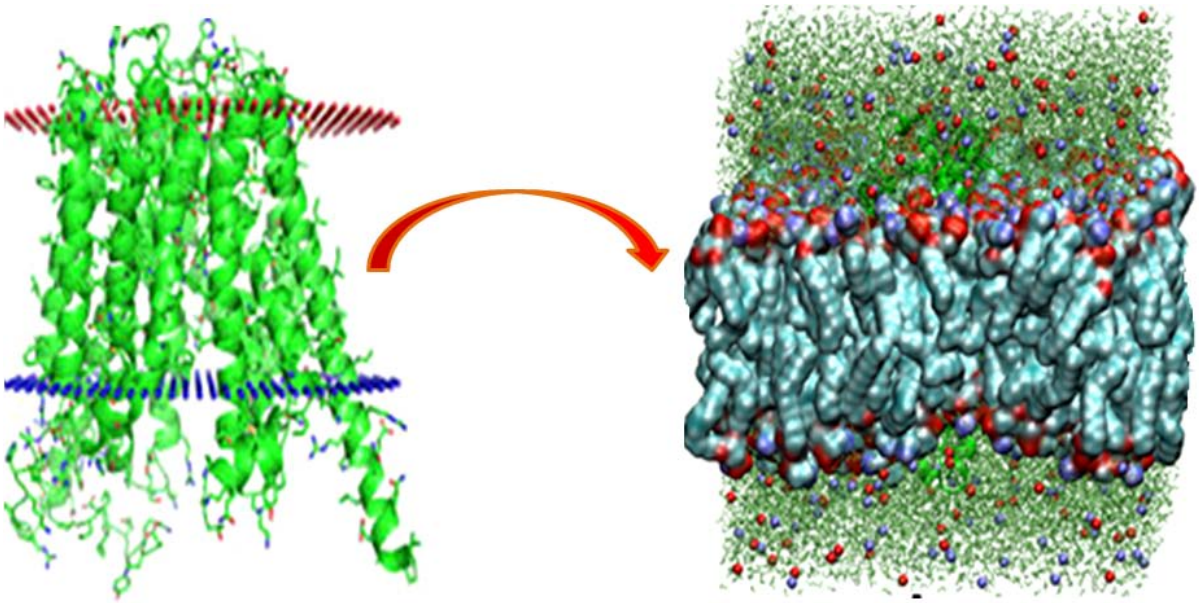
Universitat Autònoma de Barcelona

Structure and dynamics of a membrane symporter: The Melibiose Permease



Vidhya Ravi Madapusi

2016

Universitat Autònoma de Barcelona

Structure and dynamics of a membrane symporter: The Melibiose Permease

Memòria presentada per **Vidhya Ravi Madapusi**

per optar al grau de Doctor.

Els treballs presentats han estat dirgits pel

Dr. Josep Bartomeu Cladera Cerdà i el Dr. Esteve Padrós i Morell,

i realitzats en la Unitat de Biofísica

del Department de Bioquímica i de Biologia

Molecular i el Centre d'Estudis en Biofísica (CEB)

Facultat de Medicina

de la Universitat Autònoma de Barcelona

Vist i plau dels directors de la tesi:

Dr. Esteve Padrós i Morell **Dr. Josep Bartomeu Cladera Cerdà**

“You are never too old to set another goal

Or to dream a new dream”

- C.S. Lewis

3. MATERIALS AND METHODS

Results and Discussion

ABBREVIATIONS

Abbreviations

Amino Acid name	Three letter code	Single letter code
Alanine	Ala	A
Arginine	Arg	R
Asparagine	Asn	N
Aspartic Acid	Asp	D
Cysteine	Cys	C
Glutamic Acid	Glu	E
Glutamine	Gln	Q
Glycine	Gly	G
Histidine	His	H
Isoleucine	Ile	I
Leucine	Leu	L
Methionine	Met	M
Phenylalanine	Phe	F
Proline	Pro	P
Serine	Ser	S
Threonine	Thr	T
Tryptophan	Trp	W
Tyrosine	Tyr	Y
Valine	Val	V

ABBREVIATIONS

General Abbreviations

ABC	ATP-binding cassette superfamily
ABS	absorbance spectrum
AEBSF	4-(2,Aminoehtyl) benzenesulfonyl fluoride hyrdrochloride
ATP	Adenosine-triphosphate
ATR-FTIR	attenuated total reflection-Fourier transform infrared
Cys-less	MelB free of all intrinsic cysteines
cryo-EM	cryo-electron microscopy
cm^{-1}	Unit of FTIR (resolution)
DDM	dodecyl-maltoside
D^2G	Dns^2-S-Gal, 2'-(N-dansyl)-aminoethyl-1-thi-D-galactopyranoside.
DTT	1,4-dithio-threitol
EDTA	ethylenediaminetetraacetic acid
FRET	Forster energy transfer
F_0-F_1	F-type ATPase
FTIR	Fourier transform Infrared
g	gram
GLUT1	Glucose transporter

ABBREVIATIONS

GPH	gluoside-pentoside-hexuronide:cation symporter family
GlpT	glycerol-3-phosphate antiporter from E. *coli*
GPRCR	G protein coupled receptors
h	hour
ITD	iodide transporter
ISO	inside out
IR_{diff}	IR difference spectrum
KPi	potassium phosphate buffer ($K_2HPO4 + KH_2PO_4$)
lacY	gene encoding for lactose permease from E. *coli*
LacS	lactose permease from *S. thermophiles*
LacY	lactose permease from E. *coli*
$LeuT_{AA}$	sodium-leucine transporter from *Aquifex aeolicus*
LB	Luria Bertani broth
mel	melibiose [α- D-galactopyranosyl-(1-6)- α-D-glucopyranose
melA	gene encoding for the α-galactosidase from E. *coli*
melB	gene encoding for the β-galactosidase from E. *coli*
$MelB_{EC}$	gene encoding for the melibiose permease from E. *coli*
$MelB_{ST}$	gene encoding for the melibiose permease from *S.typhimurium*

ABBREVIATIONS

MES	2-(*N-morpholino*) ethansulfonic acid
MFS	major facilitator superfamily
min	minute
MW	molecular weight
µL	microlitre
µM	micrometer
mL	milliliter
mM	millimolar
MW	molecular weight
NaCl	sodium chloride
ng	nanogram
nm	nanometer
NMR	nuclear magnetic resonance
PAGE	polyacrylamide gel electrophoresis
pCMBS	p-chloromercuribenzene sulfonate
psi	pressure bar per square inch

ABBREVIATIONS

R^2	correlation coefficient
s	second
SDS	sodium dodecyl sulfate
TM	transmembrane
UV-Vis	ultraviolet- visible
α-NPG	*p-nitrophenyl-α-D-6-galactopyranosie*

1. INTRODUCTION

1.1 The membrane proteins and their importance

Nutrient uptake is the essential process for all the living beings and it is mainly done through the cell membranes. These biological membranes consist of double layer of lipid molecules in which various proteins are embedded, the 'membrane proteins'. Membrane proteins play a vital role in import and export of various substances, transduction of signals and communication with other cells. As shown in the Fig. 1.1, some membrane proteins extend across the bilayer as: a single alpha helix (1), multiple α helices (2), or as a rolled-up β sheet (a β barrel) (3). Some of these "single-pass" and "multipass" proteins have a covalently attached fatty-acid chain inserted in the cytosolic lipid monolayer. Other membrane proteins are exposed at only one side of the membrane (4). Others are attached to the bilayer solely by a covalently attached lipid chain—in the cytosolic monolayer (5), or to phosphatidylinositol in the external monolayer (6). Finally, many proteins are attached to the membrane only by noncovalent interactions with other membrane proteins (7, 8) [Alberts *et al.,* 2002].

Statistical analysis shows that around 20 to 30% of the open reading frames encode for the integral membrane proteins. Recently, transporters are gaining enormous interest to the researchers, since an important number of human diseases are related to defects in membrane transport [Wallin *et al.,* 1998].

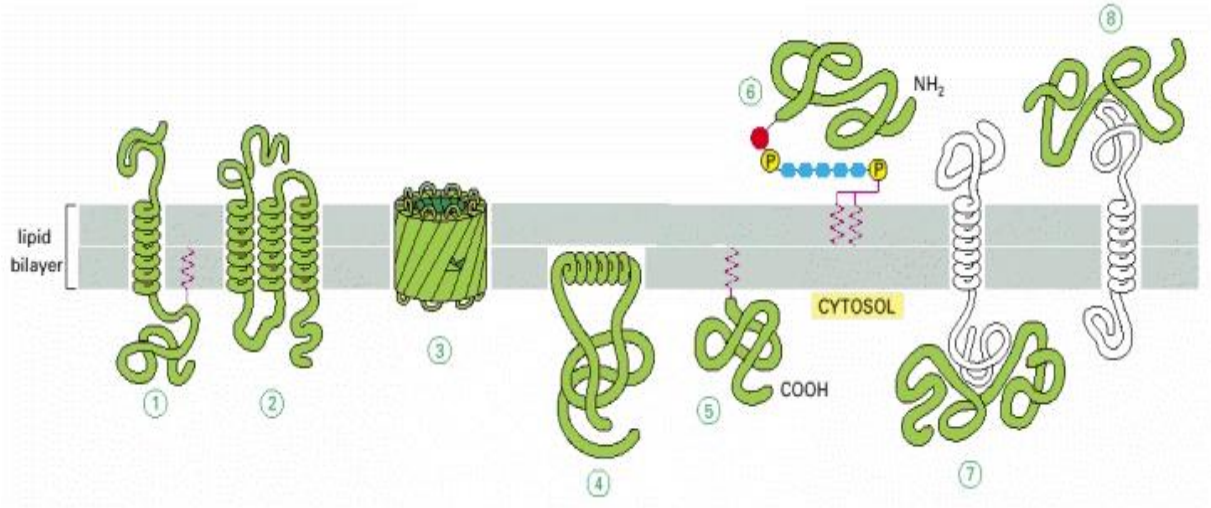


Fig. 1.1: Various ways in which membrane proteins associate with the lipid bilayer [Alberts *et al.,* 2002]

There are five main classes of membrane transport proteins as shown in Fig. 1.2:

1. **Channels:** These proteins allow the diffusion of molecules through the pore (e.g. glycerol channel).
2. **Primary Active Transporters:** These are the proteins which need energy for the transport, mainly ATP (e.g. ABC superfamily).
3. **Facilitative transporters:** The protein facilitates the transport of the molecule/ion without any use of energy. This is otherwise known as Uniporters (e.g. glucose transporter 1, GLUT1).
4. **Secondary Active Transporters:** This type of protein utilizes the energy stored in the concentration gradient which was created by pumping protons/ions out of the cells. If the ion and the substrate are transported in the same direction, the protein is called a symporter or co-transporter (e.g. the Melibiose Permease and SGLT1). Antiporters/Exchangers refer to the transport of ligands in the counter direction (e.g. Na^{+}/Ca^{2+} transport).

5. **Group Translocation:** In this case, the substrate is transported and concomitantly phosphorylated.

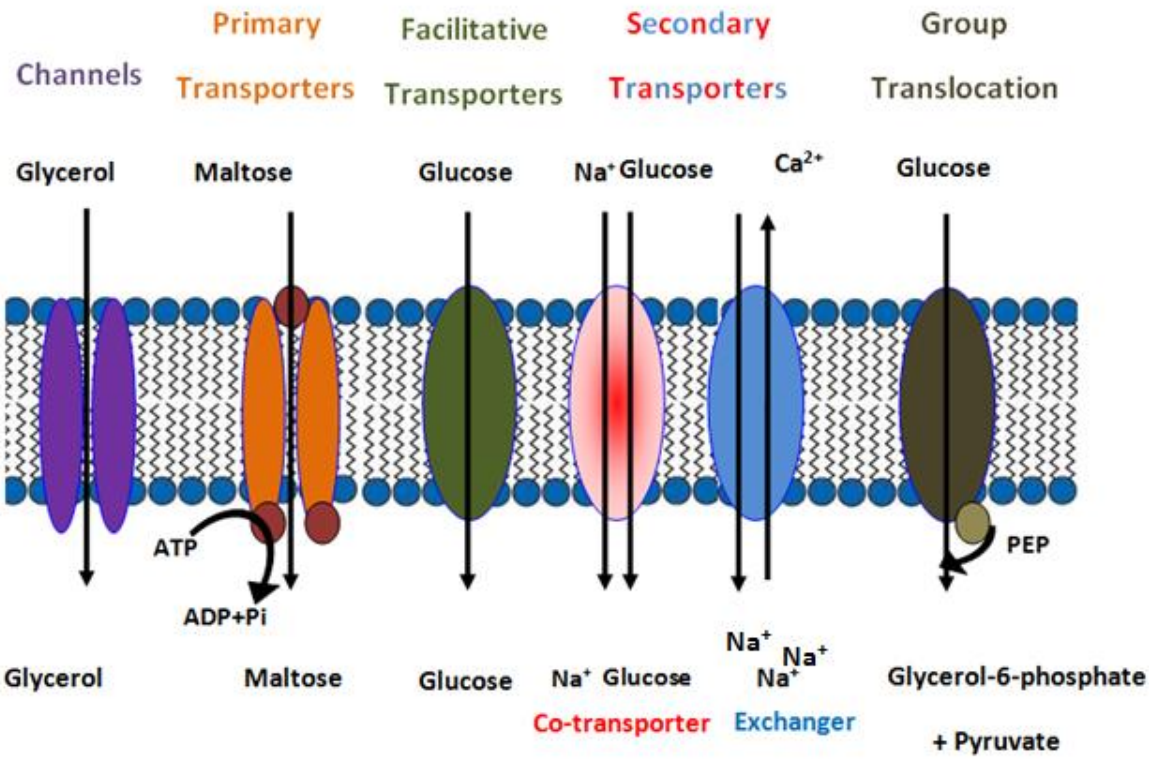


Fig. 1.2: Schematic diagram illustrating the different types of membrane transport proteins.

1.1.1. Solute Sodium Symporters (SSS)

From Mitchell´s chemiosmotic theory [Mitchell 1957], it is clearly known that ions play an important role in energy transduction across the biological membrane. Initially it was thought that Na^+ cotransport was a unique feature of animal cells alone. But, it has been shown that even bacterial cells use Na^+ cotransport. In E. *coli*, with the use of a Na^+ electrochemical gradient several substrates like proline, glutamate, serine/threonine, pantothenate and melibiose are transported by PutP, GltS, SstT, PanF and MelB as shown in Fig. 1.3 [Winkelmann 2008].

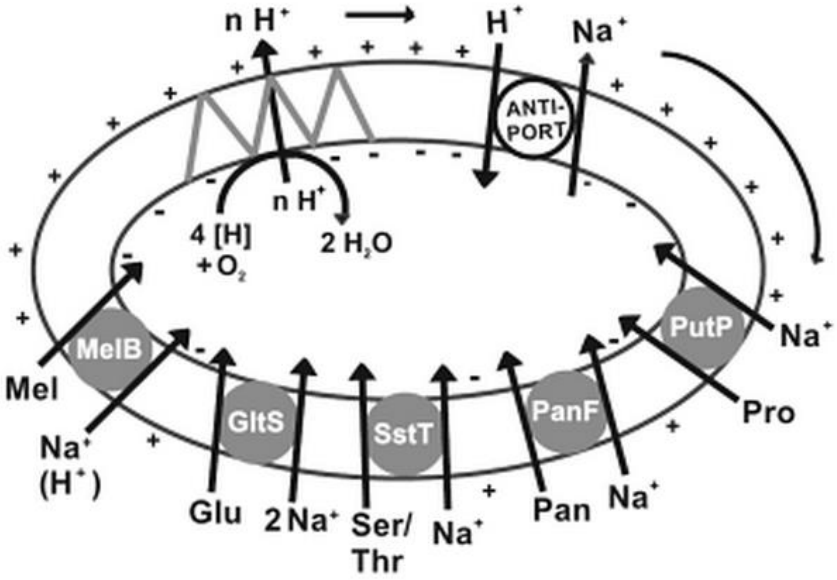


Fig. 1.3: The model represents the Na^+ dependent substrate uptake system in E. *coli*. The cells generate a primary H^+ cycle which is subsequently used to establish a Na^+ cycle which involves the action of Na^+/H^+ antiporters and the indicated Na^+/substrate symport systems [Winkelmann 2008].

Na^+ cotransport uses the energy of a Na^+ electrochemical gradient to drive the substrate accumulation. Bacteria generate the electrochemical gradient of Na^+ ($\Delta\mu_{Na+}$) by three different mechanisms as shown below (Krulwich 1990):

1. Majority of the bacteria use H^+ pumps to generate an electrochemical gradient, ATPases, where the $\Delta\mu_{H^+}$ generated by respiration is utilized to generate the Na^+ gradient using Na^+/H^+ antiport.

2. In some cases, bacteria utilize the Na^+ translocating decarboxylases. e.g., *Klebsiella aerogenes* and *Salmonella typhimurium* use the energy from decarboxylation to pump Na^+ through the membrane.

3. In some other cases, bacteria directly possess the Na^+ respiratory pump. e.g., *Propionogenium modestum.*

Two types of Na^+ cotransport have been described: simple Na^+ cotransporters, e.g., proline permease and cotransporters systems that can use $Na^+/Li^+/H^+$, e.g. melibiose permease from E. *coli.* Solute Sodium Symporters are important in human physiology and disease where mutations in glucose (SGLT1) and Iodide symporters (NIS) results in congenital metabolic disorders like Glucose-Galactose-malabsorption (GGM) and Iodide transport defect (ITD). Nowadays SGLTs are being targeted in drug trials for type II diabetes [Poulsen *et al.,* 2015]. Due to the complexity of purifying the homologues transport proteins in humans, many studies are carried out in prokaryotes to understand its function. For example, melibiose was the first substrate shown to be transported by Na^+ symport in enteric bacteria.

1.1.2. The major facilitator superfamily (MFS)

The Major Facilitator Superfamily (MFS) is the largest known super families of secondary transporters that has 10,000 sequenced members with 400- 600 amino acyl residues in length and possess 12, 14 or 24 putative α-helical membrane. The analysis of their primary sequences revealed that within any single family, sequence similarity is highly significant [Law *et al.,* 2008]. Initially, MFS was believed to uptake sugars alone, but the further studies concluded that drug efflux systems and Kreb cycle metabolites also belong to this family [Pao *et al.,* 1998]. The MFS family can transport their substrates in parallel or anti parallel mode, i.e. Uniport, symport or antiport.

1.1.3. The Melibiose Permease

The Melibiose Permease (MelB) belongs to the Glycoside (or galactoside)-pentoside-hexuronide (GPH): cation symporter family (2.A.2) [Poolman *et al.,* 1996, Saier 2000a]. The GPH family shows similarity with the other members of Major facilitator superfamily (MFS) as shown in the Fig. 1.4. The presence of the MFS motif R K (X G) R K has been described in MelB. This suggests that the GPH family is a member of the MFS superfamily [Saier 2000b; Reddy *et al.,* 2012]. GPH family members have a size of around 400 to 500 amino acids.

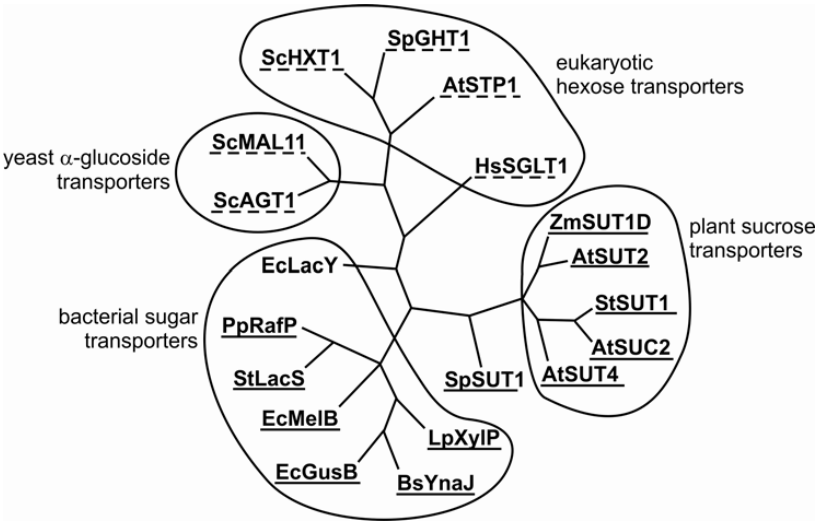


Fig. 1.4: Phylogenetic tree of sugar transporters from Bacteria, Yeast, Humans and Plants. The solid underscoring line indicate the members of GPH family and the dashed underscoring line indicate the members of MFS family [Saier 2000a].

The Melibiose Permease is found in the following organisms: *Salmonella typhimurium* [Mizushima *et al.,* 1992], *Klebsiella pneumonia* [Wilson *et al.,* 1992], *Enterobacter aerogenes* [Okazaki *et al.,* 1997] and *Escherichia coli* [Yazyu *et al.*, 1984].

The Melibiose Permease is encoded by the *melB* gene in the melibiose operon, which has two structural genes, *melA* and *melB. MelA* (α-galactosidase) is responsible for the hydrolysis of melibiose into galactose and glucose (Fig. 1.5). It has been shown that the melibiose permease contains 473 amino acids (70% of the residues are nonpolar) [Pourcher *et al.,* 1995] with a molecular weight of 52 kDa. The molecular mass assessed by the SDS polyacrylamide gel is around 39 kDa [Botfield and Wilson 1988] which is due to the high hydrophobicity of the protein in the solubilized state. MelB is a basic protein at pH 7 due to the excess of positive charges.

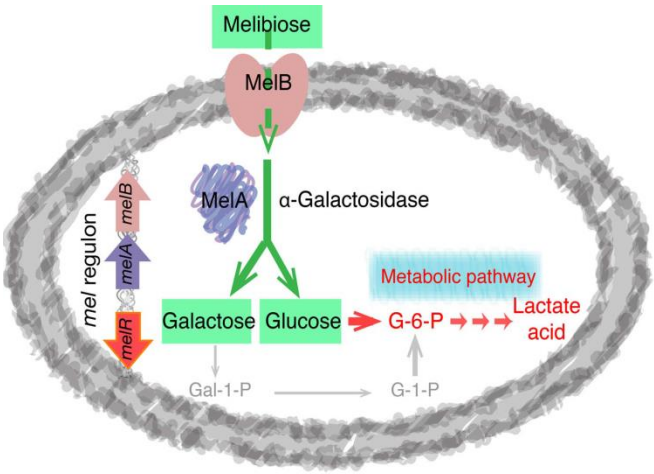


Fig. 1.5: Metabolic pathway of melibiose in *E. coli* [Tikonova *et al.,* 2015]

A versatile feature of *E. coli* MelB is it's broader than normal cation specificity [Pourcher *et al.,* 1995] which can use either Na^+ or H^+ or Li^+ as the coupling cation. The stability of the ternary complex cation-sugar-permease varies according to the chemical identity of the coupling ion ($H^+ < Li^+ < Na^+$). Sugars in α-configuration (e.g., melibiose and raffinose) can use all three cations (Na^+, Li^+ and H^+) whereas sugars in the β-configuration can use only Na^+ and Li^+ [Wilson *et al.,* 1987]. FTIR spectroscopy suggested that the permease is dominated by α-helical components around 50% and have β-structures (20%) and other additional components assigned to turns, 3_{10} helices and non-ordered structures (30%) [Dave *et al.,* 2000]. Cysteine scanning mutagenesis was primarily used in studying the structure function relationship of membrane proteins.

In MelB, a fully active permease devoid of the four native cysteine residues (Cys-less) [Weissborn *et al.,* 1997] has been shown to have similar transport properties as the wild type [Abdel-Dayem *et al.,* 2003; Meyer-Lipp *et al.,* 2006]. This was also confirmed by infrared spectroscopy by studying the binding of substrates [Leon *et al.,* 2009].

1.2. A Journey towards solving the 3D Structure of MelB

1.2.1. Prediction of 12 -Transmembrane helices

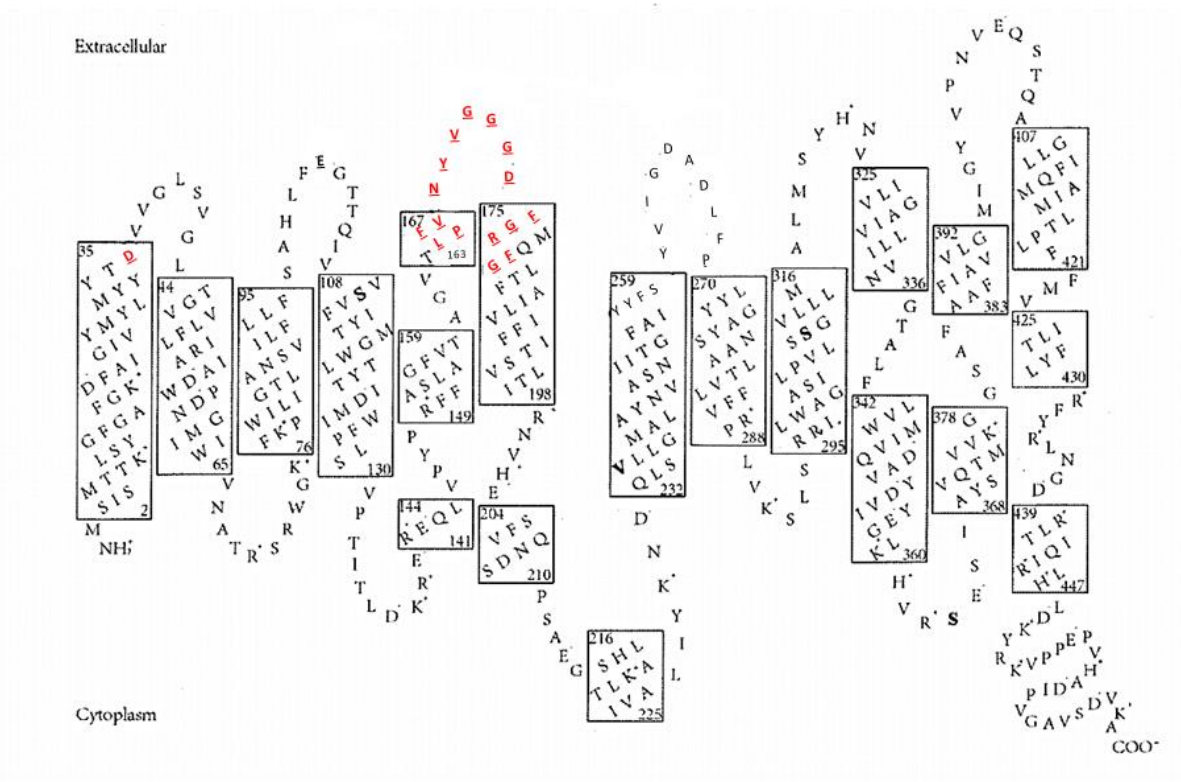


Fig. 1.6: Secondary structure of Melibiose Permease [Ethayathulla *et al.,* 2014]. The residues highlighted in red were studied in this thesis for Na^+/sugar binding (Leu-164 to Gly-178).

Before the success of crystallization, hydropathy profiling of the primary amino acid sequence was done to predict the number of helices in MelB. Initially, it was reported that MelB had 10 long hydrophobic segments traversing the membrane [Yazyu *et al.,* 1984]. Later on, it was suggested by Botfield and Wilson (1989) to have 10 or 12 transmembrane helices.

Then, a model predicted that MelB contained 12 transmembrane helices in a α-helical conformation connected by hydrophilic loops with N and C termini facing the cytoplasm [Pourcher *et al.,* 1990]. This prediction was further supported by the PhoA (alkaline phosphatase) fusion analysis [Botfield *et al.,* 1992, Pourcher *et al.,* 1996], proteolytic

cleavage experiments [Gwizdek *et al.,* 1997]. Fig. 1.6, shows the recent 2D-structure of MelB after the solved crystal structure by Ethayathulla *et al.,* 2014.

1.2.2. 2D crystals of MelB and its projection Map

The number of membrane protein structures account for less than one percent of all deposited structures in the PDB due to the problems in overexpression, purification and crystallization. In general, overexpression of membrane proteins yields less protein compared to the soluble proteins. An alternative method of study was cryo-electron microscopy [Hacksell *et al.,* 2002]. With this method, an asymmetric protein unit composed of two domains lining a central cavity was shown using 2D crystals (Fig. 1.7). 2D crystals were produced by removing the detergents from *E. coli* MelB using either the combination of bio-beads and dialysis or by dialysis alone. This resulted in the formation of 2D crystals of 49 Å*37 Å in flattened tubes.

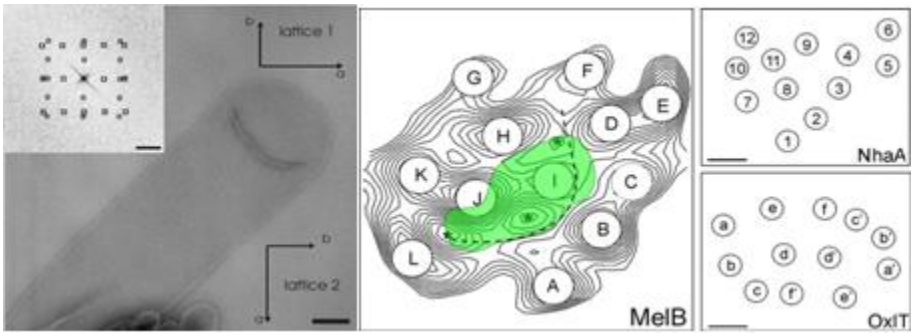


Fig. 1.7: Image taken from Hacksell *et al.,* 2002 representing the crystal of MeB **(A)** and the protein dense regions marked with letters (A-L) **(B)** protein dense regions of the NhaA antiporter and the OxlT structures.

This was followed by the production of three-dimensional structure of melibiose permease by cryo-electron microscopy at a resolution of 10 Å from 2D crystals [Purhonen *et al.,* 2005]. The crystals were formed in the presence of both substrates (Na^+ and melibiose). The structure shows a heart-shaped molecule (Fig. 1.8), which was closed in one end and open in the other, much like structure of LacY [Abramson *et al.,* 2003] and GlpT [Huang *et al*., 2003].

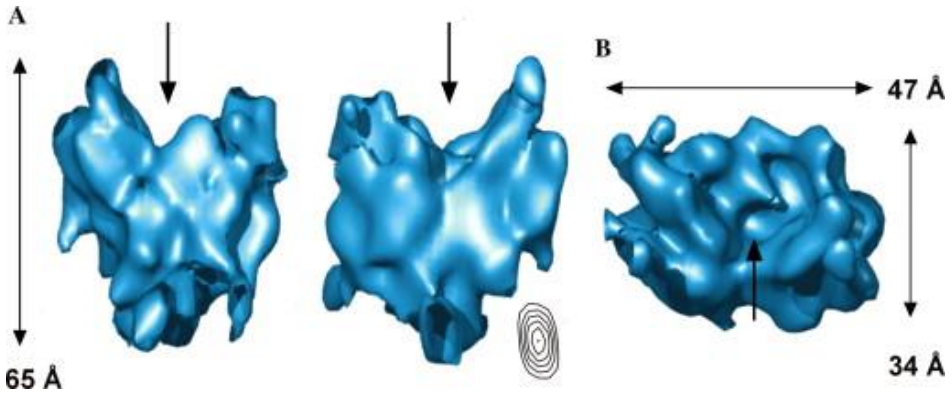


Fig. 1.8: 3D structure of MelB at 10 Å showing a heart-shaped molecule [Purhonen *et al.,* 2005].

1.2.3 Previous evidences about helix packing

Cysteine scanning mutagenesis was a good method to study the helix packing and its importance in the three dimensional structure. A functionally active permease devoid of four cysteines showed similar properties to Wild type [Weissborn *et al.,* 1997]. Zeng *et al.,* 1996, suggested that at least six helices were needed to create a channel to transport disaccharides. In order to find and localize a putative aqueous channel in MelB, some helices were systematically studied by the approach of Cys-scanning mutagenesis and probed by the water soluble SH-reagent p-choloromercuribenzene sulfonate (pCMBS). pCMBS inhibition studies showed that helix I is entirely inside the aqueous channel as shown in Fig. 1.9A and 9B [Ding and Wilson 2001a]; 30-35% of the residues in helices II (Fig. 1.10) [Matsuzaki *et al.,* 1999] and XI (Fig 1.11) [Ding and Wilson 2000] and helix IV is not in the aqueous channel [Ding *et al.,* 2001].

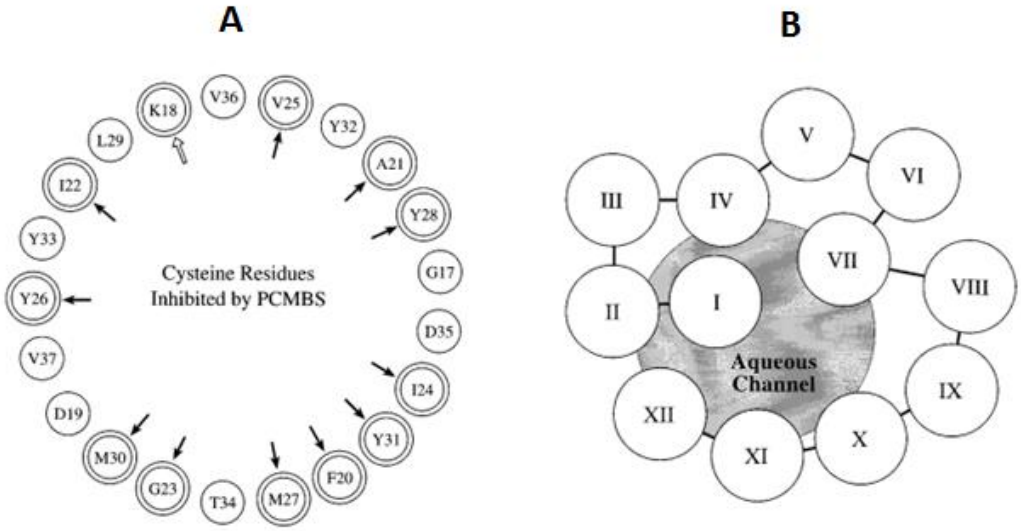


Fig. 1.9: (A) A helical wheel plot of residues in helix I. Residues inhibited by pCMBS are indicated with a black arrow and open arrow indicates no transport activity. **(B**) A proposed arrangement of helices showing that the helix I is surrounded by aqueous medium [Ding and Wilson 2001a]

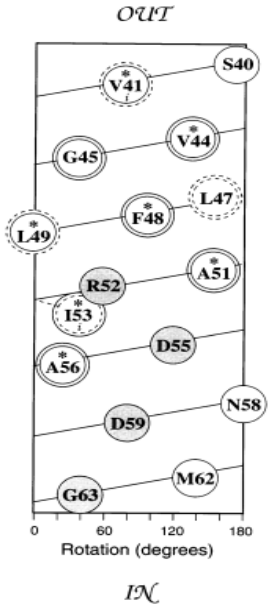


Fig. 1.10: A two-dimensional depiction of the hydrophilic face of transmembrane Helix II. The half of Helix II is shown that includes the positions where substitution of the wild-type amino acid by cysteine results in sensitivity to pCMBS. The cysteine substitutions that resulted in greater than 80% inhibition of transport after reaction with pCMBS are indicated by a second solid line around the residue name. The positions that are protected from inhibition by melibiose are indicated by an asterix (*) above the residue name. [Matsuzaki *et al.,* 1999]

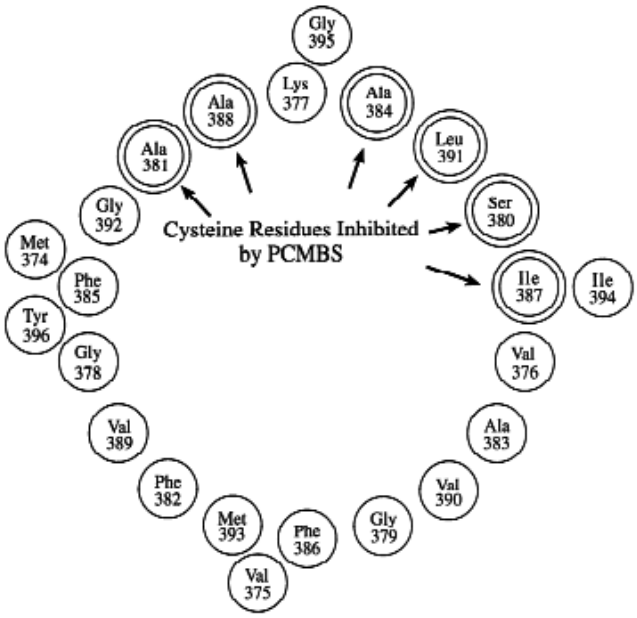


Fig. 1.11: (A) A helical wheel plot of residues in helix XI viewed from periplasmic surface of the membrane. Residues inhibited by PCMBS are indicated by arrow. [Ding and Wilson 2000]

Apart from PCMBS analysis, additional evidence for helix packing and proximity was inferred from analysis of second site revertants. Second site revertants are those in which first mutation resulted in severe loss of activity while a second mutation at a different site restored partial or full activity. Double mutation sites in a restored active mutant may implicate proximity between the two helices in which the two mutations occur. Some of the studies led to the following suggestions:

- **Helix I** (cytoplasmic end) is close to **Helix IV** (cytoplasmic end) (K18C/M123C) and to **Helix VII** (Y31C/S234L) [Ding and Wilson 2001a].

