

ETD Archive

2013

The Objectification of Women in Cane

Claudia M. Davis
Cleveland State University

Follow this and additional works at: <https://engagedscholarship.csuohio.edu/etdarchive>

 Part of the [English Language and Literature Commons](#)
How does access to this work benefit you? Let us know!

Recommended Citation

Davis, Claudia M., "The Objectification of Women in Cane" (2013). *ETD Archive*. 836.
<https://engagedscholarship.csuohio.edu/etdarchive/836>

This Thesis is brought to you for free and open access by EngagedScholarship@CSU. It has been accepted for inclusion in ETD Archive by an authorized administrator of EngagedScholarship@CSU. For more information, please contact library.es@csuohio.edu.

**STRUCTURAL BASIS OF LMAN1 CARGO CAPTURE IN ER
& RELEASE IN ERGIC**

VAIJAYANTI DAS

Master of Science in Physiology

University of Calcutta

August, 2002

Submitted in partial fulfilment of requirements for the degree

MASTER of SCIENCE in BIOLOGY

at the

CLEVELAND STATE UNIVERSITY

JULY, 2012

This thesis has been approved for
The Department of Biological, Geological and Environmental Sciences and for
The College of Graduate Studies of
Cleveland State University

by

_____ Date: _____

Bin Zhang, Ph.D., CCF & Adjunct faculty of BGES, CSU

Major Advisor

_____ Date: _____

Anton Komar, Ph.D., BGES, CSU

Advisory Committee Member

_____ Date: _____

Barsanjit Mazumder, Ph.D., BGES, CSU

Advisory Committee Member

DEDICATION

I would like to dedicate this thesis to those three special personalities who make my world. My parents (Mr Sudhangshu Das and Mrs. Rina Das) have been my constant source of inspiration. Their love and faith in me kept me going at trying times. Another person from whom I find constant solace and strength is my soul mate Mr. Shantanu Bhowal. His words of optimism and inspiration help me overcome all the hurdles of life.

ACKNOWLEDGEMENTS

None of us got where are solely by pulling ourselves up by our bootstraps. We got here because somebody – a parent, a teacher, a guide – bent down and helped us pick up our boots.
~ Thurgood Marshall

It is for Almighty who has held my hand and showered me with His blessings in this journey, without which nothing would have been possible.

It would be relatively easy for me to name all those individuals who have helped me to accomplish this, but I won't be able to thank them enough. In the last 3 years, I have met so many personalities who have contributed in their own way to help me get this done. It is with a deep sense of gratitude that I would thank my advisor Dr. Bin Zhang for giving me an opportunity to work in his lab under his close supervision, precious guidance, and patience.

Dr. Chunlei Zheng is another such person without whom this project would not have been possible. I would extend my sincere thanks to him for his close guidance and support at every step. I also appreciate the encouragement I received from my lab members and friends which kept me going.

I would never have been able to finish my thesis without the guidance of my committee members Dr. Anton Komar and Dr. Barsanjit Mazumder. I am thankful to Dr. Saurav Misra and Dr. Richard Page for their help and support in solving the crystal structure of ligand bound LMAN1.

STRUCTURAL BASIS OF LMAN1 CARGO CAPTURE IN ER & RELEASE IN ERGIC

VAIJAYANTI DAS

ABSTRACT

Proteins destined for secretion, after folding and the initial modification in the endoplasmic reticulum (ER), are transported to the ERGIC and then to Golgi stacks, before being sorted into their final destinations. Collectively these proteins that move through the secretory pathway are known as cargo proteins. Some of the soluble cargo proteins require cargo receptors to be efficiently packaged into vesicles exiting the ER. Evidence for the presence of such receptors in mammals came from the studies of a genetic disorder called the combined deficiency of factor V (FV) and factor VIII (FVIII) or F5F8D. F5F8D patients have mutations in either of the two genes – *LMAN1* or *MCFD2*. *LMAN1* and *MCFD2* form a receptor complex that facilitates the ER to Golgi transport of FV and FVIII.

Distinct binding sites for FV and FVIII exist on *LMAN1* and *MCFD2*, but the specific location of these binding sites are not known. Ambiguity exists on the mechanism of cargo release by *LMAN1* and *MCFD2*. Like FV and FVIII, other *LMAN1* cargos (cathepsin C, Z and α -1 antitrypsin) are also glycosylated. Thus the sugar binding activity of carbohydrate recognition domain (CRD) of *LMAN1* is presumably important for the recognition of the cargo glycoprotein.

Using biochemical approaches we determined the locations of sugar binding sites and their importance in cargo binding. We successfully purified the CRD of

LMAN1 using a denaturation/refolding protocol. Using isothermal titration calorimetry (ITC) assay, we confirmed that our purified CRD is a functional protein by its ability to bind to 2 calcium ions. ITC assay also provided insight on the effect of pH on binding ability of CRD to calcium as well as revealed the binding affinity of CRD towards its ligand α 1-2 manno-*bio*se. In collaboration with Dr. Saurav Misra's lab, we solved the structure of manno-*bio*se bound CRD which revealed key amino acids involved in ligand binding. Mutations of these amino acids in LMAN1 were constructed to reconfirm their importance in sugar binding *in vivo*. These biochemical studies will provide a useful insight into the mechanism of cargo binding in the ER and release in the ERGIC and Golgi.

TABLE OF CONTENTS

ABSTRACT.....	v
LIST OF FIGURES.....	xi

CHAPTERS

I. INTRODUCTION

1.1. The early secretory pathway.....	1
1.1.1. ER-GOLGI INTERMEDIATE COMPARTMENT (ERGIC).....	3
1.1.1.1. ERGIC is defined by LMAN1.....	3
1.1.2. Cargo proteins and cargo receptors- incorporating proteins to COPII vesicles.....	4
1.1.3. Trafficking of receptor dependent soluble and transmembrane cargo proteins.....	5
1.1.3.1. Cargo receptors in mammals.....	6
1.2. Combined deficiency of factor V and factor VIII.....	6
1.2.1. Phenotypic Changes associated with F5F8D	7
1.2.2. Changes in the molecular level in F5F8D patients	7
1.2.3. Hemostasis and Blood Coagulation Factors	8
1.2.3.1. Factor V.....	8
1.2.3.2. Factor VIII.....	9
1.2.4. Genetic Predisposition of F5F8D.....	10
1.2.4.1. Mutations in <i>LMAN1</i>	10
1.2.4.2. Mutations in <i>MCFD2</i>	11
1.3. Lectin Mannose-binding 1(LMAN1).....	12
1.3.1. Structure of LMAN1.....	13
1.3.1.1. Domains of LMAN1.....	14

i. Carbohydrate recognition domain (CRD).....	14
A. Similarity to leguminous lectins.....	14
B. Crystal structure of CRD of LMAN1.....	16
C. Comparison between rat LMAN1 and Isolectin 1.....	19
ii. LMAN1 Stalk.....	21
iii. Transmembrane domain (TMD) of LMAN1.....	22
iv. Cytoplasmic domain of LMAN1.....	22
1.3.2. LMAN1- A potential ER exit sorting receptor.....	23
1.3.3. Cargo recognition and binding by LMAN1.....	23
1.3.4. Other potential cargo receptors in mammals.....	26
1.4. Multiple coagulation factor deficiency 2(MCFD2): Partner of LMAN1 in cargo capture and trafficking.....	27

II. MATERIALS AND METHODS

2.1. Plasmid Construction.....	30
2.1.1. Designing the fusion protein.....	30
2.1.2. Designing the construct pET15b-CRD.....	32
2.1.2.1. Making the insert.....	32
2.1.2.2. Choosing the vector.....	35
2.1.2.3. Ligation	36
2.1.2.4. Cloning.....	37
2.2. Constructing mutant plasmids with pED vector.....	37
2.3. Protein Expression and Purification	38
2.3.1. Expression Host Transformation using the construct pET15bCRD....	38
2.3.2. Induction of Lysogen.....	38
2.3.3. Confirmation by SDS-PAGE.....	39
2.3.4. Preparation of inclusion bodies.....	39

2.3.5. Washing of inclusion bodies.....	40
2.3.6. Protein refolding	40
2.3.7. Protein purification.....	42
2.4. Expression and purification of MCFD2.....	43
2.5. Isothermal Titration Calorimetry.....	44
2.6. Crystallization of the CRD monomer with mannose α (1-2) mannose.....	46
2.7. Cell culture and transfection of COSI cells.....	46
2.8. Mannose binding assay.....	47
2.9. Western blot.....	48

III RESULTS AND DISCUSSION

3.1. Expression of the CRD of LMAN1 as a GST fusion protein and purification by GST column.....	50
3.1.1. GST-CRD Plasmid Construction.....	50
i. Amplification of GST from pSJ6 vector using Pfu.....	50
ii. Ligating amplified GST with vector containing wild type CRD.....	51
3.1.2. Purification of the fusion protein GST-CRD.....	52
3.2. The CRD of LMAN1 expression and purification by denaturation & refolding approach.....	54
3.2.1. Preparation of inclusion bodies and release of target protein pET15bCRD.....	56
3.2.2. Purification of refolded protein.....	58
3.3. Analysis of the interaction of the CRD with Ca^{2+} and the ligand at different conditions by isothermal titration calorimetry (ITC).....	61
3.3.1. Titration of the purified CRD with Ca^{2+} by the ITC assay.....	62
3.3.2. Titration of the purified CRD with 2α -mannobiose by ITC assay.....	64
3.3.3. Titration of the purified CRD with 2α -mannobiose at different saturation level of the CRD with Ca^{2+}	66
3.3.4. Effect of pH on the Ca^{2+} binding affinity of the CRD.....	69

3.4. Crystal Structure of the CRD of LMAN1 bound to mannose α (1-2) mannose.....	71
3.4.1. Comparison of the different crystal structures of the CRD of LMAN.....	74
3.4.1.1. Comparison of the CRD of LMAN1 with CRD of VIP36.....	75
3.5. Mannose Binding Assay.....	77
3.6. Purification of MCFD2.....	78
IV. CONCLUSIONS AND FUTURE DIRECTIONS	
4.1. Conclusions.....	81
4.2. Future Directions.....	84
BIBLIOGRAPHY.....	88
APPENDICES.....	93

LIST OF FIGURES

1. Model representing the secretory, endocytic, lysosomal and vacuolar pathways in mammalian cells.....	3
2. Genotype phenotype correlation in F5F8D	7
3. Schematic Representation of Factor V and Factor VIII structures and their similarity.....	8
4. Mutations found in LMAN1 gene.....	11
5. Mutations found in MCFD2 gene.....	12
6. Different Domains of LMAN1.....	13
7. Detailed representation of the Different domains of human LMAN1 showing signal peptide, CRD, helical domain and transmembrane domain.....	14
8. Shared sequences of two typical lectins	16
9. Sequence alignment of rat LMAN1 and VIP36.....	17
10. Crystal Structure of the CRD of Rat LMAN1.....	19
11. Comparison of the ligand binding sites of Isolectin 1 and p58/LMAN1.....	20
12. The hexameric form of LMAN1 bound to MCFD2.....	22
13. Biosynthesis of N linked core glycan and recognition of potential cargo by LMAN1.....	24
14. Cargo receptor function and recycling of LMAN1.....	26
15. MCFD2 domain structure showing the EF hand domains.....	29
16. pET15b vector sequence.....	36
17. Schematic Representation of the various solubilization and refolding strategies..	41
18. Amplified GST.....	51
19. Ligated colonies following Nco1 digestion.....	51
20. Coomassie Staining of SDS-PAGE of GST column purified wild type fusion GST-CRD fusion protein.....	52

21. Amplification and ligation of the CRD of LMAN1 into pET15b vector	55
22. Coomassie Staining of SDS-PAGE of pET15bCRD fractions.....	55
23. Chromatogram from Superdex 75 column used for purification of the CRD of LMAN1.....	59
24. Chromatogram of Protein Standard.....	60
25. Coomassie Staining of SDS-PAGE of sample fractions collected from gel filtration chromatography of CRD of LMAN1.....	61
26. ITC result from titrating the CRD of LMAN1 with calcium.....	63
27. ITC result from titrating the CRD of LMAN1 with 2 α mannobiose.....	64
28. Chemical Structure of Man9(GlcNAc) ₂	66
29. Titration of the CRD of LMAN1 with 30mM mannobiose at 0.4mM Ca ²⁺ saturation level.....	67
30. Titration of the CRD of LMAN1 with 30mM mannobiose at 2mM Ca ²⁺ saturation level.....	67
31. ITC result from titrating the CRD of LMAN1 with Ca ²⁺ at pH 7.4.....	69
32. ITC result from titrating CRD of LMAN1 with Calcium at pH 6.5.....	70
33. ITC result from titrating CRD of LMAN1 with Calcium at pH 6.0.....	70
34. Crystal Structure of the CRD of LMAN1 bound to ligand Mannose α (1-2) mannose and the key amino acids showing involved in interacting with the ligand.....	73
35. Overlay of different ribbon models of the CRD of LMAN1.....	74
36. Comparing the ribbon model of CRD of LMAN1 and CRD of VIP36.....	76
37. Western Blot from Mannose Binding Assay.....	77
38. Chromatogram from purification of MCFD2.....	79
39. Coomassie staining of SDS-PAGE of sample fractions collected during Superdex 75 purification of MCFD2.....	80

CHAPTER I

INTRODUCTION

1.1 The early secretory pathway

Up to one third of cellular proteins traverse through the secretory pathway to reach their final destination, including the extracellular space, the plasma membrane and various internal organelles. The secretory pathway is a defined route consisting of a series of membrane compartments. The newly synthesized proteins destined to be transported by the secretory pathway are translated from the poly-ribosome and sorted co-translationally into the lumen of the endoplasmic reticulum (ER), which constitutes the first component of the secretory pathway. In the ER, the nascent proteins are folded, oligomerized or conjugated to glycan moieties as part of their post translational modifications. Only the properly folded proteins exits the ER at specific ER exit sites and traverse through the secretory pathway. At these ER exit sites (also called transitional ER), budding of the COPII (coat protein complex II) coated vesicles take place. Vesicles are small fluid filled sac like structures that are involved in cross talk between different compartments of the cell. COPII coated vesicles carry soluble and membrane proteins for anterograde transport (Kondylis V et al, 2009).

The COPII coat consists of a small GTPase Sar1 along with Sec 23/24 and Sec13/31 protein complexes (Baines A & Zhang B, 2007; Hughes H et al, 2008). The vesicle formation in vivo requires the presence of two additional proteins Sec12p and Sec 16p (Futai E et al, 2004; Supek F et al, 2002). Sec 12 is a guanine nucleotide exchange factor (GEF) for Sar1. It is because of the Sec12p that COPII assembly takes place only on the surface of the ER (Soderholm J et al, 2004) and it is Sec16p that restricts the COPII vesicle formation to the ERES (Watson P et al, 2006 ; Bhattacharya D & Glick B.S, 2007).

Proteins that travel through the secretory pathway are conjointly called as “*cargo proteins*” (Baines A & Zhang B, 2007). The cargo proteins, carried by COPII vesicles eventually reach the stacks of Golgi from the ER where further processing of the protein takes place (Kondylis V et al, 2009). The *early secretory pathway* (Figure 1) refers to the transport of proteins from the ER to the Golgi (Tekaya H-B, 2004). But before the proteins are delivered to the Golgi apparatus, the COPII vesicles fuse to ER-Golgi intermediate compartment called the ERGIC (Kondylis V et al, 2009; Baines A & Zhang B, 2007). The supplementary transport factors and proteins that carry ER retrieval signals, also known as ER resident proteins, are recycled back to the ER (known as retrograde transport) via interaction with COPI vesicles (Lippincott-Schwartz J & Liu W, 2006). This is necessary so that they can bring about successive series of vesicle formation and thereby anterograde protein transport.

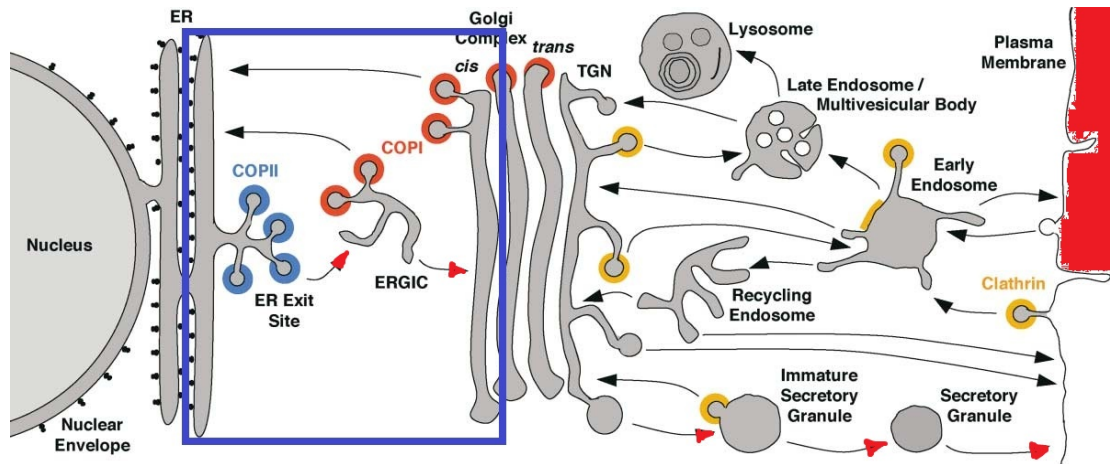


Figure 1: Model representing the secretory, endocytic, lysosomal and vacuolar pathways in mammalian cells.
 The early secretory pathway is highlighted by the purple box. Traffic through the early secretory pathway is same for all newly synthesized proteins *enroute* to their final intra or extracellular destination. COP-coat protomer complex; ER- endoplasmic reticulum; ERGIC: ER-Golgi intermediate, TGN: Trans Golgi network (Modified from Bonifacino J.S et al 2004)

1.1.1 ER-Golgi intermediate compartment (ERGIC)

ERGIC is a tubulovesicular structure between the ER and the Golgi and it was earlier proposed to be the specialized domain of the ER or the cis-Golgi. Subsequent studies revealed that it is not continuous with the cis-Golgi or the ER and has evolved in parallel with the increasing complexity of the secretory pathway in higher eukaryotic organism. It is the site of segregation of the secretory proteins for anterograde transport, via packaging into COP II coated transport vesicles. Here the proteins are subjected to further processing and quality control (Hauri H.P et al, 2000; Stephens D.J & Pepperkok R, 2001).

1.1.1.1 ERGIC is defined by LMAN1 (also known as ERGIC-53)

LMAN1 was first identified as a novel protein in the ERGIC that was recognized by the monoclonal antibody G1/93 raised against purified Golgi membrane of Caco-2 cell line. Single and double immunoelectron microscopy

detected that this monoclonal antibody localized and labeled the tubulovesicular compartment lying close to *cis*-Golgi. The protein identified by G1/93 has a molecular weight of 53kD. Experiments showed that this protein, dimerized right after its synthesis, followed by the formation of hexamers (Schweizer A et al, 1988). Because of its localization it was named as ERGIC-53. Research over the years has established that LMAN1 acts as cargo receptors for FV/FVIII and other glycoproteins but the molecular mechanism of cargo capture and release still remains to be determined.

1.1.2 Cargo proteins and cargo receptors – incorporating proteins to COPII vesicles.

Proteins that lack sorting signals are transported out of the ER or other cellular compartments by a nonspecific method called the “bulk flow. This phenomenon of bulk flow is observed mostly in cases when proteins are abundantly expressed. However the less abundant cargo proteins are sorted and concentrated before being recruited to the carrier vesicles in a selective manner (Martinez-Meranguez G et al, 1999; Oprins A et al, 2001). Studies indicate that cargo proteins in the ER that participate in this selective fashion of transport exhibit specific signals which help to recruit them into the COPII vesicles. These cargo proteins are classified into two types – (i) proteins that are able to bind to COPII coat directly and (ii) proteins that require special adapters to link them to COPII vesicles. The first group includes the transmembrane proteins that extend from one side of the ER membrane to the other side to easily reach the COPII coat proteins. These cargo proteins can act as “*cargo receptors*” for the second group which requires special receptors to link them to COPII vesicles (Baines A & Zhang B, 2007). Transport of cargo proteins by LMAN1 falls in the second category where the cargos are unable to bind to vesicles directly

and requires LMAN1 to link them to the COPII coated vesicles. Multiple paralogs of COPII proteins in mammals introduce higher degree of diversity in transport vesicles, including tissue specific and cell specific vesicles. Hence they are able to channelize a variety of cargo proteins to selective transport (Baines A & Zhang B, 2007)

The main difference between the two models of protein transport is that in the bulk flow model, properly folded proteins are nonspecifically packaged into COPII vesicles but the incompletely folded proteins are selectively retained in the ER. In the receptor mediated transport model, correctly folded cargo proteins are selectively recognized by their receptors and actively recruited to the site of vesicle budding (Thor F et al, 2009).

1.1.3 Trafficking of receptor dependent soluble and transmembrane cargo proteins

Soluble cargo proteins cannot interact directly with COPII coat protein. In the receptor-mediated transport model, trafficking of these cargo proteins are facilitated by specific transmembrane adapters or receptors. These receptors help packaging of cargo proteins into the COPII vesicles (Baines A & Zhang B, 2007).

First such evidence of receptor mediated transport of soluble cargo proteins was observed in the yeast *Saccharomyces cerevisiae*. In this organism, a conserved transmembrane protein Erv29p was essential for the packaging and transport of glycosylated pro- α -factor (gp α f). When Erv29p was deleted, it led to defective ER-Golgi trafficking of gp α f (Belden W.J & Barlowe C, 2001). Further studies revealed the first known export signal of gp α f was recognized by Erv29p (Otte S & Barlowe C, 2004). Two other cargo receptors have also been identified in yeast – Erv26p and Erv14p. Defects in these receptors led to aberrant trafficking of corresponding soluble

cargo. Research also showed that Emp46p and Emp47p are two homologous proteins in yeast that contain carbohydrate recognition domains and metal ion binding sites. They form a heteromeric complex in the ER and this complex is essential for the effective transport of Emp46p (Baines A & Zhang B, 2007). Emp47p is thus considered to be the cargo receptor for Emp46p.

1.1.3.1 Cargo receptors in mammals

So far the best studied example of cargo receptor in mammals is the LMAN1-MCFD2 protein complex. This came into light from the finding that mutations in either *LMAN1* or *MCFD2* genes are responsible for a rare, hereditary blood coagulation disorder called combined deficiency of factor V and factor VIII (F5F8D), suggesting the potential involvement of LMAN1 and MCFD2 in selective cargo transport of FV and FVIII from the ER (Zhang B, 2009). The LMAN1-MCFD2 receptor complex is believed to be the only known example to date for the trafficking of soluble cargo proteins. The involvement of a transmembrane component along with a soluble cofactor determines the uniqueness of this secretory pathway.

1.2 Combined deficiency of factor V and factor VIII (F5F8D)

F5F8D is an autosomal recessive blood coagulation disorder first reported in 1954 by Oeri J et al. Cases of F5F8D has been reported throughout the world. Mostly associated with consanguineous marriage, F5F8D is a relatively rare disease affecting 1:1,000,000 individuals in the general population (Spreafico M & Peyvandi F, 2008). However certain populations such as the Middle Eastern Jews or non-Jewish Iranians are at higher risk of developing this disease (Zhang B, 2009)

1.2.1 Phenotypic changes associated with F5F8D

F5F8D is characterized most commonly by mild to moderate bleeding symptoms beginning in childhood. Bleeding symptoms include easy bruising, epistaxis, menorrhagia and excessive bleeding following surgery, labor or tooth extraction. Because of the mild bleeding symptoms, F5F8D is often underdiagnosed. Excessive bleeding symptoms can be controlled by the administration of fresh frozen plasma and recombinant FVIII (Zhang B & Ginsburg D, 2006; Spreafico M & Peyvandi F, 2008).

1.2.2 Changes in the molecular level in F5F8D patients

F5F8D patients are diagnosed with prolonged prothrombin time (PT) and partial thromboplastin time (PTT) [Zhang B, 2009]. In F5F8D, there is a synchronized reduction of plasma factor V (FV) and factor VIII (FVIII) levels. Reduction in platelet FV level has also been observed. Usually a dip of 5-30% of FV and FVIII activities takes place in F5F8D patients but in some cases it goes as low as less than 1% or as high as 50% (Figure 2).

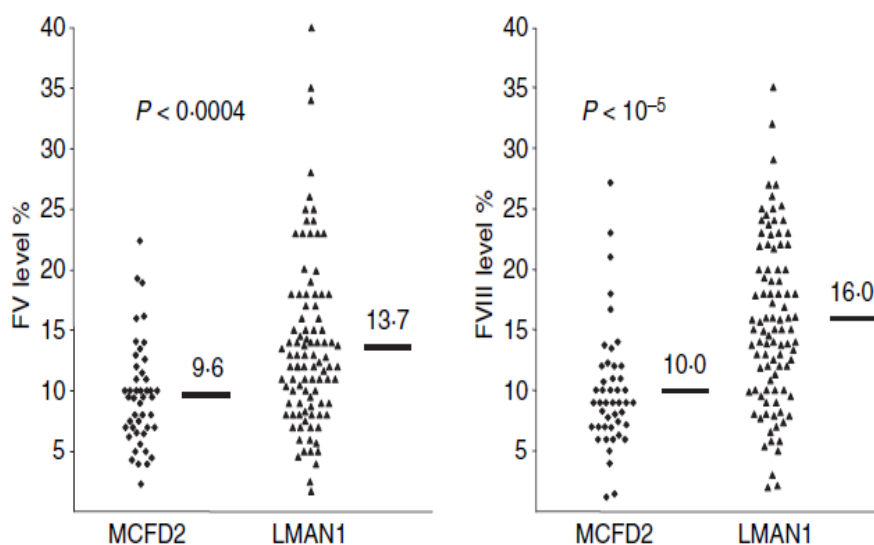


Figure 2. Genotype Phenotype Correlation in F5F8D: Correlation of LMAN1 and MCFD2 with FV/FVIII levels. Short bars and numbers indicate average values. FV and FVIII levels are lower in F5F8D patients with MCFD2 mutations compared to patients with LMAN1 mutations. (Represented from Zhang B et al, 2008)

1.2.3 Hemostasis and blood coagulation factors

Hemostasis is a well-coordinated process to stop bleeding, regulated by the sequential activation of serine proteases in the coagulation cascade. This cascade of blood clotting factors ultimately leads to the fibrin formation. FV and FVIII are essential cofactors for the proteases factor X and factor IX respectively.

1.2.3.1 Factor V

Factor V, a 330 kDa polypeptide is synthesized primarily in hepatocytes and megakaryocytes and it circulates at a concentration of 20nM. The gene for FV is located on chromosome 1 at q21-25. Also known as the labile factor, FV is the inactive precursor of factor Va. FV (pro-cofactor) undergoes several post translational modifications – sulfation, phosphorylation and glycosylation. It is activated by proteases such as thrombin, factor Xa and plasmin. Thrombin cleavage is most critical in terms of the biological function of the molecule. It cleaves the pro-cofactor to generate the non-covalently linked light and heavy chain of FVa. Sulfation is required for efficient thrombin cleavage, activation and for full procoagulant activity.

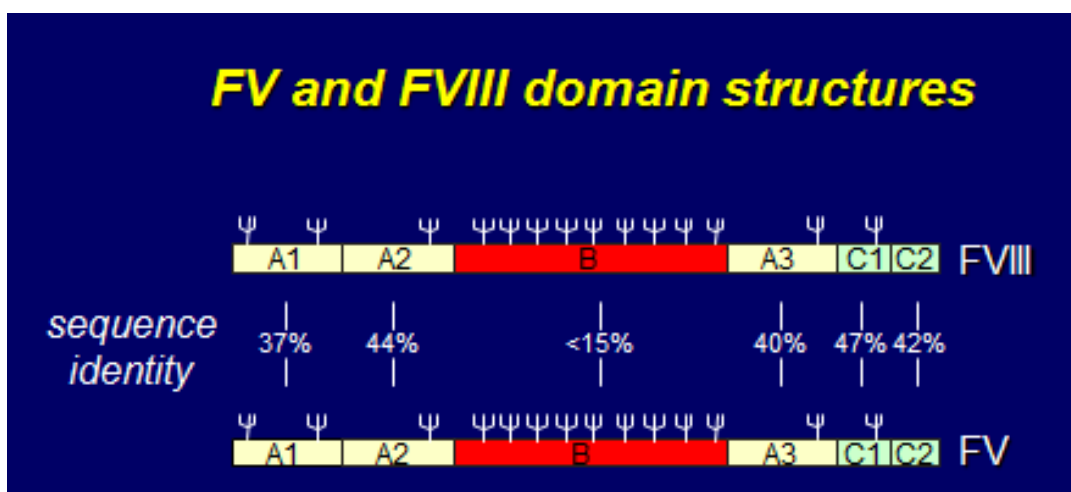


Figure 3. Schematic representation of factor V and factor VIII structures and their similarity.

