1-1-2011

A novel in vivo protein refolding technique

Yuefei Huang
Wayne State University,

A NOVEL IN VIVO PROTEIN REFOLDING TECHNIQUE

by

YUEFEI HUANG

DISSERTATION

Submitted to the Graduate School

of Wayne State University,

Detroit, Michigan

in partial fulfillment of the requirements

for the degree of

DOCTOR OF PHILOSOPHY

2011

MAJOR: BIOCHEMISTRY AND
MOLECULAR BIOLOGY

Approved by:

_____

Advisor Date

_____

_____

_____

_____

DEDICATION

To

My parents, Zhengjian Huang and Linying Zhang

ACKNOWLEDGEMENTS

I would like to express my sincere gratitude to my advisor, Dr. Jianjun Wang, for training, support and advice he provided me during my Ph.D. studies.

I would like to thank my committee members, Dr. Bharati Mitra, Dr. Christine Chow and Dr. Zhe Yang. Your advice and help are much appreciated.

I would like to thank Dr. Marilynn Doscher for your help and support, especially to us foreign students.

Thanks to all the group members: Qianqian, Arun, Congmin, Bin, Youcai, Yunhuang, Yonghong, Victoria, Chris, Nan, Leo, Carrie and more. It's been a wonderful experience working with you all. Thank you for your friendship and great memories.

Thanks to all my friends: Yanlin, Song, Yong, Jianbo, Zhigang and Hui. With you guys, I have wonderful years at Wayne State University.

Finally, I would like to thank my parents, who have always been there supporting me and instilling in me the value and the importance of education. Special thanks to my dearest husband Eric Hales: thank you for your unconditional love and support through this process. Thank you for giving professional advice on my thesis.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

AD	Alzheimers disease
apoE	Apolipoprotein E
ApoER2	Apolipoprotein E receptor 2
APC	Adenomatous polyposis coli
ATF6	Activating transcription factor 6
BP	β-propeller/EGF domain of LRP5/6
BP1-LRP6	β-propeller/EGF domain I of LDL receptor-related protein 6
BiP/Grp78	Binding immunoglobulin Protein/glucose regulated protein 78
BSE	Bovine spongiform encephalopathies
CamK2	Calcium-calmodulin kinase 2
CD	Circular dichroism
CJD	Creutzfeld Jakob disease
CK1	Casein kinase 1
CNS	Central nervous system
COP	Cytosolic coat protein
CPP	Cell-penetrating peptides
CRD	Cysteine-rich domain
Dab1	Disabled-1
DKK1	Dickkopf homolog 1
DMEM	Dulbecco minimum essential medium
Dvl	Dishevelled

E.coli	*Escherichia coli*
ECL	Enhanced chemiluminescence
EDEM	ER degradation-enhancing 1,2-mannosidases-like protein
EGF	Epidermal growth factor
eIF2α	α-subunit of the translation initiation factor 2
ER	Endoplasmic reticulum
ERAD	ER-associated degradation
ERAF	The Endoplasmic reticulum-associated folding
FBS	Fetal bovine serum
FH	Familial hypercholesterolemia
Fz	Frizzled
Grp57	Glucose regulated protein 57
GT	Glycoprotein glucosyltransferase
GlucI	Glucosidase I
GSK3	Glycogen synthase kinase 3
HIC	Hydrophobic interaction chromatography
HRP	Horseradish peroxidase
HSV-1	Herpes simplex virus 1
Ire 1	Inositol requiring enzyme-1
IPTG	Isopropyl-β-D-thiogalactoside
JNK	c-JUN N-terminal kinase
LA	LDLR type A
LBD	Ligand binding domain

LEF	Lymophoid enhancer factor
LDL	Low-density lipoprotein
LDLR	Low-density lipoprotein receptor
LRP	Low density lipoprotein receptor-related protein
MESD	Mesoderm development
NMR	Nuclear magnetic resonance spectroscopy
O.D.	Optical density
OD_{600}	Optical density at 600nm
OST	Oligosaccharyltransferase complex
PAGE	Polyacrylamide gel electrophoresis
PBS	Phosphate buffer saline
PCP	Planar cell polarity
PDI	Protein disulfide isomerase
PEI	Polyethylenimine
PERK	Protein kinase R like *ER* protein kinase
pI	Isoelectric points
PKC	Protein kinase C
PLC	Phospholipase C
PMP22	Peopheral myelin protein 22
PPI	Peptidyl-prolyl *cis-trans* isomerases
PTMs	Post-translational modifications
QC	Quality control
RAP	LDL receptor-associated protein

ROCK	Rho-associated kinase
SDS	Sodium dodecyl sulfate
SEC	Size exclusion chromatography
TCF	T cell factor
TSEs	Transmissible spongiform encephalopathies
UPR	The unfolded protein response
UGGT1	UDP-glucose:glycoprotein-glucosyltransferase
VLDLR	Very low density lipoprotein receptor
2 x YT	2 times yeast extract tryptone
YWTD	Tyr-Trp-Thr-Asp
Wg	Wnt/Wingless

CHAPTER I

Review of the Literature

1.1 Protein folding

Proteins are the most abundant biological macromolecules within a cell and carry out vital functions. The sequence of amino acids in a protein defines its primary structure. However, the linear chain of amino acids only becomes a functional protein when it folds into its three-dimensional form.

1.1.1 Protein sequence and protein structures

All proteins are polymers of amino acids (*8*). Each protein polymer – also known as a polypeptide – consists of a sequence of 20 different L-α-amino acids, also referred to as residues (*9*). To be able to perform their biological function(s), proteins fold into one or more specific spatial conformations, driven by a number of non-covalent interactions such as hydrogen bonding, ionic interactions, Van der Waals forces, and hydrophobic packing (*10, 11*).

Proteins have four structural levels. The primary structure refers to the amino acid sequence of the polypeptide chain, which is held together by covalent bonds (peptide bonds). The secondary structure refers to highly regular local sub-structures. Two main types of secondary structures are the α-helix and the β-sheet. These secondary structures have a regular geometry, being constrained to specific values of the dihedral angles ϕ and ψ on the Ramachandran plot. The tertiary structure refers to three-dimensional structure of a single protein

molecule. The α-helices and β-sheets are folded into a compact globular structure driven by the non-specific hydrophobic interactions and is stable only when the parts of a protein domain are locked into place by specific tertiary interactions, such as salt bridges, hydrogen bonds, and the tight packing of side chains. Quaternary structure is a larger assembly of several protein molecules (subunits). These quaternary protein subunit assemblies are stabilized by the same non-covalent interactions as the tertiary structure.

1.1.2 Protein folding

Protein folding occurs through an intermediate form, known as secondary structure. The most common secondary structure elements are the rod-like α-helix and the plate-like β-pleated sheet. They are stabilized by non-covalent interactions. Then, in turn, these secondary structural elements further interact, fold, and coil to produce the tertiary structure that contains functional regions (domains).

Although it is possible to deduce the primary structure of a protein from a gene sequence, the tertiary structure or the quaternary structure of a protein cannot be predicted just by its primary amino acid sequence. Meanwhile, inside the cell, the folding rate is significantly high, in a second time-scale (*12*). Therefore, to understand how a protein sequence folds into its functional structure inside of the cell becomes one of the most important tasks for biochemists.

One of most popular models for protein folding is the hydrophobic collapse model. This model proposes that, since the hydrophobic effect is the major driving force for protein folding, formation of a hydrophobic residue cluster that repels water must be the first step of folding. The polypeptide chain tends to collapse to a compact state. Then the rate-limiting steps in folding can be the reorganization of inter-residue interactions within a more-or-less disordered collapsed state (*13*). This process involves complex kinetics. Some molecules may form in the initial stages of folding between contacting residues, which allows rapid formation of native structure. Others may generate folding intermediates, which eventually will be reorganized to form the native conformation.

1.1.3. Protein misfolding diseases

Protein folding *in vivo* is of remarkably high fidelity, but sometimes defects can occur resulting in protein misfolding. In the 1990s, scientists started to understand that protein misfolding is involved in the development of many diseases (Table 1-1). Protein folding diseases can be divided into two groups. In the first group, excessive quantities of misfolded proteins aggregate forming amlyoid plaques, a hallmark of amyloidoses, such as Alzheimer's and Creutzfeldt–Jakob diseases (*14*). In the other group, genetic mutations lead to incomplete folding of a protein, which affects its function. This might, for instance, happen to the tumor suppressor p53, the loss of function of which results in cancer (*14*).

Table 1-1: Representative protein folding diseases (*15*)

Disease	Protein	Site of folding
Hypercholesterolaemia	Low-density lipoprotein receptor	ER
Cystic fibrosis	Cystic fibrosis transmembrane regulator	ER
Phenylketonuria	Phenylalanine hydroxylase	Cytosol
Huntington's disease	Huntingtin	Cytosol
Marfan syndrome	Fibrillin	ER
Osteogenesis imperfecta	Procollagen	ER
Sickle cell anaemia	Haemoglobin	Cytosol
α1-Antitrypsin deficiency	α1-Antitrypsin	ER
Scurvy	Collagen	ER
Alzheimer's disease	Amyloid α-peptide/tau	ER
Parkinson's disease	α-synuclein	Cytosol
Scrapie/Creutzfeld-Jakob disease	Prion protein	ER
Familial amyloidoses	Transthyretin/Lysozyme	ER
Retinitis pigmentosa	Rhodopsin	ER
Cataracts	Crystallins	Cytosol
Cancer	p53	Cytosol

The common characteristic of all amyloidoses is the collection of plaques (amyloids) of insoluble protein in the extracellular tissue, which cannot be degraded by proteases. Their structures are highly ordered and give them crystal-like properties. Beta-pleated sheet of identical proteins are densely packed and form long filaments (fibrils). About 20 different proteins can associate with different diseases with unique features (*16*). The plaques can be transported through the bloodstream to different tissues forming amyloid deposites, systemically. However, the localized amyloids are more important for clinical research because they mainly affect the central nervus system, causing neurodegenerative diseases.

The general mechanism of amyloid formation reveals that polypeptide chains can adopt multiple conformational states and interconvert between them on a wide range of time scales. The network of equilibria, which link several of the most important states, is schematically illustrated in Figure 1-1. Through partially folded intermediates, the protein can compact into a monomer or associate to form oligomers or higher aggregates. Oligomers may be functional, such as actin, myosin and microtubules. The majority of these proteins will be degraded, usually under carefully controlled conditions and as a part of normal biochemical processes, with their amino acids often being recycled. However, partially folded or unfolded intermediates may enable aggregation (Figure1-1). Some of these initial aggregates simply dissociate, while others may reorganize to form stable oligomers, initiating amyloid structures (*15*).

Transmissible spongiform encephalopathies (TSEs), which include mad cow disease (bovine spongiform encephalopathy, BSE) and Creutzfeld Jakob disease (CJD) in humans, are special forms of amyloidoses in which the victim's brain degenerates to a structure that looks like a porous sponge. These conditions seem to occur when normal protein particles called prions misfold. The normal prion is also called PrP^{c}, a component of the membrane of healthy nerve cells. When it is properly folded, it remains soluble and degradable. However, when it is misfolded in a particular way, it can take on an infectious, incorrectly folded three-dimensional form (PrP^{sc}) (*17, 18*). The infectious prion can be transmitted in the diet and trigger a domino effect in healthy prion, forcing it to adopt its incorrectly folded form.

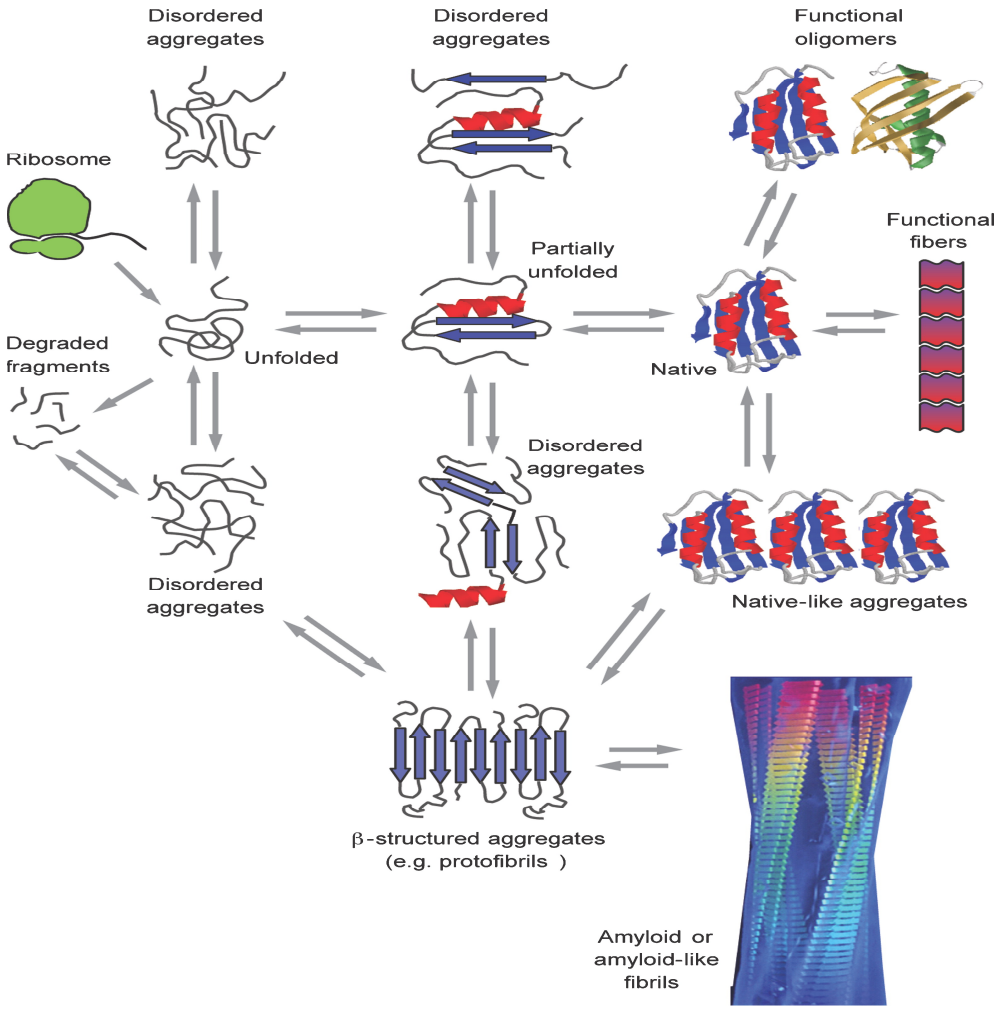

Figure 1-1: A schematic representation of the protein conformational states that are in interconversion. The transition from β-structured aggregates to amyloid fibrils can occur by addition of either monomers or protofibrils (depending on protein) to preformed β-aggregates. Many of the various states of proteins are utilized functionally by biology, including unfolded proteins and amyloid fibrils, but conformational diseases will occur when such regulatory systems fail, just as metabolic diseases occur when the regulation of chemical processes becomes impaired (*6*).

Another group of protein folding diseases is caused by lack of a correctly folded protein. Protein p53 occupies an important position in the cancer resistance network. Normally, the p53 system is switched off or in stand-by mode. But when the cells become stressed or damaged, p53 is activated to prevent genetic mutations, which can cause tumor formation marked by uncontrolled division and proliferation of cells. Mutations of p53 gene lead to misfolding of p53 and allow abnormal cell growth in an uncontrolled manner (*19*). This type of mutation in p53 is thought to occur in 50% of all cases of cancer and as many as 95% of all cases of lung cancer (*20*).

1.1.4. The endoplasmic reticulum-associated folding (ERAF) and the quality control (QC) system inside mammalian cells

The endoplasmic reticulum (*ER)* is a large, membranous organelle that is in physical contact with almost all other organelles inside the cells, including the nucleus, Golgi and plasma membrane (*21, 22*). In the ER, once the primary sequence of a nascent protein has been assembled and released from the ER-associated ribosome through the Sec61 channel, it has to undergo posttranslational modification to become fully functional. The main posttranslational modifications are the formation of disulfide bonds, conformation folding, and protein glycosylation. While folding or assembling, intermediates expose their hydrophobic surfaces, unpaired cysteines or immature glycans. Very high concentrations of chaperone proteins in the lumen of the ER, such as oxidoreductases, interact with the intermediates and promote their folding. These

chaperone proteins work in concert and form complexes, which efficiently process proteins and ensure their proper folding.

1.1.4.1. ER chaperones and folding enzymes (foldases)

About 333 proteins of the ER protein folding machinery are known. These include the lectins, N-glycan processing enzymes, protein disulfide isomerases, prolyl *cis-trans* isomerases, and the molecular chaperones (*23*). Table 1-2 gives a brief list of these proteins. ER-resident foldases such as lectins (calnexin and calreticulin) and the general ER-chaperone BiP play key roles in regulating ER-associated folding (ERAF) and degradation (ERAD) of proteins and will be discussed in detail below.

The canonical ER lectins are the type 1 membrane protein calnexin and the closely related ER luminal protein calreticulin. They are two prominent ER chaperones, calcium requiring and involved in processing Glc1Man9 N-glycans, which are attached to nitrogen (N) residues on nascent proteins. The N- glycans are generated by the action of glucosidase II on the Glc2Man9 glycan after processing of Glc3Man9 by glucosidase I, or by UGGT1 mediated addition of glucose moiety (Figure 1-2). They serve a wide array of functions including stabilizing proteins against denaturation, enhancing solubility, orienting proteins within membranes, adding structural rigidity, regulating protein turnover, and mediating pathogen interactions (*24*).

Table 1-2: Principal ER chaperones and foldases (*12*).

Family	Main members	Functions
HSP70	Bip/GRP78	Chaperone
HSP90	Endoplasmin/GRP94	Chaperone
HSP40	ERdj1-ERdj7	Co-chaperone
Lectin	Calnexin, Calreticulin	Glycoprotein quality control
	EDEMs 1,2,3	Glycoprotein degradation
	OS-9, XTPB-3	Glycoprotein degradation
Glycan processing enzymes	UGGT	Folding sensor
	ER glucosidases I, II	Trims off glucose from N-glycan
	ER mannosidases	Removes terminal mannoses
Peptidyl-prolyl isomerases	Cyclophilin B, FKBP13/23/65	Isomerize peptide bonds
Ero1	Ero1α, Ero1β	Oxidation for disulfide bonds
Oxidoreductases	PDI, ERp72, ERp57, plus many others	Disulfide bond formation and isomerization

BiP is an abundant ER chaperone. BiP contains an N-terminal ATPase domain, stimulated by ER resident J-domain co-chaperones, of which there are seven (ERdj1-7) (*25*). When BiP binds to ATP, it adopts an open conformation. ATP hydrolysis causes BiP to adopt a conformation with higher affinity for substrate binding. The ADP-bound closed state of BiP can be re-opened, with the help of the nucleotide exchange factors GRP170 and Sil1. BiP is usually the first chaperone to bind the nascent chain to facilitate translocation of the nascent chain into the *ER* lumen (Figure 1-2). ERdj1 and 2 regulate the nascent chain binding. BiP is also found in a multi-protein complex that includes several protein disulfide isomerases and peptidyl-prolyl *cis-trans* isomerases such as protein

disulfide isomerase (PDI) and peptidyl-prolyl *cis-trans* isomerases (PPIs), respectively (*26*). The association of chaperones with foldases in protein complexes was proposed to direct their activities towards the nascent polypeptide chain (*27*).

PDI is the primary oxidant of cysteine thiols in the ER, and one of the most abundant *ER* proteins. The expression of PDI is rapidly induced under different *ER* stresses (*28*), indicating a crucial role for PDI in protein folding quality control. Under different microenvironments of reduction potential, it can make (oxidize), break (reduce), and re-arrange (isomerize) disulfide bonds.

PPIs, also called rotamases, catalyze the slow isomerization of Xaa-proline peptide bonds that are found in both *cis* and *trans* in folded protein. Proline isomerization has been identified as a rate-limiting step (*29*). Correction of Xaa-proline bond orientation happens when a protein has reached an almost completely folded conformation (*30, 31*).

1.1.4.2. The ER folding machinery

Proteins destined for the secretary pathway and eventually the plasma membrane, have an extra intrinsic signal sequence on the N-terminus that is exposed as the protein begins to emerge from the ribosome. The signal sequence is first recognized and bound to a signal recognition particle (SRP), such as SRP54. SRP binds to translating ribosome and the emerging signal sequence with high affinity (*32*). Once the SRP binds to the ribosome and

recognizes the signal sequence, the complex is then transported to the ER membrane where it binds to the SRP receptor (SR) (*33*). The SRP/SR complex positions the ribosome over a translocon and the protein being translated enters the lumen of the ER as it is being translated (*32, 33*). The translocon is a complex that is composed of the microtubule tethering protein Climp63, the oligosaccharyltransferase complex (OST), and glucosidase I (GlucI) (*34*).

Upon entering the ER lumen, proteins without N-X-S/T consensus, the N-glycosylation site, directly undergo the folding process facilitated by BiP, PDI, and PPI. However, most secreted proteins are first tagged with N-glycan (Glc3Man9GlcNAc2) at N-X-S/T consensus. It has been proposed that the glycosyl moieties act as a sophisticated quality control tag for protein folding inside the *ER*. The enzymes glucosidases I and II cleave terminal glucose (Glc) residues from newly synthesized N-Glycan. The Glc1Man9 N-glycans are recognized by the chaperones calnexin and calreticulin. Calnexin and calreticulin with the bound protein disulfide isomerases ERp57 establishes a calnexin cycle to accelerate the folding of Glc1Man9 N-glycan protein (*35*). When the glucose residue is cleaved, calnexin and calreticulin are released from the protein. If the protein is folded into its proper conformation upon release of the lectins, it is sorted into COPII vesicles by cargo recruiting lectins ERGIC53, VIP36, and VIPL that reside in the *ER* to Golgi intermediate compartment, called ERGIC (Figure 1-2).

Additional candidate protein has been proposed for cargo sorting. One such candidate is the p24 family of type I *ER* membrane proteins that is highly

conserved in eukaryotes (*36*). They are abundant constituents of the *ER* and COPII vesicles and their C-termini binds to COPII coat proteins (SEC23).

1.1.4.3. ER quality control system

The integrated process of protein folding and targeted protein degradation are known as the ER quality control (QC) system (*37*). In this system, all proteins are subjected to a "primary" quality control check that monitors their conformation. A secondary quality control mechanism relies on the intrinsic signal of proteins and facilitates their transport, such as the process to send misfolded proteins for *ER*-associated degradation (ERAD). The unfolded proteins are retained in the *ER*. In order to promote a renewed cycle of folding, unfolded glycoprotein can be recognized by UDP-glucose:glycoprotein-glucosyltransferase (UGGT1) in the calnexin cycle. UGGT1 adds one glucose moiety from UDP-glucose to the unfolded proteins, which bear high content of mannose (Man9) in their N-glycans. By rebinding with calnexin and calreticulin, the misfolded proteins re-enter the calnexin cycle and undergo further folding catalyzed by ERp57 (*38*) (Figure 1-2).

Despite the presence of chaperones, some proteins persistently fail to be folded and processed properly, which can lead to the formation of non-functioning protein aggregates. *A* 1, 2-mannosidase I will cleave a mannose residue (Man) from the oligosaccharides of these proteins (*39*). This leads to the recognition by the *ER* degradation-enhancing 1, 2-mannosidases-like protein (EDEM). Then the glycoproteins are targeted for degradation by ER-associated

degradation (ERAD) or the ubiquitin- proteasome system through the disposal pathway (*40*) (*37*).

1.1.4.4. *ER* associated degradation (ERAD)

ERAD is a process by which misfolded *ER* proteins are detected and prevented from progressing along the secretary pathway by *ER*-resident factors and directed to the translocon for retrotranslocation (or dislocation) into the cytosol, where they undergo ubiquitin- and proteasome-dependent degradation (*41*). Acting as “mannose timer”, EDEM (*ER* degradation-enhancing α-mannosidase-like protein) is responsible for removing the glycoproteins, which could not be productively folded, from the calnexin cycle. Futile recycling of terminally misfolded proteins is prevented by EDEM, most likely EDEM3, which successively cleaves mannose moieties from the glycan, eventually removing the mannose that is normally re-glucosylated by UGGT. In complex with the protein disulfide isomerase ERdj5 and BiP (*42*), EDEM produces Man7 N-glycan with a terminal α 1,6-linked mannose, providing the targeting signal for ERAD (*43*). The formation of the Man7 N-glycan is a key step in directing glycoprotein substrates for ERAD. The lectin OS9 and the related XTP3-B recognize the Man7 N-glycan and bring them to form complex with ubiquitin protein ligase, the Hrd1 E3/SEL1L protein, and with the *ER* HSP90 homolog GRP94 and BiP, which help sequester the misfolded protein away from other interactions until retrotranslocation (*44*). Eventually, the misfolded proteins are retrotranslocated, ubiquitinated, and degraded by the proteasome.

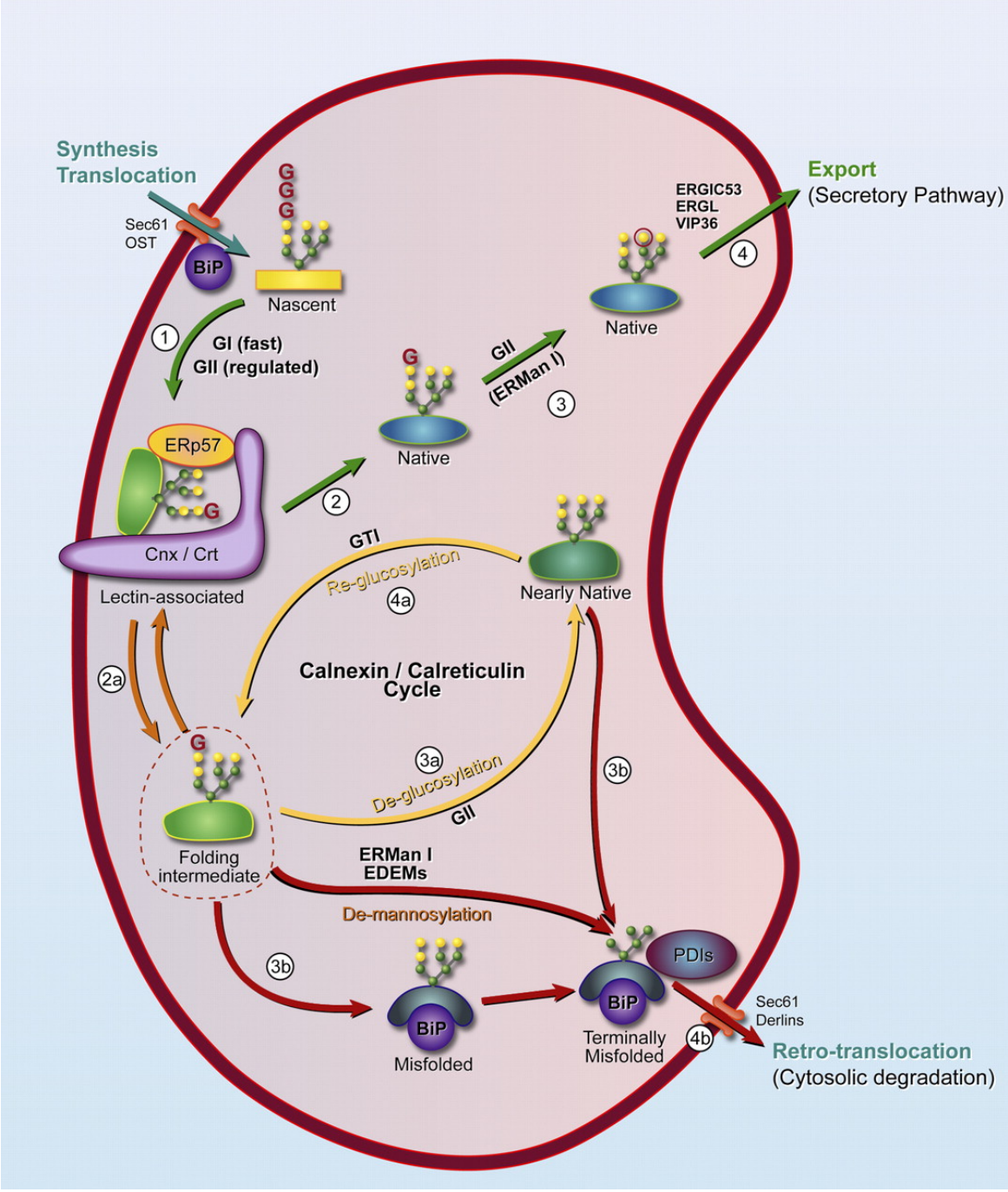

Figure 1-2: ER quality control of newly synthesized glycoproteins (*45*).

Nascent chains enter the ER lumen through the Sec61 complex and are core glycosylated by the oligosaccharyltransferase (OST). The two terminal glucose residues are rapidly trimmed by sequential action of the glucosidase I and II (step 1). Mono-glucosylated N-glycans mediate the initial association of the folding polypeptides with the ER lectin chaperones calnexin and/or calreticulin, resulting in the exposure to the glycoprotein-dedicated oxidoreductase ERp57. It is likely that most glycopolypeptides are released from calnexin/calreticulin/ERp57 in a native, transport competent state (step 2). They are rapidly deglucosylated and partially demannosylated (step 3) and eventually sequestered in transport vesicles for export from the ER (step 4). For a fraction of newly synthesized glycoproteins, folding is not completed in a single round of association with calnexin/calreticulin (step 2a). The folding intermediate released from the lectin chaperones is deglucosylated (step 3a), but its forward transport is inhibited by GT1. GT1 adds back a glucose residue (step 4a) only to glycoproteins with nearly native conformation. This second chance of binding to calnexin/calreticulin facilitates additional folding attempts which likely consists disulfide reshuffling. Glycopolypeptides released from calnexin and displaying major folding defects are ignored by GT1 (step 3b). Rather, they attract BiP, which are extensively demannosylated and dislocated across the ER membrane for proteasome-mediated degradation (step 4b).

1.1.5. The unfolded protein response (UPR)

Accumulation of misfolded proteins in the ER triggers the unfolded protein response (UPR) that signals compensatory mechanisms to keep *ER*

homeostasis, including up-regulation of ER chaperones and down-regulation of protein translation (*46, 47*) (Figure 1-3).

The UPR is a highly conserved intracellular pathway that was originally discovered in yeast. The proximal sensor of *ER* stress in yeast is the inositol requiring enzyme-1 (Ire1). Ire 1 contains a unique endoribonuclease activity in its cytosolic domain, which carries out the unconventional splicing of Hac1 mRNA (*48*). Hac1 mRNA is found constitutively in the cytosol, where its native nucleotide sequence is inefficiently recognized and translated. In response to ER stress, this transmembrane protein is dimerized and *trans*-autophosphorylated to be activated (*49*). The activated endoribonuclease on Ire1 splices out a specific intron causing a frame-shift in the Hac1 nucleotide sequence. The new nucleotide sequence is recognized and efficiently translated into the transcription factor Hac1p (*48*). Hac1p subsequently enters the nucleus and induces the genes which aid in folding and degradation of unfolded/misfolded proteins.

Although there are α- and β-alleles of both IRE1 and ATF6 in the mammalian genome, only IRE1 α is expressed in all tissues and only ATF6 α signals the UPR. IRE1 α is selectively expressed in intestinal epithelial cells and it is not known which genes are regulated by ATF6 α. IRE1 activation elicits an endoribonuclease function that induces non-conventional splicing of *Xbp1* mRNA. Splicing of *Xbp1* mRNA, the only known splicing substrate of IRE1, removes a 26-base intron that alters the translation reading frame to produce a highly active bZiP transcription factor that activates genes encoding ER protein chaperones, lipid biosynthetic enzymes, and ERAD functions (*50-52*). Upon

release from BiP, ATF6 traffics to the Golgi complex where it is cleaved by the S1P and S2P processing enzymes to produce a cytosolic fragment that activates transcription of genes providing complementary and overlapping functions with those activated by XBP1 which restores productive ER protein folding and increases ERAD. Indeed, cells deleted in either *Ire1 α*, *Xbp1 or Atf6 α* are defective in ERAD (*53, 54*).

The UPR plays an important role in functions to regulate a cell's folding capability in response to developmental demands or physiological changes. Especially, in cell types with a high secretary load or an *ER* compartment that is uniquely susceptible to stress, the UPR becomes vital for cell survival and function (*50, 55*). In response to the ER stress, UPR proteins act in a coordinated fashion to initially decrease general protein translation and subsequently increase transcription of genes that encode proteins that aid in both protein folding and degradation. Therefore, pathological conditions interfering with *ER* homeostasis and resulting in prolonged activation of the UPR could lead to various diseases (*56, 57*).

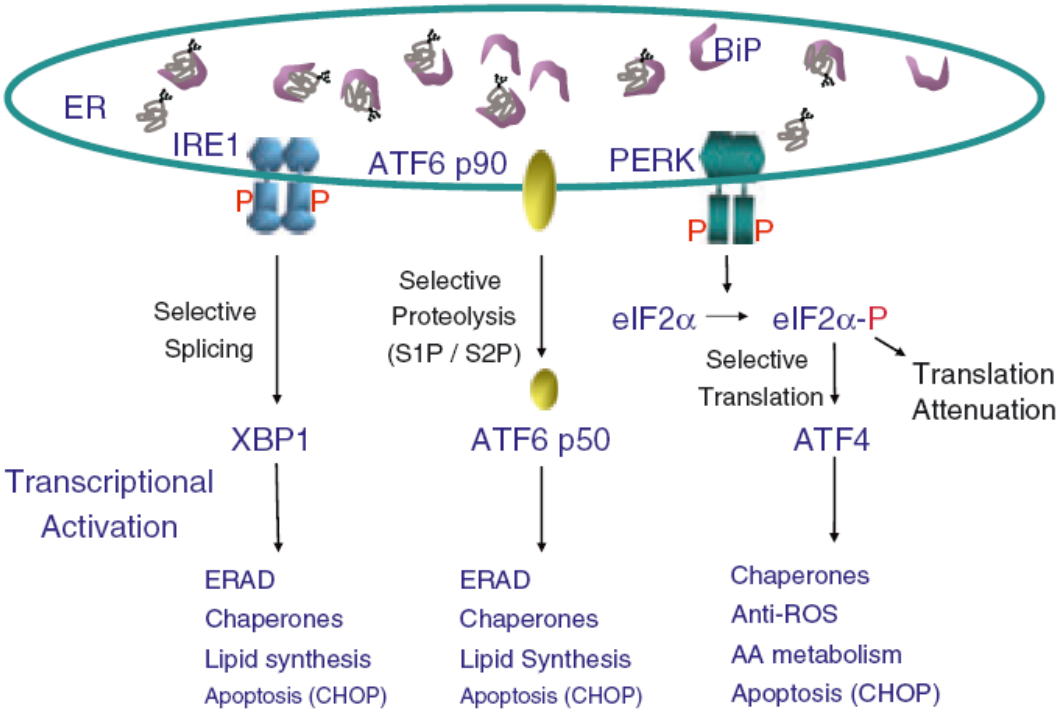

Figure 1-3: Signaling the unfolded protein response (UPR) (*58*).

The UPR in mammalian cells is far more complex than that in yeast and appears to act both at the level of transcriptional as well as translational regulation (*3*). Mammalian cells have three *ER* membrane bound proximal sensors. They are the mammalian homologue of yeast IRE1, protein kinase R like *ER* protein kinase (PERK), and activating transcription factor 6 (ATF6) (Figure 1-3) (*4, 5*). The three UPR sensors are maintained in an inactive state through interaction with the protein chaperone BiP. It is proposed that as unfolded proteins accumulate, bind, and sequester BiP, they promote BiP dissociation from PERK, IRE1, and ATF6 (*1, 7*). BiP release from IRE1 and PERK permits their homodimerization, trans-autophosphorylation, and activation. Activated PERK phosphorylates the α-subunit

of the translation initiation factor 2 (eIF2α), leading to rapid and transient inhibition of protein synthesis (*1*). eIF2 is a heterotrimeric GTPase required to bring initiator methionyl tRNA to the ribosome for AUG initiation codon selection (*2*).

1.2. Protein expression and refolding technologies

Biomedical and structural biology studies of proteins require large quantities of purified proteins. Over the past several decades, protein expression systems (*Escherichia coli,* baculovirus-mediated insect cell, *Saccharomyces cerevisiae* and several mammalian based systems) have been developed for large-scale recombinant protein expression. Each of these systems has its strength and weakness (yield, proper folding, post-translational modifications, etc).

1.2.1. Prokaryotic expression system

Escherichia coli is the simplest and by far the most widely used organism for protein expression. Our lab also developed a very high yield bacterial protein expression system that produce nearly gram quantity of pure recombinant proteins from one liter of bacterial culture (*59*). Many recombinant proteins have been successfully produced in bacteria; however, reports note that ectopically expressed proteins especially those that are cystein-rich, have basic isoelectric points (pI), and that are less than 10 kDa or greater than 50 kDa have overall reduced success rates in traditional purifaction schemes (*60*). Other factors such

as improper folding, protein aggregation (inclusion body formation), and lack of proper eukaryotic post-translational modifications (PTMs) of the ectopically expressed proteins could potentially affect their conformation, stability, and function. All of these possible limiting factors reduce the usefullness of bacterial protein expression systems, especially for eukaryotic proteins where the success rates are typically lower (*61-64*).

1.2.2. Eukaryotic expression system

The yeast strain, *Saccharomyces cerevisiae* is an expression system that has the advantage that can produce proteins in a more native eukarytoic environment but with reduced yield, a common drawback of eukarytoic expression systems. Another drawback is the *S. cerevisiae* glycosylation pattern, which is very different from its mammalian counterpart, and commonly involves hypermannosilation that potentially affects the proper folding and activity of heterologously expressed mammalian proteins.

Insect cell is another common recombinant protein production system. Appreciated advantages of insect cells are the robust and relatively inexpensive cell culture and the fact that most eukaryotic PTMs are executed properly (*65*). However, the baculovirus expression system requires multi-step process to make the viruses and the need to maintain high virus titers which sometimes can be quite challenging (*66*). Furthermore, the protein expression yield of this insect cell system is low.

For expressing physiological proteins, including therapeutic proteins, mammalian cell lines, such as CHO or HEK293, are widely used because they offer a cellular environment closer to the native condition. However, the yield of this system is quite low. The expression levels of purified recombinant protein are usually in the microgram quantity (*67*) and expression cost of mammalian cell protein expression is quite high. This expression system weighs the fidelity of PTMs, correct folding and processing in mammalian cells against their slow doubling rates, potentially inefficient transfection, overall lower productivity and their dependence on expensive reagents, such as serum (*68*).

1.2.3. Protein refolding *in vitro*

Protein studies and protein therapeutics need large quantities of properly folded and biologically functional proteins. Genetic engineering allows rapid and reliable production of heterologous proteins via recombinant methods, in particular, bacterial expression methods. These methods are used to induce cells to overexpress a target protein. Overexpressed proteins, however, often form “inclusion bodies”, the macroscopic aggregates in cells, especially for the prokaryotic expression system. Therefore, the ultimate success of overexpression strategies sometimes depends on the refolding rate of properly folded native states of aggregated proteins in the bacterial inclusion body. Therefore, protein refolding *in vitro* is important.

The first milestone to explore protein refolding *in vitro* emerged when Christian Anfinsen and his colleagues in the 1960s demonstrated that simply removing denaturant is sufficient for a fully denatured protein to refold (*69*). This work demonstrated that all the driving forces and guiding information are embedded in the primary amino acid sequence of a protein and no external template is required. Also, this study showed, just like other processes in nature, protein folding needs energy, a process following the universal thermodynamics law. Intrinsic protein refolding steps include secondary structure formation, long-range tertiary contacts, interaction of hydrophobic groups, interaction of hydrogen bond donor/acceptor, side-chain packing, and other processes. Conformational entropy and several major interactions, such as the hydrophobic effect, hydrogen bonds, electrostatic interactions, and van del Waals forces, contribute to protein refolding.

Conformational entropy is the major energy term opposing protein folding because entropy is a measure of the degree of randomness or disorder in a system. As the folding reaction proceeds, the conformational entropy eventually decreases as the unfolded state of protein has the highest conformational entropy, whereas the folded protein has the lowest conformational entropy. The loss of entropy starts to be compensated by the favorable hydrophobic and other interactions.

Based on the theory, a protein always folds into the conformations, which achieve the lowest possible energy. An energy landscape of protein folding (Figure 1-4) provides a statistical description of the folding process. It proposes

that proteins fold in an ordered way and the whole structure ensemble process is an energy-biased folding pathway (*70-72*). However, this funneled energy landscape theory fails to provide clues as to the guidance to the folding reaction happening inside living cells. Even though the folding reaction can proceed through a multitude of distinct routes to reach the lowest energy native state as described in the energy landscape theory (*73-77*), proteins do not find their native conformation from an endless number of potential three-dimensional probabilities that it could randomly fold into. The protein folding process in living organisms occurs in a second time scale (*78*). Localized interactions with molecular chaperones and enzymes inside of cells guide and speed up the folding process.

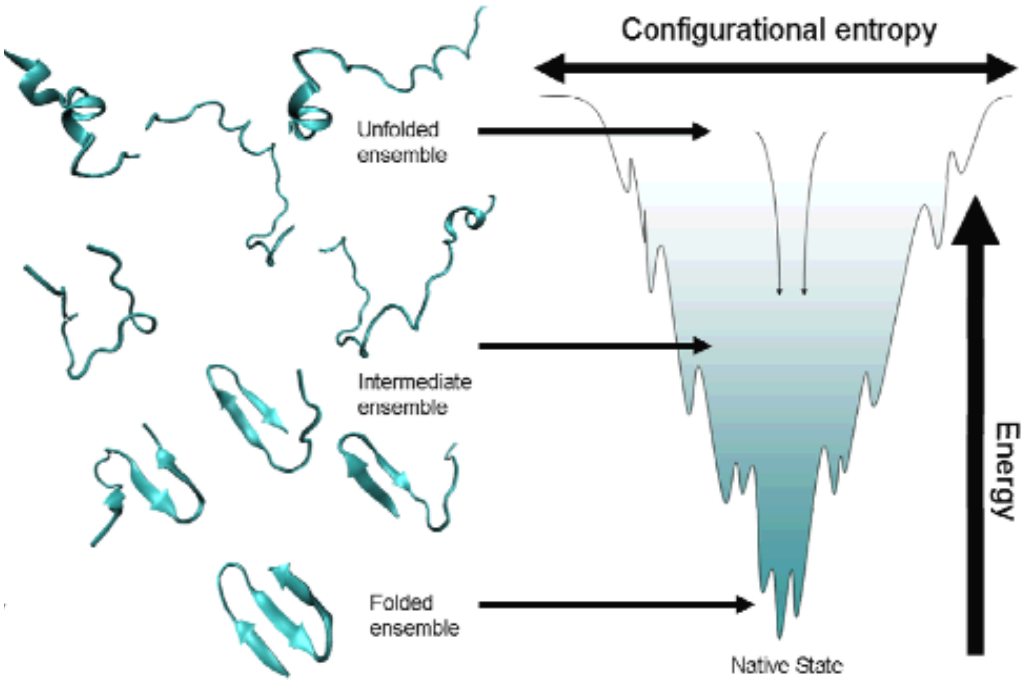

Figure 1-4: The energy landscape depiction of protein folding (*72*) .