- **Helix II** is close to **Helix IV** (D55S/G177D; R52S/W116R) [Franco and Wilson 1999; Wilson et al., 1995] and also to **Helix VII** (R52S/S247R) [Franco and Wilson 1999] and **Helix X** (R52S/T338R) [Franco and Wilson 1999].

- **Helix IV** was shown to be in close contact with **Helix XI** (D142S/V375A) [Wilson and Wilson 1998].

- **Helix X** was proposed to be close to **Helix V** or **Loop 4-5** (R149C/V343G) [Abdel-Dayem *et al.,* 2003].

- **Loop 10-11** is in close proximity with the cytoplasmic part of **Helix I** (A350C/I22N) and to the periplasmic part of **Helix VIII** (A350C/F268L) [Ding 2004].

- **Helix XI** is close to **Helix I** (K377V/F20L, K377C/I22S) [Ding and Wilson 2000a, Franco *et al.,* 2001] **Helix II** (K377C/D59A; K377V/D59V) [Ding and Wilson 2000a], **Helix V** or **loop 4-5** (L391C/A152S) and finally to **C-terminal loop** (F385/R441) (Fig. 13) [Ding and Wilson 2000a].

- Further crosslinking studies by Ding and Wilson 2001c confirmed the proximity between helices I and XI.

1.2.4. Threading analysis: A 3D model of MelB

When membrane crystallization was hard to achieve, some of the computational methods like threading made a breakthrough in structural biology. The 3D structure model of MelB was threaded through the crystal structure of lactose permease of E.coli (LacY) [Yousef and Guan 2009], as shown in Fig. 1.12. This model contains 442 residues, including all 12 transmembrane helices and connecting loops, with no steric clashes and superimposes well with the template structure. The 3D model indicates that MelB consists of two pseudo symmetrical domains with 6 helices each, with the cation binding site mostly formed by the N-terminal transmembrane domain and the sugar-binding site by the C-terminal domain

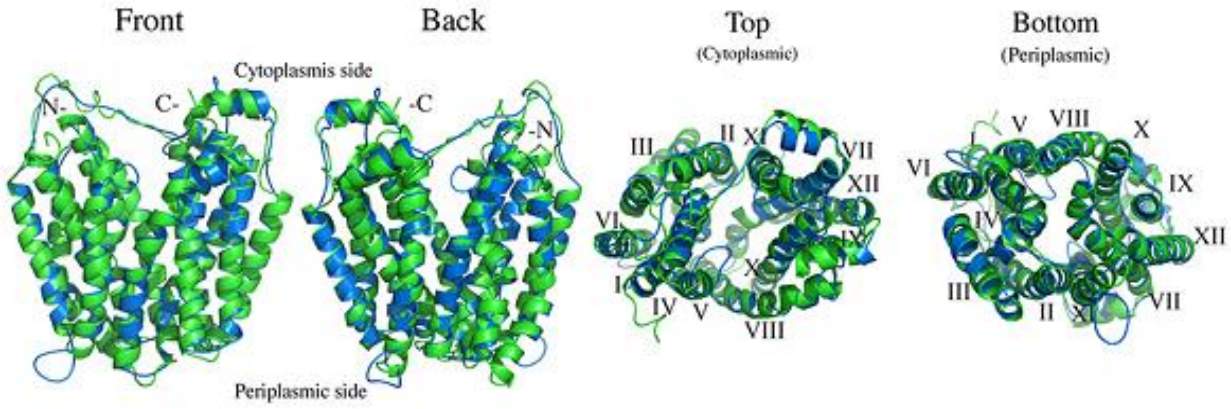


Fig. 1.12: 3D- model of melibiose permease. Comparison between the threading model of MelB and the crystal structure of LacY [Yousef and Guan 2009].

1.2.5. 3D structure of MelB

The three dimensional structure of MelB from Salmonella typhimurium ($MelB_{St}$) PDB code: 4M64 was solved by the molecular replacement method and was refined to a resolution of 3.35 Å [Ethayathulla *et al.,* 2014]. This structure confirms that the MelB adopts typical MFS fold, organized in amino and carboxy-terminal six transmembrane α-helix bundles, with a connecting cytoplasmic central loop containing two short helices (CH1 and CH2) and the cytoplasmic C terminal tail with another short helix (CH3).

As predicted by electron microscopy of $MelB_{Ec}$ [Purhonen *et al*., 2005], both N and C-termini are in the cytoplasmic side of the membrane. The solved structure has four molecules (Mol: ABCD) (as shown in Fig. 1.13) in which Mol-A and Mol-C have a conformation different from Mol-B and Mol-D. Where, Mol-A is the state exhibiting a periplasmic facing conformation with a partially occluded internal cavity due to the interactions between helices I and VII in the N and C-terminal domains respectively.

This partially occluded conformation (Mol-A) is closed on the cytoplasmic side by interdomain contacts between inner Helix IV/X, V/VIII and II/XI. Substrates are not seen in the omit maps.

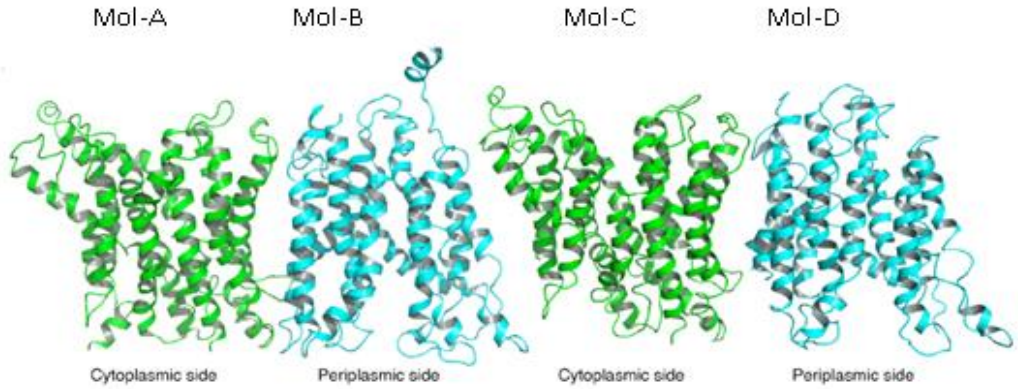


Fig 1.13: Overall fold of Mol-ABCD in the asymmetric unit shown in cartoon models [Ethayathulla *et al.,* 2014].

1.2.6. Substrate selectivity of MelB

As mentioned before, MelB has a unique feature in using $Na^+/Li^+/H^+$ as a coupling cation depending on the configuration of the transported sugar. Sugars like methyl-β-galactoside, lactose and TMG with a β-configuration are transported only with Na^+/Li^+ whereas sugars with a α-configuration can be co-transported with either $H^+/Na^+/Li^+$. A brief explanation of substrate selectivity is shown in Table 1.1 [Tsuchiya *et al.,* 1980, Wilson *et al.,* 1987]. Melibiose is a α-carbohydrate and it can be found in beans, legumes, seeds and soy-products. The structure of melibiose is as shown in Fig. 1.14.

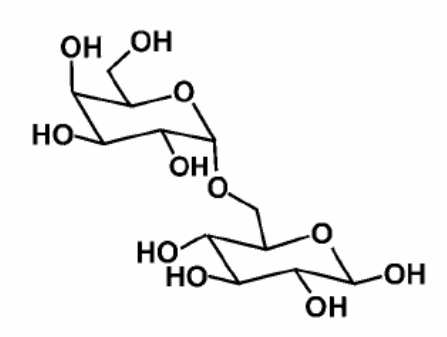


Fig. 1.14: Structure of α-D-galactopyranosyl (1-6) α-D-glucopyranose (Melibiose).

Substrates	$MelB_{EC}$	$MelB_{SY}$	$MelB_{KP}$	$LacS_{ST}$	$GusB_{EC}$
α-Galactosides:					
Melibiose	H^+, Na^+, Li^+	Na^+, Li^+	H^+	H^+	NT
ß-Galactosides:					
TMG	Na^+, Li^+, (H^+)	Na^+, Li^+	H^+, Li^+	H^+	NT
Lactose	Na^+, Li^+	ND	Li^+	H^+	NT
Monosaccharides:					
D-galactose	H^+, Na^+, (Li^+)	ND	ND	H^+	NT
α/β-Glucuronides	ND	ND	ND	ND	H^+

Table 1.1: Represents the cation selectivity of members of GPH family [Poolman *et al.,* 1996]

1.2.6.1 Cation binding sites in MelB

Indirect Evidences

Site directed mutagenesis has suggested that the carboxylate chains of Asp-19 (Helix I), 55, 59 (Helix II) and Asp-124 (Helix IV) participate in Na^+ coordination, since replacing each of these residues with neutral residues prevents the use of Na^+ or Li^+ as co-substrates [Ding and Wilson 2001b, Pourcher *et al.,* 1993]. Other mutations such as Y28F, N87A and Y120F exhibited reduced Na^+ affinity [Zani *et al.,* 1994]. Furthermore, mutagenesis of Pro-142 with Ser caused the MelB carrier to lose H^+ coupling [Kawakami *et al.,* 1988].

Mutant in $MelB_{EC}$	H^+ coupling	Na^+ coupling
Wild Type	+	+
D55C	*	-
D55E	+	*
D59N/C	-	-
D59E	+	+
D124N/C	-	-
D124E	*	+
K377C/V/D	-	-
D19C	-	+

Table 1.2: Shows the cation selectivity in MelB mutants. "+" indicates the ion coupling."-" indicates no coupling with the ions. "*" indicates that binding is observed but less than WT. Data were taken from Hama *et al.,* 1993, 1994, Botfield and Wilson 1988, Zani *et al.,* 1993 and Fuerst *et al.,* 2015].

Chimeric studies provided additional information regarding the cationic binding site. Hama and Wilson, 1993, constructed a hybrid MelB carrier by exchanging various

regions of *K.pneumoniae* MelB and *E. coli* MelB and showed that 81 residues of the *E. coli* protein are essential for Na^+ and melibiose transport, but only Asn-58 is important for Na^+ recognition.

Work done by Poolman et al., 1996, showed that the residues of the hydrophilic faces and cytoplasmic halves of Helix I, II and IV of MelB are part of the cation binding site. Table 1.2 shows the cation selectivity in MelB mutants before the evidence from crystal data.

Direct evidence from crystal data

$MelB_{ST}$ crystal data [Ethayathulla *et al.*, 2014] has shown that the three Asp residues at 55, 59 (Helix II) and 124 or 121 (Helix IV) form a trigonal bi-pyramidal geometry along with Tyr-120, Thr-121 (Helix IV) and Thr-373 (Helix XI) which is important for the cationic selectivity properties. From the PROPKA prediction it is shown that the Asp-59 is buried in a hydrophobic environment with a pKa of 9.0. The results are in agreement with the previous biochemical approaches.

1.2.6.2 Sugar binding sites in MelB

Indirect Evidences

Hydrogen bonding interactions of -OH groups with Asp, Asn, Glu and Gln along with the nonpolar interactions of the carbon backbone with aromatic residues in the initial binding of a ligand are an important form of interaction between the sugar and the protein. The binding of Na^+ clearly affects the binding of sugar by means of adjustment of the spatial structure to enhance the affinity for sugar [Abdel-Dayem *et al.,* 2003, Leon *et al.,* 2005]. It was suggested that the binding site of the cation is always close to the sugar binding site [Yamashita *et al.,* 2005, Faham *et al.,* 2008].

pCMBS inhibition studies revealed the importance of Helix XI residues (S380C, A381C, A384C, I387C, A388C, and L391C) indicating that they might lie close to the sugar binding sites [Ding and Wilson 2001a]. Cysteine scanning mutagenesis of Loop 10-11 also suggested that the residues Thr-373 and Val-376 lie close to the substrate binding site [Ding *et al.,* 2004]. Also, Asn-335 and Thr-338 from Helix X seemed to be involved in

sugar binding [Basquin 2001]. Spectroscopy studies suggested that D19C completely abolishes sugar binding but binding of cation remains unaffected. And, FRET studies showed that the Trp-64 lying in N-terminal Loop 2-3 is close to the sugar binding site [Cordat *et al.,* 1998].

Direct evidence from crystal data

A water filled cavity is observed between residues Asp-19 (Helix I), Tyr-120, Asp-124 and Trp-128 (Helix IV) which along with Arg-149 (Helix V), and Lys-377 (Helix XI) could be an important candidate for sugar binding (Fig. 1.15) [Ethayathulla *et al.,* 2014]

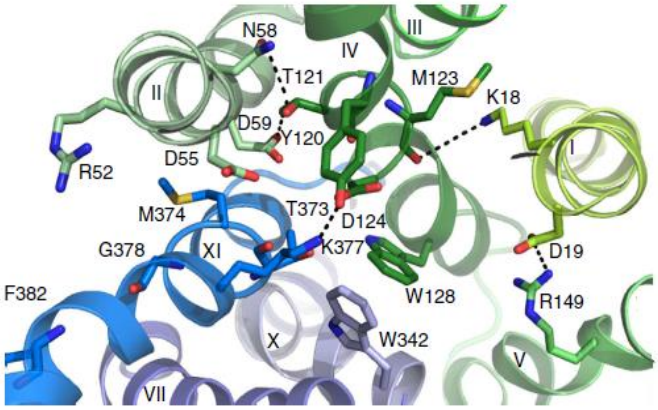


Fig. 1.15: Representation of Internal cavity in Mol-A: 4M64.pdb [Ethayathulla *et al.,* 2014].

Apart from this, a salt bridge between Asp-19 and Arg-149 and CH/π-interactions between Trp-128 and Tyr-120 with the pyranosyl ring of the sugar may play a critical role in sugar selectivity. The fact that Arg-149, when mutated to cysteine, largely remains in an inward facing conformation and is not capable of re-orienting its binding sites to the periplasmic side [Lin *et al.,* 2013] validates its importance in sugar binding.

Ethayathulla *et al.,* 2014 proposed that Na^+ binding leads to the movement of Helix IV, which connects both co-substrate sites. Asp-124 in Heix IV helps in optimizing the pyramidal shape of the cavity with Asp-55 and Asp-59 and aligns the Tyr-120 and Trp-128 for aromatic stacking, which increases the affinity for the sugar. Previous studies done by Cordat et al*.,* 2000 and Ganea *et al.,* 2011 suggested that the communication between cation and sugar binding sites involves Helix IV, Loop 4-5 and Helix V domains.

1.2.7 Ionic locks in MelB

The inside-closed structure of MelB exhibiting several multiple hydrophobic interactions on the cytoplasmic side is stabilized by interdomain electrostatic interactions in the cytoplasmic side, designated as ionic locks (L) as shown in Fig 1.16, involving three Arg residues: **(L1)** Arg-295 (Helix IX), **(L2)** Arg-141 (Helix V) and **(L3)** Arg-363 (Loop 10-11) (Fig. 20) [Ethayathulla *et al*., 2014].

Lock-1 (L1): Holds Helix V close to the C-terminal domain by forming several hydrogen bonds between Arg-295 and Gln-143 (Helix V) and Pro287 (Helix VIII).

Lock-2 (L2): Arg-141 stabilizes Helix X by forming four hydrogen bonded ion pairs with Asp- 351 and Asp-354.

Lock-3 (L3): Hold the N-terminal in outward facing conformation, Arg-363 forms ion pair and hydrogen bonding interactions with the backbone atoms of Val-204, Asp-208 (N-terminal side of the loop $_{6\text{-}7}$) and Gly-74 (loop $_{2\text{-}3}$).

Replacing any of these residues with cysteine resulted in a conformationally compromised mutant that failed to transport but retain the affinity for co-substrates binding.

Lock-4 (L4): To stabilize the outward facing conformation, Asp-35 (Helix I) forms salt bridge with Arg-175 (Helix VI) on the periplasmic side.

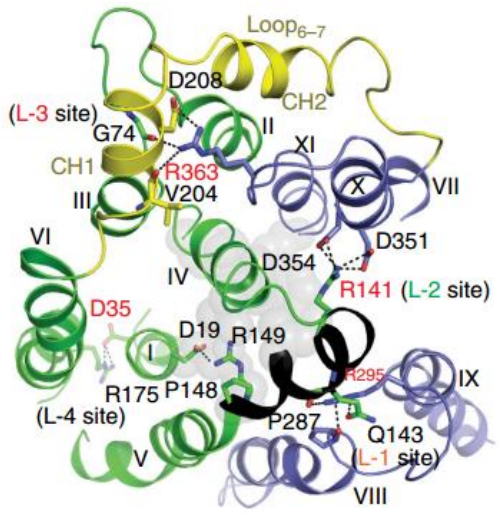


Fig. 1.16: Ionic locks in Melibiose Permease. Residues forming ionic locks are shown in sticks connected by the broken line [Ethayathulla *et al.*, 2014].

1.2.8 Alternating access mechanism in MelB

According to the model proposed by Jardetzky, called the "alternating Access mechanism ", the protein is open towards one side, binds the substrate, closes and re opens in the other side to release the substrate [Jardetzky 1966]. It seems that the alternating access mechanism is similar for most of the MFS permeases.

The mechanism of of Na^+/melibiose transport involves eight different steps [Ethayathulla *et al.,* 2014] (Fig. 1.17):

1. **Inward Na^+ bound state**: Na^+ binds to the cation binding site from the cytoplasm which induces the movement of Helix IV which further increases the affinity for the sugar.

2. **Inward Na^+ and Melibiose bound state:** Sugar binding initiates the locking process which results in the occluded state.

3. **Occluded State:** This is stage where both the cytoplasmic and the periplasmic sides are closed by the hydrophobic patch locks.

4. **Outward Na^+ and Melibiose bound state:** Interactions between the ionic locks such as Arg-295 and Arg-141 facilitate a conformational transition from an occluded to an outward state.

5. **Release of Melibiose:** Bound melibiose is released to the periplasmic side which is followed by the release of the cation.

6. **Release of Na^+:** Structural changes in the cation binding sites releases the cation.

7. **Occluded empty state:** Empty state occluded on both sides.

8. **Inward Empty state:** The state where cavity is open to the cytoplasm and the next transport event begins.

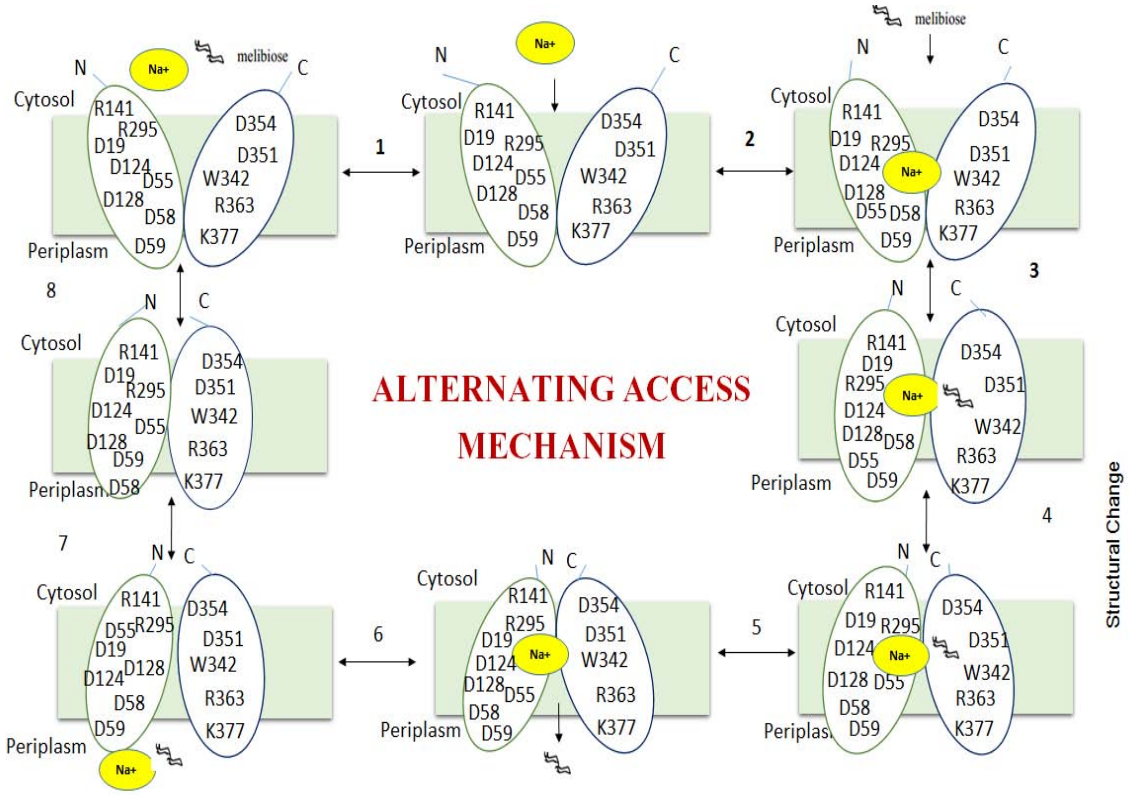


Fig. 1.17: Representation of Na^+/Melibiose symport showing the kinetic steps involved in the transport cycle. The residues shown here are important for the cation/sugar binding.

1.3. Substrate binding studies using biophysical techniques

1.3.1. Fluorescence Resonance Energy Transfer (FRET)

Insights into co-substrate induced structural changes of MelB can be obtained from FRET spectroscopy. Fluorescence detection is highly sensitive and is now a dominant methodology extensively used in biotechnology. The fluorescence signal of the protein is mainly from the emission of tryptophan at an excitation wavelength of 290 nm. Contributions from other aromatic groups are negligible.

The FRET is an interesting methodology, since the energy transfer occurs between the intrinsic tryptophan of MelB and the dansyl sugar Dns^2-S-Gal, a fluorescent analog of melibiose. Dns^2-S-Gal (D^2G) carrying a fluorescent group is a ligand of MelB. The D^2G probe is an indicator for polarity changes [Maehrel *et al.,* 1998]. FRET occurs if the emission spectrum of a donor (Trp) overlaps with the excitation spectrum of an acceptor (Dns^2-S-Gal), in close proximity less than 50 Å.

This method was useful in calculating the distance between donor and the acceptor [Cordat *et al.,* 1998]. It was shown that the Trp-64 in the N-terminal loop region and Trp-299 (Helix IX) and Trp-342 (Helix X) in the C-terminal side contribute most to the FRET [Cordat *et al.,* 1998]. In the presence of D^2G, sample excitation at 297 nm gives rise to a signal at 465 nm indicating energy transfer. In the WT, the signal of D^2G bound to MelB in sodium free medium is red shifted by up to 25 nm from that recorded in the presence of NaCl (Fig. 1.18) [Maehrel *et al.*, 1998].

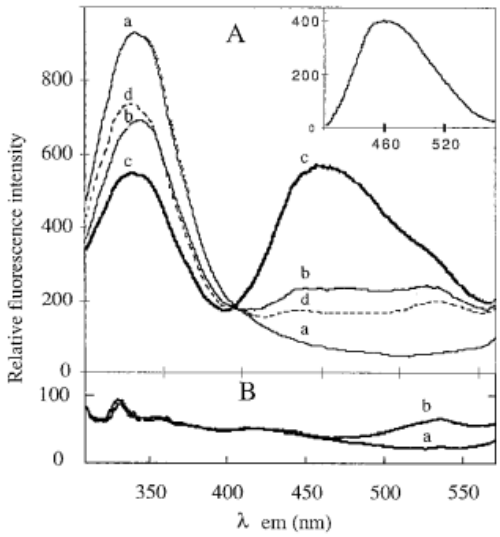


Fig. 1.18: FRET in WT Melibiose permease. **(A)** FRET of MelB, a. only protein , b. MelB with 15 µM Dns^2-S-Gal, c. MelB with 15 µM Dns^2-S-Gal + 10mM NaCl, d. Melb with 15 µM Dns^2-S-Gal + 10 mM NaCl +150 mM Melbiose. **(B)** Negative control studies only with E.coli lipids [Maehrel *et al.,* 1998].

1.3.2 Infra-red Spectroscopy

Infrared (IR) spectroscopy is a useful technique in the study of protein conformation and dynamics. The possibilities of the technique become apparent specially when applied to large proteins in turbid suspensions, as is often the case with membrane proteins. The properties of membrane proteins make difficult the use of structural techniques to study their conformation. Thus, even if X-ray crystallography is the method of choice for detailed protein structure analysis, obtaining three-dimensional crystals of membrane proteins is not easy.

Infrared spectroscopy is useful, either as an alternative option or as a complement to the high-resolution techniques in membrane protein studies, because of the wealth of information provided. Furthermore, the lipid environment does not perturb the resolution or sensitivity of the spectra and can also be the object of simultaneous study, thus providing information on the lipid-protein interactions. The increased sensitivity of modern infrared spectrometers has also allowed the use of sampling techniques other than transmission.

Thus reflection techniques, and particularly attenuated total reflectance (ATR), have been applied to study structural aspects of membrane proteins and to measure the orientation of membrane components through the use of polarized IR radiation [Arrondo *et al.,* 1999]. This sampling technique was developed by Harrick *et al.,* 1967 and later applied to membranes [Fringeli *et al.,* 1981].

ATR-FTIR spectroscopy is an emerging technology used in examining protein secondary structure and structural changes [Surewicz *et al.,* 1993]. It delivers useful insights on the structure and on the structural changes induced by binding of substrates to MelB.

1.3.2.1 ATR-infrared difference spectra of MelB

The attenuated total reflectance infrared difference spectroscopy permits the study of the interaction of substances with membrane proteins by the alternating perfusion of a buffer solution. The difference spectra gives information about the conformational change in all types of secondary structure α-helices, β-sheets and loops of MelB due to substrate binding [Leon *et al.,* 2006].

Fortunately, the substrates like sodium and melibiose don't introduce additional absorbance bands in the amide region between 1700 and 1500 cm^{-1}. pH of the buffer plays an important role and small difference could cause an inflation of proteoliposomes layers.

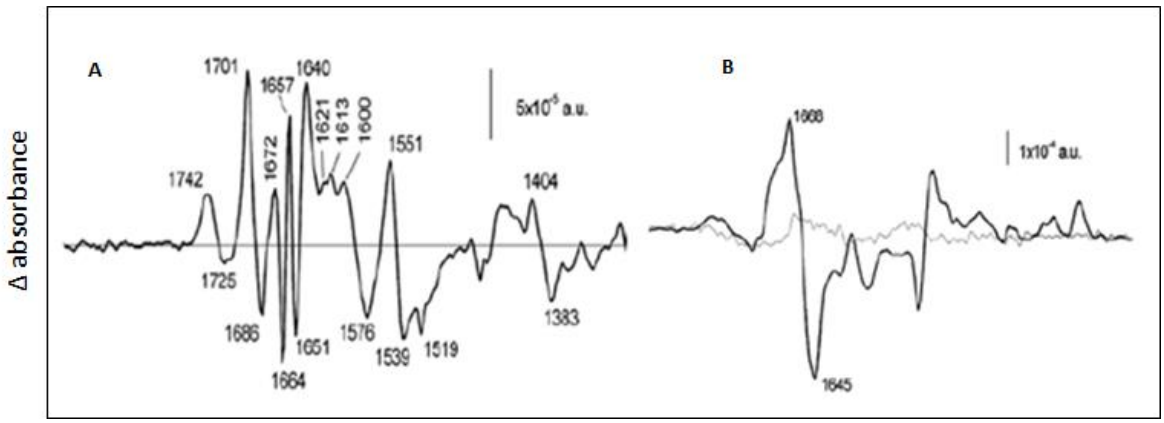


Fig. 1.19: Substrate induced secondary structural changes in melibiose permease. **(A)** Sodium induced IR_{diff} spectrum in Cys-less**. (B**) Melibiose dependent IR difference spectra in the presence of Na^+ in Cys-less. **Thin line:** Sacchrose induced IR difference spectrum [Leon *et al.,* 2005]

In the sodium dependent difference spectrum of Cys-less, the peaks at 1649, 1657 and 1663 cm^{-1} correspond to α-helix and peaks at 1553 and 1539 cm^{-1} belong to amide II region (Fig. 1.19 A). And, in the melibiose induced difference spectra peaks at 1666, 1659 and 1652 cm^{-1} belongs to α-helix; 1695, 1688, 1680 and 1672 cm^{-1} belongs to β-sheets and 1641 cm^{-1} to a reverse turn or 3_{10} helices (Fig. 1.19 B) [Leon *et al.,* 2005]

The difference spectrum of mutants was compared with Cys-less to know the effect of substrate binding and the conformational change in the mutant. Previous studies by Ding and his co-workers postulated that the mutation of residue Lys-377 was deficient in sugar transport. Recently, a study done by Fuerst *et al.,* 2015 using ATR-FTIR revealed that the mutants of Lys-377 were unable to bind both Na^+ and melibiose (Fig. 1.20).

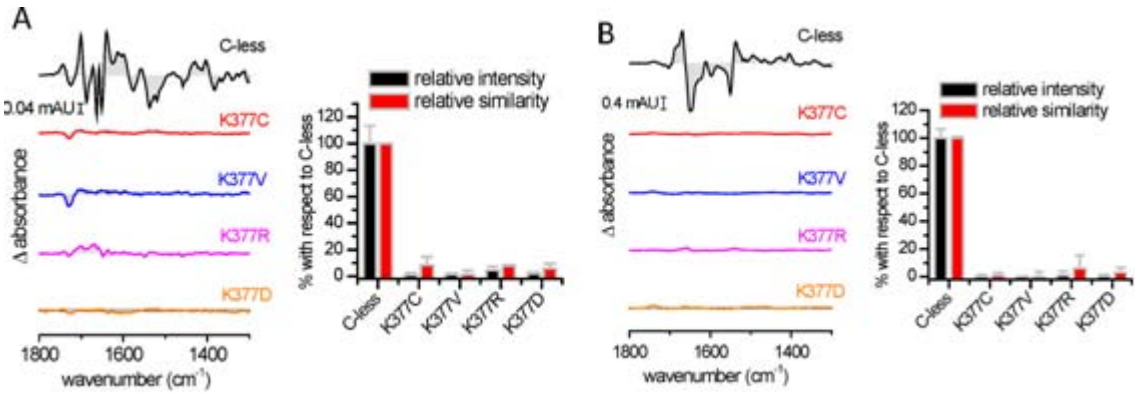


Fig. 1.20: Substrate induced infrared difference changes in Lys-377 mutants. **(A)** 50mM Na induced IR_{diff} spectra and its respective linear regression graph depicting spectral similarity and intensity. **(B)** 50mM Melibiose induced IR_{diff} spectra in the presence of 10mM Na^+ [Fuerst *et al.,* 2015]

A detailed study done by Granell *et al.,* 2010 showed the importance of several Asp residues in substrate-induced conformational changes (Fig. 1.21).

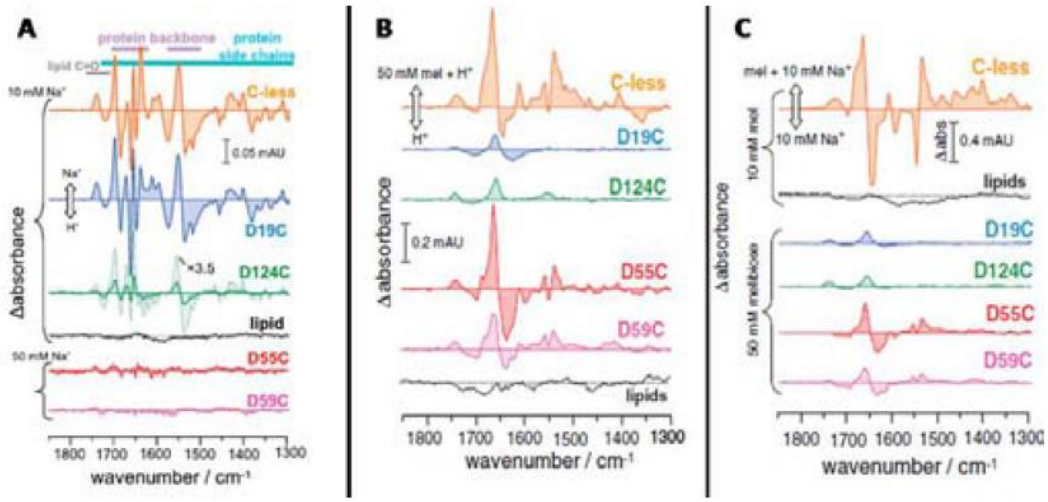


Fig. 1.21: IR_{diff} spectra for Melb mutants. **(A)** Sodium-induced IR_{diff}, **(B)** Melibiose-induced IR_{diff} in the presence of proton. Melibiose-induced IR_{diff} in the presence of sodium [Granell *et al.,* 2010]

1.4. Transport assays using radioactive sugar

Integral membrane proteins are important candidates for the potential drug target. And, the transport assays are essential to screen the drugs to target these proteins. In this study, homogeneity of the sample is more important than the quantity of the sample. In the case of melibiose permease, care was taken to prepare the sample without the presence of any cations. The transport assay experiments are usually three step process where: (i) the sample is prepared, (ii) initiation of the uptake by adding radioactive substrates and finally iii) terminate the uptake at desired time [Xie *et al.,* 2008].

The advantage of studying transport mechanism in intact cell is that the transport proteins are in its natural conformation and it has complete system to support transport. While the disadvantage is the presence of metabolic mechanism which can modify or degrade the substrate.

In the recent work done by Amin *et al.*, 2014, it has illustrated the role of three cytoplasmic Arg residues by transport studies (Arg-295 in Helix IX, Arg-141 in Helix V and Arg-363 in $loop_{10\text{-}11}$) (Fig. 1.22). Cysteine scanning mutagenesis of the three Arg residues has completely abolished the melibiose transport, while the double mutagenesis helped to slightly recover the transport process. Studying the role of Arg residues in MelB is highly beneficial in giving more information for other sugar transporters with structural similarity, such as the glucose transporter-1 (GLUT1). GLUT1 is a uniporter that facilitates the transport of glucose in the membranes of mammalian cells. Glucose transporter-1 (GLUT1) deficiency syndrome is an autosomal dominant haplo-insufficiency disorder, leading to a reduced glucose transport into the brain (Seidner *et al.,* 1998). This has been caused by 13 missense mutations particularly a novel missense mutation of Arg-333 to Gln.

In GLUT1, Arg-333 and Arg-153 are one of the crucial residues of the highly conserved amino acid motif Arg-X-Gly-Arg-Arg between Helix VIII/XI. These motifs are conserved in many of the members of the major facilator superfamily (MFS). Mutation of these residues will result in decreased glucose uptake in the brain [Leen *et al.,* 2010].

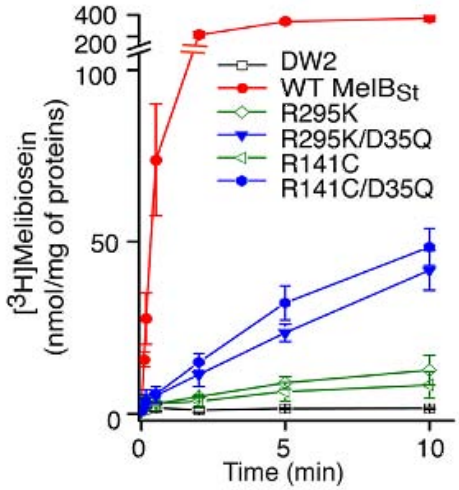


Fig. 1.22: [^{3}H] Melibiose transport assay were carried out in *E. coli* intact cells of two Arg residues in the cytolplasmic side along with the WT $MelB_{ST}$ as a positive control and DW2 as a negative control. [Amin *et al.,* 2014].

1.5. Molecular Dynamics simulations – A current state of art

Understanding the mechanisms of membrane proteins at the atomic resolution is still a challenge to current experimental conditions. The crystal structures often provide clues to unravel the working mechanism of membrane proteins, but proteins often show multiple conformational states during the biological processes and the structural information of different states of the same membrane protein is limited [Weng *et al.,* 2014]. Thus, molecular dynamics (MD) simulations are being combined with experimental studies to track the system behavior across a vast spatiotemporal domain- with atomic precision and timescales up to milliseconds. The field of membrane protein simulation has matured during recent years due to the increased computational facilities.