Activated factor Va interacts with factor Xa to form the prothrombinase complex which in turn leads to the formation of thrombin and from thrombin, ultimately fibrin or the blood clot is formed. Deficiency of FV leads to an uncommon pathologic condition called the parahemophilia (Mann K.G & Kalafatis M, 2003).

1.2.3.2 Factor VIII

Also called the antihemophilic factor, FVIII is a ~ 280kDa molecule found in both hepatocytes and endothelial cells. The Factor VIII is structurally and functionally quite similar to FV. The gene for FVIII is located in X chromosome (Xq28). Factor VIII circulates as inactive procofactor but after limited proteolysis it gains full cofactor activity (Bos M & Camire R, 2010). The major role of factor VIII is the conversion of factor X to its activated form Xa by acting as a cofactor for factor IX in the presence of calcium ion and phospholipids. In blood, it mostly exists as a stable complex with von Willebrand factor. After cleavage by thrombin, it dissociates from the complex and takes part in the coagulation cascade. Normal plasma levels of factor VIII range from 50% to 150%.

Deficiency in factor VIII leads to hemophilia A, a hereditary bleeding disorder affecting about 1 in every 5000 males in United States. Hemophilia A represents about 80% of the total hemophilia cases. FVIII level below 1% is a bleeding disorder often associated with “severe bleeding episodes” that can lead to severe consequences if left untreated (Camire R, 2009).

FV and FVIII share similar domain structures – three sets of domain A, one B domain and two sets of C domain (Figure 3). Sequence identity between the A and C domains of these two coagulation factors is very high. B domains of FV and FVIII

share little sequence homology but are both heavily glycosylated with 9 potential glycosylation sites in the heavy chain (Mann K.G & Kalafatis M, 2003).

1.2.4 Genetic Predisposition of F5F8D

Genetic studies like homozygosity mapping and positional cloning on patients with F5F8D identified mutations in either of the 2 genes *LMAN1* (also known as ERGIC 53 or p53) or *MCFD2*. In 70% cases it is caused by mutations in *LMAN1* and in 30% cases it is caused by mutations in *MCFD2* (Zhang B et al, 2006). F5F8D caused by *LMAN1* mutation is clinically indistinguishable from the disease caused by mutation in *MCFD2*, though it has been generally noticed that patients with *MCFD2* mutation has lower FV and FVIII levels (Figure 2). However from the FV and FVIII levels, it cannot be predicted which kind of mutation a patient would have.

1.2.4.1 Mutations in *LMAN1*

So far 32 mutations have been identified in *LMAN1* gene, spanning its 13 exons that lead to F5F8D. The majority of mutations are null mutations resulting from normally frameshift or non-sense mutations that produce shortened protein differing in function from the normal *LMAN1*. One F5F8D patient showed no mutations in the exon or exon-intron junction but had mutation in the regulatory region of the gene (Zhang B et al, 2006; 2008). Two missense mutations in *LMAN1* are known – one substitutes methionine with threonine thereby abolishing protein translation and the second one changes arginine to cysteine (Nufer O et al, 2003). This conversion of arginine to cysteine leads to the production of an unstable protein (Zhang B, 2006). The mutations occurring in *LMAN1* are shown in Figure 4.

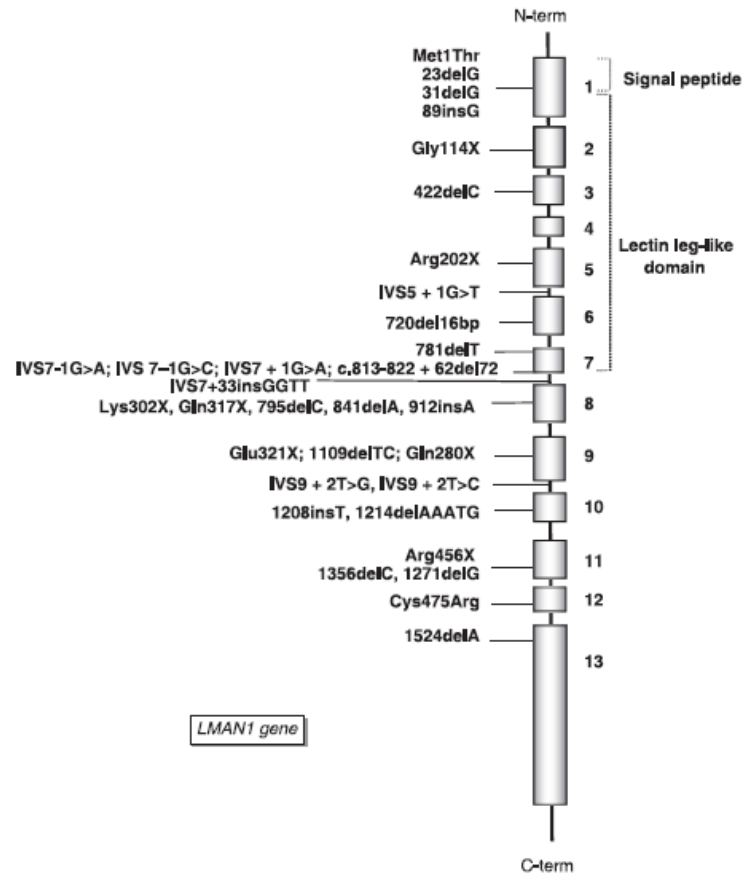


Figure 4. Mutations found in *LMN1* gene (Represented from Spreafico M et al, 2008)

1.2.4.2 Mutations in *MCFD2*

There were 17 mutations known to occur in *MCFD2* that led to F5F8D. Nine of these are deletion or splicing mutations that resulted in shortened protein products and seven were missense mutations which changed conserved amino acids (Zhang B 2009; Zhang B et al, 2011). The other kind of alteration was a non-sense mutation that deletes 8.5kb in the promoter region of the gene along with the last three amino acids of *MCFD2* thereby disturbing the protein structure (Nyfeler B et al, 2007). *MCFD2* mutations that cause F5F8D are shown in Figure 5.

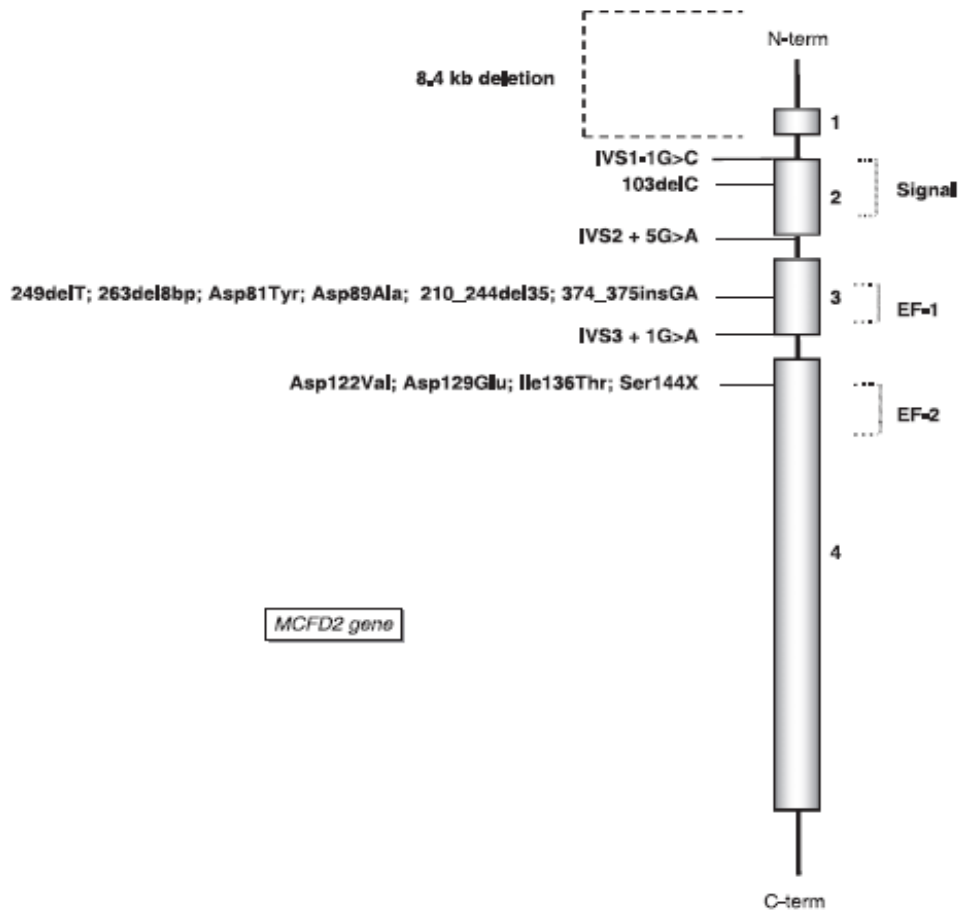


Figure 5. Mutation in *MCFD2* gene (Represented from Spreafico M et al, 2008)

1.3 Lectin Mannose-binding 1 (LMAN1)

LMAN1 is a nonglycosylated, type I single spanning transmembrane protein, of molecular weight 53kD. It cycles between the ER and the ERGIC/cis-Golgi. LMAN1 serves as the indicator protein to define organelle boundaries and has been established as a marker for the ER-Golgi intermediate compartment (ERGIC). LMAN1 is conserved across species, from vertebrates such as human and mouse to invertebrates such as *Xenopus laevis* and *Saccharomyces cerevisiae* (Hauri H-P et al, 2000).

Human *LMAN1* gene is localized in the long arm of chromosome 18 (18q21.3-18q22). The mRNA coding for this protein is 2768 bp with 13 exons, of varying lengths ranging from 59bp (exon 7) to 1244bp (exon13). Exon 13 is the longest one but it codes for only amino acids 499-510 and the remaining more than 1200bp lies in the 3' UTR (Nichols W.C et al, 1999).

LMAN1 is ubiquitously expressed in all tissue types but generally has higher expression in secretory cell types (Zhang B et al, 2011). Within the cell, LMAN1 either exists as six disulfide linked monomers or three disulfide linked dimers joined noncovalently forming a hexamer (Neve E.P et al, 2005). Endogenous LMAN1 exist in a Ca^{2+} dependent complex with MCFD2 in a 1:1 stoichiometry. However, MCFD2 is not necessary for oligomerization of LMAN1 (Zhang B et al, 2005).

1.3.1 Structure of LMAN1

Human LMAN1 consists of about 510 amino acids, which includes an N terminal signal sequence of 30 amino acids residues, a carbohydrate recognition domain (CRD) of about 240 residues a stalk region of 210 amino acids, a transmembrane domain of 18 residues and a short cytoplasmic domain of 12 residues (Figure 6 & Figure 7). The stalk region of LMAN1 consists of a series of four α helices that are believed to form a coiled coil domain with two conserved cysteine residues, Cys 466 and Cys 475.



Figure 6. Different domains of LMAN1 (Zhang B, 2009)

Each domain of LMAN1 has their respective functions which are quite diverse from one another and is necessary for the proper functioning of the protein molecule (Figure 7).

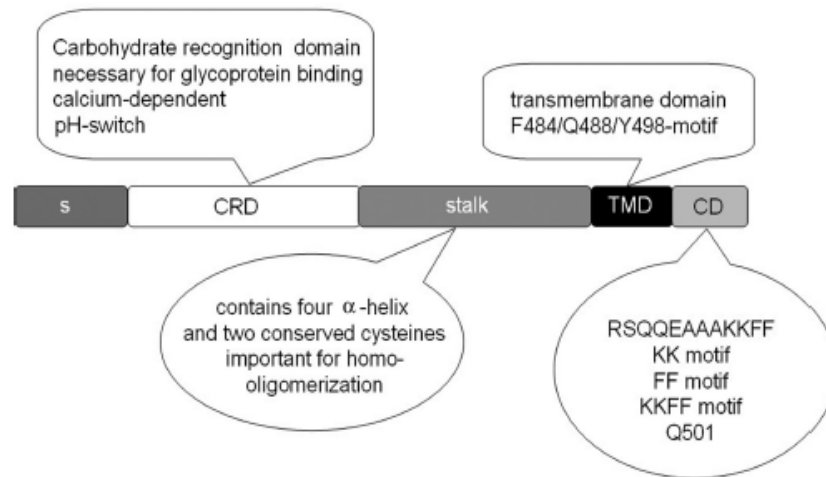


Figure 7. Detailed representation of the different domains of human LMAN1: S-Signal peptide, CRD-carbohydrate recognition domain, TMD-transmembrane domain and CD-cytoplasmic domain (Zhang Y.C et al, 2009)

1.3.1.1 Domains of LMAN1

The different domains of LMAN1 are as follows:

i) Carbohydrate Recognition Domain (CRD) of LMAN1

As the name indicates, it is that part of LMAN1 which interacts with sugar molecules of the cargo proteins. It is the ~200 residue segment in the luminal domain and bears homology with the CRD of various leguminous plant lectins.

A. Similarity to leguminous lectins

Studies revealed sequence homologies of LMAN1 (19-24%) with leguminous lectins (Itin C et al, 1996). The core structure shared the same basic architecture as the leguminous lectins. Leguminous lectins (Latin: *lectus* means “to gather” or “select”)

are a class of proteins that bind to sugar (or carbohydrate) binding proteins without modifying them, unlike glycoproteins which itself contain covalently linked sugars. Immense studies at the molecular and structural level on this class of lectin revealed that they were synthesized within the secretory organelles of the plant and mostly homo-oligomerized to form dimers or tetramers. These plant lectins were capable of agglutinating cells and bound to a variety of glycoproteins having varied carbohydrate structure. Concanavalin A was the first plant lectin to be crystallized and was found to bind and precipitate polysaccharides like glycogen and starch (Cummings R.D et al, 1999).

Usually leguminous lectins were metalloproteins that bind to metals ions Ca^{2+} , Mg^{2+} or Mn^{2+} . It was observed that if these metal ions in leguminous lectins were chelated with reagents like EDTA, lectins failed to bind to the glycoconjugates. These lectins had common consensus sequences which helped in their easy identification (Figure 8). They had antiparallel beta pleated sheets and a carbohydrate binding pocket consisting of two cavities that worked in coordination with one another. In one cavity lied the metal ions and the other had the monosaccharide. In between these two cavities, were conserved residues – an aspartate (Asp/D) and an Asparagine (Asn/N). These residues interacted with the metal ions and also were hydrogen bonded to the monosaccharide. The metal ions were believed to serve a pivotal role in orienting the conserved residues as well as maintain their proper distance from the monosaccharide (Itin C et al, 1996).

LMAN1 exhibited all these properties of a leguminous lectin hence it was considered to be a putative new class of animal L-type lectins associated with the secretory pathway. Like leguminous lectins, LMAN1 is Ca^{2+} dependent and mutation in the conserved residues of asparagine and aspartate in LMAN1 abolished its ability

to bind to mannose column. Hence, it was predicted that these two residues might be important for the lectin activity of LMAN1. This was also based on the fact that LMAN1 bore similarity to VIP-36 which is a membrane protein and was found to bind to mannose *in vivo* (Itin C et al, 1996).

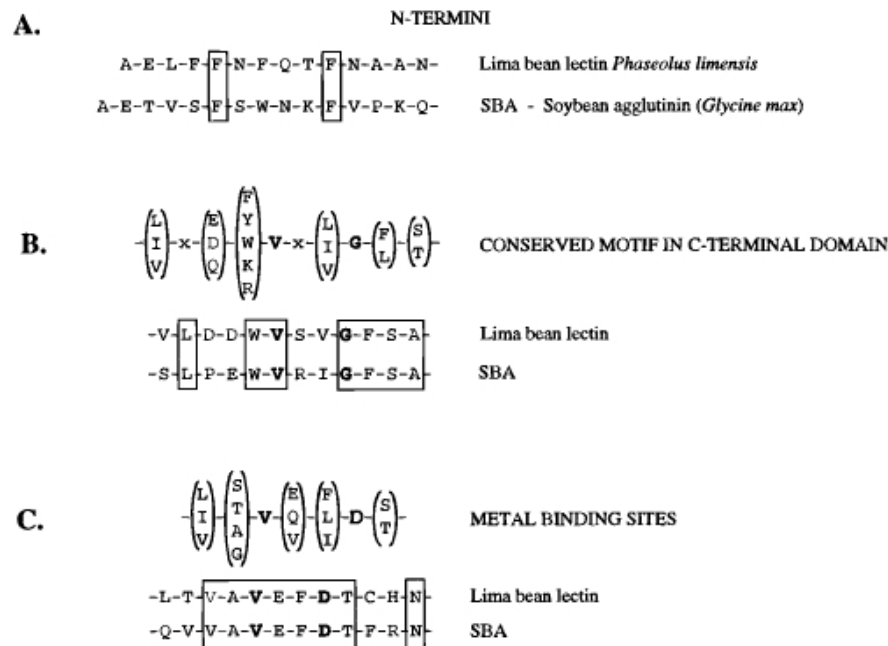


Figure 8. Shared sequences of two typical lectins – lima bean lectin and soybean agglutinin (Cummings R et al ,1999).

B. Crystal structure of the CRD of LMAN1

Crystal structure of the carbohydrate recognition domain of rat p58 which is the homologue for human LMAN1 was first solved by Velloso et al in 2002. They observed that the CRD was bulbous in shape, had one concave and one convex β -sheet packed into a β sandwich and the folds and ligand binding sites were similar to leguminous lectins. Though the oligomeric form of LMAN1/p58 binds to mannose monosaccharide but mannose was not the correct ligand for LMAN1. It was believed to bind to Man₉ of glycoproteins (Velloso L.M et al, 2002).

From the sequence homology (Figure 9), it was seen that conserved asparagine and aspartate residues in p58 - N164 and D129 corresponds to N156 and D121 of human LMAN1 and these might be crucial for binding to mannose. The other conserved residues of LMAN1 are Trp51, Trp78, Trp110, Trp128, glycines in loop between β strands, two disulfide bonded cysteines (Cys 466 and Cys 475) and proline (Velloso L.M et al, 2002). However the crystal structure of the CRD that was solved first was without the calcium ion and it was known that the proper functioning of LMAN1 requires the presence of this metal ion. Hence, Velloso later on solved the crystal structure of the CRD in presence of Ca^{2+} (Figure 10).

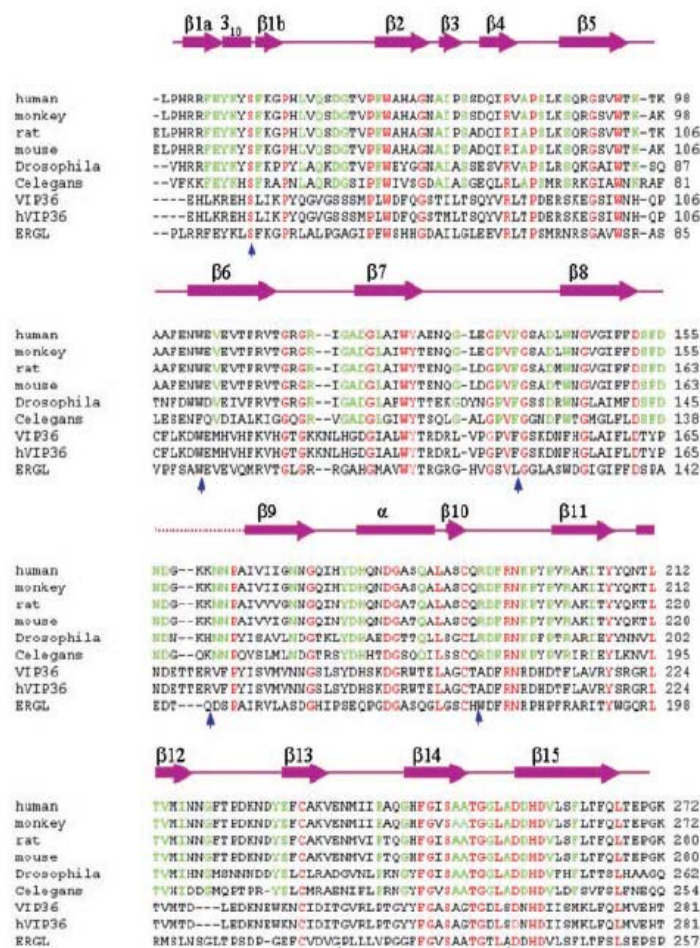


Figure 9. Sequence alignment of rat LMAN1(p58) and VIP-36 like CRD. Secondary structure elements are shown as arrows and rectangles, representing β strands, dashed line are the disordered regions. Blue arrowheads are the splice sites (Velloso L.M et al, 2002).

Leguminous lectins that bound to two calcium ions contained two metal ion binding sites –S1 and S2. In studies that were carried out on Concanavalin A, it was observed that initially the metal ion binds to site S1 and then to the second site S2. This binding of metal ions caused large structural changes in the CRD resulting in an alteration from the unlocked to the locked form of the protein. This step was considered crucial for carbohydrate binding (Figure 10). The crystal structure of the CRD of p58 behaved likewise. (Velloso L.M et al, 2003).

When the crystal of the CRD of p58 was grown in the presence of Ca^{2+} , it revealed well defined electron density compared to the calcium free form where the residues were disordered. Large conformational changes were observed between the calcium bound crystals from the calcium free form. These changes primarily involved the two regions of the protein that contain all the calcium ion coordinating residues. These two major sites were – (i) loop between residues 161 and 172 (referred as loop1) and (ii) loop between residues 179 and 193 (referred as loop 2); most prominent change was observed in the residue N164 (Figure 10). The crystal structure of the CRD in presence of Ca^{2+} ions showed the presence of a site similar to S2 of leguminous lectins and it was called M2. Despite the similarity between LMAN1 and the leguminous lectins, electron density map showed the absence of the S1 site in LMAN1. Mutagenesis studies at the histidine and glutamate residues present in the side chain of S1 also confirmed the absence. But it was observed that there is another site, referred as M1, located close to the loop between 161 and 172 which represented a metal binding site (Velloso L.M et al, 2003).

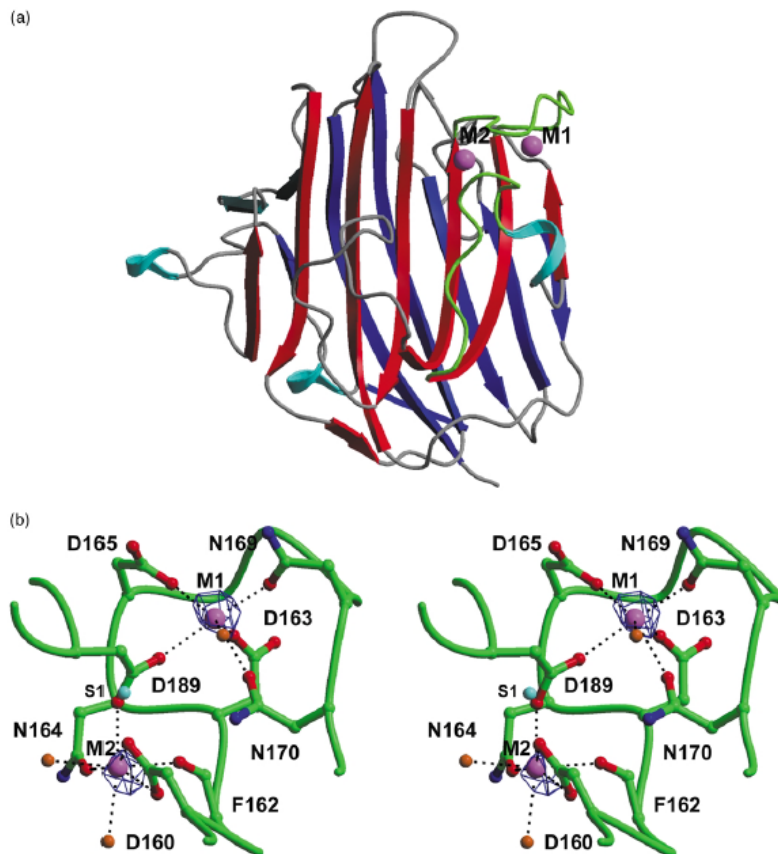


Figure 10. (a)Crystal Structure of CRD of Rat LMAN1 showing the location of bound Calcium ions (pink spheres). (b) Loops changing conformation upon binding to metal ions (shown in green) (Velloso L.M et al, 2003).

C. Comparison between rat p58/LMAN1 and Isolectin 1

Isolectin 1 is a leguminous lectin that was best illustrated and most similar to LMAN1 structurally. When LMAN1 was superposed on Isolectin 1, it was observed that the site M2 corresponded to the site S2 of leguminous lectin and they were only 1 Å apart. But the site M1 did not correspond to the site S1 as their positions were 6.4 Å apart. Residues Asp160, Asn 164, Phe162 and two water molecules that acted as calcium ion ligands were identical in p58/LMAN1, Concanavalin A and Isolectin 1 (Velloso L.M et al, 2003).

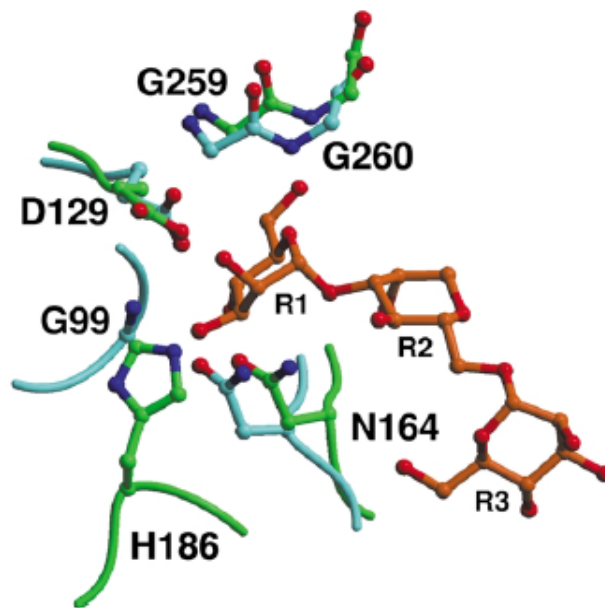


Figure 11. Comparison of the ligand binding sites of Isolectin 1 and p58/ rat LMAN1. Residues hydrogen bonding to the Man (α 1-3)4 residue of Isolectin-1 are shown in light blue and their counterparts in p58 are shown in green (Represented from Velloso L.M et al, 2003)

The ligand binding sites of leguminous lectins were highly conserved across different proteins. A model (Figure 11) was proposed comparing the ligand binding sites of octasaccharide complex of Isolectin 1 and p58/LMAN1. The three different ligands used in the modeling study were Man(α 1-3)4, Man(β 1-3)3 and Man(α 1-6)4. According to the model, the ligand binding sites for one ligand Man(α 1-3)4 were remarkably similar between Isolectin 1 and p58/LMAN1. This was the only mannose that hydrogen bonded to Isolectin 1. Except Gly99, all other hydrogen binding residues for this ligand had their counterparts in LMAN1 and were identical in nature and in similar position. Gly99 of Isolectin 1 was the only residue that did not have a counterpart in LMAN1 but was present in Isolectin 1 and Concanavalin A. Absence of this residue in LMAN1 rendered it with a more open ligand binding site compared to leguminous lectins (Velloso L.M et al, 2003).