1.2.4. *In vitro* protein refolding techniques

Protein misfolding frequently happens when expressing recombinant proteins in a foreign cell as well as inside mammalian cells under certain physiological conditions. Experimental approaches have been developed to refold aggregated proteins into their biologically active forms. The following briefly summarizes these protein refolding techniques:

1.2.4.1. Direct dilution

The simplist of these procedures is the direct dilution method. Diluting the protein reduces intermolecular interactions, denaturants, etc. that could contribute to protein aggregation. However, a major drawback of this technique is that the protein is highly diluted and needs to be concentrated. Concentrating may contribute to the reaggregation of the protein.

1.2.4.2. Dialysis

Similar to the direct dilution technique, dialysis can remove potential denaturants allowing the unfolded protein to achieve its native conformation. In contrast to the direct dilution method, the change from denaturing to native buffer conditions occurs gradually allowing the protein to slowly reach its native state. However, through this process the unfolded protein may pass through several different folding intermediates, some of which are unproductive and even could

protentiate protein-protein interactions promoting aggregation (*79*). As the protein concentration remains relatively stable during this procedure, a potential benefit; however, may contribute to promoting the aggregation and thus making protein aggregation a frequent problem of dialysis as compared to the direct dilution method. Additionally, refolding yields can be negatively affected by non-specific adsorption of protein to the membrane. However, for some proteins and with the appropriate denaturant removal rates, adapted to the requirements of the target protein, high refolding yields at high protein concentrations can be obtained (*80*).

1.2.4.3. Chromatographic methods for protein refolding

1.2.4.3.1. Protein refolding based on size exclusion chromatography

Similar to both the direct dilution method and dialysis, buffer exchange for denaturant removal can also be carried out using size exclusion chromatography (SEC). SEC has several major advantages. The denaturant can be gradually removed through use of a concentration gradiant allowing the denatured protein to refold. As with dialysis, protein aggregation can be a problem. However, the power of SEC allows enrichment of the properly folded protein over the mis-folded/aggregated intermediates by taking advantage of the mobility differences based on size, shape, etc. of the target versus the folding intermediates. Additionally, SEC also allows concomitent concentration and purification of the folded protein. However, reduced yield of the folded protein, due to aggregation,

could occur. Also, the success of SEC depends on the properties of the chromatographic resin to allow efficient separation of the renatured target protein from different folding intermediates, misfolded protein, and aggregates that might form during the refolding process (*81*).

1.2.4.3.2. Assisted protein refolding using affinity chromotography

In this technique, the unfolded protein is attached to a solid support (matrix material of the column) through a variety of methods such as 'tagging' the protein (eg. Histidine tag) or using ion-exchange or hydrophobic chromotography based methods if the biochemical properties of the protein are known. Proteins containing artificially engineered peptide tags such as the hexahistidine tag can bind to immobilized metal ions. After binding, the matrix-protein complex is brought to refolding conditions by replacing the denaturing buffer condition with a native one. Finally, the refolded protein can be detached from the matrix, *e.g*. in the case hexahistidine-tagged proteins by elution with EDTA or imidazole or by buffers with high ionic strength in the case of proteins bound by ionic interactions (*82*). Due to the selective binding, matrix-assisted refolding can combine the renaturation of the target protein along with its purification (*83*). However, copurification of the aggregated/mis-folded proteins side-by-side with the refolded protein could occur and may become representative. If predominant, these aggregated/misfolded proteins could interfere in strutural studies and thus need to be removed and the purification can be accomplished with SEC.

1.2.4.3.3. Refolding using hydrophobic interaction chromatography

Hydrophobic interaction chromatography (HIC) has also been successfully used for protein refolding with concomitant removal of contaminating proteins during the renaturation process (*84*). Unfolded proteins are applied to the column at high salt concentrations and refolded and eluted with a decreasing salt gradient. It has been proposed that refolding is facilitated during HIC because unfolded proteins adsorb at high salt concentrations to the hydrophobic matrix and, thus, are not prone to aggregation. Additionally, hydrophobic regions of the protein that adsorb to the HIC matrix form micro domains around which native structure elements can form. During migration through the column, the protein will pass through several steps of adsorption and desorption, controlled by the salt concentration and hydrophobicity of the intermediate(s), resulting finally in the formation of the native structure (*85*).

1.2.4.4. Physical conditions and chemical additives that influence protein folding

Several physical variables, such as temperature, pressure, and solution phases, can impact protein folding. Temperature has general effects on the rate of protein folding. In general, higher temperatures increasing the rate of folding but also the propensity towards aggregation whereas lower temperatures reduce the folding rate and can increase the amount of properly folded proteins (*86*). For example, growing bacteria at temperatures slightly lower than 37℃ during

the protein induction phase can reduce inclusion body formation (and hence aggregation) for some proteins (*87*).

Another physical property affecting protein folding is pressure. Protein aggregation can actually be reversed by increasing pressure up to 3 kbar. However, above 5 kbar the benefit may be lost as a loss in secondary structure may occur (*88*). Slowly reducing the pressure back to ambient condition may aid the folding process (*89, 90*).

At some point, *in vivo*, some proteins, such as integral membrane proteins, come into direct contact with or are transported through lipid membranes (*91*). Detergents and phospholipids, in the form of micelles and liposomes can aid protein refolding *in vitro* by mimicking the effect of a bilayered membrane. Such methods can solubilize purified membrane proteins enabling their proper folding. Interestingly, in an alternative approach, denatured hydrophilic proteins can be forced to fold as independent units in a reverse micelle-type system. Once transferred to this solution, the protein tries to avoid the organic phase. After reaching the hydrophilic core of the reversed micelle, proteins can refold as a single molecule (*92*). Both the micelle and reverse micelle approaches aid protein folding of integral membrane and hydrophlic proteins, respectively, mainly be preventing unwanted intermolecular protein-protein interactions that can lead to aggregation.

1.2.4.5. *In vitro* approaches to fold proteins containing disulfide linkages

Folding proteins that harbor disulfide bonds require special considerations. Reducing agents such as dithiothreitol or β-mercaptoethanol must be added in order to break the incorrect disulfide linkages. Slowly removing the reducing reagent, such as during dialysis or direct dilution, can enable the protein to sample several different disulfied linked states until the most stable one (native) is found. Although this method works well for proteins with few disulfide bounds, proteins with multiple cysteine residues require special considerations as incorrect stable interactions are more likely to occur. For example, the ligand-binding domain of apolipoprotein E receptor 2, which has 42 cysteine residues, is extremely difficult to refold *in vitro* (*93*). Similar to the *in vivo* environment, approaches utilizing protein chaperones to aid in the folding of multiple disulfide linked proteins has found some promise in folding these difficult proteins. Protein disulfide isomerase (PDI) is a folding catalyst that assists disulfide bond formation *in vivo* and has found success for aiding disulfide bond formation during *in vitro* protein folding. Although PDIs significantly increase the refolding rate, much of the protein may still not be properly folded (albeit much improved) resulting in reduced recovery of the properly folded protein. Additionally, residual concentrations of chaotropic agents in the refolding buffer, especially guanidinium hydrochloride, drastically reduce PDI activity and hence the protein yield (*94, 95*).

Although attempts have been made to mimic the *in vivo* environment to enable protein folding *in vitro*, the *in vivo* environment is still ideal.

Overexpressing proteins in bacteria and refolding the proteins *in vitro* may produce greater yields; however, some eukaryotic proteins cannot be refolded with this approach. Alternatively, in an eukaryotic environment, these proteins can be properly folded (aided by the eukaryotic folding machinery, 1.1.3.) but at a dramatically reduced yield. For this reason, our lab has created a novel protein delivery method that takes advantage of the high yield of the bacterial and the high fidelity of the eukaryotic expression systems to refold difficult proteins, such as the ligand-binding domain of apolipoprotein E receptor 2, to their native structure.

1.3. Protein delivery technology

Protein delivery techniques deliver biologically active proteins inside live cells for cell biology studies and therapeutical applications. The most popular means to deliver proteins into cells include electroporation, microinjection, the construction of viral fusion proteins, and the use of cationic liposome.

Electroporation was initially used to introduce foreign DNA into eukaryotic and bacterial cells (*96*). Scientists later found that proteins could also be introduced into cells using the same method (*97-99*). Cells are electroporated in a buffered solution harboring the purified protein of interest. High-voltage electric pulses result in the formation of small pores in the cell membrane. Proteins enter the cell via these small pores or during the process of membrane reorganization as the pores close and the cells return to their normal state. The efficiency of

delivery depends upon the strength of the applied electrical field, the length of the pulses, temperature, and the composition of the buffered medium. Although this method is successful with a variety of cell types, it has a few major drawbacks. The overall yield of the protein after repurification from the transduced cells is often negligible and the process can be damaging, causing cell death.

Another method used in the delievery of macromolecules (protein or nucleic acid) into live cells is microinjection. Microinjection was originally used to introduce DNA directly into the nucleus of a cell where it can integrate directly into the host cell genome, creating an established cell line bearing the gene of interest (*100*). Proteins, such as antibodies and mutant proteins, can also be directly delivered into cells via microinjection inorder to determine their cellular effects (*101, 102*). This method has the advantage of directly introducing proteins into specific cellular compartments; thereby, bypassing and preventing exposure of the delivered protein to extreme microenvironments found within various cellular compartments, such as the low-pH endosome. However, microinjection requires extensive training and specialized equipment, which is normally not available in most laboratories.

Several viral proteins or peptides have the ability to travel through cell membranes independent of classical receptor- or endocytosis-mediated pathways. Membrane fusion proteins, such as the HIV-1 TAT proteins, the herpes simplex virus 1 (HSV-1) DNA-binding protein VP22, and the *Drosophila* Antennapedia (Antp) homeotic transcription factor, have the ability to penetrate through the plasma membrane and into cells (*103*). The small protein

transduction domains (PTDs), also known as cell-penetrating peptides (CPPs), from these proteins can be fused to other proteins of interest to aid their delivery into cells (*103*). Sequence alignment of the PTD shows a high basic amino acid content (Lys and Arg), which may facilitate interaction of PTD with negatively charged lipids on the cell surface. The advantage of this method is that protein entry is rapid and works with many cell lines (*104*). However, the drawback to all of the PTD-mediated protein delivery systems is that the transduction domain must be covalently attached to the protein being delivered. Thus, it is necessary to design special vectors and limit the size of the cargo protein because of the relatively large PTDs. Also, denaturation or inactivation of proteins was observed after PTD fusion, especially for TAT. Reduced yield after re-purification of the PTD protein fusion can also be a major problem since degredation can occur if the PTD-fused protein is misfolded, thus significantly reducing the delivery efficiency. Furthermore, the PTD-based protein delivery method does not have a targeting capability and cannot specifically delivery a protein into the desired intracellular compartment, raising concern about the physiological relevance of the delivered proteins (*104*).

Liposomes are known for their ability as vehicles to deliver macromolecules into cells (*105, 106*). Cationic liposome can spontaneously and efficiently form complexes with negatively charged cargo, facilitating the delivery. This strategy has been successfully applied to protein delivery (*107*). But the physical characteristics of the cargo, such as the electric charges and hydrophobicity, can

influence the extent of the interaction between the cargo and the cationic liposome and ultimately impact the delivery efficiency.

Our lab recently developed an innovative protein delivery technology, based on the QQ-reagent, a polyethylenimine (PEI)–based cocktail, with a unique formulation of lipids and enhancers (Figure 1-5). It is compatible with most types of cell culture media and non-covalently binds the protein cargo of interest. This protein delivery technology is rapid and straightforward. The protein of interest can be produced by highly efficient *E. coli* expression systems and purified under denaturing conditions from inclusion bodies. The QQ-cocktails can efficiently dissolve the target proteins in delivery buffer that contains urea and modify them for protein delivery. Another important advantage of QQ-protein delivery technique is that it can deliver labeled protein inside mammalian cells, which makes *in vivo* NMR and fluorescence studies potentially possible. More importantly, the QQ-protein delivery technique has several novel features that significantly distinguishes it from all the other protein delivery techniques currently available (Figure 1-5) The novel features of the QQ-protein delivery technique ensure the physiological relevance of the delivered proteins, which is indistinguishable from the endogenous proteins. Thus, the QQ-protein delivery technique can have many applications in cell biology.

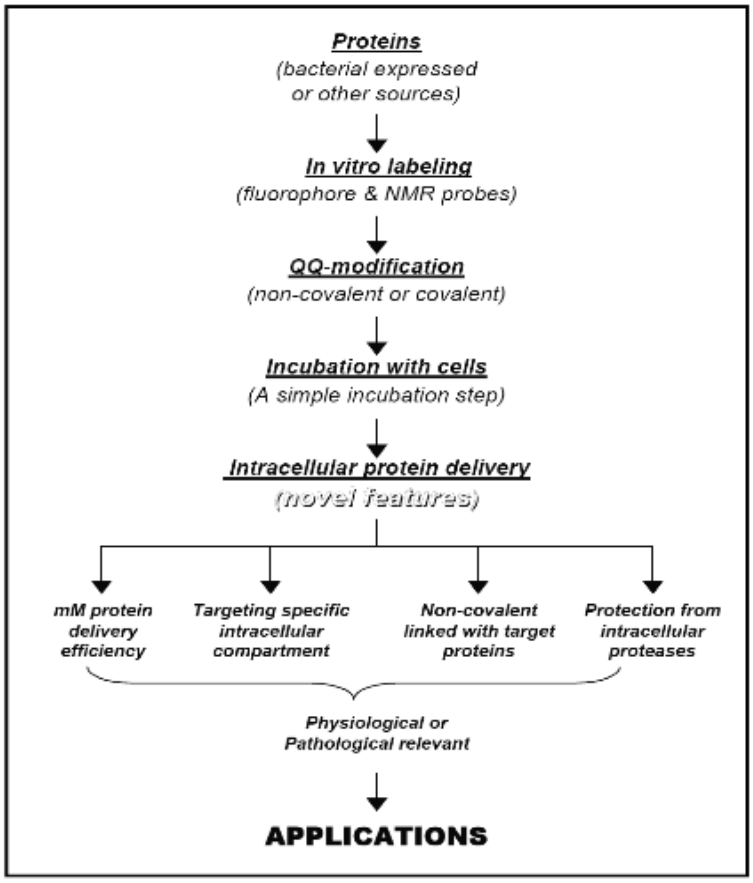

Figure 1-5: The technical flow-chart of the QQ- protein delivery system.

1.3.1. The QQ-reagent

Dr. Qianqin Li in our lab recently develoed a very efficient QQ- protein delivery system (the following data on QQ- reagents were done by Dr. Qianqian LI). In this protein delivery system, the QQ- reagent plays the key role. It is a cocktail of polyethylenimine (PEI) with unique combinations of lipids and enhancers. Table 1-3 shows examples of different formulations of the QQ-reagent. The mechanism of protein delivery by QQ-reagent is unclear. However,

PEI is well known for its efficient binding to negatively charged DNA and mediating the transfection of eukaryotic cells (*108*). PEI is highly basic harboring many amino groups. The amino groups are protonated at neutral pH and non-covalently bind negatively charged macromolecules to form a cationic complex. The cell surface is normally negatively charged due to the presence of glycoproteins, proteoglycans, and sulfated proteoglycans (*109-111*). These features can mediate interactions between the cell membrane and the cationic complex and promote uptake of the QQ-modified proteins (*112, 113*). The lipid composites and enhancers facilitate the solubility of the delivered proteins, such as membrane proteins, and help the formation of delivery vesicles, greatly enhancing the delivery efficiency of QQ-reagent protein delivery system.

Table 1-3: Recipes for different QQ reagents from the stock solution of cation reagents, lipids and enhancers (Courtesy of Dr. Qianqian Li)**.**

	Cation Reagents					Lipids				Enhancers	
QQ Reagent [a]	2K µl	8K µl	12K µl	25K µl	60K µl	DOTAP µl	DOPE µl	POPC µl	DMPE µl	DMSO µl	MG132 µl
QQ1	200					150	150			50	5
QQ2	200			50	5	100	100			50	5
QQ3	100			100		150	150			50	5
QQ4	100	50	50	50		100	100			50	5
QQ5	100	100		50		100	100			50	5
QQ6	100	50	50	50				100	100	50	5
QQ7	100	50	50	50		50	50	50	50	50	5
QQ8	50	50	50	50	50			200	200	50	5
QQ9		50	50	100			200	200		50	5
QQ10		100	100			150	150	150	150	50	5

a. 2 kDa, 8 kDa, 12 kDa, 25 kDa and 60 kDa represent the different molecular weights of PEI.

b. Total volume is 5 ml when preparing the QQ reagents. We also dissolve the protein at a concentration of 0.2-8.0 mg/ml, depending on protein solubility and the purpose of the studies.

c. In our hand, 2 kDa produces the least cellular toxicity whereas 60K produces the most cellular toxicity.

A concentration of 50 ug/ml of the QQ modified protein is optimal for relatively high protein delivery efficiency. A mixture of QQ modified protein, serum-free medium, and protease inhibitor is incubated with the mammalian cells, which have been seeded for 2-5 days to reach 80-90% confluency, for 0.5-4 hours, application dependent. The cells are extensively washed to remove the undelivered protein. With complete growth medium, the cells are ready for further experimentation.

1.3.2. High protein delivery efficiency

QQ-protein delivery has been tested using different proteins, covering a wide range of molecular weights, isoelectric points (pIs), and solubility, including two membrane proteins. Table 1-4 lists the 17 proteins, their properties, and their protein delivery efficiency using QQ-protein delivery.

Table 1-4: Proteins, their properties, and protein delivery efficiency using QQ-reagent protein delivery system (Courtesy of Dr. Qianqian Li)**.**

Proteins	Size (residues)	MW (kDa)	pI	Solubility	SDS-PAGE [a] (observable)	Conc [b] (mg/ml)	Transduction Efficiency [c]	Source
LBD-apoER2	298	31.9	4.4	Fair	Weak	0.5	Good	bacterial
BP1 of LRP6	307	34.4	5.1	Good	Yes	1	Good	bacterial
MESD(12-155)	144	16.6	5.1	Good	Yes	1	Good	bacterial
Human apoAI	243	28.1	5.4	Good	Yes	1	Good	bacterial
Human apoE(1-183)	183	21.2	5.4	Good	Yes	2	Good	bacterial
Mouse apoAI	240	27.9	5.6	Good	Yes	1	Good	bacterial
Human apoE	299	34.2	5.6	Good	Yes	1	Good	bacterial
Ferritin	183	20.0	5.6	Good	Yes	5	Good	native
MESD	195	22.0	5.9	Good	Yes	1	Good	bacterial
Oct4/Pou5f1	360	38.6	5.9	Good	Yes	0.3	Good	bacterial
LCAT	416	47.1	6.1	Poor	Yes	0.1	Good	bacterial
Hemoglobin	147	16.0	7.3	Good	Yes	3	Good	native
RAP	323	37.8	7.4	Good	Yes	1	Good	bacterial
PMP22 [d]	160	17.9	7.8	Poor	No	< 0.1	Fair	bacterial
PLP1/DM20 [d]	242	26.3	7.9	Poor	No	< 0.1	Fair	bacterial
Sox2	317	34.3	9.7	Good	Yes	0.3	Good	bacterial
MESD(150-195)	46	5.1	10.0	Good	Yes	1	Good	bacterial

[a] Protein band can be clearly visible by an SDS-PAGE of the cell lysate.
[b] This is the protein concentration used for protein transduction.
[c] The protein transduction efficiency is judged by the visibility of the protein band on an SDS-PAGE of the cell lysate after protein transduction. For the two membrane proteins, PMP22 and PLP1/DM30, the protein transduction is judged using fluorescence imaging.
[d] Both PMP22 and PLP1/DM20 are membrane proteins that contain four putative transmembrane domains.

Figure 1-6 shows the SDS-PAGE (Left panel) and Western blots (right panel) of QQ-modified receptor-associated protein (RAP) that was delivered into Hela cells. RAP is an ER resident molecular chaperone of the LDLR family members, including apoER2 and VLDLR. RAP binds tightly to the receptors at neutral pH, acting as a universal antagonist of the ligand binding domains. It also escorts the LDLR family members in the early stage of secretary pathway in the ER, reducing premature ligand binding and receptor aggregation. Detailed background information will be given in Chapter II. For now, RAP will be used as

an example to demonstrate the efficient protein delivery of the QQ-reagent. The grey arrow shows the protein band of bacterially expressed RAP, while the black arrow shows the protein band of RAP in the Hela cell lysate after protein delivery (Figure 1-6). Compared with the control Hela cell lysates (without protein delivery) (lane 1), an additional band of roughly 40 kDa was observed in the Hela cell lystes (lanes 2 and 3) of cells preincubated with QQ-modified RAP. The Western blot (Right panel) confirms that this band corresponds to RAP (compare lane 1 (control) to lane 2). In addition, the band intensity of lane 2 was estimated to be at least twice as strong as that of lane 3, where 0.25 mg/ml of purified bacterially expressed RAP was loaded. This suggests that at least 0.5 mg/ml of the QQ-delivered RAP was recovered, demonstrating the efficiency by the QQ-reagent.

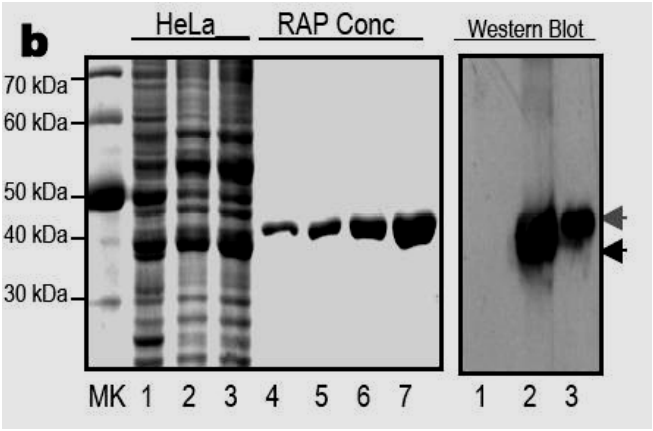

Figure 1-6: High protein delivery efficiency of QQ-reagent modified proteins (Courtesy of Dr. Qianqian Li).

Left: A 10% SDS-PAGE quantitatively showing high efficiency of the QQ-delivered RAP. Lane 1: Hela cell lysate without RAP delivery. Lane 2: 5 ul Hela cell lysate with RAP delivery for 1 hour. Lane 3: 10 ul Hela cell lysate with RAP delivery for 1 hour. Lanes 4-7: 60, 120, 250 and 500 ug/ml of RAP (20 ul). Right: A Western blot using an anti-RAP antibody. Lane 1: Hela cell lysate without RAP delivery. Lane 2: 5 ul Hela cell lysate with RAP delivery. Lane 3: 0.25 mg/ml bacterial expressed RAP in 5-ul. Gray arrow: the bacterial expressed RAP. Black arrow: the HeLa cell loaded RAP.

1.3.3. Protection of the QQ-delivered proteins from intracellular protease degradation

In order to see the stability of QQ-modified protein inside of mammalian cells, MESD, a 195-residue specific protein chaperone for the LDLR family (*114*) (*115*), was QQ-modified and tested for stability against protease degradation by adding protease cocktail into MESD solution (Figure 1-7), which is either modified without (lanes 1-4) or with (lanes 7-10) the QQ-reagent. As compared to without protease treatment (lane 5), the protein modified with the QQ-reagent (lanes 7-10) showed significantly less degredation as compared to the absence of the QQ-modification (lanes 1-4). These results suggest that the QQ-reagent can protect MESD from protease degradation and theoretically from intracellular-mediated degradation. Basically, the QQ-reagent may protect the modified protein from degradation upon entering into the cell; however, the underlying

mechanism is not well understood. This may enable the protein to refold and to obtain high yields of repurified protein.

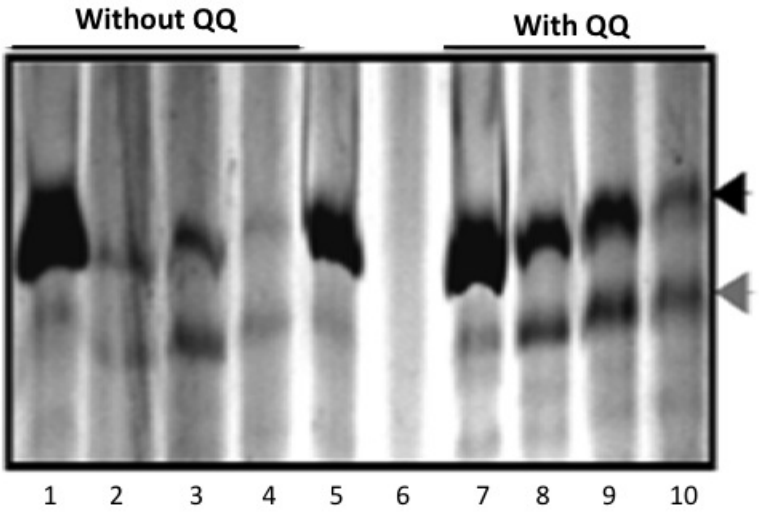

Figure 1-7: QQ-reagent protects MESD from the protease-mediated degradation *in vitro* (Courtesy of Dr. Qianqian Li).

Lane 1-4: MESD without QQ-modification. Lane 7-10: the QQ-modified MESD. While reacting with 1.4 mU/ml protease cocktail for 0 hour (Lane 1 and Lane 7), 1 hour (Lane 2 and Lane 8), 2 hours (Lane 3 and Lane 9), 4 hours (Lane 4 and Lane 10), QQ -modified MESD showed relatively higher stability than the unmodified one. Lane 5: MESD without reacting with protease cocktail. Lane 6: The QQ-reagent.

1.3.4. Targeting capability of the QQ-protein delivery technique

In order to visualize the trafficking of delivered proteins, in this case MESD, a resident ER protein, inside mammalian cells and to detect their final location, a small red fluorophore (ArrayIt Red640, Molecular Weight = 799.8 kDa)

was linked to MESD. The labeled MESD is modified with QQ-reagent and delivered inside of Hela cells. After 2 hrs post-delivery, MESD is predominantly localized to the perinuclear areas, which are potentially the ER (Figure 1-8A). To further confirm that this perinuclear localization of MESD was , in fact, to the ER, a co-localization experiment was performed. Transient transfection of a GFP-ER marker (green) cDNA into BSC-1 cells is conducted following the published protocol (*116, 117*). After 72 hrs, the transfected cells were incubated with QQ-modified MESD labeled with ArrayIt Red640 (red). After 2 hrs following protein delivery, the cells were subjected to fluorescence imaging, showing a merge of both the red and green chanels, along with the light microscopy image (Figure 1-8a). Yellow fluorescence was detected around the perinulear locations, indicating MESD was indeed localized to the ER.

The QQ-reagent was also able to correctly direct the membrane protein peripheral myelin protein 22 (PMP22) and the transcription factor Oct4 to their appropriate intracellular compartments, the plasma membrane and the nucleus, respectivley. PMP-22 was labeled with the ArrayIt Red640 (red) and delivered into Hela cells using QQ-protein delivery. Fluorescence imaging of a Hela cell shows that a majority of the red fluorescence signals are localized to the plasma membrane (Figure 1-8B and b). To study intracellular location of the QQ-delivered Oct4, the bacterially expressed Oct4 was labeled with a small fluorophore DyLight 488 (green), QQ-modified, and delievered into human fibroblast cells. Fluorescence imaging data showed that the delivered Oct4 is

predominantly localized to the nucleus, as it colocalized with the DNA stain, DAPI (blue) (Figure 1-8 C and c).

Taking together, these results show that QQ-reagent delievered proteins within the cell as evidenced by the correct *in situ* residential location of the QQ-modified MESD, PMP22, and Oct4 to the ER, the plasma membrane, and the nucleus, respectivley; otherwise, they would not be localized as such. This data also suggests that the QQ-reagent does not influence the trafficing of these proteins. Also, since these proteins were directed to their correct *in situ* cellular compartments argues that these proteins are likely to be properly folded; however, further experiments are needed to confirm this point. Overall, the QQ-reagent shows promise as a protein delievery system that can sucessfully deliever proteins into cells.

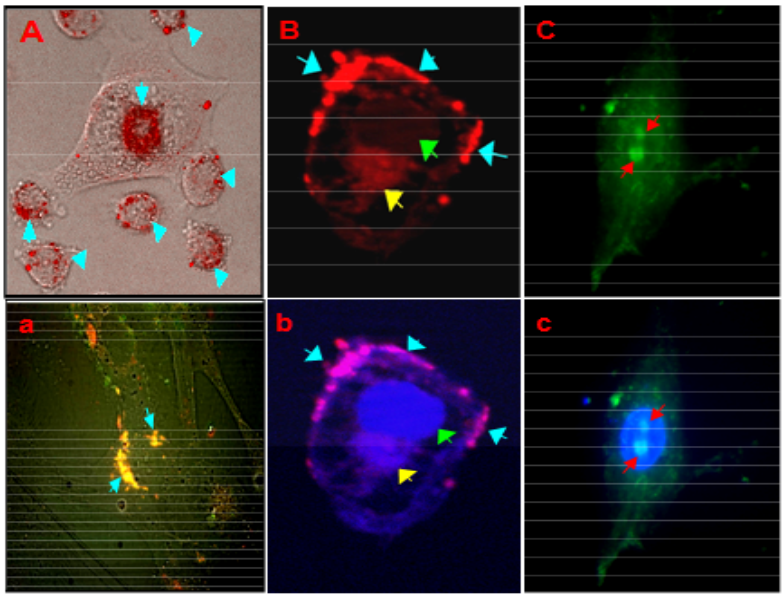

Figure 1-8: **Targeting capability of the QQ protein delivery technique** (Courtesy of Dr. Qianqian Li)**.**

(A) Fluorescence confocal image of the intracellular location of ArrayIt Red640 labeled MESD (red), merging with the light microscopy image of the same HeLa cell. The arrows showed the perinulear localization of delivered MESD. **(a)** Fluorescence confocal image of the co-localization of GFP-*ER* marker (green) and ArrayIt Red640 (red) labeled MESD in BSC-1 cells, after QQ-protein delivery. **(B)** Fluorescence confocal image of a typical HeLa cells after PMP22 transduction. PMP22 was labeled with ArrayIt Red640 (red) fluorophore. The arrow showed the plasma membrane location of PMP22. **(b)** A merging fluorescence confocal image of both Rhodamine and CFP channels for the same HeLa cell. **(C)** Fluorescence confocal image of human fibroblast cell transducted with DyLight 488 (green) labeled Oct4. The arrows showed its nucleus localization. **(c)** A merging fluorescence confocal image of both DyLight 488 (green) labeled Oct4 and nucleus staining DAPI (blue) for the same fibroblast cell.

1.4. An efficient bacterial expression system that produces gram quantity of pure recombinant proteins

The gram-negative bacterium *E. coli* offers a means for rapid, high yield, and economical production of recombinant proteins. However, high-level production of functional eukaryotic proteins in *E. coli* may not be a routine matter, sometimes it is quite challenging. Techniques to optimize heterologous protein overproduction in *E. coli* have been explored for host strain selection, plasmid copy numbers, promoter selection, mRNA stability, and codon usage, significantly enhancing the yields of the foreign eukaryotic proteins. Our lab has

been working on optimizations of bacterial expression conditions and media with a focus on achieving very high cell density for high-level production of eukaryotic proteins. Two high-cell-density bacterial expression methods have been explored, including an autoinduction introduced by Studier (*118*) recently and a high-cell-density IPTG-induction method developed in our lab, to achieve a cell-density OD_{600} of 10–20 in the normal laboratory setting using a regular incubator shaker. Several practical protocols have been implemented with these high-cell-density expression methods to ensure a very high yield of recombinant protein production. With these methods and protocols, a yield of 14–25 mg of NMR triple-labeled proteins and 17–34 mg of unlabeled proteins from a 50-mL cell culture is routinely achieved for all seven proteins tested. Such a high protein yield used the same DNA constructs, bacterial strains, and a regular incubator shaker and no fermentor is necessary. More importantly, these methods permit for consistently production of such a high yield of recombinant proteins using *E. coli* expression (*59*).

One of the major challenges to achieve high yield protein production is to achieve a high cell density bacterial expression under routine laboratory setting using a regular incubator shaker. A hybrid bacterial expression method has been developed in our lab that utilizes rich medium to achieve a high cell density before IPTG induction, while maintaining the advantage of the tightly controlled induction by IPTG in minimal medium (*59*). This hybrid expression method allows us to reach a high cell density with a final OD_{600} that is 5 to 10 fold higher than that of the regular IPTG induction method.

High cell density culture systems, especially under the non-fermentation, laboratory conditions, often cause a low or even no protein production with a high cell density culture, since bacterial cells experience stress at a high cell density using a regular incubator shaker which does not control the O_2 level, pH and nutrients of the expression medium. To solve this problem, several protocols have been developed to ensure stability of plasmids inside bacterial cells during high cell-density bacterial expression. These protocols include: (1) Proper starting culture; (2) double colony selection; and (3) optimized time course and temperature after IPTG-induction (*59*). In addition, cell areation and medium pH during high cell-density expression are also optimized. With these protocols, a gram/liter pure recombinant protein is achieved using our bacterial expression method.

1.5. Summary of literature review

Protein folding promotes a protein's native conformation and biological function. Misfolding can cause severe protein misfolding diseases, such as Alzheimer's disease, diabetes, cancer, and prion diseases. To study the protein folding mechanism, techniques including X-ray crystallography and nuclear magnetic resonance spectroscopy (NMR) are critical; however, large quantities of properly folded and biologically functional proteins are required. The present large-scale heterologous protein expression systems using *E.coli*, yeast, insect, or mammalian cell lines have inherited weaknesses, often causing protein misfolding. *In vitro* protein refolding techniques are the other option to obtain

properly folded and biologically active proteins at relatively high speed and low cost. But the efficiency of *in vitro* protein refolding is greatly affected by the aggregation and misfolding reactions happening in the same refolding system. Therefore, a more simple and cost-effective protein refolding technique is needed.

Our lab recently developed a novel QQ- protein delivery technique. By using the highly efficient QQ-protein delivery technique, we can deliver large quantities of bacterially expressed proteins inside of mammalian cells. The QQ-reagent non-covalently attaches to the delivered protein, efficiently transports the protein across the cell membrane, helps the delivered protein to evade degredation, and targets the delivered protein to its *in situ* cellular location.

Our lab has also developed a highly efficient recombinant protein expression method in an *E. coli* protein expression system. This method bases on a hybrid high cell-density IPTG-induction bacterial expression with several practical protocols to ensure high-yield production of protein. This novel bacterial expression method allows us to routinely obtain nearly gram quantity of pure recombinant proteins from a one-liter bacterial culture.

The ER is the cellular machinery of protein translation and protein folding. High density of folding chaperones and enzymes exist in the ER, efficiently assisting protein refolding. Simultaneously, the stringent quality control (QC) system can guarantee only correctly folded proteins are sent to their final destinations. By using our developed recombinant protein expression method

and QQ-protein delivery technique, we can deliver large quantity of recombinant proteins inside of mammalian cells. The ER protein folding machinery may be able to refold them with the help of foldases and chaperones. The folded protein can be repurified from the cells. The quality and quantity of purified refolded protein can meet the need of many applications, such as structural biology and protein therapeutics.

1.6. Research goals

The goal of my dissertation reseach was to develop an *in vivo* protein refolding technique that allows efficient refolding of bacterially expressed proteins using the intracellular folding machinery of mammalian cells. The QQ-protein delivery technique allows us to deliver the bacterially expressed proteins into the ER, where the protein folding machinery of the cell may properly refold the QQ-delivered proteins efficiently by recruiting chaperones and foldases. The ER quality control system ensures that only the properly refolded, bacterially expressed proteins will achieve maturity with the misfolded proteins being degraded. Thus, this *in vivo* protein refolding technique may provide an efficient strategy to refold large quantity of bacterially expressed proteins that are possibly misfolded.

In particular, this technique may be specifically useful for the efficient refolding of secreted proteins, since these proteins will be targeted for secretion into the cell culture medium once they are properly folded and posttranslationally

modified. Therefore, purification of the refolded proteins will be easy to achieve if we use serum-free cell culture medium. In addition, by properly designing the protein sequences, we can target protein domains and critical segments of an intact protein causing them to be secreted by the cell. Thus, we can take advantage of these domains and critical protein segments to allow secretion of the protein of interest, which will be already somewhat pure as it is seperated from the rest of the cellular millue, making it rather straightforward to efficiently refold and purify them in large quantity. Finally, this *in vivo* protein refolding technique may also have the capability to generate properly posttranslationally modified proteins inside the ER and Golgi, sites where some posttranslational modifications can occur.

To develop this *in vivo* protein refolding technique, two protein domains: beta- propeller/EGF domain I of LDL receptor- related protein 6 (BP1-LRP6) and the ligand-binding domains of apolipoprotein E receptor 2 (LBD-apoER2), will be tested. BP1-LRP6 and LBD-apoER2 are protein domains of two large receptors of the low-density lipoprotein receptor (LDLR)-family, with molecular weight of 34-36 kDa. Both proteins have large numbers of cysteins that form complex intracellular disulfide bonds, which are important for their structural rigidity as well as their biological function. In particular, LBD-apoER2 contain 298-residues that has 42-cysteines, forming 21 intracellular disulfide bonds. This protein domain also contains 7 Ca^{2+}-binding sites. Refolding this protein domain is quite challenging. Thus, these two protein domains are very good candidates to test this *in vivo* protein refolding technique.

Chapter II

Refolding of the Ligand-Binding Domain of Apolipoprotein E receptor 2 (LBD-apoER2) inside mammalian cells

2.1. Introduction

The aim of this project is to develop and optimize an *in vivo* protein refolding technology that can be used to efficiently refold bacterially expressed and potentially misfolded proteins using the ER folding machinery of mammalian cells. Previously, our lab demonstrated that the misfolded proteins can be delivered into mammalian cells using QQ-protein delivery and these delivered protein can be refolded inside of cells. The refolded proteins display biological functions and followed the same intracellular trafficking pathways (*119*). These observations provide the solid foundation for this *in vivo* protein refolding technology.

However, this *in vivo* protein refolding technique was in its early stage and needed to be optimized so that it can be used to efficiently refold large quantities of misfolded proteins and to produce properly folded, biologically functional proteins. In addition, our optimization also focuses on proteins with complex folding since the refolding of these kind of proteins are very challenging and in most cases, the *in vitro* protein refolding techniques cannot be used to properly refold these proteins. Our rationale is to utilize the intracellular folding machinery of mammalian cells to efficiently refold these challenging proteins. Indeed, cells have developed an efficient folding machinery and quality control systems,

ensuring that only the proper folded proteins can escape the ER-associated degradation system to reach their intracellular destinations or to be secreted outside of the cells for function.

The QQ-protein delivery technique developed in our lab ensures that we can efficiently deliver bacterially expressed proteins into the ER for refolding. This solved a major problem for this project, since the QQ-protein delivery provides us with the targeting capability to specifically and efficiently deliver the proteins of interest inside the ER, such that these proteins have a chance for refolding. The QQ-reagents are designed to non-covalently bind to the target proteins and are hence reversible, enabling the unmodified protein, to have an opportunity to be refolded by the ER resident machinery.

Our lab has been working on low-density lipoprotein receptor (LDLR) family for several years. We are particularly interested in two domains of the LDLR family: the ligand-binding domain and the YWTD β-propeller domain, since these two domains are the two essential domains for biological functions of this protein family. In order to perform research on these proteins, we first used bacteria to prepare recombinant domains using the high cell-density bacterial expression method recently developed by our lab (*59*). We found that although our bacterial expression method produced large quantities of recombinant proteins, these bacterially expressed domains were misfolded and not biologically active. This is not surprising since both of the domains contain many cysteines that form multiple intramolecular disulfide bonds. In addition, these

domains also contain many Ca^{2+}-binding sites. These structural features make it very challenging to properly fold these domains.

However, these two domains of LDLR family provide good candidate proteins for us to optimize our *in vivo* refolding technology. In this chapter, I will focus on the refolding technique aimed at refolding the ligand-binding domain of apolipoprotein E receptor 2 (LBD-apoER2), an LDLR family member. I optimized several different parameters in an attempt to properly fold LBD-apoER2 including: the best QQ-reagent/LBD-apoER2 ratio for efficient delivery into the ER, different cell culture conditions to facilitate refolding, and purification procedure to isolate the protein after refolding. Through finding the best conditions of this *in vivo* protein refolding technique, I was able to obtain properly refolded LBD-apoER2. Although the refolding efficiency of the QQ-delivered LBD-apoER2 was only up to 20%, successive rounds of delievery enabled me to get large quantitites of protein that I later purified. Most important of all, I was also able to demonstrate that the purified LBD-apoER2 could bind to the ligands of apoER2, suggesting that I obtained biologically active and hence properly folded protein using this *in vivo* protein refolding technique.

2.1.1. LDL receptor superfamily

Lipoprotein particles mediate the transport of lipophilic molecules in both the peripheral circulation system and within the central nerve system (CNS) to maintain lipid homeostasis. These particles are typically spherical complexes of

lipids and apolipoproteins. They can be recognized by the low-density lipoprotein (LDL) receptors located on the cell surface for uptake.

The low-density lipoprotein receptor (LDLR) superfamily consists of over ten known endocytic receptors, all of which share homology with the LDL receptor. Most of these receptors are multifunctional, binding and endocytosing many structurally and functionally distinct ligands. Together with the LDL receptor, these receptors form the LDLR superfamily, primarily owing to their structural homology, as well as functional redundancies.

