible that during cell lysis and immune isolation some non-covalent interactions were disrupted leading to an underestimation of the amount of aggregates. Alternatively, in addition to aggregates, there may have been
significant amounts of misfolded heavy chain monomers that were incompetent for assembly

Overall, the findings suggest that one function of calnexin is to bind to free K\(^{b}\) heavy chains and prevent inappropriate molecular interactions that lead to aggregation.

### 4.3.3 Castanospermine prevents calnexin interaction with class I molecules in mouse cells

The *Drosophila* expression system provides an excellent means to study the effects of calnexin on class I biogenesis in the absence of calnexin-inhibiting drugs that could potentially complicate the interpretation of results. Although the preceding *Drosophila* experiments clearly implicate calnexin as a molecular chaperone that facilitates class I heavy chain folding and subsequent assembly, we wished to determine if calnexin functions in a similar manner in mouse cells where the details of class I-calnexin interactions were originally established (Degcn and Williams, 1991; Degen et al., 1992). Castanospermine is an oligosaccharide processing inhibitor that prevents the formation of the Glc1Man9GlcNAc2 oligosaccharide required for initial binding of calnexin to most glycoproteins (Hammond et al., 1994; Ware et al. 1995).

Consequently, this inhibitor provided a means to study calnexin functions during class I biogenesis in mouse cells. Although the potential existed for castanospermine to affect processes other than those mediated by calnexin, the validity of the data could be assessed in the context of the results obtained using the *Drosophila* system.

We first tested the ability of castanospermine to prevent the interaction of calnexin with class I heavy chains. Mouse EL4 cells were radiolabeled with [\(^{35}\)S]Met in the presence or absence of castanospermine and then calnexin or class I heavy chains were immunoisolated from cell lysates (Figure 8A). In the absence of castanospermine treatment, anti-calnexin
Figure 8. Effect of castanospermine on heavy chain-β2m assembly in mouse cells.

(A) Murine EL4 cells that express Kβ, Dβ, and β2m were pre-incubated for 1 h in Met-free medium in the absence or presence of 1 mM castanospermine. Cells were subsequently radiolabeled in the presence of castanospermine with [35S]Met for 20 min and then were lysed in digitonin lysis buffer. Calnexin-associated molecules were isolated with anti-calnexin antiserum (lane 1). Class I molecules were isolated with antibodies recognizing either Kβ alone (anti-β antiserum, lane 2), Dβ alone (mAb 28-14-8S, lane 3), or both Kβ and Dβ (anti-HC antiserum, lane 4). In lane 6, calnexin immunoprecipitates were disrupted by heating to 90°C in 0.2% SDS-PBS and, following addition of a 10-fold excess of NP-40, reprecipitated with a combination of anti-β, anti-HC, and 28-14-8S antibodies (lane 6). Lane 5 is blank. The mobilities of calnexin, Kβ, and Dβ molecules are indicated.

(B) EL4 cells were pre-incubated for 1 h, radiolabeled with [35S]Met for 5 min, and then chased for the indicated times in the absence or presence of castanospermine. Cells were lysed in digitonin lysis buffer and Kβ molecules were isolated with antibodies recognizing either total (anti-β) or β2m-associated (mAbs Y3 and B3-24-3 combined) heavy chains (top panel). Dβ molecules were isolated from replicate samples using antibodies recognizing either total (mAb 28-14-8S and anti-HC combined) or β2m-associated (mAb B22-249-R1) heavy chains (bottom panel). Proteins were digested with endo H and analyzed by reducing SDS-PAGE. The mobilities of the endo H-sensitive and -resistant heavy chains are indicated.

(C) The results in panel B were quantified by densitometry and the amount of β2m-associated heavy chains at each time point was expressed as a percentage of the total heavy chains present in the pulse sample.
antibodies recovered calnexin and a series of associated proteins, two of which had the same electrophoretic mobilities as K\textsuperscript{b} and D\textsuperscript{b} heavy chains (Figure 8A, left panel, compare lane 1 with lanes 2-4). Disruption of anti-calnexin immune complexes with SDS and re-isolation using antibodies specific for class I heavy chains clearly showed that both K\textsuperscript{b} and D\textsuperscript{b} molecules were present in complexes with calnexin (Figure 8A, left panel, lane 6). Consistent with previous reports (Hammond and Helenius, 1994), castanospermine treatment resulted in a dramatic reduction in the amount of proteins co-immunoprecipitating with calnexin and, in particular, the two bands corresponding to class I heavy chains were absent (Figure 8A, right panel, lane 1). Disruption of anti-calnexin immune complexes and re-isolation with anti-heavy chain Ab confirmed the absence of calnexin-class I complexes (Figure 8A, right panel, lane 6). Synthesis of class I heavy chains was normal in castanospermine-treated cells as evidenced by direct immune isolation with anti-class I Ab (Figure 8A, compare lanes 2-4 in both panels).

4.3.4 Castanospermine inhibits assembly of class I molecules

Initially, the effects of castanospermine treatment on the assembly of class I heavy chains with β2m was examined. EL4 cells were subjected to a pulse-chase analysis in the absence or presence of castanospermine and the efficiency of assembly was determined by immunoprecipitation of both total and β2m-associated heavy chains (Figure 8B). Castanospermine treatment resulted in a dramatic decrease in the assembly of both K\textsuperscript{b} and D\textsuperscript{b} heavy chains with β2m as evidenced by the low level of β2m-associated heavy chains relative to untreated cells. Densitometric analysis of the fluorograms revealed that whereas both K\textsuperscript{b} and D\textsuperscript{b} molecules assembled almost quantitatively (100% for K\textsuperscript{b} and 80% for D\textsuperscript{b}) during 20 min of chase in the absence of
castanospermine, there was a 4-5 fold decrease in assembly (25% for K\textsuperscript{b} and 15% for D\textsuperscript{b}) when calnexin interactions were inhibited with castanospermine (Figure 8C). To assess the generality of this effect, similar experiments were performed using murine MDAY-D2 cells that express class I molecules of d-haplotype. Consistent with the results obtained with EL4 cells, castanospermine treatment resulted in a 2-3 fold decrease in the assembly of D\textsuperscript{d} heavy chains with \( \beta_{2m} \) (Figure 9).

As shown in Figure 8B, it was also apparent that the small population of K\textsuperscript{b} and D\textsuperscript{b} heavy chains that assembled in castanospermine-treated cells was transported more rapidly to the Golgi apparatus relative to assembled molecules in untreated cells. The apparent half-times for acquisition of endo H resistance in castanospermine-treated cells were roughly 20 min for K\textsuperscript{b} and 30 min for D\textsuperscript{b} compared to about 30 min for K\textsuperscript{b} and >>>30 min for D\textsuperscript{b} in control cells. It should be noted that the formation of endo H-resistant molecules in castanospermine-treated cells requires the action of a Golgi endo-mannosidase to convert glucosylated precursor oligosaccharides to a truncated form that can be acted on by Golgi glycosyltransferases (Moore and Spiro, 1990). Since the endo-mannosidase may be inefficient (Moore and Spiro, 1992), ER to Golgi transport rates may be even more rapid in the drug-treated cells. The enhanced transport observed in the presence of castanospermine was consistent with earlier findings, from our laboratory, in Drosophila cells where, in the absence of calnexin, heterodimers were transported more rapidly out of the ER than when calnexin was co-expressed (Jackson et al, 1994). Thus, the functions of calnexin in facilitating class I assembly and retaining incompletely assembled class I molecules that were established in Drosophila cells are fully reproducible in mouse cells treated with castanospermine. This congruency suggests that the
Figure 9. Effect of castanospermine on heavy chain-β2m assembly in MDAY.D2 cells.

(A) MDAY.D2 cells were pre-incubated for 1 h, radiolabeled with [35S]Met for 5 min, and then chased for the indicated times in the absence or presence of castanospermine. Cells were lysed in digitonin lysis buffer and Ld molecules were isolated with antibodies recognizing either total (anti-HC and 28-14-8s) or β2m-associated (30-5-7s) heavy chains. Proteins were digested with endo H and analyzed by reducing SDS-PAGE. The mobilities of class I heavy chains are indicated. Note: anti-HC antibody cross-reacts with Kd and Dd whereas 30-5-7s is Ld-specific.