The main difference between LMAN1 and Isolectin 1 was their localization. LMAN1 is mostly restricted to ER and ERGIC but Isolectin 1 is present in the

extracellular space. Because of this difference in localization, the ligands for these two kinds of lectins were also different. Isolectin 1 and other leguminous lectins recognize complex sugar structures that were extensively modified in the stacks of Golgi but LMAN1 recognizes $\text{Man}_8(\text{GlcNAc})_2$ as ligand. LMAN1 ligand differed from the complex sugar structures recognized by leguminous lectins in the number of branches as well the terminal sugar residues. When protein exit ER and enter the stacks of Golgi for further modification, the sugars are all mannose and carries three branches but when the protein exits the Golgi where it has undergone further modification, proteins carry only two branches with N-acetylglucosamine, galactose or sialic acid at the termini. Probably because LMAN1 has more exposed ligand binding site, it is able to connect with more branched structures. This may be a plausible reason for the M1 site in LMAN1 to be positioned differently from the S1 site in leguminous lectins (Velloso L.M et al, 2003).

ii) LMAN1 stalk

As shown in Figure 7, the stalk consisted of a series of four α helices expected to form a coiled coil and two conserved cysteine residues of LMAN1, Cys 466 and Cys 475 located close to the transmembrane domain. These conserved residues were initially believed to be critical for dimerization and hexamerization of LMAN1 (Figure 12). But further research revealed that the conserved cysteines might not be crucial for oligomerization of the protein. It was stated that the helical part of LMAN1 was essential for hexamerization of the protein and LMAN1 existed as SDS sensitive and SDS resistant hexamers (Neve E et al, 2005). However, the role of cysteine residues is still unclear.

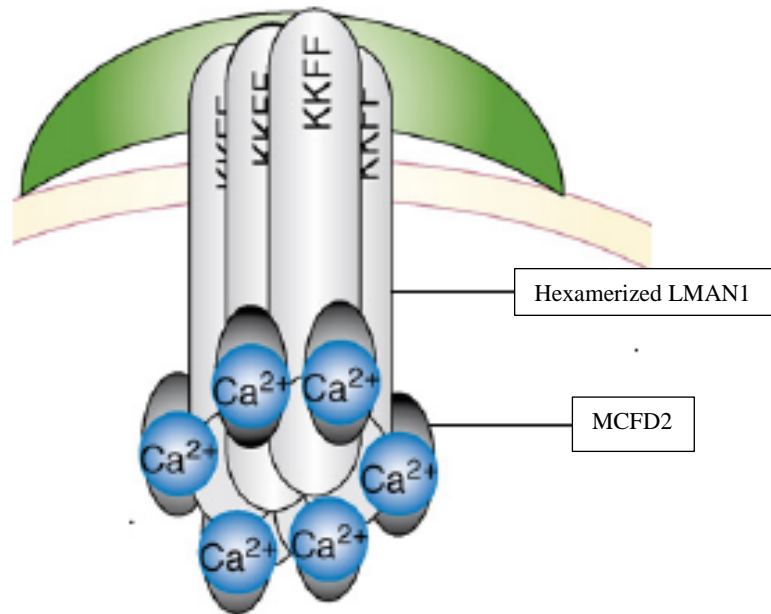


Figure 12. The hexameric form (gray cylinders) of LMAN1 bound to MCFD2 (Baines A & Zhang B, 2007)

iii) Transmembrane domain (TMD) of LMAN1

TMD of LMAN1 consists of 18 amino acids. This length of TMD affects ER retention. Experiments showed that an optimum of 21 amino acids is necessary for efficient ER exit but the length of TMD of LMAN1 is sub optimal. However, TMD contains a motif of three residues F₄₈₄/Q₄₈₈/Y₄₉₈. Q₄₈₈ is unique for human LMAN1. This motif may also contribute to the hexamerization of LMAN1 (Nufer O et al. 2003).

iv) Cytoplasmic domain of LMAN1

The cytoplasmic domain of LMAN1 with amino acid sequence RSQQEAAKKFF plays the most important role in the correct targeting of the protein in the ER-ERGIC-Golgi recycling pathway. The last four amino acid residues of the above mentioned sequence acts as determinants for this function. The double lysine (KK) residue act as ER retention and retrieval signal via interaction with COPI

and thereby leading to efficient endocytosis of LMAN1 from the cell surface when it is overexpressed (Schindler R et al, 1993). The double phenylalanine (FF) residue is important for ER exit mechanism via interaction with COPII. As a result, it considerably weakens the ER retention signal and at a steady state there is a significant level of the LMAN1 localized outside ER (Zhang Y.C et al, 2009).

1.3.2 LMAN1 – A potential ER exit sorting receptor

Any sorting receptor has to be a transmembrane protein that has a cytoplasmic domain interacting with COPI and COPII proteins and an ER luminal domain that interacts with cargo. LMAN1 structure satisfies all requirements for a sorting receptor with its CRD, cytoplasmic ER exit and retrieval signals and transmembrane domain (Itin C et al, 1996). LMAN1 has a long half-life of several days. This may contribute to its function of a transport receptor for glycoproteins (Zhang Y. C, 2009).

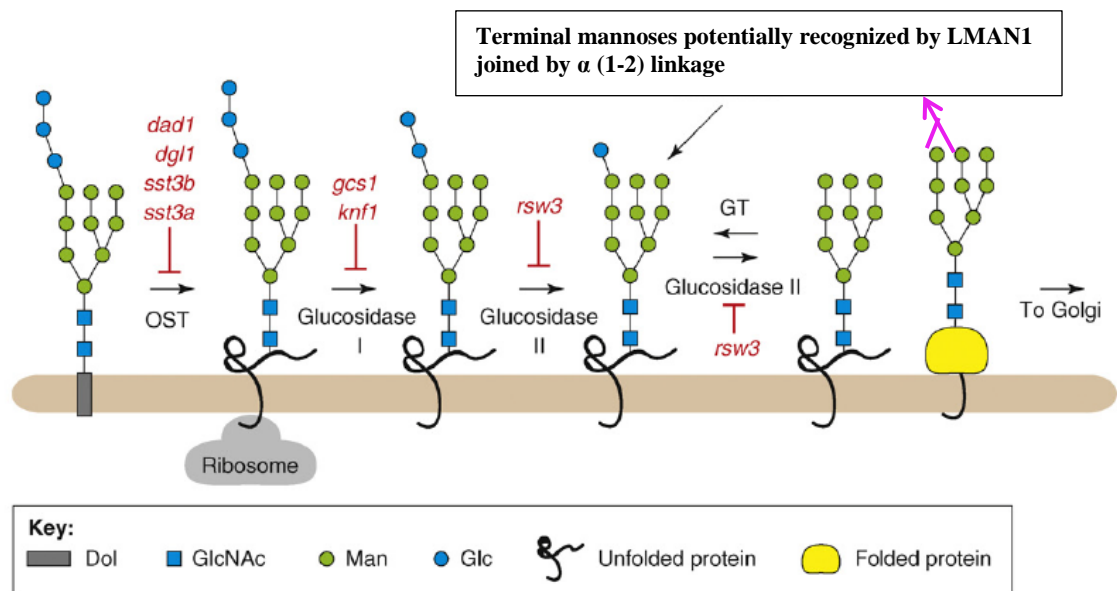
1.3.3 Cargo recognition and binding by LMAN1

LMAN1 with MCFD2 form a receptor complex and helps in the transport and synthesis of blood coagulation factor V and factor VIII. Other potential cargos for LMAN1 are the lysosomal proteins cathepsin C, cathepsin Z and α -antitrypsin (Baines A & Zhang B, 2007; Nyfeler B et al, 2008).

The cargos for LMAN1 –FV, FVIII, Cat C, Cat Z or α -antitrypsin, are all glycoproteins. They undergo trimming and processing at the early secretory pathway in order to be transported by LMAN1. This trimming and processing that takes place makes all proteins structurally and functionally diverse. It is the heterogeneity of the glycan moieties in the glycoproteins that renders them this diversity. They differ in the sugar composition, number and size of branches in the sugar tree and this

distinction is observed across cell types, tissues and species. However, the glycan moiety that conjugates to the nascent polypeptide right after its synthesis in the ER is homogenous and relatively simple. This glycan moiety is called “core glycan” and this does not make the glycoproteins much different from one another (Figure 13) (Helenius A et al, 2001).

When these glycoproteins reach the medial stacks of Golgi, they undergo extensive alteration by non-uniform trimming and processing. It is the terminal glycosylation that takes place in the vertebrate and plant Golgi complex which introduces significant heterogeneity in these glycoproteins. As a result of this switch from structural uniformity in the ER to the gain of diversity in the Golgi complex, there occurs a change in glycan function. These glycan moieties conjugated to the proteins plays a role in promoting protein folding, quality control and sorting events (Figure 13). The complex sugar chains on the glycoproteins as a result of modification in the Golgi assign more specific and novel functions (Helenius A et al, 2001).



TRENDS in Plant Science

Figure 13. Biosynthesis of N-linked core glycans and recognition of potential cargo by LMN1 (Pattison R & Amtmann A, 2009)

In the biosynthesis of N-linked glycans, initially the process begins in the cytosolic surface of the ER membrane when a 14 saccharide “core” unit is put together, by the addition of one sugar at a time to dolichylphosphate. After two N-acetylglucosamines and five mannoses are added to the precursor, the oligosaccharide flips to the luminal side where seven more glucose units are added. Oligosaccharide enzyme complex transfers this core unit to the nascent polypeptide chain, after the last three glucoses are conjugated. Immediately after the transfer, the last three glucose units are trimmed off by glucosidase I and II, followed by further trimming of terminal mannoses by one or more different ER mannosidases. Glucosyltransferase, together with glucosidase II can reglucosylate glucose free chains and thus establish a deglucosylation-reglucosylation cycle. After the action of glucosidases, calnexin-calreticulin cycle folds the protein. At this step, the newly synthesized protein is folded (Figure 13).

Meanwhile LMAN1 monomers synthesized in the lumen of the ER are anchored to the ER membrane by their transmembrane domain (Figure 14) (Hauri H-P et al, 2000). These monomers are then oligomerized. Keeping in mind the localization of LMAN1 mostly in the ERGIC and also sometimes in ER, it is likely that LMAN1 captures its cargo- $\text{Man}_8\text{-N-Glycans}$ derived from $\text{Man}_9(\text{GlcNAc})_2$ (Figure 13) after the initial processing and trimming in the ER. LMAN1 recognizes and binds the terminal mannoses. However it should be noted that the exact ligand for LMAN1 is still ambiguous. It has been reported to bind to both Man_9 as well as Man_8 structures (Appenzeller C et al, 1999) but in either case it seems logical that LMAN1 would interact with the terminal mannoses joined by α (1-2) glycosidic bond of Man_9 or Man_8 glycoproteins. After LMAN1 conjugates with its cargo glycoprotein, the

complex of cargo-LMAN1 is recruited to budding vesicles by interaction of FF residues in the cytoplasmic domain of LMAN1 with COPII coat (Figure 14).

As a result, vesicular transport to the ERGIC takes place. In the ERGIC, the cargo is released and LMAN1 is recycled back to ER via interaction of KK residues and COPI coat in a retrograde transport. The folded protein is then directed for anterograde transport to the Golgi complex (Hauri H-P et al, 2000).

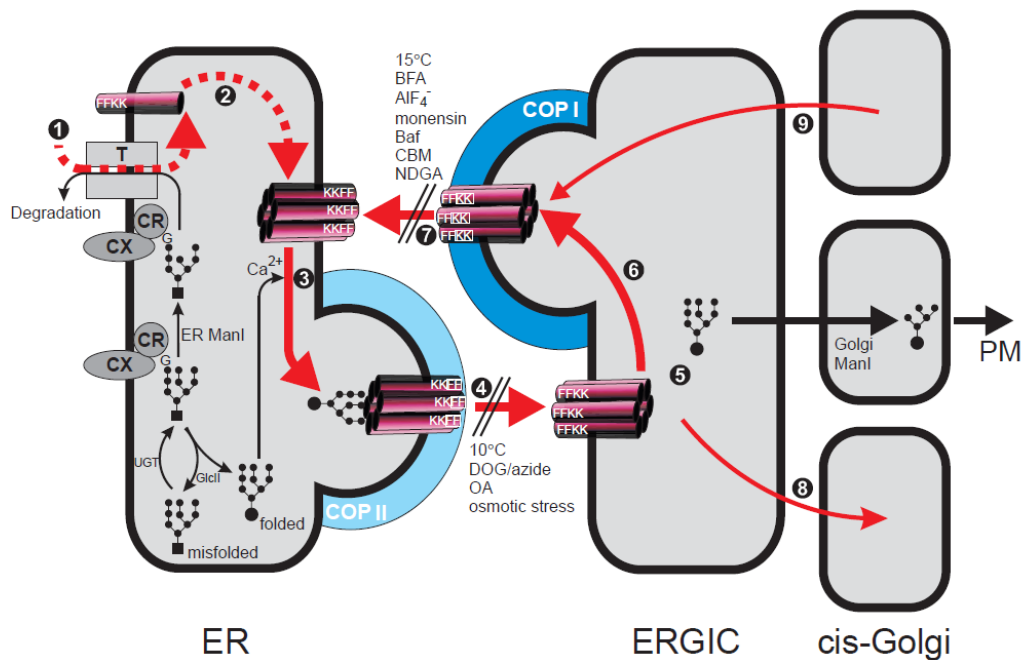


Figure 14. Cargo receptor function and recycling of LMAN1 (Hauri H-P et al, 2000)

1.3.4 Other potential cargo receptors in mammals

Apparently there are no other definitive cargo receptors known in mammals except the LMAN1- MCFD2 complex. But research revealed the presence of some other receptors in yeasts and it is likely that their homologs might serve as cargo receptors in higher eukaryotes. One such example is the p24 protein family but the function of the mammalian homolog varies from that of the yeast. One member of this

family, p21 is a transmembrane trafficking protein of amyloid precursor protein and regulates secretase activity (Orsi L et al, 2003).

LMAN1 is a member of animal L-type lectin family of protein, the other members of this family includes LMAN1-like protein(LMAN1L), LMAN2(also called VIP36) and LMAN2 like protein(LMAN2L or VIPL) (Nufer et al, 2003). LMAN2 and LMAN2L are promising cargo receptors based on some early evidences (Neve et al, 2003). BAP31 or the B cell receptor associated protein is another of the potential cargo receptors which might be involved in selective transport of MHC-I molecules (Zhang Y & Williams D, 2006).

1.4. Multiple coagulation factor deficiency 2 (MCFD2): Partner of LMAN1 in cargo capture & trafficking

MCFD2 is a small 16-kDa protein, soluble in nature having an N terminal sequence containing 42 amino acids and 2 calmodulin like EF hand domains at its C terminus. The EF hand domain contains a Ca^{2+} binding loop flanked by 2 α helices. In MCFD2 there are total 4 α helices and they are designated as helix 1 through 4. The EF hand domains of MCFD2 are connected by a linker region which is largely unstructured and of 10 amino acids long (Guy J et al, 2008).

MCFD2 homologs exist in different species of eukaryotes like rat, cow and fish. NMR studies have revealed that Ca^{2+} induces proper folding of the EF hands of MCFD2. In absence of Ca^{2+} the protein became disordered and the characteristics resembled that of an unfolded protein. Secondary structure or the existence of a hydrophobic core of the protein could not be determined in absence of calcium. However, in presence of the metal ion, both secondary and tertiary structures existed (Figure 15) (Guy J et al, 2008).

LMAN1 forms a Ca^{2+} dependent complex with MCFD2 in a 1:1 stoichiometry. The β 1 strand of the CRD of LMAN1 binds to the EF hand domain of MCFD2 and the LMAN1-MCFD2 complex interacts with cargo proteins FV and FVIII. Interaction of the cargo molecules with this complex is independent of the glycosylation state of the coagulation factors but is necessitated by calcium concentration in the ER (Zhang B et al, 2005). Unlike LMAN1, MCFD2 lacks the ER exit and retrieval motifs and the localization of MCFD2 in the ERGIC is dependent on LMAN1. LMAN1 has several other potential cargos in addition to FV and FVIII, but the role of MCFD2 in trafficking of these cargo proteins is not clear. MCFD2 missense mutants identified in F5F8D patients failed to form complex with LMAN1, however it was still able to interact with FVIII (Zhang B et al, 2005). This suggested that interaction of MCFD2 with the blood coagulation factors was independent of its interaction with LMAN1 (Zhang B et al, 2005). As mentioned earlier 30% of the F5F8D patients have mutations in MCFD2, from this it is definite that this protein also play a pivotal role in the cargo capture but the molecular mechanism and in depth function of this protein is still unknown.

Mutagenesis studies and circular dichroism spectroscopy revealed that the EF hand domains of MCFD2 are sufficient and necessary for interaction with LMAN1 (Zheng C et al, 2010a). Crystal structure of the CRD-MCFD2 complex confirmed the interaction of the EF hand domains of MCFD2 with the N-terminus of LMAN1 (Nishio M et al, 2010, Wigren E et al, 2010). EF hand domains of MCFD2 also mediate its interaction with FV and FVIII, but the sites of interactions are different from those involved in LMAN1 binding. From the more profound decrease in FV and FVIII levels observed in F5F8D patients with MCFD2 mutations (Figure 2), it seemed like MCFD2 plays a more direct role in cargo receptor function. It is likely that

MCFD2 recruits the cargo and brings FV/FVIII in close proximity of LMAN1 for further interactions.

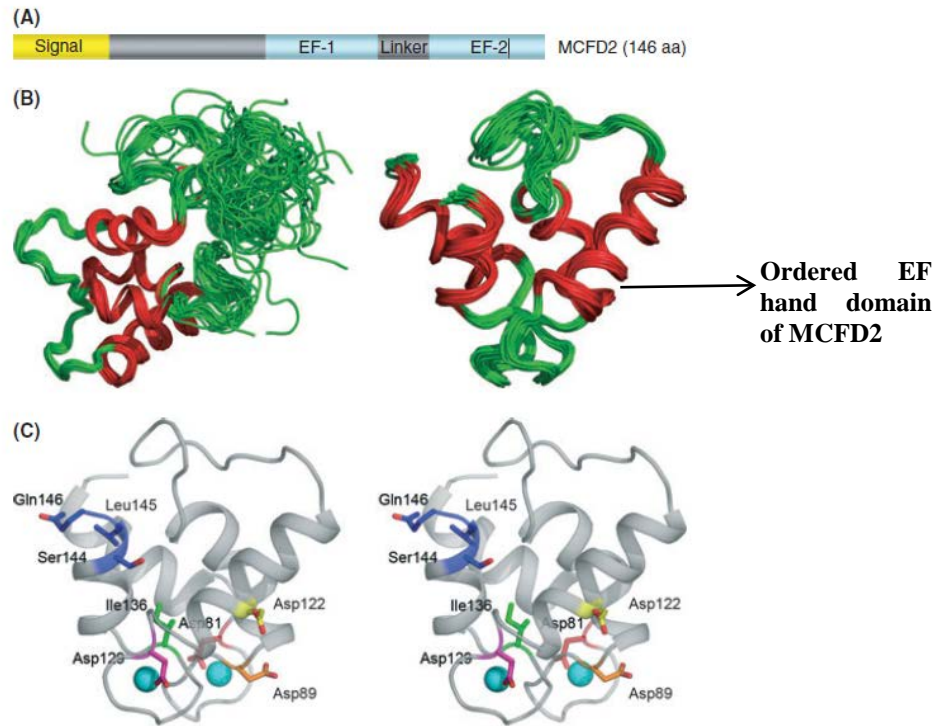


Figure 15. (A) MCFD2 domain structure. (B) NMR structure of MCFD2- left: complete structure, right : EF hand domain. (C) Stereo view showing the mutations in ordered EF hand domain (shown in color) of MCFD2. The cyan spheres represent Calcium ions (Guy J et al, 2008)

CHAPTER II

MATERIALS AND METHODS

2.1 Plasmid construction

To understand the mechanism of cargo release by LMAN1, we need to purify its functional carbohydrate recognition domain (CRD) which is the part of the molecule that binds to its cargo. For this purpose we designed our construct to express the CRD in *E. coli*.

2.1.1 Designing the fusion protein

Initially we created a fusion protein where our target protein (CRD) was fused with GST to enhance its solubility and facilitate in the purification process by affinity chromatography. For designing the fusion construct, we used pSJ6 plasmid that is derived from pET41a (Novagen) strain of *E. coli* and contains a GST tag, His₈ tag and a TEV protease cleavage upstream of BamH1 restriction site. The GST coding sequence was amplified from pSJ6 plasmid using forward primer **tataccatgggatcccctatactaggttattgg** and reverse primer **tataccatggcatccgattttggaggatgg** which incorporated Nco1 site at the start and end position.

Amplified GST was digested with Nco1(New England BioLabs) and was ligated to Nco1 digested pET15b vector into which residues 31-263 corresponding to the CRD of human LMAN1 was cloned between BamH1 and Nde1 sites. The amino acid sequence of the CRD that was cloned for our research is:

dgvggdpava lphrrfeyky sfkqphlvqs dgtvpfwaha gnaipssdq i rvapslksqr gsvwtktkaa
fenwevevtf rvtgrgriga dglaiwyaen qglegpvfgs adlwngvgif fdsfdndgkk nnpaiviign
ngqihydhqn dgasqalasc qrdfrnkpyv vrakityyqn tltvminngf tpdkndyefc akvenmiipa
qghfgisaat ggladdhdvl sfltfq

Nco1 digested plasmids were ligated such that GST is in the N terminal of the CRD with quick ligase enzyme (Promega) and the ligated plasmid was transformed into DH5 α cells (Invitrogen) followed by extraction of the plasmid (pET15b-GST-CRD) from the inoculated culture using Qiagen Miniprep kit. The extracted plasmid samples were confirmed by sequencing analysis before using them for transformation into BL21 cells to produce the fusion protein. Colonies from BL21 cells were inoculated in overnight express instant TB medium (Novagen) at 16°C for 60 hours. Auto induction media at lower temperature for a longer period of time enhances the solubility of the target protein. The cells were then harvested and lysed by sonication (with 550 Sonic Dismembrator from Fisher Scientific, total sonication time 15 minutes, with 15 seconds “on” and 15 seconds “off” pulses) using 50 mM Tris-HCl, 250 mM NaCl, and 20 mM CaCl₂, adjusted to pH 8.0. The lysed cell suspension was centrifuged at 20,000g for 20 minutes at 4°C.

The supernatant was purified using GStap 4B 1ml column from GE Life Sciences using the same buffer as mentioned earlier for binding and washing. The flow rate for the entire purification process was optimized to 0.5 ml/min for maximum

binding. After washing the column with 30 column volume (CV) of wash buffer (to remove nonspecific proteins), the recombinant protein bound to the column was eluted using 50 mM Tris-HCl, 250 mM NaCl, 20 mM CaCl₂, 50 mM L-Arginine and 20 mM reduced Glutathione (GSH), adjusted to pH 8.0.

The eluted sample was then subjected to overnight digestion at 4°C by Thrombin (5 U/mg from Sigma) to cleave the GST tag. Simultaneously the sample was also dialyzed overnight to remove GSH. The dialyzed and digested sample was then re-purified using the same GSTrap column and the buffer that was initially used for binding and washing. The purified target protein was collected in the flow-through fraction. At each step of purification, all fractions were collected to be analysed by electrophoresis to determine the correct molecular weight and the purity of the target protein sample. 12% Precast Tris-Hepes gel (Thermo Scientific) was used for electrophoresis of the samples using Hepes buffer and SeeBluePlus2 Prestained Protein Standard 1X (Invitrogen).

2.1.2 Designing the construct pET15b-CRD (wild type)

2.1.2.1 Making the insert

The CRD of human LMAN1 corresponding to amino acids 31-266 was amplified by forward primer LMAN1-CRD-S: GAGTC TCATAT Ggacg gcgt gggag gagac that introduced Nde1 site at the beginning and reverse primer CRD-AS-L: GACTG GATCC tcaag gtca gtaa ctgaaag incorporating a BamH1 site at the end. This differed from the other construct that we used for creating fusion protein by having 3 extra amino acids – T, E and P at the C terminal end. Hence this new CRD had sequence: dvggdpava lphrrfeyky sfgkphlvqs dgtvpfwaha gnaipssdqj rvapslksqr gsvwtkkaa fenwevevtf rvtgrgriga dglaiwyaen qglegpvfgs adlwngvgif fdsfdndgkk

nnpaiiign ngqihydhqn dgasqalasc qrdfrnkpyp vrakityyqn tltvminngf tpdkndyefc
akvenmiipa qghfgisaat ggladdhdvl sfltfq**tep**

F-549 S Phusion Hot Start High- Fidelity DNA Polymerase (2U/ μ l) along with F-518 5X Phusion HF buffer containing 7.5 mM MgCl₂ from Finnzymes now known as Thermo Scientific is used for amplification.

We set up 100 microliter reaction for the amplification of the CRD of LMAN1. Techne TC 512 served as thermal cycler for this purpose.

The pipetting protocol for setting up the PCR is:

Component	100 μ l Reaction	Final Concentration
H ₂ O	66 μ l	
5X Phusion HF buffer	20 μ l	1X
10 mM dNTPs	2 μ l	200 μ M each
10 μ M Forward Primer	5 μ l	0.5 μ M
10 μ M Reverse Primer	5 μ l	0.5 μ M
Template DNA	1 μ l	40 ng
Phusion Hot Start II DNA Polymerase	1 μ l	0.02 U/ μ l

We followed a 3 step cycling process for the amplification. The CRD of LMAN1 is 720 base pairs; hence accordingly the annealing temperature and extension temperature was calculated. The amplification process was:

Cycle Step	Temperature	Time	Cycles
Initial Denaturation	98°C	30 s	1
Denaturation	98°C	15 s	
Annealing	59°C	30 s	25-35
Extension	72°C	25 s	
Final extension	72°C	10 min	1

Since our amplified PCR product- the longer form of CRD was in the specified range of 10-100kB of QIA quick PCR Purification Kit, it was purified using this kit to remove primers, nucleotides, polymerases and salts. The DNA was finally eluted from the QIAquick column using 50µl of double distilled H₂O. Concentration and quality of the eluted plasmid is estimated by Nanodrop ND-1000 spectrophotometer and is found to be 97.1 ng/µl.

50 µl of the eluted plasmid was set up for digestion with NdeI (New England BioLabs, strength 20,000 u/ml) and BamHI (New England BioLabs, strength 20,000 u/ml) with NE Buffer 3 (New England BioLabs, 10X concentration) and purified BSA (New England BioLabs, 100X) at 37° C for three and half hours.

Component	Amount
Amplified CRD long form	34 µl
H ₂ O	7.5 µl
Bam H1	1.5 µl

NdeI	1.5 μ l
NE buffer 3	5 μ l
BSA	0.5 μ l

The digested product was again purified using the QIAquick PCR purification kit. Concentration of the digested product was estimated by Nanodrop spectrophotometer and concentration was 60 ng/ μ l. This would serve as the *insert* for the vector used for expressing the CRD of LMAN1.

2.1.2.2 Choosing the vector

For cloning the insert, we decided to use pET15b vector from Novagen (Cat. No. 69661-3). This carried an N terminal His tag sequence followed by a thrombin site and three restriction sites.