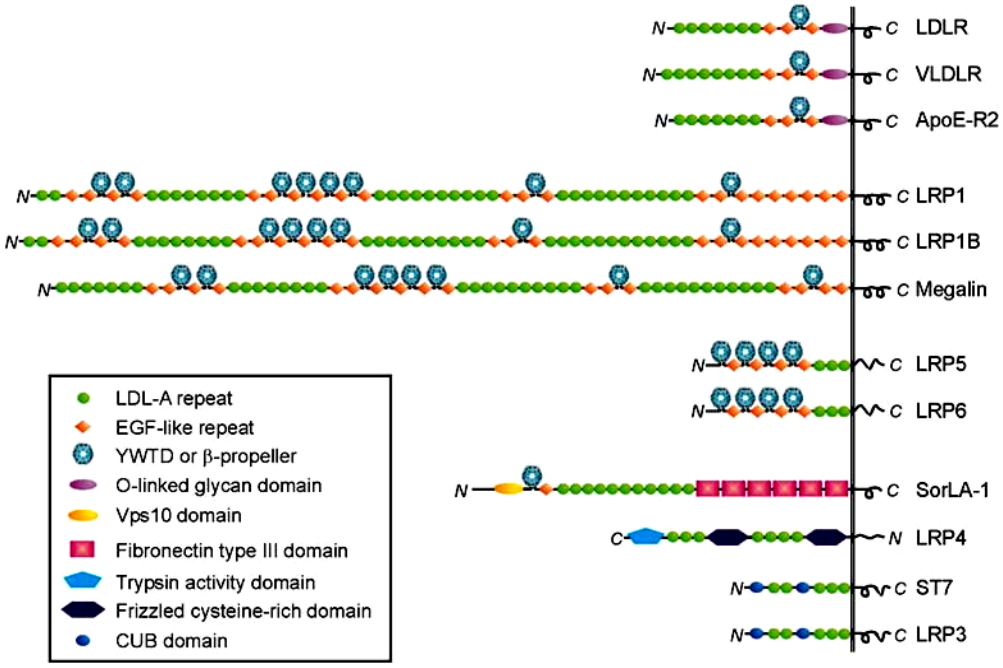

Figure 2-1: Domain organization of human LDLR family members (*120*).

Domains characteristic for the LDLR family are discussed in detail in the text. All family members have a type I topology (i.e., N-terminus in the ER lumen or the extra-cellular space and the C-terminus in the cytosol) with the exception of LRP4. The cytosolic domains contain one or multiple copies of NPXY, YXXL and di-leucine based sequences that mediate internalization and intracellular sorting.

Figure 2-1 shows the domain organization of human LDLR-family. All the members of the LDL receptor superfamily are composed of five major domains --- complement-like ligand binding repeats, epidermal growth factor (EGF) –like repeats, YWTD β-propeller repeats, a transmembrane domain, and a short cytoplasmic tail (*121*). A primary feature of this family is the presence of cysteine-rich ligand-binding repeats, which form the ligand-binding domain (*122-124*). These ligand-binding repeats, each of which is about 40 amino acids in length and contains six cysteine residues, are also found in a number of complement components and are therefore also referred to as complement-type repeats (CR) or the LDL receptor type-A (LA) repeats.

The complex structures of LDLR superfamily members indicate that biosynthesis of these receptors must require the assistance of molecular chaperones. Indeed, studies with LRP, megalin, and VLDLR have shown that RAP serves as a specialized chaperone to assist LDLR superfamily members to fold correctly and then traffic safely within the early secretary pathway (*125, 126*).

2.1.2. LDL receptor superfamily - structures

Extensive structural studies have been performed on individual domains of LDL-receptor. For the ligand-binding domain, four ligand binding repeat structures have been individually solved; only one is by X-ray crystallography, the other three are by NMR (*127*). For example, repeat 1 consists of a beta-hairpin structure, followed by a series of a beta turns with many of the side chains of the acidic residues, including the highly conserved Ser-Asp-Glu motif, clustered on one face of the module (Figure 2-2) (*128, 129*). Other ligand binding repeats display a similar structure with subtle changes (*127*). The EGF precursor homology repeat consists of 392 amino acids; there are three cysteine-rich repeats (A, B and C), two at the N-terminus, and one at the C-terminus of YWTD β-propeller repeats. In the middle, the YWTD repeats form a six-bladed β-propeller (Figure 2-2) (*130*).

In 2002, Rudenko et al. determined the crystal structure of the LDL receptor domain I and II, residues 1-699, at pH 5.3 corresponding to the ligand-released state (*131*) (Figure 2-3). Two histidines (H562, H586) in the β-propeller point to repeats 4 and 5, and H190 from repeat 5 protrudes to the β-propeller blades. Mutation of the three histidines did not affect the LDL-binding ability, but LDL receptor loses its ability to release the bound ligand (*132*). It was hypothesized that within the low pH of the endosome, the ligand-binding domain of the receptor dissociates from its ligand, and transforms from its open state to the closed state. Repeats 4 and 5 bend to contact the EGF precursor domain via their calcium-binding loop (*131*), and the ligand binding domain flexibility for the

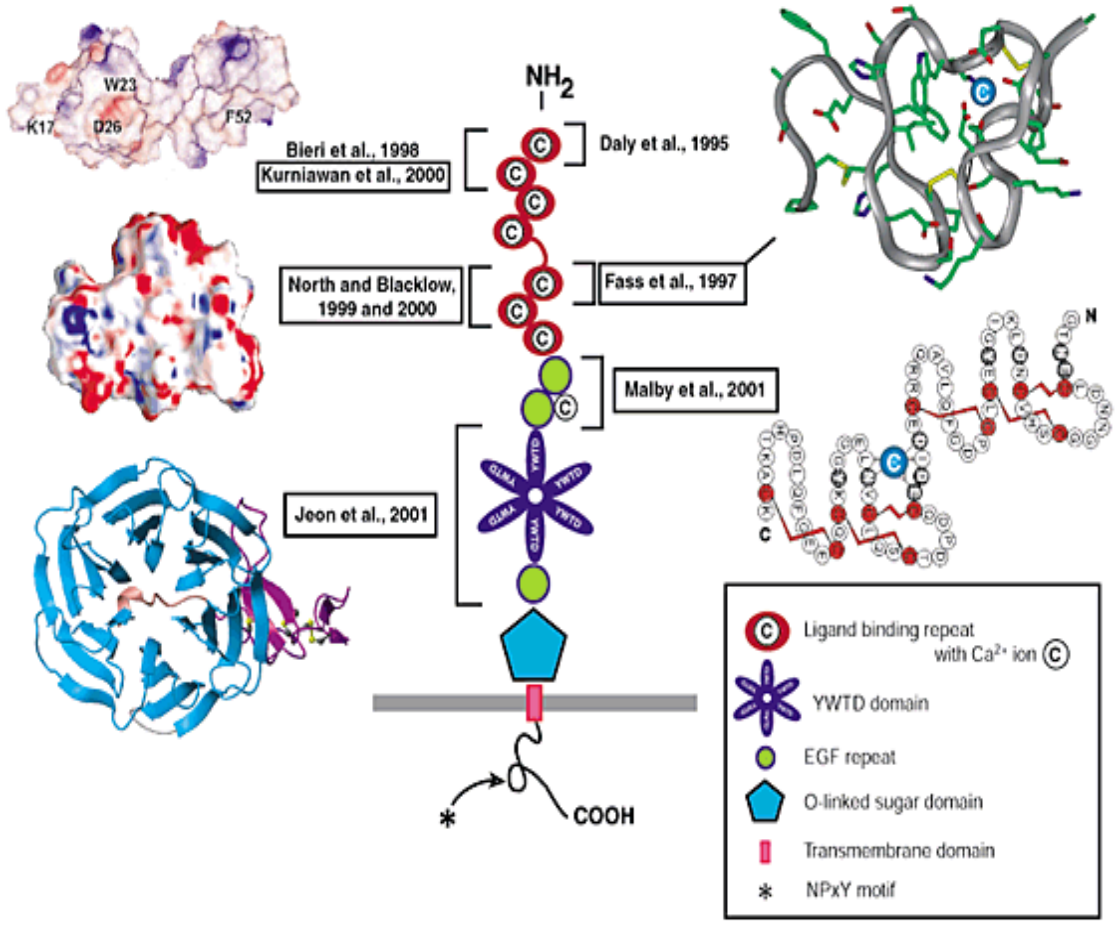

Figure 2-2: Domain organization of the LDL receptor and regions within the extracellular domain for which structural information has been obtained (*133*).

The O-linked sugar domain is thought to act mainly as a hydrophilic buffer zone that keeps bound lipoprotein particles away from the lipid bilayer of the plasma membrane. The NPxY motif in the cytoplasmic tail of the LDL receptor is required for clustering and internalization. Boxed references next to structural representations indicate the cited publication from which the respective structure has been reproduced.

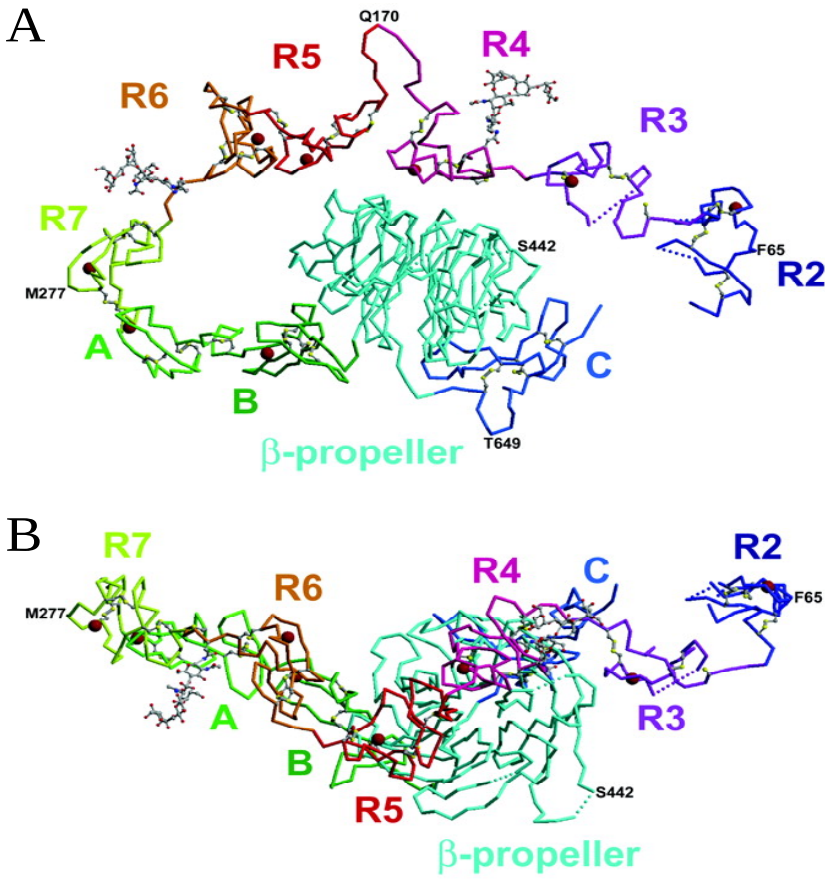

Figure 2-3: The crystal structure of LDL receptor residue 1-699 at pH 5.3 (*131*).

(**A**) C^{α} trace of LDL-R monomer. Modules are colored according to their boundaries with the ligand-binding domain containing R2 (residues 44 to 85), R3 (85 to 124), R4 (124 to 170), R5 (170 to 212), R6 (212 to 254), R7 (254 to 294), and the EGF precursor homology domain containing A (294 to 332), B (332 to 377), β propeller (377 to 643), and C (643 to 693). Regions of poor backbone connectivity are dashed. Calcium ions are indicated as red spheres, and disulfide bonds and carbohydrates on Asn^{135} and Asn^{251} are shown in gray as ball and stick (sulphur atoms, yellow; oxygen, red; nitrogen, blue; and carbon, gray). (**B)** The view is rotated 90°from that of **(A)**.

open-closed state transformation is provided by the linker regions between seven repeats instead of the connection of two domain (*132*).

2.1.3. LDLR superfamily – functions

The function that is most commonly associated with this evolutionarily ancient family is cholesterol homeostasis. In humans, excess cholesterol in the blood is captured by low-density lipoprotein (LDL) and removed in the liver by the endocytosis of the LDL receptor (*122, 124, 134*). Recent evidence indicates that the members of the LDL receptor gene family are active in the cell signalling pathways between specialized cells in many, if not all, multicellular organisms (*135, 136*).

Endocytosis is the process by which cells absorb molecules (such as proteins) by engulfing them. The LDLR family plays a critical role in mediating this important cellular process (Figure 2-4). The first step of endocytosis is for molecules or ligands to bind to LDL receptors exposed at the cell surface. These transmembrane proteins have the ligand-binding domain that recognizes and binds to the ligands. The portion of the plasma membrane with bound ligand is internalized by endocytosis. A drop in the pH (from ~7 to ~5) causes the LDL to separate from its receptor. The vesicle then pinches apart into two smaller vesicles: one containing free ligands such as LDLs; the other containing now-empty receptors. The vesicle with the LDLs fuses with a lysosome to form a secondary lysosome. The enzymes of the lysosome then release free cholesterol

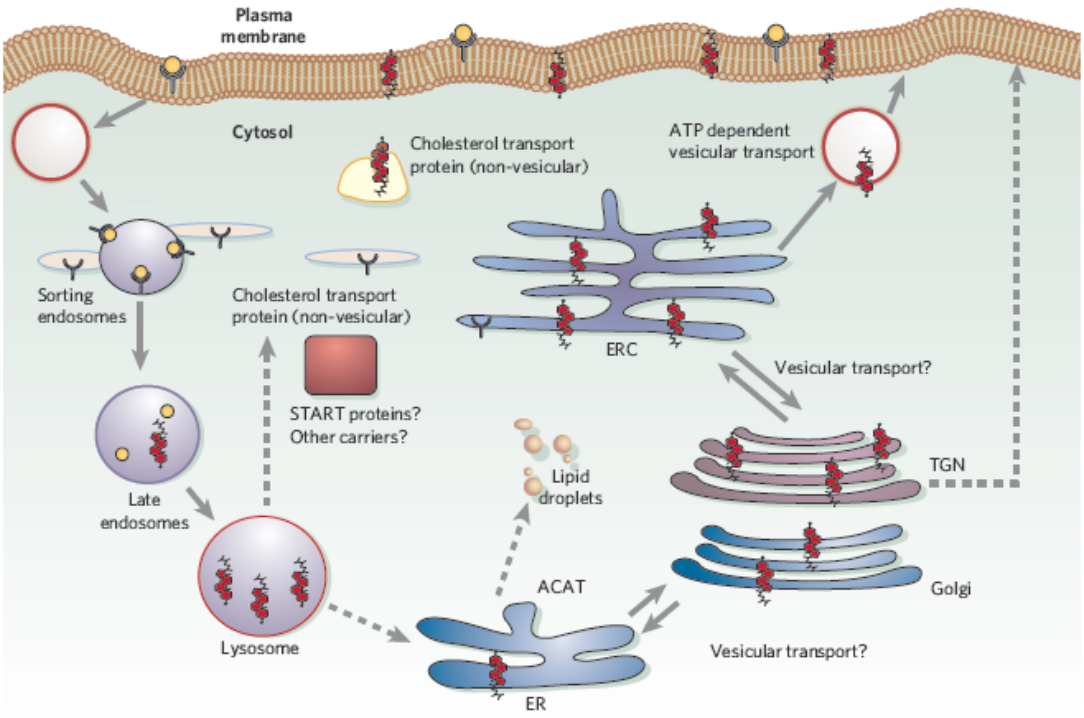

Figure 2-4: LDLR biosynthesis and transport (*137*).

Briefly, LDLR is synthesized by the ER-associated ribosome and is folded inside the rough ER and post-translationally modified inside ER and Golgi. The properly folded and post-translationally modified LDLR translocates on the cell surface for function. Upon binding to LDLR, the LDL-LDLR complex is endocytosed inside the cells via clathrin-coated vesicles, which then travels to endosome. The low pH inside endosome releases the LDL from LDLR. The released LDLR is recycled back to the cell membrane for function, where the LDL is transported to lysosome for degradation.

into the cytosol. The vesicle with unoccupied receptors returns to and fuses with the plasma membrane, turning inside out as it does so (exocytosis). In this way

the LDL receptors are returned to the cell surface for reuse. Various members of the LDLR family display functional similarities (*122, 124, 134*).

One ligand, apolipoprotein E (apoE)-containing lipoprotein, binds most of the receptors in the family, suggesting a role in lipid metabolism. Also, most members of the LDLR family bind a 39 kDa receptor-associated protein (RAP) (*138*). RAP is an endoplasmic reticulum (ER)-resident chaperone that functions in receptor folding and trafficking along the early secretary pathway (*125, 138, 139*). Upon binding to the receptors following their translation, RAP promotes proper folding and disulfide bond formation for members of the LDLR family. In addition, RAP universally antagonizes the ligands binding to the receptors. Because this later feature of RAP is believed to be important for escorting the receptors and preventing pre-mature ligand binding during their exocytic trafficking, the recombinant form of RAP has been extensively used in structural and functional studies into the biology of the LDLR family members (*138*).

Recently, several independent studies have also demonstrated roles for LDLR family members in cellular signaling (*135, 136*). For example, cellular signaling through the VLDLR and/or apoER2 is important for the Reelin/disabled pathway, which participates in neuronal cell migration during embryonic development (*140*). In addition, LRP6 has been shown to be required for the Wnt signaling pathway during embryonic development (*141-143*). Furthermore, a set of cytoplasmic adaptor and scaffold proteins containing PID or PDZ domains, including mammalian Disabled-1 (mDab1), mDab2, FE65, JNK-interacting

protein JIP-1 and JIP-2, PSD-95, CAPON, and SEMCAP-1, binds to the cytoplasmic tails of members of the LDLR family (*136, 140, 144-146*).

Members of the LDLR play roles in the pathogenesis of human diseases. A classic example is mutation of the LDLR itself in the forms of familiar hypercholesterolemia (FH) (*147*). The crucial role of the LDLR in cholesterol homeostasis is indicated by the more than 1500 human mutations that have been found in patients with familial hypercholesterolemia (FH) (*148, 149*). LRP5 and LRP6 are required for activation of the canonical Wnt signaling pathway for cell proliferation and differentiation. Loss of *Lrp5* is associated with delayed mammary gland development and Wnt1 induced tumorigenesis. Functional mutations in the Wnt co-receptor LRP5 also lead to osteoporosis-pseudoglioma, while gain of function mutations lead to high bone mass disorders (*150-153*). VLDLR and ApoER2 are obligate components of Reelin signaling pathway essential for neuronal migration during development (*154, 155*). ApoER2 immunoreactivity in human hippocampus is present exclusively in neurons and is increased in Alzheimers Disease (AD) (*156*). The c-Jun N-terminal kinase (JNK) pathway may transduce Aβ neurotoxicity, and apoER2 binds JNK-interacting proteins, thus suggesting a possible involvement of apoER2 in AD pathophysiology (*145*).

2.1.4. ApoER2

ApoER2 (apolipoprotein E receptor 2, also called LRP8, is a member of the LDLR family (*157*). It is located on chromosome 1p34 and contains 19 exons,

which span a genomic region of about 60 kb (*157*). ApoER2 is predominantly expressed in brain and placenta (*158*). However, alternative splicing of the apoER2 gene transcripts can give rise to different varieties of the apoER2 proteins from species to species or even from tissue to tissue within the same organism.

ApoER2 consists of five domains that resemble those of the LDLR and the very low-density lipoprotein (VLDL) receptor (*157*). Expression of apoER2 in 293 cells demonstrated that it is endocytosis competent (*159*). However, direct comparison of the cytoplasmic domains of LRP, LDLR, VLDLR, and apoER2 overexpressed in CHO cells suggested a role in signaling rather than in endocytosis (*160*). Studies have shown that apoER2 participates in transmitting the extracellular Reelin signal to intracellular signaling processes initiated by Disabled-1 (Dab1) (*161*). Reelin is an extracellular protein essential for the development of laminated cortical brain structures in vertebrates (*154*). Upon high-affinity binding to Reelin at the ligand-binding domain, apoER2 transmits the Reelin signal to its intracellular domain and this signal underlies learning and memory in the adult brain. ApoER2 appears to be the dominant Reelin receptor, at least in the forebrain.

Later studies showed Reelin binds with high affinity to both apoER2 and VLDLR (*155, 161, 162*), while Dab1 binds to the cytoplasmic NPxY motif of both receptors through a PTB domain (*144, 146*). Importantly, all of the mice deficient in both apoER2 and VLDL, or their ligand Reelin, or the adaptor protein Dab1 show markedly high levels of hyperphosphorylated tau protein, which are found

to accompany the formation of the neurofibrillary tangles that are one of the pathological hallmarks of Alzheimers Disease (AD) (*163*). ApoER2 knockout mice display disturbed neuronal organization in the hippocampus (*164*). ApoER2 immunoreactivity in human hippocampus is present exclusively in neurons and is increased in AD (*156*). The c-Jun N-terminal kinase (JNK) pathway may transduce Aβ neurotoxicity, and apoER2 binds JNK-interacting proteins, thus suggesting a possible involvement of apoER2 in AD pathophysiology (*145*).

2.1.5. Receptor-Associated Protein (RAP)

The complex structure of LDL receptors is highlighted by the presence of clusters of cysteine-rich, ligand-binding repeats (*165, 166*). Recent studies have shown that, under physiological condition, a 39 kDa receptor-associated protein (RAP) RAP serves as a molecular chaperone to assist the lipoprotein receptor folding and inhibit pre-mature ligand binding through LDL receptor secretary pathway (*125, 126*). Human RAP is comprised of 323 amino acids. RAP has a classical ER-signal sequence, a carboxyl-terminal tetrapeptide (HNEL) that is similar to the *ER*-retention consensus sequence (KDEL) (*165, 166*). The intracellular distribution of RAP was analyzed and quantified, showing localization primarily within the *ER* (70%) and the *Golgi* network (24%) (*167*).

The most dramatic effect RAP is the inhibition of binding and/or uptake of all other known LRP ligands (*168-173*). This ability of RAP distinguishes this protein from other LRP ligands, which seldom inhibit one another’s binding or

uptake. In addition to LRP, RAP also binds to other members of the LDLR superfamily and inhibits ligand interactions. Studies have shown that RAP exhibits high-affinity binding (K_D~1-10 nM) for LRP, megalin, the VLDLR, apoER2/LRP8 and LRP11/sorLA-1 (*174*). But relatively lower binding affinity (K_D~250 nM) was found for the LDLR (*175-179*).

RAP dissociates from the receptors at the lower pH environment of the *Golgi,* which is consistant with its requirement of binding to LRP and Ca^{2+} at neutral pH (*169*). Because of the N-terminal KDEL-like consensus, after dissociating from receptors in the *Golgi,* RAP will be recycled back to the *ER* by COPI coated vesicles (Figure 2-5) (*139*). All these special properties of RAP define a novel class of molecular chaperones that selectively protect endocytic receptors from ER associated degradation (ERAD). RAP protects receptors such as, the LDLR superfamily, by associating with the ligand binding domain of the members of receptors early in the secretary pathway, reducing their ligand binding capacity thereby preventing premature ligand-binding induced aggregation, and their subsequent degradation in the ER. Figure 2-5 depicts how RAP escorts receptors trafficking from the *ER* to the *Golgi* apparatus.

2.1.6. Challenge of refolding LBD-apoER2

It is still unclear how individual domains of LDLR family folds into a functional receptor. Disulfide bond formation and calcium incorporation are two

dominant characteristics of ligand-binding domain of LDLR (LBD-LDLR) folding in the ER. The newly synthesized LBD-LDLR polypeptide chains include seven

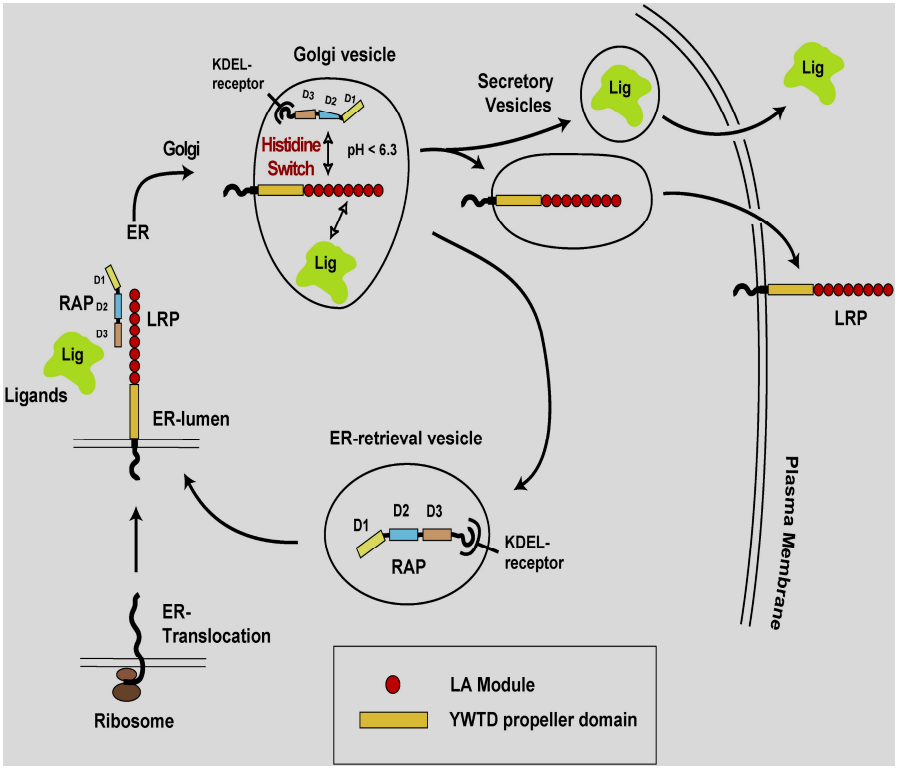

Figure 2-5: ER export and retrieval cycle of RAP- Role in LRP secretion (*126*).

LDL receptor-related proteins (LRP) are type I membrane proteins that undergo cotranslational folding during translocation into the ER. Association with RAP prevents premature ligand (Lig) binding in the ER. After transition into Golgi compartments, RAP undergoes a conformational change, which is initiated by the low-pH-induced charge reversal of histidines (Histidine switch). This partial unfolding of RAP may also facilitate its interaction with the KDEL receptor, which mediates the retrieval of RAP to the ER. The mildly acidic environment of the Golgi further suppresses ligand interaction in the absence RAP, thereby ensuring proper sorting and secretion of LRP and cosecreted ligands.

ligand binding repeats that were demonstrated to be folded rapidly into compact structures, containing non-native disulfide bonds linking distant regions of the protein (*180, 181*). With time, the aid of molecular chaperones, and foldases in the ER, the non-native disulfide bonds are reshuffled, allowing extension of the molecule. Ultimately, in the native conformation, disulfide bonds only exist between cysteine residues within individual ligand binding repeats (*131*). General chaperones recognize and interact with unfolded, partially folded or misfolded proteins due to common features such as exposed stretches of hydrophobic amino acid residues, thereby preventing aggregation. The relatively oxidizing environment of the ER may support disulfide bond formation, but the actual redox reactions are catalyzed by oxidoreductases of the protein disulfide isomerase (PDI) family (*182-184*).

Incorporation of calcium ions is the second major characteristic of LBD-LDLR folding. These domains also have very high-affinity calcium binding sites. NMR studies on the folding of repeat R5 demonstrated a strict correlation between native disulfide bond formation and calcium binding (*185*). In the absence of disulfide bonds, the calcium binding capacity is negligible. When the distal disulfide bonds (non-native disulfide bonds) between cysteines II and V, and IV and VI are formed, calcium will bind, but maximum affinity requires a native structure.

Specialized chaperones, RAP and MESD/Boca may facilitate folding of different domains of LDLR relatives. Whereas RAP improved the folding of the

ligand binding domain of LDLR family (*125, 186*), MESD/Boca mainly act on LDLR containing EGF-like repeats and YWTD β-propeller structures (*115, 187*).

The total seven ligand-binding repeats in the LBD-aopER2 contains 42 cysteines that form 21 disulfide bonds. Each ligand binding repeat in the ligand binding domain (LBD) has three disulfide bonds. In addition, each ligand-binding repeat also contains a Ca^{2+}-binding site. Because of the complex array of the disulfide bonding formation and Ca^{2+}-binding within the LBD-aopER2, it is impossible to refold the protein *in vitro* into the native conformation, especially without the help of molecular chaperones and folding enzymes. Therefore, refolding bacterially expressed LBD-apoER2 becomes extremely challenging. To our knowledge, only one report claimed successful refolding of the bacterially expressed recombinant LBD-LDLR *in vitro* by dialyzing the protein under redox conditions (*188*). We made several attempts to refold the structurally similar LBD-apoER2 under this published condition but found that this *in vitro* refolding condition could not successfully refold LBD-apoER2 into its native conformation.

This problem drove us to explore other methods to refold bacterially expressed recombinant LBD-apoER2. We found that our *in vivo* protein folding technology is able to properly refold bacterially expressed apoER2 inside mammalian cells. We further optimized this *in vivo* protein refolding technique and achieved up to 20% refolding efficiency. This allows us to prepare milligram quantity of properly refolded LBD-apoER2 that is biologically functional. This is the main focus of this chapter.

2.1.7. A novel *in vivo* protein refolding technique

Our lab recently developed a novel *in vivo* protein refolding approach to efficiently refold large quantities of bacterially expressed proteins into their properly folded and biologically functional conformations. This technique is based on the fact that the mammalian cells have comprehensive protein folding machinery to assist complex protein folding.

Protein folding in mammalian cells occurs in the ER, an extremely macro-molecularly crowded environment harboring lots of folding enzymes and molecular chaperones (*188*). Most chaperones, such as the Hsp70s, can recognize and bind to hydrophobic residues and/or unstructured backbone regions of non-native proteins, not only to block intermolecular aggregation but also to prevent or reverse intramolecular misfolding. Folding enzymes, such as disulphide isomerases and peptidyl–prolyl isomerases, can catalyze rate-limiting steps in protein folding (*12*).

Not only does the ER provide a unique folding environment, but it also exerts crucial quality control functions (*188*). Certain chaperones of the Hsp100 family even have the ability to unfold non-native proteins or to disrupt small protein aggregates by an ATP-dependent mechanism. The ATP-dependent unfolding is a prerequisite for the proteolytic degradation of misfolded proteins (*37, 38*). Proteins that fail to fold or assemble are not allowed to proceed further downstream (to the Golgi). Instead, they undergo ER-associated degradation (ERAD).

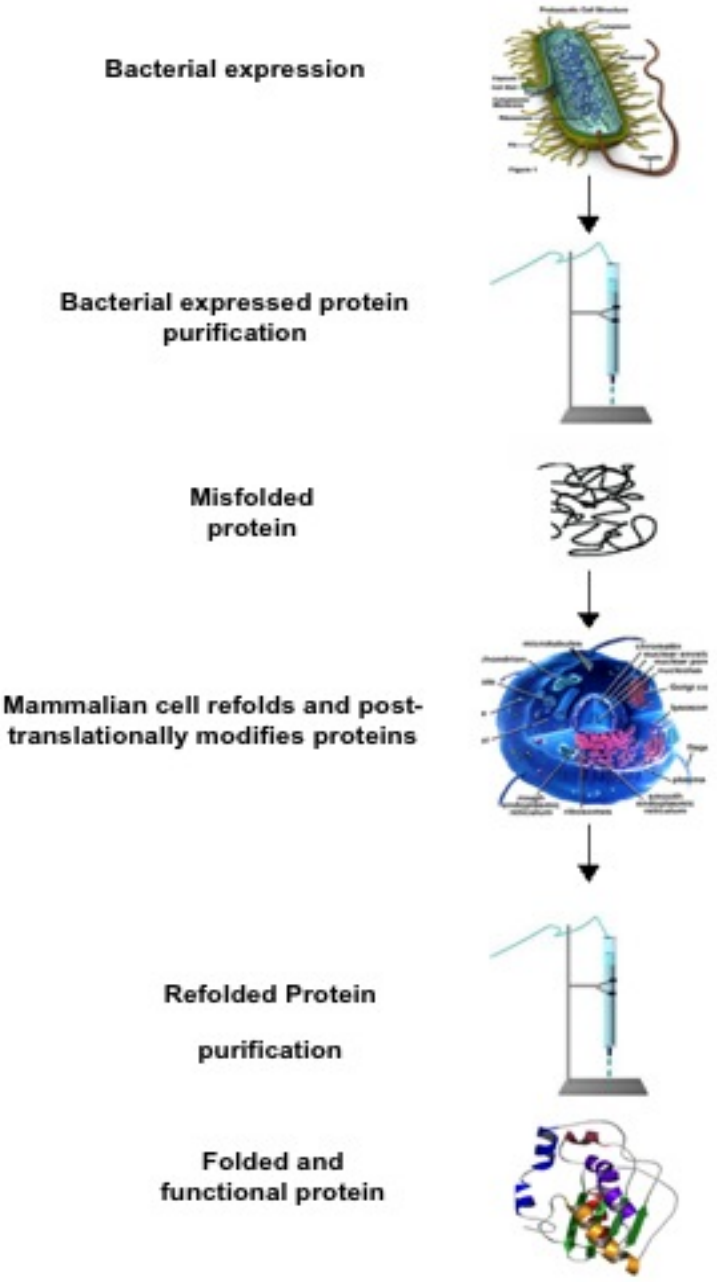

Figure 2-6: The technical flow chart of the novel *in vivo* protein refolding technique developed by our lab.

The target protein is first expressed in *E.coli* and then is subjected to affinity column purification. The bacterial expressed protein is misfolded. The QQ-reagent modifies the bacterial expressed protein and delivers it inside of mammalian cells for protein refolding and post-translational modification. Properly folded protein will be purified to obtain functional proteins.

Our refolding method takes advantage of this comprehensive protein folding machinery inside mammalian cells by using our highly efficient QQ-protein delivery technique and the high-yield bacteria expression method. As shown in Figure 2-6, first, a very high yield of bacterially expressed proteins is achieved by our advanced bacterial expression system. We purify bacterially expressed recombinant proteins by affinity chromatography. If the purified proteins are not properly folded and functional, we perform the *in vivo* protein refolding procedure. By using QQ-reagent to non-covalently modify the bacterially expressed proteins, we deliver QQ-modified proteins to mammalian cells for *in vivo* protein refolding and post-translational modification. The properly refolded protein will be sent to its targeted intracellular compartments or will be secreted into the cell culture medium, depending on its intrinsic signal. We then lyse the cells or collect the culture medium to purify the properly refolded protein. Our biophysical studies and function assay data suggest that the purified refolded proteins adopt proper conformations and are biologically functional.

2.2. Material and methods

2.2.1. Strain, plasmid and media

E.coli stain ER2556 [genotype: F ⁻ λ ⁻ fhuA2 [lon] ompT lacZ::T7 geneI gal sulA11 Δ(mcrC-mrr)11::IS10 R(mcr-73::miniTn10)2 R(zgb-210::Tn10) 1 (TetS) endA1 [dcm]; *New England Biolabs*] was used for molecular cloning. *E. coli* strain BL21(DE3) [F- *ompT hsdS*$_B$ (r_B-m_B-) *gal dcm* (DE3 [*lacI lacUV5-T7 gene1 ind1 Sam7 nin5; New England Biolabs*] was used for protein expression and purification. LBD-apoER2 with an N-terminal 6-histidine tag was constructed using a pET30a-sHT bacterial expression vector which is engineered from pET30a (+) (*Novagen*).

2.2.2. DNA manipulation

The original long his-tag of the pET30a (+) vector (*Novagen*) was replaced by a short his-tag containing six histidine plus two serine residues by mutagenesis. Mutagenesis was carried out using the QuickChange^{TM} site-directed mutagenesis kit (Stratagene, CA). The mutations were confirmed by DNA sequencing.

LBD-apoER2 was subcloned into the engineered pET30a-sHT vector. The confirmed positive DNA constructs were first transformed into *E.coli* ER2566 competent cells for plasmid DNA replication and glycerol stock purposes and then were transformed into BL21(DE3) cells for protein expression.

2.2.3. Protein expression and purification

Protein expression was carried out using *E.coli*. BL21 (DE3) cells. A 10-15 ml overnight culture in a rich medium, such as 2x YT, was made from a glycerol stock. Next morning, the overnight culture was diluted into the LB medium with a ratio 1:100. The culture was gown in LB medium with 30 ug/ml kanamycin at 37 ℃ until it reached an OD_{600} of 0.8-1.0. The cells were spun down and re-incubate into a minimal M9 medium (kanamycin, 30 ug/ml) at 1:1 volume ratio. The culture was continued at 20 ℃ for one hour. A sample of the cu lture was taken to check its OD_{600} to confirm the continuous growth of cells under lower temperature and the minimal M9 medium. Then expression of the recombinant LBD-apER2 was induced by IPTG (0.5 mM, final concentration). The culture continued overnight at 20 ℃.

The cells were harvested by centrifuge next morning. The cell pellet was re-suspended in the binding buffer (10 g/ml cells, 20 mM Tris-HCl, 500 mM NaCl, 2.5 mM imidazole, 6 M urea at pH 8.0) of His tag affinity purification system and sonicated on ice. To separate the cell pellet from the supernatant, the cell pellet was centrifuged. The pellet was again re-suspended in the binding buffer with 6 M urea and sonicated on ice. This procedure was repeated three times to ensure that all bacterially expressed LBD-apoER2 in the inclusion bodies was solubilized in the buffer system. All supernatant fractions were combined and applied on a His•Bind® resin column (Novagen). The column was washed with a large amount of binding buffer with 6 M urea and the washing buffer containing 20-30 mM imidazole for 100-500 ml. The purified recombinant LBD-apoER2 was eluted

from the column using the elution buffer with 1.0 M imidazole and 6 M urea. The eluted protein was extensively dialyzed against 10 mM ammonium bicarbonate. After dialysis, the protein solution was lyophilized and the pure LBD-apoER2 protein powder was obtained, which was weighed and stored at -20℃ for future experiments.

2.2.4 Isotope-labeled protein expression and purification

High cell-density expression method (*59*) was used to greatly improve isotope labeled protein yield at the same time to reduce the costs. The transformed cells were inoculated into LB medium for overnight cultures, which was then diluted by 100 volume folds into the LB medium at the next morning and cultured at 37 ℃ in the presence of kanamycin (30 ug/ml). When the OD_{600} reached around 1.5-2.0, we switched the cell culture to isotope-labeled M9 minimal medium by gently spinning down cells and re-suspended the cell pellet into the medium. Isotope-labeling is achieved by expressing LBD-apoER2 in this special isotope-labeled M9 minimal medium, in which NH_4Cl is replaced by 0.1% $^{15}NH_4Cl$ and glucose is replaced by 0.2% ^{13}C-glucose and H_2O is replaced by D_2O for $^{2}H/^{13}C/^{15}N$ triple-labeled samples, respectively. After switching the medium, the cells were cultured for another 1.0-1.5 hours for growth recovery and medium exchange at 37 ℃. Then 0.8 mM IPTG was ad ded to induce protein production. The cell culture was incubated in 20 ℃ for overnight expression. The cells were harvested next morning. Protein purification was carried out as same as unlabeled protein described above. The dialyzed protein solution was

lyophilized to get the pure isotope-labeled LBD-apoER2 protein powder. The protein powder was weighed and was stored at -20 ℃ fo r future experiments.

2.2.5 Cell culture

HeLa CLL-2 cells were cultured in Dulbecco minimum essential medium (DMEM) containing 8% fetal bovine serum (FBS). The cells were maintained at 37℃ in humidified air containing 5% $CO_{2.}$ For protein delivery, HeLa CLL-2 cells were incubated to reach 80% confluence (about $6x10^6$ cells).

2.2.6. Protein delivery time course

5 mg of bacterially expressed LBD-apoER2 was modified using 4 ml of QQ-reagent for overnight at 4℃ with gental shaking. QQ-reagent contains 15 mg/ml PEI (1300kDa) at pH8.0. Modified LBD-apoER2 was centrifuged at 5000 rpm for 5 minutes to remove precipitation. Then the modified LBD-apoER2 was mixed with serum-free DMEM to reach the total volume of 50 ml and the final protein concentration of 0.1 mg/ml and QQ-reagent at 1.2 mg/ml. 10 flasks (75 cm^2) of 80% confluent HeLa cells were incubated with the protein delivery medium and protease inhibitor for 10 hours. Each hour, triplicate samples of the protein delivery medium were taken. After 10 hours of incubation with the protein delivery medium, the protein delivery medium was removed and HeLa cells were gently washed with PBS buffer, three times. Then fresh serum-free DMEM

medium was added for another 12 hours. After 12 hours, samples of the medium were taken (in triplicate). All the samples were analyzed further.

2.2.7. DNA transfection and fluorescence spectroscopy

The *ER*-GFP marker DNA (15 ul, 0.5 mg/ml) or Golgi-CFP marker DNA (20 ul, 0.5 mg/ml) was mixed with Escort lipid for 2 hours (20 ul, 0.5 mg/ml). The mixture was added into 200 ul DMEM cell culture medium and incubated for 20 minutes at room temperature. This DMEM medium was then mixed with 1 ml DMEM containing HeLa cells and 5% FBS and incubated for 3 hrs. The cells were washed several times and incubated in a DMEM medium with 10% FBS for 72 hours before fluorescence imaging.

The fluorescence spectroscopy was carried out using a PTI QuantaMaster QM-7/2003 spectrofluorometer at room temperature, with an excitation at 250 nm. To deconvolute fluorescence spectra, we carried out synchronous fluorescence spectroscopy using a program provided by PTI. In particular, the excitation was between 250-500 nm and emission spectra were collected in both 300-550 and 500-750 nm.

2.2.8. Detect functional LBD-apoER2 in the time course media by Western blotting and ligand-blotting assay

2.2.8.1. Western- blotting

The recombinant LBD-apoER2 contains His-tag. Equal amount of the time course medium samples were subjected to SDS-PAGE under reducing conditions. Following transferring to the PVDF membrane, successive incubation with mouse anti-His tag monoclonal antibody (Sigma) and horseradish peroxidase (HRP)-conjugated secondary antibodies (Amersham Life Science) were carried out according to the manufacturer's specifications. The immunoreactive proteins were then detected using the enhanced chemiluminescence (ECL) system.

2.2.8.2. The ligand-blotting assay

The ligand-blotting assay can detect and quantify the properly refolded LBD-apoER2, while the bacterially expressed LBD-apoER2, which is in non-native conformation, should give a negative signal. The working principle of the ligand-blotting assay lies in the fact that the ligand-binding domain of apoER2 is fully functional on nitrocellulose under a physiological relative condition. If LBD-apoER2 is subjected to such procedure, the immobilized LBD-apoER2 can then be visualized and/or quantified following incubation with ligands that can be detected by standard Western blotting procedures.

The ligand used is the 39 kDa receptor-associated protein (RAP), which constitutes a non-lipoprotein ligand with high-affinity (K_D~1-10 nM) for the ligand-binding domain of all known members of the LDL receptor gene family (*174*). RAP is a chaperone for the LDL receptor family members. It associates with the

members of the LDLR by binding to newly synthesized and properly folded the ligand-binding domain of the receptors, preventing premature ligand-binding induced aggregation and subsequent degradation in the *ER* (*125*). RAP also escorts the receptors trafficking from the *ER* to the *Golgi* apparatus. In the ligand-blotting assay of LBD-apoER2, the advantage of using RAP is the high binding affinity between RAP and folded LBD-apoER2 that it is sufficient and easy for the detection by standard methods such as western blotting via the His tag of RAP. And the insensitivity of RAP to detergents, but not lipoproteins, allows detergent-containing washes to facilitate cleaner backgrounds.