In the work done by Fuerst *et al.,* 2015, MD simulations of *E. coli* MelB embedded in a POPE bilayer were carried out using the coordinates of Model A of $MelB_{ST}$ (PDB code: 4M64) as a template. A general workflow of MD simulation is illustrated in Fig. 1.23. A good dynamic stability of the simulation was demonstrated by a RMSD of 2.2 Å at 2 ns (Fig. 1.24 right). A bidendate ionic bond 3.7 Å is revealed between the NZ atom of Lys-377 and the CG atom of either Asp-59 or Asp-55 in Helix II (Fig. 1.24 left). This study suggests that the role of Lys -377 is to stabilize the inner region of MelB by compensating the repulsive interactions between the anionic chains of Asp-55 and Asp-59. In this sense, earlier studies have shown the importance of Asp-55 and 59 in ligand binding and the inability of the mutants to bind the cation. These results also postulated that the Lys-377 could act as a ligand to the sugar.

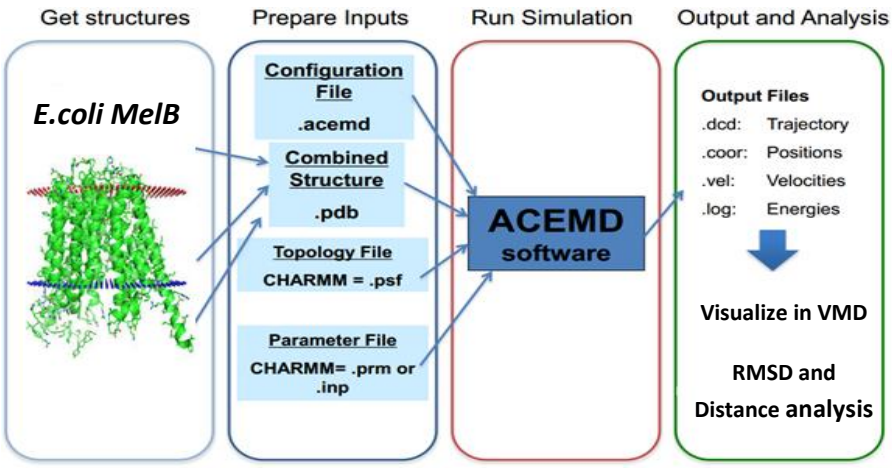


Fig. 1.23: Illustrates the general workflow of molecular dynamics simulation.

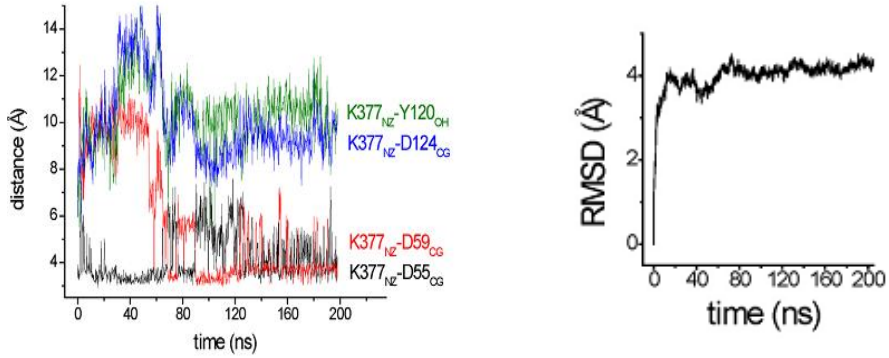


Fig. 1.24: (Left) Time course of the specified distances between residues interacting with Lys-377. It is evident that Lys-377 remains close to Asp-55 and Asp-59 and moves away from Asp-124 and Tyr-120 during the course of simulation. **(Right)** Root mean square deviation of the structure from initial to final stage of the simulation [Fuerst *et al.,* 2015].

2.OBJECTIVES

Why study Melibiose Permease?

Membrane proteins play a vital role in maintaining the physiology of the organism. An important subset of these proteins comprises active transport proteins, which use energy to transport the substrate against its concentration gradient. An outstanding example is the sugar transporter proteins. These proteins use the energy from ion gradients (usually sodium) or proton to transport the sugar molecules into the cytoplasm. It is therefore, a cation-sugar co-transport. The defect in the expression and/or function of membrane co-transporters leads to the manifestation of numerous clinical disorders. So, it is very important to study the structure, function and dynamics of membrane proteins.

The melibiose permease (MelB) of E. *coli,* is a particularly interesting protein for studying the transport mechanism, as it transports a variety of sugars into the cell catalyzed by an ion-motive-force. Interestingly, this transporter is capable of using three different cations which makes it more versatile compared to other members of the same family. Furthermore, studies carried out in this can offer some data to understand the functioning of homologous transport proteins in humans, which are difficult to study due to the complexity of purification.

Although a crystal structure of MelB *Salmonella typhimurium* and an important amount of data on E. *coli* MelB has been published, the molecular detail of the transport mechanism and the involvement of important parts of the protein is still unknown. Thus, the main objective of this thesis is to study the role of glycine rich periplasmic loop 5-6 residues along with a part of Helix V and Helix VI in E. *coli* MelB using an experimental and computational approach. This general objective was planned to be achieved through the following specific objectives:

1. Preparation of the following point mutants: L164C, P165C, F166C, and V167C in **Helix V**; N168C, Y169C, V170C, G171C, G172C, G173C, and D174C in **loop5-6***;* R175C, G176C, F177C, and G178C in **Helix VI***.*

2. Purification, reconstitution and preparation of native vesicles for the above mutants.

3. Characterization of substrate binding using Fluorescence and FT-IR spectroscopies.

4. Transport assay using radioactive melibiose sugar.

5. Finally, to study the atomic details of MelB conformational changes using Molecular Dynamics simulations (MD).

3. MATERIALS AND METHODS

3.1 Bacterial Strain and Plasmids

The full length MelB DNA containing 473 amino acids with six histidine tag at the end of C-terminus from *E. coli* K-12 strain, was inserted into the vector pK95ΔAHB (a derivative of pKK223-3 (Pharmacia, Uppsala, Sweden). The plasmid vector had the ampicillin resistance tag to prevent additional bacterial contamination and also the wild type or mutated MelB gene, under the control of the Tac promoter [Mus-Veteau *et al.,* 1996] (Fig. 3.1). This was transformed into E.*coli* DW2R (recA$^-$) strain (melA$^+$, ΔmelB, ΔlacZY) [Botfield *et al.,* 1988] which was devoid of internal melibiose and lactose permease, but contains α-galactosidase.

Competent cells were prepared using the calcium or rubidium chloride method and plasmid purification was implemented using the Qiagen Miniprep Kit (Qiagen, Hilden, Germany). Simultaneously, the plasmid was also transformed into the shuttle strain DH5α after the sequence analysis to maintain the transformation efficiency.

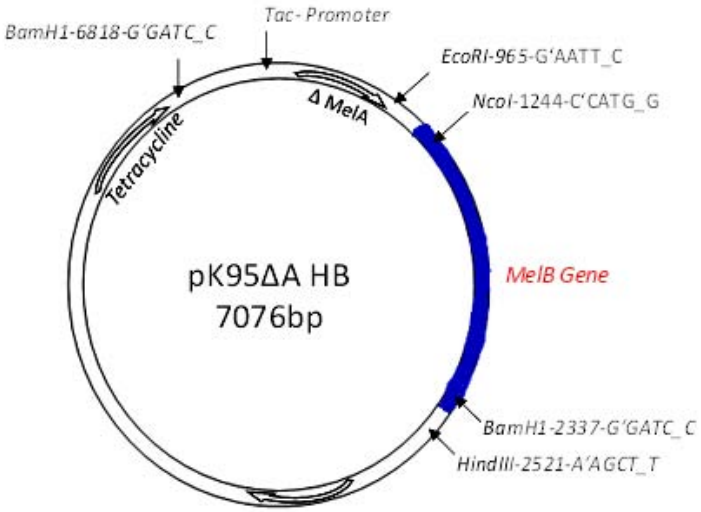


Fig. 3.1: pK95ΔAHB plasmid containing MelB gene.

3.1.1 Cysteine-scanning mutagenesis

The mutants described in this study were constructed using Cys-less MelB, which is devoid of its four native Cysteines, namely, Cys-235 mutated to Valine and Cys-110,310, 364 to Serine [Weissborn *et al.*, 1997] as a genetic background. The oligos of 27 to 30 nucleotides were designed using the PRIMERX online software tool and were ordered from Invitrogen^{TM}. The mutants were obtained using the polymerase chain reaction by running two single reactions in parallel [Edelheit *et al.*, 2009].

25 µL of forward and reverse reaction mixture containing 500ng of plasmid DNA along with 1X buffer solution (20 mM Tris-HCl (pH 7.5), 0.1 mM of dNTP's, 0.75 mM $MgCl_2$, and 0.5 µM forward-reverse oligos were added to 0.2 mL tubes. Shortly before starting the PCR reaction 0.02 units of KOD polymerase (Novagen, Merck, Germany) was added and hot-start activated. The samples were carefully placed in the PCR thermos cycler (VWR, USA) and run with the following program cycle as indicated in Table 3.1:

1.	Initial denaturation	95˚ C for 5min.
2.	30 cycles of DNA denaturation and primer annealing	95 ˚ C for 50sec 55 ˚C for 50sec
3	DNA Synthesis	68 ˚ C for 7.5min

Table 3.1: Steps followed in two-sided parallel PCR reaction.

At the end of the PCR reaction, the two parallel reaction outputs were mixed carefully (50µL in total) and heated to 95 ˚C for 5min to separate the PCR product from the template. It was then further cooled down to 37 ˚C for reannealing the denatured plasmid templates and PCR products [Edelheit *et al.,* 2009].

Finally, the end product was incubated at 37 ˚C with Dpn1 enzyme to digest the undesired methylated parental plasmid. The plasmid was kept at 4 ˚C until it was transformed into the E.coli DH5α strain. The purpose of DH5α transformation is to enlarge the plasmid quantity and to methylate the plasmid. The direct transformation into DW2-R will lead to insufficient PCR generated DNA. For sequencing, two to three single colonies were picked from each mutant plate and cultured in small volumes of 10mL LB medium at 37 ˚C, overnight. It was followed by spinning down the cells at 4000g for 10min for the plasmid extraction using Qiaprep® Miniprep Kit, and the concentration was checked by UV-Visible spectrophotometery.

Around 100ng of plasmid were used for DNA sequencing and it was carried out at CRAG genomics, UAB, Barcelona. The Table 3.2 shows the residues mutated in TM helix of Melibiose permease.

Mutation	Position in TM Helix
Leu164-Cys	Helix V
Pro165-Cys	Helix V
Phe166-Cys	Helix V
Val167-Cys	Helix V
Asn168-Cys	Loop V-VI
Tyr169-Cys	Loop V-VI
Val170-Cys	Loop V-VI
Gly171-Cys	Loop V-VI
Gly172-Cys	Loop V-VI
Gly173-Cys	Loop V-VI
Asp174-Cys	Loop V-VI
Arg175-Cys	Helix VI
Gly176-Cys	Helix VI
Phe177-Cys	Helix VI
Gly178-Cys	Helix VI
Asp35-Cys	Helix I
Arg175/Asp35-Cys	Helix VI/Helix I

Table 3.2: Site directed mutagenesis in MelB. The residues were individually changed to Cysteine in each C-DNA. Mutation# refers to three letter amino acid changed to desired cysteine. TM- transmembrane

3.1.2. Transformation and melibiose fermentation

In this study, the plasmid pK95ΔAHB, containing 100ng of wild type or mutant DNA was transformed into *E. coli* DH5α or DW2R cells using the heat shock protocol by Mandel and Higa, 1970. The cells along with the plasmid DNA were incubated for 30 min on ice, followed by incubation at 42˚C for 1.5 min followed by 3 min on ice. 900µL of LB medium were kept on ice. The plasmid DNA was added to the LB medium and incubated on a shaker at 37˚C for 60 min. The cells were then centrifuged at 4000g for 10 min, after which 700µL of the supernatant was discarded and the cells were re-suspended with the remaining 200µL of the medium. Finally, this suspension was plated on to either LB medium containing 0.1 mg/mL Ampicillin (DH5α) or MacConkey agar plate containing 0.1 mg/mL ampicillin, 0.1 mg/mL tetracycline, 10mM Melibiose (DW2R) and incubated at 37 ˚C/30 ˚C overnight (15 h).

MacConkey agar is a selective and differential media used for the isolation and differentiation of non-fastidious gram-negative rods. Gram-negative bacteria growing on the media are differentiated by their ability to ferment the sugar (melibiose), this is indicated by bright red colonies on the plate which mean that melibiose disaccharide is metabolized into glucose and galactose by the α-galactoside encoded by melA resulting in a pH change. Gram-negative bacteria that grow on MaConkey agar but do not ferment sugar appear colorless on the medium and the agar surrounding the bacteria remains relatively transparent (Fig. 3.2).

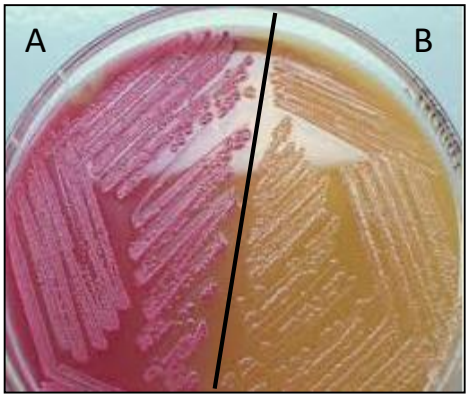


Fig. 3.2: MacConkey Agar plate. **(A)** Shows the colonies observed in expressing the WT of Melibiose Permease. **(B)** The colonies observed expressing the D35C mutant after incubation for 14h at 30 ˚C.

3.1.3. Growth of cells in Minimal Media-M9

After incubation, plates with colonies were preserved at 4 ˚C until use. A single colony was selected and grown overnight in 12 mL of LB medium with 0.1 mg/mL Ampicillin and 0.1 mg/mL tetracycline at 30 ˚C at 200 rpm for 14h. The mini culture medium was added to 400 mL of M9 minimal medium (1X), pH 7.5, along with 0.1 mg/mL Ampicillin, 1 mM $MgSO_4$, 0.1 mM $CaCl_2$, 0.34 M thiamine-HCL and incubated at 30 ˚C, 200 rpm for 7h. The midi culture was further diluted 20-fold times into the final large 8 litre cultures of M9 media containing 100 mg/mL Ampicillin, 1 mM $MgSO_4$, 0.1 mM $CaCl_2$, 1 mM thiamine-HCL , and grown at 30 ˚C overnight, under continuous shaking at 200 rpm until the OD_{600nm} reaches 1.8 to 2.0. Cells were harvested by centrifugation at 3700 g and washed twice with the re-suspension buffer (Box. 1) to remove any cell debris**.**

M9 Medium 10X (1L):

7.5 M Na_2HPO_4, 3.5 M KH_2PO_4, 0.5 M NaCl, 1M NH_2Cl, 54mM Glycerol, 2 M Casamino acid pH 7.5.

Resuspension Buffer (1L):

50mM Tris-HCl, 50mM NaCl, 5mM β-mercaptoethanol pH 8.0.

*M9 media is the primary source of carbon.

Box. 1: Buffers used for growth and resuspension of the cells.

3.2 Membrane Protein Purification

It is divided into two stages:

i) Preparation of Inverted membrane vesicles.

ii) Solubilization, purification of MelB using Ni^{2+}-NTA resin and reconstitution:

3.2.1 Preparation of Inverted Membrane Vesicles (IMV)

The harvested cells of wet weight of approximately 20-25g were re-suspended using resuspension buffer thoroughly. The suspended pellet was incubated in the cold room (4˚C) with 1 mM lysozyme and 15 mM EDTA (dissolved in water and NaOH respectively) with continuous stirring for 30 min. This step was carried out to destabilize the outer membrane and it was followed by the addition of 20 µg/mL DNAse, RNAse and 15mM MgSO4 to the solution and stirred for 5min at 4 ˚C. Usually longer incubation after the addition of DNase and RNase was avoided since it made the suspension sticky. The micro-fluidizer (Model 110S) was washed thoroughly with ice cold MilliQ water and re-suspension buffer and the geometrically fixed interaction chamber were maintained cold using ice packs. The membranes were disrupted by passing through a micro-fluidizer at 25,000 psi for three times and the Inverted Membrane Vesicles (IMV) was

collected by centrifuging at 100,000 g for 60 min. Finally, the supernatant was discarded and the tubes containing the inverted membrane pellet was re-suspended in a volume ratio of 1:1 (Pellet: Buffer-1 2X) (Box 2) and stored at -80 ˚C for a maximum of one week.

Buffer 1 (2X):
20 mM Tris-HCl, 1.2 M NaCl, 10 mM β-mercaptoethanol, 20 %(v/v) glycerol.

Box. 2: Buffers used for resuspension of the membrane pellets.

3.2.2 Solubilization and Purification of MelB using Ni resin

The frozen cells were thawed and the IMV suspension was distributed into four 35mL centrifuge tubes. The tubes were filled with Buffer 1 (1X) and centrifuged at 100,000 g for 60 min. The list of buffers used can be found in Box 3. The supernatant was discarded and the pellet was re-suspended with an approx. 20 mL of Buffer 9/10 and solubilized by adding 1% DDM (n-Dodecyl β-D Maltoside) and 0.1 mM AEBSF protease inhibitor. This protein-detergent-lipid mixture was kept on the orbital shaker for 2 hr at 4 ˚C and the formation of detergent micelles were visually observed by the gel like aspect of the suspension. 20 µL aliquots were collected for SDS PAGE analysis in each step. The mixture was centrifuged at 100,000 g for 60min. In the meanwhile, 1 mL of Ni^{2+}-NTA resin was washed with MilliQ (6 times), Buffer-1 1X (3times) and equilibrated in Buffer A until use. After the centrifugation, the supernatant containing the membrane protein was preserved and incubated with 1 mL of Ni^{2+}-NTA resin at 4 ˚C for 1 hr in an orbital shaker to remove any insolubilized protein.

This solution was then transferred to 50 mL falcon tubes and centrifuged at 3700g for 15 min and the Ni^{2+} pellet was washed twice with Buffer A. This membrane protein bound with Ni resin was loaded into a 15 mL chromatography column and washed again with 15 mL of Buffer A. This can be followed by exchange of membrane protein to detergent of choice with 5 column volumes of Buffer B. But, in this study membrane protein was solubilized and eluted with β-DDM throughout. The column was washed with 5 columns volume of Buffer C to reduce the concentration of Na^{+} and finally the protein was eluted with Buffer D containing 100mM Imidazole in 1.5 mL tubes.

The concentration of protein was calculated from UV/Vis spectrum using the Equation 1. Fractions containing > 0.15 mg/ml were pooled and 20 µL aliquots were collected for SDS purity check. The whole process involving the purification of membrane protein is illustrated in the Fig 3.3.

$$\text{Protein Concentration (mg/mL)} = \frac{Abs_{280} * \text{Total volume of the aliquots}}{\text{Extinction Coefficient of MelB}}$$

$$\text{Extinction Coefficient of MelB is } 1.5.mol^{-1}.cm^{-1}$$

Equation 1: Shows the formula to calculate protein concentration.

Buffer 9/10 2X:

20mM Tris-HCl, 600mM NaCl, 5mM β-mercaptoethanol, 10 %(v/v) glycerol, 10mM Imidazole, 10mM Melibiose pH 8.0

Buffer A:

20mM Tris-HCl, 600mM NaCl, 5mM β-mercaptoethanol, 10 %(v/v) glycerol, 10mM Imidazole, 10mM Melibiose pH 8.0, 0.2% (w/v) DDM pH 8.0.

Buffer B *:

20mM Tris-HCl, 600mM NaCl, 5mM β-mercaptoethanol, 10 %(v/v) glycerol, 10mM Imidazole, 10mM Melibiose pH 8.0 , 0.1% (w/v) detergent of choice pH 8.0.

Buffer C:

20mM Tris-HCl, 100mM NaCl, 5mM β-mercaptoethanol, 10 %(v/v) glycerol, 10mM Imidazole, 10mM Melibiose pH 8.0 , 0.1% (w/v) DDM pH 8.0.

Buffer D:

20mM Tris-HCl, 100mM NaCl, 5mM β-mercaptoethanol, 10 %(v/v) glycerol, 100mM Imidazole, 10mM Melibiose pH 8.0 , 0.1% (w/v) DDM pH 8.0.

Box. 3: Buffers used for protein solubilization (Buffer 9/10), wash and detergent exchange (Buffer A, B, C), and elution (Buffer D).

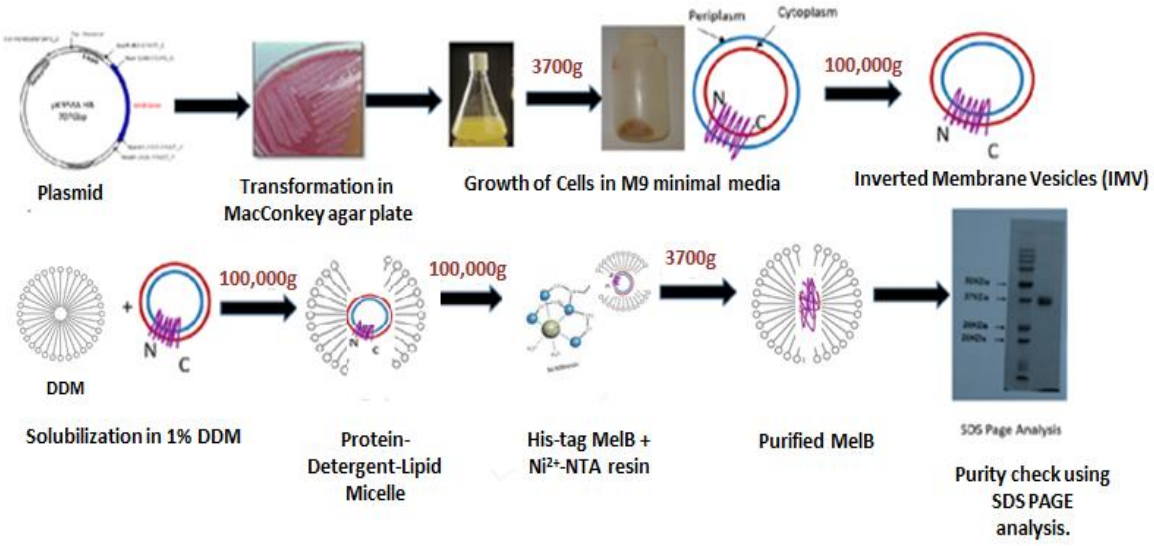


Fig. 3.3: The scheme illustrates the stages involved in the process of purification.

3.2.3 Reconstitution with bio-beads

Around, 30 mg of E.coli total lipids (Avanti Polar lipids Inc. Alabaster, AL, USA) was weighed and dissolved in chloroform with the constant supply of nitrogen air to prevent any oxidation of lipids. Bio-beads SM-2 (Bio-Rad, Hercules, USA) were washed six times with MilliQ water and three times with reconstitution buffer. Evaporation of chloroform in a rotary evaporator (Heidolph, Germany) resulted in a thin film on the small round bottomed flask. This was followed by the addition of reconstitution buffer (20mM MES, 100mM KCl, pH 6.6) to the flask and the mixture was vortexed for 10 min at room temperature.

These lipid-buffer mixtures containing large multi-lamellar vesicles were disrupted by 25 times extrusion using 0.1 µm (100 nm) polycarbonate membranes to get homogeneous uni-lammelar vesicles. The lipid vesicles were added at the ratio of 2:1 (lipid: protein) to the purified protein with 0.1% detergent. This mixture was incubated at 4 ˚C for 10min, followed by the three subsequent additions of bio-beads to the final concentration of 180 mg/mL (Fig. 3.4).

This mixture was incubated at 4 ˚C overnight, for the complete removal of detergent. The bio-beads were then separated by gravity in a column and flow-through containing the proteoliposomes was centrifuged at 310,000 g for 30min.

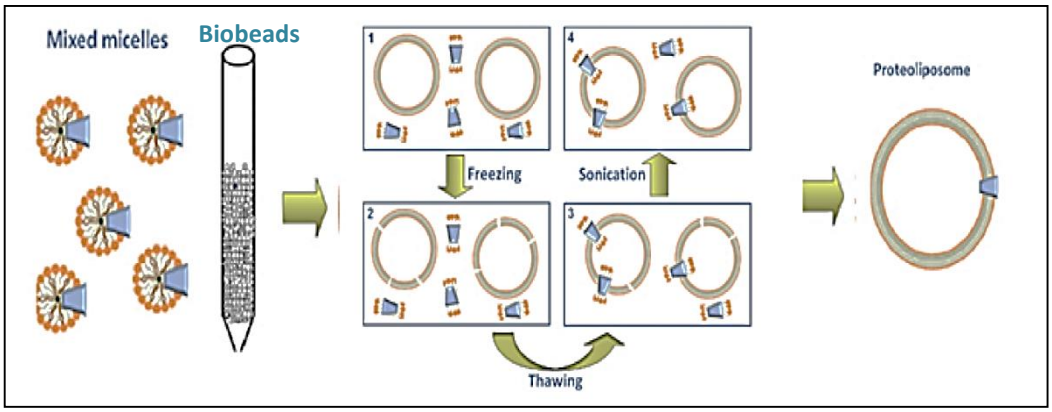


Fig 3.4: Illustration of membrane protein reconstitution process. **(A**) Mixed micelles containing lipid, protein and detergent complex was mixed with biobeads. **(B)** Freeze-thaw-sonication cycles to remove any Na+ contamination. **(C)** protein with liposomes: Proteoliposmes [Scalise *et al.,* 2013]

Pellets were re-suspended in 1mL of re-constitution buffer, flash frozen in liquid nitrogen and thawed immediately followed by sonication in an ultrasonic cleaner (Fungilab, Keyland Court, NY , USA) for two times 20s each. This was again centrifuged at 310,000g for 7min and this procedure was repeated once again. With this procedure, proteoliposomes samples can be free of both the substrates (Sodium and Melibiose).

Finally, the pellet was re-suspended in the buffer to obtain a final lipid concentration of 20 mg/mL. Aliquots were collected for SDS PAGE analysis to check the purity of the sample. Usually, the samples were used for the fluorescence studies before storage at -80 ˚C.

3.3 Preparation of Native Inside-Out Vesicles

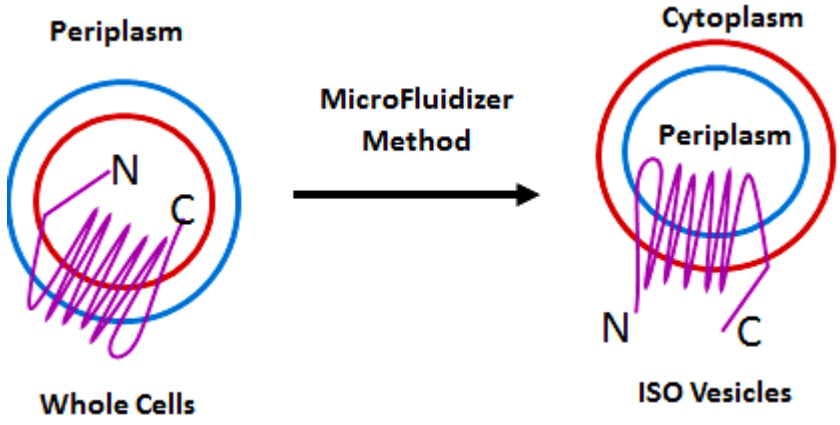


Fig. 3.5: Illustration of native membrane vesicles preparation

The Inside-out Vesicles were prepared as using the high pressure method as described in [Pourcher *et al.,* 1995]. The 3 g of frozen cells were thawed and washed thoroughly with the wash buffer (Box 4) at 4000g for 12 min. The pellets were re-suspended again with the same buffer and 15mM EDTA and 1 mg/mL lysozyme were added to the final concentration and the suspension was stirred vigorously for 30 min at 4 ˚C. Before passing the sample through the micro fluidizer, 20 µg/mL of DNase and of RNase were added to the suspension. The cells were broken using the micro fluidizer. This was followed by centrifugation of the suspension at 4000g for 10min at 4 ˚C, and, the supernatant containing the ISO was collected carefully and ultra-centrifuged for 100,000g for 60 min. This process was repeated two times and each time the pellets were re-suspended in the re-suspension buffer (Box. 4). The quantity of MelB protein was determined using Western blot with Histidine-Tag-specific reagent His-Probe^{TM}-HRP. A brief explanation of the process is shown in Fig. 3.5.

Wash Buffer:
50mM Tris pH8.0
Re-Suspension Buffer:
0.1M Kpi, 10mM K-EDTA pH 7.0

Box 4: Buffers used in Native Inside-Out Vesicles preparation.

3.4. Fluorescence Resonance Energy Transfer (FRET)

During the past 20 years, there has been a remarkable growth in the use of fluorescence techniques in biological sciences. Right now, fluorescence is a dominant methodology used in biochemical and biophysical studies. FRET is a fluorescence technique that monitors the distance between the fluorescent probes that are attached to the macromolecules. It can be useful in observing interactions between molecules or conformational changes within the same molecule. The mechanism behind FRET is the transfer of energy from the exited state of the donor to the exited state of the acceptor when they are in close contact. This process occurs whenever the emission spectrum of the Donor (D) overlaps with acceptor spectrum (A) as shown in Fig. 3.6. The efficiency of energy transfer for a single donor-acceptor pair at a fixed wavelength is given by equation 2:

$$E= 1/1+(r+R_0)^6, \quad \textbf{(eq.2)}$$

Where r is the distance between the fluorophores and R_0 is the distance at which the efficiency of energy transfer is 50 %. [Thevenin and Lazarova 2012].

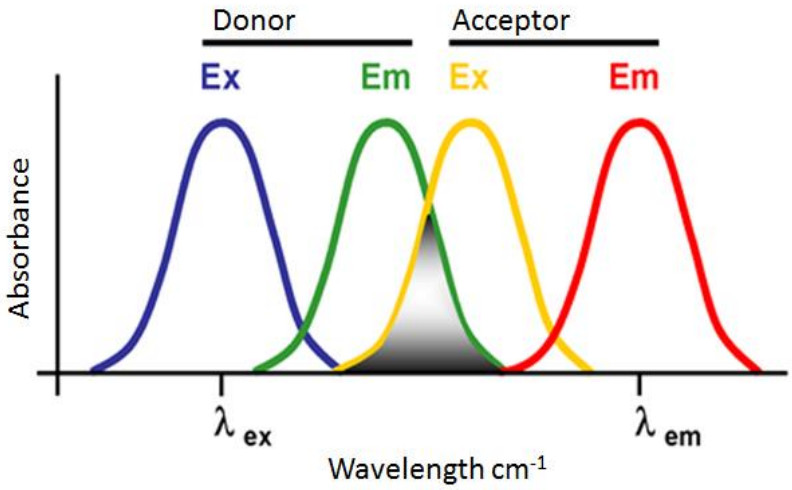


Fig. 3.6: Diagram of the overlapping spectra of a pair of FRET donor and acceptor.

(Source: Wikipedia)

In this study, the fluorescent sugar 2'-(n-dansyl) aminoethyl-1-thio-β-D-galactopyranoside (Dns^2-S-Gal or D^2G) was used as a fluorescent acceptor and the energy transfer is observed between Trp residues (donor) in the protein to D^2G.This sugar was a generous gift from Dr. Gerard Leblanc. This fluorescent dansyl-group has its excitation maximum at 335 nm where no amino acid absorbs and emits at 520 nm. D^2G binds firmly to the sugar binding site and inhibits the transport. Therefore, it acts as a competitive inhibitor to the natural Melibiose.

3.4.1. FRET in proteoliposomes and native vesicles

FRET signal in proteoliposomes and native vesicles was measured by setting the Trp excitation to 290 nm and recording the emission from 310 nm to 570 nm in a 1 cm quartz cuvette (Hellma GmbH & Co. KG, Germany) with a QuantaMaster^{TM} spectrofluorometer and data was processed with the Felix 32 software (Photon Technology International).

In all the measurements, proteoliposomes samples were sonicated for two times 10 sec each to decrease the aggregation. All the measurements were carried out at 20˚C. Before proceeding to the FRET measurements, an initial blank scan was measured using buffer (0.1 M KPi, 0.1 M KCl pH7.0) alone. If any peak apart from Raman spectrum was observed, the cuvette was washed thoroughly with the Helmanex cuvette solution. And after that, 20 µg of the samples were added to the cuvette containing the buffer, followed by acquiring protein emission spectrum, after that 16 µM D^2G was added to the MelB proteoliposomes suspension and 10 mM of NaCl. Finally 10mM melibiose was added subsequently to the suspension to check the D^2G displacement.

3.5 Infrared Spectroscopy

The term Infrared refers to the electromagnetic spectrum falling between 0.7 µm to 1000 µm. However, the region between 2.5 µm to 25 µm is primarily used in biophysical studies (4000 to 400 cm^{-1}). It is a useful technique to study the structure of the proteins [Goormaghtigh *et al.,* 1994a; Tatulian *et al.,* 2003]. One of its advantages over other spectroscopic tools is the absence of problems related to light scattering and does not require any external probe. There are three generations of IR spectrometers:

1. The first generation IR spectrometer invented in 1950s utilized prism as the monochromator and is no longer in use due to its several disadvantages like: narrow scan range, poor repeatability and strict use of sample's water content.
2. The second generation IR spectrometer introduced in 1960s used gratings as the monochromator and it was much better than the first generation. But, it still had disadvantages like: low sensitivity, low scan speed and poor wavelength accuracy which made it out of date.
3. The third generation IR spectrometer, Fourier Transform infrared spectrometer, uses the concept of interferometer. This is exceptionally powerful and is used very commonly.

3.5.1 Fourier Transform- Infrared Spectroscopy (FT-IR)

An FT-IR uses a system called Michelson interferometer to collect a spectrum. The Michelson interferometer consists of two mirrors, source, beam splitter, a laser and a detector. The energy from the source goes from the source to the beam splitter which splits the beam into two parts. One part is transmitted to a moving mirror; one reflected to a fixed mirror where the optical path is constant, as shown in Fig. 3.7. The moving mirror moves back and forth at a constant velocity.

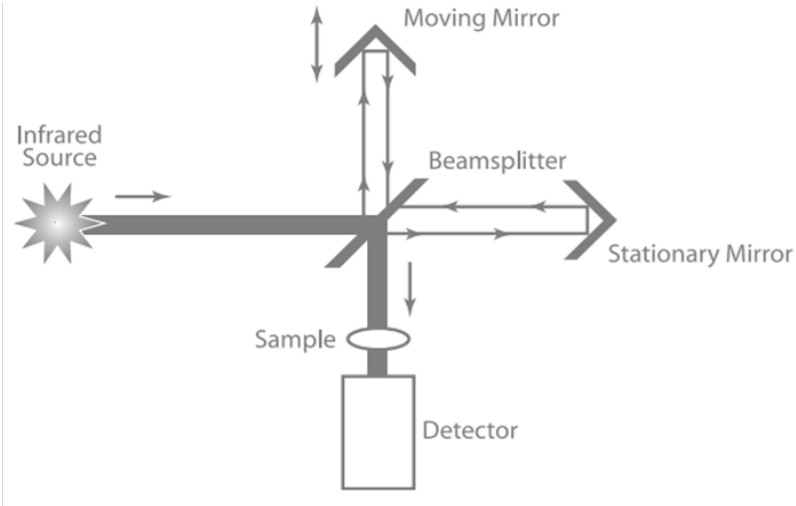


Fig. 3.7: Schematic representation of Michelson interferometer.

(Source: Thermofischer)

The two beams are reflected from the mirrors and recombined at the beamsplitter. The beam from the moving mirror has travelled a different distance than the beam from the fixed mirror. When the beams were combined a constructive and destructive interference pattern was generated known as interferogram. The interferogram then goes from beam splitter to the sample, where some energy is absorbed and some is transmitted. The transmitted portion reaches the detector which is usually cooled by liquid nitrogen.

To obtain the infrared spectrum, the detector signal is sent to the computer, and an algorithm called a Fourier transform is performed on the interferogram to convert it into a single beam spectrum. A reference or "background" single beam is also collected without a sample and the sample single beam is ratio-ed to the background single beam to produce a transmittance or "%T" spectrum. This transmittance spectrum can be converted to absorbance by taking the negative $\log_{10}$ of the data points as given in equation 3.

$$A = -\log_{10} T = -\log_{10}\left(\frac{S_{sample}}{S_{reference}}\right) \quad \textbf{(eq. 3)}$$

The following are the advantages of Fourier Transform infrared spectroscopy:

1. Many scans can be combined in a shorter time than one scan on a first generation spectrometers. This results in faster data collection.
2. It is mechanically simple with only one moving mirror.
3. It provided precise measurement method which requires no external calibration.