We already had the construct where the regular CRD (without three extra amino acids) was ligated into pET15b vector. This vector was digested (described in the previous section) with NdeI and BamHI which removed the regular CRD from the vector. The concentration of the vector was 300 ng/ μ l and a 50 μ l reaction was set up as described earlier. After digesting the vector for three hours at 37°C water bath, the sample was electrophoresed on a 0.8% agarose gel. The double enzyme digestion with BamHI and NdeI resulted in 2 bands on the agarose gel – a shorter band which is the CRD 720 bp and a bigger band of the vector running at ~ 5.7 kB (comparing with 1 kB plus ladder from Invitrogen). The empty vector was then gel extracted using QIAquick Gel Extraction Kit and concentration is again estimated with Nanodrop spectrophotometer. It was 12.5 ng/ μ l.

pET-15b sequence landmarks	
T7 promoter	453-469
T7 transcription start	452
His-Tag coding sequence	362-380
Multiple cloning sites (<i>Nde</i> I - <i>Bam</i> H I)	319-335
T7 terminator	213-259
lacI coding sequence	(866-1945)
pBR322 origin	3882
<i>bla</i> coding sequence	4643-5500

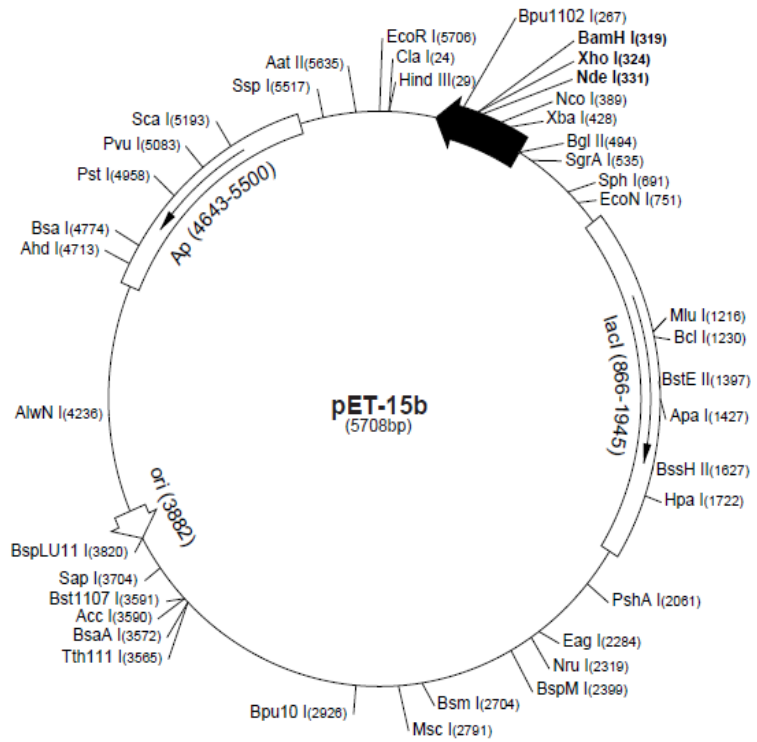


Figure 16. pET15 b vector sequence

2.1.2.3 Ligation

The digested insert and the empty vector was ligated using T4 DNA Ligase (3 u/μl from Promega) and 2X ligation buffer (Promega) incubating the samples at room temperature for 10 minutes. The volume of vector and insert used for ligation was in the ratio of 1:3. One control tube was set up with just vector and ligation mix but without the insert. 3 μl of the ligation mix from each tube was then used for transforming 40 μl of DH5 α competent cell from Invitrogen. The transformed samples were incubated in ice for 30 minutes followed by heat shock at 42°C water bath for 25 seconds and 2 minutes incubation in ice. Finally, 500 μl of SOC medium (Sigma) was used to recover the cells and was then incubated in a 37°C shaker for one

hour. The transformed cell suspension (100 µl) was then spread on an ampicillin plate and kept in a 37°C incubator overnight.

2.1.2.4 Cloning

Several colonies of DH5α cell transformed with the CRD in pET15b vector were picked up from the ampicillin plate and inoculated in 3 mls of LB with 3µl of the ampicillin. This was allowed to grow overnight in 37°C shaker followed by plasmid extraction with QIAprep Miniprep Kit. These plasmids were then digested with Nde1 and BamH1 to screen for the right insert. The digested samples were electrophoresed on a 1.5% agarose gel to confirm the presence of the correct insert. Finally the desired colonies were sent for sequencing with primer TAATACGACTCACTATAGG. Of the six colonies, only one carried plasmid DNA with the correct insert and the vector. This is our wild type pET15bCRD construct.

2.2 Constructing mutant plasmids with pED vector

For preparing mutant LMAN1 in mammalian cells, we used a different vector called pED. Flag tagged full length LMAN1 was derived from pED-LMAN1 (Moussalli M et al, 1999) by replacing the LMAN1 signal sequence with that of calreticulin and introducing a flag tag after the signal sequence. This was used as the template for generating mutant plasmids. Point mutations were introduced at S88A (forward primer: caa att cga gta gca cca tct tta aaa gcc caa aga ggc tcag & reverse primer: ctg agc ctc ttt ggg ctt tta aag atg gtg cta ctc gaa ttg), H-178A (forward primer: atg gac aaa tcc att atg acg ctc aaa atg acg ggg cta gtc & reverse primer: gac tag ccc cgt cat ttt gag cgt cat aat gga ttt gtc cat), F154Y (sense primer: gga atg gtg ttg gaa tat ttt ttg att ctt atg aca atg atg gaa aga aaa ata & antisense primer: tat ttt tct ttc cat cat tgt cat aag aat caa aaa ata ttc caa cac cat tcc), G251A/252A (sense primer: gaa tat ctg ctg caa

ctg cag ctc ttg cag atg acc atga & antisense primer: tca tgg tca tct gca aga gct gca gtt GCA gca gat attc) using QuikChange II XLSite-Directed Mutagenesis Kit (Agilent technologies/Stratagene). All mutations were confirmed by sequencing with primer aacgc gactg caggt cgaca tgctg ctatc cgtgccgc.

2.3 Protein expression and purification

2.3.1 Expression host transformation using the construct pET15bCRD

The construct pET15bCRD was used for transforming BL21 cells from Invitrogen which served as the host. BL21 was one of the several different host strains available as DE3 lysogens. Transformation procedure was same as described earlier. The only difference was now for transformation into an expression host, 1 μ l (~ 1ng) of plasmid in sterile water was used. The transformants were streaked on a plate for single colonies.

2.3.2 Induction of lysogen

After the target plasmid was established in BL21 cells, expression of the target DNA (pET15bCRD) was induced by the addition of IPTG to a growing culture. For full induction, we used a final concentration of 1 mM of IPTG.

To prepare for induction, a single colony from a freshly streaked plate was inoculated into 4 mls of LB containing ampicillin. This was allowed to grow in a 37°C shaker, 230 rpm until the OD₆₀₀ reaches 0.6-1.0. 500 μ l of this culture was then allowed to inoculate 500 mls of LB at 37°C shaker until the OD₆₀₀ reaches 0.8-1.0. 1 mM IPTG was then added and incubated for 4-5 hours more. A small volume of uninduced sample was set aside for control.

2.3.3 Confirmation by SDS-PAGE

Bacterial pellets were harvested from small volumes (1ml each) of induced and uninduced samples on a table top centrifuge 5415 R (Eppendorf). The pellet was lysed by sonication using lysis buffer [50 mM Tris-HCl and 150 mM NaCl (pH 7.4)], supernatant was electrophoresed on a reducing gel to confirm the expression of the target protein, the CRD of LMAN1.

2.3.4 Preparation of inclusion bodies

Bacterial pellet was harvested from the remaining of the induced culture by centrifugation at 6,000g for 15 minutes in a Sorvall RC 6 centrifuge from Thermo Scientific. After washing the pellet with water once, it was resuspended in lysis buffer (50 mM Tris-HCl, 150 mM NaCl, pH 7.4). For complete resuspension the volume of lysis buffer used was optimized to 7 ml/gm wet weight of bacterial pellet. About 1 mg of lysozyme (Fisher Scientific BP535-1) per 10 ml was added to the resuspended culture and incubated in ice (4°C) for 30 minutes with occasional swirling. The bacterial cells were then lysed by sonication with 550 Sonic Dismembrator from Fisher Scientific at power 4 on a 10 seconds “on” and 10 seconds “off” cycle for a total duration of 15 minutes at 4°C. After cell disruption, cellular DNA is digested by adding 10 units of DNase I (1u/ml from 2500u/ml stock of DNase I) and 3 mM of MgCl₂ were added and incubated at room temperature for 20 minutes. This combination of enzymatic and mechanical cell lysis would render in maximum cell disruption. Finally the inclusion bodies were harvested from the cell lysate by a Sorvall RC6 centrifuge at 26,000g for 30 minutes at 4°C.

2.3.5 Washing of inclusion bodies

Once we had isolated the inclusion bodies with the target protein, we needed to wash it in order to reduce contaminant levels. We used two different types of wash buffer for this process and referred it as Wash Buffer 1(1.5 M NaCl, Triton X-100 2%, 20 mM EDTA, pH 7.0) & Wash buffer 2(0.1 M Tris-HCl, 20 mM EDTA). The inclusion bodies were completely suspended in 30 mls of Wash buffer 1 and collected by centrifugation at 6,000 g for 15 minutes at 4°C. This step was repeated twice for each of the buffers. This stringent method of washing would yield 80% pure inclusion bodies.

2.3.6 Protein refolding

In order to obtain functional target protein, we needed to solubilize the inclusion bodies in detergents or surfactants to release the protein followed by refolding of the solubilized protein to give them native three dimensional conformations. There were different methods for solubilization and refolding that can be followed. In the following page, there is a schematic representation of the various solubilization and refolding strategies (Figure 17).

A. Solubilization: For our purpose, we suspended inclusion bodies in 6 M GuHCl such that the protein concentration is 5-8 mg/ml and incubated at room temperature for 30 minutes. The concentration of solubilized protein was determined by Bio-Rad Protein Assay based on the principle developed by Bradford. We used microtiter plates and recorded the absorbance on an absorbance microplate reader (Spectraplus Max 384 from Molecular Devices) set at 595 nm. The standard curve was prepared using serial dilutions of BSA from 2 mg/ml to 10 mg/ml. The supernatant was

collected after centrifuging the samples on a table top ultracentrifuge (5415 R from Eppendorf).

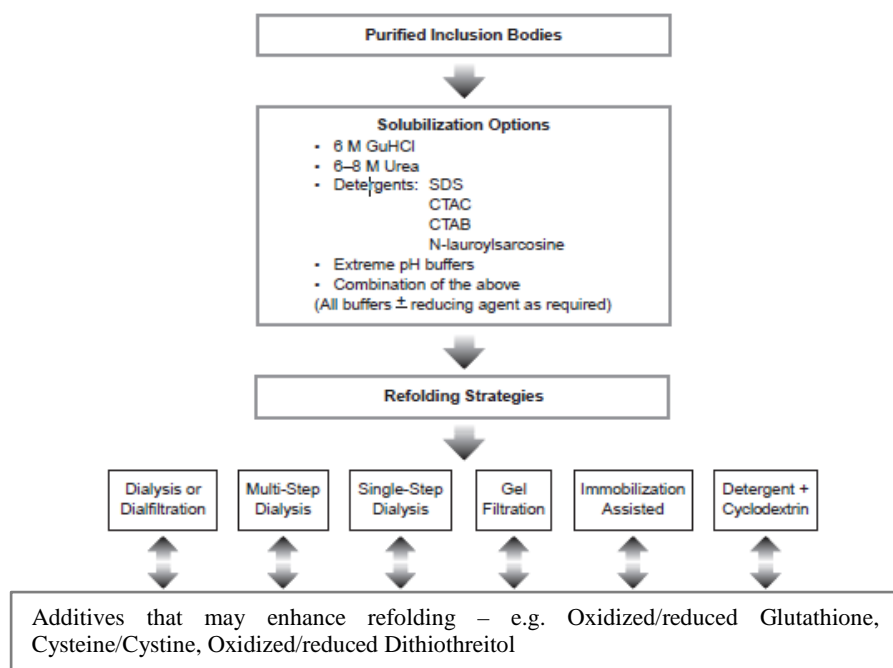


Figure 17. Schematic Representation of various solubilization and refolding strategies (Represented from Novagen Protein Refolding Kit manual)

B. Renaturation of released protein: Refolding was carried out in a redox environment using oxidized and reduced glutathione. After several tries, the refolding buffer that worked best for our protein was composed of 50 mM Tris-HCl, 0.4 M L-arginine, 5 mM reduced glutathione and 0.5 mM oxidized glutathione (pH – 8.0). The volume of refolding buffer was optimized such that finally the solubilized protein would be diluted 150 folds (~120 µg/ml). Solubilized sample was added to refolding buffer in pulses (1 ml per injection) over a period of 36 hours and the entire protocol was carried out in 4°C. Since we were dealing with large volumes of buffer, the refolded protein sample needed to be concentrated by Amicon Ultra 15 Centrifugal Filter Units with 10 kDa cut off from Millipore. Final volume of the refolded protein was reduced to 1ml at concentration of ~ 8-10 mg/ml.

2.3.7 Protein purification

We used gel filtration chromatography (size exclusion chromatography) as the preferred method for our experiment to remove the nonspecific contaminants. Our target protein, the purified CRD monomer has a MW of 28.6 KDa; hence Superdex 75 column was chosen for purification purpose as it is efficient for separating proteins in the range of 3-70 KDa. We passed our refolded and concentrated sample through Superdex 75 column (10/300Gl, Code No. 17-5174-01) in Hepes buffer (10 mM Hepes, 150 mM NaCl, pH 7.4) using the AKTA purifier UPC900 (GE healthcare).

The entire purification process was carried out at 4°C. The column volume was 24 mls and initially it was equilibrated with 2 column volume (CV) each of ddH₂O and the Hepes buffer. After the equilibration, the purification process was started through the software (Unicorn) controlling the purifier. In each injection, 500 µl of the concentrated sample was loaded in the purifier using a (Hamilton) syringe of similar volume. All the peaks recorded in the chromatogram were collected for analysis by SDS-PAGE. After the purification process was over, we used the same column for separating protein standard (Gel Filtration Standard from Bio-Rad).

All samples were electrophoresed on a NuPAGE 4-12% Bis-Tris gel (Invitrogen) using MES buffer (50 mM MES, 50 mM Tris –base, 3.5 mM SDS, 0.8 mM EDTA) and SeeBluePlus2 Prestained Protein Standard 1X (Invitrogen). The gel was stained with Coomassie blue G-250 for 1 hour and destained with ddH₂O for 3-4 hours. The SDS-PAGE results confirmed that the peak 2 from the purification is definitely our desired protein CRD monomer.

The purified sample corresponding to peak 2 had a volume of ~ 5 mls which was further concentrated using Millipore Amicon Ultra 10K 4 mls tube. Quantity of

the protein in the concentrated sample was estimated by Bio-Rad Protein Assay. It was observed that the yield of purified CRD is ~ 50% compared to the concentration of the solubilized inclusion bodies. Hence the steps of protein expression through purification were done in batches, repeating it several times to get enough amount of protein for further biochemical experiments.

2.4 Expression and purification of MCFD2

MCFD2 was ligated to pET15b vector in a similar way as described in Section 2.1. Likewise it was cloned in DH5 α cells (procedure same as section 2.1.6). Eventually the construct was expressed in BL21 cells and the soluble fraction of the cell lysate was harvested (procedure same as in section 2.2).

For purification of MCFD2 we used a modified approach of the QIAexpressionist kit which is based on the principle of immobilized metal affinity chromatography (IMAC) as our construct has 6x histidine tagged on the N terminal end. Bacterial pellet from 250 ml culture was obtained by centrifugation at 6,000 g for 20 minutes. The pellet was then resuspended in 2 ml/gm wet weight of lysis buffer (50 mM Tris-HCl, 50 mM NaCl and 10 mM imidazole adjusted to pH 7.5). The resuspended pellet was incubated at 4°C for 30 minutes with ~1 mg of lysozyme and 1 protease inhibitor cocktail tablet (Roche Diagnostics) followed by lysing the cells with a sonicator (550 sonic Dismembrator from Fisher Scientific) set to power “4” in a 10 seconds “on” and “off” cycle for a total duration of 8 minutes. The crude cell lysate was treated with 10 units of DNase per 10 mls and 3 mM MgCl₂ and incubated at 4°C for 15 minutes. MCFD2 being a soluble protein, the supernatant was collected after centrifugation (Sorvall RC 6 centrifuge) at 26,000 g for 30 minutes at 4°C. From 500 ml of bacterial culture, the volume of clear cell lysate obtained was ~10 ml.

Meanwhile 1 ml of the 50% Ni-NTA slurry was added to a 15 ml falcon tube, centrifuged at 1,000 rpm for 2 minutes and the supernatant is decanted. The slurry was then completely resuspended in 4 ml of lysis buffer for equilibration and then the buffer is decanted by centrifuging for 2 minutes at 1,000 rpm. Clear cell lysate was added to the equilibrated slurry and mixed gently by shaking (~200 rpm) at 4°C for 60 minutes. The lysate-Ni NTA was then loaded in a column and flow-through collected for SDS PAGE analysis. This was followed by washing the lysate-Ni NTA mix with 4 ml of wash buffer 2 (50 mM Tris-HCl, 50 mM NaCl, 20 mM imidazole, pH 7.5) and wash fraction was collected in separate tubes for every 1 ml for SDS-PAGE analysis. Finally MCFD2 is eluted 4 times with 1 ml elution buffer (50 mM Tris-HCl, 50 mM NaCl and 250 mM imidazole adjusted to pH 7.5). Each elution fraction is collected in separate tubes.

SDS-PAGE analysis on a 4-12% gel (NuPAGE 4-12% Bis-Tris gel from Invitrogen) with MES buffer revealed that the 1st and 2nd elution fractions contain our protein of interest. The first and second elution fractions were then combined and the volume was reduced to 1ml by Amicon Ultra Centrifugal Filter Units with 3 KDa cut off. The concentrated sample was digested with thrombin overnight (5u/mg of protein) to remove the histidine tags and then purified with superdex 75 column. Functionality test was then carried out with the purified protein by ITC assay as described in Section C. MCFD2 is similar to the CRD with 2 Ca²⁺ binding sites.

2.5 Isothermal titration calorimetry (ITC)

For this assay, we used a single injection instrument ITC₂₀₀ from Microcal. Initially a control experiment was done by injecting 200 µl of the same Hepes buffer used for gel filtration chromatography in the sample cell but without the

macromolecule (the purified CRD monomer) as well as to the reference cell and is titrated against 5 mM Ca^{2+} dissolved in same buffer. The control experiment was followed by injecting 200 μl of the purified CRD monomer (200 μM) in the sample cell; the reference cell contains only the buffer minus the macromolecule and titrated with 5 mM Ca^{2+} . The ligand concentration should be ideally 10-100 folds higher than the macromolecule. The parameters of the titration were entered into the software program controlling data acquisition. The number, volume, and length of time of injection are critical. The total number of injections was set to 20 and each time 2 μl of the ligand was injected into the sample cell. The experiment was carried out at room temperature (25°C). The space between each injection was about 3 minutes. Each peak displayed in the monitor represented a heat change and this heat change was plotted for each injection to get the binding curve. The binding curve was then analysed with the appropriate binding model to determine by binding constant, stoichiometry and entropy.

ITC assay was also used for determining the binding constant of the CRD with its ligand mannobiose 2 α (Sigma) where the functional CRD (200 μM) in 10 mM Hepes, 150 mM NaCl and 5 mM Ca^{2+} , pH 7.4 served as the titrant and the titre is 30 mM of mannobiose 2 α . To check the effect of pH on the refolded and functional CRD, 200 μM CRD was titrated with 2 mM Ca^{2+} but at different pH of the buffer. Three buffer conditions were used – 10 mM Hepes, 150 mM NaCl at pH 6, 6.5 and 7.4. We also titrated the purified CRD with 30mM mannobiose 2 α at different Ca^{2+} concentrations (2 mM and 0.4 mM).

2.6 Crystallization of CRD monomer with mannose α (1-2) Mannose [commercially 2 α mannobiose]

Our purified CRD monomer was in 10 mM Hepes, 150 mM NaCl at pH 7.4. For crystallization, the CRD was required to be in a buffer with 10 mM Hepes, 50 mM NaCl, and 4 mM CaCl₂ at pH 7.4. So a crystallization buffer was prepared with 10 mM Hepes and 6 mM CaCl₂, adjusted to pH 7.4. Now 2 parts by volume of this crystallization buffer was mixed with 1 part by volume of the purified CRD monomer. This was then concentrated to 10 mg/ml and divided into two tubes. To one tube, 100 mM D (+)-mannose (Sigma) was added and to other tube 10 mM of mannobiose 2 α (Sigma) is added. These samples were used for crystallization of CRD with its ligand.

2.7 Cell culture and transfection of COS1 cells

Vials containing COS1 cells from liquid nitrogen were quickly thawed at 37°C water bath with intermittent shaking. The contents of the vial were then pipetted gently into a 15ml falcon tube containing 5 ml Dulbecco's Modified Eagle medium supplemented with 10% fetal bovine serum and penicillin-streptomycin. The contents were then spun at 1,000 rpm for 5 minutes and the supernatant discarded. The cell pellet was washed once more with the 10 ml of media. Finally the cells were seeded on a 10 cm plate with ~8 mls of media that was warmed at 37°C water bath.

After 24 hours the growth media for cells were changed and after 48 hours, cells were given their first passage. When the cells were about 100% confluent and healthy, they were split again in a 1:4 ratio into five 100 mm plates. The following day COS1 cells were about 80% confluent and ready for transfection.

Five different plasmids were used for transfection – wild type pED flag tagged full length LMAN1 as a positive control and 4 mutants – S88A, H178A, F154Y, double mutant G251A/252A. We used 4 µg of plasmid for each plate and Fugene 6 (Promega) served as the transfection reagent. After the media of the five culture plates (when ready for transfection) were replenished with 7 ml of fresh media and kept back in the incubator; 2 ml of serum free media (for each 1 µg of plasmid, 100 µl of serum media is required) was taken in a tube. 60 µl of Fugene 6 (for every µg of plasmid, 3 µl of transfection reagent used, in total we had 20 µg of plasmid) was added to the tube containing serum free media. The mix was allowed to sit at room temperature for 5 minutes and then 412 µl of aliquot added to five tubes labelled with the corresponding mutants. 4 µg of each plasmid added to each tube and incubated at room temperature for 20 minutes. The contents of each tube were added to different culture plates drop wise and the plates were placed back in the incubator.

2.8 Mannose binding assay

Mannose binding assay was performed essentially as described in Zheng C et al, 2010b after making some modifications. Forty eight hours after transfection of COS1 cell with pED Flag LMAN1: wild type and the mutants, the cell culture plates were brought back and the media were removed. The plates were washed with ice cold 1X PBS and then 1 ml of homogenizing buffer (10 mM Tris [tris(hydroxyl methyl aminomethane)]-HCL, pH 7.4, 150 mM NaCl, 1 mM CaCl₂, and 1 mM MgCl₂) with one protease inhibitor cocktail tablet was added to each plate. The cells from each plate were scraped off and lifted to separate tubes. The cells were homogenized with a ball bearing homogenizer at 18-µm clearance 20 times. To check whether the cells were properly disrupted or not, 20 µl of cell homogenate is mixed with Trypan blue in a 1:1 ratio and observed under microscope on a hemocytometer.

The remainder of the cell lysate was cleared by centrifugation at 600g for 10 minutes. Each step was carried on ice. The supernatant from each tube was then transferred to Beckman-Coulter 4.7 ml optiseal polyallomer centrifuge tubes and centrifuged at 100,000 g for 1 hour at 4°C (Beckman Coulter Ultracentrifuge). The pellet was solubilized in the lysis buffer (10 mM Tris-HCl, pH 7.4, 150 mM NaCl, 10 mM CaCl₂, 1 mM MgCl₂, and 1% Triton X-100 with protease inhibitor) on ice for 1 hour. This resuspended pellet was again centrifuged at 100,000 g for 1 hour at 4°C. The supernatant was dialyzed against binding buffer (10 mM Tris-HCl, pH 7.4, 150 mM NaCl, 10 mM CaCl₂, 1 mM MgCl₂ and 0.15% Triton X-100) overnight at 4°C followed by incubation of the dialysate with 40 µl of equilibrated mannose beads (D mannose agarose, cross linked 4% beaded agarose from Sigma) overnight. The overnight incubated samples were centrifuged at 10.2 rcf at 4°C for one minute. The beads were washed three times with 1ml of binding buffer for every wash. The pellet was resuspended in 40µl loading dye (SDS and DTT), boiled at 95°C for 5 minutes and the supernatant was used for SDS-PAGE and Western blot.

2.9 Western blot

The samples were electrophoresed on a 4-12% NuPAGE Bis-Tris gel using MES (Sigma Aldrich) buffer [50 mM MES, 50 mM Tris –base, 3.5mM SDS, 0.8mM EDTA]. Small aliquots of samples prepared after dialysis was used as loading control as it was the start point of mannose bonding assay. Transfer was done using iBlot Dry Blotting system (Invitrogen) using 7 minutes for the transfer to a nitrocellulose membrane. After the preliminary blocking with 5% non-fat dry milk, overnight incubation of the membrane was done using mouse anti-Flag antibody (Sigma Aldrich) in 1:2000 dilution. After washing with PBS-T , it was followed by 1 hour incubation with secondary antibody (anti-mouse polyclonal conjugated with HRP

from Sigma Aldrich) in 1:10000 dilution, the membrane was washed with PBS-T three times, each wash being 10 minutes. Finally the expressed proteins were detected by treating the membrane for 5 minutes with a mixture of Lumigen PS-3 detection reagent Solution A (1 ml) and Solution B (25 μ l). The membrane was then exposed to X-ray film at various lengths of time.

CHAPTER III

RESULTS AND DISCUSSION

3.1 Expression of the CRD of LMAN1 as a GST fusion protein and purification by GST column

3.1.1. GST-CRD plasmid construction

To elucidate the exact ligand binding site and study the interaction of the carbohydrate recognition domain (CRD) of LMAN1 with its cargo glycoproteins we decided to purify the CRD of LMAN1 before proceeding with the necessary biochemical experiments. The CRD of LMAN1 is mainly expressed in inclusion bodies even under conditions optimized for enhanced solubility. We reasoned that tagging with GST at the amino terminus could enhance the solubility of the CRD. The GST tag would also facilitate the purification process by affinity chromatography using immobilized glutathione. The fusion plasmid construction is a two-step procedure (details in chapter 2.1) -

i. Amplification of GST from pSJ6 vector using Pfu

The GST open reading frame was amplified from pSJ6 plasmid which was a modified form of pET41a vector (figure 18) (details in section 2.1).

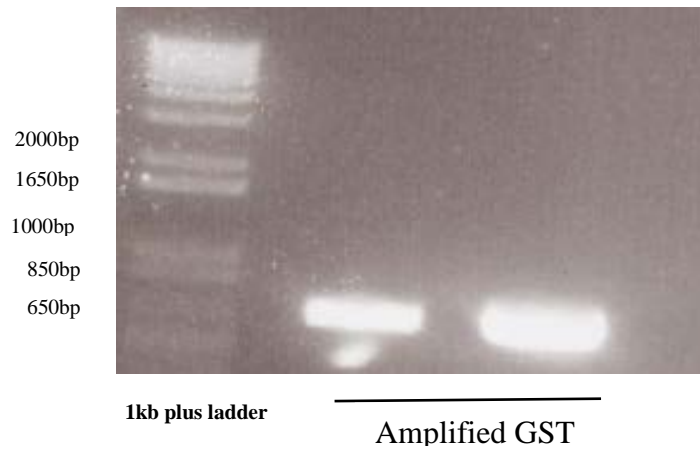


Figure 18. Amplified GST is about 700bp.

ii. Ligating amplified GST with vector containing wild type CRD

After we amplified GST with the high fidelity enzyme *Pfu*, it was ligated in frame to pET15b vector containing the wild type CRD of LMAN1 to generate GST-CRD, a fusion gene encoding the fusion protein (details in Chapter 2.1).

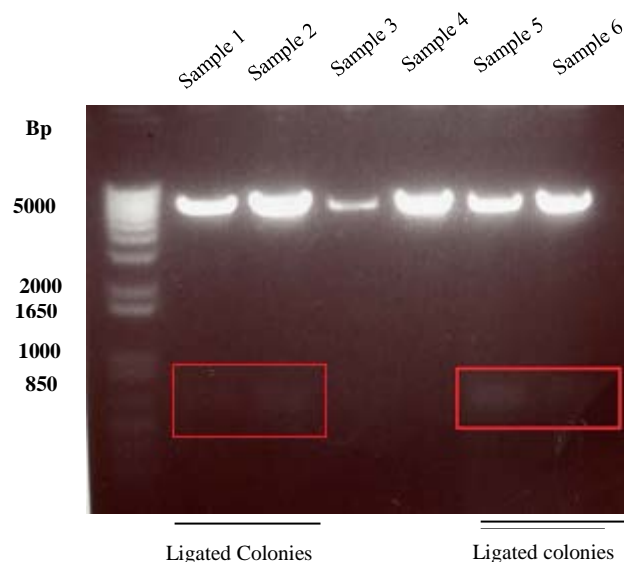


Figure 19. Ligation of amplified GST into pET15b vector with CRD and digestion with Nco1. Sample 5 and 6 showed ligated colonies following NCo1 digestion. Band of GST is visible around 700bp. The four ligated colonies were sequenced to confirm the fusion of GST to the CRD. Only Sample 1 had the correct insert and the vector.