Another advantage of the ligand-blotting assay is the whole procedure can be conducted under a physiological condition and there is no denaturants to disrupt the protein-protein interaction. Traditional far western blot assays need to go through steps such as SDS-PAGE gel and transferring step. Fractionating protein samples on an SDS-PAGE gel may cause the denaturation of protein samples. Transferring protein samples from the gel onto a solid support membrane such as PVDF or nitric cellulose membrane is assisted by the transfer buffer, which contains denaturant methanol and a relatively high voltage. The methanol in the transfer buffer and the heat produced while transferring may also denature the protein. These procedures may disrupt protein-protein interaction between the properly folded LBD-apoER2 and RAP by denaturing and produce false-negative results. Therefore, we skip the SDS-PAGE gel running step and the transferring step. We directly dot the PBS solution of LBD-apoER2 onto a nitric cellulose membrane (dot blot) and incubate with the probe RAP. Then we

apply standard Western blotting procedure to detect the bound RAP and indirectly detect our properly refolded LBD-apoER2 in the protein samples (Figure 2-7).

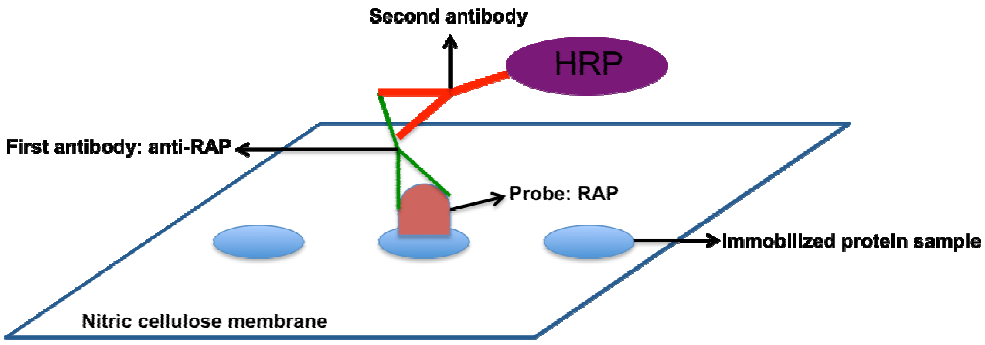

Figure 2-7: Ligand-blotting assay.

Three ul of samples was dot-blotted onto nitrocellulose membrane. After being dried for 30 minutes, the membrane was blocked in 3% non-fat milk PBS solution at room temperature for 1 hour with gentle agitation. For probing purposes, a 2 mg/ml RAP PBS solution was diluted in 10 ml 3% non-fat milk PBS solution to reach the final concentration of 20 ug/ml. After decanting the blocking buffer, the membrane reacted with RAP for 2 hours at room temperature. The membrane was then washed five times with 20 ml 1x PBS per 5 minutes at room temperature. The rabbit anti-RAP (Abnova) antibody was diluted in 3% non-fat milk PBS solution at the ration of 1:1000. The membrane was probed with rabbit anti-RAP antibody for 1 hour at room temperature with gentle agitation. The membrane was washed as described above. Next, the rabbit anti-RAP-antigen

complexes were detected with goat anti-rabbit IgG (HRP-conjugate) (1:10,000), followed by ECL detection. A replicated membrane probed with the secondary antibody alone did not display any signals (data not shown). Represenative data sets are shown. For presentation purposes, the squares containing each dot blot were arranged in a single column or a row for figures although they were derived from the same dot blot.

2.2.9. Refolding of LBD-apoER2 inside of HeLa cells

Five mg of bacterially expressed LBD-apoER2 was modified using 4 ml of QQ-reagent for overnight at 4 ℃ with gentle shaking. Modified LBD-apoER2 was centrifuged at 5000 rpm for 5 minutes to remove precipitation. Then the modified LBD-apoER2 was mixed with serum-free DMEM to reach the final volume and concentration of 50 ml and 0.1 mg/ml of the LBD-apoER2 and 1.2 mg/ml of the QQ-reagent, respectively. Ten flasks (75 cm^2) of 80% confluent HeLa cells were incubated with the protein delivery medium and protease inhibitor for 3-4 hours. After protein delivery, the protein delivery medium was removed and HeLa cells were gently washed three times with PBS. Then fresh serum-free DMEM medium was added for overnight (around 12 hours). The next morning, the overnight incubation serum-free DMEM medium, with the secreted and properly refolded LBD-apoER2 from HeLa cells, was collected and subjected to purification.

2.2.10. Purification of the refolded LBD-apoER2 from Hela cells

Protein purification of the folded LBD-apoER2 secreated from the HeLa cells was performed under native conditions without the use of denaturing reagents such as 6 M urea. The folded LBD-apoER2A was diluted into binding buffer that contained a high NaCl concentration (1M) to increase its solubility and to decrease the non-specific binding of other proteins to the nickel chelating column during purification.

The collected overnight incubation DMEM medium was 3-fold diluted with binding buffer (20 mM Tris-HCl, 1M NaCl and 10 mM imidazole at pH 8.0) and then loaded onto a His-bind resin column. The column was washed with another 100 ml of binding buffer followed by 100 ml of washing buffer (20 mM Tris-HCl, 1M NaCl and 20 mM imidazole at pH 8.0). The refolded LBD-apoER2 was finally eluted using 100 mM imidazole in a 200 ml volume. The eluate was dialyzed against 10 mM ammonium bicarbonate to remove imidazole and salts. After extensive dialysis, the elution was lyophilized. The pure refolded LBD-apoER2 powder was weighed and stored at -20 ℃.

2.2.11. Detection of refolded LBD-apoER2 by western-blotting

The refolded LBD-apoER2 protein powder was subjected to western blotting using an anti-His tag antibody (Sigma) and anti-apoER2 antibody (Abnova). Equal amount of the refolded LBD-apoER2 and bacterially expressed LBD-apoER2 were loaded onto SDS-PAGE gels. Samples were run on the gels

under both non-reducing and reducing conditions and then transferred to PVDF membranes; duplicate membranes were prepared. One PVDF membrane was incubated with mouse anti-His tag monoclonal antibody and the other membrane was incubated with mouse anti-apoER2 antibody. Both were then probed with HRP-conjugated secondary antibodies (Amersham Life Science) according to the manufacturer's specification. The immunoreactive proteins were detected by ECL.

2.2.12. Detection of the biological function of the refolded LBD-apoER2 by the ligand-blotting assay

The refolded and the bacterially expressed LBD-apoER2 protein powder were first dissolved in PBS buffer individually. Three ul of each were dotted onto the nitric cellulose membrane with RAP as the positive control. After drying the membrane for 30 minutes, the membrane was subjected to the ligand-blotting assay (see section 2.2.8.2. for details).

2.3. Results

2.3.1. Bacterial expression and purification of the recombinant LBD-apoER2

Using our high-yield *E.coli* protein expression system, large amount of recombinant LBD-apoER2 was produced. However, the recombinant LBD-apoER2 includes 42 cysteines. In the *E.coli* protein expression systems, these

cysteines cannot form proper disulfide bonds as seen in the native conformation of LBD-apoER2. They randomly form non-native intramolecular or intermolecular disulfide bonds. Most of the time, dimer, trimer or even higher oligomers were formed. Large amount of aggregates formed inclusion bodies inside *E.coli*. We used high concentrations of denaturant, such as 6 M urea, to extract the expressed recombinant LBD-apoER2. Due to the complicated conformations of the bacterially expressed recombinant LBD-apoER2, we normally could not detect the expected protein band at its molecular weight of 34 kDa on a 12% SDS-PAGE gel. Instead, we frequently detected a slower mobility protein band of approximately 50 kDa. This protein band was found only in the whole cell lysates of IPTG induced bacterial cells with absence of the reducing reagent, DTT (Figure 2-8 left panel, compare lanes 1 and 2). This suggested that this slower mobility band may be the LBD-apoER2. We performed Western blotting to confirm that the identity of this protein band (data not shown).

The bacterially expressed LBD-apoER2 protein powder was dissolved in PBS buffer at the concentration of 0.1 mg/ml. Unexpectedly, the band around 50 kDa seen in the whole cell lysates induced with IPTG (Figure 2-8 left panel, lane 1) was no longer detected, instead a diffuse band between 30 kDa and 40 kDa was now observed (right panel of Figure 2-8, lane 1). Changing the buffer condition or the drying step (lyophilization) could have affected the folding/disulfide linkages, thus affecting the mobility. Though the band was running around its expected molecular weight of 34 kDa, it does not mean the

bacterially expressed LBD-apoER2 was properly folded. The band was diffuse, suggesting it represents alternative folded states of the protein. Adding DTT

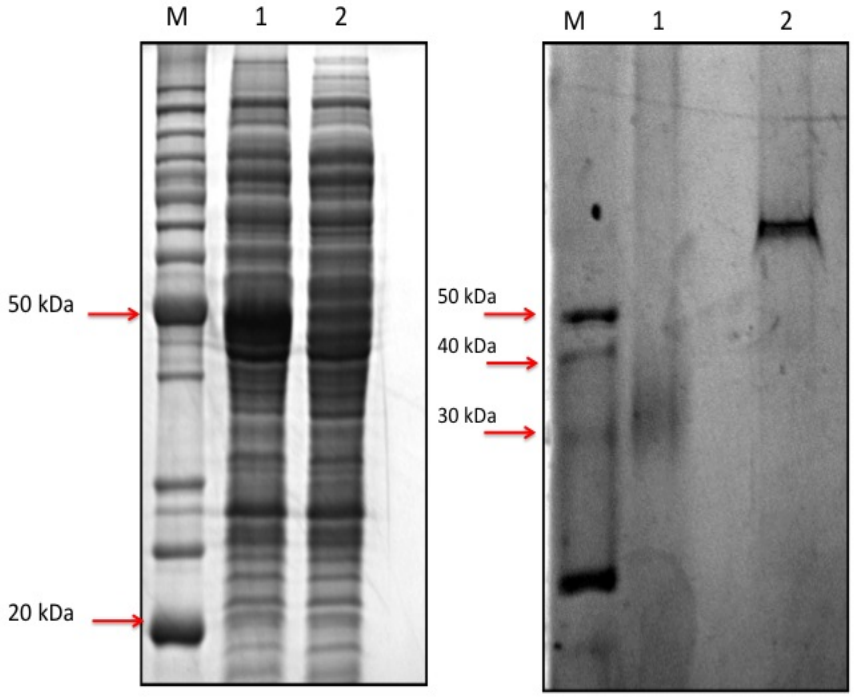

Figure 2-8: Expression and purification of bacterially expressed LBD-apoER2.

The left panel is the 12% SDS-PAGE stained with Coomassie blue, showing IPTG induced expression of LBD-ApoER2 in *E. coli*. Whole cell lysates (without DTT) induced with (lane 1) or without (lane 2) IPTG are shown. A protein molecular weight marker, lane M, was loaded with the 20 and 50 kDa bands pointed out with arrows for reference. The right panel is a Coomassie blue stained, 12% SDS-PAGE showing the eluted fraction from nickel chelating column that contains the purified LBD-ApoER2. Lane 1, the eluted fraction without DTT, showing a diffuse protein band at 34 kDa. Lane 2, the eluted fraction containing 10 mM DTT. The 30, 40, and 50 kDa bands of the molecular weight marker (lane M) are indicated for reference.

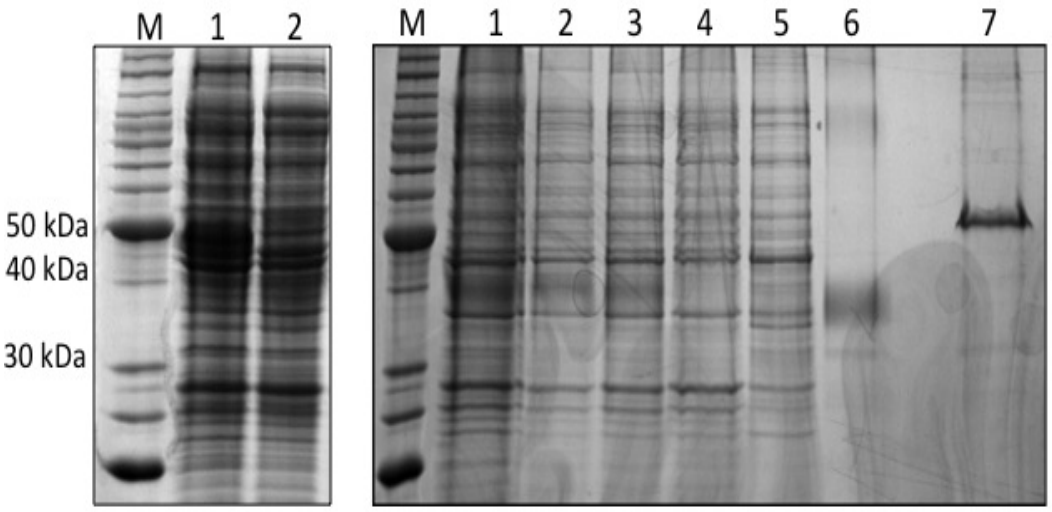

Figure 2-9: **Expression and purification of isotope-labelled, bacterially expressed LBD-apoER2.**

The left panel is the 12% SDS-PAGE showing IPTG induced expression of isotope labeled LBD-apoER2 in the minimal medium containing $^{15}NH_4Cl$, ^{13}C- glucose and D_2O. Lane 1, *E. coli* cell lysates following induction of protein expression. Lane 2, the *E. coli* cell lysates without IPTG induced expression as negative control. The right panel is the 12% SDS-PAGE showing the purification using His•Bind column. Lane 1, the 1st soluble fraction of inclusion bodies in 6 M urea. Lane 2, the 2nd soluble fraction of inclusion bodies in 6 M urea. Lane 3, the flow-through of the lysates after loading onto the nickel chelating column. Lane 4, the flow-through of binding buffer wash. Lane 5, the flow-through of washing buffer wash. Lane 6, the eluted fraction. Lane 7, the eluted fraction in a reaction mixture containing 10 mM DTT.

(100 mM) resulted in a single band of greater than 50 kDa (Figure 2-8, right panel, lane 2). DTT could have released the unstable disulfied bonds (diffuse character) in the protein and linearized the peptide chains, resulting in greatly

reduced mobility observed (Figure 2-8, Right panel, lane 2). We repeatedly found this characteristic of LBD-apoER2 in the SDS-PAGEs. Therefore, this mobility character of LBD-apoER2 with or without DTT on SDS-PAGE provided evidence for the bacterially expressed LBD-apoER2.

The recombinant LBD-apoER2 in the inclusion bodies of *E.coli* was very difficult to dissolve in the normal nickel chelating purification buffers. Six M urea was applied throughout the whole protein purification procedure and was removed through dialysis. A high yield of purified recombinant LBD-apoER2 was obtained after the lyophilizing step (50 mg/liter). In order to perform NMR studies of LBD-apoER2, we also prepared isotope labeled LBD-apoER2 (Figure 2-9). The protein produced the same character as in the unlabled protein (compare Figures 2-8, right panel, lanes 1 and 2 to Figure 2-9, Right panel, lanes 6 and 7). The yield of isotope labeled protein was lower than unlabeled protein; however, by using the high-yield protein expression system in *E.coli*, 60.04 mg of isotope labeled LBD-apoER2 was obtained for further experiments.

2.3.2. Highly efficient protein delivery by QQ-reagent

To refold bacterially expressed LBD-apoER2 inside HeLa cells, we delivered the bacterially expressed LBD-apoER2 into HeLa cells for *in vivo* refolding in the *ER*. Using QQ-protein delivery, we can deliver different proteins. However, due to the different biochemical properties of different proteins, an optimization step is necessary for an individual protein for efficient protein delivery.

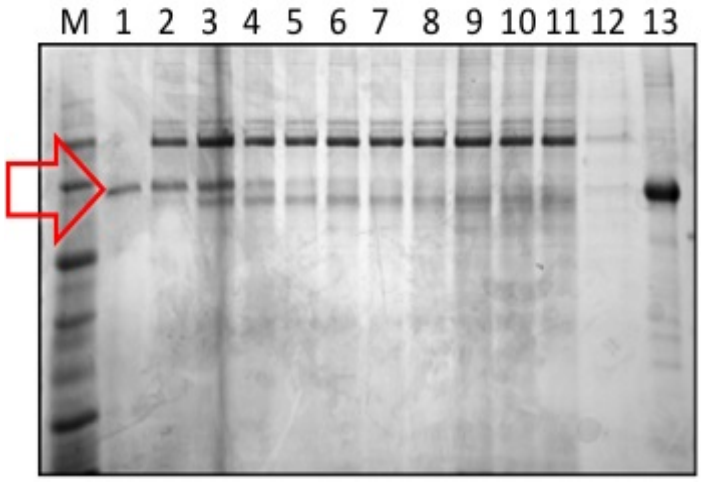

Figure 2-10: A Coomassie stained 12% SDS-PAGE showing a time course of the QQ-protein delivery of LBD-apoER2 into HeLa cells.

Bacterially expressed LBD-apoER2 was modified using QQ-reagent at the concentration of 0.1 mg/ml. HeLa cells were incubated with the protein delivery medium. Aliquots of the cell culture medium were taken at 0, 1, 2, 3, 4, 5, 6, 7, 8, 9 and 10 hours and then the protein delivery medium was exchanged for fresh serum-free DMEM medium for another 12 hours incubation. The protein samples run under reducing conditions are shown. Lanes 1-11 were the samples from 0 to 10 hours incubation. Lane 12, serum-free DMEM medium sample after another 12 hours incubation. Lane 13, the bacterial expressed LBD-apoER2 served as a positive control and positional control for the QQ-modified protein before delievery and is essential the same as lane 1 except that the protein in lane 13 is not modified with QQ-reagent.

We found that at the concentration of 0.1 mg/ml, the QQ-modified LBD-apoER2 could be delivered efficiently while HeLa cells maintained their healthy morphology. Higher concentrations were also tried. But either higher concentration of QQ-modified LBD-apoER2 or higher content of QQ reagent in the delivery medium introduced higher cellular toxicity, causing the rapid appearance of membrane blebs on the surface of most HeLa cells and subsequent apoptosis (data not shown).

The QQ-modified LBD-ApoER2 was delievered into Hela cells and samples of the cell culture medium containing both the remaining undelivered protein as well as secreted protein was collected hourly over a 10 hour peroid. All samples were subjected to SDS-PAGE. Under reducing conditions, focused and sharp protein bands of LBD-apoER2 were detectable on SDS-PAGE (Figure 2-10). The results showed that the protein band (red arrow) density of QQ-modified LBD-apoER2 was decreased after 3 hours (Figure 2-10, lanes 1-4), while a new faint band of faster mobility gradually appeared in the later time points (Figure 2-10, lanes 3-11). The slower mobility band that gradually dissipates is the original modified protein while the faster mobility band is likely the secreted LBD-ApoER2. After an additional 12-hour incubation with a regular cell culture medium that did not harbor modified protein, neither band could be detected and were likely beyond the limit of sensitivity for this method (Figure 2-10, lane 12).

This result demonstrated that 3 or 4 hours were the optimal protein delivery time for LBD-apoER2, as during this time the original protein band was

hardly detectable (Figure 2-10, lanes 4 and 5), suggesting that it has been taken up by the cell. This slower mobility band was replaced by the faster one at 4 hrs and thereafter (Figure 2-10 lanes 5-11) suggesting it to be the secreted protein. After an additional 12 hours of incuation, the secreted protein could not be detected.

After protein delivery, the properly refolded LBD-apoER2 should be secreted into the medium in the later time point including overnight incubation. The LBD-apoER2 only includes the ligand-binding domain of apoER2 with the trans-membrane domain removed. Without the transmembrane and cytosolic domains of apoER2, this mini-receptor will be secreted into the culture medium. We thought that the new lower faint band observed in the later time points was the refolded LBD-apoER2.

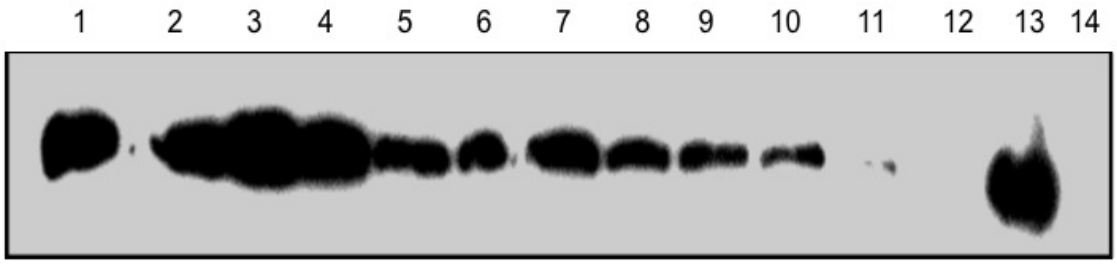

Figure 2-11: Western-blotting to detect His-tagged LBD-apoER2 during the same time course of QQ-delivery of LBD-apoER2.

An anti his-tag antibody was used. Lanes 1- 11 were the samples from 0 to 10 hours. Lane 12, serum-free DMEM medium sample after another 12 hours of incubation. Lane 13, the bacterially expressed LBD-apoER2 as a positive control. Lane 14, RAP full length without His-tag as a negative control.

To confirm the result of this time course, all the medium samples were subjected to Western blotting using an anti His-tag antibody with bacterially expressed LBD-apoER2 as a positive control and non-His-tagged RAP as a negative control (Figure 2-11). The result showed the same pattern of QQ-modified LBD-apoER2 in the delivery medium as the SDS-PAGE shown in Figure 2-10. At 3 or 4 hours, the density of LBD-apoER2 protein bands was sharply decreased and gradually disappeared by 10 hours after protein delivery (Figure 2-11, lanes 4-11). However, the secreated LBD-apoER2 in the later time points could not be detected by these Western blotting conditions (compare the trend of the faster mobility band in Figure 2-10 to Figure 2-11). This may be due to different conformations adopted by the refolded protein, whereby the His tag could be buried and inaccessible to the anti His-tag antibody. Our ligand-blotting data confirmed this suggestion (see 2.3.3). Up to this point, our data indicate that the QQ-modified LBD-apoER2 was efficiently delivered into HeLa cells and that the optimal LBD-apoER2 delivery time was 3 or 4 hours.

2.3.3. The LBD-apoER2 reaches to the ER after QQ-protein delivery.

Bacterially expressed LBD-apoER2 is misfolded, forming a mixture of different oligomers (Figure 2-12, panel A, lane1), which are inactive for ligand-binding. We designed specific experiments based on the following considerations: 1) The refolding process of LBD-apoER2 happens in the ER (*189*). 2) The refolding may occur in two steps: monomer formation and refolding.

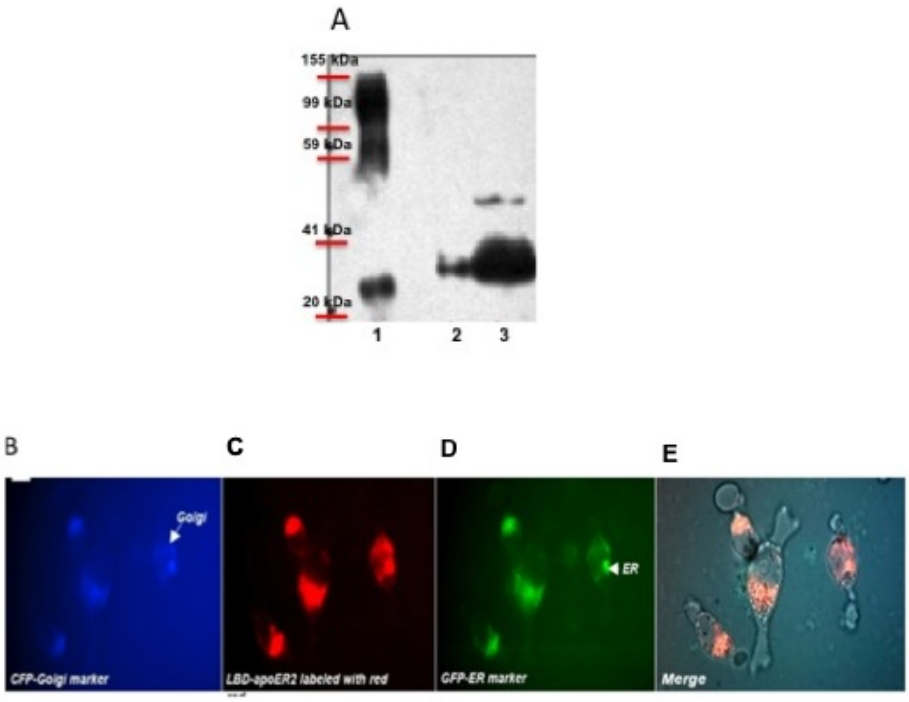

Figure 2-12: A: A Western blot of LBD-apoER2 using a 12% SDS-PAGE and an anti-His-tag antibody.

Lane 1: bacterial expressed LBD-apoER2 without DTT. Lane 2: the refolded LBD-apoER2 from HeLa cells. Lane 3: bacterial expressed LBD-apoER2 with 20 mM DTT. **Co-localization experiments:** B, fluorescence image of CFP channel (blue). C, fluorescence image of rhodamine channel (red). D, fluorescence image of FITC channel (green). E, Merge of the fluorescence images from all three channels with the light microscopy image (Courtesy of Dr. Qianqian Li).

Braakman *et al* previously suggested that the newly synthesized endogenous LDLR folded in a coordinated, non-vectorial pathway that included non-native disulfide bond formation and reshuffling (*189*). In our case, once the

bacterially expressed LBD-apoER2 is delivered into cells, the folding machinery has to break up the oligomers to form a monomer for refolding. The correctly refolded LBD-apoER2 should traffic from the ER to the Golgi, and eventually secrete into the cell culture medium via the plasma membrane.

We transfected HeLa cells with both CFP-Golgi and GFP-ER marker DNAs and incubated the transfected cells for 72 hours before protein delivery. The transfected DNA labels the ER with green and the Golgi with blue fluorescence (*116*). We also labeled LBD-apoER2 with DyLight 649 and delivered this red fluorescence labeled protein into the transfected cells. The cells were washed and incubated in a regular cell culture medium for another 6 hours and subjected to fluorescence imaging (Figure 2-12, panel B,C,D,E). Clearly, a merge of CFP (Panel B, Golgi location) with FITC (Panel D, ER location) and rhodamine channels (Panel C, LBD-apoER2 intracellular location) shows merging fluorescence (Panel E), demonstrating that the delivered LBD-apoER2 is in the ER and Golgi for refolding.

2.3.4. Optimizing the *in vivo* protein refolding protocol for LBD-apoER2

Our time course data suggested that the bacterially expressed LBD-apoER2 was efficiently delivered into the HeLa cells. The delivered, bacterially expressed LBD-apoER2 should be inside the *ER* for refolding, which takes time. The properly refolded LBD-apoER2 should follow the secretary pathway for secretion into the serum-free medium, while the persistent unfolded LBD-apoER2 would be degraded by the ERAD, since the transmembrane domain of LBD-

apoER2 was removed. This raised an important question as to when was the best time point to re-purify the properly refolded LBD-apoER2 from the cell culture medium for the highest yield of refolding.

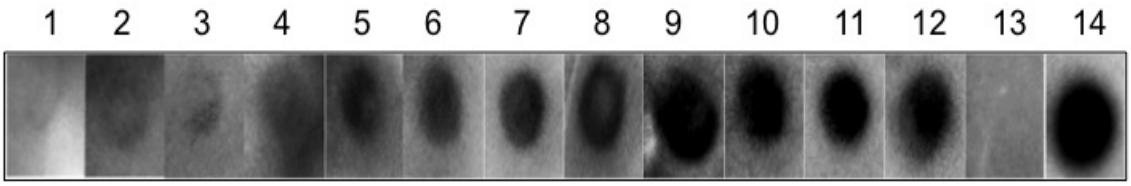

Figure 2-13: Ligand-blotting assay to detect RAP binding to the functional LBD-apoER2 after refolding.

Three ul of the medium were spotted onto a nitrocellulose membrane. The blot was incubated with 20 ug/ml full-length RAP after blocking. The bound RAP was probed by rabbit anti-RAP antibody followed by goat anti-rabbit IgG. Lanes 1-11 were the time points of 0 to 10 hours. Lane 12, serum-free DMEM medium sample after another 12 hours incubation. Lane 13, 0.1 mg/ml bacterial expressed LBD-apoER2 as a negative control. Lane 14, RAP full-length as a positive control.

We addressed this question using a ligand-blotting assay with these time course samples. Ligand-blotting can be used to specifically detect properly folded LBD-apoER2 due to its binding to RAP. While the bacterially expressed LBD-apoER2 is delivered into HeLa cells, their protein folding machinery will try to refold the delivered LBD-apoER2, and the properly refolded protein will be secreted from the cells. This is a continuous process. Our ligand-blotting assay

result (Figure 2-13) showed that the refolded LBD-apoER2 was indeed secreted into the medium and gradually accumulated in the medium. At 0-3 hours, nearly no positive dot-blot signals were observed (lanes 1-4), which was confirmed by the negative control of Lane 13, the bacterially expressed LBD-apoER without *in vivo* refolding. As time proceeded, significant intensity of the dot-blot signals were observed at 4-7 hours (lanes 5-8). At 8-hours, the strongest signal was observed which was maintained all the way to 10-hours (lanes 9-11). In addition, strong dot-blot signals were also observed in the medium of another overnight cell culture (lane 12). This data indicated that the cell culture medium during this time course contained the secreted and properly refolded LBD-apoER2 and the refolded LBD-apoER2 reached the highest concentration at 8-hour after QQ-delivery. In addition, the amount of the refolded protein was continuously accumulated in the cell culture medium during an overnight cell culture with QQ-protein delivery.

Although the refolded LBD-apoER2 could not be detected in the Coomassie stained gel after an additional overnight incubation (Figure 2-10 lane 12), the dot blot clearly demonstrated that the refolded protein was indeed present in the overnight medium (Figure 2-13, lane 12). As the overnight medium did not harbor any QQ-modified protein, coupled with the lack of signal from the undelivered bacterially expressed protein (Figure 2-13, lane 13), the similar trend between the fast mobility band (Figure 2-10, lanes 3-11) and the RAP dot blot (Figure 2-13) strongly argue that the faster mobility band observed

(Figure 2-10) is likely the functional and refolded LBD-apoER2, being responsible for the RAP binding (Figure 2-13).

We quantified the ligand-blotting data using the Kodak 1D image analysis software. As shown in Figure 2-14, the amount of the refolded LBD-apoER2 kept increasing until the 8 hours time point where it reached its peak. After 8 hours, the amount of the refolded LBD-apoER2 in the incubation medium remained constant.

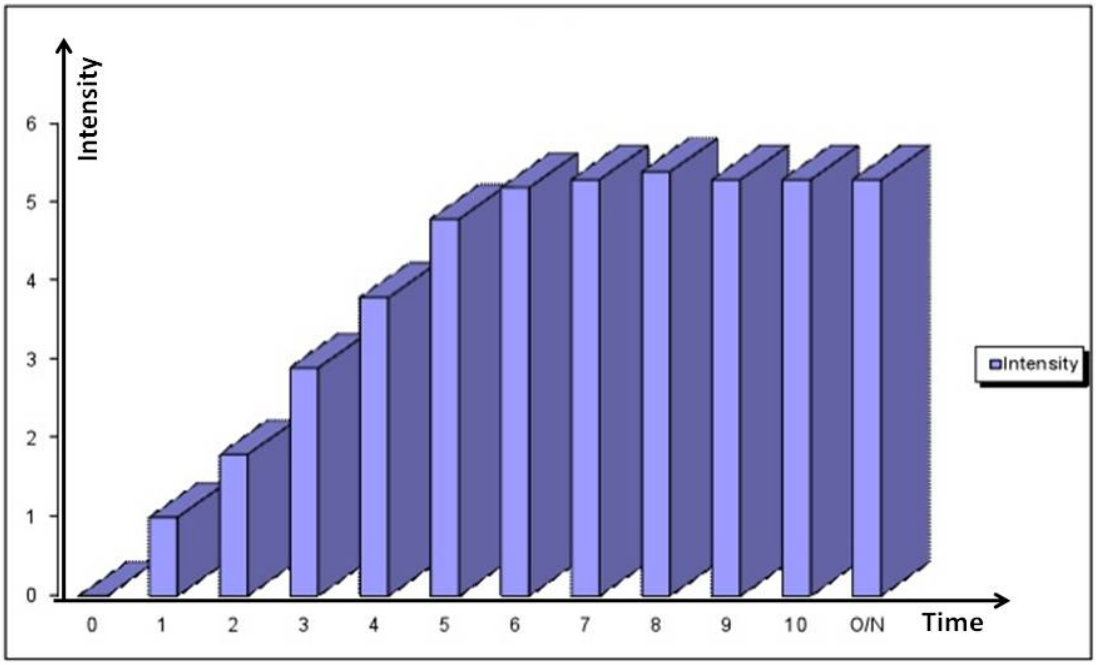

Figure 2-14: The bar graph of signals quantified from the ligand-blotting assay (Figure 2-13), showing the accumulation of the secreted and refolded LBD-apoER2 in the serum-free medium.

Since this *in vivo* refolding is a complicated process, containing protein refolding, ERAD degradation of consistently misfolded protein, and secretion of the folded protein, it is difficult to identify the best timing to obtain the highest yield of the refolded protein. However, our data clearly indicated that the secretion of the refolded LBD-apoER2 peaked at 8 hours after protein delivery. Eight to 10 hours incubation time for *in vivo* refolding was the optimal condition.

2.3.5. Refolded LBD-apoER2 was secreted into culture medium.

For LBD-apoER2 *in vivo* refolding, our optimal experimental condition was 4-hour QQ-protein delivery, followed by an additional 8-10 hours for protein refolding and secretion (Figure 2-15). Under this condition, most bacterial recombinant LBD-apoER2 was delivered inside HeLa cells by the QQ-protein delivery. The delivered LBD-apoER2 was refolded by the ER refolding machinery and most refolded LBD-apoER2 was secreted into the cell culture medium. We focused on cell culture medium to take advantage of the secreted LBD-apoER2 since it should be properly refolded and relatively more pure as compared to the LBD-apoER2 inside HeLa cells that might contain mixed population of refolded and misfolded LBD-apoER2.

To confirm this, medium samples at different time points were taken during protein refolding, and HeLa cells were also lysed afterwards. All samples were subjected to a SDS-PAGE with or without DTT. HeLa cells were also incubated with the protein delivery medium that only contained QQ-reagent

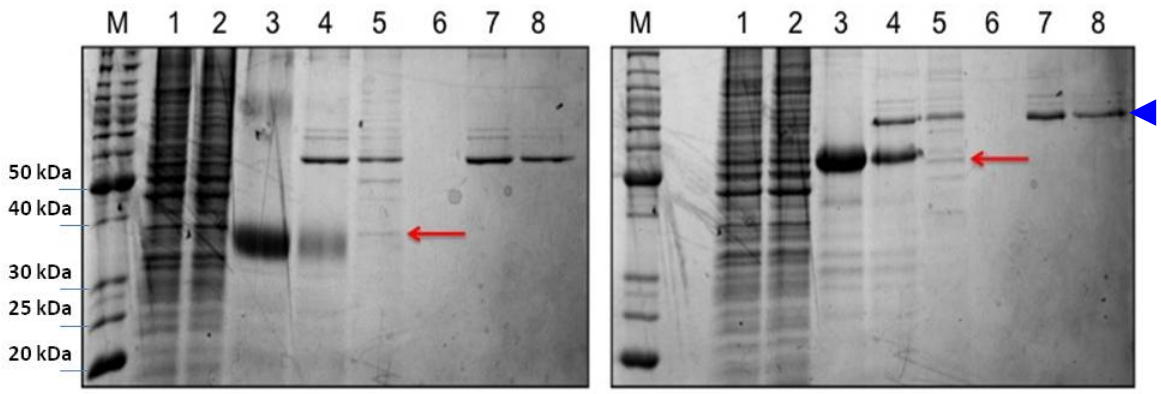

Figure 2-15: 12% SDS-PAGEs, showing the *in vivo* protein refolding.

HeLa cells were incubated with the protein delivery medium for 4 hours. The protein delivery medium was exchanged to fresh serum-free DMEM medium without QQ-LBD-apoER2 and cultured for another 10 hours. The samples were subjected to SDS-PAGE analyses without (left) and with (right) DTT. Lane 1, the HeLa cell lysates after 4-hour protein delivery and 10-hour protein refolding. Lane 2, the control HeLa cell lysates without protein-delivery. Lane 3, the protein delivery medium with 0.1 mg/ml bacterial expressed LBD-ApoER2 before adding HeLa cells. Lane 4, the protein delivery medium after 4 hours post protein delivery. Lane 5, the culture medium after 10 hours of protein refolding. Lane 6, the cell culture medium with QQ-reagent without adding HeLa cells as a negative control. Lane 7, the control HeLa cell culture medium with the QQ-reagent only after 4 hour culture. Lane 8, the serum-free DMEM medium after another 10 hour with the control HeLa cells. The blue arrow head indicates a secretion protein from HeLa cells. The faint yet properly refolded LBD-ApoER2 band that is secreated into the medium is noted by a red arrow.

under the same experimental condition. The cell culture medium of this control experiment were also taken as a negative control.

Consistent with our previous data, LBD-apoER2 was not detectable in the HeLa cell lysates (lane 1) and was only detectable in the cell culture medium (lane 5) (Figure 2-15). This result confirmed the suggestion that protein refolding did not lead to the accumulation of the LBD-apoER2 inside HeLa cells. Instead, the properly refolded LBD-apoER2 was secreted into the cell culture medium. In addition, our data indicated that the QQ-protein delivery method efficiently sent the LBD-apoER2 inside HeLa cells since the LBD-apoER2 band intensity was significantly decreased (lane 4). Accordingly, Lane 5 showed detectable refolded LBD-apoER2 after 10-hour cell culture (red arrow). The higher molecular weight bands shown on the gels (lanes 4, 5, 7, and 8, Blue Arrow Head) came from a HeLa cell secretary protein, which persistently existed in the medium during *in vivo* protein refolding. This result was confirmed by the negative controls on the SDS-PAGE gels (lanes 6-8).

2.3.6. The yield of refolded LBD-apoER2

The incubation medium (serum-free) was collected after refolding and the properly refolded LBD-apoER2 was subjected to protein purification using the nickel chelating column. Our results showed that the biochemical properties of the properly folded LBD-apoER2 were different from the bacterially expressed protein. The solubility of the folded LBD-apoER2 was lower and the binding

capability to the nickel chelating column was decreased. This is consistent with the possibility that the his-tag may be partially occluded in the refolded protein and could be responsible for the lack of signal in the Western (Figure 2-11) and the decreased binding observed despite the his-tag being present. This forced us to optimize the purification protocol for the refolded LBD-apoER2. Three main modifications were carried out: (1) the volume of purification buffer was increased to reduce the refolded LBD-apoER2 concentration and to decrease potential possibility of precipitation; (2) high concentrations of NaCl in the purification buffer were included in order to avoid non-specific protein binding to the His•Bind column and to increase the solubility of the refolded LBD-apoER2; (3) decreased concentrations of imidazole in the elution buffer because we found that high concentrations of imidazole would cause irreversible aggregation and precipitation of the refolded LBD-apoER2 during dialysis. The lower concentration of imidazole was sufficient to elute the refolded protein due to the weaker interaction with the column resin.

Even though we optimized protein delivery, protein refolding, and protein purification conditions throughout this *in vivo* protein refolding technique for LBD-apoER2, the final yield of the purified refolded LBD-apoER2 remained low. We obtained a yield less than 10% of the refolded LBD-apoER2 compared to the imput amount (QQ modified, bacterially expressed protein used for delivery) as shown in Table 2-1.

This is not surprising since LBD-apoER2 contain 21 disulfide bonds and 7 Ca^{2+}-binding sites. The folding of this protein is challenging. It has been

demonstrated by others that only 30% of the newly synthesized LDLR was properly folded and the rest proteins were degraded by ERAD due to misfolding (*189*). Our data confirmed this result for our refolding, since we could only recover less than 10% of the bacterially expressed LBD-apoER2 (Table 2-1). This suggested that most of the QQ-delivered LBD-apoER2 was degraded due to the limits of the ER folding capability. We re-purified the bacterially expressed LBD-apoER2 from the protein delivery medium after 4 hours of delivery. The re-purified bacterially expressed LBD-apoER2 protein could be re-used in the next round of *in vivo* protein refolding experiments to increase the refolding efficiency. Those consistently misfolded LBD-apoER2 were subjected to ERAD degradation due to ER quality control system. Another possible reason for this low yield was the purification due to the reduced affinity of the refolded LBD-apoER2 for the His-tag resin. After refolding, LBD-apoER2 underwent conformation changes that might make the his-tag buried and not available to bind to his-tag resin, causing a low yield of the refolded LBD-apoER2.

Table 2-1: Three representative yields of refolded LBD-apoER2.

Refolded LBD-apoER2 protein powder	Refolded percentage	Recovered bacterial expressed LBD-apoER2 protein powder	Total bacterial expressed LBD-apoER2 in the protein delivery medium
0.36 mg	7.2%	0.71 mg	5 mg
0.49 mg	9.8%	0.73 mg	5 mg
0.24 mg	6%	0.84 mg	5 mg

2.3.7. The refolded LBD-apoER2 is biologically functional

To further confirm the purified refolded LBD-apoER2 powder is properly folded and biologically functional, we dissolved the protein powder in PBS buffer with 1 mM $CaCl_2$ and performed Western blot and ligand blotting. As shown in Figure 2-16 A-C, with (lanes 3 and 4) or without reducing reagent DTT (lanes 1 and 2), the refolded LBD-apoER2 (lanes 1 and 3) showed the same mobility as the bacterially expressed one (lanes 2 and 4) on the 12% SDS-PAGE gel (Panel A). Western blotting (Panel B and C) showed positive signals for the refolded LBD-apoER2 as well as the bacterially expressed one, identifying the protein powder was LBD-apoER2. Interestingly, the LBD-ApoER2 antibody only recognized the protein when it was under reducing conditions (panel B, lanes 3 and 4) suggesting that the epitope is hidden in the refolded (lane 1) or misfolded proteins (lane 2). However, the His tag remained accessible as its was detected in all lanes (panel C). Likely, the properly refolded LBD-ApoER2 was not detected in Figure 2-11 due to its trace amount but was clearly detected here (Figure 2-16, panel C). Ligand-blotting assay (panel D) of the refolded LBD-apoER2 protein powder showed a positive signal, similar to the RAP dot-blot as the positive control. However, the bacterially expressed LBD-apoER2 protein powder showed a negative signal. This result confirmed that the refolded LBD-ApoER2 powder displayed a proper conformation that bound to RAP. Thus, the refolded LBD-apoER2 protein powder is biologically functional unlike the bacterially expressed one.