3.5.2 FTIR-attenuated total reflectance (ATR) spectroscopy:

Attenuated Total Reflection – Fourier Transform Infrared (ATR-FTIR) Spectroscopy was introduced in 1960. ATR, works by directing the IR beam into a crystals of relatively high refractive index. The crystals used are generally made of elements such as, germanium or zinc selenide. The IR beam reflects from the internal surface of the crystal and creates an evanescent wave, which projects orthogonally into the sample in contact with the ATR crystals. This so called evanescent wave has the same frequency as the incoming light, but the amplitude of the electric field decays exponentially with the distance from the interface. Some of the energy of the evanescent wave was absorbed by the sample and the reflected radiation is returned to the detector. The refractive index of the crystal and the sample is the important criteria in ATR-FTIR and is given by the equation 4,

$$\theta_c = \sin^{-1}(n_2/n_1) \quad \textbf{(eq. 4)}$$

Where n_2 is the refractive index of the sample, n_1 is the refractive index of the crystal and θ_c is the critical angle.

A good or pure ATR spectrum is observed when we exceed the critical angle. In ATR, optical properties of the sample and the crystal, the wavenumber of the light and the number of reflections determine the penetration depth as indicated in Fig 3.8 (i.e., how far the IR electromagnetic field penetrates on average in the sample in each reflection).

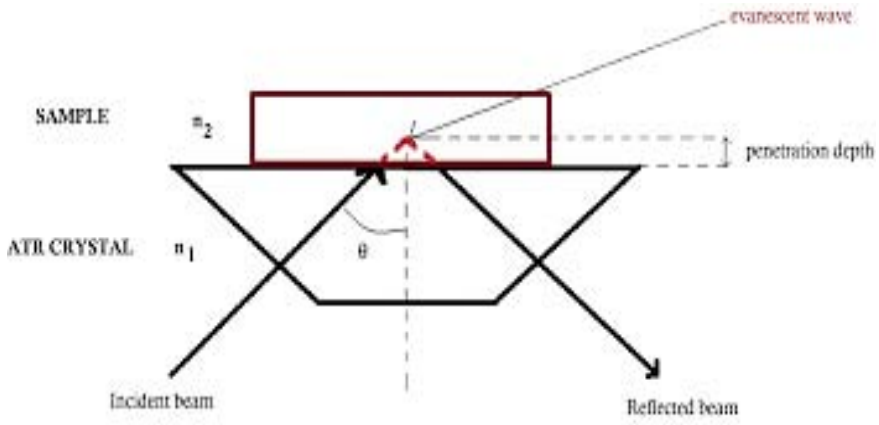


Fig. 3.8: Schematic diagram of a horizontal ATR sampling accessory illustrating the important parameters [Simonescu *et al.,* 2012]

The major benefits of ATR-FTIR are: it requires minimal or no sample preparation prior to spectral measurements due to the penetration depth of the IR beam and possibility to exchange the solvents without disturbing the protein film which can give information about protein function in different conditions.

In this work, the trapezoidal crystal made of Germanium (Harrick Scientific Products Inc., New York, USA) was used for the acquisition of all the difference spectra. It had a total length of 5 cm and 2 mm in thickness with a critical angle of 22^{o}. The angle of incidence was 45^{o} with 25 internal reflections, out of that 13 of them in the side of the crystal facing the sample. Another important aspect to be considered is the sample orientation.

The easiest sample preparation of ATR-FTIR spectroscopy is by spreading a drop of the sample using a pipette tip, usually 20 µL containing 100-150 µg of the protein on the crystal surface. The suspension was slowly evaporated under a gentle N_2 flow to generate a uniformal sample film. These techniques were very convenient for studying the reconstituted membrane protein like MelB.

3.5.3 Band Assignments

1. Amide Vibrations

The peptide group (Fig 3.9), the structural repeat unit of proteins, gives up to 9 characteristic bands named amide A, B, I, II ... VII. The smallest peptide group had been used as a model to find the modes of vibration of the amide group. Among all the amide vibrations, Amide A, Amide I and amide II are major bands of the protein infrared spectrum.

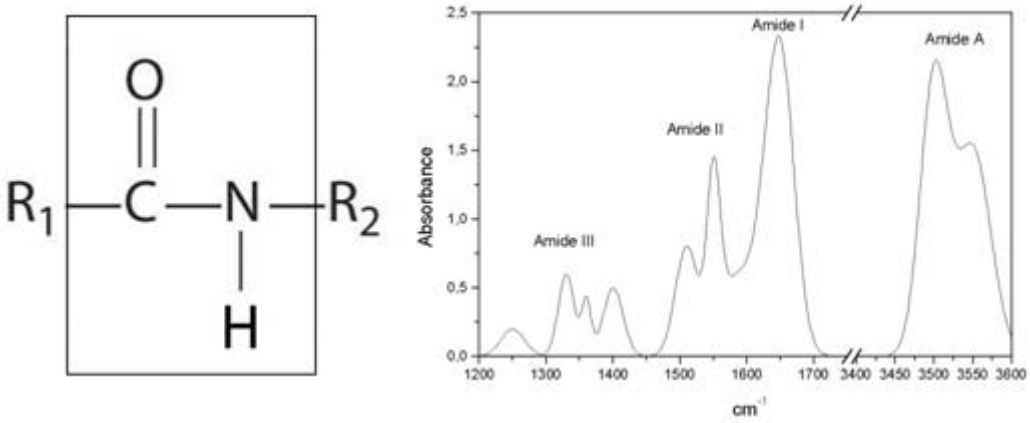


Fig. 3.9: (Left) Amide group. **(Right)** The graph illustrates the band assignments of the amide. [Krimm *et al.,* 1986].

Amide A: 95% of its signal corresponds to the N-H stretching vibration. This mode of vibration does not depend on the backbone conformation but is very sensitive to the strength of a hydrogen bond. It has wavenumber between 3225 and 3280 cm^{-1} for hydrogen bond lengths between 2.69 to 2.85 Å [Krimm *et al.,* 1986].

Amide I: is the most intense absorption band in proteins. It is primarily governed by the stretching vibrations of the C=O (70-85%) and C-N groups (10-20%). Its frequency is found in the range between 1600 and 1700 cm^{-1}. The exact band position is determined by the backbone conformation and the hydrogen bonding pattern.

Amide II: is found in the 1510 and 1580 cm^{-1} region and it is more complex than amide I. Amide II derives mainly from in-plane N-H bending (40-60% of the potential energy). The rest of the potential energy arises from the C-N (18-40%) and the C-C (about 10%) stretching vibrations.

Amide III: is very complex bands dependent on the details of the force field, the nature of side chains and hydrogen bonding. Therefore these bands are only of limited use for the extraction of structural information [Krimm *et al.,* 1986].

2 Amino acid side chain vibrations

The side chains are involved in the protein interactions, so the presence of band arising from amino acid side chains provides valuable information like protonation, cation coordination and hydrogen bonding, etc.

A thorough investigation of side chain vibration in the region between 1800 and 1400 cm^{-1} by Venyaminov *et al.,* 1990 , had suggested that among the 20 amino acids only 9 (Asp, Asn, Glu, Gln, Lys, Arg, Tyr, Phe, His) show a significant absorbance in the region . The amino acid side chain vibrations are summarized in the Table 3.3.

Amino Acid	Vibration	cm^{-1}
Asp	-COO st as	1574
	-COOH st	1716
Glu	-COO st	1560
	-COOH st	1712
Arg	$-CN_3H_5^+$ st as	1673
	st s	1633
Lys	$-NH_3^+$ bd as	1629
	bd s	1526
Asn	-C=O st	1678

	-NH_2 bd	1622
Gln	-C=O st	1670
	-NH_2 bd	1610
Tyr	ring-OH	1518
	ring-O	1602
His	ring	1596
Phe	ring	1494

Table 3.3: The amino acid side chain vibrations (according to Venyaminov *et al.,* 1990).

3.6 Experimental set-up for ATR-FTIR

In this work, all the infrared difference spectra (IR_{diff}) were acquired using Varian Bio-Rad FTS 7000 (Bio-Rad, Hercules, CA, USA) spectrometer equipped with a N_2 liquid cooled mercury-cadmium-telluride (MCT) detector.

The temperature of the sample was maintained at 25°C using an in-built casing which consisted of four nozzles, two of which is connected to the thermal bath to maintain the temperature and other two for influx and efflux of the reference (OFF) and substrate (ON) buffer. The ATR compartment was constantly purged with liquid N_2 (Leon *et al.,* 2005). The detailed explanation of this is given in (Lorenz-Fronfria *et al.*, 2012).

Before setting up ATR-IR difference spectra experiments, the ATR crystal was washed with mild detergent (Helmanex) and followed by distilled water and dried carefully before acquiring the spectra.

The spectra were collected using the ¨Rapid Scan¨ option in the Varian software. Collecting data involves the following steps:

1. Define all scan parameters.
2. Align and calibrate the spectrometer.
3. Collect a background.
4. Collect the sample data.

The scan parameters were assigned in the ¨Electronics page¨ of Rapid Scan. For all the experiments, speed of the spectrometer is set as ¨40 KHz¨, Under sampling Ratio (UDR) which is the number of laser signal zero crossings before collecting a data sample as ¨2¨, Filter as ¨30 KHz¨, Resolution as ¨4 cm^{-1} / 2 cm^{-1} ¨ and Sensitivity as 1.5, it should be set so that the center burst maximum is about -8 Volts.

Lastly, Scans to co-add was used as ¨1000¨ where number of scans add together to form the spectrum. This improves the signal-to-noise ratio in the spectrum. In the optics page, Aperture was set as ¨OPEN¨ followed by setting ¨Asymmetric¨ in the advanced page.

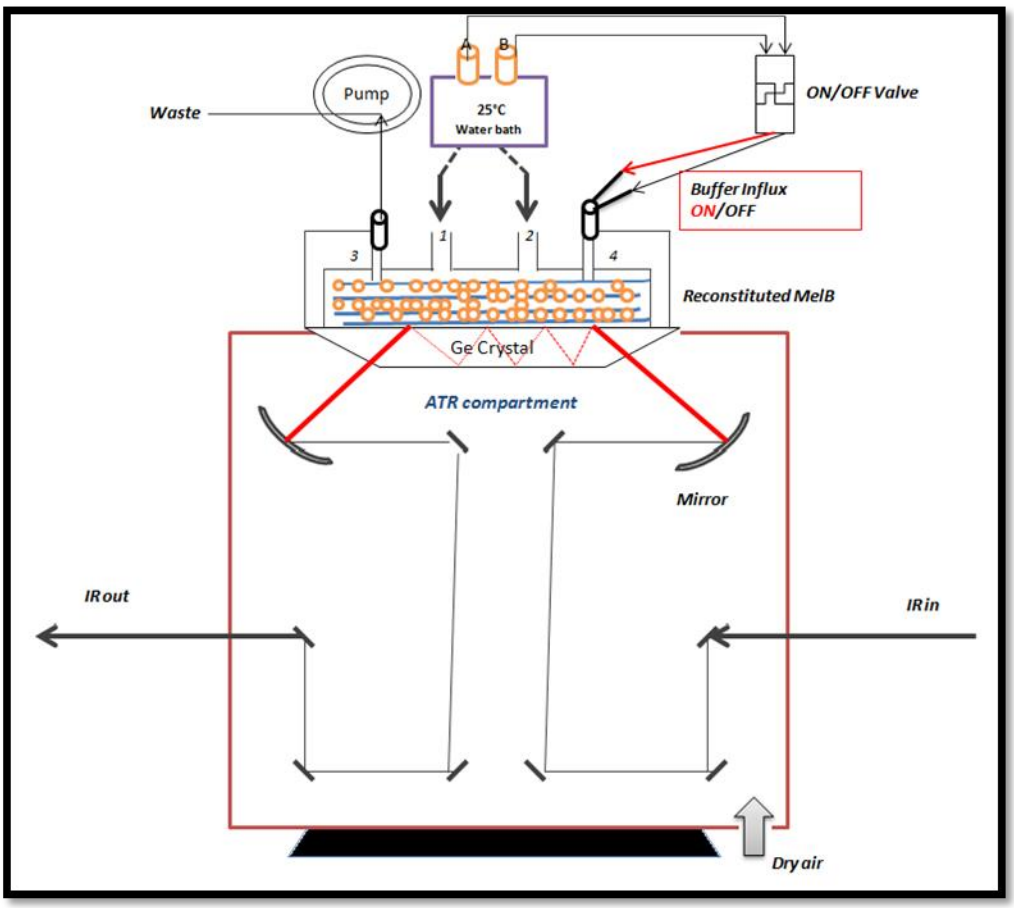


Fig. 3.10: The above scheme illustrates the geometric set up of ATR-IR difference spectra acquisition.

For each experiment, 20 µL of proteoliposomes containing 200 µg of protein was pipetted and gently distributed on a Germanium crystal. This layer of proteoliposomes was air dried and residual water was evaporated with a nitrogen gas stream. The hydration level of the sample was examined using the absorption band at 3400 cm^{-1}. The dried sample was used for acquiring an absorbance spectrum at a resolution of 2 cm^{-1} accumulating 1000 scans. This spectrum allowed us to observe the protein: lipid ratio and thus to evaluate the protein reconstitution into liposomes. A structural analysis at amide I region gives information about how far the mutant protein was conserved with respect to wild type. A band at 1620 cm^{-1} is for intermolecular β-sheets which increase upon aggregation [Jackson and Mantsch 1995].

The effect of substrate binding on each mutant with different substrates was validated using the IR difference spectra with the ATR-IR technique. As indicated in the Fig 3.10,

the MelB proteoliposomes sample was deposited on the germanium crystal and perfused constantly with substrate (ON) and reference (OFF) buffer at a rate of 1.5 ml/min (as indicated in Table 3.6). During the course of experiment, absorbance spectrum of the sample was collected indicating the quality of the protein film. Some of the mutants lost its quality after two days of perfusion with the buffers which is indicated by reduced amide I band with respect to lipid. Only in the case of Cys-less, the proteoliposomes film were very stable and was used to collect IR difference spectra for all the buffers condition (Na^+. H^+, Mel.H^+, Mel.Na^+) i.e. one set of experiment. For each MelB mutant and wild type, the spectra accumulation lasted for 7 to 24 hours depending on the substrate used. For, Na^+ substrate; 25 difference spectra of 12,500 scans and for the sugar 25,000 scans were accumulated. An example of averaged scans of Cys-less proteoliposomes in the presence of sugar and cation is indicated in Fig 3.11.

	Difference Spectra	Substrate Buffer(ON)	Reference Buffer(OFF)
Stock Buffer 20 mM MES , 100 mM KCl, pH 6.6 (10X)	Na^+ .10 mM vs K	**2min** 20 mM MES,110 mM KCl,10 mM NaCl pH 6.6	**10min** 20 mM MES,110 mM KCl, pH 6.6
	Na^+ .50 mM vs K	**4min** 20mM MES,110 mM KCl, 50 mM NaCl pH 6.6	**30min** 20 mM MES,150 mM KCl pH 6.6
	Mel. 10 mM. Na^+ .10 mM vs Na^+.10 mM	**4min** 20mMMES,100mMKCl,10 mM NaCl , 10mM Melibiose pH 6.6	**10min** 20 mM MES,100 mM KCl,10mM NaCl pH 6.6
	Mel. 50 mM. Na^+ .10 mM vs Na^+.10 mM	**4min** 20 mM MES,100 mM KCl,10 mM NaCl , 50 mM Melibiose pH 6.6	**30min** 20 mM MES,100 mM KCl,10 mM NaCl , 10 mM Melibiose pH 6.6
	Mel 50 mM vs K	**4min** 20 mM MES,100 mM KCl , 50 mM Melibiose pH 6.6	**30min** 20 mM MES,100 mM KCl pH 6.6

Table 3.6: Represents the buffers and conditions used to acquire IR difference spectra of MelB.

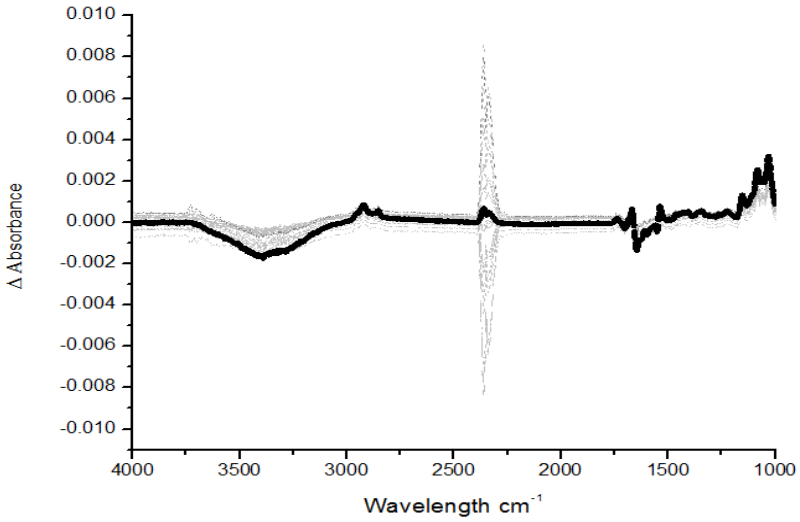


Fig. 3.11: The above spectra shows the average of 25 different scans for Cys-less mutant induced in the presence of Mel. 10mM. Na^+.10 mM vs Na^+.10 mM.

3.6.1 Corrections applied to the Infrared spectra

1. Absorbance Spectra of the Sample (Abs_{Sample})

The absorbance spectrum of the sample usually contains contributions from water which shows an intense band at 3400 cm^{-1} and an intense band at 1650 cm^{-1}. So, this was manually subtracted using absorbance spectra of the buffer (as shown in Fig. 3.12). This subtraction value will give the information about the relative volume occupied by the buffer and by the lipid-protein complex. A subtraction factor of 0.88 means that about 85% of the volume probed by the IR beam was occupied by the buffer and the remaining 15% was due to protein and lipid film, since the penetration depth of the sample is about 2.4% higher than that of the buffer.

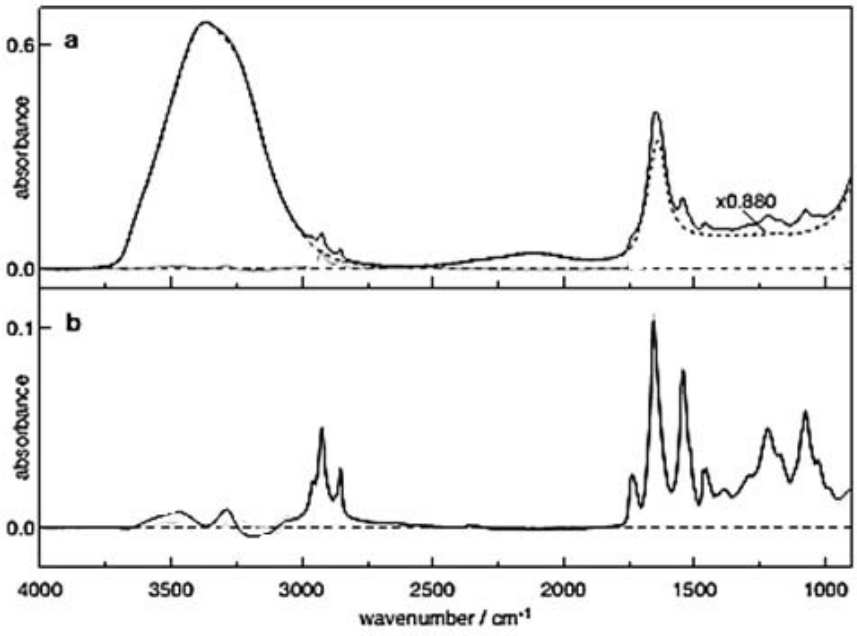


Fig. 3.12: Correction of absorbance spectra of the reconstituted MelB. **(A**) Shows the absorbance spectrum of the sample in black and absorbance spectrum of the reference buffer in dotted lines. **(B)** Shows the corrected absorbance spectra of the sample [Lorenz- Fronfría *et al.,* 2012]

2. Infra-red Difference Spectra:

To validate the spectra of each MelB mutant, several contributions had to be removed from the IR_{diff} spectra.

a. Correction for melibiose induced difference spectra:

In these difference spectra, the contributions for negative water band at 3400 cm^{-1}, melibiose absorption at 1200-1000 cm^{-1} were corrected simultaneously by using the IR_{diff} spectrum of the buffer containing melibiose (Fig 3.13 A and B).

Then the swelling of protein film was eliminated (Fig 3.13 C). The absorbance spectrum might change over time due to swelling of the protein film. Swelling means hydration level of the film changes and the distance to the reflecting surface is increased, hence the absorbance spectrum intensity decreases. To subtract this effect of the film swelling,

a single difference spectrum was computed resembling the change of deflation of the sample over time.

Thus, the absorption of the CH_2 group from the lipids at 2900 cm^{-1} was eliminated. Finally, difference spectrum was corrected for the effect of water vapor (not shown in the Fig 3.13) resulting in a final melibiose-induced IR_{diff} between wavenumbers 1900 and 1300 cm^{-1}. [Lorenz- Fronfría *et al.,* 2012].

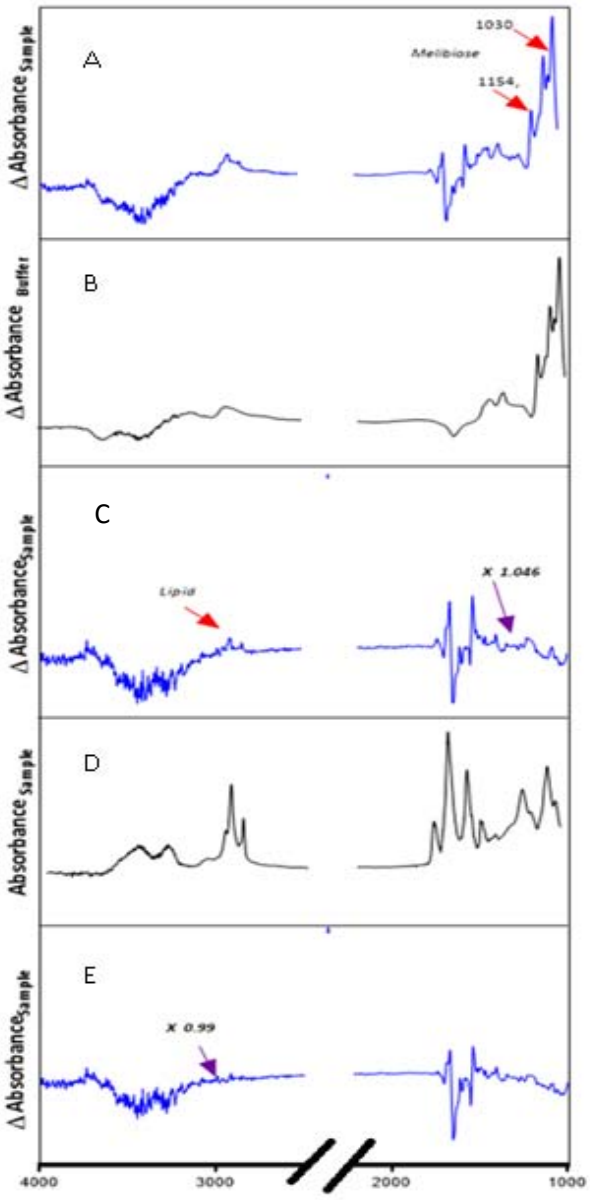


Fig. 3.13: The steps involved in correcting IR difference spectra of P165C in the presence of Na^+10mM Mel. **(A**) The average difference spectra of P165C (Na^+10mM. Mel 10mM vs Na^+10mM) which has to be corrected for melibiose contribution and lipid swelling. **(B**) The average difference spectra of the buffer alone. **(C**) Subtraction factor of 1.046 was used to correct the Melibiose (A-B). **(D)** Absorbance spectra of the sample. (E) Subtraction factor of 0.99 was applied to remove the lipid band (C-D). **(E**) The final corrected difference spectra of P165C.

b. Corrections for Na^+ induced difference spectra:

The correction for Na^+ induced difference spectra is quite similar to the above process with melibiose, except for the fact that Na^+ does not absorb IR like melibiose. So, the band of Na^+ is not visible as melibiose, but it slightly alters the vibration pattern of the water. Since, this alteration is highly negligible, preference was given to the correction of protein lipid swelling. The raw Na^+ induced difference spectra showed negative lipid band and positive water band which was corrected using buffer-minus-sample absorbance spectrum.

Finally, to eliminate the effect of unspecific contributions of Na^+ in the water absorbance region, the sample difference spectrum was subtracted from IR_{diff} of Na^+ buffer [Lórenz-Fronfría *et al.,* 2012]. A detailed process of this correction is illustrated in Fig 3.14.

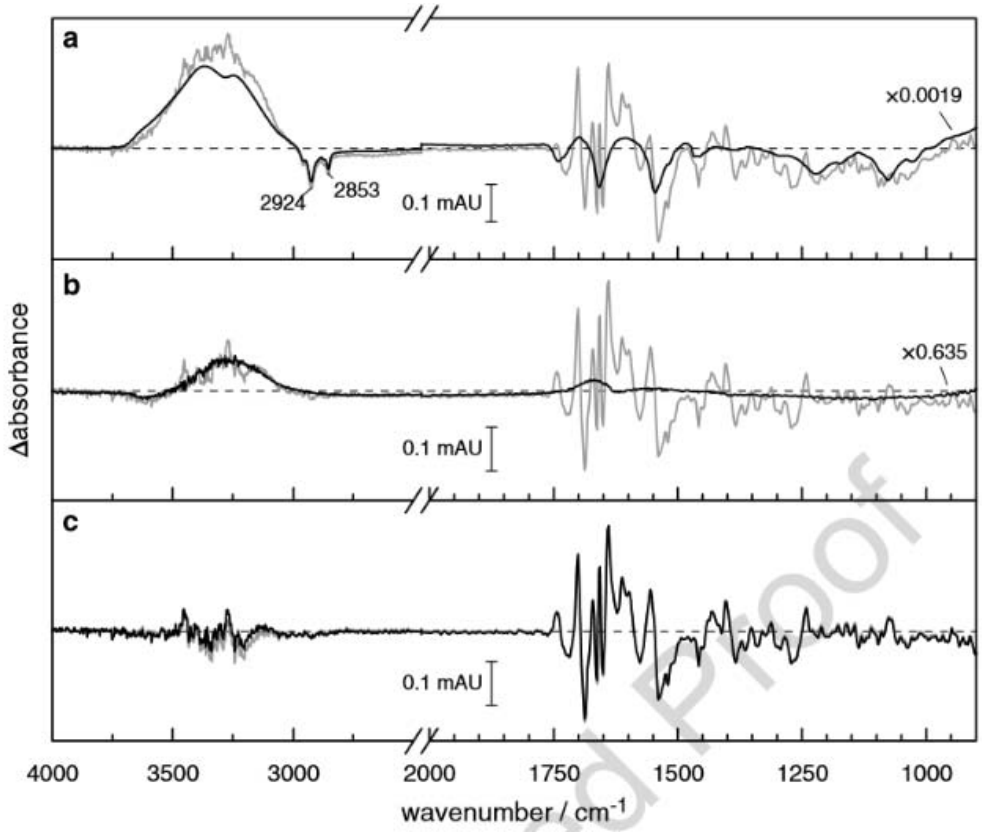


Fig. 3.14: Correction steps to obtain the IRdiff spectrum corresponding to Na+ binding to Cys-less MelB. **(a**) Raw Na+-induced IRdiff difference spectrum of Cys-less MelB (gray line). Scaled IRdiff difference spectrum for the sample film swelling (black line). **(b)** Na+-induced IRdiff spectrum of MelB corrected for the film swelling (gray line). Scaled Na+-induced IRdiff spectrum of the buffer (black line). **(c)** Final Na+-induced IRdiff difference spectrum of MelB corrected for all the unspecific spectral contributions, [Lórenz et *al.,* 2012]

3.6.2 Quantitative analysis of the spectra

The quantitative analysis was performed by linear regression method for the absorbance and difference spectra in the region 1700/1900 to 1300 cm^{-1}. It provided two values, the slop, m, and the correlation coefficient, R^2.The slope m, determines the relative intensity of the spectral features to the reference spectrum and the correlation coefficient, R^2 represents the similarity of the spectrum with respect to Cys-less.

The slope, m, value from the absorbance spectrum was used to determine the quantity of protein deposited on the ATR crystal which in turn was used to normalize the IR_{diff} spectra since protein film varies each time. The correlation analysis in the Fig 3.15, shows the normalized absorbance spectra of R175C mutant with respect to Cys-less.

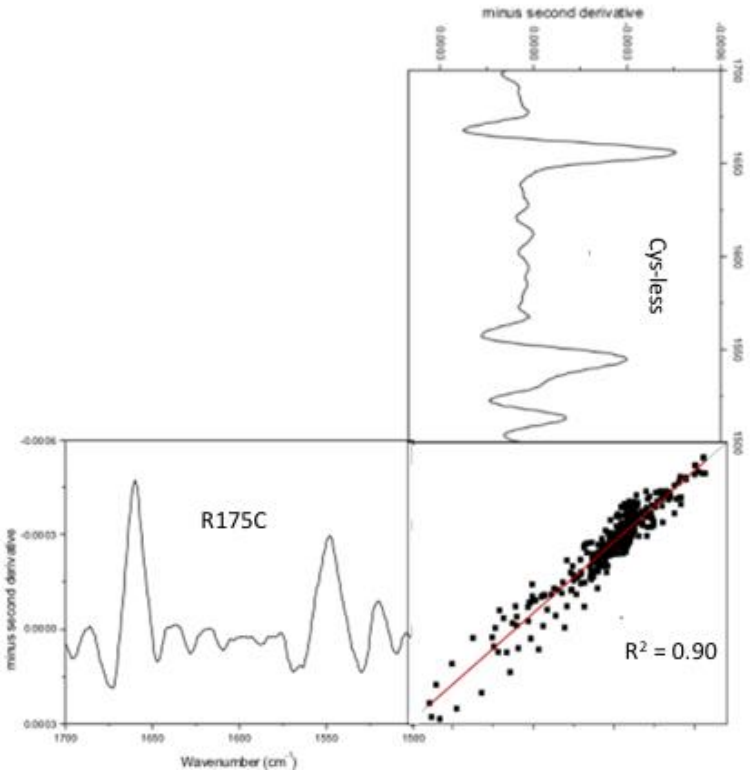


Fig. 3.15: Illustration of correlation analysis of two absorbance spectra using their second derivative spectra. The parameter m was used to normalize the spectra with respect to Cys-less and R^2 corresponds to similarity of the structural components between different forms of MelB

As shown in Fig. 3.16 quantitative comparison of two normalized difference spectra (input and a reference spectrum) was performed by linear regression in the 1900-1300 cm^{-1} interval on the first derivative of the difference spectra [Granell *et al.,* 2010].

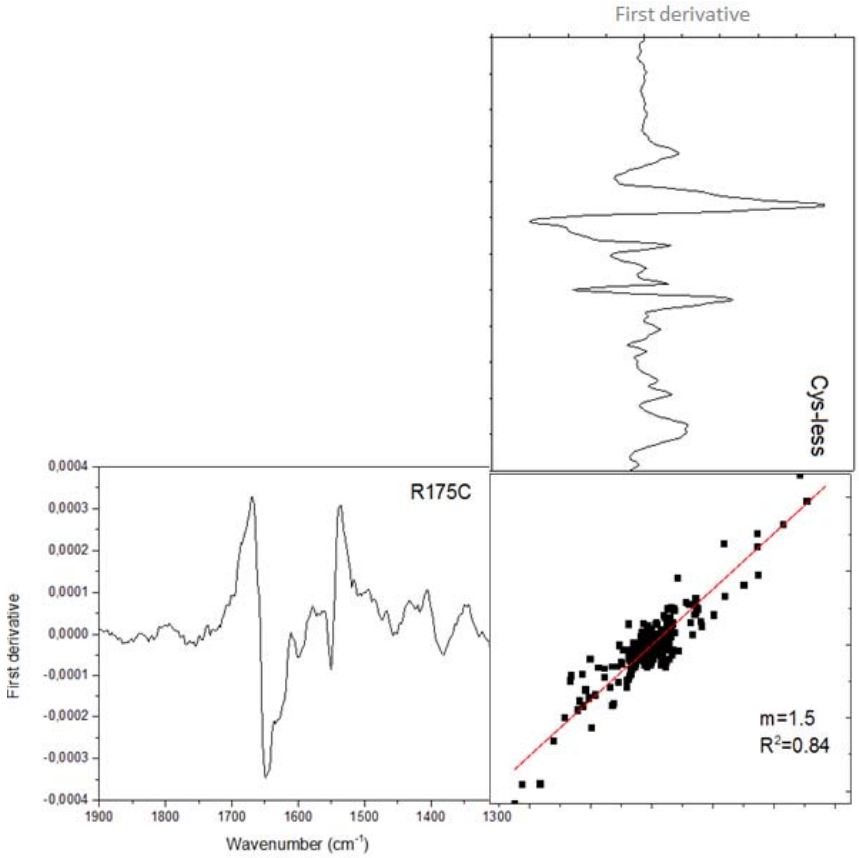


Fig. 3.16: Illustration of correlation analysis of two Infrared difference spectra using their first derivative spectra. The two parameters, R^2 and m, corresponds to similarity and intensity of the IR difference spectra between mutants and Cys-less.

3.7 Analysis of the IR_{diff} spectra using color gradient SI map

The analysis of SI map plots the differences in shape with respect to differences in intensity derived from IR_{diff} spectra. Each mutant can be represented on the SI map characterized by a pair of co-ordinates (s,i). Mutants lying on the map from the diagonal upwards have a difference spectrum differing in intensity with respect to Cys-less and mutants lying on the map from the diagonal downwards have a difference spectrum dominated by differences in shape.

The s and i coordinates can be used to define a similarity index as s+i. In this way, we are defining an index that has two components: one component corresponding to the degree of similarity in shape (s) and other component corresponding to the degree of similarity in intensity (i). The similarity index is intended to be a measure of the total degree of similarity. The lower the similarity index the more dissimilar are two spectra. For example, spectra with pair of coordinates (100, 0) will have the same similarity index as spectra with a pair (0,100). Maximum similarity (identity) corresponds to the pair of coordinates (100,100), i.e., to a similarity index of 200. Maximum dissimilarity is given by the pair (0, 0), a similarity index of 0 (Fig. 3.17).

From the way in which the similarity index has been defined, it follows that a spectrum with a given dissimilarity percentage in intensity will reflect the same degree of substrate binding perturbation as a spectrum with the same dissimilarity percentage in shape. For example, a mutant with coordinates (90, 10) and a mutant with coordinates (10, 90) will represent the same degree of similarity and will assume to have an equivalent effect on substrate binding. The first mutant (90, 10) gives a difference spectrum with a reduced intensity but with a very similar shape to the spectrum of Cys-less. The second mutant (10, 90) has almost the same intensity but an almost completely different shape. In both the cases the magnitude of the effect on the ´native binding´ is considered to be equivalent.

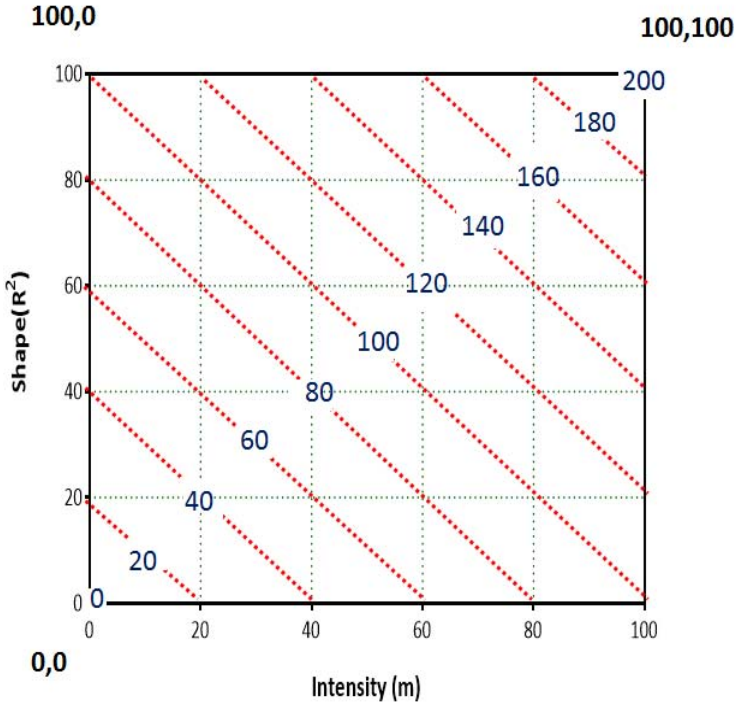


Fig. 3.17: Illustrates the SI map as a sequence of diagonals corresponding to points with the same value of the similarity index.

Finally, the similarity scale was defined using a color gradient code (Fig. 3.18). Here, the color intensity is an indication of the degree by which the mutation of the residue into a Cys affects the infrared spectrum and therefore the substrate binding. Using this, the mutants in the helix and loops were further classified based on three groups (high, intermediate and low) with respect to the effect on substrate binding.

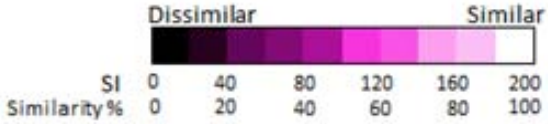


Fig. 3.18: Color gradient scale of SI map indicating the highly similar, dissimilar and intermediate regions due to substrate binding.