3.1.2. Purification of the fusion protein

The majority of the GST fusion protein accumulated in the cytoplasmic fraction was captured by the affinity medium and impurities removed by washing. The wild type fusion protein expressed from pET15b-GST-CRD was purified as mentioned in Chapter 2.1. The four most critical factors that were optimized during the protocol:

- The composition and pH of the binding and elution buffer
- The flow rate of the sample during purification that allows optimum binding to the column.
- The temperature at which the experiment was carried out.
- Ideal way to cleave the GST tag from CRD after purification

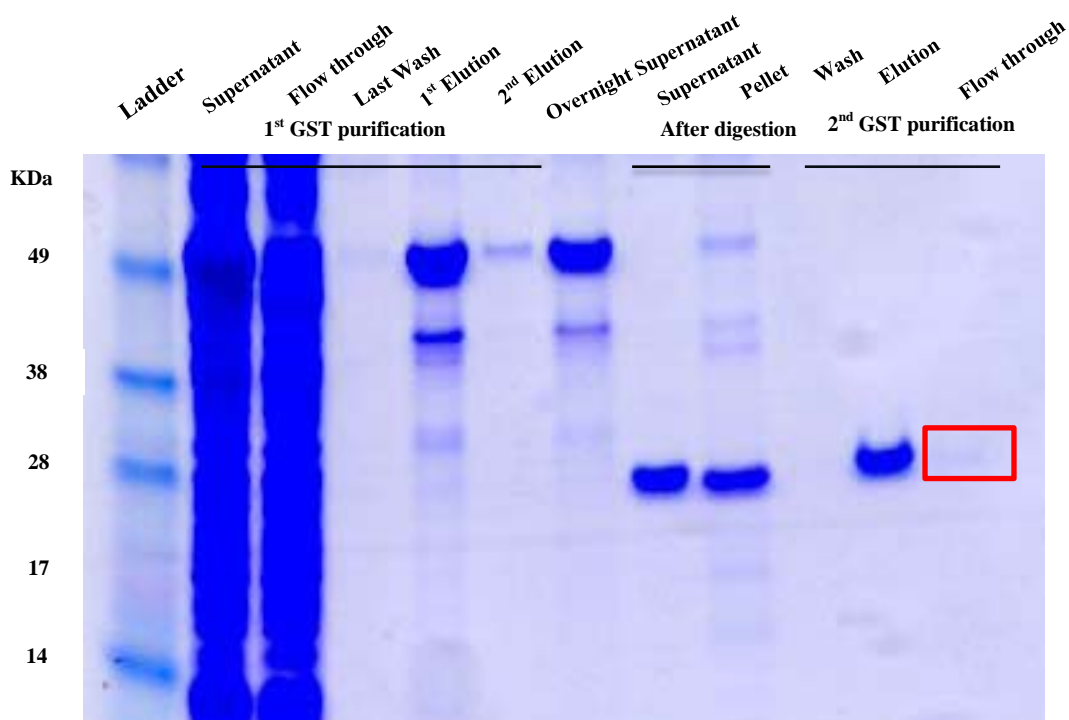


Figure 20. Coomassie Staining of SDS-PAGE of wild type GST-CRD fusion protein purified by GST column. The band of purified protein CRD is indicated by the red square. Signal strength of purified CRD is weak.

After the first round of purification of the fusion protein by the GST column, apparently we observed no significant precipitation when the eluted sample was stored overnight at 4°C. To be sure of any protein loss due to precipitation, we centrifuged the sample before proceeding with thrombin digestion. A small volume of the supernatant was used for SDS-PAGE analysis while the rest of sample was digested with thrombin off-column overnight at 4°C. As observed in lane 7 of Figure 20, the protein loss was minimal as indicated by the strong signal of the fusion protein around 50 KDa. Overnight digestion with thrombin resulted in large amount of precipitation and we electrophoresed both the supernatant and the pellet after centrifugation (lanes 8 & 9 of figure 20). When the digested sample (supernatant) was purified for the second time by GST column, our target protein (the purified CRD) was collected in the flow-through fraction and elution fraction yielded just the GST. The signal for the elution fraction was very strong after second round of purification as observed in lane-11 of figure 20. However, when the flow through fraction was analyzed, we observed a very weak band in lane-12 around 28.6 KDa. While carrying out the purification process we observed that considerable amount of protein was precipitating at each step of the protocol despite the efforts for optimization. From this we concluded that though we were successful in purifying the soluble CRD of LMAN1 but the yield was too low to carry out further biochemical experiments.

This poor yield was possibly due to aggregation of the CRD after thrombin treatment. As was mentioned earlier, the CRD alone is mainly expressed in inclusion bodies. Even though GST tag improved solubility of the fusion protein, cleavage of GST tag by overnight treatment with thrombin resulted in aggregation of the CRD. Removal of GST is likely making the protein unstable causing some conformational changes. Hence the majority of the initially soluble CRD becomes insoluble after

digesting with thrombin. This can be the reason why we observed significant amount of precipitate post digestion.

3.2 The CRD of LMAN1 expression and purification by the denaturation & refolding approach

Since we were unable to get sufficient amount of the soluble CRD using the GST fusion protein, we attempted a new approach. Our target protein CRD normally was expressed in the insoluble fraction as inclusion bodies in denatured form, so we decided to refold it and give it a native functional three dimensional conformation. For this approach, we decided to modify a method that has been reported in a literature (Nishio M et al, 2010). The CRD of LMAN1 that was used for this purpose is slightly different from the one used for creating the fusion protein. This new construct had residues 31-266 corresponding to the CRD of human LMAN1 cloned into pET15b vector. Hence this was longer than the previously used CRD by 3 amino acids – T, E and P at the C terminal end which may contribute to increase the stability of the protein. This DNA fragment cloned into empty pET15b plasmid was confirmed by sequencing and found to be in frame, then it was expressed into BL21 cells (figure 21) [details in Chapter 2.1].

A control experiment was performed with the wild type protein to make sure that the protein was expressed and of right molecular weight (figure 22). Bacterial culture was lysed by sonication and total, soluble and insoluble fractions were electrophoresed on a 4-12% Bis-Tris gel using MES buffer and stained with Coomassie.

Amplification of CRD and ligation into empty pET15b

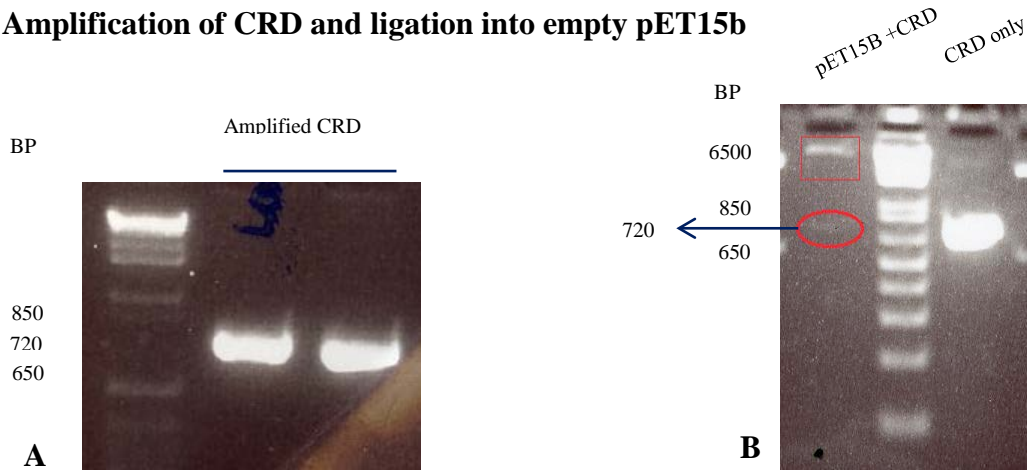


Figure 21. A. CRD of LMN1 amplified by Phusion Hot Start DNA Polymerase and electrophoresed on a 2 % agarose gel with 1kb plus DNA Ladder. B. Amplified CRD is ligated into pET15B Vector and digested with Nde1 and BamH1, electrophoresed on 2% gel agarose gel to confirm the presence of insert. The CRD alone (encircled by red ellipse) in 720 base pairs and vector is ~6.5kB (encircled by red square). Amplified CRD alone is also digested with enzymes and used as control (right lane beside 1kb plus ladder).

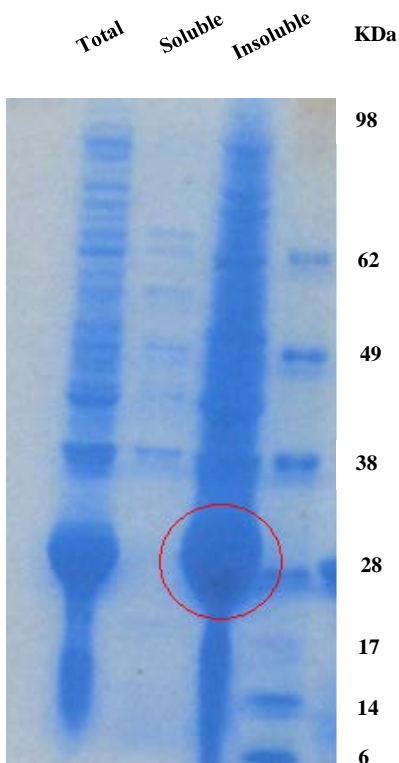


Figure 22. Coomassie R 250 staining of SDS-PAGE of the total, soluble and insoluble fractions of pET15BCRD expressed in BL21. The insoluble fraction shows that the target protein (marked by the red ellipse) is expressed and of the right molecular weight 28.6 kD.

3.2.1 Preparation of inclusion bodies and release of target protein pET15bCRD

Inclusion bodies are dense aggregates of misfolded polypeptides localized intracellularly in bacterial cytoplasm or sometimes in the periplasmic space. Aggregating characteristic of a protein or the dysfunction of the cellular machinery which under normal conditions is responsible for proper folding of the polypeptide contributes to the formation of inclusion bodies. Sometimes the inherent nature of a protein determines if it would be expressed in the soluble form or accumulates as inclusion bodies. In order to retain the functional 3D configuration, these misfolded polypeptides must be released from the inclusion bodies and refolded (Middelberg A.P.J, 2002).

There are different approaches for release and refolding of the target proteins from the inclusion bodies (details in Chapter 2). One of the most commonly used methods for refolding is by dialysis as it is less time consuming. Hence, we harvested the inclusion bodies from bacterial cell lysate (details in Section 2.2) and followed the Novagen protein refolding kit for solubilization and refolding of inclusion bodies. According to this kit, dialysis was the preferred method for refolding. The kit used CAPS buffer along with 0.3% N-lauroylsarcosine to achieve solubility of the inclusion bodies followed by dialysis with Tris buffer. The problem that we encountered was yield of the refolded protein was poor and it was observed that large amount of it grouped and precipitated during the refolding process. It seemed like dialysis was not efficient enough to reduce the detergent concentration from our target protein and thereby enhance refolding. We carried out turbidimetric assay which confirmed beyond permissible levels of detergent in the protein. Hence we observed large amount of aggregated protein contributing to lower yield.

The other popular method for refolding of protein is by dilution. One of the important factors that contribute to improving the yield of target protein from inclusion bodies is the purity. Presence of contaminants in inclusion bodies is likely to reduce the yield of the refolded protein. So we modified our approach and developed a new method for washing, solubilization of inclusion bodies and refolding of the target protein. We followed a stringent wash method to harvest pure inclusion bodies. Enzymatic digestion of DNA, repeated washing with detergent (Triton X-100) and chelating agent (EDTA) yielded acceptably clean inclusion bodies. For solubilization we preferred 6 M guanidine hydrochloride (GuHCl) over any other denaturants like 8 M urea because GuHCl is stronger and is thereby able to disintegrate otherwise robust inclusion bodies. Moreover, urea has a tendency to decompose to cyanate easily (Clark E.D.B et al, 1999).

The complexity of protein refolding lies in the fact that the exposed surface of the protein interacting with solvent should be hidden and other molecules of the protein should be able to interact with each other by formation of hydrogen bonds, salt bridges and van der Waals interactions (Pierce M et al, 1999). Successful *in-vitro* protein refolding is initiated by reduction of denaturant concentration. We followed the pulse renaturation or dilution method for protein refolding as this was more efficient in removal of the denaturant and thereby prevent loss of protein by aggregate formation. Addition of solubilized sample into large volume of the refolding buffer in pulses over a period of 36 hours at 4°C not only diluted the solubilized protein sample but also reduced the amount of denaturant/detergent used for solubilization. Hence the repetitive dilution of unfolded protein would minimise the tendency to form aggregates. Moreover, L-arginine in the refolding buffer acted as a suppressor of protein aggregation by enhancing solubility of partially structured folding

intermediates. Since the CRD of LMAN1 contains one disulfide bond, refolding in redox milieu facilitated associated disulfide bond formation. Hence oxido-shuffling system with reduced and oxidized glutathione (GSH/GSSG) was used for oxidative refolding of the CRD. We optimized the conditions and found that 150 folds dilution of the protein (120 μ g/ml) in refolding buffer gave the maximum yield.

3.2.2. Purification of refolded protein

It was likely that the refolded protein was contaminated with several other nonspecific bacterial proteins in the inclusion bodies, hence it was extremely important to purify it before using it for biochemical studies. To do this we used a gel filtration column equilibrated with HEPES buffer. Since our biochemical studies were mostly based in and around physiological pH levels this buffer seemed ideal for our purpose, as several articles suggested that HEPES works better for a wider pH range (6.6-9) compared to Tris which is good for pH 7.0 and above. Superdex-75 column was chosen for this purpose as it is efficient in separating proteins in the range of 3 KDa–70 KDa and the CRD has a molecular weight of 28.6 KDa. During protein refolding by pulse renaturation, as the protein was diluted 150 folds, it resulted in a large volume of refolding buffer. Hence for each batch of protein purification, the volume was reduced by concentrating it to 1 ml before loading the sample in a Superdex 75 column. The only downside in pulse dilution method was that it was time consuming and required a large supply of the concentration device. Details of the purification process are in Chapter 2.

A chromatogram was generated for each batch of protein purification by the Superdex column showing UV absorbance by each protein present in the refolded sample which is a mixture of target protein as well as nonspecific proteins. Proteins

were separated based on their molecular weight, with the largest protein being the first to exit the column.

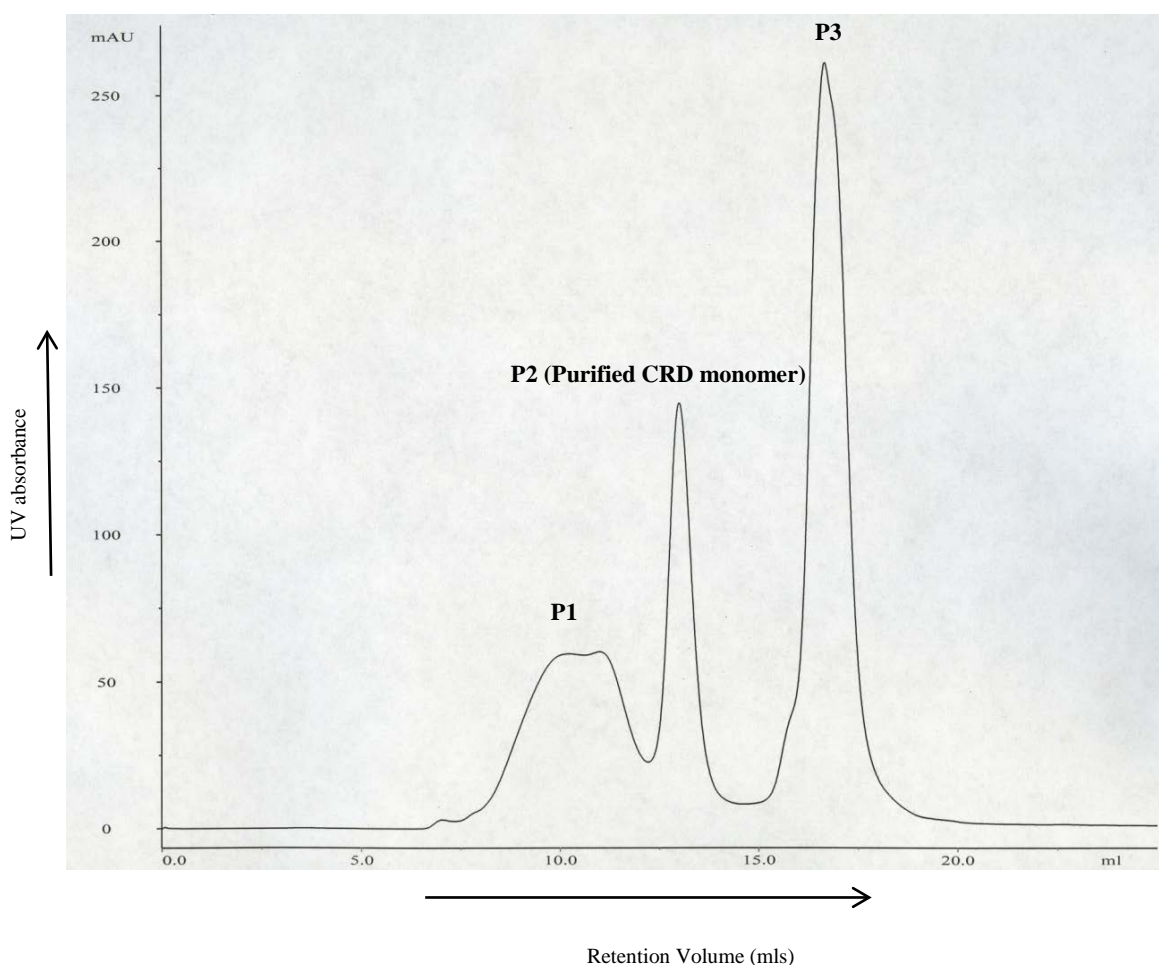


Figure 23. Chromatogram from Superdex 75 column used for purification of CRD of LMAN1. The refolded and concentrated protein sample is loaded on the Superdex 75 column and fractions' corresponding to each peak was collected.

In the chromatogram (figure23), there were 3 peaks. The x axis showed the volume of the buffer passing through the column and at what volume of the buffer, the proteins were exiting the column. The UV absorbance on the y axis increased with the increase in concentration of each protein. Each of the fractions corresponding to a peak was collected for analysis by SDS-PAGE which gave us an idea about the purity of the sample and also confirmed the correct molecular weight of the CRD. Before performing the SDS-PAGE, we carried out the size exclusion chromatography of

protein standard (figure24) which is a commercially available mixture of proteins of known molecular weight. This informed us about the retention volume of purified CRD monomer, so that for subsequent purifications we know which peak to collect.

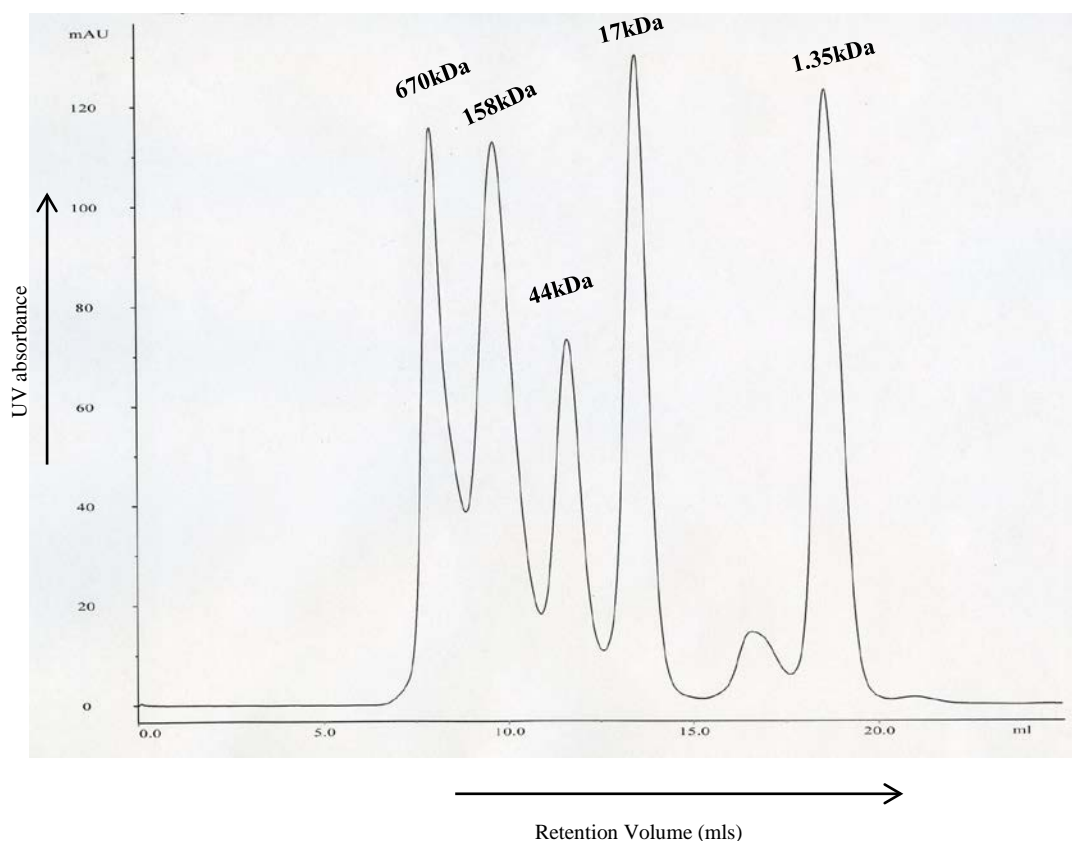


Figure 24. Chromatogram of Protein Standard. Comparing the chromatogram of protein standard to that of CRD purification, since M.W of CRD of LMAN1 is 28.6kD, it is likely to eluted in between 44kD (3rd peak) and 17kD (4th peak). This is the reason we labeled Peak 2 of Figure 22 as purified CRD monomer.

After comparing with the protein standard and SDS-PAGE of the collected fractions confirmed that peak 2 from the Superdex 75 column was indeed purified CRD monomer (figure25). Though SDS-PAGE results showed a considerably strong band of the CRD in Peak 1, the retention volume of the Peak 1 fraction as observed during chromatography was around 10 ml. Compared with the protein standard, peak 1 likely contained larger protein of MW greater than 44 KDa. It is possible that peak 1 consisted of mostly aggregates of CRD which denatured during reducing PAGE.

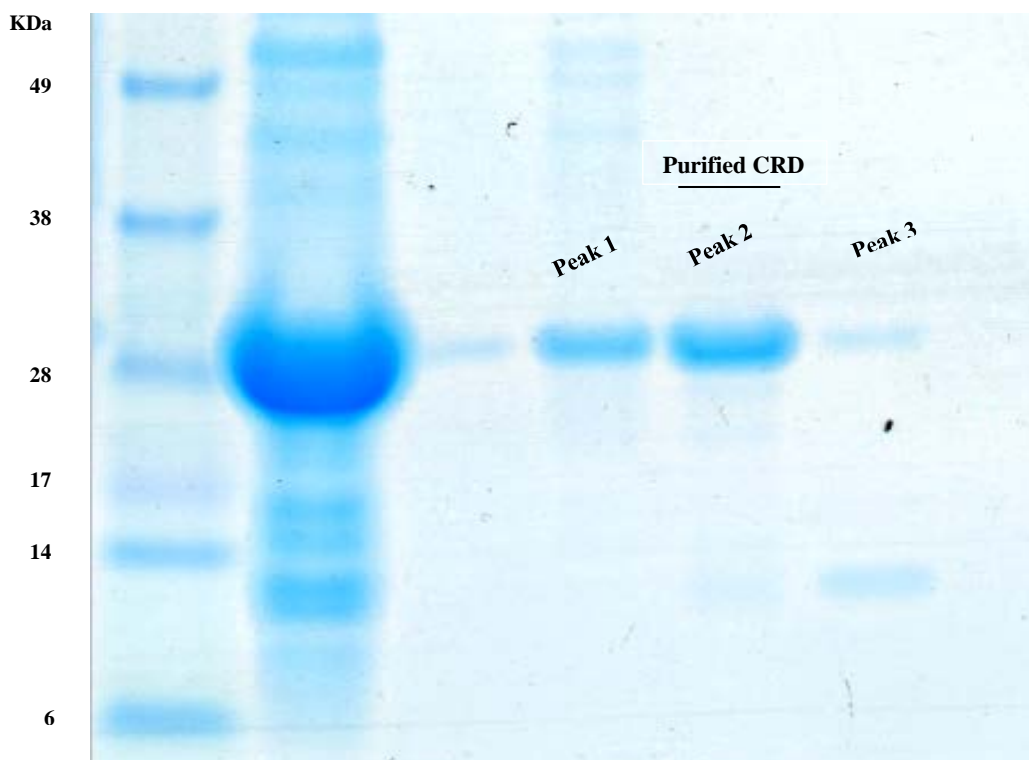


Figure 25. Coomassie Staining of SDS-PAGE (using MES buffer) of the sample fractions collected from gel filtration chromatography of the CRD of LMN1.

3.3 Analysis of the interaction of the CRD with Ca^{2+} and the ligand at different conditions by isothermal titration calorimetry (ITC)

The ITC assay is an efficient tool for quantitative determination of the thermodynamic aspects of any biochemical reaction. Whenever two molecules interact, either it is an endothermic reaction or an exothermic reaction. ITC assay is able to detect the heat released or absorbed during a reaction which is recorded over a time frame. For any ITC experiment, three factors are extremely critical- the ligand concentration, the macromolecule concentration and the choice of buffer. We accurately determined the concentration of the ligand and the purified CRD and optimized the buffer conditions before each experiment. Different parameters can be measured in a single ITC experiment like the stoichiometry of binding (n), the binding

constant or the association constant (K_a), entropy (ΔH_b), enthalpy (ΔS_b) (Pierce et al, 1999). However, for our purpose we determined the stoichiometry of binding (n) and the binding constant (K_a) which provided us the information on the molar ratio of interaction between the biomolecules and the strength of interaction respectively.

3.3.1 Titration of the purified CRD with Ca^{2+} by ITC assay

It was necessary to carry out a functionality test on the purified sample of the CRD to verify its correct folding before continuing with other biochemical experiments or utilizing it for solving the crystal structure of CRD of LMAN1 bound to mannose. It is known that the CRD of LMAN1 has 2 Ca^{2+} binding sites (Velloso L et al, 2003). If each batch of the purified samples of CRD monomer can bind to 2 Ca^{2+} ions, then it can be said that the sample is properly folded and functional. We used the ITC assay for this purpose. To determine the efficiency of the refolding process we titrated the purified CRD of LMAN1 (200 μ M) with 5 mM Ca^{2+} at a neutral pH of 7.4. In figure 26 (top panel), the peaks that we observed represent the heat change that took place when Ca^{2+} ion interacted with the CRD in the sample cell. During the initial phase of the experiment, the peaks were higher indicating that the heat liberated is more, it gradually diminished as more and more ligand (Ca^{2+}) was injected into the sample cell containing the CRD and the system progressively reached saturation. The raw data was then analysed by the Origin Software that came with the ITC device and fitted into an appropriate binding model to get the binding curve (Figure 26, bottom panel). The ITC result indicated that the stoichiometry of binding between the CRD and Ca^{2+} is 2 with a negligible error. Thus it can be inferred that our refolding method yielded a protein that is functional and regained its native confirmation. Not only did the ITC assay confirmed the presence of two binding sites on CRD for Ca^{2+} (Velloso L et al, 2003) it also provided additional information on the

binding affinity between the CRD and Ca^{2+} . Binding affinity (K_B) is defined by the dissociation equilibrium constant (K_D , unit is M) which is the inverse of association equilibrium constant (K_a , unit is M^{-1}). For tight binding complexes, K_a is in the range of 10^9 - 10^{10} M (Pierce M et al, 1999) and hence the corresponding K_D (dissociation equilibrium constant) is small. The smaller the K_D value, slower is the rate of dissociation between the interacting molecules and hence stronger is the binding affinity (K_B). In figure 26, we see that the K_a is $2.17\text{E}^4 \text{M}^{-1}$ and thereby in the order of 10^4M^{-1} . Hence $K_B = 1/ K_a = 2.17\text{E}^4 \text{M}$ or in the order of 10^4M indicating the intensity of binding is moderate. For each batch of purified protein we carried out the ITC assay as a quality control experiment before utilizing the sample for other biochemical procedures (figure 26).