(B) The results in panel A were quantified by densitometry and the amount of β2m-associated heavy chains at each time point was expressed as a percentage of the total heavy chains present in the pulse sample.
effects of castanospermine are due primarily to inhibition of calnexin interactions.

4.3.5 Castanospermine impairs folding of class I heavy chains

In Drosophila cells lacking calnexin, the inefficient assembly of class I molecules could be attributed largely to inefficient folding of the heavy chain as evidenced by a reduction in a conformational epitope defined by mAb 28-14-8S (for Dβ) and by increased aggregation of heavy chains (for Kb). To determine if a similar situation existed in castanospermine-treated mouse cells, we initially attempted to monitor acquisition of the 28-14-8S epitope in β2m-deficient murine RI E-ββ cells. These experiments proved unsuccessful due to very rapid degradation of Dβ heavy chains upon castanospermine treatment (see Figure 13). However, Dβ molecules expressed in EL4 cells, appeared more stable in the presence of castanospermine and were more amenable to this study. As shown in Figure 10A and 10B, the 28-14-8S epitope was acquired rapidly in untreated cells being present in 80% of total heavy chains following the 5 min pulse radiolabeling and reaching 100% between 10 and 20 min of chase. By contrast, only about 45% of total heavy chains acquired the epitope in castanospermine-treated cells. This result was remarkably similar to the 2 to 3-fold reduction in the formation of the 28-14-8S epitope observed in Drosophila cells lacking calnexin, again suggesting that the effect of castanospermine is likely attributable to a loss of heavy chain interaction with calnexin and further supporting a chaperone function for calnexin in facilitating heavy chain folding.

The reduction in heavy chain folding as defined by the 28-14-8S epitope was considerably less than the 5-fold reduction in Dβ assembly with β2m observed following castanospermine treatment (Figure 8C). If the reduced assembly was due primarily to the inability of heavy chains to acquire a conformation competent for association with β2m, then the 28-14-8S epitope was
Figure 10. Castanospermine treatment impairs folding of $\text{D}^b$ heavy chains in mouse cells.

(A) Mouse EL4 cells were pre-incubated for 1 h, radiolabeled for 5 min with $[\text{35S}]\text{Met}$, and chased for the indicated times all in the absence or presence of castanospermine. Digitonin lysates of cells were incubated with antibodies recognizing either total $\text{D}^b$ heavy chains (mAb 28-14-8S and anti-HC combined) or $\text{D}^b$ heavy chains with a folded a3 domain (mAb 28-14-88) and isolated proteins were digested with endo H and analyzed by reducing SDS-PAGE. Note that since the anti-HC antiserum does not discriminate between $\text{D}^b$ and $\text{K}^b$, both heavy chains appear in the total heavy chain pool but can be discriminated following endo H digestion. Only the mobilities of the endo H-sensitive and resistant $\text{D}^b$ heavy chains are indicated.

(B) The results in panel A were quantified by densitometry and the amount of 28-14-8S reactive heavy chains at each time point was expressed as a percentage of the total heavy chains present in the pulse sample.
not an accurate reflection of an assembly-competent conformation. To determine whether additional indications could be obtained for improper folding of D\textsuperscript{b} heavy chains following castanospermine treatment, D\textsuperscript{b} molecules were isolated with mAb 28-14-8S from control and drug-treated EL4 cell lysates and were analyzed by non-reducing SDS-PAGE for assessment of aggregation state. Interestingly, there was an increase in the formation of heavy chain dimers in castanospermine-treated cells (Figure 1, compare non-reducing panels + castanospermine). These dimers were sensitive to reduction indicating that they were covalently linked via inter-chain disulfide bonds (Figure 11). Furthermore, they were present in the 5 min pulse sample establishing that they form rapidly following heavy chain synthesis. Similar disulfide-linked homodimers have been described previously for non \beta\textsubscript{2}m-associated class I heavy chains (Capps et al., 1993). Therefore, at least some of the D\textsuperscript{b} heavy chains that have acquired a 28-14-8S reactive conformation exhibit characteristics of improperly folded molecules. There was no evidence for the formation of higher order oligomers or aggregates either in this SDS-PAGE analysis or by velocity density gradient centrifugation (Figure 12).

Velocity density gradient centrifugation was also used to assess whether aggregates of K\textsuperscript{b} heavy chains, analogous to those observed in Drosophila cells lacking calnexin (Figure 7B), formed in castanospermine-treated EL4 cells. Such aggregates were observed although they were less abundant. Approximately, 15% of K\textsuperscript{b} heavy chains were recovered as aggregates in fractions 4-16 whereas in untreated cells, even upon longer exposure, there were no K\textsuperscript{b} molecules in the denser fractions (Figure 12).

As described above, in Drosophila cells lacking calnexin, the degradation of D\textsuperscript{b} heavy chains whether free or \beta\textsubscript{2}m-associated was markedly accelerated relative to cells expressing
Figure 11. Effect of castanospermine treatment on the aggregation state of D\textsuperscript{b} molecules.

Mouse EL4 cells were radiolabeled in the absence or presence of castanospermine as described for Figure 6. At the indicated chase times, aliquots of cells were lysed in digitonin lysis buffer and D\textsuperscript{b} molecules were isolated with mAb 28-14-8S. Proteins were resolved by reducing and non-reducing SDS-PAGE and the mobilities of dimeric and monomeric D\textsuperscript{b} heavy chains are indicated.
EL4 cells were incubated for 24 hours with 400u/ml IFNγ to enhance expression of class I molecules. During this same incubation, cells were treated with or without 1mM castanospermine. The cells were subsequently radiolabeled for 15 minutes, lysed in RIPA buffer and subjected to glycerol gradient centrifugation as in Figure 4. Immunoprecipitation of Kβ and Dβ molecules from each fraction (combination of anti-HC, 28-14-8S and anti-8 antibodies) was followed by endo H digestion. The mobilities of Dβ and Kβ heavy chains are indicated. On longer exposure Kβ molecules can be seen as far as fraction 13 in castanospermine treated cells.

Figure 12. Aggregation state of class I molecules in EL4 cells
calnexin (Figures 3 and 4). In contrast, K\textsuperscript{b} molecules were considerably more stable in the absence of calnexin (Figure 2). To examine if a similar phenomenon occurred in castanospermine-treated mouse cells, the half-times of turnover were determined in pulse-chase biosynthetic experiments. Consistent with the observations made in \textit{L. pneumophila} cells, free D\textsuperscript{b} heavy chains expressed in R1E-D\textsuperscript{b} cells were degraded very rapidly following castanospermine treatment ($t_{1/2}$ = 20 min) compared to no detectable degradation over 20 min in untreated cells. Furthermore, the differential stability of D\textsuperscript{b} versus K\textsuperscript{b} heavy chains was apparent when EL4 cells were treated with castanospermine. Whereas both heavy chains were degraded very slowly under control conditions, the half-times of turnover were 60 min and 100 min for D\textsuperscript{b} and K\textsuperscript{b} heavy chains, respectively, following incubation with castanospermine (Figure 13). The increased turnover in response to castanospermine treatment was also observed for L\textsuperscript{d} molecules in MDAY cells (Figure 13).

4.3.6 Castanospermine reduces cell surface expression of class I molecules

The results presented above suggest that calnexin plays an important role as a molecular chaperone facilitating class I biogenesis in the ER. Since calnexin's influence is detectable at multiple levels including heavy chain folding, subunit assembly, and degradation of assembly intermediates, it was of interest to determine how these various effects are ultimately manifested in the expression of class I molecules at the cell surface. EL4 cells were treated in the absence or presence of castanospermine for 24 hours and the expression of B2m-associated molecules at the cell surface was determined by flow cytometry (Figure 14). The extended period of treatment with castanospermine was necessary to allow for turnover of class I molecules present at the cell
Figure 13. Castanospermine increases turnover of class I molecules in murine cell lines

Murine cell lines were pre-incubated for 1 hour, pulse radiolabeled for 5 minutes and chased for up to 100 minutes all in the presence or absence of castanospermine. Various class I molecules were immunoprecipitated, resolved on SDS-PAGE and amounts quantified by densitometry. The results are presented in two ways. In A, heavy chain levels in the presence of castanospermine are plotted as a percentage of the levels in the absence of drug treatment. In B, the approximate half times of turnover are given.
surface prior to treatment. No reduction in heavy chain or $\beta_2 m$ synthesis occurred during this time (data not shown). Figure 14 depicts the fluorescence data obtained following incubation of cells with $\beta_2 m$-dependent mAbs followed by FITC-conjugated secondary antibody, and shows a reduction in fluorescence intensity when cells were treated with castanospermine. Mean fluorescence intensities were calculated and the level of surface expression following castanospermine was determined as a percentage of the level of expression on untreated cells. For both $K^b$ and $\Omega^b$ the surface expression of $\beta_2 m$-associated molecules decreased to 30% of control values. The reduced surface expression can most likely be attributed to the decreased assembly efficiency that occurs in castanospermine-treated cells (Figure 8), since unassembled heavy chains either aggregate or are largely degraded intracellularly. These results suggest that calnexin interactions in the ER translate into more efficient cell surface expression of class I molecules. Since surface expression is a prerequisite for presentation of peptide ligands to cytotoxic T cells, calnexin is likely to contribute significantly to the efficiency with which a cell can present endogenous antigens.