3.8 Computational approach in the membrane transporter

Membrane transporter plays a vital role in active exchange of materials across the cell membrane in an energy dependent or electrochemical gradient manner. However, high resolution experimental structures of membrane protein complexes are relatively scarce due to the following reasons [Arinaminpathy *et al.,* 2006],

1. Membrane transport proteins are present in the cell at low concentrations and the overexpression is less easy when compared with soluble proteins.
2. Difficulty in using high resolution experimental techniques with the proteins in their natural lipidic environment.
3. Difficulty in the use of the detergent in the place of lipids can denature the protein or impede in the crystallization by phase separation. Optimization of detergent concentration takes a long time.

Transporters describe cycles with many different conformational forms and it is extremely difficult to have crystal structures of different intermediates. Thus, the combination of experimental methodologies with computational approaches like Molecular dynamic and Molecular docking studies are the most commonly used to understand the conformational dynamics of membrane protein in atomic detail. Usually a high resolution crystal structure or a homology modeled structure is used as a template. In this work, the crystal structure of $MelB_{ST}$ PDB ID: 4M64 [Ethayathulla *et al.,* .2014] was used as a starting template. Among the four models of the structure file, Mol-A (an active state) was used. Once the structural model is available, it can be used to generate a Protein/Membrane complex which can be later used in computer simulation methods to provide key insights into the general nature of protein motions and aspects of motion linked to the functions of proteins in their native state.

3.8.1 Preparation of the Protein/Membrane complex

The Protein/Membrane complex was generated using the automated web program called Bilayer Builder in Input Generator inside CHARMM-GUI (WW.Charmm-gui.org) [Jo *et al.,* 2008]. Steps involved are:

Step 1: Protein Manipulation:

In this study, $MelB_{ST}$ has been used as a template to mutate the necessary amino acids for those of E.coli MelB (60 different residues compared to $MelB_{ST}$) using the option-Mutation in Membrane Builder.

Step 2: Alignment of the Protein:

The protein molecule was aligned along the Z-axis and the hydrophobic region of the protein was placed along Z=0 [Jo *et al.*2013].

Step 3: System Size Determination:

System size was specified by assigning the specific lipid type using the homogenous lipid bilayer option. Using Geometry option, specified "X=90", Box type="Rectangular" and lipid type="POPE" (1-palmitoyl-2-oleoyl-sn-glycero-3-phosphoethanolamine).

Step 4: System Generation:

The membrane was built around the protein using the replacement method and the ions and water are added in the following step. Once, the protein was well inserted into the lipid bilayer, the following output files were downloaded -"download.tgz" "Step4_lipid.pdb" and Step4_lipid.psf".

Step 5: Solvation and Ionization:

In order to examine the possible conformational changes in the native environment, it is necessary to solvate the system using the "**Add Solvation Box**" plugin called Extension in the Visual Molecular Dynamics (VMD version 1.9.2). In the periodic box, min and max values of X as -5, -5(Å), Y as -5, -5(Å) and Z as +15, +15(Å) were entered. This generated a box with a layer of water of 15 Å in the direction of Z with respect to the protein and -5 Å less water molecules with the most extreme atoms in to X and Y.

The advantage of solvation is that it can deal with periodic boundary conditions, where atoms leaving one side of the cell, may re-enter from the opposite side (Fig. 3.19).

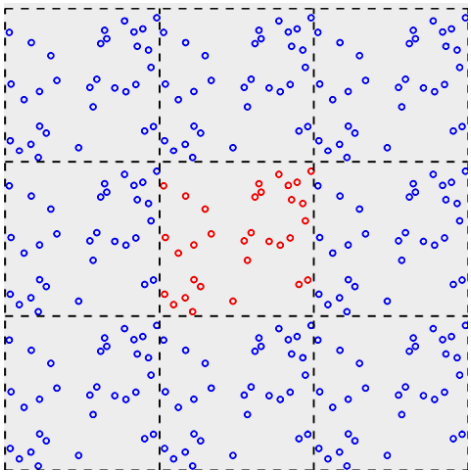


Fig. 3.19: The particles in the simulations are shown in red, while their mirror images are shown in blue. A particle which goes out from the simulation box by one side is reintroduced in the box by the opposite side.

Periodic boundaries were used to eliminate surface effects. The dimensions of periodic boundary cell must be large enough to prevent any undesirable interactions of the protein with its mirror image.

Solvation was followed by Ionization using the **"Auto-ionize"** window in the Extension pull down. Solvate.pdb and Solvate.psf were used as the input files. NaCl was chosen and "Neutralize and set NaCl concentration to"= 0.15 M. This generated a system with the protein inserted into the lipid membrane, water and ions.

3.8.2 Minimization and Equilibration in NAMD

The generated system was minimized and equilibrated using NAMD (NAnoscale Molecular Dynamics) [Kale *et al,* 1999] with the specific parameter files. Four input files are needed prior to the process:

1. A .pdb file containing the initial coordinates or positions of each atom of the system,
2. A. psf file also known as Structure file which defines the bond type,
3. The CHARMM27 force field parameter and topology files: where the parameter file (par_all36_prot22_na_lipid_carb_ethers_cgneff_water_ions.prm) contains all numerical constants needed to evaluate forces and energies and the topology file (top_all27_prot_lipid.rtf, toppar_water_ions.str) defines which atoms are connected to one another through chemical bonds.
4. Lastly, a user-defined configuration files (.conf) which contains the simulation protocol: time step, initial velocities (or temperature), number of minimization/equilibration steps, file paths, boundary conditions, restraints, integration parameters, etc.

If a file containing the initial atomic velocities were not provided, NAMD can randomly generate a Maxwell-Boltzmann distribution of velocities corresponding to a given temperature. NAMD uses the Verlet leapfrog method of integration to calculate atomic trajectories.

The goal of energy minimization is to find a nearest energy minimum conformation, by adjusting the structure to the force field, particular distribution of solvent molecules, and relaxing the possible steric clashes created by guessing coordinates of atoms during generation of psf file. Here, this was done in 100000 steps in periodic boundary conditions where interactions like steric clashes were removed, which may cause the

numerical integers to become unstable in the process of dynamics. The minimized structure has small forces on each atom and therefore serves as an excellent starting point for molecular dynamics simulations.

Once the input files (above mentioned) were available, minimization was carried out using NAMD 2.1 in windows. It took nearly 7 hours to finish one set of minimization using a desktop computer with Central Processing Unit (CPU) card. Following energy minimization, equilibration was done to thermalize the system and release the restraints. The energy minimized output was used as an input along with the topology, parameter and .PSF files. Equilibration was done by slow heating the system from 0 to 298 K in a step of 2 K with a given time step of 1 fs which ran for 10 ps (100000 numsteps) in the NVT thermodynamic ensemble. In this canonical ensemble, amount of substance (N), volume (V) and temperature (T) are constant. Here, the energy of endothermic and exothermic processes is exchanged with the thermostat. In this work, the Langevin dynamics thermostat was used which balances friction with random noise to drive each atom in the system towards a target temperature.

Command line used in Windows for Minimization and Equilibration:

namd +p4 Minimization/Equlibration.conf (refer Appendix)

(+p4 designates 4 processors and this can be changed accordingly, .conf files contains all the instructions to the program)

3.8.3 Molecular Dynamics Simulations using ACEMD

Molecular dynamics simulations (MD) are used for monitoring time dependent processes of molecules by solving Newton's equation of motion in what is known as,

$$F_i = m_i.a_i, \qquad \text{(eq.5)}$$

Where F_i is the force, m_i is the mass, and a_i is the acceleration of atom i.

The force on atom i can be computed directly from the derivative of the potential energy V with respect to the coordinates r_i as given in equation 6 and 7,

$$F_i = -dV/dr_i. \qquad \text{(eq.6)}$$

$$m_i\, d^2r_i/dt^2 = -dV/dr_i. \qquad \text{(eq.7)}$$

MD simulations were carried out using the program called ACEMD (Accelerating biomolecular dynamics) which was optimized to run on compute unified device architecture (CUDA) compatible NVIDIA GTX680 -Graphics processing unit (GPU). The CUDA programming model, an extended C-like language for GPUs, abstracts the implementation details of the GPU so that the programmer may easily write code that is portable between current and future GPUs.

The use of GPU card accelerates the simulation, visualization and analysis by factors ranging from 10X to 100X faster than contemporary multi-core CPU cards because it is comprised of hundreds of cores that can handle thousands of threads simultaneously. GPU-accelerated desktop workstations can now provide performance levels that used to require a cluster, but without the complexity involved in managing multiple machines or high-performance networking.

Users of NAMD and VMD can now perform many computations on laptops and modest desktop machines which would have been nearly impossible without GPU acceleration. For example, NAMD users can easily perform simulations on large systems containing hundreds of thousands and even millions of atoms [Stone *et al.,* 2009].

This explains the use of ACEMD instead of NAMD for dynamic simulations. The input files for the dynamics were taken from the output of the equilibration along with parameter and topology files. In this work, MD Simulations were performed under NPT condition for 250ns, where the temperature, pressure and the number of atoms were kept constant. Fig 3.21 illustrates the steps involved in MD simulations using ACEMD.

Command line used in Linux to run ACEMD:

acemd dynamics.conf (refer Appendix)

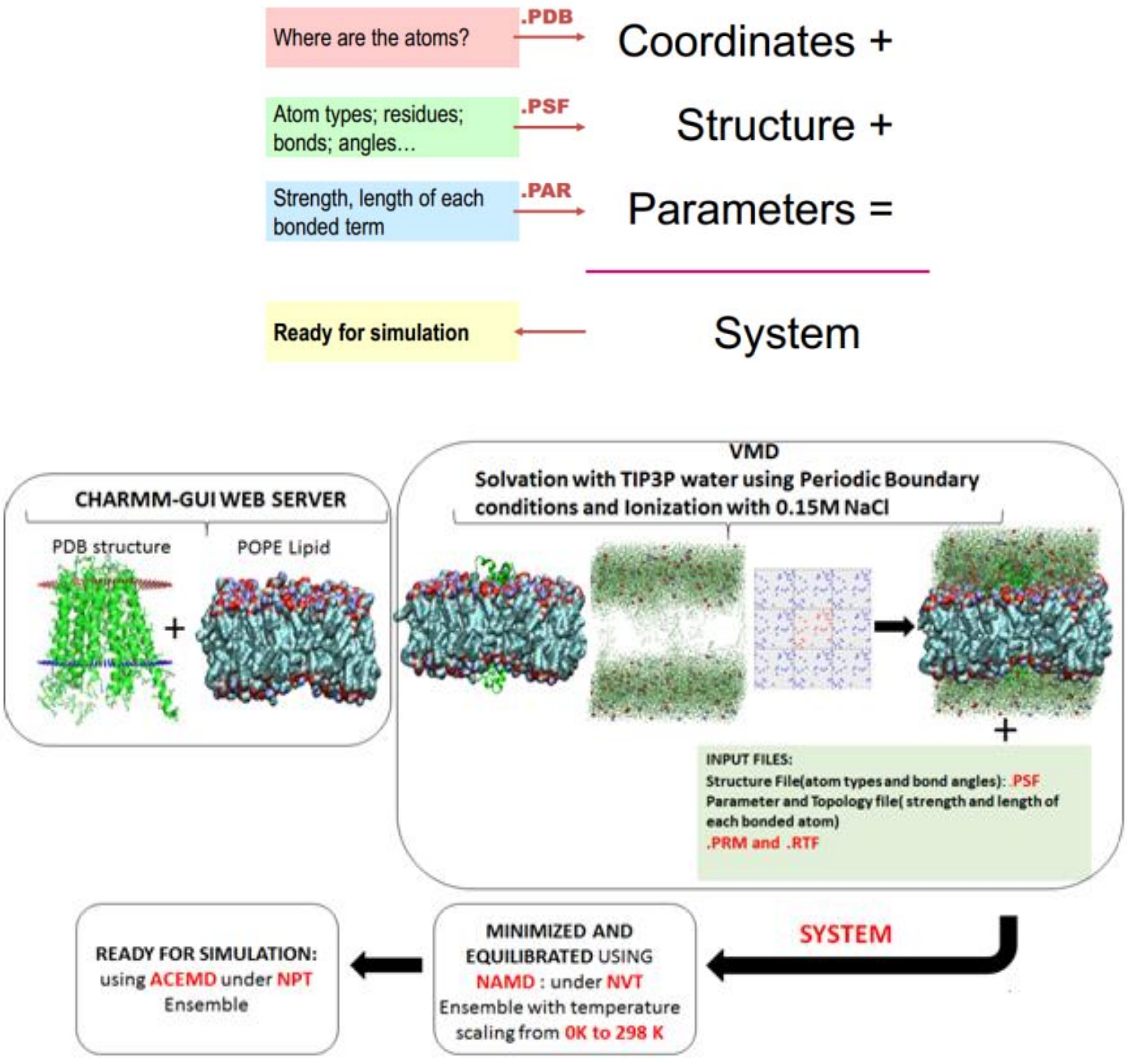


Fig. 3.21: (Top) Coordinates, structure and parameter files used as input in MD simulations. **(Bottom)** The flowchart illustrating the methods involved in the MD simulation using ACEMD.

3.8.4 Analysis in MD Simulations

The output file, .dcd containing a list of coordinates and velocities of the atoms in the simulation was loaded along with the .psf file in the VMD Visual Molecular Dynamics (VMD version 1.9.2) [Humphrey *et al.,* 1996] and it was displayed in the ¨OpenGl" window.

Deviation from X-ray crystal structure

One of the important factors to be analyzed after dynamics is whether the protein is stable and close to the experimental structure. The standard way to do is by measuring of Root Mean Square Deviation/displacement (RMSD) of all the heavy atoms. RMSD is a parameter indicating the stability of the entire atom selection (e.g. protein backbone or trace) along time. Usually this helps in evaluating whether the system has reached equilibrium and also estimates how long one should run a simulation to achieve equilibrium.

In this work, this was usually calculated using the plug-in in the VMD extension called ¨RMSD visualizer tool¨ by selecting the trace of the system. And the generated output file was exported in .dat format and analyzed/plotted using Origin 9 software.

Fluctuations Analysis

RMSD gives no indication on which part of the protein is more stable/variable. Thus, calculating the Root mean square-fluctuations (RMSF) is useful for characterizing local changes along the protein chain. It is the time average of RMSD for each residue; it is also called RMSD per residue. The peaks in the plot indicate areas of the protein that fluctuate the most during the simulation. Usually core protein regions have low RMSF and exposed loops have high RMSF values.

RMSF was calculated in this work, by writing a small script in the VMD Tk console (as mentioned in Appendix) and a text output file was generated which was further plotted using Origin 9.

Hydrogen bond and salt-bridge distance analysis

In this work apart from calculating basic properties like RMSD and RMSF, the output of the simulation was also used to analyze more specific details about the protein such as hydrogen bond and salt bridge distances between two atoms over the course of simulations. This can provide detailed information about how the helices/loops move towards/away from each other during the course of simulation.

Target residues were manually selected in VMD and checked for the nearby interactions with other residues. And the distances were represented using a ¨Label¨ option in VMD. The files were exported as .dat file and plotted using Origin 9.

3.9 Transport assay in Melibiose Permease

E.coli DW2-R cells containing the WT or the mutated MelB were expressed overnight in 25 mL of LB-medium containing 100 µg/mL ampicillin, 100 µg/mL tetracycline and 0.5% glycerol until it reaches an OD_{600} of 1.2.

The cells were centrifuged at 3700g for 10min and the pellets were washed four times with the wash buffer to remove any Na^+ contamination (Guan *et al.,* 2011). Finally the pellets were re-suspended using the re-suspension buffer and adjusted to an Abs_{600} of 2.2 (0.7 mg/mL).

2µL of [^{3}H] melibiose with a specific activity of 10 mCi/mmol at a final concentration of 0.4 mM either in the presence or absence of 20mM NaCl was added to the tubes. The reaction was started by adding 50 µL of sample to the above tubes at different time s (0, 2, 5, and 10 min). The reaction was stopped at a given incubation time by diluting 60-fold (3mL) using the quenching buffer and fast filtrated using Grade GF/C microfiber sheets (WHATMANN) with a pore size of 1.3 µM and washed with the same quenching buffer. Radioactive melibiose accumulated by the cells was maintained in the filter while the non-accumulated melibiose just passed through the filter.

The filters were kept inside the scintillation tubes along with 3 ml of OPTIPHASE 'HISAFE' (PerkinElmer, Inc.), and the intracellular [^{3}H] melibiose was measured in a scintillation analyzer (PerkinElmer Tri-Carb 2810 TR). OptiPhase HiSafe is a general purpose liquid

scintillation cocktail. It combines very high counting efficiency with moderate to high sample holding capacity for a wide range of aqueous and non-aqueous solutes. In all the experiments, a blank value was measured using the [^{3} H] Melibiose without the protein.

Wash Buffer:
0.1M KPi pH 7.5
Resuspension Buffer:
0.1M KPi pH 7.5, 0.01M MgSO4
Quenching Buffer:
0.1M Kpi pH 5.5, 0.1M KCl, 0.01M MgSO4

Box 5: Buffers used in Transport assay measurements.

4. RESULTS (PART-1)

Secondary transporters couple the free energy stored in an ionic gradient to the movement of solutes across the cell membrane. Here, binding of the first substrate will control the binding of the second substrate. In other words, it can be evident that the both substrate binding sites should be close to each other. This phenomenon is seen in many co-transporters like LeuT [Yamashita *et al.*, 2005], MhP1 [Shimamura *et al.,* 2010] and LacY [Abramson *et al.,* 2003]. From previous studies, it was also proved that in MelB the Na^+ binding sites are primarily located in the N-terminal domain of the protein (Poolman *et al.,* 1996) while the sugar binding sites are located in both N and C-terminal domains.

From the recent crystal data from $MelB_{ST}$ (4M64.pdb) it has been shown that the three Asp residues at positions 55, 59 (Helix II) and 124 or 121 (Helix IV) form a trigonal bi-pyramidal geometry with Tyr-120, Thr-121 (Helix IV) and Thr-373 (Helix XI) which is important for the cationic selectivity properties. Residues Asp-19 (Helix I), Tyr-120, Asp-124, Trp-128 (Helix IV), Arg-149 (Helix V) and Lys-377 (Helix XI) found near the water filled cavity are important candidates for sugar binding [Ethayathulla *et al.,* 2014]. Cysteine scanning mutagenesis done by our group for Helix V revealed that it is an important helix for the function of MelB. One of the interesting mutants studied in this group was R149C, which is a key residue for the reorientation of MelB and fixes the MelB in an inward facing conformation [Lin *et al.,* 2013]. Some of the other residues like Ala-155, Thr-159 and Gly-161 had a major impact when mutated to Cysteine, suggesting that these residues could also be involved in substrate binding [unpublished results Lin.Y 2012].

In this study, Cysteine scanning mutagenesis was done for four residues at the end of **Helix V** (^{164}**LPFV**^{167}) along with **loop 5-6** (^{168}**NYVGGGD**^{174}) and four residues in **Helix VI** (^{175}**RGFG**^{178}). Followed by this, other mutation was also done for one of the charged residues in Helix I (Asp35) which was shown to form a salt bridge interaction with Arg175 in Helix VI (Ethayathulla *et al.,* 2014). The loops connecting the transmembrane helices are important for their structural organization and protein stability [Tastan *et al*., 2009].

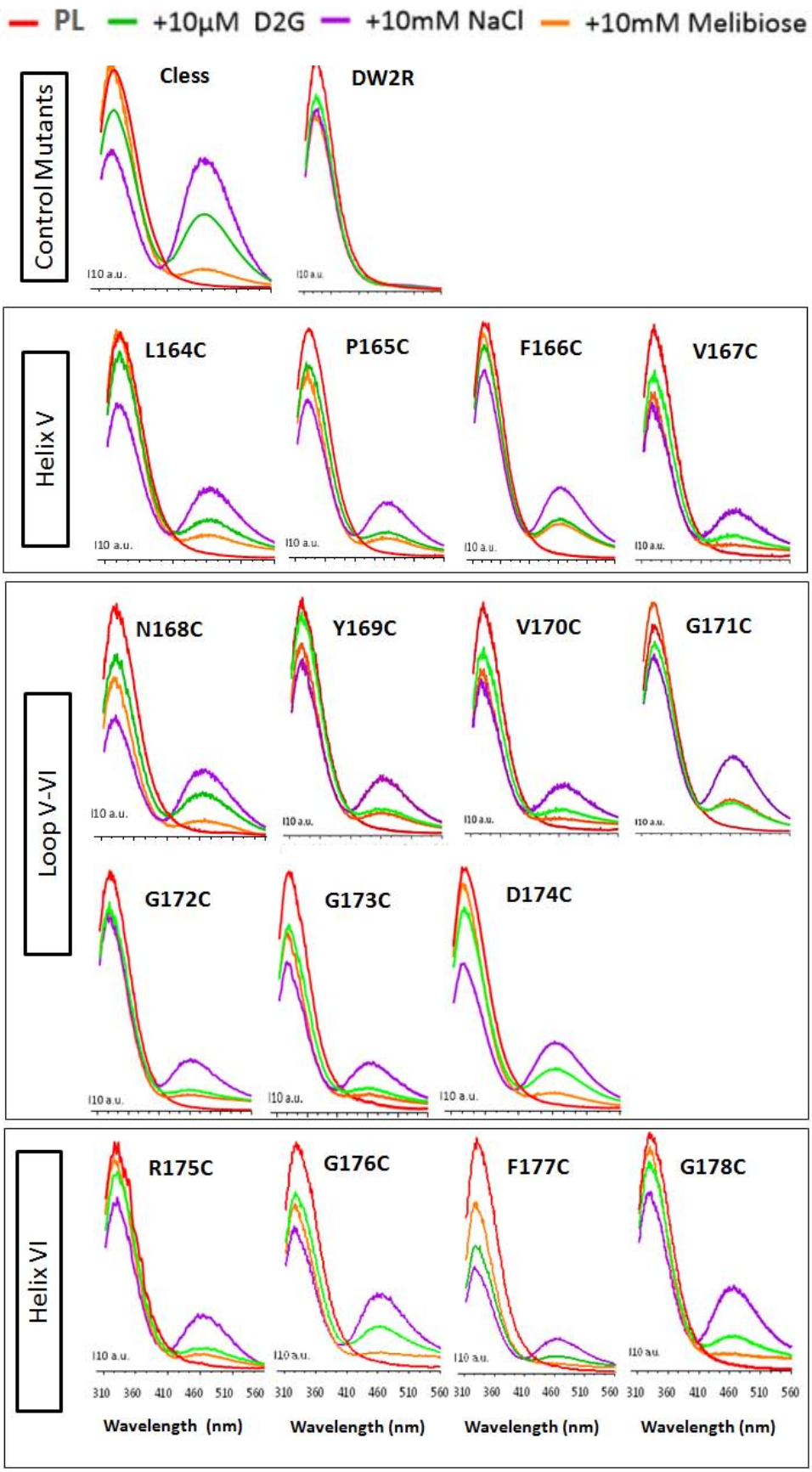
4.1. Fluorescence resonance energy transfer (FRET) analysis

Substrate binding studies can be done using FRET by measuring the energy transfer from Trp residues to the fluorescent sugar analog (exited at 290 nm) [Maehrel *et al.,* 1998]. As mentioned before, this is a stepwise process where the energy of Trp fluorescence signal initially decreases around the region of 330nm and increases in the region around 460nm when D^2G is added to the sample (indicating the transfer of energy). This was followed by the addition of Na^+ resulting in the further reduction of Trp fluorescence (330 nm) followed by a further increase of signal around 460 nm indicating that Na^+ enhances the affinity for the sugar. Finally, the addition of the native sugar melibiose resulting in the decrease of the FRET signal (460 nm) implies that the fluorescent sugar and the native melibiose sugar compete for the same binding site [Maehrel *et al.,* 1998, Cordat *et al.,* 1998]. This method is useful to determine the capability of substrate binding in specific mutants in both reconstituted protein and native vesicles forms.

4.1.1 Effect of FRET in Proteoliposomes

As shown in the Figure 4.1, there is no clear FRET signal from Trp side chains to D^2G bound to the sugar-binding site obtained from D35C (Helix I) and R175C/D35C (salt-bridge) mutants. The mutant F177C in Helix VI shows a very weak energy transfer from Trp-side chain to D^2G bound to sugar binding site. So, from these results we can conclude that the mutation of Asp-35 and Asp-35/Arg-175 eliminates the D^2G/Na^+ coupled binding in proteoliposomes and the mutation of Phe-177 highly affects it.

Other mutants, i.e., L164C, P165C, F166C, and V167C in Helix V; N168C, Y169C, V170C, G171C, G172C, G173C, and D174C in Loop5-6; R175C, G176C, and G178C in Helix VI (Fig. 1) shows a similar behavior of FRET response to the consecutive additions of D^2G, Na^+, and melibiose, confirms that all of them preserve the capability of sugar/Na^+ binding with a reduction in the intensity when compared to Cys-less, suggesting that the mutations may affect the affinity of the substrates.


PL
+10µM D2G
+10mM NaCl
+10mM Melibiose
Control Mutants
Cless
DW2R
Helix V
L164C
P165C
F166C
V167C
Loop V-VI
N168C
Y169C
V170C
G171C
G172C
G173C
D174C
Helix VI
R175C
G176C
F177C
G178C
I10 a.u.
310 360 410 460 510 560
Wavelength (nm)

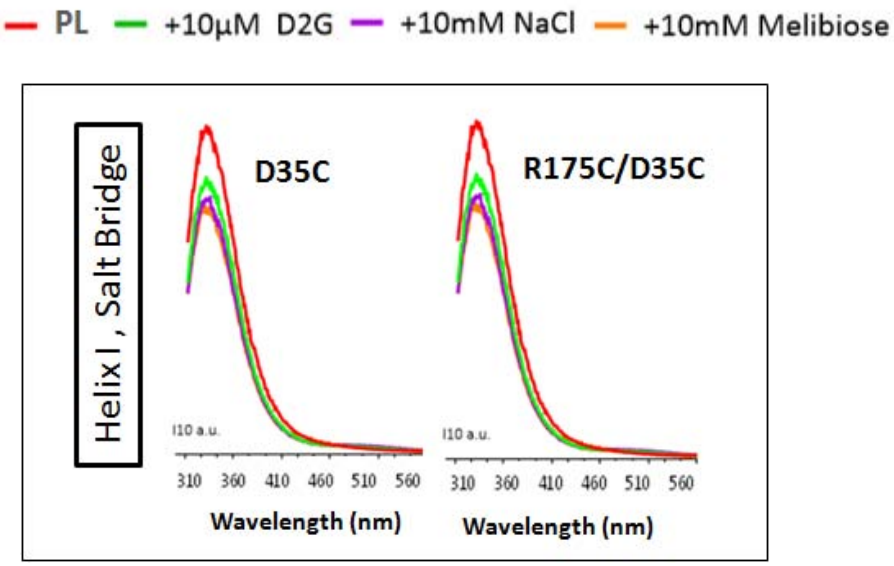


Fig. 4.1: Substrates-induced Trp-D^2G FRET signal changes of MelB mutants in proteoliposomes (PL). The proteoliposomes(PL) were in 100mM KPi and 100mM KCl pH 7.0. The representative plots of FRET changes for the different MelB mutants induced by the sequential addition of 20 µg/ml protein (**Red**), 10 µM D^2G (**Green**), 10mM NaCl (**Purple**) and 10mM Melibiose(**Orange**) into the PL suspension are shown. All the samples were illuminated at 290 nM (half-bandwidth 5), and the emission spectra were collected at 310-570 nm. The plots represented here are the average of two or three individual experiments.

The main advantage of measuring FRET in proteoliposomes is that the purity of the protein i.e. the signal recorded for FRET changes is only due to MelB. The process of obtaining proteoliposomes has been explained thoroughly in the "Materials and Methods" section. Briefly, protein has to be extracted from the native membrane by the use of detergents. The process is called solubilization. This was followed by the process of purification using Ni-affinity chromatography. Only the protein with His-tag ~~was~~ is eluted with the detergent.

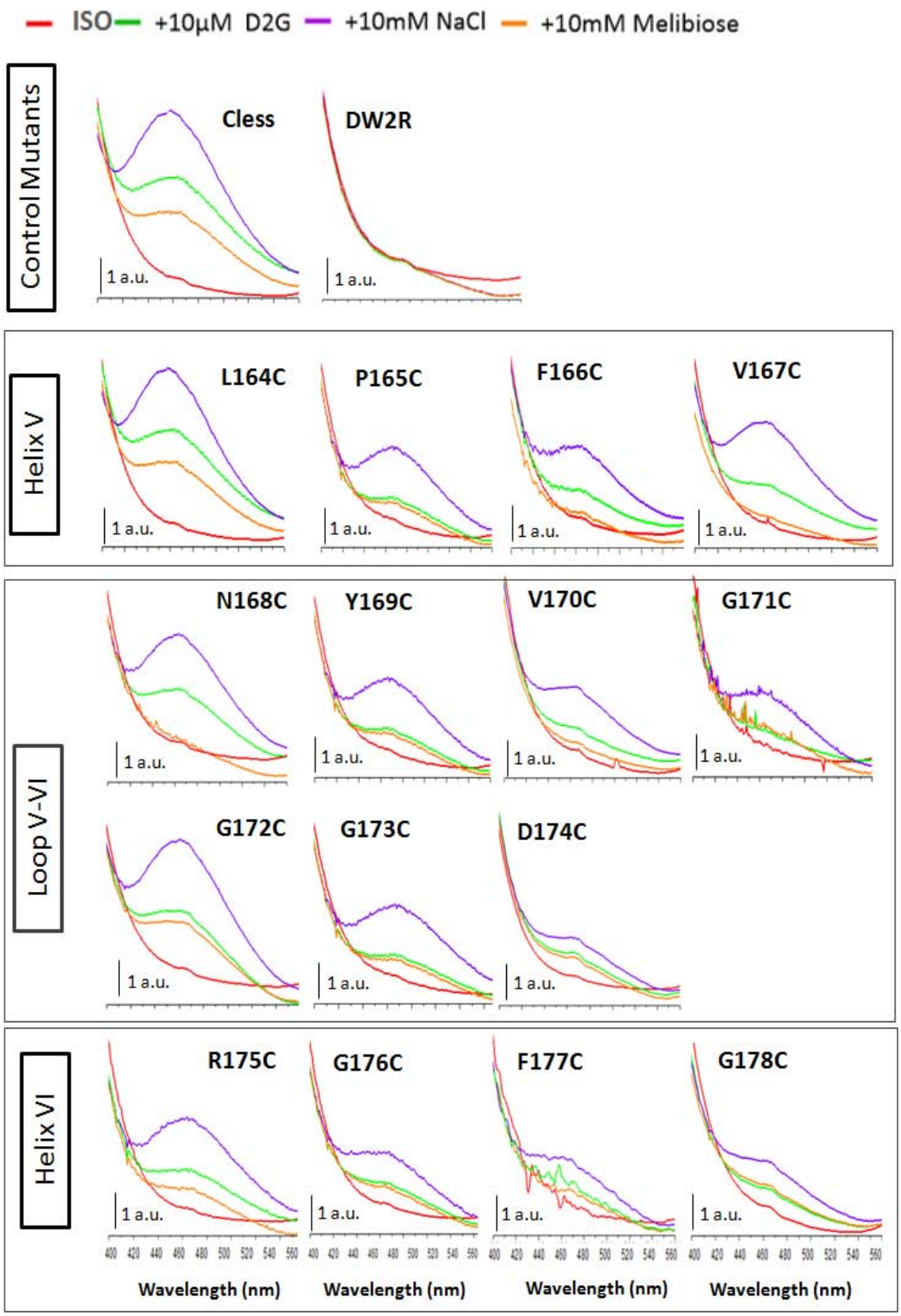
4.1.2 Effect of FRET in Inside Out Vesicles

The major disadvantage of measuring FRET changes in the previous approach is the extraction of the protein from the membrane by using detergents which might affect the protein's native conformation. So, the measure of FRET changes in native vesicles has helped us to know the effect of mutation on the substrate binding having the protein in the native membrane environment. The disadvantage of this method is that the prepared vesicles contain many other proteins besides MelB.

As shown in Fig 4.2, there is no significant FRET signal obtained for the mutants D35C and R175C/D35C (Helix I and salt-bridge mutants), suggesting that they cannot bind any substrates in the native membrane vesicles, which is in agreement with the previous results obtained from the FRET spectroscopy using proteoliposomes.

Mutants like L164C (Helix V) and G172C (loop5-6) (Fig. 4.2), present a very similar behavior to Cys-less after a consecutive addition of D^2G, Na^+, and melibiose. These results are contradictory to those obtained from proteoliposomes which had reduced affinity compared to Cys-less. These differences in behavior may be explained by supposing that the purification and reconstitution processes lead to the loss of some important structures or interactions that are key for substrate binding.

Other mutants, i.e., P165C, F166C, and V167C in Helix V; N168C, Y169C, V170C, G171C,G173C, and D174C in Loop5-6; R175C, G176C, and G178C in Helix VI (Fig. 2) show a similar behavior of FRET response to the one measured with proteoliposomes. This confirms that all of them preserve the capability of sugar/Na^+ binding with a reduction in the intensity compared to Cys-less, suggesting that the mutations may affect the affinity of the substrates.


ISO
+10µM D2G
+10mM NaCl
+10mM Melibiose
Control Mutants
Cless
DW2R
Helix V
L164C
P165C
F166C
V167C
Loop V-VI
N168C
Y169C
V170C
G171C
G172C
G173C
D174C
Helix VI
R175C
G176C
F177C
G178C
1 a.u.
400 420 440 460 480 500 520 540 560
Wavelength (nm)

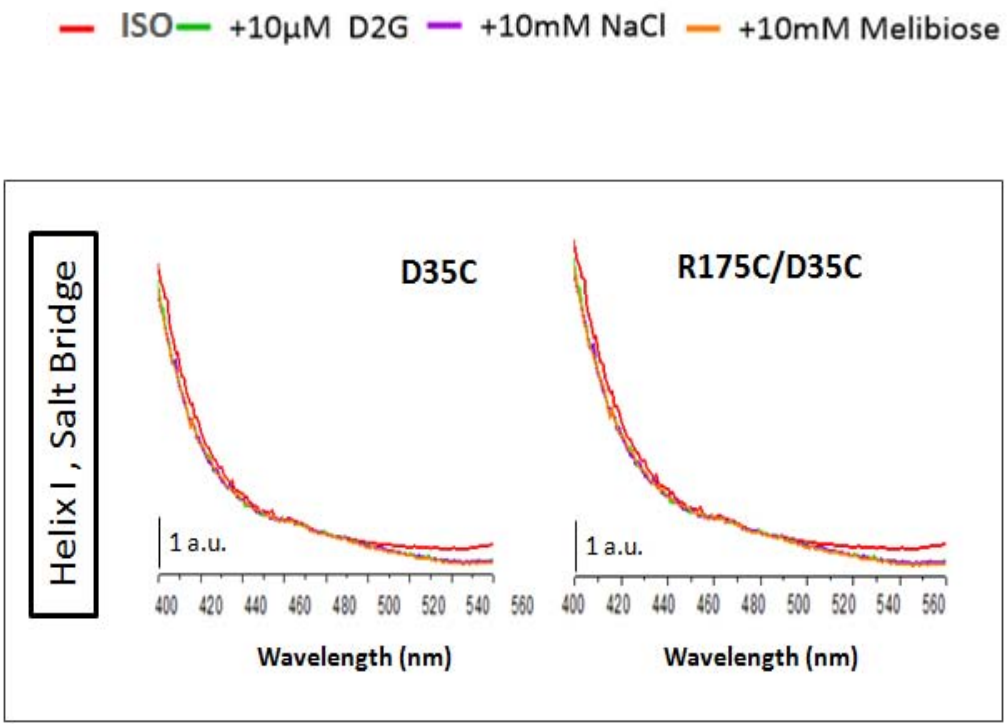


Fig. 4.2: Substrate-induced Trp-D^2G FRET signal changes of MelB mutants in ISO vesicles. The spectra are the representative plots of FRET changes of different MelB mutants in Inside-out Vesicles, induced by the sequential addition of 10 µg/ml protein in the ISO (Red), 16 µM D^2G (Green), 10 mM NaCl (Purple) and 10mM Melibiose (Orange). All the samples were excited at 290 nM (half-bandwidth 5nm), and the emission spectra were collected at 310-570 nm. The plots represented here are the average of two or three independent experiments for both ISO. Here, C-less is the positive control and DW2R (without MelB protein) is the negative control. All the spectra were normalized for the protein concentration.

Presence of MelB in ISO vesicles were accessed using Western Blot with a specific reagent called His-Probe™-HRP as in Fig 4.3.