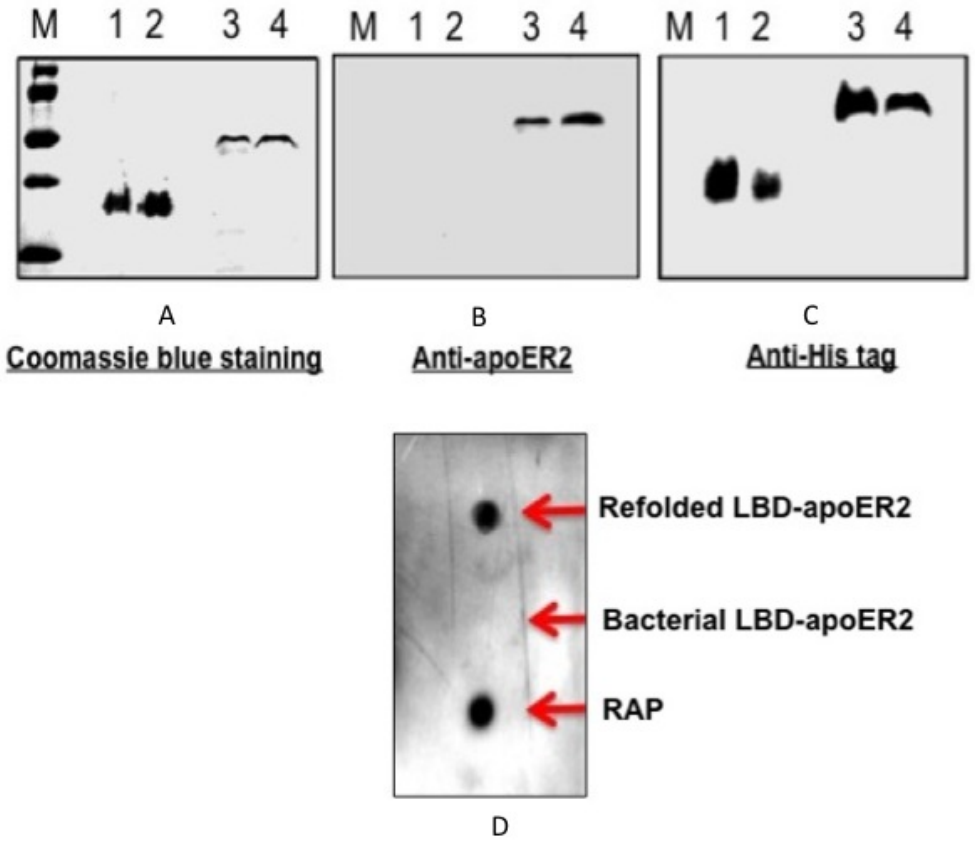

Figure 2-16: SDS-PAGE, Western blots using anti-apoER2 and anti his-tag antibodies and ligand blotting of the refolded LBD-apoER2 protein powder.

Panel A: a 12% SDS-PAGE stained with Coomassie blue. Panel B: A Western blot using anti-apoER2 antibody. Panel C: A Western blot using anti His tag antibody. In all three panels, Lane 1, the refolded LBD-apoER2 powder. Lane 2, bacterial LBD-apoER2 powder. Lane 3, bacterial LBD-apoER2 powder with 10 mM DTT. Lane 4, the refolded LBD-apoER2 powder with 10 mM DTT. Panel D: Ligand-blotting to detect the binding of RAP to the functional LBD-apoER2. The bound RAP was probed by anti-RAP antibody.

2.4. Discussion

Proteins have to adopt the correct three-dimensional structures in order to be functional**.** Obtaining the native fold of a protein sometimes is not a simple task, especially when the protein is expressed in *E. coli* to meet the need of large quantity of properly folded proteins for structural study or for therapeutics. Many proteins precipitate in the form of inclusion bodies when synthesized in *E. coli. In vitro* experimental approaches have been developed to refold the aggregated proteins into their biologically active forms. The main refolding strategies are chromatographic methods and refolding in free solution. New techniques employing additives that aid in protein refolding, such as chaperone proteins, may become more widely used. However, these new methods require rigorous testing for refolding of not only simple model proteins such as RNase A, but also for proteins with complex folding. New techniques of protein refolding are needed to increase the robustness of refolding processes and to decrease the costs for broader applications.

Our lab recently developed a novel *in vivo* protein refolding technique. Different from other *in vitro* protein refolding techniques, this protein refolding technique enables protein refolding inside of mammalian cells and is facilitated by folding chaperones and enzymes in the *ER* and the *Golgi.* The ER quality control system ensures that only properly refolded proteins continue intracellular trafficking whereas the persistent unfolded protein is degraded by ERAD or UPR (Figure 1-2 and 1-3). Properly folded and biologically functional proteins follow the same intracellular trafficking pathway as endogenous ones. Additionally, the

refolded protein can be purified from mammalian cells or the culture medium depending on the design and properties of the refolded proteins.

This *in vivo* protein refolding technique is based on a highly efficient recombinant protein expression method in *E. coli* and the QQ-protein delivery technique, which were recently developed in our lab. Specially, the unique QQ-reagent can non-covalently attach to the bacterially expressed proteins, protect the proteins from degradation, and efficiently deliver them inside of mammalian cells. The delivered proteins traffic to the *ER* and *Golgi* and undergo efficiently protein refolding by the ER folding machinery. Thus, this *in vivo* protein refolding technique utilizes ER protein folding machinery and quality control system to efficiently refold proteins and to ensure that the misfolded proteins will be degraded, thus, providing a novel protein refolding technique that is currently unavailable for the scientific community.

ApoER2 is one of the primary members of LDL receptor family, playing an important role in lipid metabolism in the brain and neurodegenerative diseases such as Alzheimer's disease. Like its homologue - LDL receptor, apoER2 has five major motifs: complement-like binding repeats/the LDL receptor type-A repeats, epidermal growth factor (EGF)-like repeats, YWTD repeats, a transmembrane domain and a short cytoplasmic tail (Figure 2-1). Studies showed that the seven complement-like repeats form a ligand-binding domain, mainly responsible for binding to the ligands such as apoE, RAP and Reelin (*161, 174, 190*). Even though no structural information of apoER2 has been reported, structure studies on other LDL receptor family members like LDL

receptor and LRPs have shown that the each ligand binding repeat in the ligand-binding domain contains six cysteines, forming three disulfide bonds per repeat (*190*). The negatively charged amino acids (Asp/Glu) at the C-terminus of each repeats are considered to be important for the ligand binding function, partially complementary to the positively charged residues on apoE or RAP (*190*). Also, these acidic residues form an important calcium coordination site, which is highly conserved. Calcium was reported to be important for receptor folding, stabilizing the secondary structure and ligand binding (*191*). Thus, the ligand-binding domain of LDLR family has a complex folding and refolding of this complex structure is very challenging.

We undertook this challenge and explored efficient refolding of the human LBD-apoER2 using this *in vivo* protein refolding technology. We believe that this is one of the ultimate measures for this *in vivo* protein refolding technology and the successful refolding of LBD-apoER2 will be a good testimony of the enabling capability of this technology. We subcloned the ligand-binding domain of apoER2 (LBD-apoER2) into a high-level expression vector which was expressed in *E. coli* using our high cell-density expression method to produce gram quantity pure recombinant LBD-apoER2 (*59*). We then optimized the *in vivo* protein refolding conditions specific for the efficient refolding of LBD-apoER2. We showed that QQ-modified bacterially expressed LBD-apoER2 could be efficiently delivered inside HeLa cells in 3 to 4 hours at 0.1 mg/ml concentration (Figures 2-10). Once the delivered LBD-apoER2 reaches the *ER* (Figure 2-12), the protein folding process begins. The bacterially expressed LBD-apoER2 undergoes

protein refolding process, which is facilitated by chaperones and folding enzymes in the *ER*. The quality of refolded protein is monitored by the QC system in the *ER.* Persistent unfolded LBD-apoER2 is degraded by ERAD or UPR. The properly folded LBD-apoER2 follows the same secretary pathway as endogenous counterpart and is secreted into the cell culture medium if it is a secretion protein.

Because LBD-apoER2 only includes the ligand-binding domain with the transmembrane and cytosolic domains removed, this mini-receptor will not locate on the cell membrane as the full-length apoER2. Instead, LBD-apoER2 will be secreted into the culture medium once they are properly refolded. Our data of a time course of refolding indicated that the refolded LBD-apoER2 was gradually accumulated in the culture medium with the peak of accumulation at 8-10-hours after protein delivery. After reaching the peak, the level of the refolded LBD-ApoER2 in the medium remained the same during this 10-hour time course. Thus, our data suggested that the refolding of the QQ-delivered LBD-apoER2 might be nearly completed by the ER folding machinery under our experimental condition. This allowed us to optimize the *in vivo* protein refolding protocol for LBD-apoER2 by QQ-protein delivery for 4-hours, followed by additional 8-10-hours of protein refolding in the regular cell culture medium for the best yield of the refolded LBD-apoER2.

Purification of the refolded LBD-apoER2 from the cell culture medium turned out to be a challenging task since the refolded LBD-apoER2 seemed to significantly reduce its binding capability to His-tag resin, thus causing a major

trouble to purify them. However, this is one of the key steps for the final yield of the refolded LBD-apoER2 protein. We had to optimize our standard purification protocol by increasing the sample loading volume and NaCl concentration in the loading and washing medium and reducing the imidazole concentration in the elution buffer. These modifications greatly improved binding capability of the refolded LBD-apoER2 to His tag resin that significantly enhanced the purity and yield of the refolded LBD-apoER2.

We probed the function of the refolded LBD-apoER2 by a ligand-blotting assay using RAP as the ligand and an anti-RAP antibody. RAP acts as a specialized chaperone inside the *ER* and binds to properly folded and functional ligand-binding domains of the LDLR family including apoER2 with nM binding affinity (*174*). No denaturing reagents or detergents were used in this ligand-blotting assay. However, 1 mM calcium ion was added since calcium was critical for RAP-binding to LBD-apoER2. The pH was adjusted to 7.2 to mimic the pH inside the *ER*. Our data clearly demonstrated that the properly refolded LBD-ApoER2 was biologically active for RAP binding, confirming the efficiency of this novel *in vivo* protein refolding technology.

This *in vivo* protein refolding technique is unique compared to the other *in vitro* protein refolding techniques. It takes advantage of the efficiency of the ER folding machinery of mammalian cells. By providing a physiological environment, the bacterially expressed misfolded LBD-apoER2 was efficiently refolded by the ER folding chaperones and enzymes. In addition, the refolding quality of LBD-apoER2 is closely monitored by the quality control system in the ER. Only the

properly refolded protein can be further subjected to post-translational modifications and eventually secreted into cell culture medium. In contrast, the persistent unfolded LBD-apoER2 is degraded by ER-associated degradation or unfolded protein response mechanism in the *ER*.

This was confirmed by the previous data of my colleague Dr. Qianqian Li, who delivered the bacterially expressed MESD inside mammalian cells (data not shown). After purification, she found two major glycosylation forms of MESD in addition to the native MESD protein. The glycosylation patterns were the same as those of the endogenous MESD. This is an important achievement since properly posttranslationally modified and refolded proteins are important to the biological activity of a protein, which is especially critical towards the production of therapeutic proteins for human disease therapies. Our results suggest that we have successfully achieved this goal for refolding of LBD-apoER2.

Chapter III

Refolding of the beta-propeller/epidermal growth factor-like (EGF) domain I of low-density lipoprotein (LDL) receptor-related protein 6 (BP1-LRP6) inside of mammalian cells

3.1. Introduction

3.1.1. LRP5/6 are members of the low-density lipoprotein receptor (LDLR) family with unique structural arrangements

The LDLR family contains more than 14 members of cell surface receptors. Although the members in this family performs diverse functions, they share common structural domains: 1) ligand-binding domain of the complement-type (CR)repeats, 2) epidermal growth factor (EGF) -like repeats, 3) YWTD (Tyr−Trp−Thr−Asp) β-propeller domains, 4) a transmembrane domain, and 5) one or more endocytic motifs within their cytoplasmic domains (Figure 2-1) (*192*). LDL receptor-related protein 5 (LRP5) and LRP6 are structurally diverse from the other LDLR family members. Rather than having large clusters of ligand binding CR repeats in the extracellular region, LRP5/6 contain only three ligand binding CR repeats that are next to the transmembrane region (Figure 2-1). Instead, the YWTD β-propeller/EGF (BP) modules occupy most of the extracellular domain and play important functional roles. Ligands of other LDLR members typically bind to clusters of ligand-binding CR repeats, whereas all identified ligands for LRP5/6 bind to the YWTD β-propeller/EGF (BP) modules (*192*). Evidence implicating YWTD domains in protein−protein interactions has emerged from several studies. LRP6 from xenopus along with its homologs in mice and the fly

(*Drosophila*) binds to the developmental signaling protein Wnt, implicated as a co-receptor in Wnt signaling (*141-143*).

The LDLR family members play important roles in diverse physiologic progresses, including lipoprotein metabolism, protease regulation, calcium homeostasis, cell migration, and embryonic development (*193*). For example, VLDLR and ApoER2 mediate Reelin signaling in neuronal cell migration during embryonic development (*140*), and LRP1 participates in PDGF-mediated signaling (*194*). However, LRP5 and 6 were identified as indispensible co-receptors of the canonical Wnt/β-caternin pathway and function primarily in Wnt signaling pathway (*141, 142*).

3.1.2. LRP5/6 are essential receptors for the canonical Wnt signaling pathway

The discovery of the common evolutionary origin of the *Drosophila* segment polarity gene *Wingless* and the murine proto-oncogene *Int-1* laid the foundation for the canonical Wnt signaling pathway (*195*). The name Wnt was coined as a combination of Wg (*Wingless*) and *Int.* The Wnt signaling pathway plays a critical role in orchestrating proper tissue development in embryos and tissue maintenance in adults (*196-198*). It has been shown that Wnt activates at least three intracellular signaling pathways including the Wnt/β-catenin (canonical Wnt pathway), planar cell polarity (PCP), and Ca^{2+}pathway (Figure 3-1) (*199*). In the PCP pathway, Wnt activates small G protein including Rho and Rac through

Dishevelled (Dsh), thereby activating Rho-associated kinase (ROCK) and JNK (*200*). The PCP pathway is involved in the regulation of tissue polarity, cell migration, and cytoskeletal arrangement. Furthermore, Wnt increases the intracellular Ca^{2+} concentration probably by activating protein kinase C (PKC) and Ca^{2+}/calmodulin-dependent protein kinase. The Wnt/Ca^{2+} pathway regulates cell adhesion and cell movement during gastrulation (*201*).

The canonical Wnt pathway, which regulates the ability of the β-catenin protein to drive activation of specific target genes, is better characterized. In the absence of a Wnt signal, β-catenin is actively degraded in the cell by the actions of a multiple-protein complex called the “destruction complex”. Within the complex the Axin and adenomatous polyposis coli (APC) form a scaffold that facilitates β-catenin phosphorylation by casein-kinase I α (CK1 α) and glycogen synthase kinase (GSK3 β). The phosphorylated β-catenin is subsequently recognized and ubiquitinylated, resulting in its proteasomal degradation (Figure 3-2) (*202*). However, in the presence of Wnt signals, β-catenin can be accumulated and be translocated to the nucleus, where it interacts with the T-cell factor/lymphoid enhancing factor (TCF/LEF) family of transcription factors, leading to the transcription of Wnt target genes (*202*). The Wnt/ β-catenin signaling pathway directs a specific set of genes that strictly control temporal and special regulation of cell proliferation, differentiation, survival, and movements (*197, 198*). Therefore, aberrant activation of this pathway can lead to tumor formation.

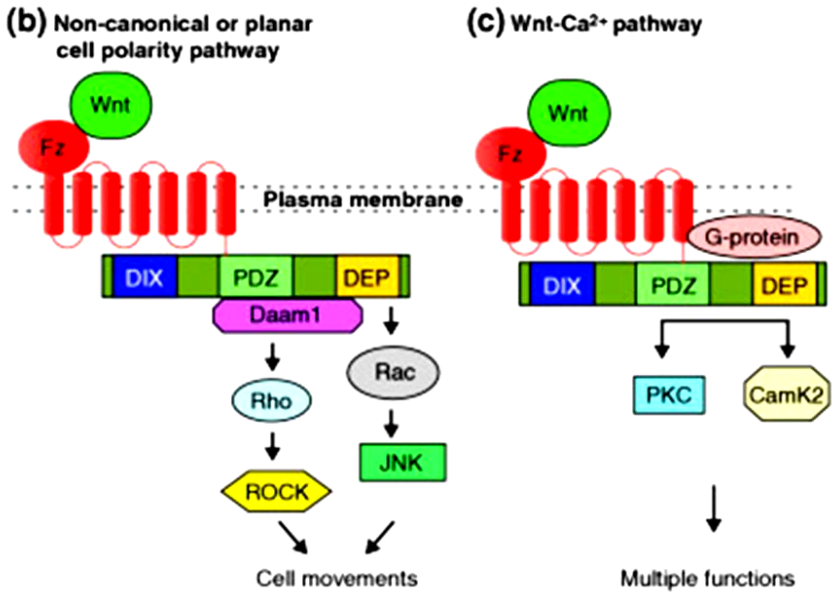

Figure 3-1: A schematic representation of the Wnt signal transduction cascade (*199*).

(b) For non-canonical or planar cell polarity (PCP) signaling, Wnt signaling is transduced through Frizzled independent of LPR5/6. Utilizing the PDZ and DEP domains of Dsh, this pathway mediates cytoskeletal changes through activation of the small GTPases Rho and Rac. **(c)** For the Wnt-Ca^{2+} pathway, Wnt signaling via Frizzled mediates activation of heterotrimeric G-proteins, which engage Dsh, phospholipase C (PLC; not shown), calcium-calmodulin kinase 2 (CamK2), and protein kinase C (PKC). This pathway also uses the PDZ and DEP domains of Dsh to modulate cell adhesion and motility. Note that for the PCP and Ca^{2+} pathways Dsh is proposed to function at the membrane, whereas for canonical signaling Dsh has been proposed to function in the cytoplasm

LRP5 and LRP6 exhibit overlapping, as well as distinct, tissue- and cell-type-specific expression patterns (*192, 203*). Mice deficient in *Lrp6* gene are embryonic lethal; the mutant embryos exhibit developmental defects that resemble the phenotypes observed for individual Wnt ligand mutants, including mid/hindbrain defects, posterior truncation, and abnormal limb patterning (*141*). In contrast, LRP5 knock out (KO) mice are viable and fertile, although they exhibit low bone mass and eye vascularization defects (*152*). Interestingly, it has been shown that overexpression of LRP6 confers significantly stronger Wnt signaling activity than LRP5 in *Xenopus* and mammalian cells (*142, 192*). Thus, it is possible that LRP5 and LRP6 exhibit functional redundancy, most likely by acting with different efficiencies in a context-dependent manner. The features of LRP5-KO mice recapitulate human osteoporosis-pseudoglioma syndrome, whereas gain-of-function mutations in *Lrp5* affects the homeostatic balance of osteoblasts and osteoclasts after birth, which leads to high bone mass (*152, 204*). Recently, LRP6 mutations and polymorphisms were found to be associated with various human diseases, including coronary artery disease, Alzheimer's disease, hyperlipidemia, and osteoporosis (*192, 205*). These studies indicate a possible predominant role of LRP5 in bone development and maintenance and a broader role of LRP6 in pathophysiology.

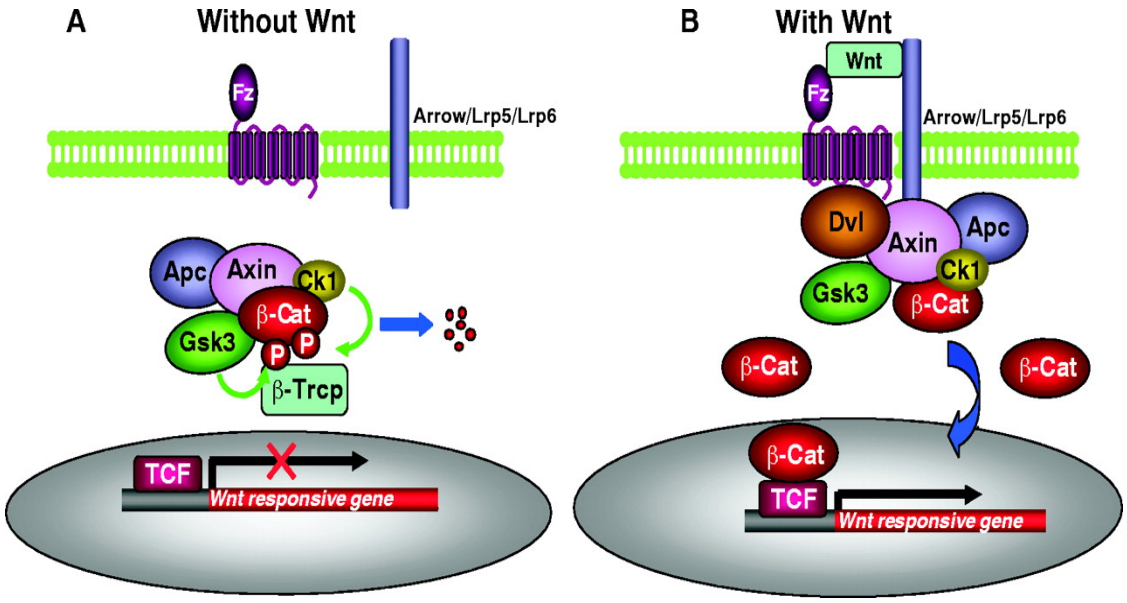

Figure 3-2: A simplified classical view of Wnt/β-catenin signaling (*192*).

(A) Without Wnt, the scaffolding protein Axin assembles a protein complex that contains Apc, Gsk3, Ck1, and β-catenin. In this complex, β-catenin is sequentially phosphorylated by Ck1 and Gsk3. Phosphorylated β-catenin is recognized by β-Trcp, which is a component of an ubiquitin-ligase complex that conjugates β-catenin with ubiquitin. Poly-ubiquitinated β-catenin is degraded by the proteosome. TCF/LEF-associated co-repressors, such as Groucho, and Axin-associated Diversin, PP2A and other proteins are omitted for simplicity. (B) In the presence of Wnt, β-catenin phosphorylation and degradation is inhibited. Accumulated β-catenin forms a nuclear complex with the DNA-bound TCF/LEF transcription factor, and together they activate Wnt-responsive genes. This signaling cascade is perhaps initiated by a Wnt-induced Fz-Lrp5/Lrp6 co-receptor complex, which recruits Axin to the plasma membrane through Lrp5/Lrp6-Axin association. Fz-associated Dishevelled (Dvl) protein may bind Axin and inhibit Axin-Gsk3 phosphorylation of β-catenin, either directly or indirectly

via Dvl-associated proteins. Lrp5/Lrp6-Axin binding may also promote Axin degradation. Either or both of these events can lead to β-catenin accumulation. This description represents one of several possibilities. The composition of the Axin complex upon Wnt stimulation is not well defined. Gsk3-binding protein (GBP/Frat), and nuclear β-catenin-associated Legless/Bcl9 and Pygopus are omitted for simplicity.

3.1.3. The YWTD β-propeller/EGF (BP) domain of LDL receptors and the molecular chaperone MESD

The number and arrangement of different LDLR domains varies greatly among different members of LDLR family; in humans, sizes range 5-fold from the smallest LDLR (95 kDa) to the largest LRP2 (522 kDa) (*206, 207*). Such large and complex proteins challenge structural studies. However, a crystal structure of LDL receptor YWTD β-propeller/EGF (BP) domain was solved as shown in Figure 3-3.

The YWTD repeat is folded into a six-bladed β-propeller domain. Each β-propeller consists of six blades arranged radially about a central axis (Figure 3-3). Along this axis is a central channel, 8−9 Å in diameter, filled with water molecules that form hydrogen bonds with backbone and side chain from the β-strands lining the channel. Each blade of the propeller has four antiparallel β-strands that are offset from the YWTD repeats. The “first” strand of the sixth blade is a C-terminal strand that follows the fifth blade to complete circularization

of the steroidal YWTD β propeller domain. The YWTD β propeller is followed by an EGF-like domain (E3) that packs against the propeller to create a continuous hydrophobic core in the mature protein (*130*). The linker connecting the propeller to E3 packs tightly against the base of the propeller and contacts predominantly hydrophobic side chains in the 1−2 loops of blades two and three, situating E3 in contact with the second and third blades of the propeller. The large interface area between the YWTD domain and the C-terminal EGF modulesuggest that other members of the LDLR family will share this organization, including the BP domain of LRP-6.

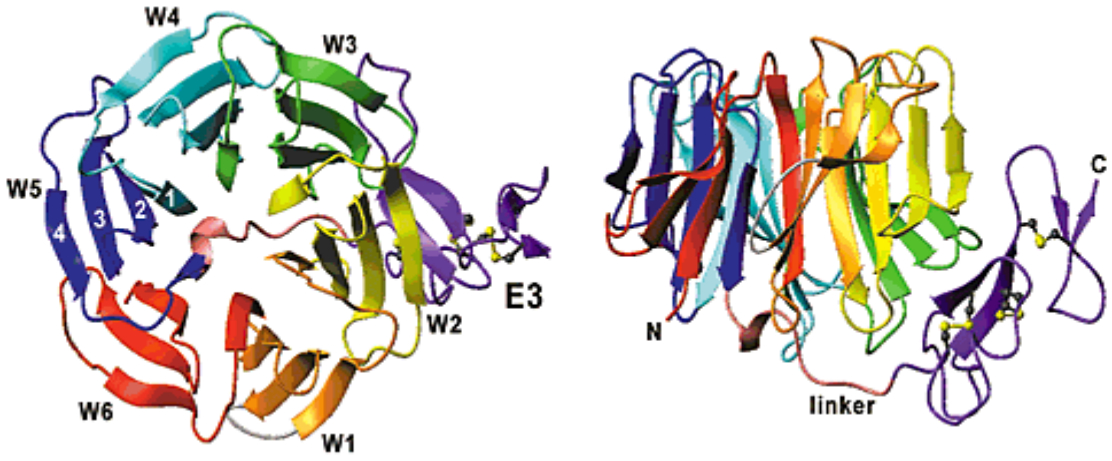

Figure 3-3: Ribbon representation of the YWTD domain and adjacent C-terminal EGF-like module (E3) of the LDL receptor, colored to point out the six YWTD repeats of the six-bladed propeller (*130*).

For LDLR and LRPs, at least two private chaperones, the LRP-receptor-associated protein (RAP) and the Boca/Mesd chaperone, are dedicated to the

folding and maturation of these multi-domain proteins as they transit through the secreatory pathway to the cell surface (*114, 115*). RAP binds the cysteine-rich ligand-binding modules of LDLR family members, preventing premature association with their respective ligands. Also, RAP escorts fully folded LDLR family members from the ER to the Golgi, where the lower pH triggers RAP to dissociate and recycle back to the *ER* via its retention signal (*126*). In *Drosophila*, Boca was shown to promote maturation and surface expression of several LDLR family members through a direct, but transient, interaction (*114, 187*). Boca/Mesd family members are localized in the ER, where they aid in the folding of YWTD β-propeller domains. *ER* chaperone proteins provide the most prevalent mechanism for achieving quality control during protein translation and translocation (*208*). Boca/MESD are important for the folding of LRP5 and LRP6,.

Our lab recently solved the high resolution NMR structure of full-length of MESD (*209*). MESD contains two structural and functional domains: a chaperone domain and an escort domain. The chaperone domain is located in residues 1-150 and highly structured, with a four-stranded anti-parallel β-sheet in the center, surrounded by two rigid helices on the top and one rigid helix in the bottom. Another flexible short helix is also above the sheet, containing several critical binding-residues to the BP domain of LRP6. The escort domain is located in residues 151-195 and is flexible, containing two short helices and an array of conserved lysine and arginine residues.

3.1.4. Challenge of refolding of BP1-LRP6

The BP1-LRP6 is a 328-amino acid fragment at the N-terminal of LRP6, including six YWTD β-propeller motifs and one epidermal growth factor (EGF)-like repeat. It is a multiple domains protein with three disulfide bonds in the EGF domain. Refolding such a complex multiple-domain protein BP1-LRP6 *in vitro* is challenging. Spontaneous refolding *in vitro* is only efficient for small, single-domain proteins that fold rapidly. multiple domain proteins often refold inefficiently due to the formation of partially folded intermediates that tend to self-associate (aggregate) into disordered complexes, driven by hydrophobic forces and inter-chain hydrogen bonding. The aggregation process irreversibly removes proteins from their productive folding pathways and only can be avoided *in vivo* protein refolding facilitated by molecular chaperones.

Previous *in vitro* protein folding studies with BP1-LRP6 have shown that not only the formation of disulfide bonds *in vitro* occurs significantly slower than *in vivo*, but also the efficiency of forming native disulfide bonds *in vitro* is dramatically lower than *in vivo*, even under the most favorable condition (*210*). During *in vitro* refolding, non-native disulfide bonds may be formed. Compared with the lower efficiency of forming native disulfide bonds *in vitro*, our *in vivo* protein refolding technique can offer much greater native disulfide bond formation efficiency.

Specific ER chaperone MESD was reported to be necessary for the folding of BP1-LRP6 (*211*). Our *in vivo* protein refolding technique can equip the refolding process with complete molecular chaperones and folding enzymes,

greatly increasing the refolding efficiency. Therefore, to refold large quantities of bacterially expressed BP1-LRP6 into its native conformation, we used our novel *in vivo* protein folding technology.

3.2. Material and methods

3.2.1. Strain, plasmid and media

E. coli cells were grown in Luria-Bertani (LB) medium (*BD*) at 37℃ and then were grown in minimal M9 medium at 20℃ for overnight. Kanamycin (30 μg/ml) was added as required. *E. coli* stain ER2556 [genotype: F ⁻ λ ⁻ fhuA2 [lon] ompT lacZ::T7 geneI gal sulA11 Δ(mcrC-mrr)11::IS10 R(mcr-73::miniTn10)2 R(zgb-210::Tn10) 1 (TetS) endA1 [dcm]; *New England Biolabs*] was used for molecular cloning. *E. coli* strain BL21(DE3) [F- *ompT hsdS*$_B$ (r_B-m_B-) *gal dcm* (DE3 [*lacI lacUV5-T7 gene1 ind1 Sam7 nin5; New England Biolabs*] was used for protein expression and purification. BP1-LRP6 with an N-terminal 6-histidine tag was expressed from plasmid pET30a (+) (*Novagen*).

3.2.2. DNA manipulation

The original long his-tag of the pET30a (+) vector (*Novagen*) was replaced by a short His tag containing six histidine plus two serine residues by mutagenesis. Mutagenesis was carried out using the QuickChange^{TM} site-

directed mutagenesis kit (Stratagene, CA). The mutations were confirmed by DNA sequencing.

Human BP1-LRP6 cDNA (kindly provided by Dr. Guojun Bu, Mayo Clinic, Jacksonville, FL, USA) was used as the template of PCR. BP1-LRP6 was constructed by subcloning PCR products into the *BamHI/XbaI* sites of the engineered pET30a(+) vector. BP1-BP1-LRP6 includes six YWTD β-propellers and one EGF repeat, referring the residues 20-324 of full-length LRP6. The constructs derived from PCR were confirmed by DNA sequencing.

The engineered pET30a (+) vector provide high protein expression level. The confirmed positive DNA constructs were first transformed into the *E. coli* strain ER2566 competent cells for plasmid DNA replication and glycerol stock production. Ultimately, they were transformed into BL21 (DE3) cells for protein expression.

3.2.3. Protein expression and purification

Protein expression was carried out using the *E. coli* strain BL21(DE3). A 10-15 ml overnight culture in a rich medium, such as 2×YT, was made from a glycerol stock of BP1-LRP6 expressing cells. The next morning, the overnight culture was diluted into the LB medium with a ratio 1:100. The culture was grown in LB medium with 30 ug/ml kanamycin at 37 ℃ until it reached an OD_{600} of 0.8-1.0. The cells were spun down and re-inoculated into a minimal M9 medium (kanamycin, 30 μg/ml) at half volume of LB medium. The culture was continued

at 20 ℃ for one hour. A sample of the culture was take n to check its OD_{600} to confirm the continuous growth of cells under lower temperature and the minimal M9 medium. Then expression of the recombinant BP1-LRP6 was induced by IPTG at the concentration of 1 mM for overnight at 20 ℃. The cells were harvested by centrifuge the next morning. The cell pellet was resuspended in binding buffer (10 g/ml cells, 20 mM Tris-HCl, 500 mM NaCl, 2.5 mM imidazole, 6 M urea at pH 8.0) of His tag affinity purification system followed by sonication while on ice. To separate the cell pellet from the supernatant, the cell pellet was spun down. The pellet was again resuspended in the binding buffer containing 6 M urea and then sonicated on ice again. This procedure was repeated three times to ensure complete solubilization of all bacterially expressed BP1-LRP6 in the inclusion bodies in this buffer system. All supernatant fractions were combined and applied on a His•Bind® resin column (Novagen). The column was extensively washed with binding buffer containing 6 M urea followed by washing with 100-500 ml of wash buffer harboring 20-30 mM imidazole. The purified recombinant BP1-LRP6 was eluted from the column using elution buffer with 1.0 M imidazole and 6 M urea. The eluted protein was dialyzed against 10 mM ammonium bicarbonate to remove the unwanted salt and imidazole in the buffer . Ultimately, the dialyzed protein solution was lyophilized to get the pure BP1-LRP6 protein powder. The protein powder was weighed and was stored at -20 ℃ for future experimental purposes.

3.2.4. Time course for QQ-protein delivery

The bacterially expressed BP1-LRP6 was modified with QQ-reagent overnight at 4 ℃ with gently shaking. Modified BP1-L RP6 was centrifuged at 5000 rpm for 5 minutes to remove precipitation. Then the modified BP1-BP1-LRP6was mixed with serum-free DMEM to reach a final protein concentration of 0.1 mg/ml. Eighteen dishes (35 cm^2) of 80% confluent (about 0.3×10^6 cells) HeLa cells were incubated with the protein delivery medium and protease inhibitor for 3 hours. Another 3 dishes of HeLa cells were incubated with QQ-reagent only (no BP1-LRP6) for 3 hours as a negative control. Samples of HeLa cells were taken in triplicate every 30 minutes to monitor the protein delivery efficiency. All the samples were subjected for SDS-PAGE gel electrophoresis with or without DTT.

3.2.5. Confocal fluorescence imaging

Zeiss Apotome microscopy system is useful for optical sectioning in multicolor applications and for high-resolution fluorescence microscopy. We use this system to monitor and visualize the protein delivery of fluorophore attached BP1-LRP6. The bacterially expressed BP1-LRP6 was first dissolved in PBS buffer, pH 7.4 (0.5-1.0 mg/ml), overnight. The solution then was spun for 10 minutes at 12,000xg at room temperature. Seventy ul of protein solution was taken and 1 µl of amine-reactive ArrayIt®Green540 was added at room temperature. Incubation was carried out for 6 hours or overnight in a cold room and then purified using a desalting spin-column at 10,000 rpm for 4 minutes to remove free dye. The purified, fluorescence labeled protein was incubated with

QQ-reagent overnight without further purification. HeLa cells were seeded three days before the experiment to reach 80% confluency. QQ-reagent modified, fluorophore labeled BP1-LRP6 in serum-free DMEM culture medium was incubated with Hela cells at 37 °C for 1 hour or 2 hours. The cells were very healthy after protein delivery, displaying normal morphology (Figure 3-7). Cells were washed with PBS extensively to remove any labeled protein from the medium. Then the Hela cells were subjected to confocal fluorescenceimaging using a Zeiss Axioplan2 Imaging fluorescent microscope with ApoTome imaging at the Microscopy and Imaging Resources Laboratory at Wayne State University, School of Medicine. Fluorescence was recorded with an AxioCam MRm camera and AxioVision software.

3.2.6. Far-Western blot

The cell lysates on a native PAGE gel were tansferred onto a nitrocellulose membrane at 380 mA for 2 hours. After transfer, the membrane was incubated at room temperature with the probe---MESD, which is the private chaperone of LRP5/6 and can specifically bind to the properly folded BP1-LRP6. The membrane was then washed three times with 20 ml PBS for 5 minutes each time, to remove unbound MESD. The membrane was incubated with anti-MESD (Abnova) for 1 hour in 2% dry milk PBS, followed by incubation with HPR-conjugated seconday antibody for 1 hour. At last, the membrane was detected

using ECL (Pierce). Using this assay, properly folded BP1-LRP6 can be detected because of the formation of the MESD/BP1-LRP6 complex.

3.2.7. Time course for protein refolding

Protein refolding inside mammalian cells takes time, especially for BP1-LRP6. We delivered large quantities of bacterially expressed BP1-LRP6 inside HeLa cells. After protein delivery, we set up a three-day protein refolding time course experiment to monitor the refolding and stability of the delivered BP1-LRP6 inside HeLa cells.

BP1-LRP6 first was modified with QQ-reagent for overnight at 4 °C. Then QQ-reagent modified BP1-LRP6 was mixed with serum-free DMEM to reach 0.1 mg/ml concentration. The protein delivery medium was incubated with 80% confluent HeLa cells for 3 hours (the optimal protein delivery condition) at 37 ℃. Then the protein delivery medium was removed and the Hela cells were gently washed with pre-warmed PBS buffer three times. Hela cell samples were taken in triplicate to monitor the protein delivery. During the three-day protein refolding time course experiment, Hela cells were incubated with fresh serum-free DMEM and every 24 hours, HeLa cell samples were taken to monitor the protein folding and the protein stability inside of the cells. One flask of cells was set aside, only loading with QQ-reagent protein delivery medium (without BP1-LRP6) as a negative control.

3.2.8. Detection of the refolded BP1-LRP6 by Western blot

The recombinant BP1-LRP6 was fused with His-tag. Protein purification was performed using nickel chelating resin column. The purified refolded BP1-LRP6 protein powder and the bacterially expressed BP1-LRP6 protein powder was dissolved in PBS buffer and the concentrations of both of the proteins were adjusted to 0.25 mg/ml, confirmed by a BCA protein assay (Pierce). Protein samples were loaded at different volumes and analyzed by SDS-PAGE with or without the reducing reagent, DTT. Following transferring to PVDF membrane, successive incubation with mouse anti-His tag monoclonal antibody (Sigma) or mouse anti-BP1-LRP6 antibody (kindly provided by Dr. Jian-Ping Jin) and horseradish peroxidase-conjugated secondary antibodies (Amersham Life Science) were carried out according to the routine Western blotting procedures. The immunoreactive proteins were then detected using the ECL system.

3.2.9. Far UV Circular Dichroism (CD) Spectroscopy

Circular dichroism (CD) spectroscopy measures differences in the absorption of left and right circularly polarized light, which arise due to structural asymmetry. Protein secondary structure can be determined by CD spectroscopy in the far UV region (190-260 nm). Three common secondary structure motifs (α-helix, β-sheet, and random coil) exhibit the characteristic and distinctive CD spectra in the far UV region. α-Helical proteins have negative bands at 222 nm and 208 nm and a positive band at 193 nm. Proteins with well-defined antiparallel β-pleated sheets have negative bands at 218 nm and positive bands at 195 nm,

whereas disordered proteins have very low ellipicity above 210 nm and negative bands near 195 nm.

CD is reported in units of ellipticity (θ), usually expressed in millidegrees, which reflects the sum of the entire molecular population in a sample solution. It can be normalized to the molar concentration of the sample as molar ellipticity.

Circular dichroism spectroscopy is commonly used to: 1) determine if a protein is folded; 2) characterize protein secondary structure and tertiary structure (using near-UV CD spectrum); 3) compare the structures of a wild-type protein with different mutants; 4) study protein stability to temperature or chemical denaturation (GdnHCl or urea) and effects of salt, pH, organic solvents; 5) study protein folding, unfolding, denaturation or aggregation; and 6) determine whether protein-lipid interactions alter the conformation of a protein.

In our studies, CD measurements were carried out on an Olis DSM 17 CD spectrophotometer (Bogart, GA), with a constant 20 °C temperature capability under computer control within ±0.2 ^{o}C. Protein concentrations were determined by the BCA protein assay (Pierce) and absorbance at 280 nm using a spectrophotometer, which gave a similar result of the protein concentration by both methods. Eight hundred μl samples in 25 mM phosphate buffer at pH 7.4 were prepared. The far UV spectra were collected at 190-260 nm with a cuvette using 250 μl sample in a 1 mm pathlength, at a protein concentration of 0.1-0.2 mg/ml. Fifteen scans between 190 and 260 nm were acquired and averaged. A base-line scan was subtracted to produce the final average scan. The value of molar ellipticity ([θ] x 10^3 deg cm^2/dmole) was plotted as a function of protein

concentration. The average α-helical, β-sheet, and other secondary structure element content in the refolded BP1-LRP6 was calculated using DICHROWEB server (*212*).

3.3. Results

3.3.1. Expression and purification of the recombinant BP1-LRP6 from *Escherchia coli (E. coli)*

BP1-LRP6 with N-terminus His tag was expressed in *E. coli*. Bacterially expressed BP1-LRP6 adopted random conformations and preferred forming oligomers by non-native disulfide bonds. In order to break the intermolecular disulfide bonds for monomer BP1-LRP6 protein band, all the samples in Figure 3-4 were reduced by DTT.