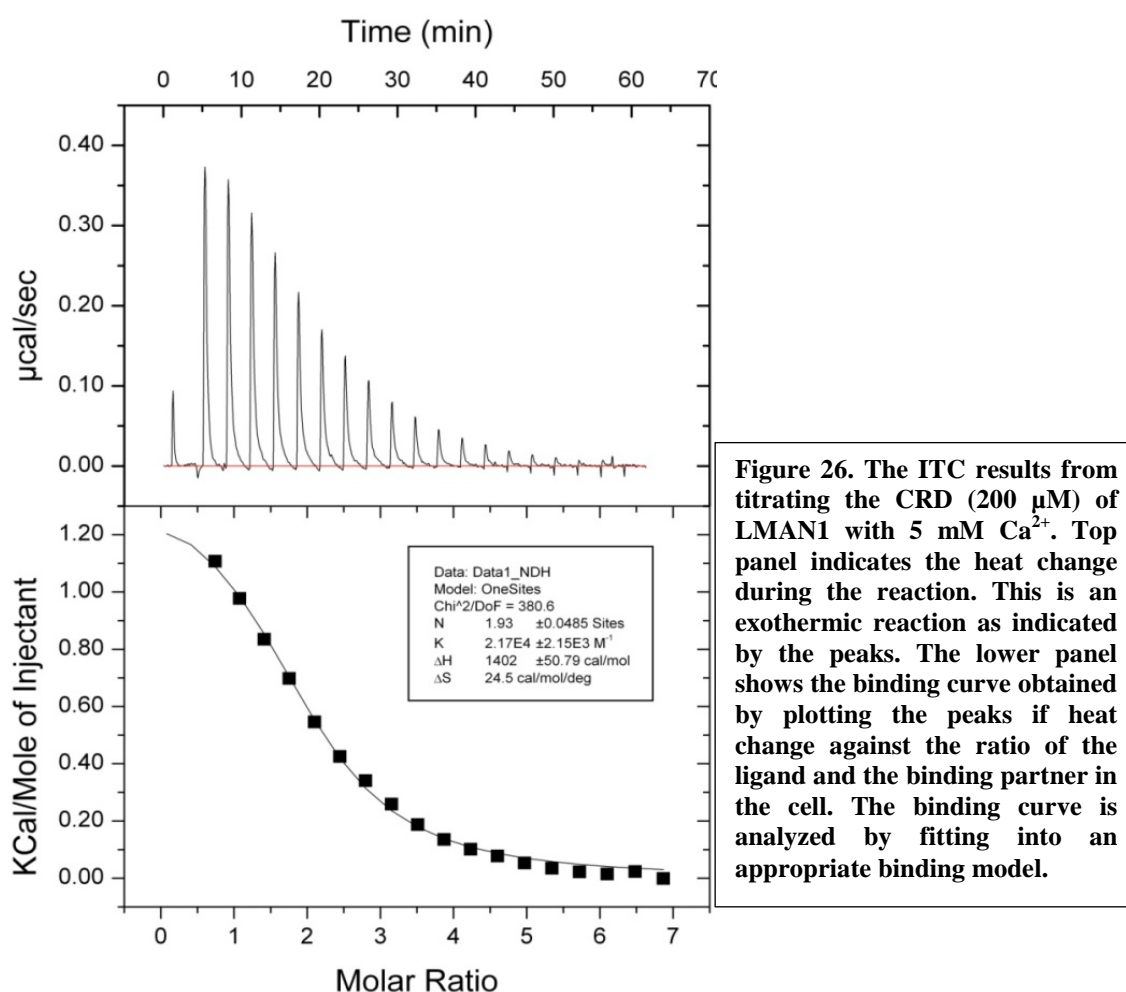


Figure 26. The ITC results from titrating the CRD (200 μM) of LMAN1 with 5 mM Ca^{2+} . Top panel indicates the heat change during the reaction. This is an exothermic reaction as indicated by the peaks. The lower panel shows the binding curve obtained by plotting the peaks if heat change against the ratio of the ligand and the binding partner in the cell. The binding curve is analyzed by fitting into an appropriate binding model.

3.3.2 Titration of the purified CRD with 2 α -Mannobiose by ITC assay

It was important to know whether the purified CRD of LMAN1 was able to bind to its ligand. Since LMAN1 binds to its cargo through the interaction of the CRD and the glycan (mannose) moiety of the cargo glycoprotein, we studied CRD-mannose interaction as a proxy for LMAN1-cargo interaction (figure 27). This was done by titrating the purified CRD (200 μ M) against 30 mM mannose- α (1-2)-mannose (a dimannose covalently linked by α (1-2) linkage). We observed it was an endothermic reaction. This experiment also provided us the information on how strongly the CRD binds to ligand (cargo) mannose- α (1-2)-mannose (commercial available as 2 α -mannobiose).

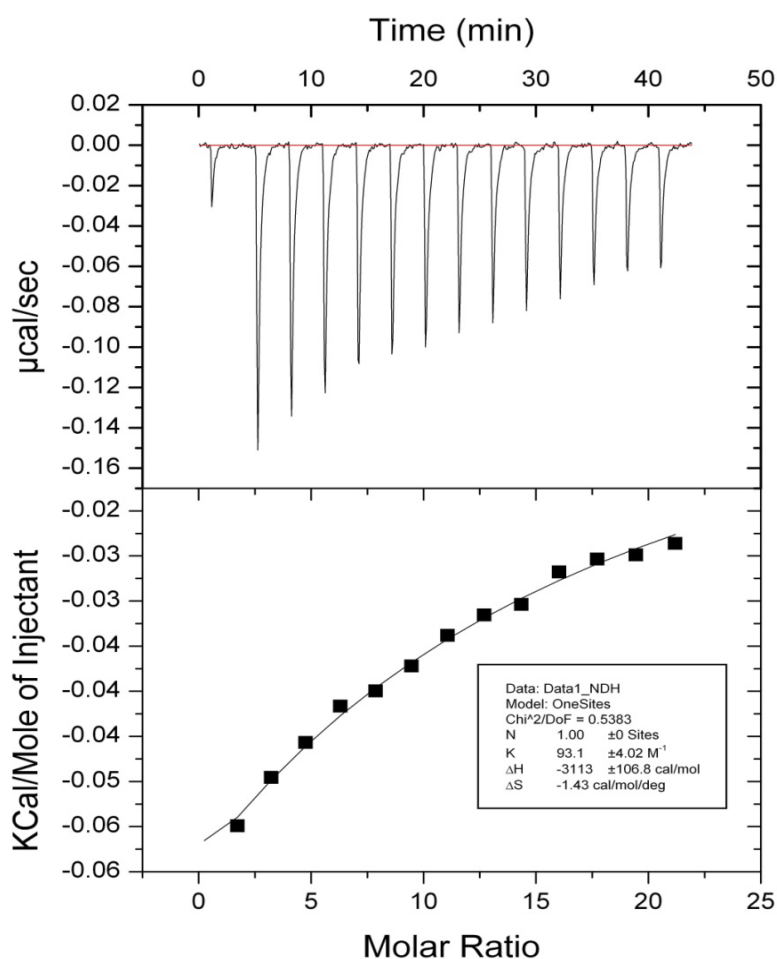


Figure 27. The ITC result from titrating the CRD of LMAN1 (in 5 mM Ca^{2+}) with 30 mM 2 α mannobiose. The lower panel shows the binding curve when the heat change is plotted against the ratio of 2 α mannobiose. The K_a is obtained by analyzing the binding curve by fitting into an appropriate binding model. $K_a = 93.1 \text{ M}^{-1}$.

As explained earlier, we can say from the ITC assay of the CRD titration with mannose- α (1-2)-mannose (figure 27), the binding affinity (K_B) between the CRD and its ligand dimannose is weak as the K_a is only 93.1 M^{-1} . The weak binding affinity between the CRD of LMAN1 and its ligand dimannose can be reasoned by the logic that tight binding between LMAN1 and its cargo might affect the release of cargo in the ERGIC.

In vivo LMAN1 always exists as hexamers which is difficult to simulate in vitro. The hexameric arrangement of the CRD of LMAN1 *in vivo* might lead to an additive effect on the binding affinity for the ligand. Moreover, VIP36 another member of the lectin family which exists as monomer and binds its ligand in the Golgi has weak affinity towards its ligand (high mannose glycans) but still plays an efficient role in post ER quality control. So it seems rational that the binding affinity between the CRD of LMAN1 and dimannose even though weak would be efficient enough for the transport of the cargo. The exact arrangement of hexameric LMAN1–MCFD2 and cargo is yet to be solved, which would provide a better understanding.

The reason why we used mannose α (1-2) mannose as the ligand for the CRD that we purified can be explained by figure 28. Here we can see that the terminal mannoses are linked by α (1-2) glycosidic bond and the α (1-3) linkage is located deeper inside the chain of glycan moiety. By frontal affinity chromatography, it was detected that VIP36, which belongs to the same family as LMAN1 has specificity for the D1 arm (figure28) (Sato T et al, 2007). So it is likely that the CRD of LMAN1 which shares homology with the CRD of VIP36 would display a similar behavior, specifically for the terminal mannoses linked by α (1-2) bond. Hence, 2α -mannobiose served as the ligand for our experiments on CRD. Moreover, the different di and tri

mannoses with α (1-3) or α (1-6) linkage are commercially not available and are difficult to synthesize in laboratory setup.

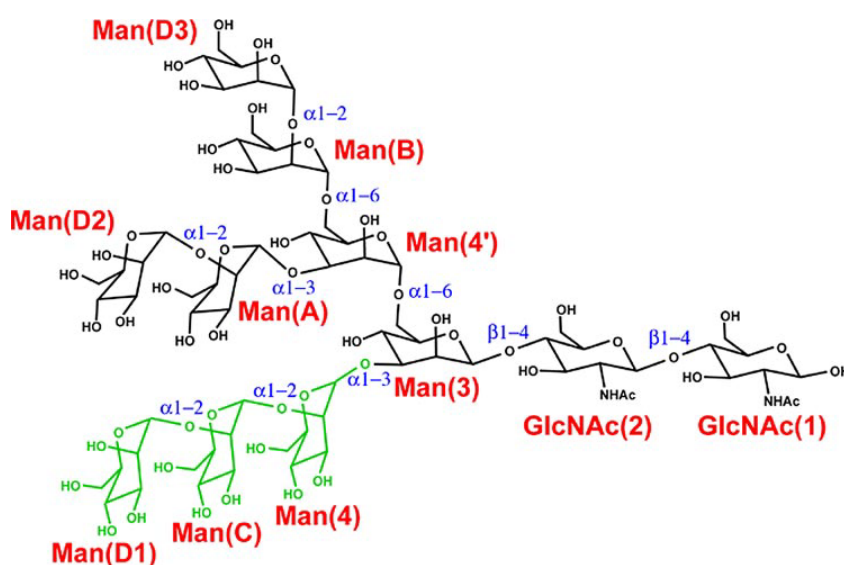


Figure 28. Chemical Structure of $\text{Man}_9(\text{GlcNAc})_2$. (Represented from Satoh T et al, 2007)

3.3.3 Titration of the purified CRD with 2α -mannobiose at different saturation level of the CRD with Ca^{2+} by ITC assay

We know that the lectin activity of LMAN1 is Ca^{2+} dependent. Hence any loss of calcium or decrease in interaction of the CRD with Ca^{2+} can initiate the release of the cargo. Since LMAN1 recycles between ERGIC and ER, it can be said that the release of cargo takes place in the ERGIC which is suggestive of the fact that ERGIC may have a lower Ca^{2+} concentration compared to ER. The Ca^{2+} concentration in the ER is 0.4 mM and 0.3 mM in Golgi but in ERGIC it remains undetected. It is possible that when the CRD of LMAN1 is less saturated with Ca^{2+} , it releases its cargo. To explore the effect of the change in saturation level of the CRD with Ca^{2+} (in the ER to trans Golgi network) on interaction with the ligand, we carried out two ITC assays which helped us to comprehend better the mechanism of cargo release by LMAN1 (figure 29 & 30).

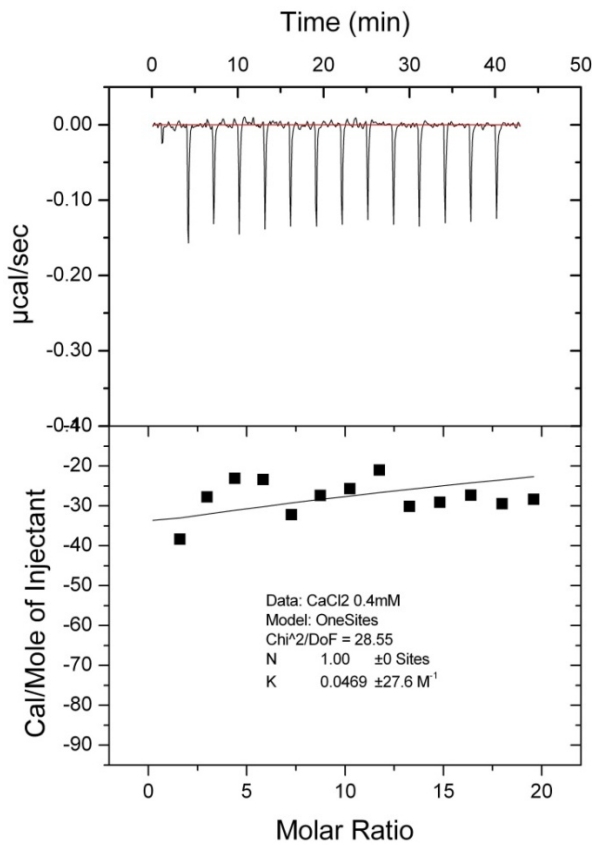


Figure 29. The ITC assay result when the CRD of LMAN1 (containing 0.4 mM Ca²⁺) is titrated with 30 mM dimannose in CRD at pH 7.4. The top panel shows practically there is no heat change and the lower panel indicates the binding curve did not have a good fit.

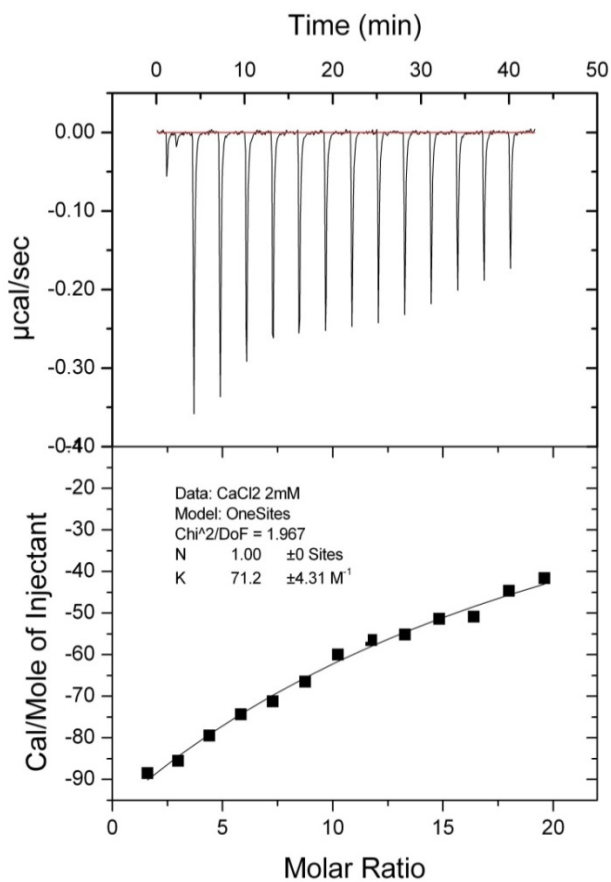


Figure 30. The ITC assay result when the CRD of LMAN1 (containing 2 mM Ca²⁺) is titrated with 30 mM dimannose at pH 7.4. The top panel shows significant heat change and the lower panel indicates the binding curve had a better fit.

Since the molar ratio in any ITC experiment is the ratio of the ligand to the macromolecule, from figure 26, it is evident that at molar ratio 2, the Ca^{2+} concentration interacting with the CRD is 0.4 mM and at molar ratio 7, the Ca^{2+} is 1.4 mM. When the molar ratio is 2, it can be said by comparing the peak heights that about 50% of the calcium binding sites of the CRD are saturated with the ligand whereas at molar ratio 7, the system has reached a steady state with no further heat exchange suggesting that there is almost 100% saturation of the calcium binding sites of the CRD. So for the ITC assays shown in figures 29 & 30, we used 0.4 mM and 2 mM Ca^{2+} (little more than 1.4 mM to attain 100% saturation) in the purified CRD sample to titrate it against 30 mM of dimannose. It is obvious from figure 29, that when 50% of the calcium binding sites of the CRD were occupied, the binding affinity between CRD and dimannose is extremely weak ($K_a = 0.0469 \text{ M}^{-1}$). Practically no binding took place between the two molecules. However, when 100% of the calcium binding sites of the CRD were occupied, the binding affinity between the CRD and dimannose increases ($K_a = 71.2 \text{ M}^{-1}$) (figure 30) and is similar to what we observed in figure 27. So from these two data it can be concluded that the CRD saturation with Ca^{2+} plays a significant role in the interaction between the CRD and its ligand dimannose. Lowering of Ca^{2+} can trigger the release of cargo by LMAN1 in the ERGIC.

Even though there is a Ca^{2+} gradient across the anterograde secretory pathway, the descent is likely to be subtle. There can be possibly other factors contributing to the cargo release mechanism.

3.3.4 Effect of pH on the Ca^{2+} binding affinity of the CRD

It is known that there is a progressive acidification in the organelles in anterograde pathway; ER has a neutral environment of pH 7.4 with lowering of pH as we proceed to trans Golgi network. To investigate the effect of dropping of pH on the calcium binding affinity of the CRD, we used the ITC assay to titrate the purified CRD with Ca^{2+} at three different pH levels – 7.4, 6.5 and 6 as a representation of the milieu of ER, cis-Golgi and trans Golgi network (figures 31,32 & 33 respectively). It was earlier hypothesized by Hauri et al (2004) that there is pH induced Ca^{2+} loss which initiates the release of cargo by the CRD of LMAN1 but there existed no direct evidence. From our ITC assay results (figures 31, 32 & 33) we observed that with the lowering of pH, the binding affinity (K_B) of the CRD to Ca^{2+} undergoes slight but not significant change suggesting that there would not be any major Ca^{2+} loss. Hence the earlier model stating the pH induced Ca^{2+} loss being the reason for cargo release was not accurate.

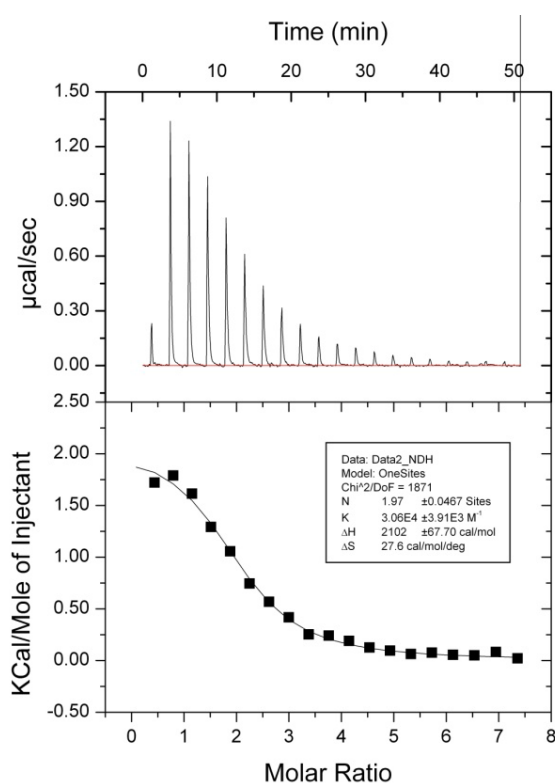


Figure 31. The ITC result from titration of the CRD monomer with 2 mM Ca^{2+} ion at nearly neutral pH 7.4 simulating the local microenvironment of ER. Binding constant (K_a) is $3.06 \times 10^4 \text{ M}^{-1}$.

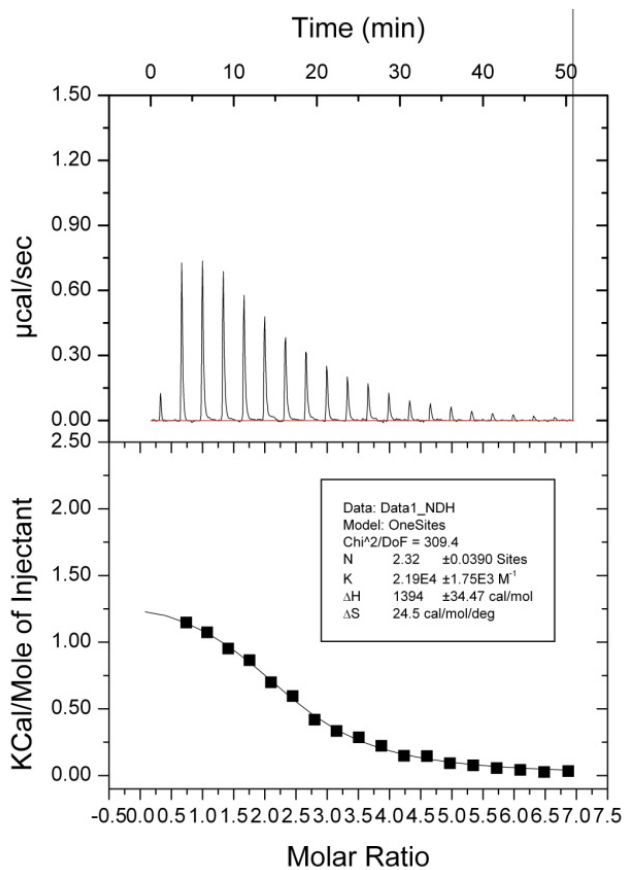


Figure 32. The ITC result from titrating the CRD monomer with 2 mM Ca²⁺ ion at slightly acidic environment of pH 6.5. This mimics the environment at ERGIC *in vitro*. K_a is 2.19*10⁴ M⁻¹.

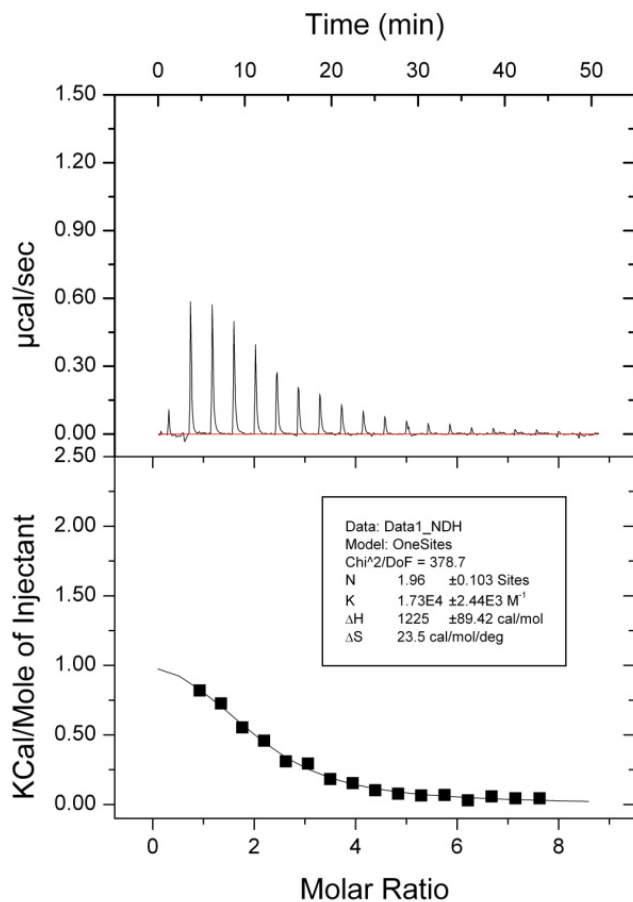


Figure 33. At acidic pH of 6.0 which reproduces the environment at Golgi, when the CRD monomer is titrated with 2 mM Ca²⁺, the K_a is even lower, it is 1.73*10⁴ M⁻¹.

We also checked the effect of pH on the binding affinity of the CRD to dimannose by the ITC assay (data not shown). It was observed that lowering in pH has no substantial effect on the binding affinity of the CRD towards its ligand. Though from these ITC assays it can be precisely inferred that Ca^{2+} concentration gradient across organelles in the anterograde pathway is the sole factor contributing to the release of cargo by the CRD of LMAN1 but it is certainly a crucial factor in the release mechanism; more decisive evidence would be to check if the LMAN1-MCFD2 complex is impervious to the Ca^{2+} saturation gradient of the CRD of LMAN1 that we used for our studies as oppose to the CRD- dimannose complex. LMAN1-MCFD2 should exist as a complex even at low concentration of Ca^{2+} (implying lower saturation level of the CRD with Ca^{2+}) which otherwise triggered the release of cargo.

3.4 Crystal structure of the CRD of LMAN1 bound to mannose α (1-2) mannose [commercially 2 α mannobiose]

With the help of molecular docking, our lab came up with an in silico model of the CRD of LMAN1 (data not shown) which predicted the presence of two potential mannose binding sites on the surface of the CRD. The molecular modeling technique also suggested some conceivable amino acid residues (R90 along with previously hypothesized N156 and D121 in the first site and S189 and Q191 in the second site) involved in interacting with the cargo and those residues (R90, S189, Q191) were located far from the Ca^{2+} binding site. According to the in silico model, mutation of these potentially critical amino acids would not disrupt the Ca^{2+} binding site unlike N156/D121, the two possibly vital amino acids involved in interaction with the glycoproteins and lying in close proximity to the Ca^{2+} binding site (Itin C et al, 1995). Though the crystal structure of the CRD of LMAN1 was solved earlier

(Velloso et al, 2003) it had no ligand bound to it. So it was absolutely necessary for us to solve the crystal structure of the CRD of LMAN1 bound to its ligand dimannose that would help us to elucidate the number of cargo binding pocket(s) of the CRD and the key amino acid residues involved in binding to the ligand.

In collaboration with Dr. Saurav Misra's laboratory at Lerner Research Institute, Cleveland Clinic, the crystal structure of CRD bound to mannose- $\alpha(1-2)$ -mannose was solved. The functional CRD of LMAN1 that we purified was used for the X-ray crystallography. The resolution of the structure was 2.42 Å and it revealed that the ligand 2 α manno- $\alpha(1-2)$ mannobiose was localized on one site involving amino acids S-88, H-178, F154, G 251/252, N156 and D121 (figure34) and it does not involve the amino acid residues (R90, S189 & Q191) predicted by the molecular modeling study. As observed in figure 34, the crystal structure of the CRD of LMAN1 bound to mannose $\alpha(1-2)$ mannose (dimannose) also provided us information about the orientation of the ligand as it fits into the ligand binding pocket. From the optical density of the crystals, it seemed logical that the mannose residue at n-1 position fits into the sugar binding pocket while the terminal mannose residue (n) lies outside of the pocket. Hence our crystal structure of the CRD of LMAN1 bound to ligand mannose $\alpha(1-2)$ mannose clears the ambiguity about the number of ligand binding pocket on the surface of CRD and shows the important amino acids involved in binding to the ligand. We also tried to crystallize the CRD with mannose $\alpha(1-3)$ mannose and mannose $\alpha(1-6)$ mannose but those crystals did not diffract any x rays.