Fig. 4.3: Western Blot analysis indicating the presence of MelB in different mutants of Inside Out Vesicles using His-Probe™-HRP. Cys-less is the positive control.

4.2. Mutagenesis and Percentage effect on substrate binding

From the proteoliposomes and inside out vesicle FRET results, we compared the effect of D^2G binding in the presence of proton (H^+) and sodium (Na^+). The results were represented as the percentage effect on substrate binding for each mutant with respect to Cys-less (Fig 4.4). The results were further categorized into three different groups: low effect (50-100%), intermediate effect (25-50%) and high effect (0-25%).

From the Table 1 and Fig. 4.4, it is very clear that mutants D35C (Helix I) and R175C/D35C (Helix VI-Helix I) have very high effect on the substrate binding with complete loss of substrate binding, while the mutant F177C in Helix VI shows a highly reduced binding affinity (less than 25%) in all the cases. This means that these residues play an important role in substrate binding. Apart from this, native vesicles of the mutant D174C (loop5-6) show a highly reduced effect on binding the sugar in the presence and absence of Na^+, which might be due to the reduced expression of MelB in vesicles compared to Cys-less (as shown in Fig 3). The mutants G172C and L164C in loop5-6 showed an intermediate binding effect in the case of proteoliposomes while the binding of substrates in the case of native vesicles is similar to Cys-less, which might be due to altered protein native conformation while solubilizing with the detergent.

Interestingly, proteoliposomes of mutants G171C, G176C and G178C showed a low effect (50-100%) on the binding of the substrates in the presence of Na^+, which means that a change in the native conformation could have happened after detergent solubilization, increasing the affinity of D^2G in the presence of sodium.

Other mutants like P165C, F166C, and V167C in Helix V; N168C, Y169C, V170C, G171C, and G173C in loop5-6; R175C in Helix VI showed an intermediate substrate binding effect (25-50%)

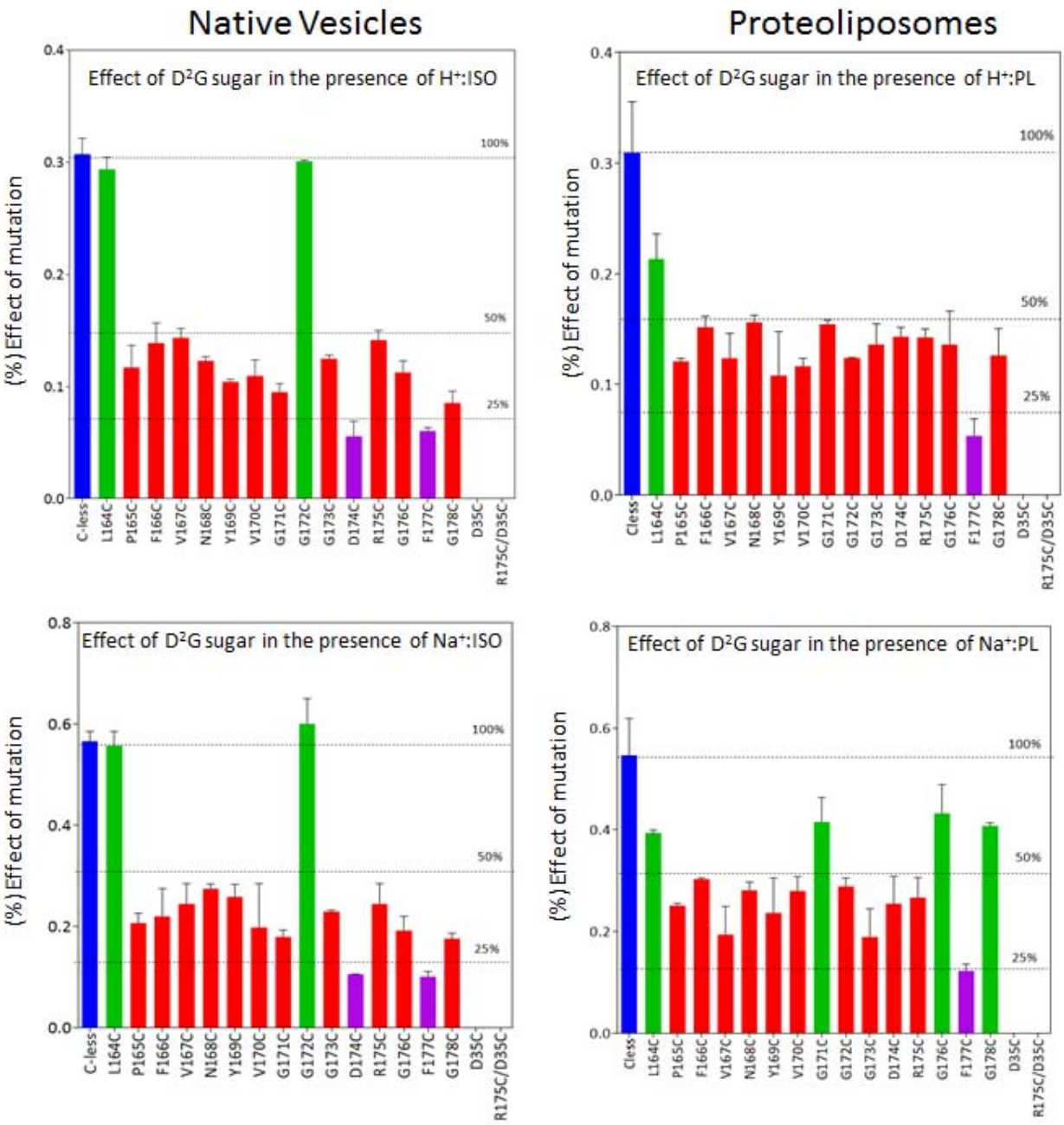


Fig. 4.4: Effect of Dansyl sugar (D^2G) binding in the presence of H+ and Na^+. **(Top)** FRET changes in Inside Out Vesicles (ISO) and Proteoliposomes (PL) in the presence of H^+. **(Bottom**) FRET changes in ISO and PL in the presence of Na. In all the cases, protein was normalized to Cys-less protein concentration (marked as blue), mutants showing low effect (50-100%) was marked as green , intermediate effect (25-50%) was marked as red, and the mutant showing high effect (0-25%) was marked as purple. The results shown here are the averages of two or three independent experiments.

Effect of Mutation	Native Vesicles	PL	Native Vesicles	PL
	D^2G in the presence of H^+	**D^2G in the presence of H^+**	**D^2G in the presence of Na^+**	**D^2G in the presence of Na^+**
50-100% **(Low effect)**	C-less, L164C, G172C	C-less, L164C	C-less, L164C, G172C	C-less, L164C, G171C, G176C, G178C
25-50% **(Intermediate effect)**	P165C -G171C, G173C-G176C, G178C	L164C-G176C, G178C	P165C-G171C, G173C, R175C, G176C, G178C.	L164C-R175C
0-25% **(High effect)**	F177C, D35C, R175C/D35C	F177C, D35C, R175C/D35C.	D174C, F177C, D35C, R175C/D35C.	F177C, D35C, R175C/D35C.

Table 1: Percentage effect of substrate binding for each mutant with respect to Cys-less in Inside Out vesicles (ISO) and Proteoliposomes (PL). Here, the mutants which shows 50-10% effect in the above bar graph are marked as 'low effect', the mutants which are in the range 25-50% are marked as 'intermediate effect' and finally the mutants which are in the range 0-25% are marked as 'high effect' due to mutation.

5. RESULTS (PART-2)

5.1 Binding analysis using ATR-FTIR

The ATR-FTIR technique is useful to determine the interaction between substrates and the protein. Previous work has shown that this technique is useful to reveal the substrate binding conformational changes in MelB [Leon *et al.,* 2006]. Since, the infra-red technique is a time consuming process, here it was applied to few important MelB cysteine mutants to detect the substrate induced conformational changes as described in Table 5.1.

Mutants	Location in MelB	Importance	Results from FRET
P165C	Helix V	Shown to be 80% conserved.	Intermediate
N168C	Loop 5-6	Only polar residue	Intermediate
D174C	Loop 5-6	Only negatively charged residue	Intermediate /High
R175C	Helix VI	Positively charged residue in the beginning of Helix.	Intermediate
D35C	Helix I	Charged residues corresponding for salt bridge interactions.	High
R175C/D35C	Helix I/Helix VI		
F177C	Helix VI	Aromatic residue.	High

Table 5.1: Shows the list of residues studied using ATR-FTIR, where PL= Proteoliposomes, ISO=Inside-Out Vesicles

The ionic, hydrophobic and polar interactions from residues in table 1 may be important for the conformational changes taking place during MelB transport. Here, substrate-induced IR_{diff} spectra were measured in response to the interaction of the coupling ion with MelB or to the interaction of the sugar with the ion-MelB binary complex [Granell *et al.,* 2010]. The difference spectra (positive and negative peaks) are obtained by subtracting the absorbance spectra in the presence of substrate from the absorbance spectra without the substrate.

Previous work done using ATR-FTIR has revealed its importance in knowing the effect of mutations on the binding of substrates to MelB by quantitative comparison of the shape and intensity of the structure sensitive amide I and amide II protein bands between 1500-1700 cm^{-1} [Leon *et al.,* 2005].

5.2. Mutagenesis and its effect on MelB structure:

FTIR is a powerful technique for the examination of protein secondary structure and structural changes [Surewicz *et al.*1993]. From previous results, it has been shown that the mutation of some of the residues alters the binding of substrates. In the present work IR absorbance spectra of structure sensitive amide I and amide II protein bands (1500-1700 cm^{-1}) for each mutant were compared with the Cys-less spectrum. The structural information of the protein can be obtained from amide I and amide II bands, mostly C=O stretching and N-H bending coupled with C-N stretching of the peptide bond, respectively.

The lipid-to-protein ratio can be deduced from the relative intensity of the C=O stretching of the phospholipids ester. The 2:1 (w/w) ratio between lipids and purified MelB represents an active and stable transporter [Granell 2009]. A low ratio indicates a failure of the reconstitution process of the solubilized protein into the liposomes which can lead to an inactive sample.

The absorbance spectrum (Fig. 5.1 A) of each mutant has very similar broad amide I and II region spectrum when compared to the reference spectrum of fully functional MelB. The second derivative (Fig. 5.1 B) of the absorbance spectrum was computed using the Savitzky-Golay algorithm for each MelB mutant to observe the changes caused by the mutagenesis. Each second derivative spectrum was compared with Cys-less to obtain

the spectral similarity (m) and intensity (R^2) values. The slope m further used to normalize the difference spectra for protein concentration (Fig. 5.1 D).

From the Fig 5.1 C. it is very clear that the mutants D174C and F177C have a very similar structure with respect to Cys-less which means mutation of this residues have not contributed to any structural major alterations. Whereas, the mutants P165C and N168C has led to a shift of amide I band at 1659 cm^{-1} to 1655 cm^{-1} which mainly corresponds to α-helices. As the difference of structure is very minuscule, the reasonable possibility would be that mutation of this residue would have altered the surroundings of the side chains that may be involved in interactions with other residues. Lastly, mutants like R175C, D35C and R175C/D35C has led to a shift and increase in intensity of band at 1659 cm^{-1} and 1635 cm^{-1} which are mainly for -helices and β-sheets.

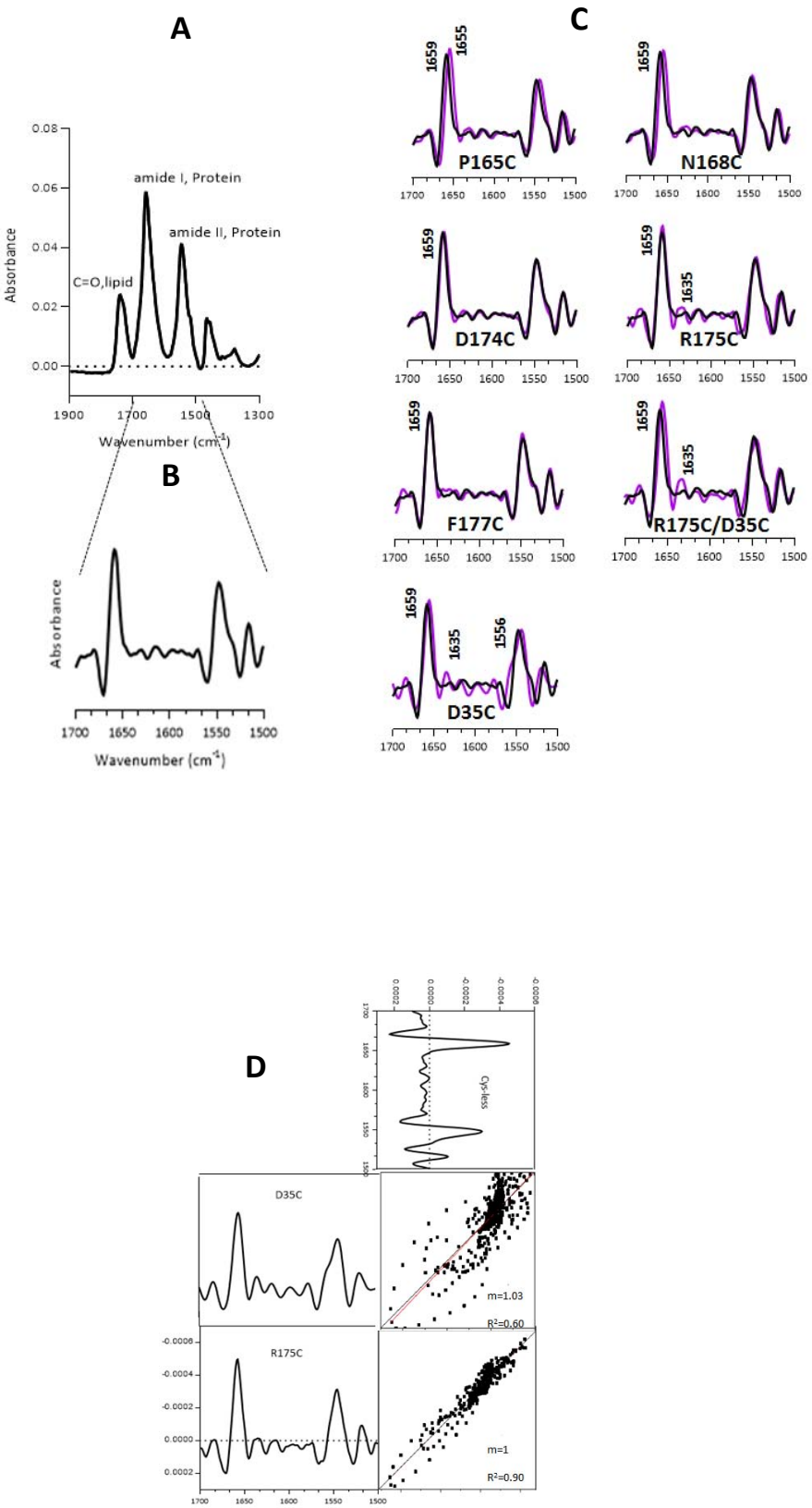

A
amide I, Protein
amide II, Protein
C=O,lipid
Absorbance
Wavenumber (cm-1)
B
Absorbance
Wavenumber (cm-1)
C
P165C
N168C
D174C
R175C
F177C
R175C/D35C
D35C
D
D35C
R175C
m=1.03
R²=0.60
m=1
R²=0.90

Fig. 5.1. Structural comparison of MelB mutants using IR spectroscopy of hydrated samples**. (A)** The raw absorbance spectrum of Cys-less at 4 cm^{-1} resolutions\ in a hydrated film. **(B).** The second derivative of the raw spectrum using Savitzky-Golay algorithm at 8 cm^{-1} showing structure sensitive amide I and amide II peaks at 1700-1500cm^{-1}. **(C)** Shows the comparison of Cys-less second derivative spectra with the MelB mutants. **(D)** The second derivative at 8 cm^{-1} is used for structural comparison between two spectra with respect to Cys-less. Linear correlation analysis was measured for amide I/II regions. The R^2 gives the similarity while the m represents their relative intensity. In this case m is always 1 since the protein concentration was normalized to Cys-less.

5.3. Structural changes in MelB due to substrate binding:

5.3.1. Sodium (Na^+) binding

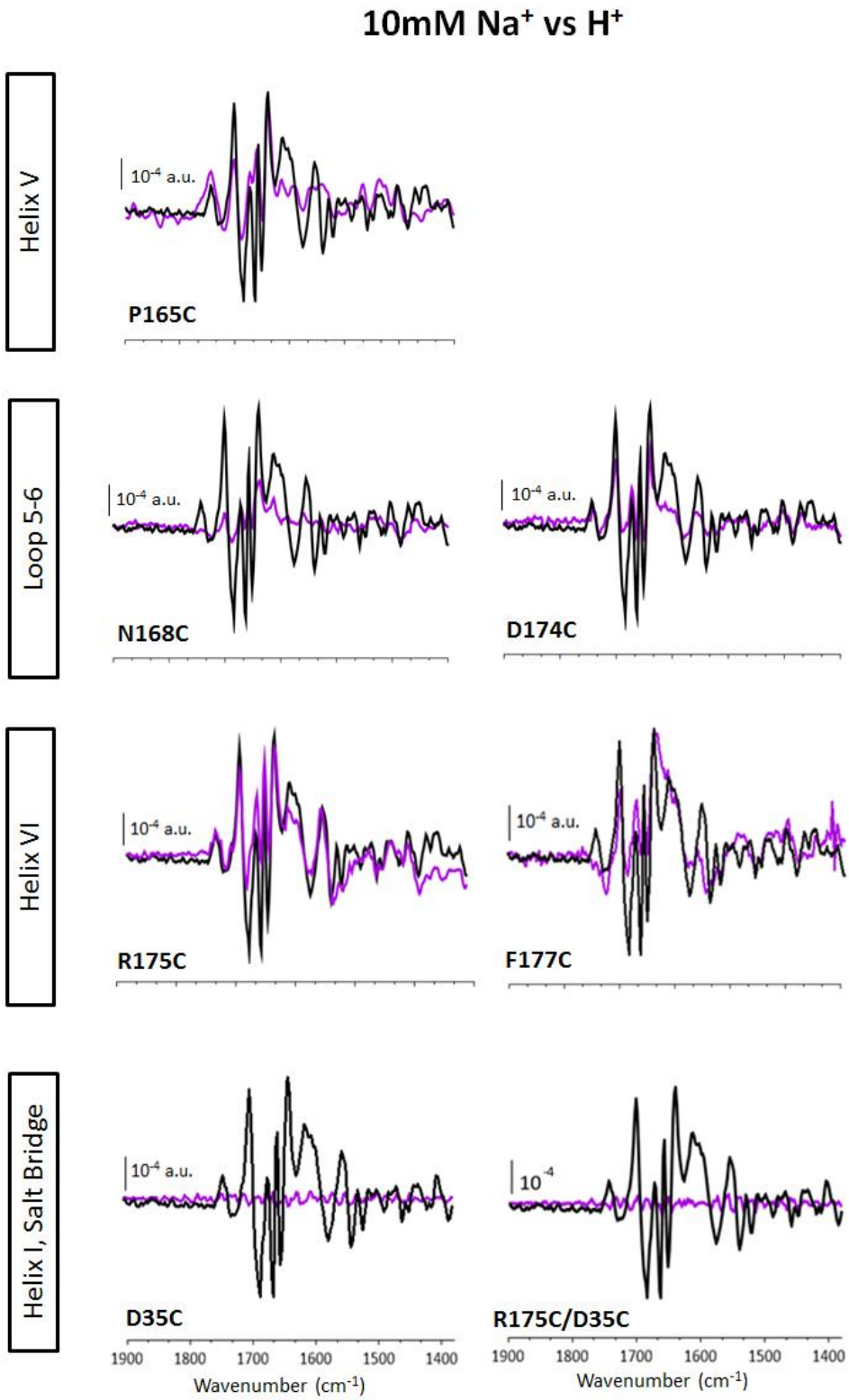
The purified MelB transporter reconstituted into liposomes was used to study the binding of substrates using Infra-red difference spectroscopy, where buffers were supplemented with different concentration of substrates. As indicated in Granell *et al.* 2010, 10mM Na^+ (which is three fold above the affinity constant) was added to the positive control C-less (Fig. 5.2 Black) which produces a reproducible difference spectrum. The difference spectrum is the subtraction between the absorption spectra in the presence of proton and the absorption spectra in the presence of sodium.

Peaks in the difference spectra not only represent Na^+ interaction with important residues but also the structural changes responsible for the increase in affinity for the sugar binding. The peaks at 1703 cm^{-1} is due to vibration of aspartic and glutamic acids. And, the peaks at 1693, 1688, 1680 and 1674 cm^{-1} are primarily provoked by β-sheet structure. Peaks at 1668, 1659 and 1652 cm^{-1} is due to α-helix. A featureless straight line spectra indicates the absence of interaction of the cation with MelB.

The difference spectra are usually obtained in the broad range from 4000-1000 cm^{-1}. But the most important protein sensitive region is amide I (~1650 cm^{-1}) and amide II bands (~1550 cm^{-1}) which provides a fingerprint to characterize the structural response of MelB mutants to Na^+ binding with respect to C-less.

As shown in the Fig 5.2A., the difference spectra obtained in the presence of 10mM Na^+ for the mutants D35C and R175C/D35C shows a straight line indicating that mutation of these residues completely abolishes sodium binding. To further confirm this effect, the difference spectra was also acquired in the presence of 50mM Na^+ (Fig 5.2B), 25 times the $K_{0.5}$ of the wild type transporter [Ganea *et al*., 2009]. These results prove that Asp-35 and the salt bridge residue Arg175/Asp35 are essential residues for Na^+ binding.

Whereas, other mutants display a Na^+-induced difference spectrum with a lack of intense peaks at 1662 cm^{-1},1657 cm^{-1} and 1651 cm^{-1} which were assigned to α-helices (Fig 5.2A). The results indicate that they retain the ability to bind Na^+, implying that these residues are not strictly essential for Na^+ binding.


10mM Na+ vs H+
Helix V
Loop 5-6
Helix VI
Helix I, Salt Bridge
10-4 a.u.
P165C
N168C
D174C
R175C
F177C
D35C
R175C/D35C
10-4
1900 1800 1700 1600 1500 1400
Wavenumber (cm-1)

A

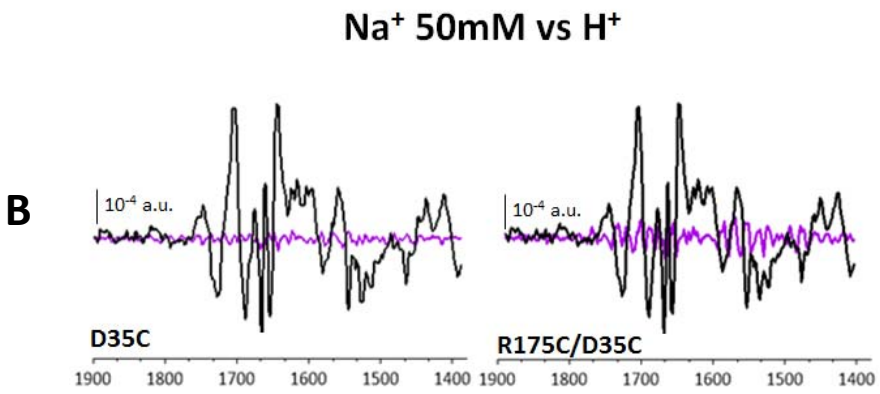


Fig. 5.2. (A) Sodium Induced IR_{diff} spectra in the presence of 10 mM Na^+ at 4 cm^{-1} resolutions of different MelB mutants (Violet) compared to Cys-less (Black). Difference spectra was obtained by constant perfusion of buffer with and without 10 mM Na^+ (20 mM MES, 110 mM KCl and 10 mM Na^+ pH 6.6). **(B)** Sodium induced IR_{diff} spectra in the presence of 50mM Na^+ (20 mM MES, 110 mM KCl and 50 mM Na^+ pH 6.6) at 4 cm^{-1}. All the difference spectra were normalized to the amount of protein probed. The difference spectra are the average of two or three independent experiments.

5.3.2. Sugar (Melibiose) binding in the presence of H^+

We then analyzed how mutation of each residue affects the transporter ability to bind the melibiose sugar in the presence of H^+. In these experiments the sugar was added at a concentration of 50mM, a value close to the half saturating concentration of Cys-less for melibiose [Meyer-Lipp *et al*., 2006].

The white colonies on the MacConkey agar plates for the mutants D35C and R175C/D35C indicated the absence of sugar metabolism in the bacteria. This was once again proved by the results obtained from FRET studies. To complement these results the binding affinity of these residues were also studied using IR_{diff} spectroscopy. As shown in the Fig. 5.3. the difference spectra of D35C and R175C/D35C mutants did not show any peak assignable to protein in its difference spectrum (1700-1500 cm^{-1}) induced by Melibiose. Briefly, the data indicate that the mutation of Asp-35 and the Asp-35/Arg-175 shows no conformational changes in response to the addition of Melibiose or it is similar to the response obtained by the E. coli lipids without MelB, which further indicates that these residues are very essential for sugar binding.

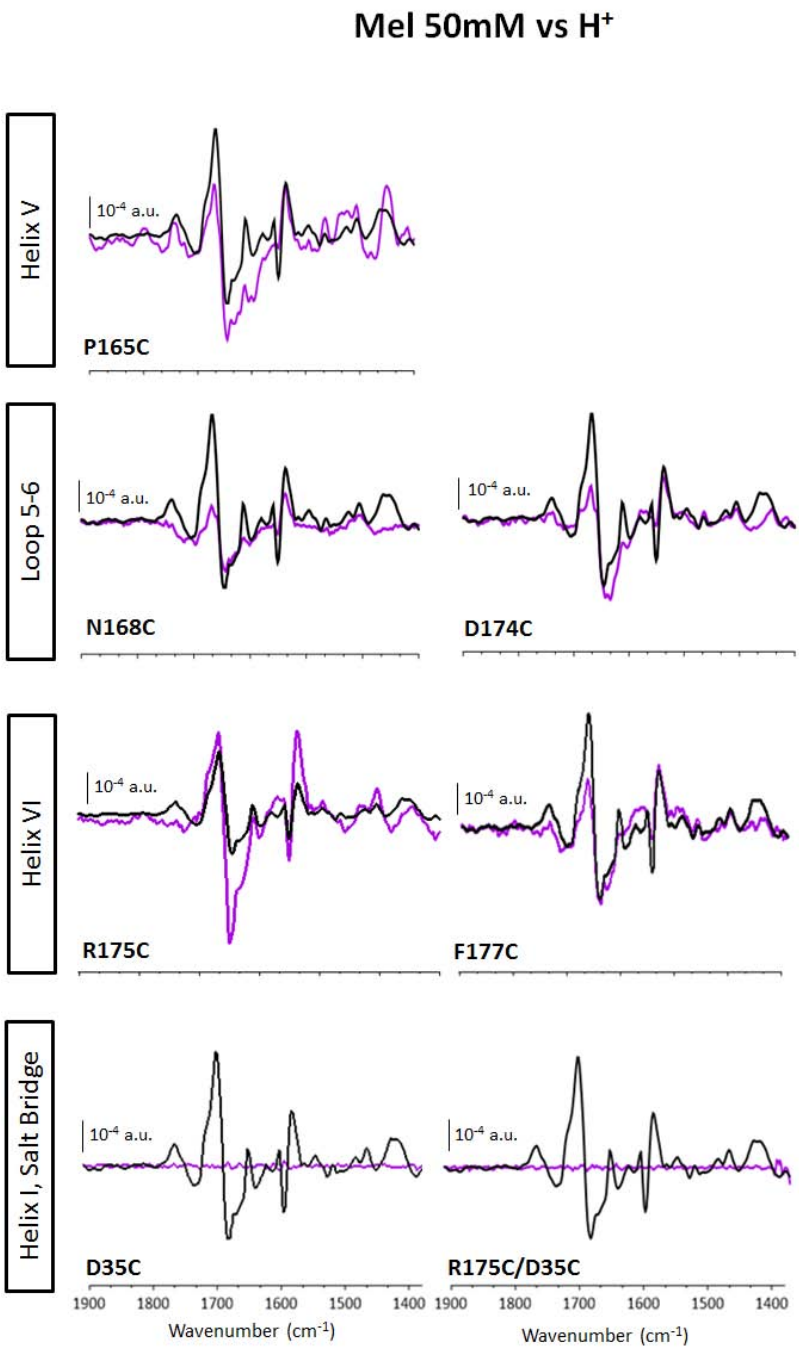


Fig. 5.3. 50 mM Melibiose Induced IR_{diff} spectra in the presence of H^+ at 4 cm^{-1} resolution of different MelB mutants (Violet) compared to Cys-less(Black). Difference spectra was obtained by constant perfusion of buffer with and without melibiose (20 mM MES, 100 mM KCl and 50mM Melibiose pH 6.6). All the difference spectra were normalized to the amount of protein probed. The difference spectra are the average of two independent experiments.

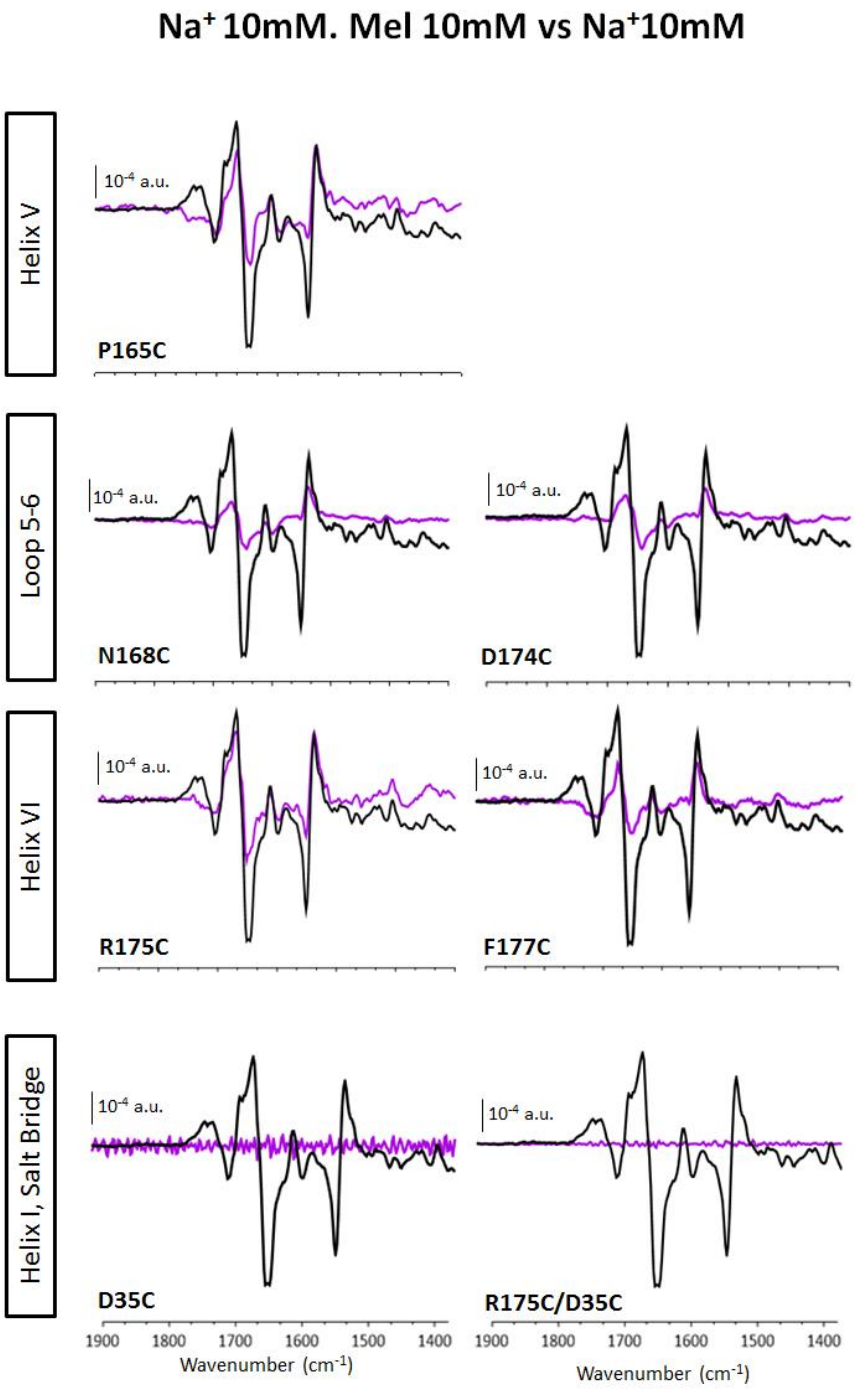
Similarly, to the Na^+ binding, other mutants display a melibiose-induced difference spectrum indicating that they retain the ability to bind melibiose, implying that these residues are not strictly essential for sugar binding.

5.3.3. Sugar (Melibiose) binding in the presence of Na^+:

We also verified if the presence of sugar generates structural changes in the mutants in the presence of co-substrate sodium. Here, the sugar was added at a concentration of 10mM in the presence of 10mM Na^+. From previous studies, it has been proven that the sugar binding affinity increases in the presence of a co-substrate cation. This IR_{diff} difference in the presence of Na^+ reflects the coupling between the sugar and cation binding sites, which corresponds to the second stage of the transport cycle after binding of Na^+. Melibiose induced difference spectra in the presence of Na^+ have further peaks at 1643 and 1631 cm^{-1}.

As seen in the Fig. 5.4A the mutants D35C and R175C/D35C again did not show any conformational rearrangements as that of C-less in the region 1700-1500 cm^{-1}, which was once again concluded by the flat IR_{diff} spectrum similar to that E. coli lipids without MelB protein. As shown in Fig 5.4B we also checked whether the mutants Asp-35 and Arg-175/Asp-35 to cysteine will show any binding affinity with the higher concentration of melibiose in the presence of 10 mM Na^+(Mel.50mM vs Na^+.10mM). But, the mutants generated a feature-less difference spectrum after subtracting the background noise like effect of buffer and the substrate. All these results indicate that these residues are very essential for substrate binding. As in the other cases, all other mutants display a difference spectra induced by melibiose in the presence of Na^+ which indicates that they retain the ability to bind melibiose in the presence of Na^+.

A


Na+ 10mM. Mel 10mM vs Na+10mM
Helix V
Loop 5-6
Helix VI
Helix I, Salt Bridge
10-4 a.u.
P165C
N168C
D174C
R175C
F177C
D35C
R175C/D35C
1900 1800 1700 1600 1500 1400
Wavenumber (cm-1)

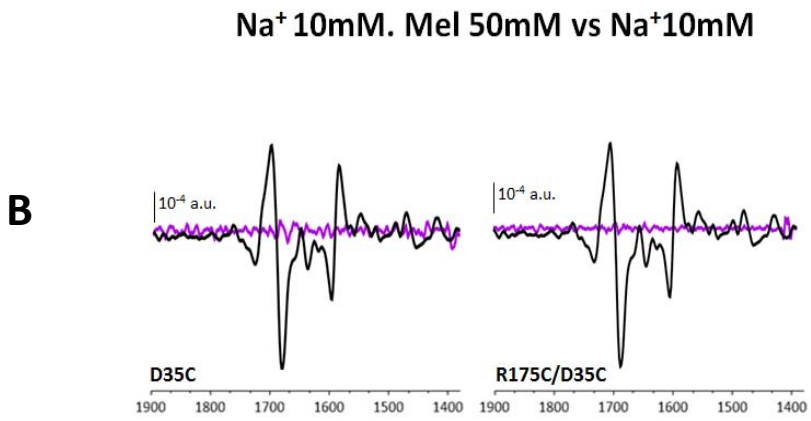


Fig. 5.4: (A) 10 mM Melibiose induced IR_{diff} spectra in the presence of 10 mM Na^+ at 4 cm^{-1} resolution of different MelB mutants (Violet) compared to Cys-less (Black). Difference spectra was obtained by constant perfusion of buffer with and without Melibiose (20 mM MES, 100 mM KCl, 10 mM NaCl and 10 mM Melibiose pH 6.6). **(B**) 10 mM Melibiose induced IR_{diff} spectra in the presence of 50mM Na^+ at 4 cm^{-1} resolution for D35C and R175C/D35C mutants. All the difference spectra were normalized to the amount of protein probed. The difference spectra are the average of two independent experiments.

5.4. Quantitative evaluation of the difference Spectrum:

The IR_{diff} spectra of the MelB mutants can be quantitatively compared as unbiased as possible by means of a linear regression analysis encompassing the structure-sensitive 1700-1500 cm^{-1} region from their difference spectra [Granell *et al.,* 2010]. The linear correlation parameter, R^2 represents the spectral similarity of a mutant response relative to Cys-less and m which is the slope of the linear correlation analysis gives the relative intensity of the spectral features in common with Cys-less intensity. A reduced intensity lower than 100% for any given mutant implies either a reduced affinity for the added substrate.

The similarity and intensity graph was plotted for the above described results of the IR_{diff} spectra.