Figure 3-4 showed the highly efficient recombinant BP1-LRP6 expression and the nickel chelating column purification process. The strong protein band of 37 kDa (BP1-LRP6) was shown in the IPTG-induced cell lysates while there was no induced protein expression band in the control (lanes 1 and 2). Large quantities of recombinant BP1-LRP6 expressed inside *E. coli* caused protein aggregation and formed inclusion bodies. We had to use 6 M urea in the purification buffer and repeated the sonication step three times in order to extract the recombinant BP1-LRP6. The extracted recombinant BP1-LRP6 in the buffer was loaded onto a nickel chelating column, followed by extensive washings to reduce non-specific proteins from binding. BP1-LRP6 was eluted out with 1 M

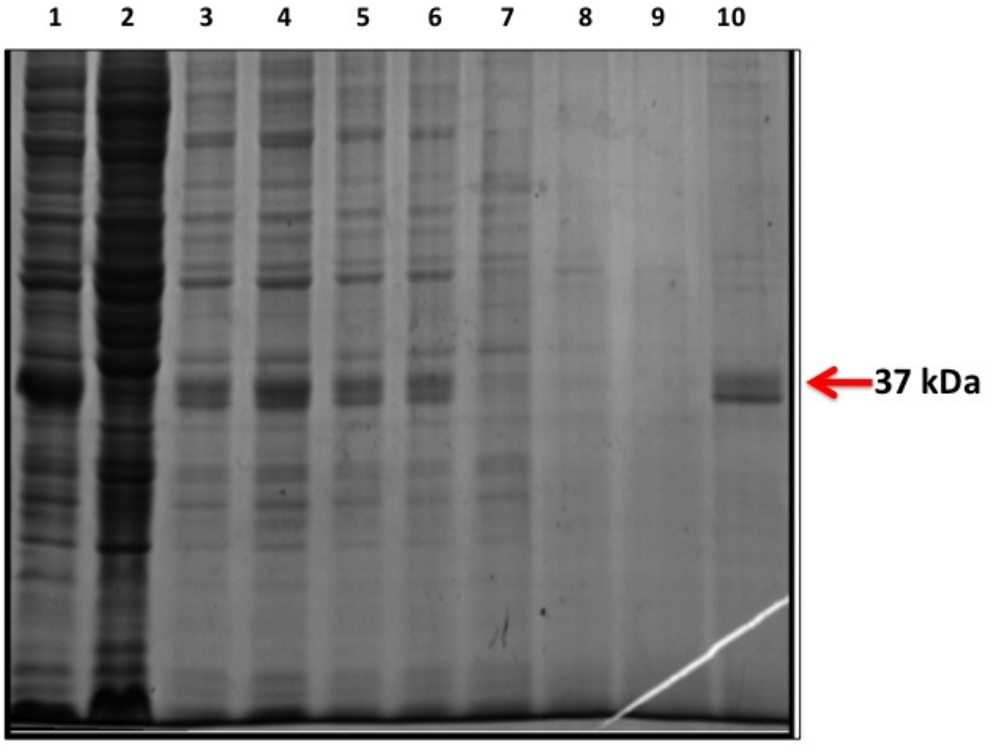

Figure 3-4: A 12% reducing SDS-PAGE gel, showing the induced expression of BP1-LRP6 in *E. coli* and its purification by using nickel chelating column.

Lane 1: *E. coli* cell lysates after IPTG induction. Lane 2: *E. coli* cell lysates without IPTG induction as a negative control. Lane 3: the cell extract before loading onto the column. Lane 4-6: the cell extract from the 1st sonication, 2nd sonication, and 3rd sonication. Lane 7: the flow through after loading. Lane 8: the flow through of the binding buffer wash. Lane 9: the flow through of the washing buffer wash. Lane 10: the elution.

imidazole. The elution was dialyzed and lyophilized to obtain pure protein powder. By using this high-yield recombinant protein expression method, we could routinely obtain about 70 mg of pure BP1-LRP6 protein powder from 500

ml cell culture. Pure bacterially expressed BP1-LRP6 was stored at -20 °C for further experiments.

3.3.2. Highly efficient protein delivery by QQ-reagent

To optimize the protein delivery condition for *in vivo* refolding of BP1-LRP6, we conducted a time course of protein delivery experiment for 3 hours. Bacterially expressed BP1-LRP6 was modified with QQ-reagent for overnight at 4 ℃. The QQ-modified BP1-LRP6 was mixed with serum-fr ee DMEM to reach the final concentration of 0.1 mg/ml. Dishes containing 80% confluent HeLa cells were incubated with the protein delivery medium for different periods of time. Every half hour, one dish of HeLa cells was taken out and was suspended in the SDS-PAGE loading buffer. All the samples were analyzed by a 12% SDS-PAGE. In order to observe the monomer band of BP1-LRP6, we ran a duplicate SDS-PAGE under reducing condition.

Figure 3-5 showed the gradual appearance of a monomer BP1-LRP6 protein band (red arrow) in the Hela cell lysates samples from 0.5 hour to 3 hours. Under a reducing condition (Figure 3-5, right panel), the intensity of the BP1-LRP6 protein band was significantly increased (red arrow). The data demonstrated that the protein delivery of BP1-LRP6 into Hela cells was linear, maximizing around 2 hours post delivery.

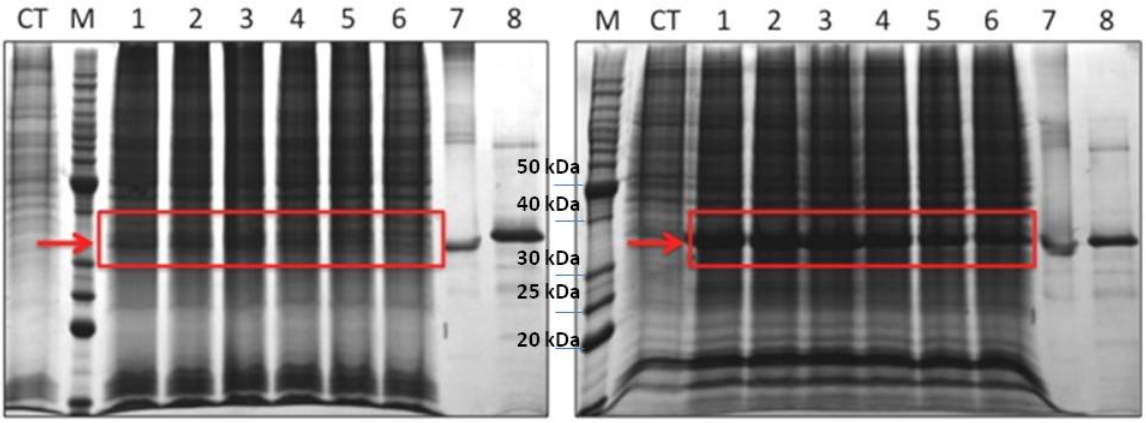

Figure 3-5: 12% SDS-PAGEs of a time course of the QQ-delivery of the BP1-LRP6 into HeLa cells.

Lane CT: the Hela cell lysis after loaded with QQ-reagent for 3 hours as a negative control. Lane M: protein marker. Lane 1-6: the Hela cell lysis after loaded with QQ-modified BP1-LRP6 for 3 hours, 2.5 hours, 2 hours, 1.5 hours, 1 hour and 0.5 hours. Lane 7: the bacteria expressed BP1-LRP6 powder. Lane 8: the bacteria expressed BP1-LRP6 powder with DTT. Left panel: Without DTT. Right panel: With DTT.

Bacterially expressed BP1-LRP6 was labeled with amine-reactive ArrayIt®Green540 and delivered into Hela cells with the QQ-reagent and imaged using confocal fluorescence imaging (Figure 3-6). Images were taken after 1 and 2 hours after protein delivery. After 1 hour protein delivery, the fluorescence images showed that most of BP1-LRP6 was located in the perinuclear location inside HeLa cells, possibly in the ER for refolding (Figure 3-6, top panel).This was confirmed by the imaging data after 2 hours of protein delievery, showing that more BP1-LRP6 proteins were located into the perinuclear area of the HeLa cells (Figure 3-6, bottom panel). This result confirmed our result of the time

course of BP1-LRP6 delivery, suggesting that 2 hour protein delivery appeared to be the optimal protein delievery time (Figure 3-5). The QQ-reagent is highly efficient for protein delivery with most of bacterially expressed BP1-LRP6 being delivered inside of HeLa cells, ready for refolding.

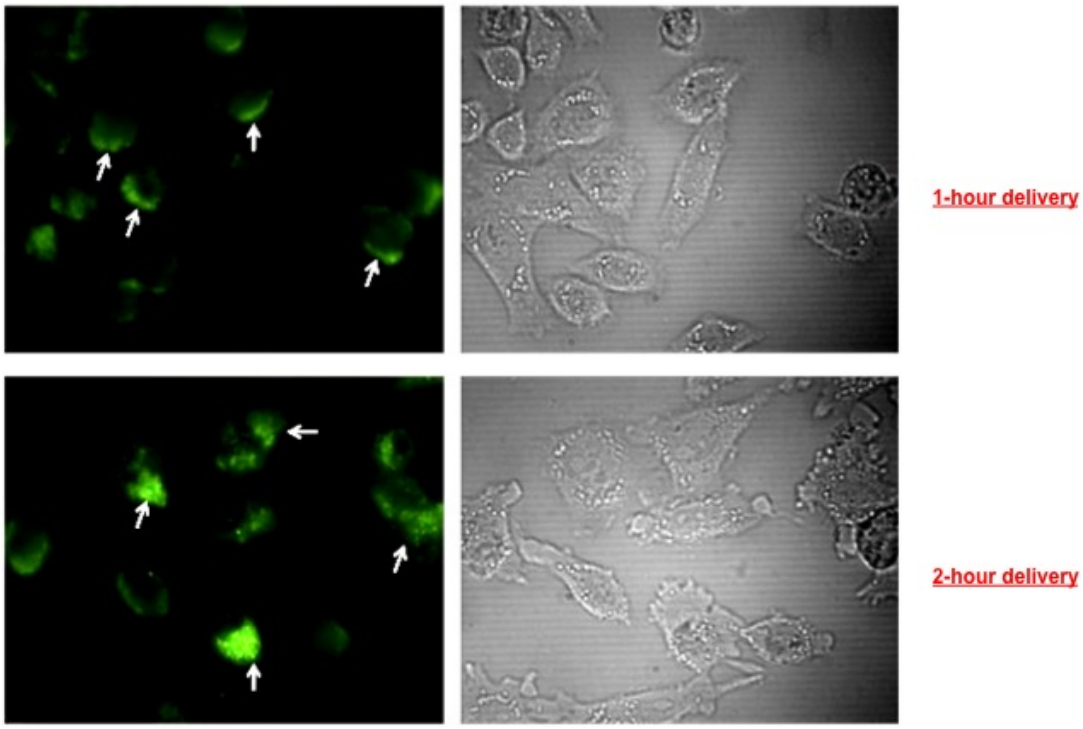

Figure 3-6: Confocal fluorescence images of the Hela cells with the delivered BP1-LRP6.

BP1-LRP6 was first labeled with small molecular fluorophore (ArrayIt® Green 540 amine-reactie dye) and then modified with QQ-reagent for protein delivery. Fluorescence images (on the left) and DIC images (on the right) were taken after 1-hour and 2-hour protein delivery.

A relatively high concentration of BP1-LRP6 was kept inside HeLa cells during refolding. After delivering bacterially expressed BP1-LRP6 for 2 hours, we changed the protein delivery medium for fresh serum-free DMEM for another overnight incubation. We took samples of the HeLa cells and ran the cell lysates sample on a SDS-PAGE. We found that BP1-LRP6 accumulated inside HeLa

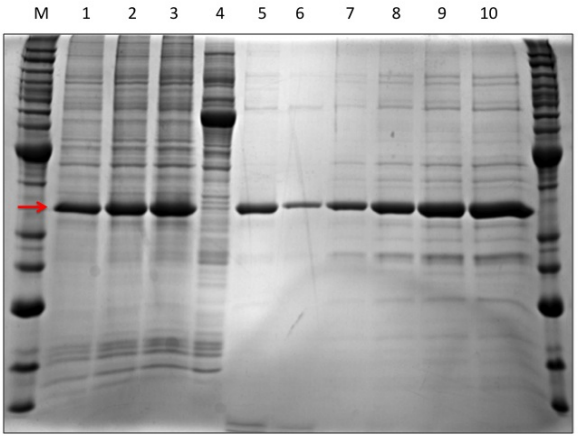

Figure 3-7: A 12% SDS-PAGE analysis to semi-quantify the delivered BP1-LRP6 inside HeLa cells.

All the samples were reduced with DTT. Lane M: protein marker. Lane 1-3: 10 ul, 15 ul and 20 ul of the Hela cell lysates sample after 2 hours protein delivery and an overnight incubation. Lane 4: the intact Hela cell lysates as negative control. Lane 5: the modified BP1-LRP6 before loading. Lane 6: the remaining undelievered BP1-LRP6 in the medium after 2 hours of protein delievery. Lane 7-10: 10 ul bacterially expressed BP1-LRP6 powder at the concentrations of 0.25 mg/ml, 0.5 mg/ml, 0.75 mg/ml and 1 mg/ml. The concentration of the delivered BP1-LRP6 was between 0.25 mg/ml and 0.5 mg/ml.

cells (Figure 3-7). To semi-quantify the amount of BP1-LRP6 inside of HeLa cells, we ran the HeLa cell lysates samples and the bacterially expressed BP1-LRP6 samples with known concentrations under the same reducing condition (Figure 3-7). We found that the delivered BP1-LRP6 inside Hela cells reached the concentration range between 0.25 mg/ml and 0.5 mg/ml (lanes 7-10, Figure 3-8). It was surprising that the protein delivery could be so efficient for BP1-LRP6 as compared to that for LBD-ApoER2 (Chapter 2). The delivered protein appeared stable inside HeLa cells even after overnight incubation. The differences in delievery efficiencies could be related to the unique biochemical properties of individual proteins (modified by QQ-reagent to differing extents, charge, etc.).

3.3.3. Inside of HeLa cells, the delivered BP1-LRP6 was under refolding process, showing increasing ligand-binding ability in the Far-western blot

MESD is an *ER* chaperone especially binding properly folded YWTD β-propeller domains of LRP5/6 and escorting them throughout the secretary pathway (*209*). We used MESD as our far-Western blot probe to specifically bind and detect the properly folded and biologically functional BP1-LRP6 in the samples.

Bacterially expressed BP1-LRP6 was modified by QQ-reagent and was delivered inside of HeLa cell for 4 hours. Each hour, we collected the sample of HeLa cells harboring delivered BP1-LRP6. All the HeLa cell samples were lysed

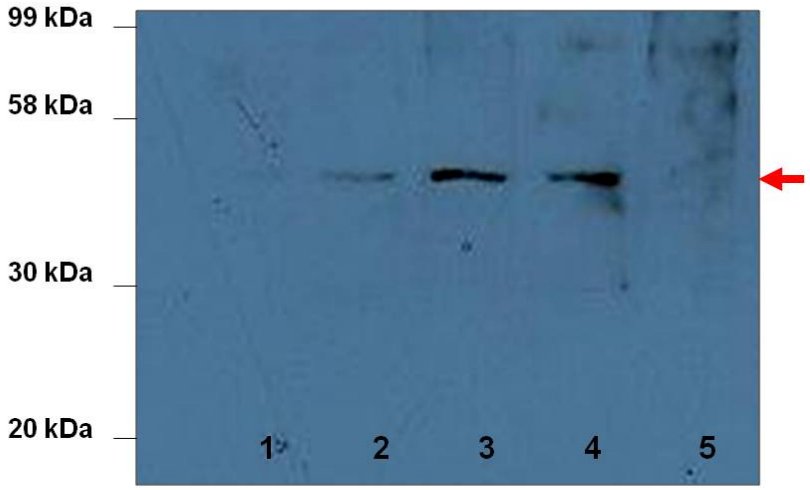

Figure 3-8: Far-Western blot (MESD as probe) of BP1-LRP6.

The HeLa cells with delivered BP1-LRP6 were collected at different time point. The samples were lysed and were subjected to far-western blot using MESD as probe. Lane1: 1 hour. Lane 2: 2 hours. Lane 3: 3 hours. Lane 4: 4 hours. Lane 5: the bacterial expressed BP1-LRP6 without refolding as negative control. The red arrow shows the band site of BP1-LRP6. (Courtesy of Dr. Qianqian Li).

and, with the bacterially expressed BP1-LRP6 sample, the HeLa cell lysates were subjected to the far-Western blot. Our result showed that the delivered BP1-LRP6 gradually underwent conformational changes inside of HeLa cells, showing stronger and stronger binding ability to its probe, MESD, as time went on (as shown in Figure 3-8). This data suggested that once the bacterially expressed BP1-LRP6 was delivered inside of HeLa cells, the refolding process began. More bacterially expressed BP1-LRP6 was refolded and increasing amounts of properly folded BP1-LRP6 appeared, showing its specifically binding

to the ligand (our probe) MESD. At this point, we confirmed that the delivered BP1-LRP6 can be refolded inside of HeLa cells. Given refolding time long enough, more BP1-LRP6 can be properly folded showing its biological function to bind MESD and eventually the delivered BP1-LRP6 can be totally refolded while the persistent unfolded BP1-LRP6 can be degraded by the active mechanisms such as ERAD or UPR.

3.3.4. Optimization of the protein refolding protocol for BP1-LRP6

Protein refolding inside Hela cells takes time. We conducted experiments to monitor the fate of the delivered BP1-LRP6 (Figure 3-9). Our recombinant BP1-LRP6 does not have a signal peptide at the N-terminus, thus the refolded BP1-LRP6 remains inside HeLa cells. We tried to prolong the *in vivo* refolding time by incubating the protein delivered HeLa cells with serum-free DMEM for three days (Figure 3-9). Every 24 hours, HeLa cell samples were taken. All the samples were incubated with 10 mM DTT before being analyzed by 12% SDS-PAGE for a monomer BP1-LRP6 protein band. Our results showed that the amount of BP1-LRP6 inside HeLa cells was significantly decreased in the first day of *in vivo* protein refolding. However, in the subsequent two days of *in vivo* refolding, the amount of BP1-LRP6 inside HeLa cells was relatively stable, showing only slight decrease. After 3 days of refolding, the HeLa cells still appeared morphologically normal (data not shown). Since a BP1-LRP6 protein band (red arrow) remained detectable and the protein refolding time was maximal to 3 days. These results suggest that 2 hours of protein delievery of the

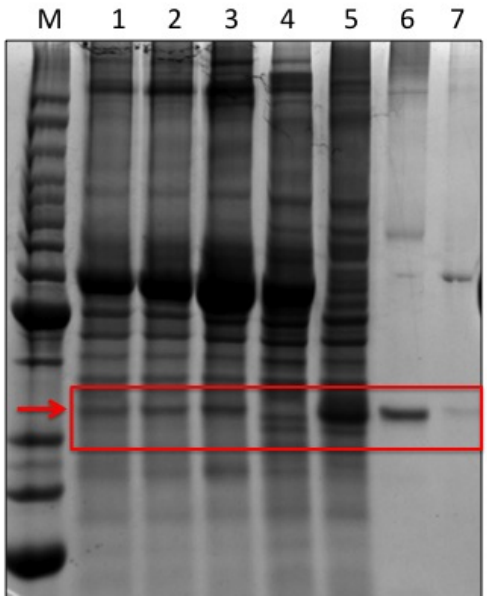

Figure 3-9: The time course experiment on refolding BP1-LRP6 inside HeLa cells.

Lane M: protein marker. Lanes 1-3: the Hela cell lysates after incubation for 3 days, 2 days and 1 day post protein delievery. Lane 4: the Hela cell lysates of negative control. Lane 5: the Hela cell lysis after 2 hours protein delivery. Lane 6: the modified BP1-LRP6 before loading. Lane 7: the remaining BP1-LRP6 in the protein delivery medium after 2-hour of protein delievery. All the samples were in the presence of 10 mM DTT.

bacterially expressed BP1-LRP6 into Hela cells by the QQ-reagent (Figure 3-5) followed by 3 days of protein refolding inside of the cells (Figure 3-9) are likely the optimized conditions for the refolding of BP1-LRP6 inside Hela cells.

Mammalian cells have developed mechanisms to actively remove toxic proteins, such as unfolded or misfolded proteins. Not only the active ATP-

consuming ER associated degradation (ERAD) and unfolded protein response (UPR) but also the unspecific proteolytic enzymes of the lysosome can actively degrade the toxic unfolded or misfolded proteins inside mammalian cells (Chapter 1). Also, mammalian cells can secrete out the toxic proteins and compact them into a less toxic amyloids stocked in the extracellular tissue. Ultimately, accumulation of these toxic proteins in the cell cytoplasm can cause apoptosis. It is unlikely that the unfolded BP1-LRP6 could escape away from such a well-organized and strict control, being kept inside healthy HeLa cells. We thought that the relatively stable stock of BP1-LRP6 inside HeLa cells after 3 days of *in vivo* protein folding was the potentially refolded BP1-LRP6.

3.3.5. The high yield of refolded BP1-LRP6

Is the refolded BP1-LRP6 properly folded? How is its structure organized? To address these questions, we need first to purify the refolded BP1-LRP6 from the HeLa cells after three-day *in vivo* protein folding (Figure 3-10). In order to keep the native conformation of the refolded BP1-LRP6 as it was inside Hela cells, no denaturing reagent was used in the buffers of nickel chelating column purification. Due to the protein-protein interactions inside HeLa cells, non-specific proteins could be co-purified (data not shown). To solve the problem, we modified the purification protocol by increasing NaCl concentration up to 1 M in the buffers. This modification greatly improved the quality of the eluted refolded BP1-LRP6. The column was extensively washed and the refolded LRP6 was eluted by an imidazole concentration gradient (lanes 7-9, Figure 3-10).

To identify the purified refolded BP1-LRP6, Western blot against anti-His tag antibody and anti-LRP6 were applied to the same quantity of the purified refolded BP1-LRP6 and the bacterially expressed one (Figure 3-11). Purified refolded BP1-LRP6 showed strong positive signals detected by both anti-His tag antibodies and anti-LRP6 antibodies, identifying the re-purified protein as the refolded BP1-LRP6 (Figure 3-11). In the anti-His tag Western blot, both the bacterially expressed and the refolded protein powders showed positive signals of BP1-LRP6 at the monomer position. Compared with the bacterially expressed BP1-LRP6, our *in vivo* protein refolding method had significantly enriched monomeric BP1-LRP6 in the sample.

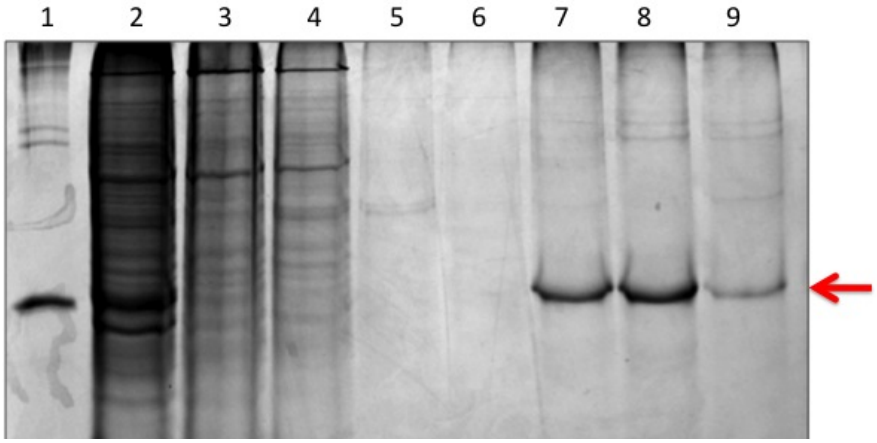

Figure 3-10: Re-purification of refolded BP1-LRP6 from Hela cells.

Lane 1: the bacteria expressed BP1-LRP6 powder. Lane 2: the Hela cell extract before loading onto His•Bind column. Lane 3: the flow through after loading. Lane 4: the flow through of binding buffer washing. Lane 5-6: the flow through of washing buffer wash (20 mM and 70 mM imidazole washing). Lane 7-9: the elution at 100mM, 250mM and 1M imidazole concentrations. All the samples were treated with 10 mM DTT before analyzed by 12% SDS-PAGE.

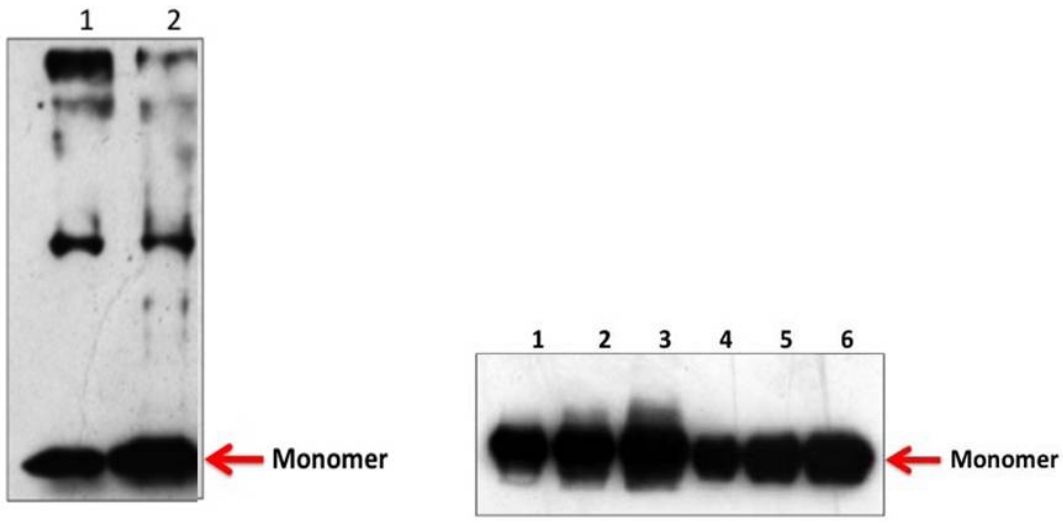

Figure 3-11: Anti-His tag (Left panel) and anti-LRP6 (right panel) western blotting to identify purified refolded BP1-LRP6.

The left panel: all the protein samples were at 50 ng/μl and without DTT. Lane 1: 10 ul (500 ng) bacteria expressed BP1-LRP6. Lane 2: 10 ul (500 ng) purified refolded BP1-LRP6. The right panel: all the protein powder sample were at the concentration of 50 ng/ul. Lane 1-3: 4 ul, 8 ul, and 12 ul (200, 400, and 600 ng, respectively) purified refolded BP1-LRP6. Lane 4-6: 4 ul, 8 ul and 12 ul bacterial expressed BP1-LRP6 at the same concentrations.

The crystal structure of the homologue of BP1-LRP6, the LDL receptor BP domain, had been solved (PDB file # 1IJQ) (*130*). In this structure, we noticed that the six cysteine residues in EGF-like domain formed three intramolecular disulfide bonds. Only the properly folded BP1-LRP6 should have structure similarities to the LDL receptor BP domain; its cysteine residues should also only form intra-domain disulfide bonds. If folded properly, the oligomeric species

(intermolecular bonds) should be greatly reduced while the monomeric band (intramolecular bonded form) (37 kDa) should be greatly enriched. This is exactly what was observed (lanes 1 and 2, the left panel of Figure 3-11). This finding suggests that this *in vivo* protein refolding technique enables efficient refolding of BP1-LRP6. Previous published data (*211*) supports this conclusion, showing that the molecular chaperone MESD was co-overexpressed with BP1-LRP6 to facilitate its folding. MESD dramatically enhanced the folding of BP1-LRP6, and the properly folded BP1-LRP6 was secreted into the medium. Only one monomer band of the secreted BP1-LRP6 was shown in their Western blot. Therefore, the monomer BP1-LRP6 represents the potential properly folded BP1-LRP6. Our purified refolded BP1-LRP6 had large percentage of the properly folded BP1-LRP6.

The *in vivo* protein refolding method was optimized specifically for BP1-LRP6. The bacterially expressed BP1-LRP6 was modified by the QQ-reagent at 0.1 mg/ml (final concentration). Protein delivery was conducted for 2 hours and protein was extensively refolded inside of HeLa cells for 3 days. After 3 days of refolding, the potentially properly refolded BP1-LRP6 was purified by His•Bind (Nickel chelating) column chromatography. A high yield of the refolded BP1-LRP6 was achieved through this *in vivo* refolding process. The yields of the refolded BP1-LRP6 from two independent purification schemes are shown in Table 3-1. We routinely obtained a yield of around 60%.

Table 3-1: The final yield of the purified refolded BP1-LRP6 using *in vivo* protein refolding technique.

QQ- modified bacterially expressed BP1-LRP6	Purified refolded BP1-LRP6	The final yield rate
30 mg	18.54 mg	62%
10 mg	5.64 mg	57%

3.3.6. Well-defined secondary structure shown in the far-UV CD spectrum of refolded BP1-LRP6

Far-UV circular dichroism (CD) spectroscopy is an excellent tool for rapid determination of the secondary structure and folding properties of proteins. CD spectra of the purified refolded BP1-LRP6 and the bacterially expressed BP1-LRP6 were recorded in 50 mM sodium phosphate buffer at pH 7.4. The results showed very different spectra for these two samples (Figure 3-12). While the spectrum of the bacterially expressed BP1-LRP6 was more characteristic of a random coil, the spectrum of the purified refolded BP1-LRP6 showed a positive CD signal at 193 nm and negative CD signal near 218 nm, suggesting a predominantly defined β-sheet structure. This result is consistent with the crystal structure of the homologue of BP1-LRP6--- the LDL receptor BP domain, in which six YWTD repeats together fold into a six-bladed β-propeller and each blade is consisted of four-stranded β-sheets. Also, in the CD spectrum of the

refolded BP1-LRP6, the negative CD peaks centered near 208 nm and 222 nm suggest that the refolded BP1-LRP6 also contains some α- helical structure.

To further quantify the result, we sent the CD data of the refolded BP1-LRP6 to the DICHROWEB server (*212*). DICHROWEB is a user-friendly interface to analyze CD spectra. As the spectra of proteins that have been characterized by x-ray crystallography are standards, DICHROWEB compared the spectra of the refolded BP1-LRP6 with the standards using programs such as CONTINLL, VARSLC, K2d, CDSSTR, and SELCON3 (*212*), assuming that the spectrum of a

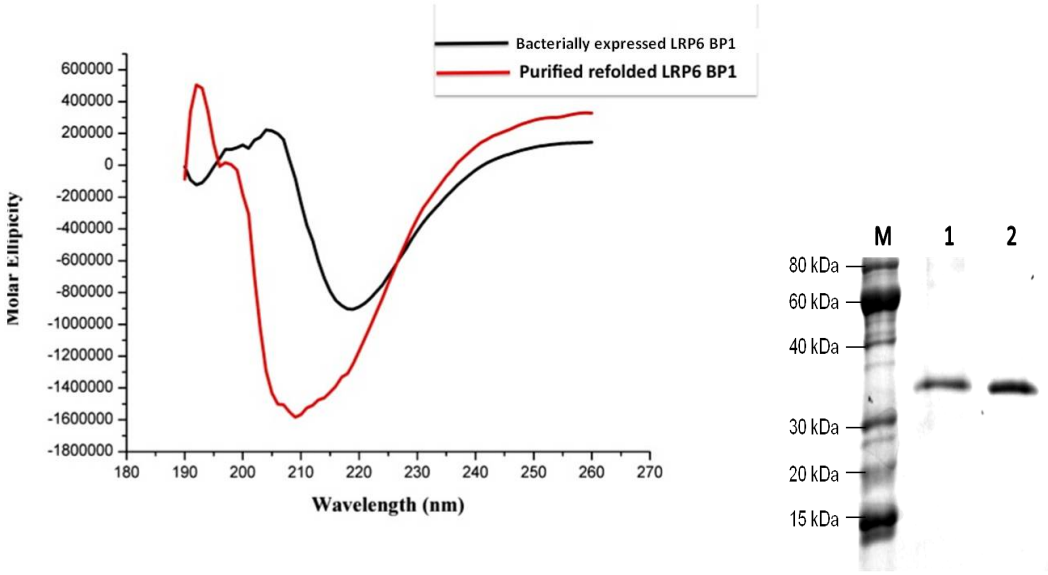

Figure 3-12: Far-UV circular dichroism (CD) spectrum of the bacterially expressed BP1-LRP6 and the purified refolded BP1-LRP6.

The right panel showed the CD samples on 12% SDS-PAGE. Lane 1 is the refolded BP1-LRP6 protein sample. Lane 2 is the bacterially expressed BP1-LRP6 protein sample. Both samples had reacted with 10 mM DTT.

protein can be represented by a linear combination of the spectra of its secondary structural elements, plus a noise term, which includes the contribution of aromatic chromophores and prosthetic groups (*213*). The calculated secondary structure element content of the refolded BP1-LRP6 was shown as Table 3-2. Compared with its homologue---the LDL receptor BP domain, the refolded BP1-LRP6 had almost the same high content of β- sheet and small percent of α- helices. The percentage numbers of the secondary element are so close that it strongly supports the idea that the refolded BP1-LRP6 adopts a similar conformation as the LDL receptor BP domain, further confirming that our refolded BP1-LRP6 was properly folded and adopted a defined and a potentially native conformation.

Table 3-2: Comparison of the secondary structure content of the refolded BP1-LRP6 (using DICHROWEB web server) with that of the BP domain of LDLR (*130, 212*).

Secondary Structure	Alpha-helix (%)	Beta-sheet (%)	Loop (%)
The refolded BP1-LRP6	**2.6**	**41.0**	**56.4**
LDL receptor YWTD-EGF	**2.0**	**46.0**	**52.0**

3.4. Discussion

LRP5 and LRP6 are unique members of LDL receptor family. Instead of having clusters of ligand binding repeats in the extracellular domain to exert

ligand-binding function, the YWTD β-propeller/EGF (BP) modules occupies large part of the extracellular domain and plays an important role in ligand binding. Especially in Wnt signal transduction, LRP5/6 acts as co-receptor of the canonical Wnt/β-catenin pathway. Malfunctions of LRP5 can cause disorder in bone metabolism (*152*). LRP6 mutations associate with coronary artery disease, Alzheimer's disease, and cancer (*192*). Structural studies on LRP5/6 are very important to understand the mechanisms of the above diseases.

BP1-LRP6 is a 328 amino acids (37 kDa) fragment, including six YWTD β-propeller motifs and one EGF-like repeat. No previous structural study on LRP6 has been reported. But, as the homologue of BP1-LRP6, the x-ray crystallography structure of LDL receptor YWTD β-propeller/EGF (BP) domain has shown that the six YWTD repeats of LDL receptor was folded into a six-bladed β-propeller domain. Each blade consisted of four-stranded β-sheets. EGF-like domain tightly packs against the second and third blades of the propeller.

Refolding the bacterially expressed BP1-LRP6 *in vitro* is challenging. BP1-LRP6 is a 37 kDa protein with multiple motifs. During *in vitro* refolding, the exposed hydrophobic residues in the refolding intermediates may cause irreversible aggregation by hydrophobic interactions or non-native hydrogen bonding. Also, without help of molecular chaperones and folding enzymes, especially RAP and MESD, it is very difficult to form native intra-repeat disulfide bonds in the EGF-like repeat of BP1-LRP6 for a proper folding.

Instead of using traditional *in vitro* protein refolding methods, we used our novel *in vivo* protein refolding technique to refold bacterially expressed BP1-LRP6. With optimized conditions, QQ-reagent efficiently delivered large quantity of bacterially expressed BP1-LRP6 inside HeLa cells within two hours (Figure 3-5 to 3-7). Protein refolding was extensively processed inside HeLa cells for three days (Figure 3-9). During these three days, the bacterially expressed BP1-LRP6 underwent refolding facilitated by the ER molecular chaperones and folding enzymes. Persistent unfolded or misfolded BP1-LRP6 could be actively degraded by the ERAD or UPR mechanism in the ER.

During BP1-LRP6 folding, MESD may play a key role to escort the properly folded BP1-LRP6 travel from the ER to Golgi. We delivered a large quantity of BP1-LRP6 inside Hela cells for refolding (Table 3-1). The quantity may be above the threshold of endogenous MESD, causing the refolded BP1-LRP6 to be retained inside of the Hela cells. Also, the refolded BP1-LRP6 was lacking a signal peptide, which may cause the difference in localization (intracellular versus secreted) of the refolded BP1-LRP6.

We used a nickel chelating column to purify the refolded BP1-LRP6 from HeLa cell lysates (Figure 3-10). We found our purified and potentially refolded BP1-LRP6 was greatly enriched for the monomeric form (Figure 3-11). The refolding efficiency and the final yield for this cysteine-rich, large-sized protein were so high that nearly 60% percent of the delivered BP1-LRP6 was refolded and re-purified. To futher purify our refolded BP1- LRP6, the protein could be applied to a size exclusion chromatography to separate the monomeric species

from the higher molecular weight oligmeric ones. A functional assay was performed to prove that the BP1-LRP6 obtained by our novel *in vivo* protein refolding technique was properly refolded and biologically functional (Figure 3-8). This provided more evidence that the refolded BP1-LRP6 was in a native conformation.

The secondary structure and folding properties of the refolded BP1-LRP6 was studied by far-UV circular dichroism (CD). Our CD data showed that the refolded BP1-LRP6 adopted a β-sheet rich conformation while the spectrum of the bacterially expressed BP1-LRP6 was more characterized as random coil (Figure 3-12). The DICHROWEB server further analyzed and quantitated the CD data (Table 3-2). The analysis showed that the refolded BP1-LRP6 and its homologue, the LDL receptor BP domain shared almost the same percentage of α-helix and β-sheet in their structure, suggesting that our refolded BP1-LRP6 adopted similar defined conformation as its homologue, the LDL receptor BP domain. Our *in vivo* protein refolding technique is unique because it gives a more physiological relevant in-cell environment for protein refolding. In this environment, not only the refolding is assisted by molecular chaperones and folding enzymes in the ER, but also the folding quality is strictly monitored by a complex quality control system inside the mammalian cells. Compared with traditional *in vitro* protein refolding methods, our method is much more efficient in large protein refolding, which has been demonstrated in the refolding of LBD-apoER2 (Chapter 2). However, endogenous ER chaperones and folding enzymes have their threshold for protein folding. This may cause the retention of

the refolded protein inside of the mammalian cells or partially refolding of the delivered protein. We confronted these problems when we challenged our *in vivo* protein refolding technique with cysteine-rich, large-size protein, BP1-LRP6. We hope we can overcome these problems by further optimizing our *in vivo* protein refolding technique or by engineering mammalian cells that can sustain large quantity protein refolding.

CHAPTER IV

Conclusion and future directions

4.1. Conclusion

Studies of *in vitro* protein refolding allow us to understand fundamental aspects of the folding mechanisms without the complications of the biological environment. From Anfinsen's RNase renaturation experiment, it has been long accepted that all the driving forces and guiding information of protein folding are embedded in the primary amino acid sequence of a protein and no external template is required. Under physiological conditions, the native state protein adopts the most thermodynamically stable conformation. *In vitro* refolding process is a good model for defining the types of intramolecular interaction that drive polypeptide folding and understanding the mechanisms of protein folding. However, *in vitro* refolding cannot accurately reflect the folding process of nascent proteins inside of cells due to the differences between protein folding in the intracellular environment and those in the test tube: 1) the intracellular environment is a highly crowded macromolecular environment; 2) folding inside of the cell is facilitated by molecular chaperones and folding enzymes and must be accomplished in the context of the synthesis of polypeptide chains on ribosomes; 3) cells are an intrinsically dynamic system. Protein folding experiments *in vitro* may not mimic the dynamics states *in vivo*. Ultimately, protein folding must be studied in a cellular context. It is necessary to develop new techniques using new *in vivo* systems to integrate the best aspects of both

the high-level bacterial expression system and the efficient cellular folding machinery of mammalian cells.

Our novel *in vivo* protein refolding technique uniquely integrates these two aspects. This technique first expresses recombinant proteins in *E. coli* with a very high-yield and then deliver the bacterially expressed pure protein directly into the *ER* of HeLa cells, using the QQ- protein delivery system, to properly refold the protein by the comprehensive protein folding machinery inside of mammalian cells. We drive the proteins into the inclusion bodies for the purpose of high yield and high purity. We then purify the recombinant proteins under denatured conditions since the bacterially expressed proteins will be refolded in the HeLa cells. The key step of this *in vivo* protein refolding technique is the QQ-protein delivery system that allows specific delivery of the bacterially expressed proteins into the *ER* for efficient refolding. This novel *in vivo* protein refolding technique offers a unique tool for protein folding study as well as for broader applications, such as therapeutical protein productions.

This thesis also presents the applications and optimizations of our novel *in vivo* protein refolding technique on two challenging proteins: LBD-apoER2 and BP1-LRP6. ApoER2 is an important LDL receptor super family member involved in ApoE metabolism in brain and the onset of Alzheimer's disease. The LBD-apoER2 contains 7 ligand binding repeats. Each repeat contains 6 cysteines forming 3 intra-repeat disulfide bonds, plus one Ca^{2+}-binding site. Bacterially expressed LBD-apoER2 formed non-native disulfide bonds, causing multiple conformations and aggregation. LRP6 is the co-receptor of Wnt signalling

pathway, involved in cell proliferation, differentiation and mobility. The homologue of BP1-LRP6, BP domain of LDL receptor, adopts a β-sheet rich conformation. In BP domain of LDL receptor, the YWTD β-propeller module is composed of six β-strand blades and the C-terminal EGF-like module is stabilized by three intra-domain disulfide bonds. Bacterially expressed BP1-LRP6 was misfolded and formed large amounts of oligomers.

By using our *in vivo* protein refolding technique, we efficiently delivered the misfolded proteins into the *ER* of HeLa cells. Under the physiological environment of *ER* and with the help of molecular chaperones and folding enzymes, the misfolded proteins were efficiently refolded into their native conformations. The refolded LBD-apoER2 displayed RAP binding function, implying its proper conformation. The refolded BP1-LRP6 also showed biological function of MESD binding ability, indicating the appearance of native conformation after refolding. Far-UV circular dichroism spectra of the refolded BP1-LRP6 indicated a β-sheet rich structure. The data analysis of the far-UV circular dichroism spectra provided indirect evidence that the refolded BP1-LRP6 adopted a remarkably similar secondary structure as its homologue --- BP domain of LDL receptor. Based on our results, we concluded that the application of our novel *in vivo* protein refolding method gave high yields of properly folded LBD-apoER2 and BP1-LRP6.

This novel *in vivo* protein refolding technique not only provides an efficient tool to obtain large quantities of properly folded recombinant proteins with stable isotopes for structural apporaches such as NMR spectroscopy, but also opens

interesting new perspectives for producing biological active polypeptides and proteins with therapeutic value on a large-scale. Moreover, this technique bridges the existing technical gap between *in vitro* protein folding and folding in living cells and can be used to gain a fundamental understanding of dynamic intracellular folding mechanisms in the living mammalian cells by studying the folding of the radioactive amino acid labelled recombinant proteins inside of mammalian cells. Or, as the fluorescence based techniques (such as fluorescent resonance energy transfer, FRET) develop, we can expect to study conformational changes of proteins in the living cell.

4.2. Future Directions

We have successfully refolded LBD-apoER2 and BP1-LRP6 into their native comformations with biologically functions by applying the *in vivo* protein refolding technique. The logical next step is to investigate the biochemical and structural characters of these properly folded proteins under physiologically relevant circumstances. One of strengthes of our *in vivo* protein refolding is that we can efficiently refold isotope-labelled bacterially expressed proteins into their native conformations. With isotope labelling, the refolded LBD-apoER2 and BP1-LRP6 can be dissolved in physiologically relavent buffers, ready for NMR structural studies. BP domains of LRP6 is involved in Wnt signalling pathway and LBD-apoER2 may be involved in the onset of Alzheimer disease. They are

potentially important drug targets. Detailed structure information on these two proteins can be very helpful for drug design.