4.3.7 Castanospermine inhibits folding and assembly of the human HLA-B27 molecule

There is some controversy as to whether calnexin interactions are as important in the biogenesis of human class I molecules as they are for mouse class I molecules. This is highlighted by the inability to demonstrate an interaction between calnexin and heavy chain-$\beta_2 m$ heterodimers in human cells although such an interaction is easily detected in mouse cells.
Figure 14. Effect of castanospermine on cell surface expression of class I molecules.

Mouse EL4 cells were cultured under normal conditions either with no additions or with two additions of castanospermine (50 μg/ml) at 24 h and 16 h prior to flow cytometric analysis. β2m-associated K^\text{b} and D^\text{b} molecules at the cell surface were detected by incubation first with mAb Y3 (for K^\text{b}) or with mAb B22-249 (for D^\text{b}) followed by FITC-conjugated goat anti-mouse IgG. Control samples omitting either the first or both antibodies were included in the analysis.
Calnexin association with unassembled human heavy chains has, however, been clearly demonstrated (Degen et al., 1992; Rajagopalan and Brenner, 1994). Human C1R cells, transfected with the class I HLA-B27 gene, were used to determine whether some of the functions of calnexin established for murine class I molecules can be extended to the human system. The folding and assembly states of B27 molecules expressed in these cells in the absence or presence of castanospermine were assessed by comparing reactivity with antibodies which recognize unassembled heavy chains (mAb HC-10), heavy chains with folded α3-domains (mAb 5H7), and β2m-associated heavy chains (mAb W6/32). Pulse-chase radiolabeling was used to follow the kinetics of formation of the various epitopes (Figure 15A and 15B). Castanospermine treatment resulted in a 3.5-fold decrease in assembly of the B27 heavy chain with β2m. Whereas 70% of heavy chains assembled over a 20 min period in control cells only 20% assembled in the presence of castanospermine. No further assembly was observed in either case beyond 20 min (data not shown). A less dramatic but reproducible difference was also observed in reactivity to mAb 5H7, which recognizes a conformational epitope in the α3 domain of the heavy chain regardless of association with β2m. About 45% of heavy chains from drug-treated cells acquired this epitope in contrast to 70% of heavy chains from control cells. These results are consistent with those presented above for murine class I molecules, suggesting that calnexin is required for efficient heavy chain folding and β2m assembly of both human and murine class I molecules. Whether there are species differences in how calnexin functions in downstream events, such as in peptide loading or interaction with the TAP peptide transporter, remain to be established.
Figure 15. Effect of castanospermine on the folding and assembly of the human HLA-B27 molecule.

(A) Human C1R cells expressing HLA-B27 were pre-incubated for 1 h, radiolabeled for 5 min with $[^35]S$Met, and then chased for the indicated times all in the absence or presence of castanospermine. B27 molecules were isolated from digitonin lysates using antibodies recognizing either unfolded (mAb HC-10), α3 domain-folded (mAb 5H7), β2m-associated (mAb W6/32) or total (all mAbs combined) heavy chains. Isolated proteins were analyzed by reducing SDS-PAGE and the mobility of the B27 heavy chain is indicated.

(B) The data from the experiment in panel A was quantified by densitometry and the various conformational forms of the B27 heavy chain were expressed as a percentage of the total heavy chain signal present at each chase time.
4.4 Discussion

Two independent approaches were used to assess the functions of calnexin in the biogenesis of class I histocompatibility molecules. Expression of mouse class I subunits in *Drosophila* cells either alone or in combination with calnexin permitted a direct evaluation of calnexin functions in a heterologous environment. In addition, the glucosidase inhibitor castanospermine was employed as an indirect means to prevent calnexin interactions within the environment of mammalian cells. The use of castanospermine to assess calnexin functions is dependent on its selectivity. The only other putative molecular chaperone that is known to be affected by castanospermine treatment is the soluble ER protein, calreticulin. Calreticulin shares a high degree of sequence identity with calnexin and it has been proposed to function as a molecular chaperone based on its transient interaction with newly synthesized glycoproteins that traverse the ER (Nauseef et al., 1995; Hebert et al., 1995). Preliminary results of experiments designed to examine whether calreticulin might interact with class I molecules are not presented here but essentially such interactions could be detected only in association with heavy chains following their assembly with β2m. Therefore, calreticulin is not involved in class I biogenesis at early stages of folding and assembly. This is in contrast to calnexin which binds during or immediately after heavy chain synthesis (Degen and Williams, 1991). Thus any effects of castanospermine treatment on the folding and assembly of class I molecules can most likely be attributed to prevention of calnexin binding

Consistent with this expected selectivity of castanospermine is the remarkable similarity in results obtained either by examining class I biogenesis in the absence of calnexin in *Drosophila* cells or by preventing calnexin binding through castanospermine treatment of mouse or human
cells. In the absence of calnexin interactions, assembly of class I heavy chains with β2m was substantially impaired. Heterodimers were formed at only 20-30% of the level observed in the presence of calnexin. The reduced assembly was largely due to misfolding of the heavy chains as evidenced by a reduction in mAb-defined conformational epitopes or by the formation of oligomers or aggregates. These findings demonstrate that calnexin facilitates the folding and assembly of class I molecules and provide the most direct evidence to date that calnexin functions as a bona fide molecular chaperone.

How calnexin promotes protein folding or assembly is the subject of considerable debate. Helenius and co-workers have proposed that the most important component of calnexin's interactions is its binding to Glc[ManGlcNAc] oligosaccharides on nascent glycoproteins. This specific oligosaccharide structure is maintained by a cycle of glucose removal and reglucosylation which is thought to continue as long as the glycoprotein remains incompletely folded. When folding is complete the glycoprotein is no longer a substrate for the glucosyltransferase that catalyzes reglucosylation and the cycle ends (Hammond and Helenius, 1994). In this view, the main function of calnexin is simply to retain incompletely folded or unassembled glycoproteins within the ER where folding enzymes and molecular chaperones such as BiP and GRP94 can assist the folding process (Helenius, 1994; Hebert et al., 1995). Although it has been clearly shown that calnexin retains non-native proteins in the ER (Jackson et al., 1994, Rajagopalan et al., 1994), its ability to facilitate class I heavy chain folding and assembly with β2m as demonstrated in the present study cannot be attributed only to retention. For example, the enhanced folding of D β heavy chains in the presence of calnexin is clearly evident by the end of a 5 min pulse radiolabeling in both Drosophila cells and in
castanospermine-treated mouse cells (Figures 6 and 10). The effect of calnexin on increasing the assembly of heavy chains with β2m is also obvious immediately following the pulse in *Drosophila* cells (Figures 2 and 5) and during the first 10 min of chase in mouse or human cells (Figures 8 and 15). Within these short time periods there are negligible differences in export of class I molecules from the ER in the absence or presence of calnexin.