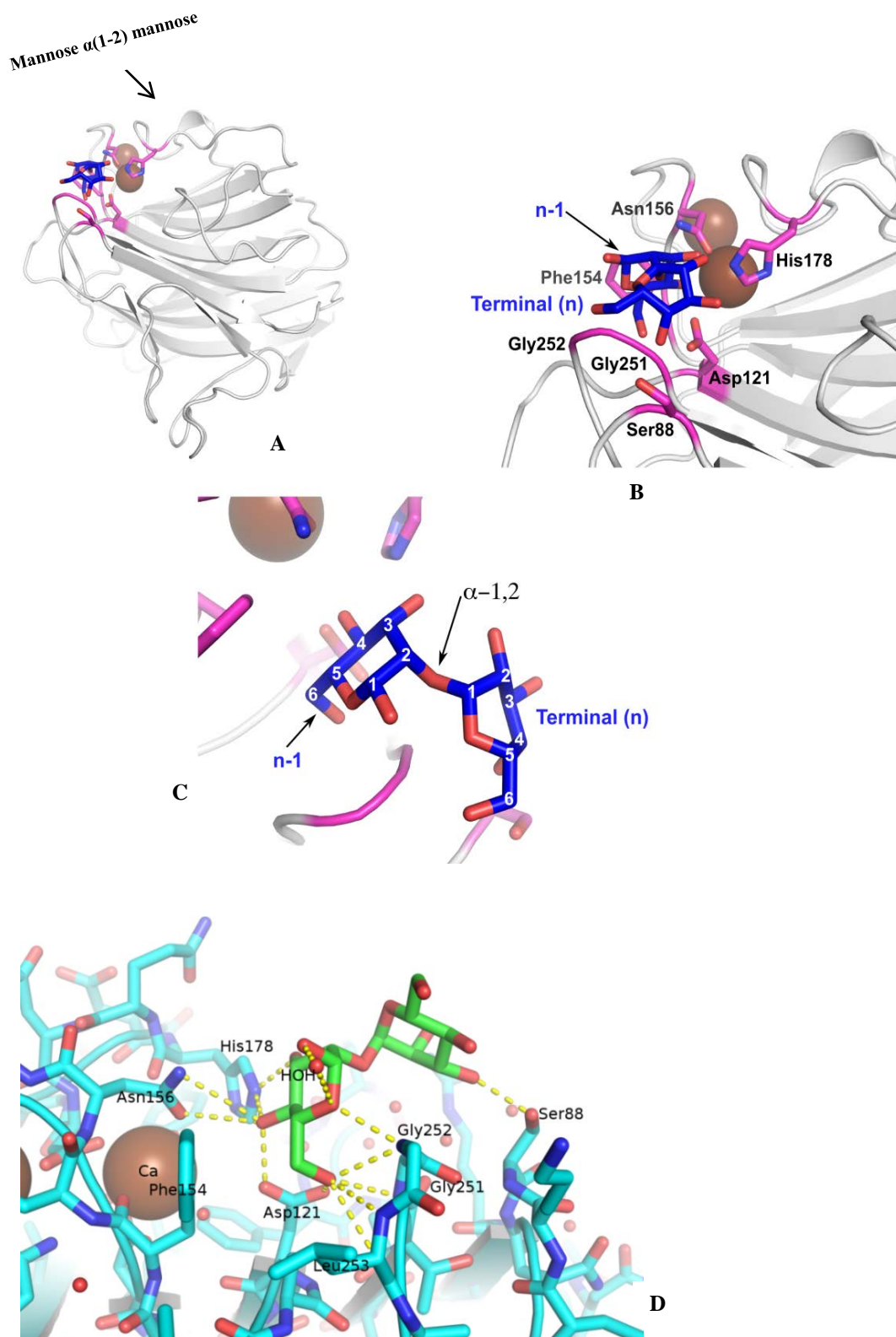


Figure 34. (A) Crystal Structure of CRD of LMN1 bound to ligand Mannose α (1-2) mannose. (B) The key amino acids involved on the surface of CRD involved in binding with the ligand. The terminal mannose of the ligand is “n” and the mannose joined to the terminal one by α (1-2) glycosidic bond is the “n-1” mannose. (C) Enlarged photo of the ligand showing α 1-2 linkage between the “n” & “n-1” mannose. The carbon molecules are numbered 1-6. (D) Zoomed image showing the interactions between the key amino acid residues and the ligand (shown in green and red). S88 interacts with the terminal mannose (n) whereas other residues (D121, F154, N156, H178 and G251/252) interact with (n-1) that lies within the pocket. Calcium ions are shown in brown, ligand shown in blue and red in A, B and C.

3.4.1 Comparison of the different crystal structures of the CRD of LMAN1

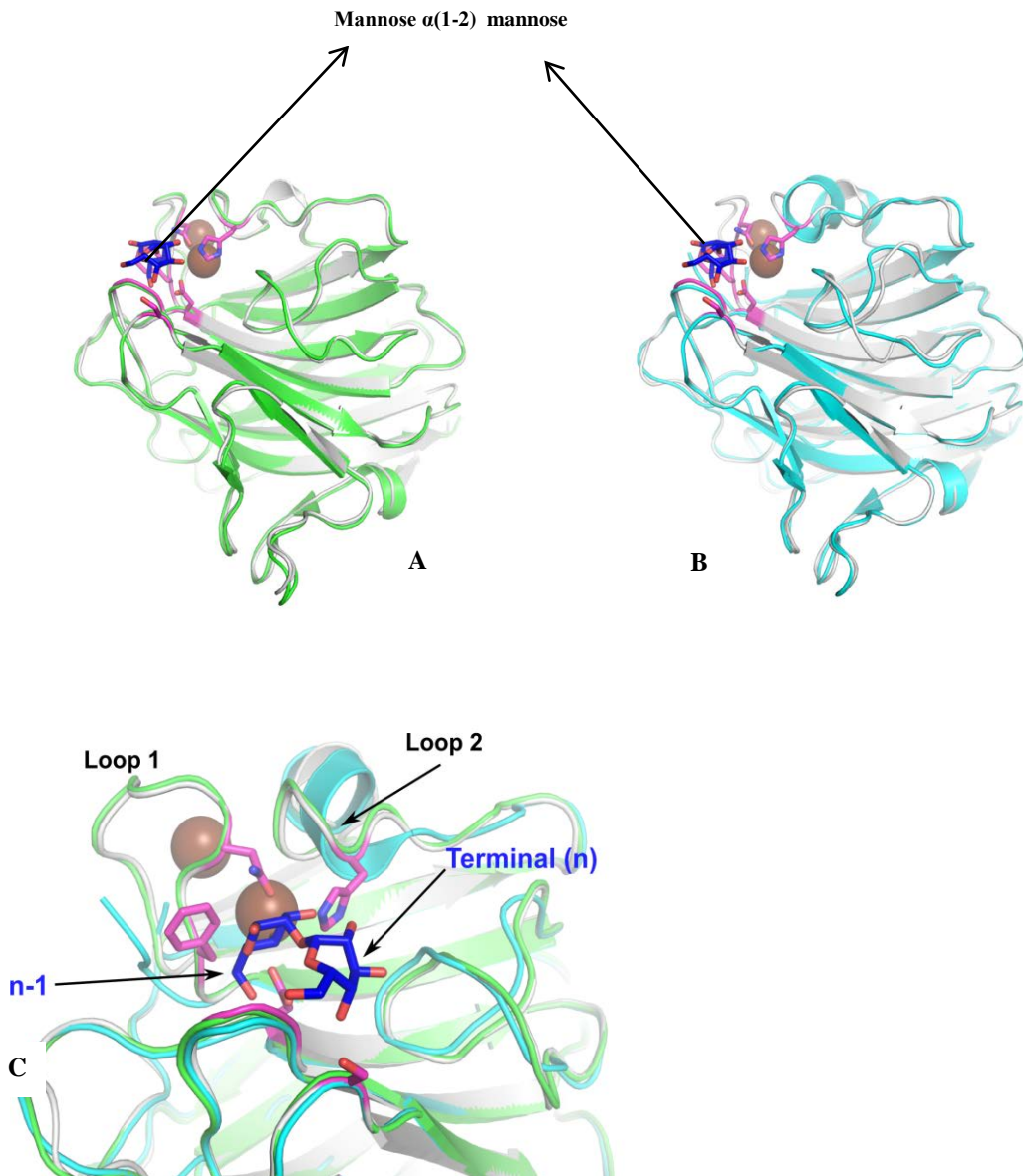


Figure 35. A. Overlaying the ribbon model of the CRD of LMAN1 bound to Ca^{2+} (Velloso et al, 2003; shown in green) on the CRD of LMAN1 bound to ligand mannose $\alpha(1-2)$ mannose and Ca^{2+} (shown in grey). B. Overlay of CRD of LMAN in absence of Ca^{2+} (Velloso et al, 2002; shown in cyan) on the CRD of LMAN1 bound to ligand mannose $\alpha(1-2)$ mannose in presence of Ca^{2+} (shown in grey). C Superposition of the three crystal structures showing no dramatic conformational changes in the structures solved in presence of calcium except in Loop 1 and Loop 2 in the structure in absence of calcium. Calcium ions are shown in brown and the ligand shown in blue sticks.

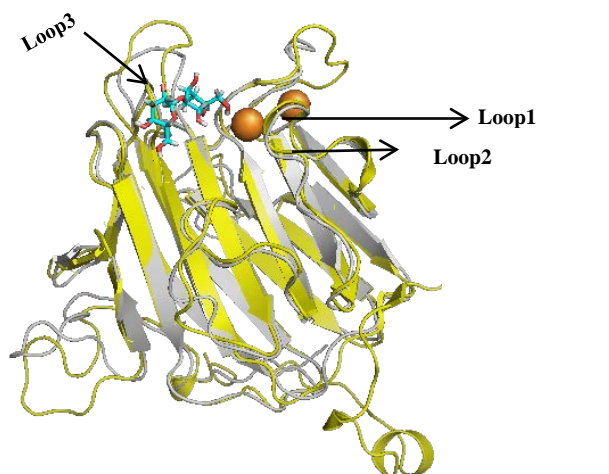
If we compared the ligand bound state of the CRD of LMAN1 to that of unbound state (figure 35A), apparently there was no dramatic conformational changes involved as both the structures were solved in presence of calcium. The structures

were superposed on one another using Pymole software; it seemed like one structure was aligned to the other quite well. However, the structure of the CRD bound to ligand when overlaid on the CRD of LMAN1 solved in absence of Calcium (figure35B), we saw that the loop 1 and loop 2 were not completely aligned (figure35C). Loop 1 remained unstructured in absence of calcium and loop 2 unfolds in presence of calcium. Except these changes there were no other significant variations in the crystal structures of CRD of LMAN1 in ligand bound and free state as well as calcium bound and free state.

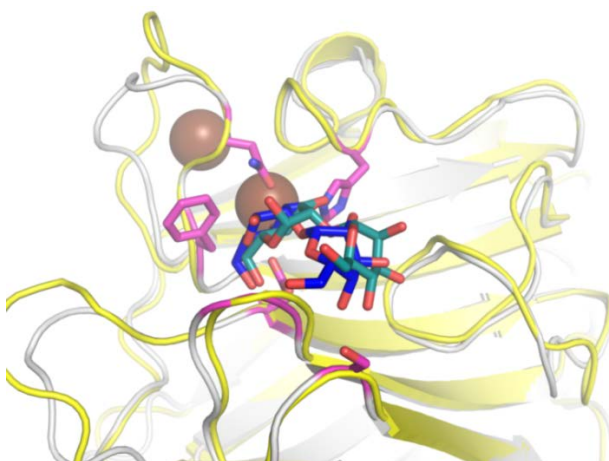
3.4.1.1 Comparison of the CRD of LMAN1 with the CRD of VIP36

VIP 36 (vesicular-integral protein) also known as LMAN2 is of 36 KDa and belonging to the family of leguminous L type lectin. As both the proteins are members of the same family, the carbohydrate recognition domain (CRD) of VIP36 shares 46% homology with the CRD of LMAN1. Despite the similarity in the core structure, there are some significant functional and structural differences. VIP36 is glycosylated whereas LMAN1 is nonglycosylated (Kamiya Y et al, 2008). The CRD of VIP 36 binds to one Ca^{2+} rather than two like the CRD of LMAN1 and there is no evidence on oligomerization of VIP36 whereas LMAN1 exists as homo hexamers. VIP36 lacks the helical domain of LMAN1. The crystal structure of the CRD in Ca^{2+} bound state of VIP36 was initially solved using the Ca^{2+} bound CRD of LMAN1 as a model. But later on, structure of the ligand bound form of the CRD of VIP36 in presence of Ca^{2+} was revealed by Satoh T et al in 2007. Our structure of the CRD of LMAN1 bound to its ligand when superimposed on the CRD of VIP36 bound to mannose using Pymole software we observed no major deviations (figure33D); though some minor structural differences were observed in loop 1(loop 1 of VIP36 is longer by two amino acids) and loop 3 of VIP36 and LMAN1 and side chain

orientation of few amino acid residues are different between the two lectins (figure36). The amino acid residues of LMAN1 involved in binding to ligand are somewhat similar to that of VIP36 (Sato T et al, 2007). Even the orientation of the ligand was same in both the proteins. LMAN1 acts as a cargo transporter binding its cargo in the ER at a neutral pH of 7.4 and releasing the cargo in ERGIC but the optimum pH required for ligand binding by VIP 36 is 6.5. VIP36 binds its cargo in the Golgi, dependent of Ca^{2+} and evidences indicate the main function of VIP 36 is post ER quality control, recognizing and recycling incompletely folded, untrimmed proteins back to ER. In spite of the resemblance that we observed from the superposed model (Figure 36), the functional differences of the two proteins are noteworthy.



A



B

Figure 36. A. Shows the overlay of ribbon model of CRD of VIP36 on CRD of LMAN1. B. Magnified view of the ligand binding site in the ribbon model comparing the crystal structure of CRD of VIP36 (shown in yellow) bound to ligand mannose α 1-2 mannose (shown in teal) and CRD of LMAN1 (shown in grey) bound to mannose α (1-2) mannose (shown in teal blue). Both the structures solved in presence of Calcium (brown sphere) and aligned well with each other without any major variation.

3.5 Mannose binding assay

The crystal structure of the CRD with mannose α (1-2) mannose revealed the amino acids that are involved in interaction with the ligand. But it was necessary to determine the most critical amino acids from this group whose mutation would lead to abolishment of mannose binding by the CRD of LMAN1. So we carried out mannose binding assay with the mutants and the wild type plasmid. The mutants used for this purpose were S88A, F154Y, H178A and G251A/252A.

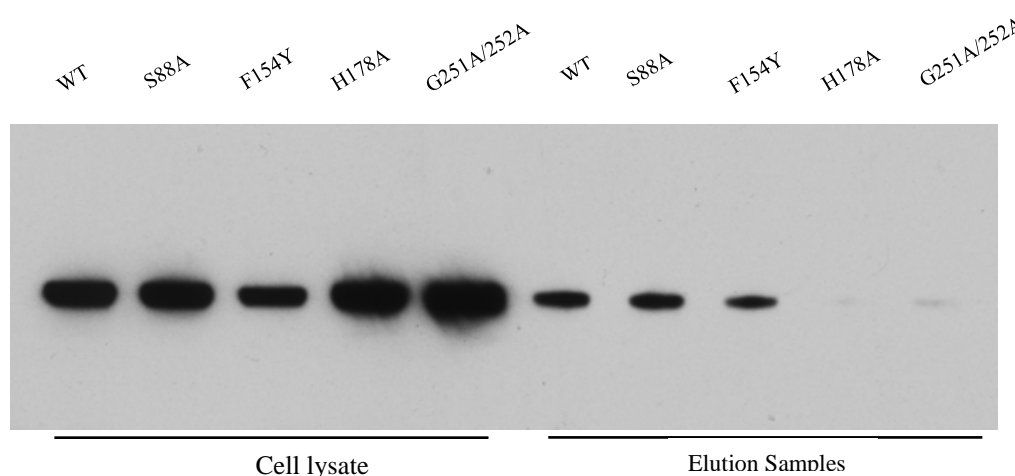


Figure 37. Result from the western blot with anti-flag antibody performed after mannose binding assay of the mutants and the wild type (WT) plasmids. The first six lanes are the samples before dialysis which served as the loading controls to check the level of protein expression by each plasmid when transfected into COS1 cells.

From Figure 37, it is evident that H178A mutant and G251A/252A double mutant has the most prominent effect of mannose binding. The intensity of the band was extremely weak for these mutants. So it can be inferred that these two amino acids on the surface of the CRD are most critical for binding to mannose and mutation in these residues leads to abolishment of the ability to interact with cargo glycoproteins.

It has earlier been observed that mutation in H-178 obliterated mannose binding (Hauri et al, 2004) but then it was not known if H-178 interacted with the

ligand. Our crystal structure showed that H-178 has direct interaction with the ligand. Moreover it was earlier found that H178 was the only basic amino acid on the surface of CRD. So it was hypothesized that this residue might act as a “proton sensor” detecting change in pH levels as the CRD of LMAN1 translocate the cargo proteins from ER to ERGIC and thereby trigger the release of cargo in the ERGIC (Hauri H-P et al, 2004). However, our in-vitro studies showed that mannose binding of the CRD is not dependent on pH levels. H-178 and G251/252 are the amino acids interacting with the mannose residue at “n-1” position and mutation in them disrupts the sugar binding site. F154 is another amino acid interacting with this residue but perhaps mutation in F-154 may not lead to obliteration of enough hydrogen bonds with the mannose and hence it does not lead to abolishment of sugar binding.

3.6 Purification of MCFD2

Cellular MCFD2 is always bound to LMAN1 in a 1:1 stoichiometry (Zhang B et al, 2005). The precise role of MCFD2 is unclear; it may affect the sugar binding ability of LMAN1 or play other important roles in the cargo capture and release process. Because it is essential for LMAN1-MCFD2 complex formation, we would like to check the integrity of LMAN1-MCFD2 complex at different conditions mimicking the milieu of ER, ERGIC and Golgi.

For biochemical experiments, we need purified MCFD2. By using the principle of IMAC (immobilized metal affinity chromatography), we purified MCFD2 by passing the crude cell lysate through the Ni-NTA resin packed in a column, followed by overnight digestion with thrombin to remove the His tag. Finally non-specific proteins were removed by Superdex 75 column (Figure 38) [details in Chapter 2]. Peak 1(P1) of Figure 38 is the peak of MCFD2. It is eluted at a position

equivalent to a dimer as reported previously (Kawasaki N et al, 2008). Each of the fractions collected by chromatographic purification was electrophoresed on a SDS gel (Figure 39). From the coomassie stained gel picture (Figure 39), it is evident that P1 is MCFD2 (MW 16KDa), running close to MW ~ 17KDa of the protein ladder. With the purified MCFD2 we would try to check if the binding ability of LMAN1 towards its ligand increases in the presence of MCFD2.

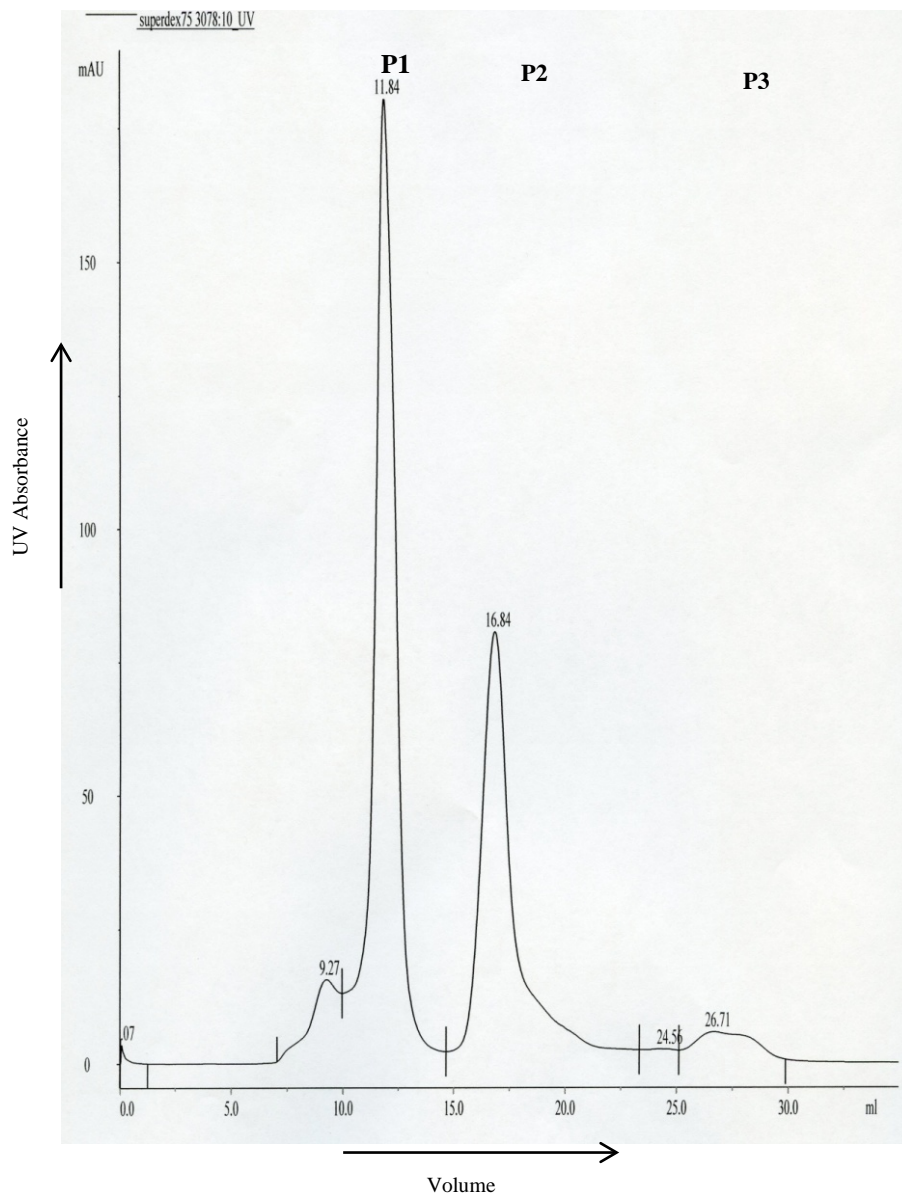
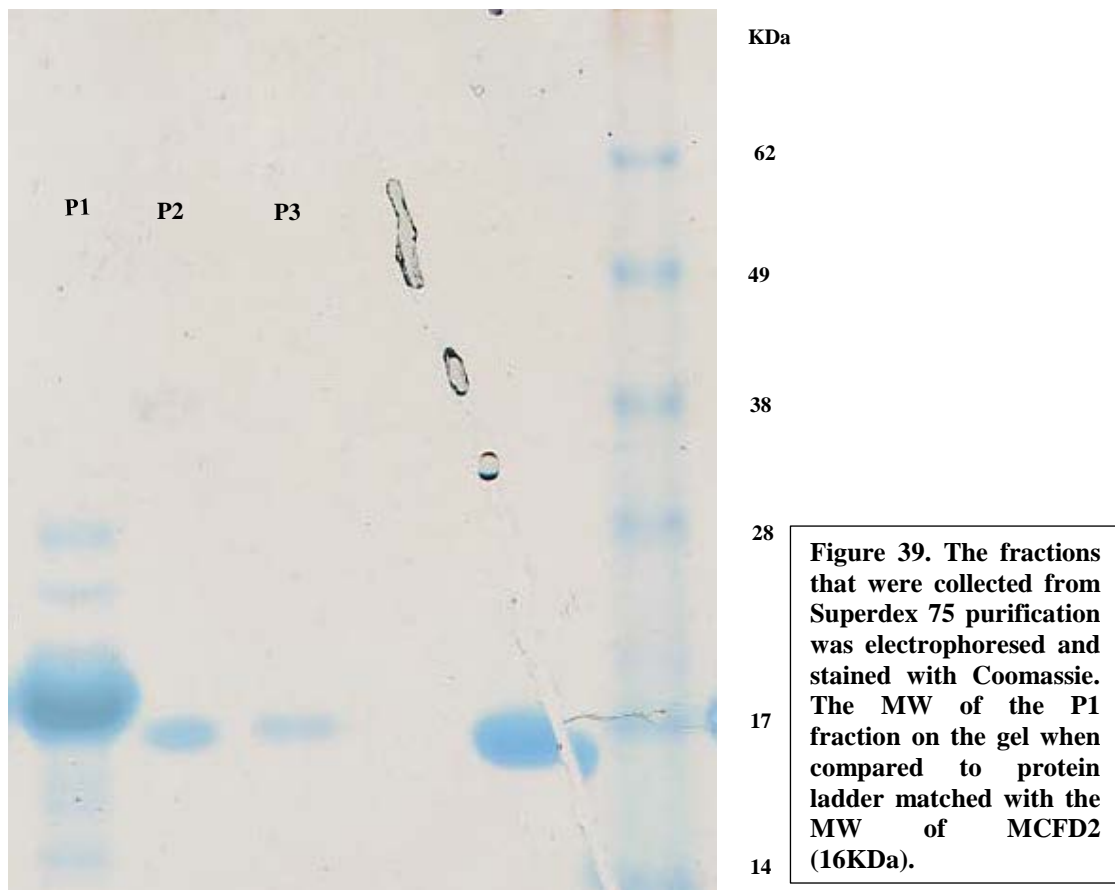


Figure 38. Chromatogram from Superdex 75 gel filtration chromatography of MCFD2 sample which was eluted from Ni-NTA column and digested with Thrombin overnight at 4°C. P1 fraction is the target protein MCFD2 after we compared the retention time and volume with that of protein standard.



It has been demonstrated earlier that the soluble form of LMAN1 bound modified surface glycans of HeLaS3 cells only in the presence of wild type MCFD2 and Ca^{2+} (Kawasaki N et al, 2008). However, X-ray crystallography revealed that MCFD2 binding site on LMAN1 is distant from the sugar binding site and the LMAN1-MCFD2 complex formation is impertinent to sugar binding (Nishio et al, 2010). To clarify the role, if any, of MCFD2 on the ligand binding affinity of LMAN1 we would like to compare the LMAN1 sugar binding ability in the presence or absence of MCFD2 by the ITC assay under physiological conditions. Being a soluble cofactor of the cargo receptor complex, MCFD2 is likely involved in other roles in the capture and release process of the cargo proteins. Biochemical experiments would help us to elucidate its functions.

CHAPTER IV

CONCLUSIONS AND FUTURE DIRECTIONS

4.1 Conclusions

Using our developed protocol of denaturation and refolding, we were able to purify the CRD monomer of LMAN1 to study the structural basis for LMAN1 cargo capture in the ER and release in the ERGIC. Using the ITC assay, we demonstrated that the purified CRD was able to bind two Ca^{2+} ions, which is consistent with earlier reports that the CRD of LMAN1 has 2 Ca^{2+} binding sites. This indicated that the purified protein retained its functionality. This ITC assay also revealed the binding affinity of the CRD with Ca^{2+} , which is of moderate intensity ($K_B=10^4$ M). Thus our denaturation refolding strategy worked well for our target protein. The purified and functional CRD was used for X-ray crystallography to obtain the ligand bound structure of the CRD. The X-ray crystal structure of the CRD that was solved in 2003 had no ligand bound to it. Uncertainties existed on the specific location, number of the binding sites and the key amino acids involved in ligand binding. The ligand for LMAN1 and its correct orientation facilitating a proper fit into the receptacle on the CRD was ambiguous. Based on information from published results, physiological

rationale and easy availability, we decided to use mannose α (1-2) mannose as the ligand for LMAN1 in our in vitro studies. Mannose α (1-2) mannose mimics the terminal structure of glycan moiety where the mannose at the “n” position is linked to mannose at “n-1” position by a glycosidic bond. From our solved crystal structure of the CRD with its ligand dimannose, it appeared that the ligand mannose α (1-2) mannose fits properly into the ligand binding pocket. The crystal structure also helped us to know the correct orientation of the ligand, the mannose at the “n-1” position lied within the pocket while the mannose at “n” position branched out. We were also able to decipher the key amino acids involved in interacting with the dimannose on the side chains of the cargo molecule. The key amino acids involved include S88, D121, N156, F154, H178 and G251/252. From the crystal structure it was evident that there is one mannose bound to the surface of the CRD of LMAN1. The amino acids involved in networking with the carbohydrate moiety were tested further by mutagenesis studies and in vivo experiments to confirm their significance in interaction with the sugar residues. Mannose binding assay was performed with the wild type LMAN1 and mutants - S88A, F154Y, H178A and G251A/252A. It was observed that H178 and G251/252 are most critical for mannose binding and mutations in these residues led to complete abolishment of the binding ability of the CRD to its ligand.

Apart from solving the crystallographic structure of the CRD bound to dimannose, we were also successful in determining the binding affinity of the CRD to its ligand. From our experiments, the binding between the CRD and its ligand mannose α (1-2) mannose is weak, which might be required to facilitate the easy release of cargo at the ERGIC. To elucidate the mechanism of cargo release by LMAN1, we checked the effect of pH on the binding affinity of the CRD to Ca^{2+} ions

by the ITC assay. It was observed that the binding affinity (K_B) between the CRD of LMAN1 and Ca^{2+} did not decrease significantly with the decrease in pH. Also the lowering of pH had no effect on the binding affinity between the CRD and the dimannose. However, we noted that the saturation level of the CRD with Ca^{2+} played a significant role in the binding affinity of the CRD with its ligand. When we titrated the CRD with dimannose at different Ca^{2+} concentrations, we observed that at a lower saturation level of the CRD (Ca^{2+} at 0.4 mM, 50% sites of the CRD saturated with Ca^{2+}); there is practically no binding between the CRD and its ligand. When the Ca^{2+} concentration in the CRD is increased to 2 mM (100% of the sites of the CRD saturated with Ca^{2+}), the binding affinity of the CRD for dimannose increased significantly.