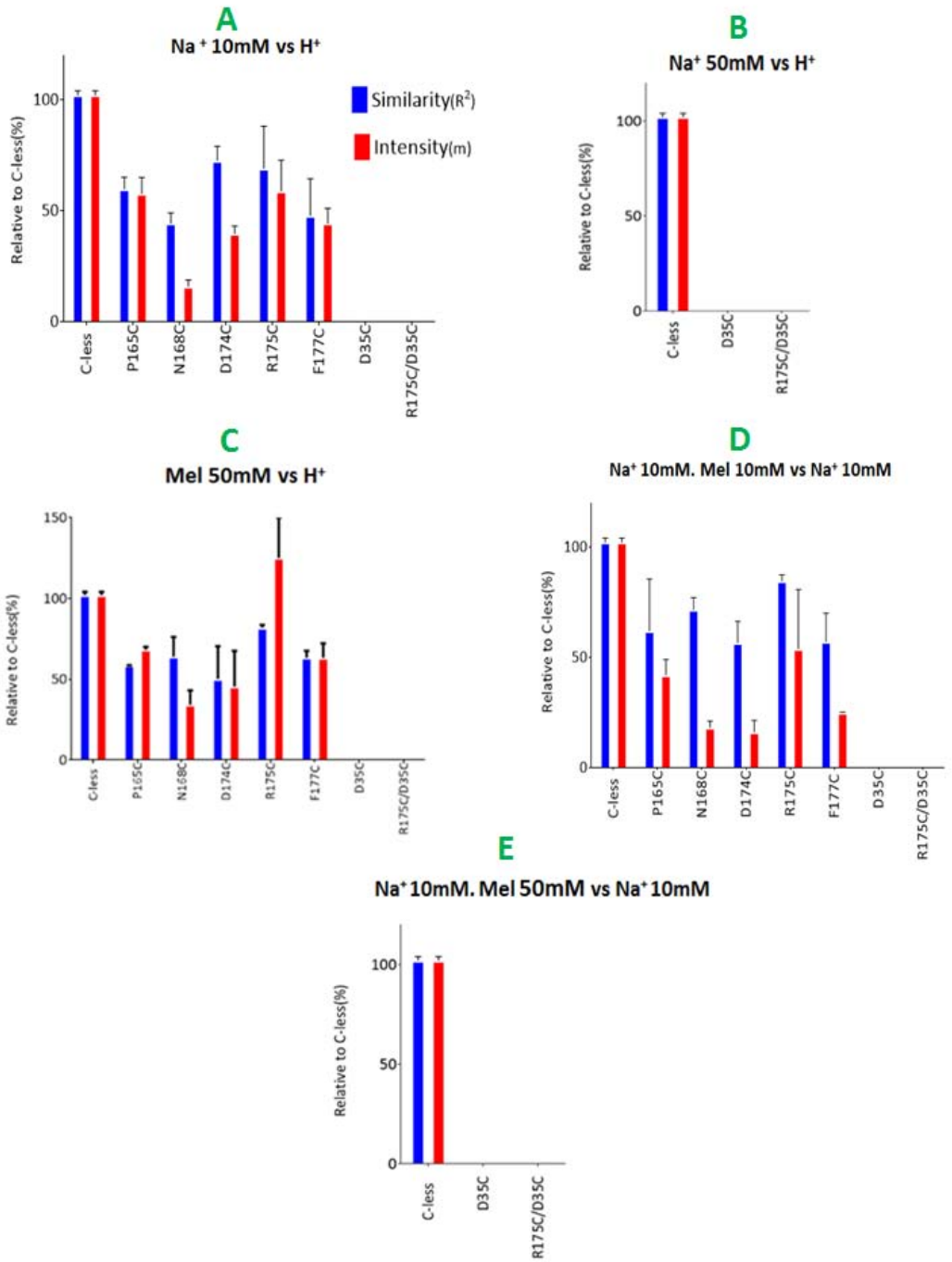


Fig. 5.5: The bar graph represents the spectral similarity (m) and intensity (R^2) of **(A) 10 mM Na^+-** induced IR difference spectra **(B) 50 mM Na^+-** induced IR difference spectra **(C) 50 mM Melibiose-** induced IR difference spectra in the presence of H^+ **(D) 10 mM Melibiose** induced spectra in the presence of Na^+ **(E) 50 mM Melibiose** induced spectra in the presence of Na^+. The error bar corresponds to one standard error of the mean.

5.4.1. Na^+. 10 mM vs H^+ and Na^+. 50 mM vs H^+

As shown in Fig 5.5A and 5.5B., two mutants D35C and R175C/D35C did not show any clear intensity and similarity values when compared to Cys-less, which means that these residues don't produce any conformational changes typical for Na^+ binding. These results indicate that these residues are inactive to Na^+ binding. So, the side chains of these residues are very important for the conformational changes necessary for Na^+ binding.

P165C, 80% conserved mutant (refer Appendix) in Helix V show moderate intensity and similarity (50%) when compared to Cys-less which means that they are not strictly essential for Na^+ binding, but this may participate in the configuration of Na^+ pocket. These results are in agreement with the previous work done by Yazyu et al., 1985, where the mutation of Pro-122 to serine caused a structural change of the carrier and alterations in cation and substrate specificities.

N168C, mutant in loop 5-6 shows a reduced similarity (~40%) and highly reduced intensity when compared to Cys-less. This result strengthens the notion that Asn residues are one of the important candidates for forming the Na^+ co-ordination network in MelB [Poolman *et al.,* 1996]. Aromatic mutant in the loop 5-6, D174C shows reduced intensity (~40%) and moderate similarity (~70%) indicates that this partially retain the Na^+ binding with less complete structural changes necessary for complete Na^+ binding when compared with Cys-less. Surprisingly, Arg175, Helix VI residue involved in salt bridge formation with D35 shows a moderate similarity and intensity (70-80%) when mutated to cysteine. The results again mean that this residue retains the binding of Na^+. Lastly, an aromatic mutant F177C in Helix VI, display reduced similarity and intensity (~40%). Therefore, this residue is not essential for Na^+ binding, although Na^+-induced structural changes is less complete when compared to Cys-less. Thus, this residue is required for full and native-like structural changes in MelB.

5.4.2. Mel 50mM vs H^+

As shown in Fig 5.5C, two mutants D35C and R175C/D35C show insignificant intensity and similarity when compared to Cys-less, which means that these residues don't produce any conformational changes typical for melibiose binding. These results indicate that these residues are completely inactive to melibiose binding.

P165C, a conserved mutant in Helix V show moderate intensity and similarity (50-65%) when compared to Cys-less which means that they are not strictly essential for melibiose binding, but this may participate in the configuration of sugar binding pocket. Unlike Na^+ binding, N168C, mutant in loop 5-6 shows an improved intensity (~30%) and moderate similarity (60%) when compared to Cys-less. This suggests that this residue is not an important candidate for sugar binding. D174C, a loop 5-6 mutant shows moderate intensity and similarity (~40-50%) indicates that this retains the melibiose binding with less complete melibiose induced structural changes when compared with Cys-less. R175C, Helix VI mutant shows a moderate similarity and high intensity (70-130%) when compared to cys-less. The results again mean that this residue retains the binding of melibiose. An aromatic mutant F177C in Helix VI, display moderate similarity and intensity (~60%). Therefore, this residue is not important for melibiose binding, although melibiose induced structural changes is less complete when compared to Cys-less.

5.4.3. Na^+ 10mM. Mel 10mM vs Na^+10mM and Na^+ 50mM. Mel 50 mM vs Na^+ 10mM

As shown in Fig. 5.5D and 5.5E (higher concentration of melibiose), two mutants D35C and R175C/D35C did not show any intensity and similarity values when compared to Cys-less. These results indicate that residues are inactive to melibiose binding even in the presence of Na^+, which clearly confirms that these residues are highly essential for substrate binding. As shown in Fig 5.5D, mutants in Helix V, loop 5-6 and Helix VI (P165C, N168C, D174C, R175C and F177C) display moderate to high similarity (50-88%) when compared to previous quantitative analysis with melibiose induced only in the presence of H^+. This means that in these residues structural changes necessary for melibiose binding increases in the presence of Na^+. The spectral intensity remains moderate for

P165C like previous case, while for the mutants N168C, D174C, R175C and F177C it reduces drastically. This implies that these residues have reduced affinity for melibiose binding in the presence of Na^+.

5.5. Analysis of spectral data using SI maps

To further complete the characterization of the mutants, we performed a comparison of that in the form of a shape/intensity map (SI map). In the SI map the relative percentage of differences in similarity with respect to C-less are plotted against the relative percent of differences in intensity. Each mutant on the SI graph is characterized by its s and i coordinates. Detailed explanation of this is given in the Materials and Methods Section. In brief, it can be explained as spectra with a pair of coordinates (50, 0) will have the same similarity index as a spectrum with a pair (0, 50). In both the cases, magnitude of effect on native binding is the same. The similarity scale defined is given in a color code. Using this code, the mutants are further classified based on three groups (high, intermediate and low) with respect to the substrate binding as indicated in Table 5.2.

5.5.1. Effect of 10mM Na^+ binding to MelB

In the Fig. 5.6B, SI map was plotted with the values of intensity and similarity from the bar graph for each mutant. It clearly shows that the mutants P165C, D174C and F177C has an intermediate effect on Na^+ binding with a similarity index of ~40-60%, and the mutant N168C, D35C and R175C/D35C has a high effect on sodium binding with the similarity index between 0-40% while the mutant R175C has a very low effect on Na^+ binding with a similarity index of more than 60% with respect to C-less.

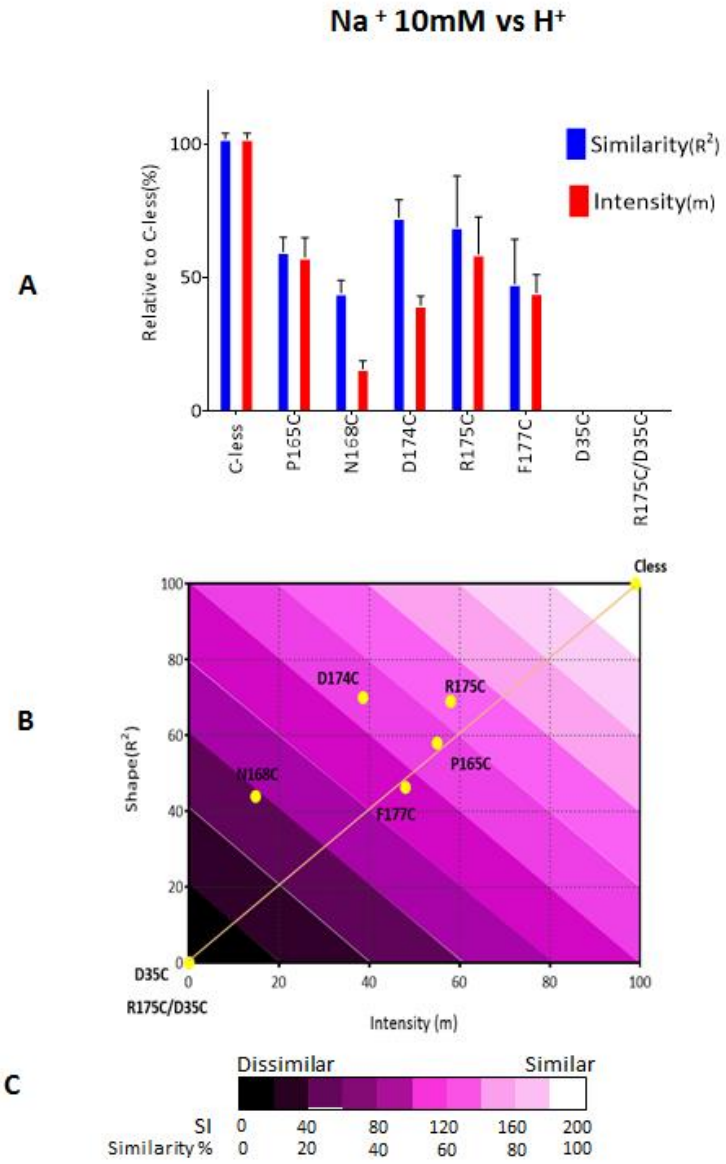


Fig. 5.6: **(A).** The bar graph represents the spectral similarity (R^2) and intensity (m) of the mutants compared to Cys-less by linear regression analysis of its first derivative spectra in the presence of 20 mM MES, 110 mM KCl, and 10mMNaCl pH 6.6. The error bar represents the variations depending on at least two independent data acquisitions. **(B)** The graph represents the SI (Shape/intensity) map. Here the shape means "similarity ". In the SI map the differences in percentage similarity from the above bar graph are plotted against the percentage intensity. The mutants lying on the map from the diagonal upwards have a difference spectra dominated by intensity and diagonal downwards shows spectra dominated by similarity. **(C)** The color scale shows the percentage of similarity- darker is more dissimilar and lighter is close to similar.

5.5.2. Effect of Melibiose binding to MelB:

As shown in Fig. 5.7B, SI map plotted shows that the mutants P165C, and F177C has an intermediate effect of sugar binding with a similarity index of ~40-60%, and the mutants D174C, N168C, D35C and R175C/D35C has a high effect on sugar binding with the similarity index between 0-40% and the mutant R175C has a very low effect on melibiose binding.

5.5.3. Effect of sugar binding to MelB in the presence of sodium:

As shown in Fig. 5.8B, SI map was plotted with the values of intensity and similarity from the bar graph for each mutant. It clearly shows that the mutants P165C, N168C, D174C and F177C has high effect on sugar binding in the presence of Na^+ with a similarity index of approx. 0-40% along with the mutant D35C and R175C/D35C. The mutant P165C shows intermediate effect with SI index of around 40-60% and the mutant R175C has a very low effect on Na^+.Mel binding with a similarity index of more than 80% with respect to C-less.

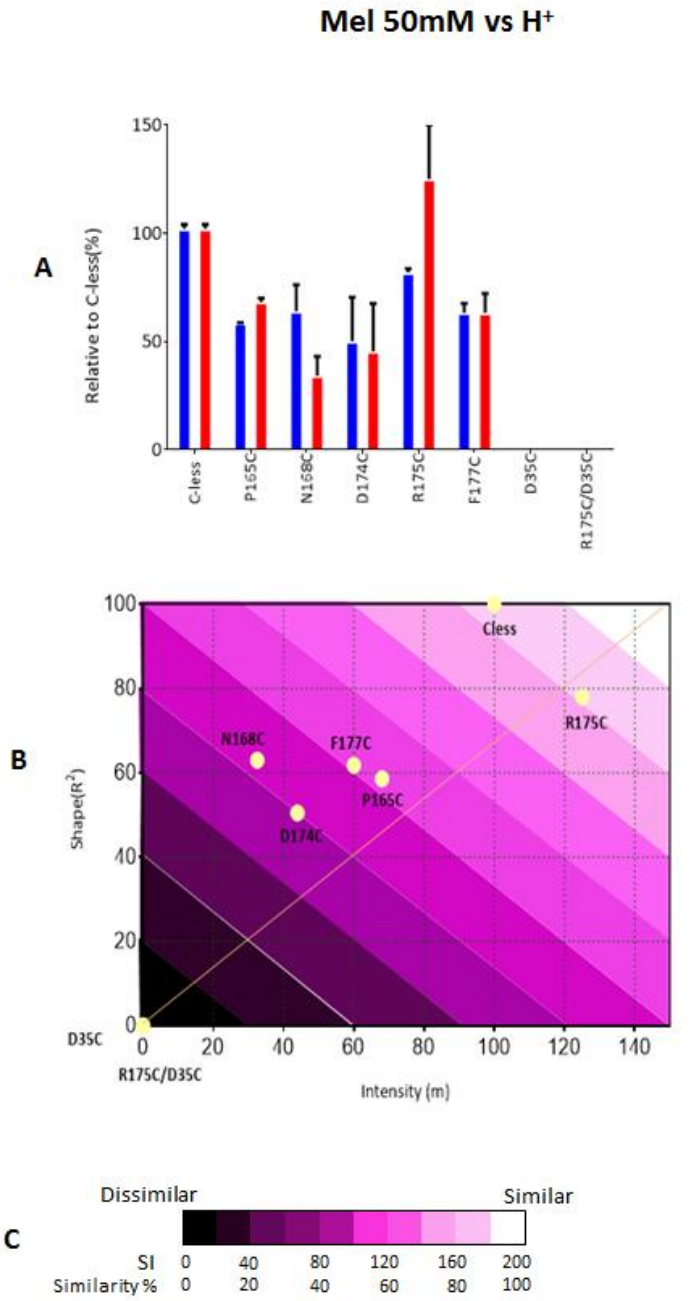


Fig. 5.7: (A). The bar graph represents the spectral similarity (R^2) and intensity (m) of the intensity of mutants compared to Cys-less by linear regression analysis of its first derivative spectra in the presence of 20 mM MES, 100 mM KCl, and 50 mM Melibiose pH 6.6. The error bar represents the variations depending on at least two independent data acquisitions. **(B)** The graph represents the SI (Shape/intensity) map. Here the shape means "similarity ". In the SI map the differences in percentage similarity from the above bar graph are plotted against the percentage intensity. The mutants lying on the map from the diagonal upwards have a difference spectra dominated by intensity and diagonal downwards shows spectra dominated by similarity. **(C)** The color scale shows the percentage of similarity- darker is more dissimilar and lighter is close to similar.

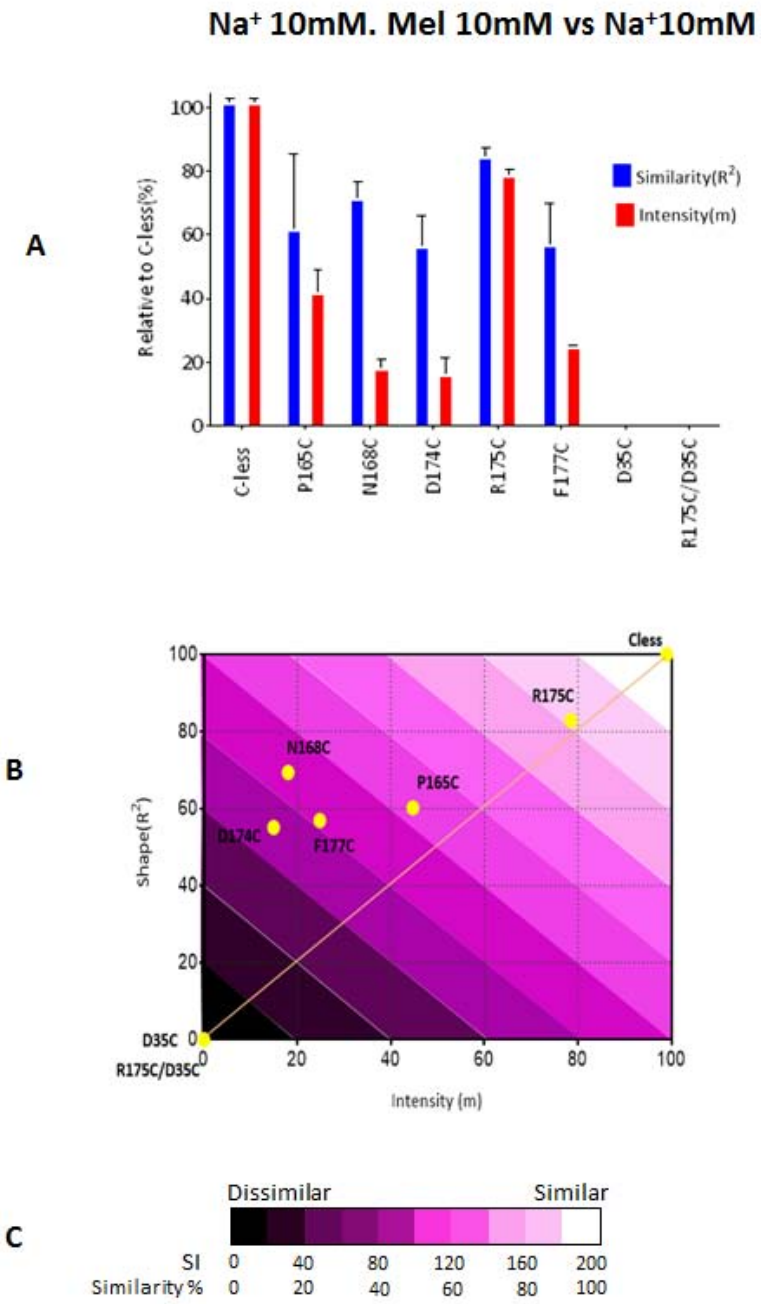


Fig. 5.8: (A). The bar graph represents the spectral similarity (R^2) and intensity (m) of the intensity of mutants compared to Cys-less by linear regression analysis of its first derivative spectra in the presence of 20 mM MES, 100 mM KCl, 10 mMNaCl and 10mM Melibiose pH 6.6. The error bar represents the variations depending on at least two independent data acquisitions. **(B)** The graph represents the SI (Shape/intensity) map. Here the shape means "similarity ". In the SI map the differences in percentage similarity from the above bar graph are plotted against the percentage intensity. The mutants lying on the map from the diagonal upwards have a difference spectra dominated by intensity and diagonal downwards shows spectra dominated by similarity. **(C)** The color scale shows the percentage of similarity- darker is more dissimilar and lighter is close to similar.

Effect on Binding	Similarity %	Na^+ Binding	Mel. Binding	Na^+. Mel Binding
High	**0-30**	R175C/D35C, D35C, N168C	R175C/D35C, D35C	R175C/D35C, D35C
Intermediate	**30-60**	P165C, F177C, D174C.	P165C, N168C D174C, F177C	P165C, N168C, D174C, F177C
Low	**60-100**	C-less, R175C	C-less, R175C	C-less,R175C

Table 5.2: Represents the effect of substrate binding in Melibiose Permease.

5.6. Transport assay in Melibiose Permease

Radioactive assays further validate how the mutation of the residues affects the substrate binding. In this study, mutants of some residues that showed weak or intermediate substrate binding in FRET and FTIR studies (N168C, R175C, D35C and F177C) were selected for transport studies. The principle behind this assay is to measure the sugar transport using radioactive [^{3}H] melibiose in the presence of Na^+.

From the Fig. 5.9, we can safely conclude that mutations of Asn-168, Arg-175 and Phe-177 shows reduced the sugar transport in the presence of cation when compared with Cys-less, which are in a good agreement with the results obtained from binding studies using spectroscopic approaches. Similarly, the mutation of Asp-35 completely abolishes the transport of sugar in MelB which are in agreement with biophysical studies. These data are in agreement with the results published by Amin *et al.,* 2014, which shows that mutant D35C in $MelB_{ST}$ completely abolished the transport activity.

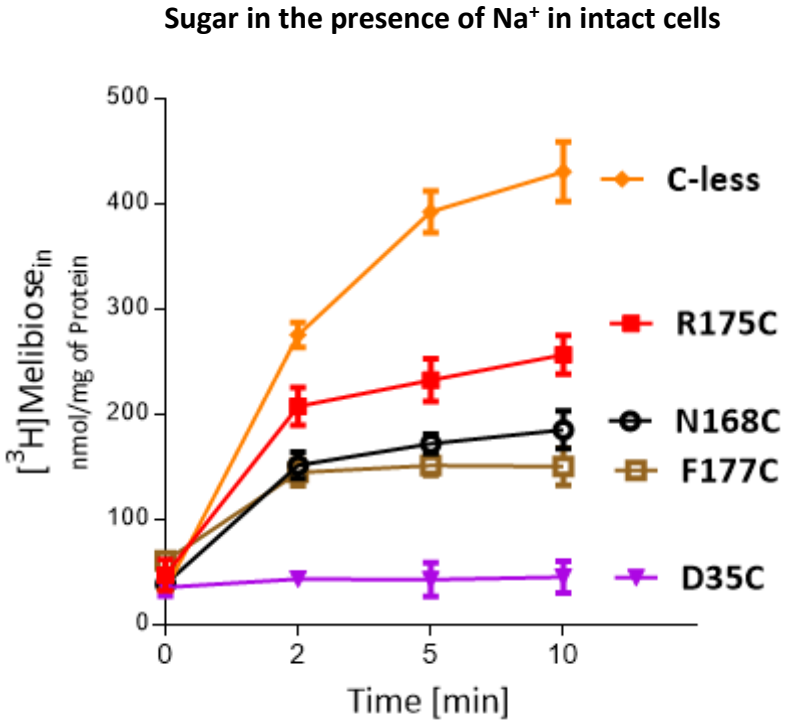


Fig. 5.9: [^{3}H] Melibiose transport assay were carried out in E.coli intact cells of the mutants In C-less (positive control), Helix VI (R175C and F177C), Loop 5-6 (N168C) and Helix I (D35C). Cells was assayed using cells (0.7mg/ml concentration) re-suspended in 0.1M KPi pH 7. Transport was started immediately after the addition of [^{3}H] Melibiose in the presence of Na^+ to a final concentration of 0.4mM, respectively. The reaction was quenched by the rapid dilution.

6. Results (Part-3)

6.1 Molecular Dynamics Simulations in MelB

The transporters are integral membrane proteins which are involved in the exchange of molecules across the cellular membranes in all the living organisms. Roughly 30% of the human genome encode for the membrane proteins (MP´s). But, MP´s are underrepresented in the protein data bank (PDB): of the 100,000 protein structures only approximately 2100 are MP. This scarcity is due to the challenges that are faced during the production of high quality protein for crystallization such as, failure of over expression or the choice of detergent for solubilization. These limitations make that it can take decades to determine the structure of MP. One good example to this is the structure of Melibiose Permease. After a long journey, recent work done by Ethayathulla *et al.,* 2014 solved the structure of MelB in *Salmonella typhimurium* (4M64.pdb).

In order to know more in detail about E. *coli* MelB, we have built an atomistic model of *E. coli* MelB using the 3D coordinates of MelB Salmonella typhimurium by mutating the required amino acids to those of *E.coli* MelB using the CHARMM GUI web program. These orthologues share 85% sequence identity and all the essential amino acids are conserved. We have embedded MelB in a 1-Palmitoyl-2-oleoyl-sn-glycero-3-phosphoethanolamine (POPE) lipid bilayer under 150 mM NaCl ionic strength and solvated with TIP3P waters [Fuerst *et al.,* 2015]. The reason behind using POPE was to mimic the experimental conditions where the major head group component of E.*coli* membrane lipids is phosphatidylethanolamine. Initially the system was energy minimized and equilibrated for 200 ps under NVT conditions using the NAMD program. This was followed by molecular dynamics simulations using the ACEMD program for a minimum of 150 ns at constant pressure and temperature using Particle Mesh Ewald method.

MelB transporter describes cycles with many different conformational forms and it is extremely difficult to have crystal structures of different intermediates. Thus, the combinations of experimental methodologies with computational approaches like molecular dynamics are commonly used to understand the conformational dynamics of membrane protein in atomic detail.

6.2. Analysis of Trajectories

6.2.1 Stability analysis: Root Mean Square Deviation (RMSD)

RMSD is the average displacement of the atoms at an instant of the simulation relative to a reference structure, usually the first frame of the simulation or the crystallographic structure. RMSD is defined as:

$$RMSD = \Sigma_N \left(\frac{(R_i - R_i^0)^2}{N} \right)$$

Where N is the total number of atoms/residues considered in the calculation and R_i is the vector position of particle i and R_i^0 is the position of the particle in the initial state.

As shown in Fig 6.1, we saw the transition from substrate-free outward partially occluded to fully open state and is characterized by large Cα distance between Tyr-33 and Tyr-256 in two out of twelve MD simulations replicas (as shown in Fig 6.1B). For reference we compared partially occluded structures with-XylE and MelB and fully open structures - with YarJ and FucP (Fig 6.1C).

From the twelve MD simulations (Fig 6.1A), we chose one occluded (Fig 6.1A S3) and one open conformation (Fig 6.1A S2) and analyzed RMSD by computing the C_α atoms using the first frame of simulation as a reference in VMD- RMSD trajectory tool (Fig 6.2). In S3, the RMSD curves did not change significantly from the starting structure and stabilized after 60 ns.

This means that the protein structure is very stable and does not diverge from the initial structure (Fig. 6.2). In the case of simulation 2, the RMSD curve has a higher value than the initial structure around 60 ns which clearly proves that the protein has undergone some higher structural changes than those of simulation 3, in agreement with the opening of the transporter.

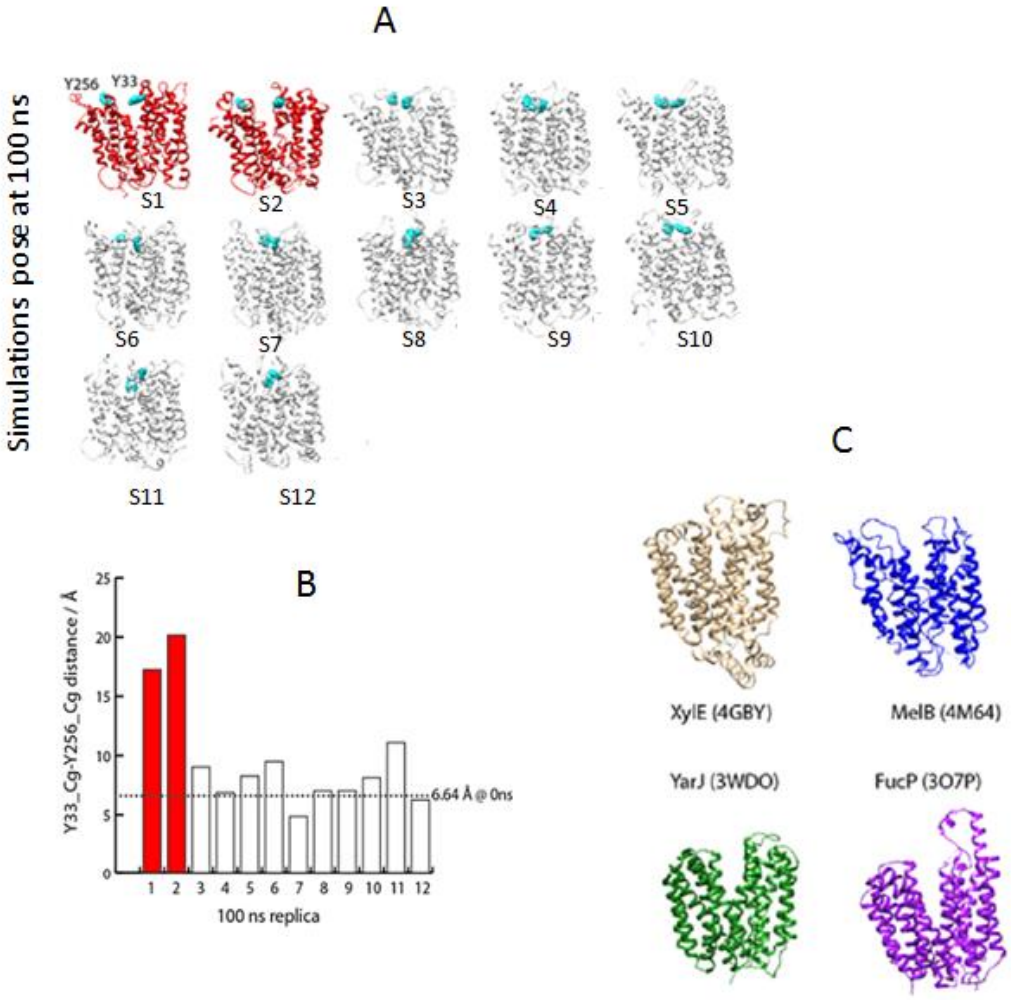


Fig 6.1: A. comparison of the 100 ns structures of all 12 replicas. **B.** The distance between the carbon gamma of Tyr33 and Tyr256 is plotted as a measure of extracellular opening, where the dash line indicates the distance at 0 ns. **C**. The crystal structures of MelB, XylE, YarJ and FucP are plotted for comparison

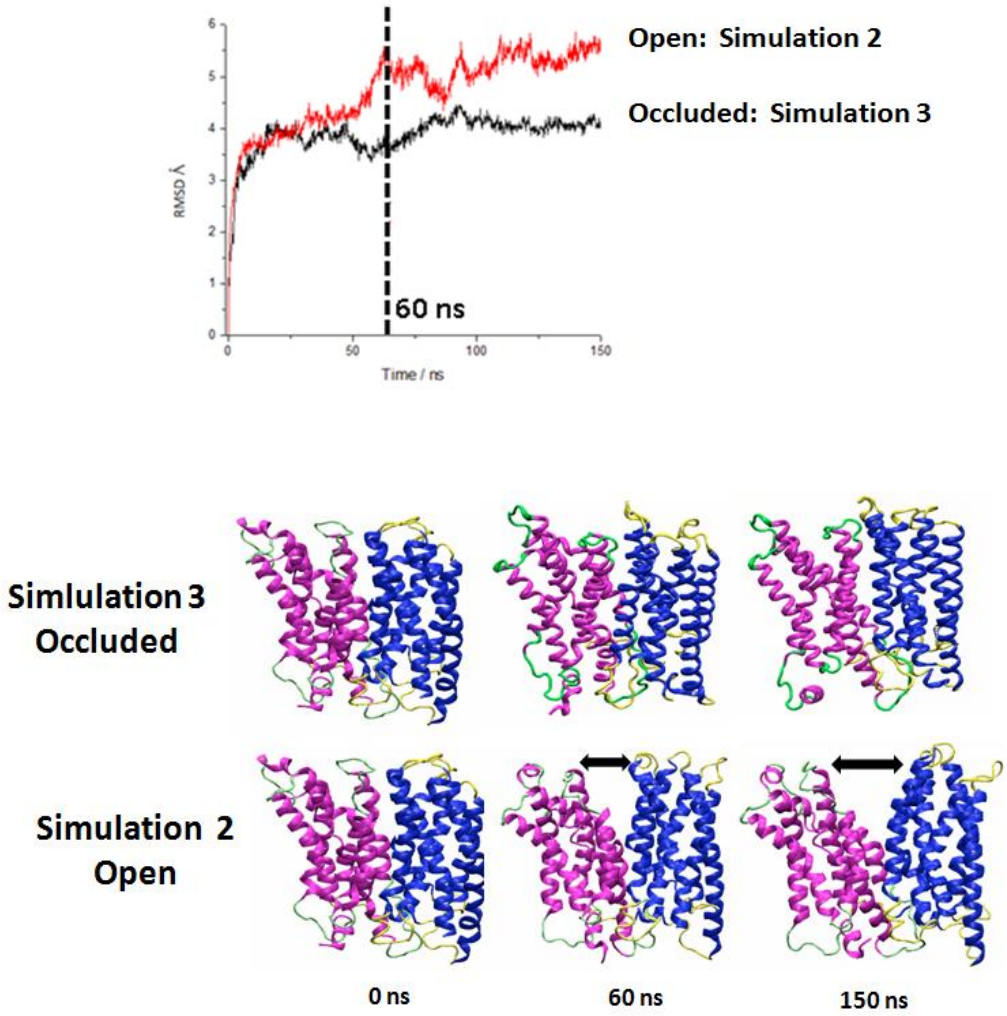


Fig. 6.2: MD Stability Analysis in 12-TM Wild type E.*coli* MelB: The N and C-terminal domains are colored in blue and magenta, respectively. **Top**: Backbone Cα RMSD of Simulation 2 is compared Simulation 3. **Bottom:** Cartoon representing different stages in simulations. Simulation 3 shows the consistent partially occluded conformation of the protein throughout the dynamics whereas simulation 2 shows the opening of the MelB transporter from 60ns.

6.2.2 Fluctuation Analysis: Root Mean Square Fluctuations (RMSF)

The root mean square fluctuation was calculated for each residue using VMD for the above simulations (occluded and open form) to characterize the protein dynamics. Here, RMSF describes fluctuations of each Cα atom of the protein averaged over the simulation time. As shown in the Fig. 6.3 the general fluctuations of protein in simulation 2 and simulation 3 are similar except in the region of Helix V and Helix 7-8. This means that these residues have undergone major conformational changes during the transition from partially occluded to open conformation.