How then does calnexin effect its functions? There is now considerable evidence that polypeptide associations constitute a substantial component of the overall interaction between calnexin and glycoproteins. Numerous examples exist of calnexin binding to proteins that either naturally lack N-linked oligosaccharides (Rajagopalan et al., 1994) or that have lost oligosaccharides through mutagenesis or treatment with tunicamycin (Arunachalam and Cresswell, 1995; Loo and Clarke, 1994; Carreno et al., 1995a). Furthermore, sites of interaction between calnexin and class I heavy chains have been mapped in proximity to the transmembrane region of the heavy chain where no N-linked glycans are located (Margolese et al., 1993; Carreno et al., 1995a). Finally, when complexes of calnexin and diverse membrane or soluble glycoproteins are digested with endoglycosidase H, all of the glycoprotein's oligosaccharide chains are accessible to the enzyme and can be removed without any dissociation of the deglycosylated protein from calnexin (Ware et al., 1995 (Chapter III); Arunachalam and Cresswell, 1995; Zhang et al., 1995). Based on these findings it is likely that calnexin possesses a polypeptide binding site in addition to its site for binding the Gal\{Man9GlcNAc2\} oligosaccharide. Consequently, calnexin may function in a manner analogous to other molecular chaperones, facilitating protein folding by binding to structural elements exposed in unfolded proteins (such as hydrophobic surfaces) and preventing inappropriate intermolecular
associations that lead to aggregation.

The current results provide support for this view. In the absence of calnexin extensive aggregation of K\textsuperscript{b} heavy chains was observed in \textit{Drosophila} cells and, to a lesser extent, in castanospermine-treated mouse cells. Significant formation of aberrant disulfide-linked D\textsuperscript{b} heavy chain dimers was also observed in mouse cells treated with castanospermine although the production of large aggregates was not detected. The formation of D\textsuperscript{b} aggregates may have been largely precluded by the rapid degradation of unassembled D\textsuperscript{b} heavy chains observed both in \textit{Drosophila} cells lacking calnexin and in castanospermine-treated mouse cells. Little degradation of D\textsuperscript{b} heavy chains occurred when they were associated with calnexin. Therefore, depending on the intrinsic sensitivity of a glycoprotein to degradation, interaction with calnexin may prevent the formation of aggregates or stabilize the polypeptide chain against proteolytic attack.

Calnexin also enhanced heavy chain folding by a mechanism other than preventing aggregation or degradation. For D\textsuperscript{b} heavy chains, the acquisition of a folded \textalpha3 domain as detected with mAb 28-14-8S was stimulated 2-3-fold by calnexin during a 5 min pulse radiolabeling (Figures 6 and 11). This effect was so rapid that it could not be attributed to differential degradation of heavy chains in the absence and presence of calnexin. The enhanced folding may be due to calnexin stabilizing partially folded conformations of the heavy chain because of the observation that in detergent lysates the 28-14-8S epitope is gradually lost with time in the absence of calnexin but is retained when the heavy chain remains associated with calnexin (unpublished data).

There are conflicting reports on the role of calnexin in facilitating folding and assembly of human class I molecules. In this study castanospermine treatment of human CIR cells
transfected with HLA-B27 molecules resulted in decreased folding of human molecules as assessed by acquisition of the 5H7 monoclonal antibody epitope. In addition there was a 2-3 fold decrease in assembly of B27-β2m heterodimers. Another study using the same cells transfected with B7 demonstrated a role for calnexin in acquisition of intrachain disulfide bonds but could not demonstrate an effect on assembly of B7-β2m heterodimers (Tector and Salter, 1995). One possible explanation for the discrepancy in the data is that the level of β2m in these cells was very high and even in the absence of castanospermine treatment there was only approximately 20% assembly. This suggests that β2m may have been limiting and thereby obscuring any effects calnexin may have on assembly. Alternatively there may be differences in class I allotype requirements for calnexin. This possibility can be tested by examining the effects of castanospermine on assembly of a range of human class I molecules.

A model of how calnexin functions involves a two-step mechanism for calnexin binding that incorporates both oligosaccharide and polypeptide components in the overall interaction (Ware et al., 1995, Chapter 111). In the first step, calnexin utilizes its lectin site to bind Glc\(\alpha\)Man\(\beta\)GlcNAc\(\beta\) oligosaccharides and this interaction serves to bring the nascent glycoprotein into proximity with the membrane-constrained chaperone. For most glycoproteins this step is a prerequisite for stable binding thereby explaining the remarkable efficacy of oligosaccharide processing inhibitors or tunicamycin in preventing calnexin binding. Subsequently, in the second step, calnexin associates with the unfolded polypeptide and at this stage the lectin-oligosaccharide interaction becomes less important for maintaining the complex (evidenced by the complete accessibility of the oligosaccharides to digestion with endo H).

Based on the present findings, it is likely that calnexin functions in the second step as a typical
molecular chaperone, preventing aggregation or rapid degradation and stabilizing folding or assembly intermediates. As yet it is not clear whether calnexin functions like other chaperones in cycles of binding and release and, if so, what would regulate such cycles. Helenius and co-workers have suggested that calnexin release and re-binding are regulated solely by cycles of glycoprotein deglucosylation and reglucosylation (Hebert et al., 1995) but such a model does not accommodate a polypeptide component in the binding interaction and is difficult to reconcile with the chaperone functions of calnexin demonstrated in the present study. Alternatively, the recent finding that the ER luminal domain of calnexin binds ATP raises the possibility that binding cycles could be regulated by nucleotide binding or exchange (ATPase activity was not detected; Ou et al., 1995). Such a hypothetical cycle might involve nucleotide-induced dissociation of the glycoprotein from calnexin followed by two-step re-binding and would continue until all polypeptide sites for calnexin binding are buried in the folded molecule. In this scenario, the deglucosylation-reglucosylation cycle maintains the Glc|Man9GlcNAc2 structure necessary for the first step in calnexin binding (or re-binding) and also facilitates final dissociation when the glucosyltransferase fails to reglucosylate the native glycoprotein.

It is important to point out that calnexin's interactions are not absolutely essential for the folding of class I heavy chains and their subsequent assembly with β2m, but rather increase the efficiency of these processes. In the absence of calnexin, about 40% of mouse or human heavy chains acquire correctly folded α3 domains as judged by reactivity to a single mAb and about 20-30% of heavy chains assemble with β2m. These findings suggest that the diversity of folding enzymes and molecular chaperones within the ER may provide a degree of redundancy in assisting protein folding. This notion is supported by the fact that cells can survive without
calnexin as evidenced by their viability for prolonged periods in the presence of glucosidase inhibitors such as castanospermine. Also, the recent demonstration that class I biogenesis appears normal in a calnexin-deficient mutant cell line suggests that, in the chronic absence of calnexin, alternative chaperone pathways can fully compensate for the loss of calnexin function (Scott and Dawson, 1995; Sadasivan et al., 1995). One candidate for a compensatory chaperone is BiP which appears to associate with several unglycosylated (and presumably not calnexin-associated) mouse class I molecules, but is not associated with their glycosylated counterparts (Fraser et al., 1987; Degen, 1995). No interaction of BiP with normal mouse class I molecules has been detected although transient binding during or immediately following translation cannot be excluded (Degen et al., 1992). A recent report has demonstrated that prolonged treatment of cells with castanospermine results in increased expression of BiP and grp94 mRNA (Pahl and Baeuerle, 1995). Preliminary results in murine cells demonstrate that both grp94 and BiP are upregulated by treatment in castanospermine for 12 hours (A. Vassilakos, data not shown). Both synthetic and total protein levels increase 2-3 fold with prolonged treatment with castanospermine. Preliminary co-immunoprecipitation experiments suggest that grp94 interacts with class I heavy chains and that in the absence of calnexin and calreticulin the interaction increases several fold. Consistent with Degen et al., 1992, BiP was not found to associate with class I heavy chains.

The most recent candidate for a cooperating chaperone is calreticulin which binds both human and mouse class I molecules only following heavy chain assembly with β2m (A.
Vassilakos, data not shown, Sadasivan et al., 1996 and Van Leeuwen and Kearse, 1996c). This is in contrast to calnexin that binds to both free and P2m-associated mouse heavy chains and only free human heavy chains. It is likely that these closely related proteins share the same lectin binding site based on their comparable sensitivities to castanospermine and the oligosaccharide binding studies presented in Chapter III. Importantly, their differential selectivity for class I assembly intermediates argues further for the existence of a second polypeptide binding site, a site that is apparently less well conserved than their lectin site. It will be of particular interest to determine if calnexin and calreticulin interact with class I molecules concurrently, with overlapping binding specificity or strictly sequentially during class I biogenesis and the extent to which their functions overlap. Reconstitution of calreticulin-class I interactions in Drosophila cells will be particularly helpful in this regard.