We know that the lectin activity of LMAN1 is Ca^{2+} dependent and there is a progressive acidification of the organelles from the ER to the Golgi. LMAN1 binds to its cargo FV/FVIII in the ER, which is at relatively neutral pH of 7.4. The cis and trans Golgi network have a pH range from 6.7-6.0. The Ca^{2+} level in the ER is 0.4 mM and in the Golgi it is 0.3 mM. The pH and Ca^{2+} levels of the ERGIC where the cargo release takes place is not known, although an earlier study reported an undetectable level of Ca^{2+} . Lower Ca^{2+} concentration in ERGIC decrease the affinity of LMAN1 with cargo and this trigger the release of cargo. Our results support a model in which the relative saturation level of the CRD with Ca^{2+} plays the crucial role in the cargo release by LMAN1.

From our results, it is evident that Ca^{2+} concentration affects only the interaction between LMAN1 and the cargo. The interaction between LMAN1 and MCFD2 is much stronger (K_a in the range of 10^7 - 10^8 M^{-1}) and is less sensitive to

changes in pH and Ca^{2+} concentration ensuring the integrity of the LMAN1-MCFD2 receptor complex during the cargo release process (Kawasaki et al, 2008).

4.2 Future Directions

The crystal structure predicts that both the H-178A and G251A/252A abolish mannose binding without altering Ca^{2+} binding. To confirm that these mutants retain full Ca^{2+} binding capacity, we will purify these two mutant proteins and test their Ca^{2+} binding by the ITC assay. To do this we need to create new mutant constructs in the backbone of pET15b vector and express the corresponding protein products. These mutant proteins will be then purified by Superdex 75 column and titrated against Ca^{2+} in an ITC experiment. If the mutants show no significant change in the binding affinity for Ca^{2+} as compared to the WT CRD construct (the protein that we used for our ITC experiments), then it can be concluded that point mutations at H-178 and G251/252 disrupt the mannose binding ability of the CRD only, without disrupting the Ca^{2+} binding site. These mutant proteins can also be tested by the ITC assay to confirm their inability in binding mannose. We do not expect to observe any interaction when the mutant proteins would be titrated with dimannose.

To support our proposed model, that only Ca^{2+} concentration gradient is crucial factor contributing to the release of cargo by LMAN1, we would titrate the purified CRD against purified MCFD2 at different saturation levels of the CRD with calcium to ensure the stability of the LMAN1-MCFD2 complex across different levels of calcium. We expect to observe a strong affinity between the CRD and

MCFD2 or no significant change in K_a between the CRD and MCFD2 at the same saturation levels of the CRD with Ca^{2+} which affected the K_a between the CRD and dimannose ; thereby helping us to conclude that it is only the binding affinity between the LMAN1 and the ligand which gets affected by change in calcium concentration while the LMAN1-MCFD2 complex remains insensitive to this change.

In the in vivo and in vitro mannose binding assays we examined the effects of LMAN1 mutations on their binding to the dimannose, equivalent to the terminal mannoses of the glycan moiety attached to glycoproteins. As stated above, these mutations are predicted to affect mannose binding without disturbing the calcium binding. FV and FVIII are the known glycoprotein cargo of LMAN1 responsible for the pathology of F5F8D. Next we would like to check whether mutations in LMAN1 that abolish mannose binding will also abolish FV/FVIII binding. Failure of mannose-binding mutants to interact with FV/FVIII would indicate that the role of calcium in LMAN1 is to hold the mannose binding pocket together.

To test binding ability of LMAN1 mutants on the actual cargo proteins, COS I cells will be cotransfected with the Flag tagged WT or mutant LMAN1 (H178A or G251A/252A) and FV/FVIII. Twenty hours after transfection, cells will be metabolically labeled with [^{35}S]-methionine/cysteine for 45 minutes followed by a 30 minutes incubation in complete medium. This would allow the newly synthesized proteins to be labeled. The cells will be then treated with a crosslinking agent [DSP (dithio-bis succinimidyl propionate)] which will covalently link two interacting proteins. The interactions between LMAN1 and FV/FVIII will be assessed by immunoprecipitation using anti-Flag, anti -FV or anti-FVIII antibodies. If mannose binding activity of LMAN1 is important for interaction with FV/FVIII, in the presence or the absence of cross-linker because FV/FVIII will not be precipitated

along with mutant flag tagged LMAN1, using an anti-flag antibody. The WT flag tagged LMAN1 should show positive signal for FV/VIII by anti-flag antibody as WT LMAN1 would be able to interact with FV/FVIII and precipitate together.

The other study that we intend to do is to create a construct of LMAN1 with both the CRD and the helical domain (the entire luminal domain along with conserved cysteine residues C466 and C475) in a pET15b vector, express it into BL21 cells to have the corresponding protein. This protein would be then purified by modifying our denaturation-refolding protocol and used for crystallization to obtain the structure of the oligomerized LMAN1-MCFD2 complex. This structure will reveal how individual subunits of LMAN1 and MCFD2 are arranged in the hexamer. FV and FVIII are both large proteins containing multiple N glycosylation sites. It is important to figure out whether adjacent subunits of LMAN1 can coordinate to bind different branches of the same N-linked glycan or different oligosaccharides. Our lab has demonstrated that both LMAN1 and MCFD2 can interact with FV/ FVIII (Zheng C et al, 2010a). The oligomeric structure of the LMAN1-MCFD2 complex will shed light on how FV and FVIII cargo proteins simultaneously (or sequentially) interact with both MCFD2 and LMAN1.

In order to elucidate the role of MCFD2 on carbohydrate binding by the CRD of LMAN1, we need to first confirm that the purified MCFD2 is functional. This will be done by the ITC assay. It is known that LMAN1 and MCFD2 exist in a 1:1 stoichiometric complex. In this calorimetric assay the purified CRD is titrated against the purified MCFD2 in presence of Ca^{2+} using 10 mM HEPES and 150 mM NaCl at almost neutral pH of 7.4. If the stoichiometry of binding (n) is 1, this would indicate that the purified proteins are fully functional. If the purified proteins were non-functional then they would not be able to bind to one another in the 1:1 ratio. The

functional proteins are then allowed to form a complex in vitro and the complex will be purified by gel filtration chromatography to remove any unbound proteins. The purified complex will then be titrated against mannose α 1-2 mannose (ligand) and the association constant or K_a between the two is determined by the ITC assay. This K_a is compared with the K_a between the purified CRD (which is not complexed with MCFD2) and mannose α (1-2) mannose. If the K_a value is not higher when the CRD is bound to MCFD2, then it is likely that MCFD2 does not enhance the sugar binding ability of LMAN1.

BIBLIOGRAPHY

1. Appenzeller- Herzog C, Roche A-C, Nufer O, Hauri H-P, (2004) pH-induced conversion of the Transport Lectin ERGIC-53 Triggers Glycoprotein Release, *J Biol Chem*, Vol 279, No. 13, 12943-12950.
2. Appenzeller C, Andersson H, Kappeler F, Hauri H-P (1999) The lectin ERGIC-53 is a cargo transport receptor for glycoproteins, *Nat Cell Biol*, Vol 1
3. Belden W.J and Barlowe C, (2001) Role of Erv29p in Collecting Soluble Secretory Proteins into ER-Derived Transport Vesicles, *Science*, Vol 294, 1528-1531.
4. Bonifacino J.S and Glick B.S, (2004) Mechanism of Vesicle Budding and Fusion, *Cell*, Vol 116, 153-166.
5. Bhattacharya, D and Glick, B.S, (2007) Two mammalian Sec16 homologs have nonredundant functions in ER export and transitional ER organization, *Mol Biol Cell* 18, 839-849.
6. Baines A.C and Zhang B, (2007) Receptor-mediated protein transport in the early secretory pathway, *Trends in Biochemical Sciences*, Vol 32, No.8.
7. Bos M.H.A and Camire M.R, (2010) Blood Coagulation factors V and VIII: Molecular Mechanism of Procofactor Activation, *J Coagul Disord*, 2(2), 19-27.
8. Clark E.D.B, Schwarz E, Rudolph R, (1999) Inhibition of Aggregation Side Reactions during *in vitro* Protein Folding, *Methods in Enzymology*, Vol 309.
9. Cummings R.D et al, (1999) Essentials of Glycobiology, *Cold Spring Harbor Laboratory Press*, Ch 30, Plant Lectins.
10. Camire R.M, (2009) Hemophilia: Basic and translational science, *Expert Rev Hematol*, 3(2), 149-151.
11. Futai, E et al, (2004) GTP/GDP exchange by Sec12p enables COPII vesicle bud formation on synthetic liposomes. *EMBO J*. 23, 4146-4155.
12. Guy J.E, Wigren E, Svard M, Hard T and Lindqvist Y, (2008) New Insights into Multiple Coagulation Factor Deficiency from the Solution Structure of Human MCFD2, *Journal of Molecular Biology*, 381, 941-955.
13. Hauri H.P and Schweizer A, (1992) The endoplasmic reticulum-Golgi intermediate compartment, *Current Opinion in Cell Biology*, 4:600-608.
14. Hauri H-P, Kappeler F, Anderson H and Appenzeller C, (2000) ERGIC 53 and traffic in the secretory pathway, *Journal of Cell Science*, 113, 587-596.
15. Helenius A et al, (2001), Intracellular Functions of N-linked Glycans, *Science*, 291

16. Hughes H, Stephens DJ, (2008) Assembly, organization and function of COPII coat, *Histochem Cell Biol*, 129,129-51.
17. Itin C, Roche A-C, Monsigny M and Hauri H-P, (1996) ERGIC 53 is a functional Mannose-selective and Calcium-dependant Human Homologue of Leguminous Lectins, *Molecular Biology of the Cell*, Vol 7,483-493.
18. Kawasaki N, Ichikawa Y, Matsuo I, Totani K, Matsumoto N, Ito Y , Yamamoto K, (2007) The sugar-binding ability of ERGIC 53 is enhanced by its interaction with MCFD2, *Blood*, Vol 111, No. 4.
19. Kamiya Y, Kamiya D, Yamamoto K, Nyfeler B, Hauri H-P, Kato K (2008) Molecular Basis of Sugar Recognition by the Human L-type Lectins ERGIC-53, VIPL, and VIP36, *JBC*, Vol 283, No. 4.
20. Kondylis V, Pizette S and Rabouille C, (2009) The early secretory pathway in development : A tale of proteins and mRNAs, *Seminars in Cell and Developmental Biology* 20, Review ,817-827.
21. Lippincott-Schwartz, J. and Liu W, (2006) Insights into COPI coat assembly and function in living cells. *Trends Cell Biol*, 16, e1-e4.
22. Martinez-Meranguez, Geuze H J, Slot J W and Klumperman J, (1999) Vesicular tubular clusters between the ER and Golgi mediate concentration of soluble secretory proteins by exclusion from COPI coated vesicles, *Cell* 98,81-90.
23. Moussalli M, Pipe SW, Hauri HP, Nichols WC, Ginsburg D, Kaufman RJ, (1999) Mannose-dependent endoplasmic reticulum (ER)-Golgi intermediate compartment-53-mediated ER to Golgi trafficking of coagulation factors V and VIII. *J Biol Chem*. 274(46):32539–32542.
24. Middelberg A.P.J, (2002) Preparative Protein Refolding, *Trends in Biotechnology*, Vol 20, No. 10.
25. Mann K.G and Kalafatis M, (2003) Factor V: A combination of Dr Jekyll and Mr Hyde, *Blood*, 101, 20-30.
26. Nichols et al, (1998) Mutations in the ER-Golgi intermediate compartment protein ERGIC-53 cause combined deficiency of coagulation factors V and VIII, *Cell*, 93,61-70.
27. Neerman-Arbez M, Johnson K.M, Morris M.A, Mcvey J.H, Peyvandi F, Nichols W.C, Ginsburg D, Rossier C, Antonarakis S.E & Tuddenham E.G.D, (1999) Molecular analysis of ERGIC 53 in 35 families with combined Factor V and Factor VIII deficiency, *Blood*, 93, 2253-2260.
28. Nichols W.C, Terry V.H, Wheatley M.A, Yang A, Zivelin A, Ciavarella N, Stefanile C, Matsushita T, Saito H, de Bosch N.B, Ruiz-Saez A, Torres A, Thompson A.R, Feinstein D.I , White G.C , Negrier C, Vinciguerra C, Aktan M, Kaufman R.J,

- Ginsburg D & Seligsohn U, (1999) ERGIC 53 gene structure and mutation analysis in 19 combined factors V and VIII deficiency families, *Blood*, 93, 2261-2266.
29. Neve E.P et al, (2003) VIPL, a VIP36-like membrane protein with a putative function in the export of glycoproteins from the endoplasmic reticulum, *Exp. Cell Res*,288,70-83.
30. Nufer O, Kappeler F, Guldbrandsen S and Hauri H-P, (2003), ER export of ERGIC-53 is controlled by cooperation of targeting determinants in all three of its domains, *Journal of Cell Science* 116, 4429-4440.
31. Nufer O. et al, (2003) Profile-based data base scanning for animal L-type lectins and characterization of VIPL, a novel VIP36-like endoplasmic reticulum protein, *J Biol Chem*, 278,15886-15896.
32. Neve E.P, Lahtinen U, Petterson R. F (2005) Oligomerization and Intracellular localization of the Glycoprotein Receptor ERGIC-53 is Independent of Disulfide bonds, *JMB*, 354, 556-568.
33. Nyfeler B, Kamiya Y, Boehlen F, Yamamoto K, Kato K, Moerlose P, Hauri H-P, Neerman-Arbez M (2007) Deletion of three residues from the C terminus of MCFD2 affects binding to ERGIC 53 and causes combined factor V and factor VIII deficiency, *Blood*,111,1299-1301.
34. Nyfeler B, Reiterer V, Wendeler MW et al, (2008), Identification of ERGIC-53 as an intracellular transport receptor of $\alpha 1$ antitrypsin, *J Cell Biol* , 180(4):705-712.
35. Nishio M, Yukiko K, Mizushima T, Wakatsuki S, Sasakawa H, Yamamoto K, Uchiyama S, Noda M, McKay A.R, Fukui K, Hauri H-P, Kato K, (2010) Structural basis for the cooperative interplay between the two causative gene products of combined factor V and factor VIII deficiency, *Proc. Natl . Acad. Sci .U. S. A*, 107,9, 4034-4039.
36. Oeri J, Matter M, Isenschmid H, Hauser F, Koller F, (1954) Congenital Factor V deficiency (parahemophilia) with true hemophilia in two brothers, *Bibl Paediatr*, 1954,58: 575-88.
37. Oprins A, Rabouille C,Posthuma G, Klumperman J, Geuze HJ, and Slot JW, (2001)The ER to Golgi interface is the major concentration site of secretory proteins in the exocrine pancreatic cell, *Traffic* 2,831-838.
38. Orci L. et al, (2003) Mammalian Erv46 localizes to the endoplasmic reticulum-Golgi intermediate compartment and to *cis*-Golgi cisternae, *Proc. Natl . Acad. Sci .U. S. A* 100, 4586-4591.
39. Otte, S. and Barlowe C, (2004) Sorting signals can direct receptor mediated export of soluble cargo proteins into COPII vesicles, *Nat Cell Biol*,61189-1194.
40. Pierce M.M, Raman C.S, Nall B.T, (1999) Isothermal Titration Calorimetry of Protein-Protein Interactions, *Methods*, 19,213-221.

41. Pattison R.J, Amtmann A (2009) N-glycan production in the endoplasmic reticulum of plants, *Trends in Plant Science*, Cell Press, Vol 14, No. 2.
42. Schweizer A, Fransen J.A.M, Bachi T, Ginsel L and Hauri H.P, (1988) Identification, by a Monoclonal Antibody, of a 53-kD Protein Associated with a Tubulo-Vesicular Compartment at the cis-side of the Golgi Apparatus, *The Journal of Cell Biology*, Vol 107, 1643-1653.
43. Schindler R, Itin C, Zerial M, Lottspeich F and Hauri H.P (1993), ERGIC-53, a membrane protein of the ER-Golgi intermediate compartment, carries an ER retention motif, *Eur. J. Cell Biol*, 61, 1-9.
44. Stephens D.J and Pepperkok R, (2001) Illuminating the secretory pathway: when do we need vesicles? *Journal of Cell Science* 114,1053-1059.
45. Supek, F. et al, (2002) Sec 16p potentiates the action of COPII proteins to bud transport vesicles, *J Cell Biol*, 158, 1029-1038.
46. Soderholm, J et al, (2004) The transitional ER localization mechanism of *Pichia pastoris* Sec 12, *Dev. Cell* 6,649-659.
47. Satoh T, Cowieson N.P, Hakamata W, Ideo H, Fukushima K, Kurihara M, Kato R, Yamashita K, Wakatsuki S, (2007) Structural Basis for Recognition of High Mannose Type Glycoproteins by Mammalian Transport Lectin VIP36, *J Biol Chem*, Vol 282, No. 38.
48. Spreafico, M & Peyvandi F, (2008) Combined Factor V and Factor VIII deficiency, *Haemophilia*, 14, 1201-1208.
49. Tekaya H-B (2004) The ER-Golgi-Intermediate Compartment: dynamics and cargo sorting studied by time-lapse video microscopy, Basel.
50. Thor F, Gautschi M, Geiger R and Helenius A, (2009) Bulk Flow Revisited: Transport of a Soluble Protein in the Secretory Pathway, *Traffic*, 10: 1819-1830.
51. Velloso L.M, Svensson K, Schneider G, Petterson R. F, Lindqvist Y, (2002), Crystal Structure of the Carbohydrate Recognition Domain of p58/ERGIC-53, a Protein Involved in Glycoprotein Export from the Endoplasmic Reticulum, *The Journal of Biological Chemistry*, Vol 277, 15979-15984.
52. Velloso L.M, Svensson K, Schneider G, Petterson R. F, Lindqvist Y, (2003), The Crystal Structure of the Carbohydrate-recognition Domain of the Glycoprotein Sorting Receptor p58/ERGIC-53 Reveals an Unpredicted Metal-binding Site and Conformational Changes Associated with Calcium Ion Binding, *Journal of Molecular Biology*, 334, 845-851.
53. Watson, P. et al, (2006) Sec 16 defines endoplasmic reticulum exit sites and is required for secretory cargo export in mammalian cells, *Traffic* 7, 1678-1687.

54. Wigren E, Bourhis J-M, Kursula I, Guy J. E, Lindqvist Y (2010) Crystal structure of the LMAN1- Crd/MCFD2 transport receptor complex provides insight into combined deficiency of factor V and factor VIII, *FEBS Lett*, doi:10.1016/j.febslet.2010.02.009
55. Zhang B et al (2003) Bleeding due to disruption of a cargo-specific ER-to-Golgi transport complex, *Nature Genetics*, Vol 34, No. 2.
56. Zhang B, Kaufman R and Ginsburg D, (2005) LMAN1 and MCFD2 form a Cargo Receptor Complex and Interact with Coagulation Factor VIII in the Early Secretory Pathway, *Journal of Biological Chemistry*, Vol 280, 25881-25886.
57. Zhang B et al, (2006) Combined deficiency of factor V and factor VIII is due to mutations in either LMAN1 or MCFD2, *Blood*, Vol 107, No.5.
58. Zhang Y and Williams D.B (2006), Assembly of MHC class I molecules within the endoplasmic reticulum, *Immunol Res.* 35,151-162.
59. Zhang B, Spreafico M, Yang A, Platzer P, Callaghan M.U, Avci Z, Ozbek N, Mahlangu J, Haw T, Kaufman R. J, Marchant K, Tuddenham E.G, Seligsohn U, Payvandi F & Ginsburg D, (2008) Genotype-Phenotype correlation in combined deficiency of Factor V and Factor VIII, *Blood*, 111, 5592-5600.
60. Zhang B, (2009) Recent developments in the understanding of the combined deficiency of FV and FVIII, *British Journal of Hematology*, 145, 15-23.
61. Zhang Y.C, Zhou Y, Yang C.Z and Xiong D.S, (2009) A review of ERGIC-53:Its structure, functions, regulation and relations with diseases, *Histol Histopathol* ,24, 1193-1204
62. Zheng C, Liu H-H, Zhou J, Zhang B, (2010a) EF hand domains of MCFD2 mediate interactions with both LMAN1 and coagulation factor V or VIII, *Blood*, 115(5):1081-7
63. Zheng C, Liu H, Yuan S, Zhou J and Zhang B, (2010b) Molecular Basis of LMAN1-MCFD2 cargo receptor formation and ER-to-Golgi transport of FV/FVIII, *Blood*, Vol 116, No. 25.
64. Zhang B, Zheng C, Zhu M, Tao J, Vasievich M.P, Baines A, Kim J, Schekman R, Kaufman R, Ginsburg D (2011) Mice deficient in LMAN1 exhibit FV and FVIII deficiencies and liver accumulation of α 1- antitrypsin, *Blood*, Volume 118, Number 12

APPENDICES

APPENDIX 1

1. Construction of pET15b-GST-CRD

Restriction Digestion of pET15b-CRD with NcoI (50 μ l reaction)

Water	Vector(pET15b-CRD)	NE Buffer 3	NcoI
8 μ l	35 μ l	5 μ l	2 μ l

Restriction Digestion of Amplified GST (insert) with NcoI (50 μ l reaction)

Water	Vector(pET15b-CRD)	NE Buffer 3	NcoI
8 μ l	35 μ l	5 μ l	2 μ l

Ligation Reaction

Water	Ligation Buffer (2X)	Vector(pET15bCRD)	Insert(GST)	Ligase
1 μ l	5 μ l	2 μ l	1 μ l	1 μ l

Restriction Digestion of pET15b-GST-CRD with NcoI (20 μ l reaction)

Water	Ligated Plasmid(pET15b-GST-CRD)	NE Buffer 3	NcoI
12.5 μ l	5 μ l	2 μ l	0.5 μ l

APPENDIX 2

Construction of pET15-CRD

Restriction Digestion of Amplified CRD (insert) with Nde1 & BamH1 (50 μ l reaction)

Water	Amplified CRD	NE Buffer	BSA	BamH1	Nde1
7.5 μ l	34 μ l	3 5 μ l	0.5 μ l	1.5 μ l	1.5 μ l

Restriction Digestion of pET15b (vector) with Nde1 & BamH1 (50 μ l reaction)

Water	Vector	NE Buffer	BSA	BamH1	Nde1
7.5 μ l	34 μ l	3 5 μ l	0.5 μ l	1.5 μ l	1.5 μ l

Restriction Digestion of Ligated pET15b-CRD with Nde1 & BamH1 (20 μ l reaction)

Water	Amplified CRD	NE Buffer	BSA	BamH1	Nde1
11.8 μ l	5 μ l	3 2 μ l	0.2 μ l	0.5 μ l	0.5 μ l

APPENDIX 3

3. Crystallization and data collection

Crystals of FlnA-Ig10 were grown by hanging drop vapor diffusion at 293 K in 2 μ l drops. Sitting drops initially in a 1:1 ratio mixture. Final hanging drops in a 2:1 (protein: screen) ratio mixture (1.2 μ l drops)

Hanging drops consisted of a 1:1 ratio mixture of 1 mM (10 mg/ml) protein and a reservoir solution. Initial screening was conducted using the sparse matrix crystallization screens JCSG+ (Qiagen), ProComplex (Qiagen), Classics (Hampton Research), Index (Hampton Research) and SaltRx (Hampton Research). Subsequent optimizations identified an optimal reservoir solution composed of 0.2 M ammonium acetate, 25 % (w/v) polyethylene glycol (PEG) 3350 and 0.1 M Bis-Tris, pH 5.5. Single crystals appeared after 6 weeks and grew for an additional 2 weeks. Crystals were cryoprotected by brief transfer through reservoir solution supplemented with 20% (w/v) glycerol and flash-frozen in liquid nitrogen.

Diffraction data was collected with a Cu-K α source at 1.5418 Å using a Rigaku MicroMax-007HF generator and a Rigaku Saturn 944+ CCD detector. Data integration was performed with *XDS* (2). Data reduction and scaling were performed with *SCALA* (3). The final data set was processed to a cut-off of 2.44 Å, based on significant drops in unaveraged I/sigma (<2.7) and completeness (<90%) in shells at higher resolutions.

Structure solution and refinement

A molecular replacement search model was generated by iTASSER using the HsFlnA-Ig10 sequence (residues 1158-1252) as input. Phases were calculated by molecular replacement using the *PHASER* component of *PHENIX*. The best molecular replacement solution was subjected to automated rebuilding in *PHENIX* with *RESOLVE*, followed by iterative rounds of model building in *COOT* and refinement in *PHENIX*. The refinement protocol used isotropic atomic displacement parameters for all atoms. All molecular structure figures were generated with *Pymole*.

Validation and deposition

Stereochemical analysis of the FlnA-Ig10 structure was completed with *MolProbity*. All residues are within the favored regions of the Ramachandran plot. No poor rotamers or C β deviations were found. *MolProbity* analysis of all-atom contacts calculated a clash score of 7.09, which ranks the FlnA-Ig10 structure in the 98th percentile of 326 structures deposited in the PDB that were solved at similar resolution ($2.442 \text{ \AA} \pm 0.25 \text{ \AA}$). The *MolProbity* score, a weighted measure of stereochemical statistics, was 1.44, which ranks the FlnA-Ig10 structure in the 100th percentile of 7,752 structures of similar resolution deposited in the PDB. The atomic coordinates for Ig10 have been deposited in the PDB (accession code XXX).

Data Collection

Beam Line	APS 24-ID-E
Wavelength (Å)	0.9792
Space group	<i>P6</i>
Cell dimensions	
<i>a</i> , <i>b</i> , <i>c</i> (Å)	110.26, 110.26, 38.33
α , β , γ (°)	90, 90, 120
Resolution (Å) ^a	47.75-2.42 (2.55-2.42)
R_{merge} ^b	0.082 (0.437)
$R_{r.i.m.}$ ^c	0.086 (0.459)
<i>I</i> / σ <i>I</i>	23.0 (7.7)
Wilson <i>B</i> factor (Å ²)	43.9
Completeness (%)	100.0 (100.0)
Redundancy	10.7 (10.9)
No. of reflections	111,907
No. of unique reflections	10,429

Refinement

Resolution (Å)	38.33-2.42
No. reflections for refinement	10,427
R_{work}/R_{free}	0.202 / 0.226
No. atoms	
Protein	1,799
Water	83
Calcium	2
2 α -Mannobiose	23

Average *B* factors

Protein	38.1
Water	45.0
Calcium	47.1
2 α -Mannobiose	60.8

R.m.s deviations

Bond lengths (Å)	0.007
Bond angles (°)	0.953

Ramachandran plot statistics

Favored regions %	97.4 (221/227)
Allowed regions %	100.0 (227/227)
Disallowed regions	0.0

MolProbity validation statistics

Poor rotamers (%)	0.0
C β deviations >0.25Å	0
MolProbity clash score	13.36
MolProbity clash percentile	87 th percentile (N=330, 2.42Å \pm 0.25Å)
MolProbity score	1.76
MolProbity score percentile	98 th percentile (N=7,808, 2.42Å \pm 0.25Å)

^a Values in parentheses are for the highest resolution shell.

^b The merging *R* factor is defined as
$$R_{merge} = \frac{\sum_{hkl} \sum_i |I_i(hkl) - \overline{I(hkl)}|}{\sum_{hkl} \sum_i I_i(hkl)}$$

^c The redundancy-independent merging *R* factor (*I*) is defined as
$$R_{r.i.m.} = \frac{\sum_{hkl} \left[\frac{N}{N-1} \right]^{1/2} \sum_i |I_i(hkl) - \overline{I(hkl)}|}{\sum_{hkl} \sum_i I_i(hkl)}$$

APPENDIX 4

pED vector Map (Kaufman R, Molecular biotechnology, Vol16, 2000)