The folding of the cysteine-rich ligand binding repeats of LBD-apoER2 relies on a special *ER/Golgi* resident molecular chaperone RAP (*138*). RAP escorts the ligand binding domain and inhibits pre-mature ligand binding through the secretary pathway. RAP exerts high binding affinity to the ligand binding domain of the LDLR family (K_D~ 1-10 nM) in the *ER (125, 138, 139).* But under the acidic environment of *Golgi,* RAP is released fom the binding complex by lowering the binding affinity to ligand binding repeats. The RAP releasing mechanism in *Golgi* is still not clear. Acoording to the x-ray crystal structure of RAP-D3 and LDLR-LA3-4, A low-pH induced unfolding model was proposed (*214*). In this model, the histidine residues which are buried in the hydrophobic interior of the RAP-D3 helical bundle are protonated under the low pH environment of *Golgi* and, in sequence, trigger unfolding of the helical bundle and dissociation of RAP-D3/LDLR-LA3-4 complex. However, Gettins' group found that domain D1 of RAP also bound LA repeats of LRP with very high affinity (*215*). In their low-pH induced YWTD displacing model, the flanking YWTD "propeller" domains competitively bind to RAP at the two binding sites in RAP-D1 and RAP-D3 under the low pH environment of *Golgi.* Thus, RAP is displaced by YWTD domains, releasing from LA repeats. The debating of these two models diverges at different study objects. The low-pH induced YWTD displacing model claimed for the acuteness by applying the intact RAP and three ligand binding repeats of LRP instead of RAP-D3 and two ligand binding repeats only of LDLR.

To clear the debating, we can conduct isothermal titration calorimetry (ITC) experiments on the protein-protein interactions between the intact RAP (available in our lab) and our refolded LBD-apoER2 which contains the whole ligand binding domain (seven ligand binding repeats) at different pH environments. This research will offer direct evidences to elucidate the RAP releasing mechnism in *Golgi*.

Protein chaperones are involved in the protein folding process inside of cells. For example, Bip, the abundant general ER chaperone, usually is the first chaperone binding to the nascent peptide chain to facilitate translocation of the nascent peptide chain into the ER lumen. PDI is one of the most important ER chaperones facilitating the native disulfide bond formation (*28*). RAP is the specific folding chaperone for the ligand-binding domain of LDLR superfamily members (*125*). MESD is the private chaperone for LRP5/6 BP domain folding (*139*). In our research, the folding threshold of endogenuous molecular chaperones and folding enzymes was the main factor that limits the *in vivo* protein refolding efficiency. Our QQ- protein delivery system is powerful. We can co-deliver important molecular chaperones or folding enzymes inside of cells with the proteins of interest to overcome the folding threshold problem and to further increase the refolding efficiency of the *in vivo* protein refolding technique.

On the other hand, the host mammalian cells could also be engineered for the purpose of improving *in vivo* protein refolding efficiency. Reports have provided examples of how the growth, survival, and productivity of the host cell can be improved through use of these genetic engineering techniques. Proto-

oncogenes, cell cycle control genes (cyclins), growth factor genes (e.g., insulin-like growth factor), and anti-apoptotic genes have been inserted into cell lines for the generation of superior recombinant protein production hosts (*216*). Improvements of *in vivo* protein refolding and post-translational modification can also be achieved by genetic engineering method. Stable cell lines can be generated by integrating key genes of molecular chaperones or folding enzymes into the host cell genome. These engineered host cell lines stably overexpress these important molecular chaperones or folding enzymes and they can be used as the refolding host cell lines to enhance the refolding ability. Thus, the protein refolding efficiency by using these engineered host cells may be significantly improved.

The establishment of a basis for understanding the determinants of protein folding has the potential to transform our understanding of many of the most perplexing issues surrounding protein folding diseases, and perhaps most importantly, the true nature of misfolded pathogenic protein species. Our *in vivo* protein refolding technique has opened the field for *in vivo* analysis of protein folding diseases, which has previously only been possible *in vitro*. We believe this approach may become a fundamentally important technique that could aid in understanding the protein folding and the complex chaperone networks inside of cells and its physiological relevance could, ultimately, be of clinical benefit.

Appendix: Correspondence

Table 1-1: Representative protein folding diseases Adapted from Dobson, C.M. (2001) The structural basis of protein folding and its links with human disease. *Philos Trans R Soc Lond B Biol Sci* 356, 133-145.

Figure 1-1: A schematic representation of the protein conformational states that are in interconversion Adapted from Chiti, F., and Dobson, C.M. (2006) Protein misfolding, functional amyloid, and human disease. *Annu Rev Biochem* 75, 333-366.

With permission from Dr. Christopher M. Dobson

Re: copyright permission request

CMD44 [cmd44@cam.ac.uk]

You replied on 6/14/2011 9:15 AM.

Sent: Tuesday, June 14, 2011 12:54 AM

To: Huang, Yuefei

Dear Yuefei Huang,

I am very happy to give you permission to reproduce the figures you mention in your email below, and indeed any others from our papers, with of course due acknowledgement.

Very best wishes,

Chris

Sent from my iPhone

On 13 Jun 2011, at 16:31, "Huang, Yuefei" <yuhuang@med.wayne.edu> wrote:

> Dear Dr. Christopher M. Dobson:

>

> My name is Yuefei Huang. I am a Ph.D. student at Wayne State University. I am writing for your permission to include in my thesis Table 1 of your published papers “the structural basis of protein folding and its link with human disease“ and Figure 2 of your published papers “protein misfolding, functional amyloid, and human disease”. They are very important for the literature review part of my thesis. The thesis will be made available to the public through the graduate school of Wayne State University. In addition, the thesis will be available on ProQuest (formerly UMI), and copies of the dissertation will be sold on demand. Please grant me permission to use your work. Could you email the permission to yuhuang@med.wayne.edu?

>

> Thank you for your help.

>

> Sincerely,

>

> Yuefei Huang, Ph.D. candidate student

> Dept. of Biochemistry and Molecular Biology

> Wayne State University, School of Medicine

> Detroit, MI. 48201

> Tel: 313-577-2682

> Email: yuhuang@med.wayne.edu

Table 1-2: Principal ER chaperones and foldases. Adapted from Maattanen, P., Gehring, K., Bergeron, J.J., and Thomas, D.Y. (2010) Protein quality control in the ER: the recognition of misfolded proteins. *Semin cell Dev Biol* 21, 500-511.

With permission from CCC

ELSEVIER LICENSE

TERMS AND CONDITIONS

Jun 13, 2011

This is a License Agreement between Yuefei Huang ("You") and Elsevier ("Elsevier") provided by Copyright Clearance Center ("CCC"). The license consists of your order details, the terms and conditions provided by Elsevier, and the payment terms and conditions.

All payments must be made in full to CCC. For payment instructions, please see information listed at the bottom of this form.

Supplier Elsevier Limited

The Boulevard,Langford Lane

Kidlington,Oxford,OX5 1GB,UK

Registered Company Number 1982084

Customer name Yuefei Huang

Customer address 540 E. Canfield Detroit, MI 48201

License number 2687251485636

License date Jun 13, 2011

Licensed content publisher Elsevier

Licensed content publication Seminars in Cell & Developmental Biology

Licensed content title Protein quality control in the ER: The recognition of misfolded proteins

Licensed content author Pekka Määttänen, Kalle Gehring, John J.M. Bergeron, David Y. Thomas

Licensed content date July 2010

Licensed content volume number 21

Licensed content issue number 5

Number of pages 12

Start Page 500

End Page 511

Type of Use reuse in a thesis/dissertation

Intended publisher of new work other

Rightslink Printable License 6/13/11 4:55 PM

https://s100.copyright.com/AppDispatchServlet Page 2 of 6

Portion figures/tables/illustrations Number of figures/tables/illustrations 1

Format electronic

Are you the author of this Elsevier article? No

Will you be translating? No

Order reference number

Title of your thesis/dissertation a novel in vivo protein refolding technique

Expected completion date Aug 2011

Estimated size (number of pages) 200

Elsevier VAT number GB 494 6272 12

Permissions price 0.00 USD

Figure 1-2: ER quality control of newly synthesized glycoproteins. Adapted from Hebert, D.N., and Molinari, M. (2007) In an dout of the ER: protein folding, quality control, degradation, and related human diseases. *Physiol Rev* 87, 1377-1408.

With permission from The American Physiological Society

Request for Permission to Reproduce Previously Published Material

(please save this file to your desktop, fill out, **save again**, and e-mail to permissions@the-aps.org)

Your Name: Yuefei Huang E-mail: yuhuang@med.wayne.edu

Affiliation:

University Address (for PhD students): Dept. of Biochemistry and Molecular biology, School of Medicine, 540 E. Canfield, Detroit, MI., 48201

Description of APS material to be reproduced (check all that apply):

☑ Figure ☐ Partial Article ☐ Abstract
☐ Table ☐ Full Article ☐ Book Chapter
Other (please describe):

Are you an author of the APS material to be reproduced? ○ Yes ◉ No

Please provide all applicable information about the APS material you wish to use:

Author(s): Daniel N. Hebert, Maurizio Molinari

Article or Chapter Title: In and out of the ER: protein folding, quality control, degradation, and related

Journal or Book Title: Physiol Rev.

Volume: 87 Page No(s): 1383 1377-1408 Figure No(s): 2 Table No(s):

Year: 2007 DOI: 10.1152/physrev.00050.2006

(If you are reproducing figures or tables from more than one article, please fill out and send a separate form for each citation.)

Please provide all applicable information about where the APS material will be used:

How will the APS material be used? (please select from drop-down list) Dissertation
If "other," please describe:

Title of publication or meeting where APS material will be used (if used in an article or book chapter, please provide the journal name or book title as well as the article/chapter title):
a novel in vivo protein refolding technique

Publisher (if journal or book): ProQuest LLC

URL (if website):

Date of Meeting or Publication: 2011

Will readers be charged for the material: ○ Yes ◉ No

Additional Information:

APPROVED
By pripka at 9:08 am, Jun 14, 2011

THE AMERICAN PHYSIOLOGICAL SOCIETY
9650 Rockville Pike, Bethesda, MD 20814-3991

Permission is granted for use of the material specified above, provided the publication is credited as the source, including the words "used with permission."

Rita Scheman
Publications Manager & Executive Editor

Figure 1-3: Signaling the unfolded protein response (UPR). Adapted from Kaufman, R.J., Back, S.H., Song, B., Han, J., and Hassie, J. (2010) The unfolded protein responses is required to maintain the integrity of the endoplasmic reticulum, prevent oxidative stress and preserve differentiation in beta-cells. *Diabetes Obes Metab* 12 Suppl 2, 99-107.

With permission from CCC

Rightslink Printable License 6/12/11 2:34 PM

JOHN WILEY AND SONS LICENSE

TERMS AND CONDITIONS

Jun 12, 2011

This is a License Agreement between Yuefei Huang ("You") and John Wiley and Sons ("John Wiley and Sons") provided by Copyright Clearance Center ("CCC"). The license consists of your order details, the terms and conditions provided by John Wiley and Sons, and the payment terms and conditions.

All payments must be made in full to CCC. For payment instructions, please see information listed at the bottom of this form.

License Number 2686630362972

License date Jun 12, 2011

Licensed content publisher John Wiley and Sons

Licensed content publication Diabetes, Obesity and Metabolism

Licensed content title The unfolded protein response is required to maintain the integrity of the endoplasmic reticulum, prevent oxidative stress and preserve differentiation in beta-cells

Licensed content author R. J. Kaufman,S. H. Back,B. Song,J. Han,J. Hassler

Licensed content date Oct 1, 2010

Start page 99

End page 107

Type of use Dissertation/Thesis

Requestor type University/Academic

Format Electronic

Portion Figure/table

Number of figures/tables 1

Number of extracts Original Wiley figure/table number(s) Figure 2

Will you be translating? No

Order reference number

Total 0.00 USD

Figure 1-4: The energy landscape depiction of protein folding. Adapted from Brooks, C.L., 3rd, Onuchic, J.N., and Wales, D.J. (2001) Statistical thermodynamics: Taking a walk on a landscape. *Science* 293, 612-613.

With permission from CCC

Rightslink Printable License 6/12/11 3:19 PM

THE AMERICAN ASSOCIATION FOR THE ADVANCEMENT OF SCIENCE LICENSE

TERMS AND CONDITIONS

Jun 12, 2011

This is a License Agreement between Yuefei Huang ("You") and The American Association for the Advancement of Science ("The American Association for the Advancement of Science") provided by Copyright Clearance Center ("CCC"). The license consists of your order details, the terms and conditions provided by The American Association for the Advancement of Science, and the payment terms and conditions.

All payments must be made in full to CCC. For payment instructions, please see information listed at the bottom of this form.

License Number 2686650058529

License date Jun 12, 2011

Licensed content publisher The American Association for the Advancement of Science

Licensed content publication Science

Licensed content title Taking a Walk on a Landscape

Licensed content author Charles L. Brooks III, José N. Onuchic, David J. Wales

Licensed content date Jul 27, 2001

Volume number 293

Issue number 5530

Type of Use Thesis / Dissertation

Requestor type Other Individual

Format Electronic

Portion Text Excerpt

Number of pages requested 1

Order reference number

Title of your thesis /dissertation a novel in vivo protein refolding technique

Expected completion date Aug 2011

Estimated size(pages) 200

Total 0.00 USD

Figure 2-1: Domain organization of human LDLR family members. Adapted from Gent, J., and Braakmain, J. (2004) low-density lipoprotein receptor structure and folding. *Cell Mol Life Sci* 61, 2461-2470.

With permission from CCC

Rightslink Printable License 6/13/11 3:55 PM

SPRINGER LICENSE

TERMS AND CONDITIONS

Jun 13, 2011

This is a License Agreement between Yuefei Huang ("You") and Springer ("Springer") provided by Copyright Clearance Center ("CCC"). The license consists of your order details, the terms and conditions provided by Springer, and the payment terms and conditions.

All payments must be made in full to CCC. For payment instructions, please see information listed at the bottom of this form.

License Number 2687230878990

License date Jun 13, 2011

Licensed content publisher Springer

Licensed content publication Cellular and Molecular Life Sciences

Licensed content title Low-density lipoprotein receptor structure and folding

Licensed content author J. Gent

Licensed content date Oct 1, 2004

Volume number 61

Issue number 19

Type of Use Thesis/Dissertation

Portion Figures

Author of this Springer article No

Order reference number

Title of your thesis /dissertation a novel in vivo protein refolding technique

Expected completion date Aug 2011

Estimated size(pages) 200

Total 0.00 USD

Figure 2-2: Domain organization of the LDL receptor and regions within the extracellular domain for which structural information has been obtained. Adapted from Herz, J. (2001) Deconstructing the LDL receptor- a rhapsody in pieces. *Nat struct Biol* 8, 476-478

With permission from Natur Publishing group and CCC

This is a License Agreement between Yuefei Huang ("You") and Nature Publishing Group ("Nature Publishing Group") provided by Copyright Clearance Center ("CCC"). The license consists of your order details, the terms and conditions provided by Nature Publishing Group, and the payment terms and conditions.

All payments must be made in full to CCC. For payment instructions, please see information listed at the bottom of this form.

License Number	2702560543989
License date	Jul 05, 2011
Licensed content publisher	Nature Publishing Group
Licensed content publication	Nature Structural and Molecular Biology
Licensed content title	Deconstructing the LDL receptor — a rhapsody in pieces
Licensed content author	Joachim Herz
Licensed content date	Jun 1, 2001
Volume number	6
Issue number	8
Type of Use	reuse in a thesis/dissertation
Requestor type	academic/educational
Format	electronic
Portion	figures/tables/illustrations

Number of figures/tables/illustrations	1
Figures	Figure 1- Domain organization of the LDL receptor and regions within the extracellular domain for which structural information has been obtained
Author of this NPG article	no
Your reference number	
Title of your thesis / dissertation	a novel in vivo protein refolding technique
Expected completion date	Aug 2011
Estimated size (number of pages)	200
Total	0.00 USD

Figure 2-3: The crystal structure of LDL receptor residue 1-699 at pH 5.3. Adapted from Rudenko, G., Henry, L., Henderson, K., Ichtchenko, K., Brown, M.S., Golstein, J.L., and Deisenhofer, J. (2002) Structure of the LDL receptor extracellular domain at the endosomal pH. *Science* 298, 2353-2358.

With permission from CCC

Rightslink Printable License 6/13/11 4:15 PM

THE AMERICAN ASSOCIATION FOR THE ADVANCEMENT OF SCIENCE LICENSE

TERMS AND CONDITIONS

Jun 13, 2011

This is a License Agreement between Yuefei Huang ("You") and The American Association for the Advancement of Science ("The American Association for the Advancement of Science") provided by Copyright Clearance Center ("CCC"). The license consists of your order details, the terms and conditions provided by The American Association for the Advancement of Science, and the payment terms and conditions.

All payments must be made in full to CCC. For payment instructions, please see

information listed at the bottom of this form.

License Number 2687240570257

License date Jun 13, 2011

Licensed content publisher The American Association for the Advancement of Science

Licensed content publication Science

Licensed content title Structure of the LDL Receptor Extracellular Domain at Endosomal pH

Licensed content author Gabby Rudenko, Lisa Henry, Keith Henderson, Konstantin Ichtchenko, Michael S. Brown, Joseph L. Goldstein, Johann Deisenhofer

Licensed content date Dec 20, 2002

Volume number 298

Issue number 5602

Type of Use Thesis / Dissertation

Requestor type Other Individual

Format Electronic

Portion Text Excerpt

Number of pages requested 1

Order reference number

Title of your thesis /dissertation a novel in vivo protein refolding technique

Expected completion date Aug 2011

Estimated size(pages) 200

Total 0.00 USD

Figure 2-4: LDLR biosynthesis and transport. Adapted from Maxfield, F.R., and Tabas, I. (2005) Role of cholesterol and lipid organization in disease. *Nature* 438, 612-621.

With permission from CCC

NATURE PUBLISHING GROUP LICENSE

TERMS AND CONDITIONS

Jun 17, 2011

This is a License Agreement between Yuefei Huang ("You") and Nature Publishing Group ("Nature Publishing Group") provided by Copyright Clearance Center ("CCC").

The license consists of your order details, the terms and conditions provided by Nature Publishing Group, and the payment terms and conditions.

All payments must be made in full to CCC. For payment instructions, please see information listed at the bottom of this form.

License Number 2691510726637

License date Jun 17, 2011

Licensed content publisher Nature Publishing Group

Licensed content publication Nature

Licensed content title Role of cholesterol and lipid organization in disease

Licensed content author Frederick R. Maxfield and Ira Tabas

Licensed content date Dec 1, 2005

Volume number 438

Issue number 7068

Type of Use reuse in a thesis/dissertation

Requestor type academic/educational

Format electronic

Portion figures/tables/illustrations

Number of figures/tables/illustrations 1

High-res required no

Figures Figure 1: Intracellular cholesterol transport

Author of this NPG article no

Your reference number

Title of your thesis /dissertation a novel in vivo protein refolding technique

Expected completion date Aug 2011

Estimated size (number of pages) 200

Total 0.00 USD

Figure 2-5: ER export and retrieval cycle of RAP- Role in LRP secretion. Adapted from Herz, J. (2006) The switch on the RAPper's necklace. *Mol Cell* 23, 451-455.

With permission from CCC

ELSEVIER LICENSE

TERMS AND CONDITIONS

Jun 13, 2011

This is a License Agreement between Yuefei Huang ("You") and Elsevier ("Elsevier") provided by Copyright Clearance Center ("CCC"). The license consists of your order details, the terms and conditions provided by Elsevier, and the payment terms and conditions.

All payments must be made in full to CCC. For payment instructions, please see information listed at the bottom of this form.

Supplier Elsevier Limited

The Boulevard,Langford Lane

Kidlington,Oxford,OX5 1GB,UK

Registered Company Number 1982084

Customer name Yuefei Huang

Customer address 540 E. Canfield Detroit, MI 48201

License number 2687240812918

License date Jun 13, 2011

Licensed content publisher Elsevier

Licensed content publication Molecular Cell

Licensed content title The Switch on the RAPper's Necklace…

Licensed content author Joachim Herz

Licensed content date 18 August 2006

Licensed content volume number 23

Licensed content issue number 4

Number of pages 5

Start Page 451

End Page 455

Type of Use reuse in a thesis/dissertation

Intended publisher of new work other

Portion figures/tables/illustrations

Number of figures/tables/illustrations 1

Format electronic

Are you the author of this Elsevier article? No

Will you be translating? No

Order reference number

Title of your thesis/dissertation a novel in vivo protein refolding technique

Expected completion date Aug 2011

Estimated size (number of pages) 200

Elsevier VAT number GB 494 6272 12

Permissions price 0.00 USD

Figure 3-1: A schematic representation of the Wnt signal transduction cascade. Adapted from Habas, R., and David, L.B. (2005) Dishevelled and Wnt signaling: is the nucleus the final frontier? *J. Biol 4,2*.

With permission from Biomedcentral

ref:00D2CUt.5002GDClA:ref]

"David Roman" <david.roman@biomedcentral.com> [david.roman@biomedcentral.com]

Sent: Tuesday, June 14, 2011 5:37 AM

To: Huang, Yuefei

Dear Dr Huang

Thank you for contacting BioMed Central.

Reproduction of figures or tables is permitted free of charge and without formal written permission from the publisher or the copyright holder, provided that the figure/table is original, BioMed Central is duly identified as the original publisher, and that proper attribution of authorship and the correct citation details are given as acknowledgment.

If you have any questions please do not hesitate to contact me.

Best wishes

David Roman

Customer Service Advisor

david.roman@biomedcentral.com

www.biomedcentral.com

----------Your Question/Comment ----------

To Whom It May Concern:

My name is Yuefei Huang, a Ph.D. student at Wayne State University. I am completing my thesis, entitled " a novel in vivo protein refolding technique". I would like your permission to reprint in my thesis excerpts from the following:

Habas, R., and Dawid, I. B. (2005) Dishevelled and Wnt signaling: is the nucleus the

final frontier?, J Biol 4, 2.

The excerpts to be reproduced are: Figure 1- A schematic representation of the Wnt signal transduction cascade. The requested permission extends to any future revisions and editions of my thesis by ProQuest Information and Learning (ProQuest) through its UMI® Dissertation Publishing business. ProQuest may produce and sell copies of my thesis on demand and will make my thesis available for free internet download through the Open Access publishing method required by California State University, Long Beach. Please grant me permission to use the work. You can email the permission to yuhuang@med.wayne.edu. Thank you very much.

Sincerely,

Yuefei Huang, Ph.D. candidate student

Dept. of Biochemistry and Molecular Biology

Wayne State University, School of Medicine

Detroit, MI. 48201

Tel: 313-577-2682

Email: yuhuang@med.wayne.edu

Figure 3-2: A simplified classical view of Wnt/β-catenin signaling. Adapted from Dobson, C.M. (2001) The structural basis of protein folding and its links with human disease. *Philos Trans R Soc Lond B Biol Sci* 356, 133-145.

With permission from Biologists

Re: copyright permission request

Sue Chamberlain [sue@biologists.com]

Sent: Tuesday, June 14, 2011 4:45 AM

To: Huang, Yuefei

Dear Yuefei,

Permission is granted with no charge.

The acknowledgement should state "reproduced / adapted with permission" and give the source journalname - the acknowledgement should either provide full citation details or refer to the relevant citation in thearticle reference list - the full citation details should include authors, journal, year, volume, issue and pagecitation.

Where appearing online or in other electronic media, a link should be provided to the original article (e.g.via DOI):-Development: dev.biologists.org

Best wishes,

Sue Chamberlain

On 13/6/11 23:44, "Huang, Yuefei" yuhuang@med.wayne.edu wrote:

> To Whom It May Concern:

>

> My name is Yuefei Huang, a Ph.D. student at Wayne State University. I am completing my thesis, entitled " a novel in vivo protein refolding technique". I would like your permission to reprint in my thesis excerpts from the following:

>

> He, X., Semenov, M., Tamai, K., and Zeng, X. (2004) LDL receptor-related proteins 5 and 6 in Wnt/beta-catenin signaling: arrows point the way, Development 131, 1663-1677.

>

> The excerpts to be reproduced are: Figure 1- A simplified prevailing view of Wnt/?-catenin signaling

>

> The requested permission extends to any future revisions and editions of my thesis by ProQuest Information and Learning (ProQuest) through its UMI® Dissertation Publishing business. ProQuest may produce and sell copies of my thesis on demand and will make my thesis available for free internet download through the Open Access publishing method required by California State University, Long Beach. Please grant me permission to use the work. You can email the permission to yuhuang@med.wayne.edu.

>

> Thank you very much.

>

> Sincerely,

>

> Yuefei Huang, Ph.D. candidate student

> Dept. of Biochemistry and Molecular Biology

> Wayne State University, School of Medicine

> Detroit, MI. 48201

> Tel: 313-577-2682

> Email: yuhuang@med.wayne.edu

Figure 3-3: Ribbon representation of the YWTD domain and adjacent C-terminal EGF-like module (E3) of the LDL receptor, colored to point out the six YWTD repeats of the six-bladed propeller. Adapted from Jeon,H., Meng, W., Takagi, J., Eck, M.J., Springer, T.A., and Blacklow, S.C. (2001) Implications for familial hypercholesterolemia from the structure of the LDL receptor YWTD-EGF domain pair. *Nat Struct Biol* 8, 499-504.

With permission from Dr. Stephen Blacklow

Re: copyright permission request

Stephen Blacklow [sblacklow@rics.bwh.harvard.edu]

You replied on 6/19/2011 9:12 AM.

Sent: Friday, June 17, 2011 4:09 PM

To: Huang, Yuefei

Dear Yuefei,

You are welcome to use the figure!

Best of luck preparing and defending your thesis.

Steve

On Jun 17, 2011, at 4:05 PM, Huang, Yuefei wrote:

> Dear Dr. Stephen C. Blacklow:

>

> My name is Yuefei Huang. I am a Ph.D. student at Wayne State University. I am writing for your permission to include in my thesis Figure 2 of your published papers "Implications for familial hypercholesterolemia from the structure of the LDL receptor YWTD-EGF domain pair". It is very important for the literature review part of my thesis. The thesis will be made available to the public through the graduate school of Wayne State University. In addition, the thesis will be available on ProQuest (formerly UMI), and copies of the dissertation will be sold on demand. Please grant me permission to use your work. Could you email the permission to yuhuang@med.wayne.edu?

>

> Thank you for your help.

>

> Sincerely,

>

> Yuefei Huang, Ph.D. candidate student

> Dept. of Biochemistry and Molecular Biology

> Wayne State University, School of Medicine

> Detroit, MI. 48201

> Tel: 313-577-2682

> Email: yuhuang@med.wayne.edu

--

Stephen Blacklow, M.D., Ph.D.

Professor of Biological Chemistry and Molecular Pharmacology

Professor of Pathology

Director, MD, PhD Program Basic Sciences Track

Harvard Medical School

Dana Farber Cancer Institute, Department of Cancer Biology

44 Binney Street

Boston, MA 02115

phone: (617) 632-2453

email: sblacklow@rics.bwh.harvard.edu

http://pathology.bwh.harvard.edu/blacklow

REFERENCES

1. Harding, H. P., Zhang, Y., and Ron, D. (1999) Protein translation and folding are coupled by an endoplasmic-reticulum-resident kinase, *Nature 397*, 271-274.
2. Sonenberg, N., and Dever, T. E. (2003) Eukaryotic translation initiation factors and regulators, *Current opinion in structural biology 13*, 56-63.
3. Rutkowski, D. T., and Kaufman, R. J. (2004) A trip to the ER: coping with stress, *Trends Cell Biol 14*, 20-28.
4. Rutkowski, D. T., and Kaufman, R. J. (2004) A trip to the ER: coping with stress, *Trends in cell biology 14*, 20-28.
5. Scheuner, D., and Kaufman, R. J. (2008) The unfolded protein response: a pathway that links insulin demand with beta-cell failure and diabetes, *Endocr Rev 29*, 317-333.
6. Chiti, F., and Dobson, C. M. (2006) Protein misfolding, functional amyloid, and human disease, *Annual review of biochemistry 75*, 333-366.
7. Shen, J., Chen, X., Hendershot, L., and Prywes, R. (2002) ER stress regulation of ATF6 localization by dissociation of BiP/GRP78 binding and unmasking of Golgi localization signals, *Dev Cell 3*, 99-111.
8. Pauling, L., and Corey, R. B. (1951) The structure of synthetic polypeptides, *Proc Natl Acad Sci U S A 37*, 241-250.
9. Pauling, L., and Corey, R. B. (1951) The polypeptide-chain configuration in hemoglobin and other globular proteins, *Proc Natl Acad Sci U S A 37*, 282-285.

10. Kendrew, J. C. (1958) Architecture of a protein molecule, *Nature 182*, 764-767.

11. Kendrew, J. C., Bodo, G., Dintzis, H. M., Parrish, R. G., Wyckoff, H., and Phillips, D. C. (1958) A three-dimensional model of the myoglobin molecule obtained by x-ray analysis, *Nature 181*, 662-666.

12. Maattanen, P., Gehring, K., Bergeron, J. J., and Thomas, D. Y. (2010) Protein quality control in the ER: the recognition of misfolded proteins, *Semin Cell Dev Biol 21*, 500-511.

13. Dobson, C. M., and Hore, P. J. (1998) Kinetic studies of protein folding using NMR spectroscopy, *Nat Struct Biol 5 Suppl*, 504-507.

14. Stefani, M., and Dobson, C. M. (2003) Protein aggregation and aggregate toxicity: new insights into protein folding, misfolding diseases and biological evolution, *J Mol Med 81*, 678-699.

15. Dobson, C. M. (2001) The structural basis of protein folding and its links with human disease, *Philos Trans R Soc Lond B Biol Sci 356*, 133-145.

16. Sacchettini, J. C., and Kelly, J. W. (2002) Therapeutic strategies for human amyloid diseases, *Nat Rev Drug Discov 1*, 267-275.

17. May, B. C., Fafarman, A. T., Hong, S. B., Rogers, M., Deady, L. W., Prusiner, S. B., and Cohen, F. E. (2003) Potent inhibition of scrapie prion replication in cultured cells by bis-acridines, *Proc Natl Acad Sci U S A 100*, 3416-3421.

18. Korth, C., May, B. C., Cohen, F. E., and Prusiner, S. B. (2001) Acridine and phenothiazine derivatives as pharmacotherapeutics for prion disease, *Proc Natl Acad Sci U S A 98*, 9836-9841.

19. Nikolova, P. V., Wong, K. B., DeDecker, B., Henckel, J., and Fersht, A. R. (2000) Mechanism of rescue of common p53 cancer mutations by second-site suppressor mutations, *EMBO J 19*, 370-378.

20. Bullock, A. N., and Fersht, A. R. (2001) Rescuing the function of mutant p53, *Nat Rev Cancer 1*, 68-76.

21. Voeltz, G. K., Rolls, M. M., and Rapoport, T. A. (2002) Structural organization of the endoplasmic reticulum, *EMBO Rep 3*, 944-950.

22. Borgese, N., Francolini, M., and Snapp, E. (2006) Endoplasmic reticulum architecture: structures in flux, *Curr Opin Cell Biol 18*, 358-364.

23. Scott, M., Lu, G., Hallett, M., and Thomas, D. Y. (2004) The Hera database and its use in the characterization of endoplasmic reticulum proteins, *Bioinformatics 20*, 937-944.

24. Helenius, A., and Aebi, M. (2004) Roles of N-linked glycans in the endoplasmic reticulum, *Annual review of biochemistry 73*, 1019-1049.

25. Otero, J. H., Lizak, B., and Hendershot, L. M. (2010) Life and death of a BiP substrate, *Semin Cell Dev Biol 21*, 472-478.

26. Meunier, L., Usherwood, Y. K., Chung, K. T., and Hendershot, L. M. (2002) A subset of chaperones and folding enzymes form multiprotein complexes in endoplasmic reticulum to bind nascent proteins, *Mol Biol Cell 13*, 4456-4469.

27. Shimizu, Y., and Hendershot, L. M. (2007) Organization of the functions and components of the endoplasmic reticulum, *Adv Exp Med Biol 594*, 37-46.

28. Dorner, A. J., Wasley, L. C., Raney, P., Haugejorden, S., Green, M., and Kaufman, R. J. (1990) The stress response in Chinese hamster ovary cells. Regulation of ERp72 and protein disulfide isomerase expression and secretion, *J Biol Chem 265*, 22029-22034.

29. Schmid, F. X., and Baldwin, R. L. (1978) Acid catalysis of the formation of the slow-folding species of RNase A: evidence that the reaction is proline isomerization, *Proc Natl Acad Sci U S A 75*, 4764-4768.

30. Lilie, H., Rudolph, R., and Buchner, J. (1995) Association of antibody chains at different stages of folding: prolyl isomerization occurs after formation of quaternary structure, *J Mol Biol 248*, 190-201.

31. McPhie, P. (1980) Kinetic studies on the unfolding and refolding of pepsinogen in urea. The nature of the rate-limiting step, *J Biol Chem 255*, 4048-4052.

32. Egea, P. F., Stroud, R. M., and Walter, P. (2005) Targeting proteins to membranes: structure of the signal recognition particle, *Current opinion in structural biology 15*, 213-220.

33. Schwartz, T. U. (2007) Origins and evolution of cotranslational transport to the ER, *Adv Exp Med Biol 607*, 52-60.

34. Dejgaard, K., Theberge, J. F., Heath-Engel, H., Chevet, E., Tremblay, M. L., and Thomas, D. Y. (2010) Organization of the Sec61 translocon,

studied by high resolution native electrophoresis, *J Proteome Res 9*, 1763-1771.

35. Zapun, A., Darby, N. J., Tessier, D. C., Michalak, M., Bergeron, J. J., and Thomas, D. Y. (1998) Enhanced catalysis of ribonuclease B folding by the interaction of calnexin or calreticulin with ERp57, *J Biol Chem 273*, 6009-6012.

36. Dominguez, M., Dejgaard, K., Fullekrug, J., Dahan, S., Fazel, A., Paccaud, J. P., Thomas, D. Y., Bergeron, J. J., and Nilsson, T. (1998) gp25L/emp24/p24 protein family members of the cis-Golgi network bind both COP I and II coatomer, *J Cell Biol 140*, 751-765.

37. Wickner, S., Maurizi, M. R., and Gottesman, S. (1999) Posttranslational quality control: folding, refolding, and degrading proteins, *Science 286*, 1888-1893.

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

39. Ruddock, L. W., and Molinari, M. (2006) N-glycan processing in ER quality control, *J Cell Sci 119*, 4373-4380.

40. Tsai, B., Ye, Y., and Rapoport, T. A. (2002) Retro-translocation of proteins from the endoplasmic reticulum into the cytosol, *Nat Rev Mol Cell Biol 3*, 246-255.

41. Werner, E. D., Brodsky, J. L., and McCracken, A. A. (1996) Proteasome-dependent endoplasmic reticulum-associated protein degradation: an

unconventional route to a familiar fate, *Proc Natl Acad Sci U S A 93*, 13797-13801.

42. Ushioda, R., Hoseki, J., Araki, K., Jansen, G., Thomas, D. Y., and Nagata, K. (2008) ERdj5 is required as a disulfide reductase for degradation of misfolded proteins in the ER, *Science 321*, 569-572.

43. Hosokawa, N., Kamiya, Y., and Kato, K. (2010) The role of MRH domain-containing lectins in ERAD, *Glycobiology 20*, 651-660.

44. Meusser, B., and Sommer, T. (2004) Vpu-mediated degradation of CD4 reconstituted in yeast reveals mechanistic differences to cellular ER-associated protein degradation, *Mol Cell 14*, 247-258.

45. Hebert, D. N., and Molinari, M. (2007) In and out of the ER: protein folding, quality control, degradation, and related human diseases, *Physiol Rev 87*, 1377-1408.

46. Denisov, A. Y., Maattanen, P., Sprules, T., Thomas, D. Y., and Gehring, K. (2007) 1H, 13C and 15N resonance assignments of the bb' domains of human protein disulfide isomerase, *Biomol NMR Assign 1*, 129-130.

47. Ichijo, H. (1999) From receptors to stress-activated MAP kinases, *Oncogene 18*, 6087-6093.

48. Nadanaka, S., Yoshida, H., Sato, R., and Mori, K. (2006) Analysis of ATF6 activation in Site-2 protease-deficient Chinese hamster ovary cells, *Cell Struct Funct 31*, 109-116.

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

50. Iwakoshi, N. N., Lee, A. H., Vallabhajosyula, P., Otipoby, K. L., Rajewsky, K., and Glimcher, L. H. (2003) Plasma cell differentiation and the unfolded protein response intersect at the transcription factor XBP-1, *Nat Immunol 4*, 321-329.

51. Oda, Y., Okada, T., Yoshida, H., Kaufman, R. J., Nagata, K., and Mori, K. (2006) Derlin-2 and Derlin-3 are regulated by the mammalian unfolded protein response and are required for ER-associated degradation, *J Cell Biol 172*, 383-393.

52. Lee, A. H., Scapa, E. F., Cohen, D. E., and Glimcher, L. H. (2008) Regulation of hepatic lipogenesis by the transcription factor XBP1, *Science 320*, 1492-1496.

53. Wu, J., Rutkowski, D. T., Dubois, M., Swathirajan, J., Saunders, T., Wang, J., Song, B., Yau, G. D., and Kaufman, R. J. (2007) ATF6alpha optimizes long-term endoplasmic reticulum function to protect cells from chronic stress, *Dev Cell 13*, 351-364.

54. Yoshida, H., Matsui, T., Hosokawa, N., Kaufman, R. J., Nagata, K., and Mori, K. (2003) A time-dependent phase shift in the mammalian unfolded protein response, *Dev Cell 4*, 265-271.

55. Reimold, A. M., Ponath, P. D., Li, Y. S., Hardy, R. R., David, C. S., Strominger, J. L., and Glimcher, L. H. (1996) Transcription factor B cell

lineage-specific activator protein regulates the gene for human X-box binding protein 1, *J Exp Med 183*, 393-401.

56. Harding, H. P., Zeng, H., Zhang, Y., Jungries, R., Chung, P., Plesken, H., Sabatini, D. D., and Ron, D. (2001) Diabetes mellitus and exocrine pancreatic dysfunction in perk-/- mice reveals a role for translational control in secretory cell survival, *Mol Cell 7*, 1153-1163.

57. Scheuner, D., Song, B., McEwen, E., Liu, C., Laybutt, R., Gillespie, P., Saunders, T., Bonner-Weir, S., and Kaufman, R. J. (2001) Translational control is required for the unfolded protein response and in vivo glucose homeostasis, *Mol Cell 7*, 1165-1176.

58. Kaufman, R. J., Back, S. H., Song, B., Han, J., and Hassler, J. (2010) The unfolded protein response is required to maintain the integrity of the endoplasmic reticulum, prevent oxidative stress and preserve differentiation in beta-cells, *Diabetes Obes Metab 12 Suppl 2*, 99-107.

59. Sivashanmugam, A., Murray, V., Cui, C., Zhang, Y., Wang, J., and Li, Q. (2009) Practical protocols for production of very high yields of recombinant proteins using Escherichia coli, *Protein Sci 18*, 936-948.

60. Jeon, W. B. (2010) Retrospective analyses of the bottleneck in purification of eukaryotic proteins from Escherichia coli as affected by molecular weight, cysteine content and isoelectric point, *BMB Rep 43*, 319-324.

61. Braun, P., Hu, Y., Shen, B., Halleck, A., Koundinya, M., Harlow, E., and LaBaer, J. (2002) Proteome-scale purification of human proteins from bacteria, *Proc Natl Acad Sci U S A 99*, 2654-2659.

62. Christendat, D., Yee, A., Dharamsi, A., Kluger, Y., Savchenko, A., Cort, J. R., Booth, V., Mackereth, C. D., Saridakis, V., Ekiel, I., Kozlov, G., Maxwell, K. L., Wu, N., McIntosh, L. P., Gehring, K., Kennedy, M. A., Davidson, A. R., Pai, E. F., Gerstein, M., Edwards, A. M., and Arrowsmith, C. H. (2000) Structural proteomics of an archaeon, *Nat Struct Biol 7*, 903-909.

63. Lesley, S. A., Kuhn, P., Godzik, A., Deacon, A. M., Mathews, I., Kreusch, A., Spraggon, G., Klock, H. E., McMullan, D., Shin, T., Vincent, J., Robb, A., Brinen, L. S., Miller, M. D., McPhillips, T. M., Miller, M. A., Scheibe, D., Canaves, J. M., Guda, C., Jaroszewski, L., Selby, T. L., Elsliger, M. A., Wooley, J., Taylor, S. S., Hodgson, K. O., Wilson, I. A., Schultz, P. G., and Stevens, R. C. (2002) Structural genomics of the Thermotoga maritima proteome implemented in a high-throughput structure determination pipeline, *Proc Natl Acad Sci U S A 99*, 11664-11669.

64. Yee, A., Chang, X., Pineda-Lucena, A., Wu, B., Semesi, A., Le, B., Ramelot, T., Lee, G. M., Bhattacharyya, S., Gutierrez, P., Denisov, A., Lee, C. H., Cort, J. R., Kozlov, G., Liao, J., Finak, G., Chen, L., Wishart, D., Lee, W., McIntosh, L. P., Gehring, K., Kennedy, M. A., Edwards, A. M., and Arrowsmith, C. H. (2002) An NMR approach to structural proteomics, *Proc Natl Acad Sci U S A 99*, 1825-1830.

65. Possee, R. D. (1997) Baculoviruses as expression vectors, *Curr Opin Biotechnol 8*, 569-572.

66. Zhao, Y., Chapman, D. A., and Jones, I. M. (2003) Improving baculovirus recombination, *Nucleic Acids Res 31*, E6-6.

67. Durocher, Y., Perret, S., and Kamen, A. (2002) High-level and high-throughput recombinant protein production by transient transfection of suspension-growing human 293-EBNA1 cells, *Nucleic Acids Res 30*, E9.

68. Braun, P., and LaBaer, J. (2003) High throughput protein production for functional proteomics, *Trends Biotechnol 21*, 383-388.

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

70. Onuchic, J. N., Luthey-Schulten, Z., and Wolynes, P. G. (1997) Theory of protein folding: the energy landscape perspective, *Annu Rev Phys Chem 48*, 545-600.

71. Onuchic, J. N., Socci, N. D., Luthey-Schulten, Z., and Wolynes, P. G. (1996) Protein folding funnels: the nature of the transition state ensemble, *Fold Des 1*, 441-450.

72. Brooks, C. L., 3rd, Onuchic, J. N., and Wales, D. J. (2001) Statistical thermodynamics. Taking a walk on a landscape, *Science 293*, 612-613.

73. Socci, N. D., Onuchic, J. N., and Wolynes, P. G. (1998) Protein folding mechanisms and the multidimensional folding funnel, *Proteins 32*, 136-158.

74. Shakhnovich, E. I. (1998) Folding nucleus: specific or multiple? Insights from lattice models and experiments, *Fold Des 3*, R108-111; discussion R107.

75. Wolynes, P. G. (1997) Folding funnels and energy landscapes of larger proteins within the capillarity approximation, *Proc Natl Acad Sci U S A 94*, 6170-6175.