An examination of calnexin's role in class I biogenesis should ultimately address what the implications are for class I-restricted antigen presentation. It is clear from the results presented here that calnexin is required for the efficient assembly and subsequent expression of class I molecules at the cell surface. For the 20-30% of class I molecules that are able to assemble in the absence of calnexin, there is preliminary data indicating that they associate with the TAP peptide transporter and apparently acquire peptide ligands based on their stability in detergent lysates (W.-K. Suh, A. V., and D. B. W., unpublished data). Whether or not calnexin influences the spectrum of peptides bound is currently under investigation. Collectively, the results demonstrate that calnexin, while not absolutely necessary for class I biogenesis, plays an important role in the efficient folding, assembly, and surface expression of murine and human class I molecules and hence, in a cell's ability to present endogenous antigens.
CHAPTER V
DISCUSSION OF RESULTS AND
FUTURE DIRECTIONS
5.0 Summary of Results

Much of the early studies on the nature of calnexin interactions in the ER led to the hypothesis that calnexin functions as a molecular chaperone. Consistent with this role is the observation that calnexin binds to newly synthesized polypeptides and releases them after completion of folding and/or assembly. The data presented in Chapter IV approached the question of calnexin function in two ways. The role of calnexin in folding and assembly of class I molecules was examined in a heterologous expression system, wherein no drugs were required to produce a calnexin negative condition. In addition, function was examined in mammalian cells using castanospermine. The results from the two approaches were similar and provided clear evidence for calnexin acting as a molecular chaperone. Furthermore, the results in Chapter IV provide evidence that calnexin improves folding and assembly by preventing degradation and aggregation of class I heavy chain molecules. These observations are consistent with how other molecular chaperones act and provide a common theme as to how diverse chaperones function.

The remainder of the work presented here focused on characterizing the interactions between calnexin and newly synthesized proteins. Recent studies suggested that calreticulin also acts as a chaperone via a mechanism similar to calnexin’s and so calreticulin was included in these studies as well (see Chapter III). Results of experiments using purified recombinant calnexin and calreticulin demonstrated direct and specific binding to the Glc,Man,GlcNAc$_2$ oligosaccharide. This structure is present very early on newly synthesized glycoproteins and provides evidence for calnexin and calreticulin functioning early in the biogenesis of newly synthesized glycoproteins within the ER. In addition, substrate specificity studies, coupled with competition experiments between these proteins, suggested that they possess similar lectin sites
which bind to the Glc,Man,GlcNAc oligosaccharide with the same affinity. Further characterization of the lectin properties of calnexin and calreticulin localized the lectin domain to a central homologous region characterized by two novel, tandemly repeated sequence motifs. We also addressed whether polypeptide mediated interactions occur. Both membrane associated and soluble calnexin substrates could be completely de-glycosylated without loss of interaction with calnexin. This data provides evidence for polypeptide mediated interactions in addition to the lectin-oligosaccharide mode of interaction. These findings led us to propose a model for calnexin/calreticulin action that involves a sequential two-step mechanism encompassing both lectin and polypeptide modes of interaction (see next section).

5.1 Models of Calnexin and Calreticulin Interactions

While it has become clear that the lectin properties of calnexin and calreticulin are essential in the recognition of substrates it is not yet clear the extent to which polypeptide mediated interactions are involved. This is perhaps the greatest point of contention in the two basic working models put forward in the literature (Figure 1). We propose a model for calnexin and calreticulin function in which interaction with oligosaccharide provides initial recognition of newly synthesized polypeptide. This is followed by polypeptide mediated interactions which stabilize the complex (Figure 1A). It is envisioned that polypeptide binding is analogous to that observed with other chaperones and it is this interaction that prevents aggregation and thus enhances folding efficiency. In this model the release of substrate occurs by an as yet unidentified mechanism. ATP could be involved in this process but this is highly speculative and is suggested only by analogy to the Hsp 70 and Hsp 60 mechanisms of action. Following
A. Model of calnexin and calreticulin interactions with newly synthesized polypeptides based on the studies described in this thesis. Initial binding and re-capture after release are mediated via interactions with oligosaccharide on the substrate protein. Subsequent to the initial lectin mediated interaction, polypeptide mediated interactions occur which stabilize the complex. Mechanism of release from the chaperone is not clear.

B. In an alternative model, the interactions between calnexin and calreticulin and newly synthesized proteins are strictly via lectin binding and do not include polypeptide mediated interactions. Glucosidase II acts on the substrate while bound to the chaperone to cause release (Helenius, 1995).

Figure 1
substrate release, the action of glucosidase II trims the terminal glucose residue on the oligosaccharide. If the substrate is completely folded it will not be a substrate for UDP-glucosyltransferase and can exit the ER. In contrast, if the substrate is unfolded, UDP-glucosyltransferase will re-glucosylate the oligosaccharide and the substrate can undergo another round of binding to and release from calnexin. In the opposing model, calnexin and calreticulin binding to oligosaccharide provides the only mode of interaction with newly synthesized proteins (Figure IA). The release and re-binding cycle is regulated by glucosidase II (triglycerase release) and UDP-glucosyltransferase (re-binding). This cycle would only retain glycoproteins in the ER and, as a result, it is not clear from this model how any chaperone function would be effected.

In the first model, interaction with oligosaccharide provides initial recognition of newly synthesized polypeptide, followed by polypeptide mediated interactions which stabilize the complex. The evidence for polypeptide mediated interactions comes from three different approaches. First, a study examining interactions between various mutants of class I molecules and calnexin in murine cell lines implicated the membrane proximal regions of murine class I molecules in interaction with calnexin. In this study chemical cross-linking of detergent lysates was used to examine these interactions. The results provide evidence for interaction at sites remote from either the luminal lectin domain of calnexin or the carbohydrate sites on class I heavy chains (Margolese et al., 1993). Second, there are several reports of naturally unglycosylated proteins or mutants lacking N-linked glycosylation sites that bind to calnexin (see Table III, Chapter I). It is not clear from these studies whether these mutants bind with similar affinity or to the same extent as glycosylated substrates of calnexin. Finally, there are
several examples where oligosaccharide can be removed from calnexin associated molecules without loss of interaction (Ware et al., 1995, Zhang et al., 1995 and Arumanchalam and Cresswell, 1995). In these latter experiments calnexin complexes were treated with endoglycosidase H to remove oligosaccharide on the associated glycoproteins. Complete removal of oligosaccharide did not result in dissociation of either murine class I molecules, human class I molecules or human class II molecules from calnexin. The argument that these interactions are purely due to aggregation or trapping of molecules in detergent micelles (Helenius et al., 1997) is not likely accurate. The same observations are seen in systems using the soluble substrate, α,-antitrypsin, for calnexin (Chapter II) or in completely soluble systems with calreticulin and myeloperoxidase (Bill Nauseef, personal communication). Furthermore, the class I heavy chain molecules isolated with calnexin (Chapter III) were not likely to be aggregated given that addition of β,αm and peptide to isolated complexes results in 100% formation of trimer molecules (A. Vassilakos and E.K. Mitchell, unpublished results). In addition there is recent evidence that calreticulin interacts with polypeptide regions of proteins in a specific manner (Coppolino et al., 1995; also see Chapter III Introduction and Discussion).

In contrast to the preceding experiments, there are now two examples where lectin mediated interactions occur in the absence of peptide interactions (Rodan et al., 1997 and Zapun et al., 1997). In both systems ribonuclease was the substrate for interaction with calnexin and calreticulin. In Rodan et al., constructs of ribonuclease B were engineered with two glycosylation sites. Ribonuclease was expressed by an in vitro transcription reaction followed by translation/translocation in a reticulocyte lysate/pancreatic microsome system. Treatment with μM concentrations of deoxynojirimycin produced monoglucosylated forms of ribonuclease.
Complexes of calnexin and calreticulin with ribonuclease could then be isolated from detergent lysates by immunoprecipitation. Treatment of the lysates with peptide N glycanase F, to de-glycosylate glycoproteins, resulted in a complete loss of interaction between the chaperones and ribonuclease, suggesting only lectin mediated interactions were present. A mutant form of ribonuclease missing essential disulfide bonds (mutated Cys residues) that is more trypsin sensitive than the wild type ribonuclease was engineered to determine whether folding state affects interaction with calnexin or calreticulin. There was no difference in interaction between the wild type (folded) or mutant (unfolded) ribonuclease and calnexin. These results suggest that for ribonuclease the oligosaccharide moiety is the only recognition element.

In Zapun et al., a purified system was designed to assess interactions between ribonuclease B containing a single oligosaccharide and a soluble form of calnexin containing the lumenal domain of calnexin. Purified ribonuclease B was monoglucosylated by treatment of the protein with purified UDP-glucosyltransferase and radiolabeled UDP-glucose. Only a small percentage of the molecules were glucosylated due to the low abundance of the Man,GlcNAc$_2$ oligosaccharide on ribonuclease B. Unfolded ribonuclease could be produced by reduction of the molecules. Subsequent re-folding of reduced ribonuclease could be achieved in the presence of a redox buffer and PDI. Interaction between calnexin and different conformations of ribonuclease was examined. All forms of ribonuclease were found associated with calnexin, suggesting only lectin mediated interactions. In contrast to Rodan et al., treatment of the complexes with peptide N glycanase F was not sufficient to dissociate the complex, but treatment with endoglycosidase H was capable of effecting dissociation. Again the endoglycosidase H results suggest only lectin interactions between calnexin and ribonuclease.
Although interactions between calnexin or calreticulin and ribonuclease were demonstrated in these studies, these were weak. In Zapun et al., interactions with calnexin were only demonstrated by rapid gel filtration separation at low temperatures. In Rodin et al., an additional oligosaccharide chain was engineered into ribonuclease before interaction with calnexin and calreticulin could be seen. These observations are not consistent with the often stable interactions seen between calnexin and its in vivo substrates (Chapter II and IV). It is possible that ribonuclease may not contain the necessary peptide motifs necessary for interaction with calnexin or calreticulin. Furthermore, the observed lack of polypeptide interactions with reduced forms of ribonuclease are difficult to interpret because it is not clear the extent to which reduced ribonuclease is unfolded. Local secondary structure may mask any polypeptide sequences suitable for interaction with calnexin.

In the study from Zapun et al., reduced/unfolded ribonuclease was incubated in a redox buffer with PDI in the presence or absence of calnexin. Native gel electrophoresis at various time points demonstrated that there was no difference in the yield of correctly-folded ribonuclease either with or without soluble calnexin. Although this may seem to suggest that calnexin is not acting as a chaperone, the results are not conclusive. Firstly, calnexin is a membrane protein and so the soluble construct, used in this study, may not contain segments that are important for chaperone function. Secondly, and more importantly, in the past chaperone activity has been demonstrated under conditions where an unfolded protein would aggregate and misfold, i.e., under concentrated conditions much like the ER environment. It is under these conditions that chaperones function, by binding to exposed hydrophobic sites and preventing inappropriate inter and intra-molecular interactions. In addition, chaperones often slow the rate
of folding but ultimately increase the yield of correctly folded molecules. It would be difficult to demonstrate such an activity under conditions that favour correct folding of the protein without assistance, such as those used by Zapun et al. It remains to be seen whether other substrates, more prone to aggregation and misfolding than ribonuclease, demonstrate the same interactions with calnexin or whether polypeptide mediated interactions will be required.

The second major difference between the models presented in Figure 1 is the mechanism of release of substrate from calnexin/calreticulin. The studies of Zapun et al., and Rodan et al., address this issue in the context of the lectin only model in Figure 1B. An interesting difference between these two studies is the effect of glucosidase II on complex stability (see Figure 1 for glucosidase II action). Rodan et al., demonstrated that glucosidase II treatment of detergent lysates results in loss of association between calnexin and ribonuclease. Furthermore, if glucosidase II activity is inhibited by addition of castanospermine, calnexin complexes remain for prolonged periods. The authors suggest that these results are consistent with their model in which dissociation from calnexin and calreticulin is triggered by glucosidase II (Figure 1B). However, this is unlikely given that glucosidase II would require access to the terminal glucose that has been shown to interact with the lectin domain of calnexin and calreticulin (Chapter III, Petrescu et al., 1997). It is more likely that in the microsomal system examined, calnexin interactions are dynamic and the substrate is released and then re-binds during the incubation (i.e., in the lectin only model it is more likely that the Kₜ controlled dissociation of substrate from calnexin followed by glucosidase II action accounts for the release of substrate. This would not be surprising given the apparent low affinity of calnexin/calreticulin for oligosaccharide). During release and re-binding, the newly synthesized glycoprotein would lose the terminal
glucose via glucosidase II action and then fold such that UDP-glucosyltransferase would be less likely to recognize and re-glucosylate the oligosaccharide. As a result, a lower proportion of molecules would re-bind during the incubation. Under conditions where glucosidase II is inhibited, the monoglucosylated substrate would be stable and thus capable of re-binding to calnexin without a need for UDP-glucosyltransferase even when completely folded. A dynamic view of the observations in the microsomal system would be consistent with the apparently contradictory findings of the study by Zapun et al. In this study, oligosaccharide on ribonuclease B bound to a large excess of calnexin (driving the equilibrium to the bound only state) was not accessible to glucosidase II (Zapun et al., 1997).

In the case of the first model that involves polypeptide interactions between calnexin/calreticulin and their substrates (Figure 1A), what might constitute the trigger for release from calnexin? If a cycle of binding and release is to be evoked, much like hsp70 and hsp60 reaction mechanisms, there must be an event which results in release from these proteins. As yet there is no clear evidence for what causes release. Calnexin and calreticulin bind Ca" and there is some evidence for their binding ATP. Therefore ATP binding, ATP hydrolysis, changes in Ca" conditions etc., are potential avenues to pursue (see Introduction). In any of these scenarios there is likely to be a concomitant conformational change that could result in a lower affinity for the bound substrate. In addition to what causes release at each step of the cycle, we must also consider what ultimately causes the release from calnexin and transport out of the ER. Again the process could be looked at from a dynamic perspective. For example, as class I heavy chain molecules fold and assemble with β2m and peptide they undergo several distinct conformational changes that can be measured experimentally (Chapter IV). At each step one
could envision that the heavy chain oligosaccharides become progressively less susceptible to
the action of UDP-glucosyltransferase, which prefers unfolded substrates. Early in its biogenesis
the heavy chain would likely be a good substrate for UDP-glucosyltransferase. As a result,
the molecules assemble it is likely that re-binding to calnexin (via re-glucosylation) would be
slow and the rate of transport out of the ER would predominate. This model could be tested
in one of the cell-free microsomal systems that reconstitute ER-Golgi transport. One intriguing
possibility is that transport out of the ER occurs via binding to the mannose selective lectin
ERGIC53. ERGIC53 is a putative sorting receptor found to recycle between the ER,
intermediate compartment and the cis-Golgi (see Introduction). Retention in the ER versus
transport to the Golgi could be the result of competition between binding to
calnexin/calreticulin and the sorting receptor. Early in their biosynthesis relatively unfolded
molecules would be ideal substrates for the UDP-glucosyltransferase and therefore calnexin and
calreticulin while being poor substrates for the mannose-selective ERGIC53. As glycoproteins
fold they would become poor substrates for the UDP-glucosyltransferase and better substrates
for ERGIC53. As a result, transport of folded molecules would be favoured over unfolded
molecules. As yet there is no experimental evidence that ERGIC53 functions as a lectin in vivo
or that the putative lectin activity is necessary for transport of proteins out of the ER. Therefore,
the above mechanism remains a speculative, but nonetheless intriguing model for how lectins
may have evolved to function in ER quality control.

As a final note on the mechanism of action of calreticulin and calnexin, researchers have
found that deletion of the calnexin gene in S. pombe results in a lethal phenotype (Parlati et al.,
1995 and Jannatipour and Rokeach, 1995). In contrast, deletion of the UDP-glucosyltransferase gene, which is stress inducible, results in no observable defect (Fernandez et al., 1996). This is somewhat surprising given the putative role of the re-glucosylation enzyme in the proposed calnexin reaction cycle. One possibility is that peptide mediated interactions are sufficient in these yeast cells to provide efficient interaction with calnexin. Alternatively, the action of glucosidase II in these cells may be sufficiently slow to allow for interaction and chaperone function before loss of the terminal glucose residue occurs. Determination of the lectin activity, if any, of *S. pombe* calnexin, analysis of the glucosidase activities present and analysis of the interactions between calnexin and its substrates in these cells may provide clues as to why there are disparate requirements for calnexin and UDP-glucosyltransferase.