76. Dill, K. A., and Chan, H. S. (1997) From Levinthal to pathways to funnels, *Nat Struct Biol 4*, 10-19.

77. Bryngelson, J. D., Onuchic, J. N., Socci, N. D., and Wolynes, P. G. (1995) Funnels, pathways, and the energy landscape of protein folding: a synthesis, *Proteins 21*, 167-195.

78. Kubelka, J., Hofrichter, J., and Eaton, W. A. (2004) The protein folding 'speed limit', *Current opinion in structural biology 14*, 76-88.

79. Clark, E. D. B. (1998) Refolding of recombinant proteins, *Curr Opin Biotechnol 9*, 157-163.

80. Horowitz, P. M., and Simon, D. (1986) The enzyme rhodanese can be reactivated after denaturation in guanidinium chloride, *J Biol Chem 261*, 13887-13891.

81. Gu, Z., Zhu, X., Ni, S., Zhou, H., and Su, Z. (2003) Inhibition of aggregation by media selection, sample loading and elution in size exclusion chromatographic refolding of denatured bovine carbonic anhydrase B, *J Biochem Biophys Methods 56*, 165-175.

82. Stempfer, G., Holl-Neugebauer, B., and Rudolph, R. (1996) Improved refolding of an immobilized fusion protein, *Nat Biotechnol 14*, 329-334.

83. Cho, T. H., Ahn, S. J., and Lee, E. K. (2001) Refolding of protein inclusion bodies directly from E. coli homogenate using expanded bed adsorption chromatography, *Bioseparation 10*, 189-196.

84. Bai, J. Z., Ban, Y., Bian, J. G., Cai, X., Chang, J. F., Chen, H. F., Chen, H. S., Chen, J., Chen, J. C., Chen, Y. B., Chi, S. P., Chu, Y. P., Cui, X. Z., Dai, Y. M., Dai, Y. S., Dong, L. Y., Du, S. X., Du, Z. Z., Dunwoodie, W., Fang, J., Fang, S. S., Fu, C. D., Fu, H. Y., Fu, L. P., Gao, C. S., Gao, M. L., Gao, Y. N., Gong, M. Y., Gong, W. X., Gu, S. D., Guo, Y. N., Guo, Y. Q., Guo, Z. J., Han, S. W., Harris, F. A., He, J., He, K. L., He, M., He, X., Heng, Y. K., Hong, T., Hu, H. M., Hu, T., Huang, G. S., Huang, L., Huang, X. P., Izen, J. M., Ji, X. B., Jiang, C. H., Jiang, X. S., Jin, D. P., Jin, S., Jin, Y., Jones, B. D., Ke, Z. J., Kong, D., Lai, Y. F., Li, F., Li, G., Li, H. H., Li, J., Li, J. C., Li, K., Li, Q. J., Li, R. B., Li, R. Y., Li, W., Li, W. G., Li, X. Q., Li, X. S., Liu, C. F., Liu, C. X., Liu, F., Liu, H. M., Liu, J. B., Liu, J. P., Liu, R. G., Liu, Y., Liu, Z. A., Liu, Z. X., Lou, X. C., Lu, G. R., Lu, F., Lu, H. J., Lu, J. G., Lu, Z. J., Luo, X. L., Ma, E. C., Ma, F. C., Ma, J. M., Malchow, R., Mao, Z. P., Meng, X. C., Mo, X. H., Nie, J., Nie, Z. D., Olsen, S. L., Paluselli, D., Peng, H. P., Qi, N. D., Qian, C. D., Qiu, J. F., Rong, G., Shen, D. L., Shen, H., Shen, X. Y., Sheng, H. Y., Shi, F., Song, L. W., Sun, H. S., Sun, S. S., Sun, Y. Z., Sun, Z. J., Tang, S. Q., Tang, X., Tian, D., Tian, Y. R., Toki, W., Tong, G. L., Varner, G. S., Wang, J., Wang, J. Z., Wang, L., Wang, L. S., Wang, M., Wang, P., Wang, P. L., Wang, W. F., Wang, Y. F., Wang, Z., Wang, Z. Y., Wei, C. L., Wu, N., Xia, X. M., Xie, X.

X., Xu, G. F., Xu, Y., Xue, S. T., Yan, M. L., Yan, W. B., Yang, G. A., Yang, H. X., Yang, J., Yang, S. D., Ye, M. H., Ye, Y. X., Ying, J., Yu, C. S., Yu, G. W., Yuan, C. Z., Yuan, J. M., Yuan, Y., Yue, Q., Zang, S. L., Zeng, Y., Zhang, B. X., Zhang, B. Y., Zhang, C. C., Zhang, D. H., Zhang, H. Y., Zhang, J., Zhang, J. M., Zhang, J. W., Zhang, L. S., Zhang, Q. J., Zhang, S. Q., Zhang, X. Y., Zhang, Y. J., Zhang, Y., Zhang, Y. Y., Zhang, Z. P., Zhao, D. X., Zhao, J., Zhao, J. W., Zhao, P. P., Zhao, W. R., Zhao, Y. B., Zhao, Z. G., Zheng, J. P., Zheng, L. S., Zheng, Z. P., Zhong, X. C., Zhou, B. Q., Zhou, G. M., Zhou, L., Zhou, N. F., Zhu, K. J., Zhu, Q. M., Zhu, Y., Zhu, Y. C., Zhu, Y. S., Zhu, Z. A., Zhuang, B. A., and Zou, B. S. (2003) Observation of a near-threshold enhancement in the pp mass spectrum from radiative J/psi-->gammapp decays, *Phys Rev Lett 91*, 022001.

85. Geng, X. X., Quan, L. N., Ma, R., and Tang, L. P. (2011) [Effects of As2O3 and all-trans retinoic acid on the growth of HeLa cell line and their relation with gene NDRG1.], *Zhonghua Zhong Liu Za Zhi 33*, 8-12.

86. Yoshii, H., Furuta, T., Yonehara, T., Ito, D., Linko, Y. Y., and Linko, P. (2000) Refolding of denatured/reduced lysozyme at high concentration with diafiltration, *Biosci Biotechnol Biochem 64*, 1159-1165.

87. Carrio, M. M., and Villaverde, A. (2001) Protein aggregation as bacterial inclusion bodies is reversible, *FEBS Lett 489*, 29-33.

88. Robinson, C. R., and Sligar, S. G. (1995) Hydrostatic and osmotic pressure as tools to study macromolecular recognition, *Methods Enzymol 259*, 395-427.

89. St John, R. J., Carpenter, J. F., Balny, C., and Randolph, T. W. (2001) High pressure refolding of recombinant human growth hormone from insoluble aggregates. Structural transformations, kinetic barriers, and energetics, *J Biol Chem 276*, 46856-46863.

90. St John, R. J., Carpenter, J. F., and Randolph, T. W. (1999) High pressure fosters protein refolding from aggregates at high concentrations, *Proc Natl Acad Sci U S A 96*, 13029-13033.

91. Simon, S., Peskin, C., and Oster, G. (1992) What drives the translocations of proteins, *Proc Natl Acad Sci USA 89*, 3770 - 3774.

92. Hagen, A., Hatton, T., and Wang, D. (1990) Protein refolding in reversed micelles, *Biotechnology and bioengineering 35*, 955 - 965.

93. Li, Y., Lu, W., Schwartz, A. L., and Bu, G. (2002) Receptor-associated protein facilitates proper folding and maturation of the low-density lipoprotein receptor and its class 2 mutants, *Biochemistry 41*, 4921-4928.

94. Woycechowsky, K., Hook, B., and Raines, R. (2003) Catalysis of protein folding by an immobilized small-molecule dithiol, *Biotechnology progress 19*, 1307 - 1314.

95. Woycechowsky, K., Wittrup, K., and Raines, R. (1999) A small-molecule catalyst of protein folding in vitro and in vivo, *Chem Biol 6*, 871 - 879.

96. Neumann, E., Schaefer-Ridder, M., Wang, Y., and Hofschneider, P. H. (1982) Gene transfer into mouse lyoma cells by electroporation in high electric fields, *The EMBO journal 1*, 841-845.

97. Marrero, M. B., Schieffer, B., Paxton, W. G., Schieffer, E., and Bernstein, K. E. (1995) Electroporation of pp60c-src antibodies inhibits the angiotensin II activation of phospholipase C-gamma 1 in rat aortic smooth muscle cells, *J Biol Chem 270*, 15734-15738.

98. Nolkrantz, K., Farre, C., Hurtig, K. J., Rylander, P., and Orwar, O. (2002) Functional screening of intracellular proteins in single cells and in patterned cell arrays using electroporation, *Anal Chem 74*, 4300-4305.

99. Rui, M., Chen, Y., Zhang, Y., and Ma, D. (2002) Transfer of anti-TFAR19 monoclonal antibody into HeLa cells by in situ electroporation can inhibit the apoptosis, *Life Sci 71*, 1771-1778.

100. Capecchi, M. R. (1980) High efficiency transformation by direct microinjection of DNA into cultured mammalian cells, *Cell 22*, 479-488.

101. Abarzua, P., LoSardo, J. E., Gubler, M. L., and Neri, A. (1995) Microinjection of monoclonal antibody PAb421 into human SW480 colorectal carcinoma cells restores the transcription activation function to mutant p53, *Cancer Res 55*, 3490-3494.

102. Theiss, C., and Meller, K. (2002) Microinjected anti-actin antibodies decrease gap junctional intercellular commmunication in cultured astrocytes, *Exp Cell Res 281*, 197-204.

103. Schwarze, S. R., Hruska, K. A., and Dowdy, S. F. (2000) Protein transduction: unrestricted delivery into all cells?, *Trends Cell Biol 10*, 290-295.

104. Fenton, M., Bone, N., and Sinclair, A. J. (1998) The efficient and rapid import of a peptide into primary B and T lymphocytes and a lymphoblastoid cell line, *J Immunol Methods 212*, 41-48.

105. Felgner, P. L., Gadek, T. R., Holm, M., Roman, R., Chan, H. W., Wenz, M., Northrop, J. P., Ringold, G. M., and Danielsen, M. (1987) Lipofection: a highly efficient, lipid-mediated DNA-transfection procedure, *Proc Natl Acad Sci U S A 84*, 7413-7417.

106. Zabner, J., Fasbender, A. J., Moninger, T., Poellinger, K. A., and Welsh, M. J. (1995) Cellular and molecular barriers to gene transfer by a cationic lipid, *J Biol Chem 270*, 18997-19007.

107. Zelphati, O., Wang, Y., Kitada, S., Reed, J. C., Felgner, P. L., and Corbeil, J. (2001) Intracellular delivery of proteins with a new lipid-mediated delivery system, *J Biol Chem 276*, 35103-35110.

108. Boussif, O., Lezoualc'h, F., Zanta, M. A., Mergny, M. D., Scherman, D., Demeneix, B., and Behr, J. P. (1995) A versatile vector for gene and oligonucleotide transfer into cells in culture and in vivo: polyethylenimine, *Proc Natl Acad Sci U S A 92*, 7297-7301.

109. Mislick, K. A., and Baldeschwieler, J. D. (1996) Evidence for the role of proteoglycans in cation-mediated gene transfer, *Proc Natl Acad Sci U S A 93*, 12349-12354.

110. Behr, J. P. (1994) Gene transfer with synthetic cationic amphiphiles: prospects for gene therapy, *Bioconjug Chem 5*, 382-389.

111. Demeneix, B., Behr, J., Boussif, O., Zanta, M. A., Abdallah, B., and Remy, J. (1998) Gene transfer with lipospermines and polyethylenimines, *Adv Drug Deliv Rev 30*, 85-95.

112. Godbey, W. T., Wu, K. K., and Mikos, A. G. (1999) Poly(ethylenimine) and its role in gene delivery, *J Control Release 60*, 149-160.

113. Remy-Kristensen, A., Clamme, J. P., Vuilleumier, C., Kuhry, J. G., and Mely, Y. (2001) Role of endocytosis in the transfection of L929 fibroblasts by polyethylenimine/DNA complexes, *Biochim Biophys Acta 1514*, 21-32.

114. Culi, J., and Mann, R. S. (2003) Boca, an endoplasmic reticulum protein required for wingless signaling and trafficking of LDL receptor family members in Drosophila, *Cell 112*, 343-354.

115. Hsieh, J. C., Lee, L., Zhang, L., Wefer, S., Brown, K., DeRossi, C., Wines, M. E., Rosenquist, T., and Holdener, B. C. (2003) Mesd encodes an LRP5/6 chaperone essential for specification of mouse embryonic polarity, *Cell 112*, 355-367.

116. Lippincott-Schwartz, J., and Patterson, G. H. (2003) Development and use of fluorescent protein markers in living cells, *Science 300*, 87-91.

117. Zhang, K., Shen, X., Wu, J., Sakaki, K., Saunders, T., Rutkowski, D. T., Back, S. H., and Kaufman, R. J. (2006) Endoplasmic reticulum stress activates cleavage of CREBH to induce a systemic inflammatory response, *Cell 124*, 587-599.

118. Studier, F. W. (2005) Protein production by auto-induction in high density shaking cultures, *Protein expression and purification 41*, 207-234.

119. Li, Q., Huang, Y., Xiao, N., Murray, V., Chen, J., and Wang, J. (2008) Real time investigation of protein folding, structure, and dynamics in living cells, *Methods Cell Biol 90*, 287-325.

120. Gent, J., and Braakman, I. (2004) Low-density lipoprotein receptor structure and folding, *Cell Mol Life Sci 61*, 2461-2470.

121. Rudenko, G., and Deisenhofer, J. (2003) The low-density lipoprotein receptor: ligands, debates and lore, *Current opinion in structural biology 13*, 683-689.

122. Herz, J., and Bock, H. H. (2002) Lipoprotein receptors in the nervous system, *Annual review of biochemistry 71*, 405-434.

123. Hussain, M. M., Maxfield, F. R., Mas-Oliva, J., Tabas, I., Ji, Z. S., Innerarity, T. L., and Mahley, R. W. (1991) Clearance of chylomicron remnants by the low density lipoprotein receptor-related protein/alpha 2-macroglobulin receptor, *J Biol Chem 266*, 13936-13940.

124. Krieger, M., and Herz, J. (1994) Structures and functions of multiligand lipoprotein receptors: macrophage scavenger receptors and LDL receptor-related protein (LRP), *Annual review of biochemistry 63*, 601-637.

125. Willnow, T. E., Rohlmann, A., Horton, J., Otani, H., Braun, J. R., Hammer, R. E., and Herz, J. (1996) RAP, a specialized chaperone, prevents ligand-induced ER retention and degradation of LDL receptor-related endocytic receptors, *The EMBO journal 15*, 2632-2639.

126. Herz, J. (2006) The switch on the RAPper's necklace, *Mol Cell 23*, 451-455.

127. Fass, D., Blacklow, S., Kim, P. S., and Berger, J. M. (1997) Molecular basis of familial hypercholesterolaemia from structure of LDL receptor module, *Nature 388*, 691-693.

128. Daly, N. L., Scanlon, M. J., Djordjevic, J. T., Kroon, P. A., and Smith, R. (1995) Three-dimensional structure of a cysteine-rich repeat from the low-density lipoprotein receptor, *Proc Natl Acad Sci U S A 92*, 6334-6338.

129. Daly, N. L., Djordjevic, J. T., Kroon, P. A., and Smith, R. (1995) Three-dimensional structure of the second cysteine-rich repeat from the human low-density lipoprotein receptor, *Biochemistry 34*, 14474-14481.

130. Jeon, H., Meng, W., Takagi, J., Eck, M. J., Springer, T. A., and Blacklow, S. C. (2001) Implications for familial hypercholesterolemia from the structure of the LDL receptor YWTD-EGF domain pair, *Nat Struct Biol 8*, 499-504.

131. Rudenko, G., Henry, L., Henderson, K., Ichtchenko, K., Brown, M. S., Goldstein, J. L., and Deisenhofer, J. (2002) Structure of the LDL receptor extracellular domain at endosomal pH, *Science 298*, 2353-2358.

132. Beglova, N., Jeon, H., Fisher, C., and Blacklow, S. C. (2004) Cooperation between fixed and low pH-inducible interfaces controls lipoprotein release by the LDL receptor, *Mol Cell 16*, 281-292.

133. Herz, J. (2001) Deconstructing the LDL receptor--a rhapsody in pieces, *Nat Struct Biol 8*, 476-478.

134. Hussain, M. M., Strickland, D. K., and Bakillah, A. (1999) The mammalian low-density lipoprotein receptor family, *Annu Rev Nutr 19*, 141-172.

135. Herz, J. (2001) The LDL receptor gene family: (un)expected signal transducers in the brain, *Neuron 29*, 571-581.

136. Gotthardt, M., Trommsdorff, M., Nevitt, M. F., Shelton, J., Richardson, J. A., Stockinger, W., Nimpf, J., and Herz, J. (2000) Interactions of the low density lipoprotein receptor gene family with cytosolic adaptor and scaffold proteins suggest diverse biological functions in cellular communication and signal transduction, *J Biol Chem 275*, 25616-25624.

137. Maxfield, F. R., and Tabas, I. (2005) Role of cholesterol and lipid organization in disease, *Nature 438*, 612-621.

138. Bu, G. (1998) Receptor-associated protein: a specialized chaperone and antagonist for members of the LDL receptor gene family, *Curr Opin Lipidol 9*, 149-155.

139. Bu, G., Geuze, H. J., Strous, G. J., and Schwartz, A. L. (1995) 39 kDa receptor-associated protein is an ER resident protein and molecular chaperone for LDL receptor-related protein, *The EMBO journal 14*, 2269-2280.

140. Trommsdorff, M., Gotthardt, M., Hiesberger, T., Shelton, J., Stockinger, W., Nimpf, J., Hammer, R. E., Richardson, J. A., and Herz, J. (1999) Reeler/Disabled-like disruption of neuronal migration in knockout mice lacking the VLDL receptor and ApoE receptor 2, *Cell 97*, 689-701.

141. Pinson, K. I., Brennan, J., Monkley, S., Avery, B. J., and Skarnes, W. C. (2000) An LDL-receptor-related protein mediates Wnt signalling in mice, *Nature 407*, 535-538.

142. Tamai, K., Semenov, M., Kato, Y., Spokony, R., Liu, C., Katsuyama, Y., Hess, F., Saint-Jeannet, J. P., and He, X. (2000) LDL-receptor-related proteins in Wnt signal transduction, *Nature 407*, 530-535.

143. Wehrli, M., Dougan, S. T., Caldwell, K., O'Keefe, L., Schwartz, S., Vaizel-Ohayon, D., Schejter, E., Tomlinson, A., and DiNardo, S. (2000) arrow encodes an LDL-receptor-related protein essential for Wingless signalling, *Nature 407*, 527-530.

144. Howell, B. W., Lanier, L. M., Frank, R., Gertler, F. B., and Cooper, J. A. (1999) The disabled 1 phosphotyrosine-binding domain binds to the internalization signals of transmembrane glycoproteins and to phospholipids, *Mol Cell Biol 19*, 5179-5188.

145. Stockinger, W., Brandes, C., Fasching, D., Hermann, M., Gotthardt, M., Herz, J., Schneider, W. J., and Nimpf, J. (2000) The reelin receptor ApoER2 recruits JNK-interacting proteins-1 and -2, *J Biol Chem 275*, 25625-25632.

146. Trommsdorff, M., Borg, J. P., Margolis, B., and Herz, J. (1998) Interaction of cytosolic adaptor proteins with neuronal apolipoprotein E receptors and the amyloid precursor protein, *J Biol Chem 273*, 33556-33560.

147. Brown, M. S., and Goldstein, J. L. (1986) A receptor-mediated pathway for cholesterol homeostasis, *Science 232*, 34-47.

148. Hobbs, H. H., Russell, D. W., Brown, M. S., and Goldstein, J. L. (1990) The LDL receptor locus in familial hypercholesterolemia: mutational analysis of a membrane protein, *Annu Rev Genet 24*, 133-170.

149. Villeger, L., Abifadel, M., Allard, D., Rabes, J. P., Thiart, R., Kotze, M. J., Beroud, C., Junien, C., Boileau, C., and Varret, M. (2002) The UMD-LDLR database: additions to the software and 490 new entries to the database, *Hum Mutat 20*, 81-87.

150. Ai, M., Holmen, S. L., Van Hul, W., Williams, B. O., and Warman, M. L. (2005) Reduced affinity to and inhibition by DKK1 form a common mechanism by which high bone mass-associated missense mutations in LRP5 affect canonical Wnt signaling, *Mol Cell Biol 25*, 4946-4955.

151. Boyden, L. M., Mao, J., Belsky, J., Mitzner, L., Farhi, A., Mitnick, M. A., Wu, D., Insogna, K., and Lifton, R. P. (2002) High bone density due to a mutation in LDL-receptor-related protein 5, *N Engl J Med 346*, 1513-1521.

152. Gong, Y., Slee, R. B., Fukai, N., Rawadi, G., Roman-Roman, S., Reginato, A. M., Wang, H., Cundy, T., Glorieux, F. H., Lev, D., Zacharin, M., Oexle, K., Marcelino, J., Suwairi, W., Heeger, S., Sabatakos, G., Apte, S., Adkins, W. N., Allgrove, J., Arslan-Kirchner, M., Batch, J. A., Beighton, P., Black, G. C., Boles, R. G., Boon, L. M., Borrone, C., Brunner, H. G., Carle, G. F., Dallapiccola, B., De Paepe, A., Floege, B., Halfhide, M. L., Hall, B., Hennekam, R. C., Hirose, T., Jans, A., Juppner, H., Kim, C. A., Keppler-Noreuil, K., Kohlschuetter, A., LaCombe, D., Lambert, M., Lemyre, E., Letteboer, T., Peltonen, L., Ramesar, R. S., Romanengo, M., Somer, H., Steichen-Gersdorf, E., Steinmann, B., Sullivan, B., Superti-Furga, A., Swoboda, W., van den Boogaard, M. J., Van Hul, W., Vikkula, M., Votruba, M., Zabel, B., Garcia, T., Baron, R., Olsen, B. R., and

Warman, M. L. (2001) LDL receptor-related protein 5 (LRP5) affects bone accrual and eye development, *Cell 107*, 513-523.

153. Van Wesenbeeck, L., Cleiren, E., Gram, J., Beals, R. K., Benichou, O., Scopelliti, D., Key, L., Renton, T., Bartels, C., Gong, Y., Warman, M. L., De Vernejoul, M. C., Bollerslev, J., and Van Hul, W. (2003) Six novel missense mutations in the LDL receptor-related protein 5 (LRP5) gene in different conditions with an increased bone density, *Am J Hum Genet 72*, 763-771.

154. D'Arcangelo, G., Panuccio, G., Tancredi, V., and Avoli, M. (2005) Repetitive low-frequency stimulation reduces epileptiform synchronization in limbic neuronal networks, *Neurobiol Dis 19*, 119-128.

155. D'Arcangelo, G., Homayouni, R., Keshvara, L., Rice, D. S., Sheldon, M., and Curran, T. (1999) Reelin is a ligand for lipoprotein receptors, *Neuron 24*, 471-479.

156. Motoi, Y., Aizawa, T., Haga, S., Nakamura, S., Namba, Y., and Ikeda, K. (1999) Neuronal localization of a novel mosaic apolipoprotein E receptor, LR11, in rat and human brain, *Brain Res 833*, 209-215.

157. Kim, D. H., Magoori, K., Inoue, T. R., Mao, C. C., Kim, H. J., Suzuki, H., Fujita, T., Endo, Y., Saeki, S., and Yamamoto, T. T. (1997) Exon/intron organization, chromosome localization, alternative splicing, and transcription units of the human apolipoprotein E receptor 2 gene, *J Biol Chem 272*, 8498-8504.

158. Kim, D. H., Iijima, H., Goto, K., Sakai, J., Ishii, H., Kim, H. J., Suzuki, H., Kondo, H., Saeki, S., and Yamamoto, T. (1996) Human apolipoprotein E receptor 2. A novel lipoprotein receptor of the low density lipoprotein receptor family predominantly expressed in brain, *J Biol Chem 271*, 8373-8380.

159. Brandes, C., Kahr, L., Stockinger, W., Hiesberger, T., Schneider, W. J., and Nimpf, J. (2001) Alternative splicing in the ligand binding domain of mouse ApoE receptor-2 produces receptor variants binding reelin but not alpha 2-macroglobulin, *J Biol Chem 276*, 22160-22169.

160. Riddell, D. R., Sun, X. M., Stannard, A. K., Soutar, A. K., and Owen, J. S. (2001) Localization of apolipoprotein E receptor 2 to caveolae in the plasma membrane, *J Lipid Res 42*, 998-1002.

161. Hiesberger, T., Trommsdorff, M., Howell, B. W., Goffinet, A., Mumby, M. C., Cooper, J. A., and Herz, J. (1999) Direct binding of Reelin to VLDL receptor and ApoE receptor 2 induces tyrosine phosphorylation of disabled-1 and modulates tau phosphorylation, *Neuron 24*, 481-489.

162. Benhayon, D., Magdaleno, S., and Curran, T. (2003) Binding of purified Reelin to ApoER2 and VLDLR mediates tyrosine phosphorylation of Disabled-1, *Brain Res Mol Brain Res 112*, 33-45.

163. Brich, J., Shie, F. S., Howell, B. W., Li, R., Tus, K., Wakeland, E. K., Jin, L. W., Mumby, M., Churchill, G., Herz, J., and Cooper, J. A. (2003) Genetic modulation of tau phosphorylation in the mouse, *J Neurosci 23*, 187-192.

164. Gebhardt, C., Del Turco, D., Drakew, A., Tielsch, A., Herz, J., Frotscher, M., and Deller, T. (2002) Abnormal positioning of granule cells alters afferent fiber distribution in the mouse fascia dentata: morphologic evidence from reeler, apolipoprotein E receptor 2-, and very low density lipoprotein receptor knockout mice, *J Comp Neurol 445*, 278-292.

165. Ashcom, J. D., Tiller, S. E., Dickerson, K., Cravens, J. L., Argraves, W. S., and Strickland, D. K. (1990) The human alpha 2-macroglobulin receptor: identification of a 420-kD cell surface glycoprotein specific for the activated conformation of alpha 2-macroglobulin, *J Cell Biol 110*, 1041-1048.

166. Striekland, D. K., Ashcom, J. D., Williams, S., Battey, F., Behre, E., McTigue, K., Battey, J. F., and Argraves, W. S. (1991) Primary structure of alpha 2-macroglobulin receptor-associated protein. Human homologue of a Heymann nephritis antigen, *J Biol Chem 266*, 13364-13369.

167. Bu, G., Maksymovitch, E. A., Geuze, H., and Schwartz, A. L. (1994) Subcellular localization and endocytic function of low density lipoprotein receptor-related protein in human glioblastoma cells, *J Biol Chem 269*, 29874-29882.

168. Bu, G., Williams, S., Strickland, D. K., and Schwartz, A. L. (1992) Low density lipoprotein receptor-related protein/alpha 2-macroglobulin receptor is an hepatic receptor for tissue-type plasminogen activator, *Proc Natl Acad Sci U S A 89*, 7427-7431.

169. Herz, J., Goldstein, J. L., Strickland, D. K., Ho, Y. K., and Brown, M. S. (1991) 39-kDa protein modulates binding of ligands to low density lipoprotein receptor-related protein/alpha 2-macroglobulin receptor, *J Biol Chem 266*, 21232-21238.

170. Moestrup, S. K., and Gliemann, J. (1991) Analysis of ligand recognition by the purified alpha 2-macroglobulin receptor (low density lipoprotein receptor-related protein). Evidence that high affinity of alpha 2-macroglobulin-proteinase complex is achieved by binding to adjacent receptors, *J Biol Chem 266*, 14011-14017.

171. Nykjaer, A., Petersen, C. M., Moller, B., Jensen, P. H., Moestrup, S. K., Holtet, T. L., Etzerodt, M., Thogersen, H. C., Munch, M., Andreasen, P. A., and et al. (1992) Purified alpha 2-macroglobulin receptor/LDL receptor-related protein binds urokinase.plasminogen activator inhibitor type-1 complex. Evidence that the alpha 2-macroglobulin receptor mediates cellular degradation of urokinase receptor-bound complexes, *J Biol Chem 267*, 14543-14546.

172. Orth, K., Madison, E. L., Gething, M. J., Sambrook, J. F., and Herz, J. (1992) Complexes of tissue-type plasminogen activator and its serpin inhibitor plasminogen-activator inhibitor type 1 are internalized by means of the low density lipoprotein receptor-related protein/alpha 2-macroglobulin receptor, *Proc Natl Acad Sci U S A 89*, 7422-7426.

173. Williams, S. E., Ashcom, J. D., Argraves, W. S., and Strickland, D. K. (1992) A novel mechanism for controlling the activity of alpha 2-

macroglobulin receptor/low density lipoprotein receptor-related protein. Multiple regulatory sites for 39-kDa receptor-associated protein, *J Biol Chem 267*, 9035-9040.

174. Iadonato, S. P., Bu, G., Maksymovitch, E. A., and Schwartz, A. L. (1993) Interaction of a 39 kDa protein with the low-density-lipoprotein-receptor-related protein (LRP) on rat hepatoma cells, *Biochem J 296 (Pt 3)*, 867-875.

175. Battey, F. D., Gafvels, M. E., FitzGerald, D. J., Argraves, W. S., Chappell, D. A., Strauss, J. F., 3rd, and Strickland, D. K. (1994) The 39-kDa receptor-associated protein regulates ligand binding by the very low density lipoprotein receptor, *J Biol Chem 269*, 23268-23273.

176. Jacobsen, L., Madsen, P., Moestrup, S. K., Lund, A. H., Tommerup, N., Nykjaer, A., Sottrup-Jensen, L., Gliemann, J., and Petersen, C. M. (1996) Molecular characterization of a novel human hybrid-type receptor that binds the alpha2-macroglobulin receptor-associated protein, *J Biol Chem 271*, 31379-31383.

177. Kounnas, M. Z., Morris, R. E., Thompson, M. R., FitzGerald, D. J., Strickland, D. K., and Saelinger, C. B. (1992) The alpha 2-macroglobulin receptor/low density lipoprotein receptor-related protein binds and internalizes Pseudomonas exotoxin A, *J Biol Chem 267*, 12420-12423.

178. Medh, J. D., Fry, G. L., Bowen, S. L., Pladet, M. W., Strickland, D. K., and Chappell, D. A. (1995) The 39-kDa receptor-associated protein modulates

lipoprotein catabolism by binding to LDL receptors, *J Biol Chem 270*, 536-540.

179. Stockinger, W., Hengstschlager-Ottnad, E., Novak, S., Matus, A., Huttinger, M., Bauer, J., Lassmann, H., Schneider, W. J., and Nimpf, J. (1998) The low density lipoprotein receptor gene family. Differential expression of two alpha2-macroglobulin receptors in the brain, *J Biol Chem 273*, 32213-32221.

180. Jansens, A., van Duijn, E., and Braakman, I. (2002) Coordinated nonvectorial folding in a newly synthesized multidomain protein, *Science 298*, 2401-2403.

181. Radford, S. E. (2003) Co-translocational misfolding in the ER of living cells, *Nat Struct Biol 10*, 153-154.

182. Freedman, R. B., Klappa, P., and Ruddock, L. W. (2002) Protein disulfide isomerases exploit synergy between catalytic and specific binding domains, *EMBO reports 3*, 136-140.

183. Tu, B. P., and Weissman, J. S. (2004) Oxidative protein folding in eukaryotes: mechanisms and consequences, *J Cell Biol 164*, 341-346.

184. Frand, A. R., Cuozzo, J. W., and Kaiser, C. A. (2000) Pathways for protein disulphide bond formation, *Trends Cell Biol 10*, 203-210.

185. Koduri, V., and Blacklow, S. C. (2001) Folding determinants of LDL receptor type A modules, *Biochemistry 40*, 12801-12807.

186. Bu, G., and Rennke, S. (1996) Receptor-associated protein is a folding chaperone for low density lipoprotein receptor-related protein, *J Biol Chem 271*, 22218-22224.

187. Culi, J., Springer, T. A., and Mann, R. S. (2004) Boca-dependent maturation of beta-propeller/EGF modules in low-density lipoprotein receptor proteins, *The EMBO journal 23*, 1372-1380.

188. Dolmer, K., and Gettins, P. G. (2006) Three complement-like repeats compose the complete alpha2-macroglobulin binding site in the second ligand binding cluster of the low density lipoprotein receptor-related protein, *J Biol Chem 281*, 34189-34196.

189. Dobson, C. M. (2004) Principles of protein folding, misfolding and aggregation, *Semin Cell Dev Biol 15*, 3-16.

190. Myant, N. B. (1990) Current approaches to the genetics of coronary heart disease (CHD) including an account of work done at Hammersmith Hospital, *Boll Soc Ital Biol Sper 66*, 1015-1041.

191. Jeon, H., and Blacklow, S. C. (2005) Structure and physiologic function of the low-density lipoprotein receptor, *Annual review of biochemistry 74*, 535-562.

192. He, X., Semenov, M., Tamai, K., and Zeng, X. (2004) LDL receptor-related proteins 5 and 6 in Wnt/beta-catenin signaling: arrows point the way, *Development 131*, 1663-1677.

193. Strickland, D. K., Gonias, S. L., and Argraves, W. S. (2002) Diverse roles for the LDL receptor family, *Trends Endocrinol Metab 13*, 66-74.

194. Loukinova, E., Ranganathan, S., Kuznetsov, S., Gorlatova, N., Migliorini, M. M., Loukinov, D., Ulery, P. G., Mikhailenko, I., Lawrence, D. A., and Strickland, D. K. (2002) Platelet-derived growth factor (PDGF)-induced tyrosine phosphorylation of the low density lipoprotein receptor-related protein (LRP). Evidence for integrated co-receptor function betwenn LRP and the PDGF, *J Biol Chem 277*, 15499-15506.

195. Rijsewijk, F., Schuermann, M., Wagenaar, E., Parren, P., Weigel, D., and Nusse, R. (1987) The Drosophila homolog of the mouse mammary oncogene int-1 is identical to the segment polarity gene wingless, *Cell 50*, 649-657.

196. Wodarz, A., and Nusse, R. (1998) Mechanisms of Wnt signaling in development, *Annu Rev Cell Dev Biol 14*, 59-88.

197. Lustig, B., and Behrens, J. (2003) The Wnt signaling pathway and its role in tumor development, *J Cancer Res Clin Oncol 129*, 199-221.

198. Giles, R. H., van Es, J. H., and Clevers, H. (2003) Caught up in a Wnt storm: Wnt signaling in cancer, *Biochim Biophys Acta 1653*, 1-24.

199. Habas, R., and Dawid, I. B. (2005) Dishevelled and Wnt signaling: is the nucleus the final frontier?, *J Biol 4*, 2.

200. Adler, P. N. (2002) Planar signaling and morphogenesis in Drosophila, *Dev Cell 2*, 525-535.

201. Veeman, M. T., Axelrod, J. D., and Moon, R. T. (2003) A second canon. Functions and mechanisms of beta-catenin-independent Wnt signaling, *Dev Cell 5*, 367-377.

202. Hurlstone, A., and Clevers, H. (2002) T-cell factors: turn-ons and turn-offs, *The EMBO journal 21*, 2303-2311.

203. Angers, S., and Moon, R. T. (2009) Proximal events in Wnt signal transduction, *Nature reviews. Molecular cell biology 10*, 468-477.

204. Kelly, O. G., Pinson, K. I., and Skarnes, W. C. (2004) The Wnt co-receptors Lrp5 and Lrp6 are essential for gastrulation in mice, *Development 131*, 2803-2815.

205. Mani, A., Radhakrishnan, J., Wang, H., Mani, M. A., Nelson-Williams, C., Carew, K. S., Mane, S., Najmabadi, H., Wu, D., and Lifton, R. P. (2007) LRP6 mutation in a family with early coronary disease and metabolic risk factors, *Science 315*, 1278-1282.

206. Lillis, A. P., Van Duyn, L. B., Murphy-Ullrich, J. E., and Strickland, D. K. (2008) LDL receptor-related protein 1: unique tissue-specific functions revealed by selective gene knockout studies, *Physiol Rev 88*, 887-918.

207. Marzolo, M. P., and Bu, G. (2009) Lipoprotein receptors and cholesterol in APP trafficking and proteolytic processing, implications for Alzheimer's disease, *Semin Cell Dev Biol 20*, 191-200.

208. Anelli, T., and Sitia, R. (2008) Protein quality control in the early secretory pathway, *The EMBO journal 27*, 315-327.

209. Chen, J., Liu, C. C., Li, Q., Nowak, C., Bu, G., and Wang, J. (2011) Two structural and functional domains of MESD required for proper folding and trafficking of LRP5/6, *Structure 19*, 313-323.

210. Freedman, R. B., Brockway, B. E., and Lambert, N. (1984) Protein disulphide-isomerase and the formation of native disulphide bonds, *Biochem Soc Trans 12*, 929-932.

211. Liu, C. C., Pearson, C., and Bu, G. (2009) Cooperative folding and ligand-binding properties of LRP6 beta-propeller domains, *J Biol Chem 284*, 15299-15307.

212. Whitmore, L., and Wallace, B. A. (2004) DICHROWEB, an online server for protein secondary structure analyses from circular dichroism spectroscopic data, *Nucleic Acids Res 32*, W668-673.

213. Greenfield, N. J. (2006) Using circular dichroism spectra to estimate protein secondary structure, *Nat Protoc 1*, 2876-2890.

214. Estrada, K., Fisher, C., and Blacklow, S. C. (2008) Unfolding of the RAP-D3 helical bundle facilitates dissociation of RAP-receptor complexes, *Biochemistry 47*, 1532-1539.

215. Lazic, A., Dolmer, K., Strickland, D. K., and Gettins, P. G. (2006) Dissection of RAP-LRP interactions: binding of RAP and RAP fragments to complement-like repeats 7 and 8 from ligand binding cluster II of LRP, *Arch Biochem Biophys 450*, 167-175.

216. Arden, N., Nivtechanyong,T., & Betenbaugh, M.J. (2004) Cell engineering blocks stress and improves biotherapeutic production, *Bioprocessing 3*, 23-28.

ABSTRACT

A NOVEL *IN VIVO* PROTEIN REFOLDING TECHNIQUE

by

YUEFEI HUANG

August 2011

Advisor: Dr. Jianjun Wang

Major: Biochemistry and Molecular Biology

Degree: Doctor of Philosophy

Proteins perform their functions in their native folded states and misfolding of proteins may cause severe diseases, including Alzheimer's disease, Parkinson's disease, prion disease and diabetes. Understanding protein folding is important for us to engineer proteins to treat these diseases. For protein therapeutics, large quantities of properly folded and functional proteins are required. The current technology produces recombinant proteins using either eukaryotic or prokaryotic expression system, both of them have major problems that prevent production of large quantities of properly folded and functional human proteins for protein therapeutics.

Although the eukaryotic cells have comprehensive folding machinery that contains chaperones and folding enzymes and a complex quality control (QC) system to ensure that only properly folded proteins will be generated to perform their functions, either intracellular or extracellular, the protein yield is usually very

low. Protein production using this system is usually costly. In contrast, prokaryotic cells can be used to produce large quantities of recombinant human proteins at a low cost. However, the produced human proteins using prokaryotic cells usually misfold and are not functional due to the much simpler protein folding machinery and QC system of these prokaryotic cells. To solve this problem, the *in vitro* protein refolding technique has been developed that either mimics the intracellular redox conditions to promote protein folding at a diluted concentration or uses column chromatography to refold the misfolded recombinant proteins. Although this *in vitro* protein folding technique has some success for small proteins with simple folds, the refolding efficiency is generally very low. For large proteins of complex folds of multiple domains, this *in vitro* protein refolding technique is usually not working.

To solve these challenges, our lab recently developed an *in vivo* protein refolding technique that uses the intracellular folding machinery and QC system of the Endoplasmic reticulum (ER) of mammalian cells to refold the misfolded recombinant proteins produced using bacterial expression system. This novel technique uses the QQ-protein delivery technology developed in our lab to directly deliver bacterially expressed proteins into the ER for refolding. We showed that the intracellular folding machinery of mammalian cells had a large capacity to properly refold large quantities of misfolded bacterially expressed proteins and the QC system of the mammalian cells ensured that only properly folded proteins followed the normal intracellular trafficking pathway as their endogenous counterparts. Since the refolded proteins contain an affinity tag, we

can purify the properly refolded proteins. This *in vivo* refolded technique takes the advantage of the high yield prokaryotic expression system and the comprehensive protein folding machinery/QC system of mammalian cells to efficiently produce large quantities of properly folded and biologically functional proteins.

I optimized this *in vivo* protein refolding technique for the beta-propeller/EGF domain I of LDL receptor-related protein 6 (BP1-LRP6) and the ligand-binding domain of apolipoprotein E receptor 2 (LBD-ApoER2). These two proteins contain a large number of cysteines that form intracellular disulfide bonds. The folding of these two proteins is very challenging. I performed optimizations of experimental conditions that allow me to produce large quantities of properly folded and functional BP1-LRP6 and LBD-apoER2. The yield of refolding is about 20-60%, depending on different proteins, allowing me to produce milligram quantity of properly refolded and functional BP1-LRP6 and LBD-apoER2. The Far-UV Circular Dichroism (CD) Spectrum of refolded BP1-LRP6 showed a high percentage of beta-sheet which is consistent with the x-ray crystal structure of the beta-propeller/EGF domain of low-density lipoprotein receptor (LDLR). Refolded LBD-apoER2 showed the biological function of active binding the chaperone receptor-associated protein (RAP) in the ligand-blotting assay. My results suggested that, as a new tool, this protein refolding technique can be used to produce large quantities of properly folded and biologically functional proteins for many applications including protein therapeutics to treat human disease, structural biology and protein folding studies.

AUTOBIOGRAPHICAL STATEMENT
YUEFEI HUANG

CONTACT INFORMATION

Department of Biochemistry and Molecular Biology, 540 E.Canfield Avenue,

Wayne State University School of Medicine Detroit, MI, 48201

EDUCATION

September 2005 – current Ph.D. in Biochemistry, Wayne State University, Detroit, MI, USA

September 1997 – July 2001 B.S in Microbiology, Sichuan University, Chengdu, China.

PUBLICATIONS

- Qianqian Li, **Yuefei Huang**, Nan xiao, Victoria Murray, Jianglei Chen and Jianjun Wang (2008). Chapter 14: Real time investigation of protein folding, structure, and dynamics in living cells. *Methods In Cell Biology.* **90**:287-325.
- Victoria Murray, Jianglei Chen, **Yuefei Huang,** Qianqian Li and Jianjun Wang (2010). Preparation of very-high-yield recombinant proteins using novel high-cell-density bacterial expression methods. *Cold Spring Harb Protoc*. **8**, pdb.prot5475.