5.2 Evolution of Chaperones with Unique Organelle Specific Substrate Recognition

What has become clear from the results presented here and from other studies is that calnexin and calreticulin constitute a new chaperone family. There is some evidence to suggest that these molecules may interact with polypeptide elements on newly synthesized proteins, much the same as classical molecular chaperones. In addition, they appear to have uniquely evolved to function in the ER where glycoprotein synthesis predominates over the synthesis of non-glycosylated proteins. Oligosaccharide on polypeptides serves many functions on mature polypeptides ranging from intracellular targeting, cell-cell interactions, immunological functions and even apoptosis (reviewed in Rademacher et al., 1988, Lasky, 1995, and Gamberg and Tolvanen, 1996). Whether calnexin and calreticulin evolution exploited pre-existing glycosylation pathways to increase efficiency of interaction with substrates or vice versa remains
an interesting question for evolutionary biology. As more calnexin homologues are sequenced
this question may be explored. It is interesting that N-glycosylation and the expression of
calnexin and calreticulin appear to go hand in hand. In addition, even though complex glycans
have evolved only in multicellular organisms, the early high mannose processing events that
produce substrates for calnexin and calreticulin are conserved from yeast to mammals. Given
this observation, it would be of interest to examine whether the yeast homologues have lectin
properties similar to those observed for mammalian calnexin and calreticulin. Analysis of the S.
cerevisiae calnexin is perhaps the homologue of greatest interest. This homologue has the lowest
percentage sequence identity of all the calnexin homologues (Chapter 1, Figure 12). Although it
retains the critical motif 1 and 2 sequences the rest of the molecule is quite divergent, differing
in overall length, domain structure, and may not even be membrane associated. If this divergent
member of the family retains oligosaccharide binding it would be interesting to see whether it
can restore viability to the S. pombe calnexin-deficient strain. This may address the question of
whether the lectin domain alone is sufficient for function or whether there are other important
interactions that are missing in the S. cerevisiae calnexin. Interestingly, S. cerevisiae
also differs from other species in that KRE5, the putative glucosyltransferase, does not have high
sequence similarity to other glucosyltransferases. Whether this protein is in fact a transferase
remains to be determined. Given that both the putative calnexin and glucosyltransferase
homologues are so different in S. cerevisiae it will be interesting to see whether a calnexin-like
chaperone pathway is retained in S. cerevisiae.

Yeast do not contain calreticulin homologues and this may account for the observation
that deletion of the calnexin gene in S. pombe results in a lethal phenotype. It is odd that deletion
of the calnexin gene in \textit{S. cerevisiae} is not lethal. This suggests that lethality in the \textit{S. pombe} knockout may be due to some non-chaperone function of calnexin. Alternatively there could be a calnexin substrate in \textit{S. pombe} that is critical for viability that is absent or not essential in \textit{S. cerevisiae}. Given that yeast do not have a calreticulin homologue, the evolutionary advantage, if any, of expressing both chaperones is another interesting question for evolutionary biology.

5.3 Calnexin Versus Calreticulin - Substrate Specificity

Given the similarities in their mechanisms of interaction it is not surprising that there are a growing number of examples of shared substrates for calnexin and calreticulin, e.g., human and mouse class I molecules, GLUT 1 glucose transporter, influenza HA and the T cell antigen receptor (see Chapter 1, Table III). The extent of functional redundancy that occurs and the implications for glycoprotein biogenesis in the ER remain unknown and need to be examined. Although there may be some redundancy in function, one of the outstanding questions is what forms the basis of the observed differences in substrates between calnexin and calreticulin (see Chapter 1, Table III). Furthermore it would be interesting to examine whether these differences limit the extent to which these molecules can compensate for the loss of one of these two chaperones. There are several possible explanations for differential substrate binding. These chaperones, although sharing similar oligosaccharide specificities, may differ in their putative polypeptide binding components, much like peptide binding differences observed for Hsp 70's. This is supported by the observation that calnexin binds early in the biogenesis of human class I molecules prior to \(\beta_2\mbox{-m}\) binding while calreticulin does not bind until after formation of dimers (Sadasivan et al., 1996). Thus, there appears to be differences in which conformation is
recognized by either chaperone. This supports the possibility that the oligosaccharide can mediate initial interactions to either of these chaperones but that stable interaction is determined by polypeptide mediated interactions.

Another explanation for the observations with human class I molecules is based on the fact that the ER environment for each of the two molecules is different. Calnexin is tethered to the membrane via a transmembrane domain. In contrast calreticulin is a soluble molecule that can only associate with the membrane via interactions with membrane proteins. As a result each may encounter different regions of any given polypeptide not only as it is being synthesized but also as the protein matures. Further evidence for this comes from a study that mapped the interaction to either calnexin or calreticulin of each of the oligosaccharides in hemagglutinin (HA). This study demonstrated that there are preferences for binding to specific oligosaccharides along the HA molecule. Calnexin bound preferentially to the membrane proximal stem and calreticulin bound to the more lumenally disposed globular head of HA (Hebert et al., 1997, Helenius et al., 1997). It remains to be determined whether this preference is related to positioning in relation to the membrane or to the polypeptide regions exposed in the vicinity of each oligosaccharide being different. In another study, calreticulin was engineered with a transmembrane domain and as a result the pattern of interaction with newly synthesized glycoproteins resembled that of calnexin and not calreticulin (Wada et al., 1995). This is the most compelling evidence that calnexin and calreticulin are differentially exposed to the ER lumenal compartment.

In the field of class I biogenesis, an overriding question is why, in murine cells, dimers interact with calnexin but in the human system dimers interact almost exclusively with
calreticulin. It is possible that class I molecules differ enough between these species to give rise to differences in how they bind calnexin and calreticulin. The presence of only one oligosaccharide on human class I molecules compared to 2 or 3 on murine class I molecules may contribute to differences in binding to calnexin and calreticulin. This possibility can be tested by introducing glycosylation sites in human class I molecules or deleting them in murine class I molecules followed by analysis of interactions with calnexin and calreticulin. Alternatively, calnexin and calreticulin, while recognizing similar oligosaccharide moieties, may differ in their polypeptide binding specificity. Expression of chimeric human/mouse class I molecules followed by analysis of interactions with these chaperones would address this possibility.

Understanding the basis for these observed differences would provide a basis for understanding how each chaperone functions, spatially and temporally in class I biogenesis in both the human and mouse models.

Another area which may provide clues to calnexin and calreticulin structure/function relationships are several observed species differences. For example murine class I molecules were expressed in Drosophila cells and there was little or no interaction between these molecules and the endogenous calnexin (Chapter IV). This is in spite of the fact that the Drosophila calnexin appears to have intact motif 1 and 2 regions which are responsible for the lectin properties of these molecules (Cristodoulou et al., 1997). It would be surprising if oligosaccharide binding specificity was conserved between calnexin and calreticulin and yet lost in a calnexin homologue. This question will be addressed as soon as the Drosophila calnexin homologue is examined and its lectin properties, if any, are established. In the same study, dog calnexin was introduced into these cells and was sufficient to restore murine class I biogenesis.
As more and more calnexin and calreticulin genes are cloned and compared this provides the possibility of structure/function predictions. These can then be tested by mutagenesis and domain swap experiments.

Most of the work discussed here involves the study of calnexin and calreticulin in homologous and heterologous cell culture expression systems. In these cases a calnexin deficient environment has been achieved through heterologous expression (Drosophila system), drug treatment (mammalian cultured cells) and, in one case, through mutagenesis of a human tissue culture cell line (Scott and Dawson, 1995). As a result we are only examining the interplay between calnexin and calreticulin in an artificial environment. In cell culture, expression levels are not reflective of in vivo levels which may constantly change during development, differentiation and even as growth conditions change in the environment. This limitation of cell culture studies is highlighted by the finding that calreticulin knockout mice are not viable and die during embryogenesis (Coppolino et al., 1997), yet castanospermine treated cultured cells or cells lacking glucosidases remain viable. The result with calreticulin deficient mice suggests that either calnexin expression is absent or, alternatively, that calnexin cannot functionally replace calreticulin at a critical developmental stage. Further analysis of the lethal defect in these mice as well as determining the levels of calnexin present prior to death of the embryos will undoubtedly address this question. Furthermore it would be interesting to see if the converse is true for calnexin knockout mice.

Much of the discussion surrounding the functions of the various ER chaperones appears to focus on the level of redundancy between their functions. Calnexin, calreticulin and GRP94 all appear to interact with murine class I molecules and, in the human system, BiP is also
involved (Chapter IV and data not shown). The finding that calnexin is essential in S. pombe, which lacks calreticulin, but not when calreticulin remains (CEM-NKR, human cell line), supports the possibility of redundancy in calnexin/calreticulin function which compensates for loss of calnexin (Scott and Dawson, 1995). The finding that castanospermine treated cells are viable even in the absence of both calnexin and calreticulin interactions suggests yet another level of redundancy where other ER chaperones can compensate for the loss of these two components of the ER folding machinery (Chapter IV). Although there is likely some compensation, in castanospermine treated cells the efficiency of folding is compromised suggesting that chaperone functions are not identical. The concept that redundancy between these chaperones has evolved to prevent problems caused by the loss of a single component, is likely only partially correct. It is more likely that these chaperones act in a complementary or sequential fashion in the ER. This would create an efficient chaperone machine that co-operates to provide the maximum opportunity for newly synthesized proteins to fold correctly. This concept is supported by the recent work of Tatu and Helenius that demonstrates a network of weakly associated ER chaperones that are involved in the early stages of polypeptide biogenesis (Tatu and Helenius, 1997). This is reminiscent of the sequential interactions between unfolded proteins and hsp70/40 followed by interaction with the chaperonins in the eukaryotic cytosol and bacteria (see Chapter I). Expression of different combinations of ER chaperones along with model substrates in Drosophila cells and also genetic manipulation of chaperone populations in yeast cells are just two experimental approaches that may provide insights into the sequential and cooperative nature of ER chaperone function. These methods would complement pulse/chase radiolabelling and co-immunoprecipitation studies which identify, temporally,
5.4 Future Directions

The preceding discussion outlined some general directions to address the functions of calnexin and calreticulin and their possible overlapping roles as molecular chaperones. The work presented here leads to some immediate future studies based on the methods established in the laboratory.

The Drosophila heterologous expression system provides an excellent opportunity to examine the structure/function relationship between calnexin and calreticulin in more detail. The system provides the ability to examine not only interaction with newly synthesized glycoproteins but also assays for “quality control” and chaperone functions. Intracellular transport rates, folding and assembly efficiencies as well as aggregation and degradation rates can be examined in detail. The results of these experiments using calnexin as the exogenous chaperone were presented in Chapter II. This same system can be applied to calreticulin, to ask whether it can also act as a molecular chaperone in class I biogenesis. As demonstrated in Chapter II, examining class I biogenesis in this system is ideal in that there are distinct conformational changes during folding and assembly that are discernible with a variety of available reagents. Furthermore, this system provides an opportunity to examine calnexin and calreticulin interactions and functions either independently or when co-expressed in the same cell. This could address issues such as functional redundancy, specificity differences and the possibility of sequential interactions with the same substrate at different stages of folding and assembly.
The results presented in Chapters II and IV provide an emerging picture of the interactions between calnexin/calreticulin and their substrates. The homologous region between calnexin and calreticulin defines the lectin domain and as such is sufficient for interaction with oligosaccharide on newly synthesized polypeptides (Chapter III). The *Drosophila* expression system can be used to ask whether the lectin domain is sufficient for chaperone activity. All of the oligosaccharide binding studies were performed using soluble calnexin, and so one could ask whether this construct can function similar to the full length calnexin bound to the membrane in *Drosophila* cells. This kind of mutagenesis coupled with expression in *Drosophila* cells is ongoing in the lab. Deletion of the C-terminal cytosolic domain is also being done to ask whether the cytosolic domain (containing the phosphorylation sites) plays a role in calnexin function. In the future these studies can be expanded to try to localize the putative peptide binding site of calnexin and calreticulin. Removal of this site, followed by co-precipitation experiments would determine whether the oligosaccharide mediated interaction is sufficient to maintain a stable calnexin/substrate complex. The nature of this region could provide clues as to the differential binding specificities of these molecules. Deletion, domain swap (between calnexin and calreticulin) and site directed mutagenesis experiments of both the lectin domain and the putative peptide binding domain would lead to an understanding of how these regions function and at what step in the reaction cycle they are required (Figure 1).

While cell biological methods are useful in defining components of a biological process and establishing their function in that process there are limits to the mechanistic detail that can be elucidated. To fully understand the interactions and reactions occurring, one must develop *in vitro* assays for function. The greatest strides in understanding chaperone function at a molecular
level have come from a combination of structural analyses coupled with elegant in vitro folding assays using purified chaperones and single model substrates (Chapter I, Hsp70 and 60 families). These studies require large amounts of purified calnexin and calreticulin as well as a model substrate for folding assays. There has been success in the lab in the production of soluble calnexin and the GST fusion system is currently being used to produce large amounts of calreticulin for structural analysis. Both calnexin and calreticulin are too large for current NMR technology and as a result the approach being taken is X-ray crystallography. Coupled with structure determination of the chaperones alone, it would be of great interest to determine the structure of a co-crystal with oligosaccharide. If this method is successful it will provide exquisite details of the lectin site in calnexin and calreticulin and the molecular interactions with oligosaccharide. Based on this structure one could predict critical residues that would be targets for mutagenesis and subsequent functional analysis. Although full length calnexin and calreticulin are too large for NMR studies some of the GST fusion mutants would be of correct size once cleaved from the GST using the available Factor Xa proteolytic site. The dynamics of calcium binding, oligosaccharide binding and the effects of nucleotide could be examined.

Although the crystal structure of calnexin and calreticulin bound to oligosaccharide will give clues to the interactions present, it is not until an in vitro folding assay is established that the mechanism of action of these chaperones can be fully understood. Calnexin and calreticulin are unique chaperones in that they require a defined oligosaccharide to function properly. As a result traditional model substrates, such as luciferase and rhodanese, are not available for these studies. One method for production of model substrates is to use glycoproteins bearing Man,GlcNAc, oligosaccharides and to re-glucosylate them using purified recombinant UDP-
glucosyltransferase (Figure 1). This approach has been successful in other laboratories using
rhinonuclease as the substrate for re-glucosylation. At the moment this method appears to be best
for producing small quantities of substrate for in vitro studies (Zapun et al., 1997). Once an ideal
substrate is established, the model proposed in Chapter III, and the updated version including
calreticulin (Figure I), can be tested and compared to the other models put forward in the
literature (see above discussion on mechanism of action). The effects of ATP, UDP-
glucosyltransferase and EGTA on folding of the model substrate can be examined in detail.

Given that oligosaccharide binding appears to be a prerequisite for subsequent
polypeptide interactions there is the possibility that engagement of the lectin domain opens a
polypeptide binding region. Like DnaK conformation is affected by nucleotide binding and
interaction with Dna J. The in vitro folding assay could provide evidence for this by examining
whether addition of oligosaccharide results in binding of hydrophobic peptides to calnexin and
calreticulin. Like the Dna K example, structure determination, with or without oligosaccharide,
will indicate whether there is a conformational change upon binding to oligosaccharide. This is
an intriguing possibility because it would provide control for the binding of substrate. The
peptide binding site could open in response to the initial interaction with oligosaccharide, rather
than exposing a hydrophobic surface at an inappropriate time when substrate is not available.
This is only a possibility, but provides an example of the kinds of questions that can be
addressed with in vitro folding assays.
5.6 Conclusions

Insights into the functions of calnexin and calreticulin and the molecular interactions involved have evolved greatly in the past few years. Through the work presented here as well as from the work of others we have come to understand that these molecules comprise a new family of general molecular chaperones. They have evolved an elegant mode of interaction that exploits reactions unique to the ER and as a result sets them apart from other chaperones. This work provides a basis for future studies that will increase our understanding of the molecular interactions and reactions that occur during calnexin/calreticulin mediated folding.
REFERENCES