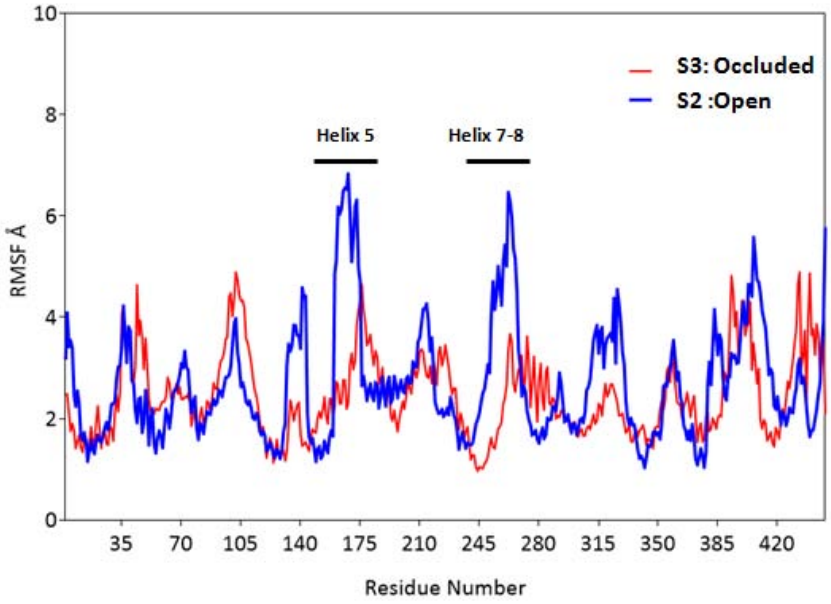


Fig. 6.3: Fluctuation analysis in WT MelB for two simulations using the backbone Cα. The simulation 3- occluded (red plot) shows the very stable fluctuation throughout the dynamics around 4Å while the simulation 2- open shows fluctuation around the marked regions indicates that these residues plays a role in opening of the structure.

6.2.3 Analysis of Interatomic distances:

The MD trajectory allows following the position of every single atom over time, and also certain interactions over time. Here, we analyze the role of target residues studied in this work in occluded and open conformation of MelB transporter.

Occluded conformation:

The study of interaction taking place along the molecular dynamics simulations can help us relate the molecular dynamics study with the experimental work by permitting to find out what mutations residues will result in breaking of the bond which in turn might be responsible for the low/no substrate binding affinity. As shown in the Fig 6.4 we have found some interactions involving the mutated residues studied in this thesis.

These results are in agreement with the experimental results where the mutation of some of these residues resulted in low substrate binding, which implies that these interactions are essential to maintain the correct native fold of the protein.

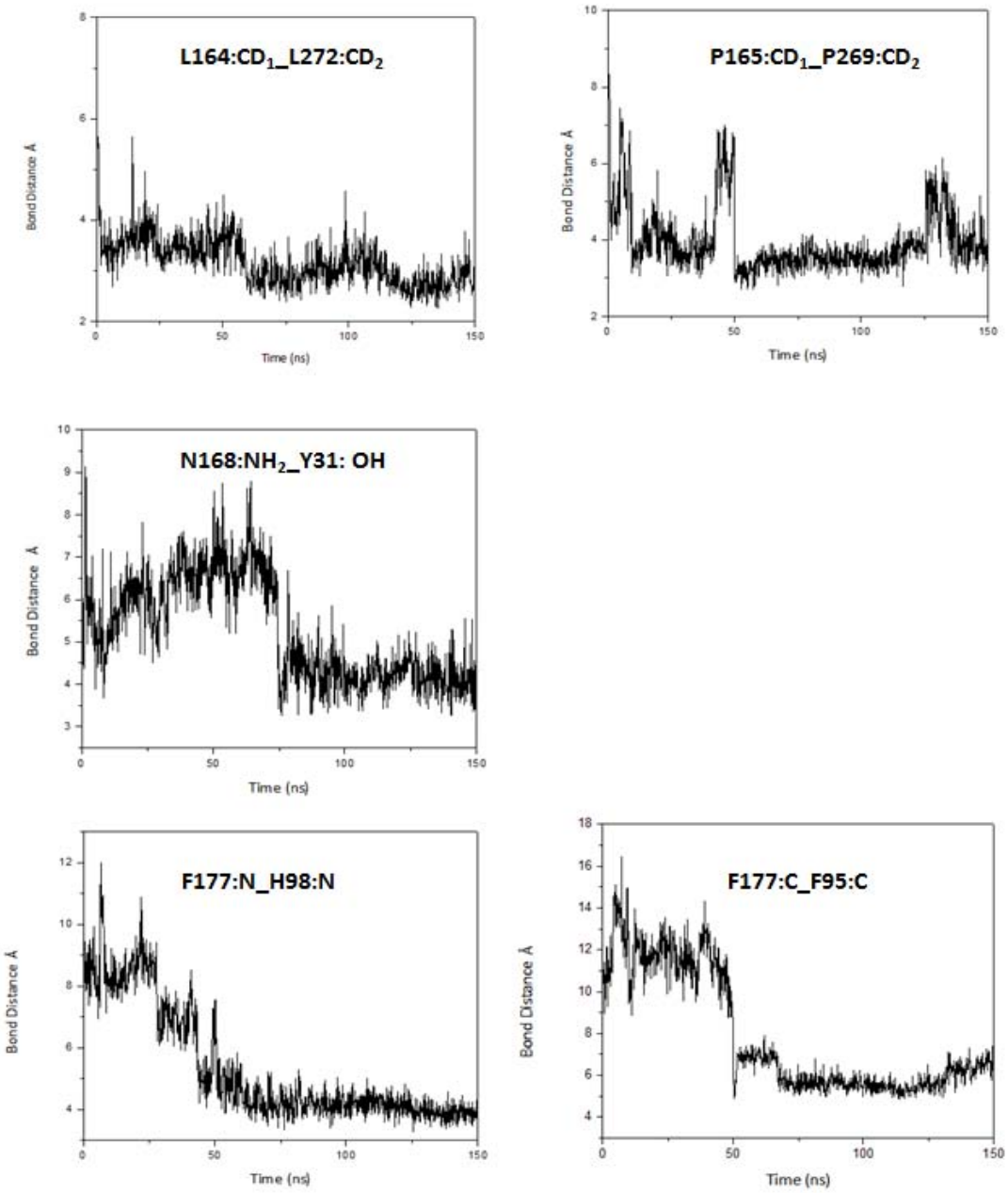


Fig. 6.4: Illustrative representation of the interatomic distances as a function of time (0 to 150 ns) for residues L164, P165, N168 and F177 in Occluded conformation. The distances were calculated using the VMD program and plotted using Origin 9.0.

Open Conformation:

Further dynamics studies, yielded a stable fully open MelB state. Similarly, interatomic distance analysis was performed for this form (Fig 6.5). The *S. typhimurium* outward partially occluded state of MelB contains a set of ionic locks and hydrophobic patches as described by Ethayathulla *et al.,* 2014.

As shown in Fig 6.5 the opening up of the structure implies the unlocking of hydrophobic interactions, mainly L164-L272. This strengthens the previous notion from experimental results that Helix V is important for the conformational changes necessary for the transport to take place [Abdel-Dayem *et al.,* 2003, Meyer-Lipp *et al.,* 2006 and Lin *et al.,* 2013].

Other interactions like P165:P269, N168:Y31, F177:H98, F177:F95 (Fig 6.5) shows larger deviation when compared with the occluded state. These results indicate that these residues might be important for closed conformation of the protein.

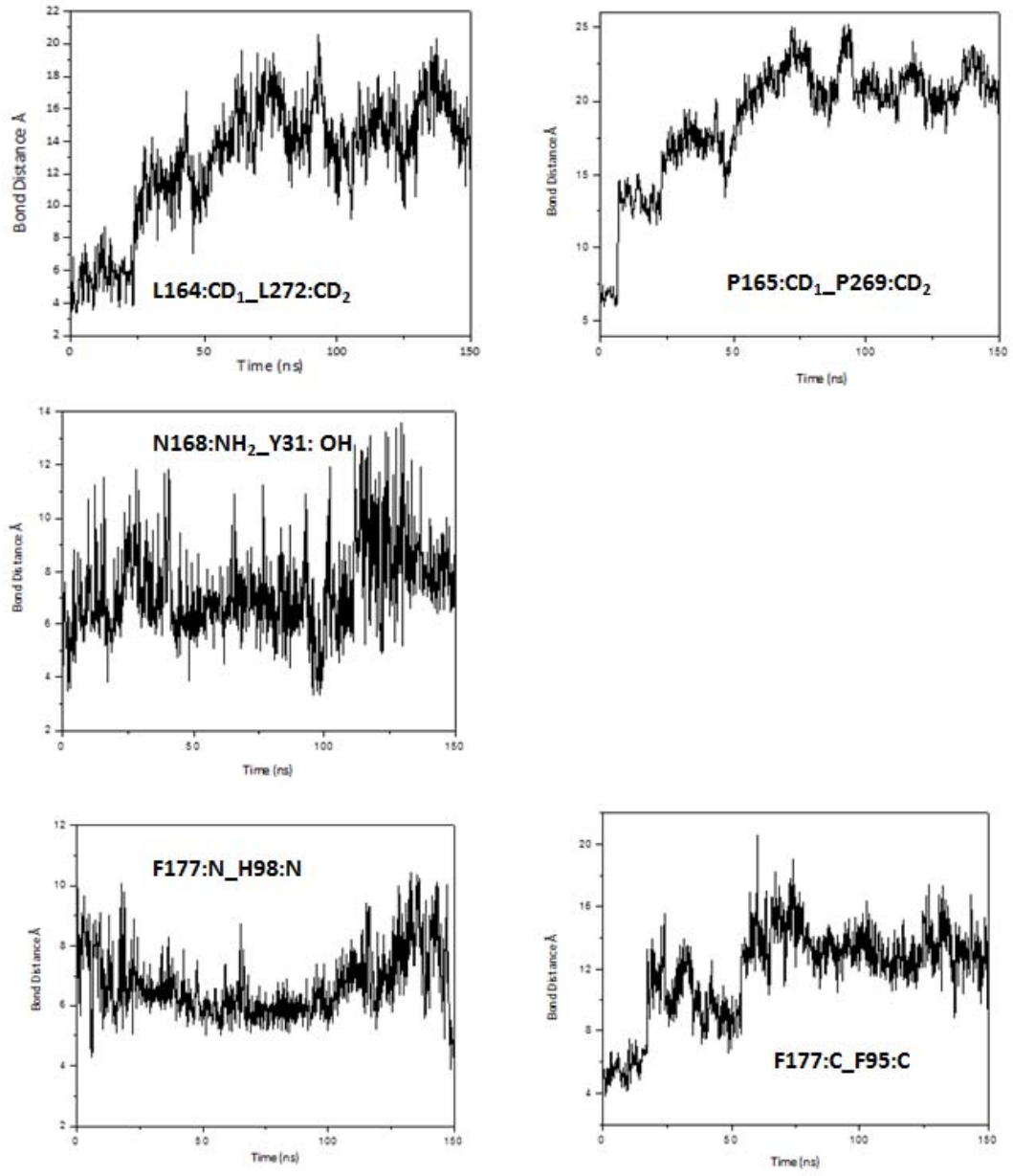


Fig. 6.5: Illustrative representation of the interatomic distances as a function of time (0 to 150 ns) for residues L164, P165, N168 and F177 in open conformation. The distances were calculated using the VMD program and plotted using Origin 9.0.

6.2.4 Role of Salt Bridges in Occluded and Open conformation

The salt bridge is the non-covalent interaction between two ionized sites. Of all the non-covalent interactions salt bridges are the strongest. And it is considered to exist only when the distance between the charged residues are less than 4 Å.

In this work, we checked the how salt bridges behave in the molecular dynamics which have resulted in occluded and open conformation (Simulation 3 and Simulation 2). As shown in the Fig 6.6, some of the important salt bridges in MelB WT transporter show a stable bond in Simulation 3 which remained in an occluded conformation, whereas in the case of Simulation 2 (open state) it resulted in a no bridge formation with a distance of > 4Å. From the above results, we can conclude that these salt bridges play an important role in maintaining the outward open conformation of the transporter.

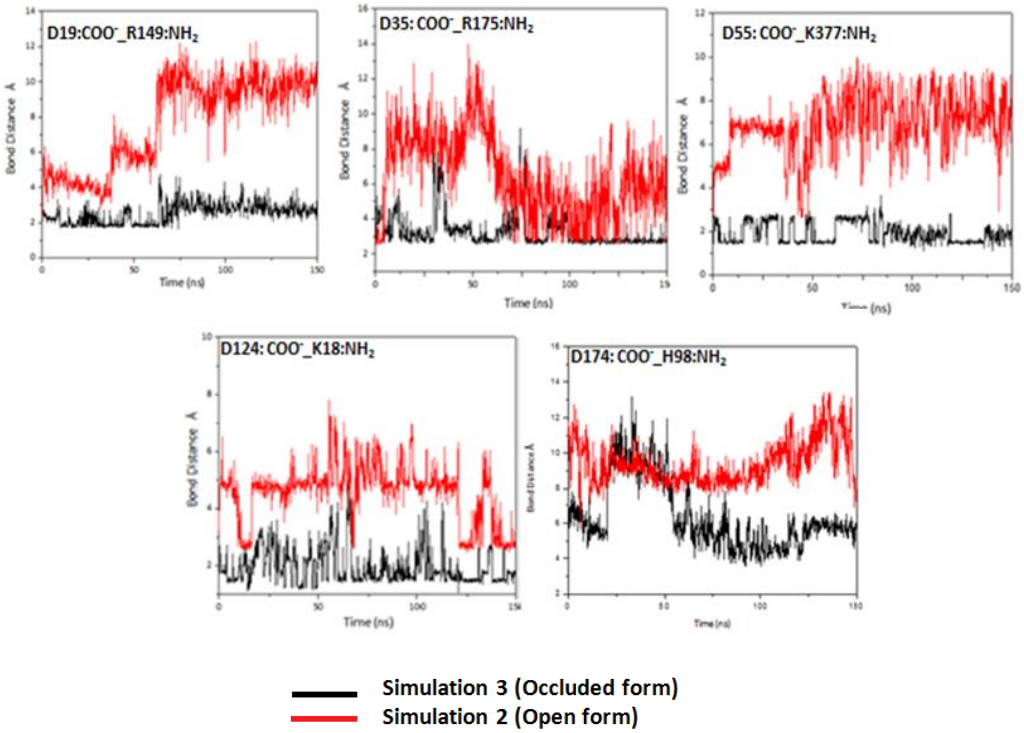


Fig 6.6: Representation of salt bridge interatomic distances as a function of time (0 to 150 ns) for residues D19, D35, D55, E101, D124 and D174. The distances were calculated using the VMD program and plotted using Origin 9.0

6.3. Comparison of N168C and Wild type MD simulations

According to Poolman *et al.,* 1996, the side chain of Asn58 could form one of the coordinates in the binding of Na^+, an interaction which is not essential for Li^+ or H^+. The role of Asn_{58} – Asp_{59} in MelB is the reminiscent of X-Glu_{61}-Glu_{62}-Y motif in subunit c of the E.coli F_0F_1-ATPase, which enables the enzyme to bind Li^+ (Zhang and Fillingame, 1995). And, Poolman et al., speculated that a similar cation binding pocket may be present in members of the GPH family and that Asn-58 and Asp-59, in addition to Asp-19, 55,124 are crucial for the co-ordination of Na^+.

These results encouraged us to know more about the role of Asn 168 in the loop 5-6. Our experimental results, gave us some clear evidence that mutation of Asn-168 to Cys resulted in a mutant which has lower efficiency in binding the substrates. To further strengthen that concept, we performed a molecular dynamics studies in the mutated form of Asn-168 to Cys.

As shown in Fig 6.7 it is very evident that in the case of wild type dynamics there is a stable ionic lock formation between the triad (Asp19, Asp55, Asp124) and the triad (Arg149, Lys377, Lys18), respectively, whereas in the case of N168C simulation all the Asp residues form a very weak ionic lock, with distances higher than about 6 Å. This result implies that mutation of this residue results in broken salt bridges which are most likely needed for the optimal substrate binding.

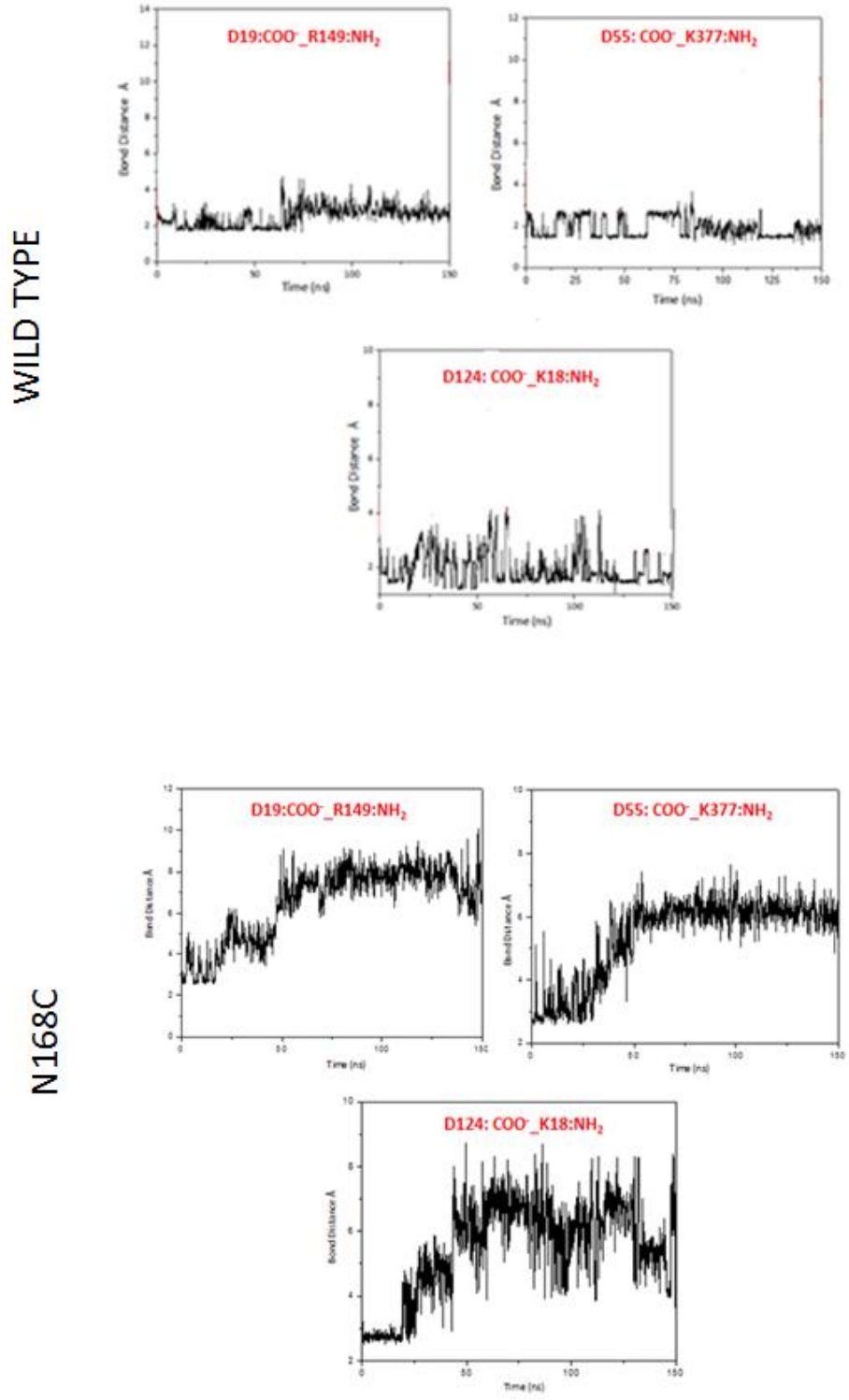


Fig. 6.7: Salt bridge interatomic distances comparison in Wild type and N168C MD simulations.

7. GENERAL DISCUSSION

Integral membrane proteins make up about one-third of all proteins in humans [Wallin *et al.,* 1998]. Nearly 25% of genes encode for a membrane protein and these include protein targets for over about 50 % of all drugs in use today. A prediction from sequence analysis has shown that the majority of membrane proteins are hydrophobic, mainly composed of trans-membrane α-helices. With respect to membrane transporters it is important to understand how transmembrane helices participate in substrate transport and which residues are implicated in substrate binding. Till now research in membrane proteins has mainly focused mainly on these elements and the connecting regions between them have received less attention. Reporting the structure of the Peptide-Proton symporter $PePT_{SO}$ Newstead *et al.,* 2011 emphasized the relevance of hairpin loops 4-5 and 10-11 as potential regulatory elements for the transport mechanism (Fig. 7.1).

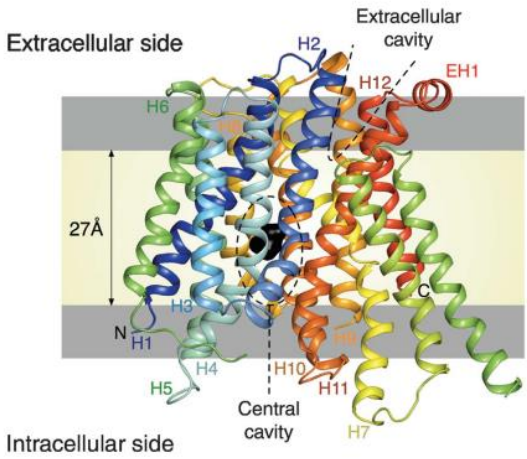


Fig. 7.1. Ribbon representation of 3D-Peptide Proton Symporter $PepT_{So}$ [Newstead *et al.,* 2011]

The transport mechanism of melibiose permease includes a series of conformational changes that let the protein to open towards the periplasmic side, allowing the binding of the cation and then the sugar, followed by reorientation of the protein opening towards the cytoplasmic side, where the sugar is released followed by the sodium.

Studies by [Abdel-Dayem *et al.,* 2003] and [Yousef and Guan 2009] have shown the importance of MelB cytoplasm loop 4-5 and Helix V for substrate binding. The 3D structure of E. coli MelB shows that the MelB has hydrophobic periplasmic loops and longer charged cytoplasmic loops. These characteristics can explain the uniform orientation and stability of MelB in proteoliposomes. A similar pattern has been described for other transporters such as LacS [Knol *et al.*, 1996] and bacteriorhodopsin [Rigaud *et al.*, 1995]. The cytoplasmic loops are always the prime determinants for the orientation of the protein when anchoring into the membrane in multi spanning membrane proteins [Gafvelin and Heijne 1994].

Previous studies, have shown that the in Lac permease, a minimum length of the cytoplasmic loop is necessary for activity, stability and insertion into the membrane [Weinglass *et al.,* 2000], and that loop 6-7 of the calcium pump was found to be important for the transport process [Menguy *et al.,* 1998].

In MelB, previous studies were centered importance of cytoplasmic loop residues for the structure function relationship but less information exists on the periplasmic loops. In the work done by Gwizdek *et al.,* 1997, it is shown that MelB membrane assembly is Sec independent which is in agreement with the hypothesis that protein with short periplasmic loops does not require the translocase machinery [Whitely *et al.,* 1994]. In brief, protein secretion in E. *coli* strictly requires the integral membrane proteins, SecY and SecE, and a membrane associated protein SecA.

So, when the crystal structure was still not available we started this work using the threading model of MelB in order to investigate the role of glycine rich periplasmic loop 5-6 (^{169}YVGGGDRGFG^{178}) along with few residues in Helix V (^{164}LPFVN^{168}) using cysteine scanning mutagenesis because Helix V was previously proved to be important for sugar binding [Ding and Wilson*,* 2001a].

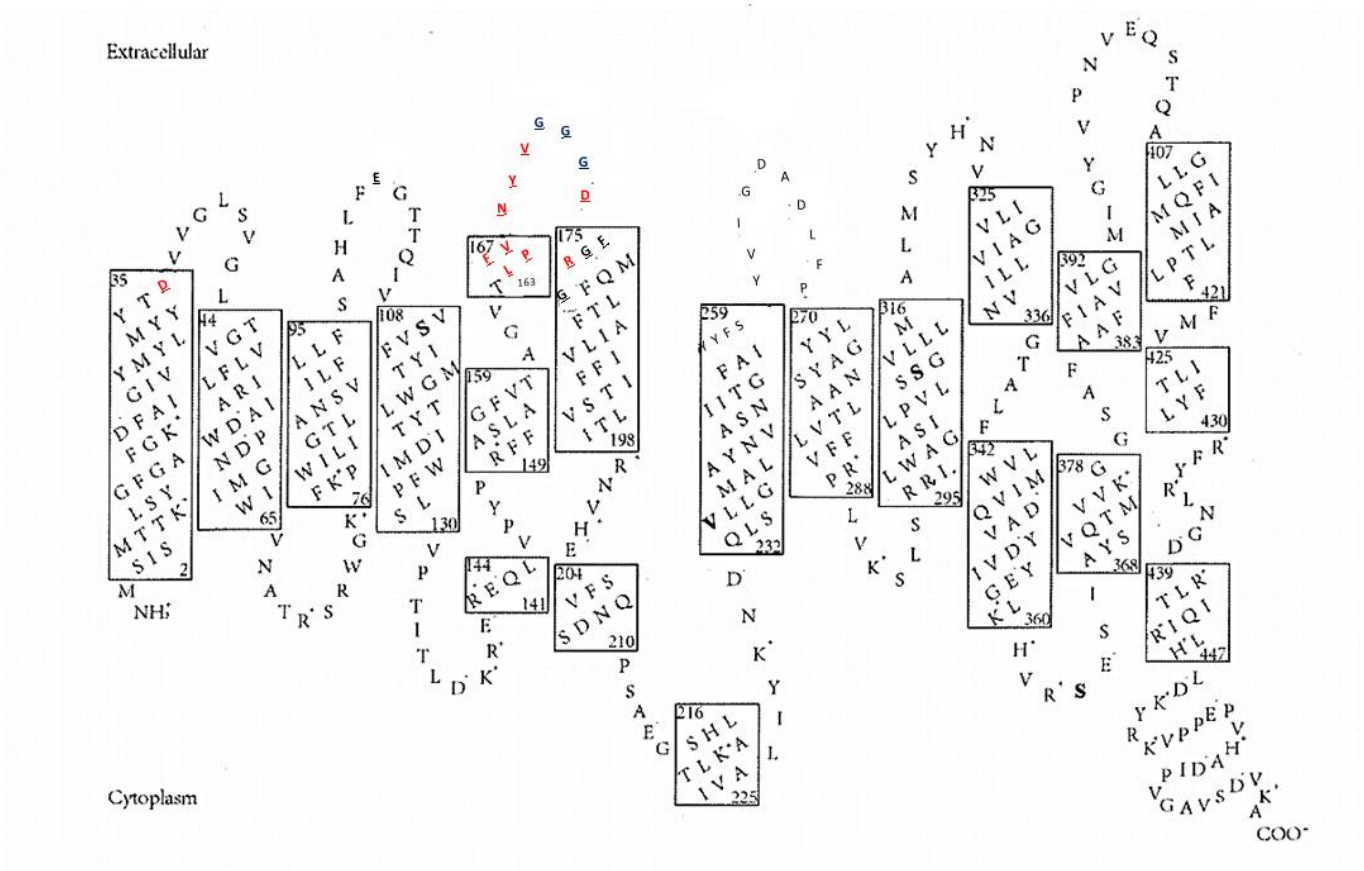


Fig. 7.2: 2D structure of *E. coli* MelB after the solved crystal structure of MelB *Salmonella typhimurium* [Ethayathulla *et al.,* 2014]. The residues highlighted in red are the target residues in this study. Three glycine residues are indicated in blue.

Later, the solved crystal structure of MelB *Salmonella typhimurium* (PDB code: 4M64, more than 85% identity to E. *coli* MelB) with partially occluded conformation open towards the cytoplasm has showed that the target residues in this work were not only present in the Helix V and loop 5-6, but are also distributed on Helix VI [Ethayathulla *et al.,* 2014]. Fig. 7.2 represents the secondary structure of the *E. coli* MelB after the solved crystal structure of MelB *Salmonella typhimurium.*

Some other insights from the crystal structure related to this study are: the presence of an Ionic lock between Helix VI (R175) and Helix I (D35) as shown in Fig. 7.3. This was later studied by Amin *et al.*, 2014 to know the role of these ionic locks in the transport activity of MelB from *Salmonella typhimurium.*

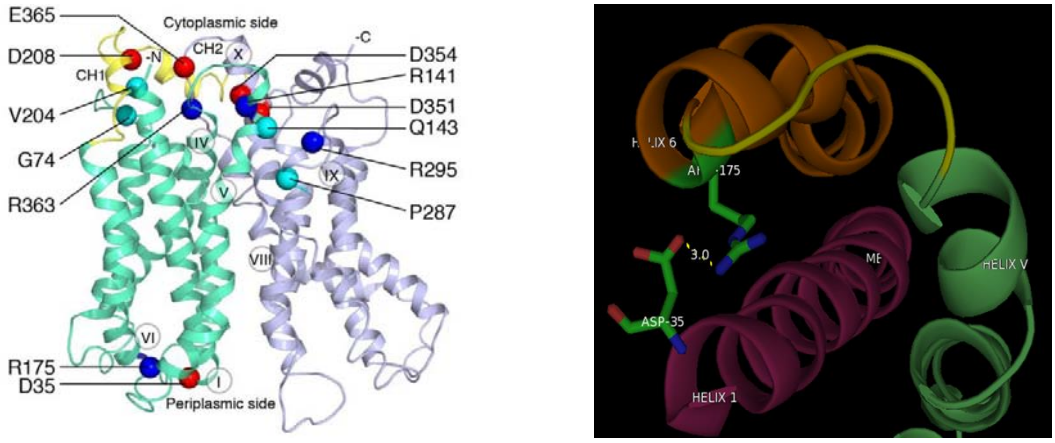


Fig. 7.3: **(Left)** Overall fold in $MelB_{ST}$ representing the positions involved in interactions with Arg residues [Amin *et al.,* 2014] **(Right)** shows the salt bridge distance between R175 and D35 in *E. coli* MelB after mutating the respective residues using Swiss-Model. And, the plot was made using VMD 1.9.2.

In the later part of the thesis point mutation was also done for D35 of Helix I and the salt bridge residue D35-R175 to know their role in E. *coli* MelB. As shown in Fig. 7.4, the PCR generated mutants of Asp-35 and Arg-175/Asp-35 showed white colonies on MacConkey agar plates which indicates that there is absence of MelB metabolism in this mutant.

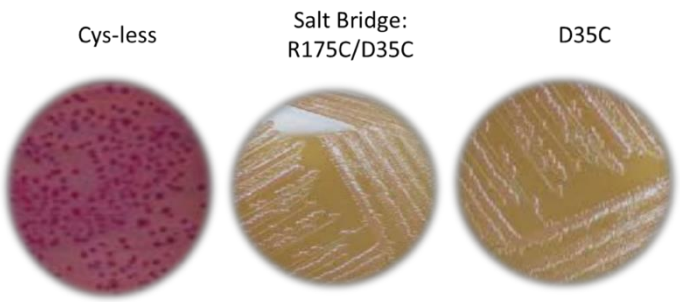


Fig. 7.4: Shows Cys-less, R175C/D35C and D35C mutant on Mac-Conkey agar plate supplemented with 10mM Melibiose grown at 30˚C. R175C and R175C/D35C displays white colonies indicating the absence of MelB metabolism while Cys-less shows bright red colonies indicating the presence of MelB metabolism.

7.1. Characterization of mutants using experimental approaches (FRET, FTIR and Transport assays)

Over expression and purification of E. *coli* Melibiose permease is a relatively straightforward procedure rendering a high yield. The well-optimized purification protocols permit the purification of considerable amounts of functional protein compared to eukaryotic membrane proteins where post-translational modifications complicate the process. In this work, solubilization of MelB from the membrane was done using DDM in the presence of a high concentration of sodium to ensure the stability of MelB.

The proteoliposomes containing the purified MelB transport protein were used to study the substrate binding characteristics in both FRET and ATR-FTIR spectroscopy. The FRET analysis permitted to approach substrate binding by measuring the binding of a sugar analog (D^2G) in the absence and presence of sodium. ATR-FTIR was used to study the three stages of alternate access mechanism as indicated in Fig. 7.5.

The substrate-induced IR_{diff} spectra of each mutant was compared to Cys-less based on the intensity (m) and similarity (R^2) correlation parameters reflecting the capability of the substrates binding to each mutant. One of the major drawbacks of this technique is that it's not time efficient and this fact prevented from measuring all the mutants that

had been studied by FRET. Thus, in this work ATR_FTIR approach was applied to some selected mutants (P165C, N168C, D174C, R175C, D35C, R715C/D35Cand F177C) as characterized by FRET analysis.

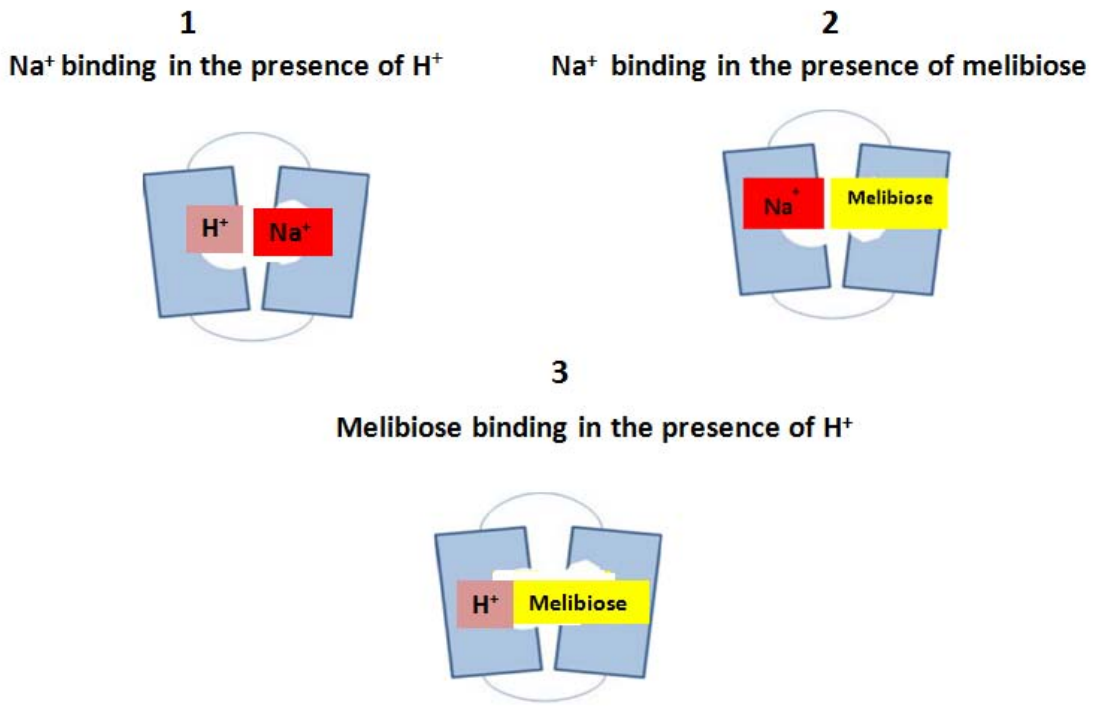


Fig. 7.5: Illustration of the three different stages of substrate binding studied using ATR-FTIR. **(1)** Na^+ binding in the presence of H^+. **(2)** Melibiose binding in the presence of Na^+ **(3**) Melibiose binding in the presence of H^+.

In order to facilitate the interpretation of the point mutations, we have developed a linear correlation method using a strategy where the mutant's infrared spectral differences in intensity and shape are plotted on a SI map (Shape-Intensity map) on which each point can be assigned a similarity index. The similarity index can be translated into a color scale which facilitates the comparison of a group of mutants with the control (Cys-less) and allows a quick evaluation of the effect of any of the point mutation on the conformational effects induced by the binding of the substrates.

Secondly, in order to evaluate any possible negative effect due to the presence of detergent, the mutants were characterized using a polymorphous Inside Out Vesicles (ISO) which comprised of a bulk of membrane proteins. This was used as a comparative study to strengthen the results from proteoliposomes. The assessment of the vesicles

always has limitations. Western blotting with the specific anti-His reagent was used as a quantification method to check the protein expression yields.

Previous studies have reported that the IR difference spectrum for native vesicles (ISO) failed to demonstrate a difference signal. The main reason must be related to the extensive quantity of membrane proteins in the vesicles which increase the background noise. Thus, ISO vesicles were was just used in FRET with D^2G. This sugar contains a galactoside-group alike melibiose, hence it binds exclusively to the sugar binding moiety of MelB and LacY. Since, our bacterial strain is devoid of the *LacY* gene, the data collected from the FRET of the vesicles can be entirely assigned to the melibiose transporter.

For all the mutants described in this thesis, the FRET (Trp -> D^2G) of MelB containing proteoliposomes (PL) and ISO vesicles was measured. In general, most of the mutants showed an intermediate effect on substrate binding. While other mutants like mutants like D174C in ISO vesicles, F177C (ISO and PL), D35C (ISO and PL), and R175C/D35C (ISO and PL), showed very high effect or almost unable to bind the substrates.

FTIR results of mutants based on SI maps shows that the regarding the difference in shape (R^2) (vertical axis of SI maps), all mutants seems to have an intermediate (some a low) effect on cation and sugar binding except for the mutants D35C and R175C/D35C that has high effect. However if we consider the horizontal axis of the SI maps where the effect on the difference spectra intensity is represented, N168C, D174C and F177C along with D35C and R175C/D35C display a highly affected Na^+ or Na^+.Melibiose binding. The high effect on the intensity of the difference spectrum can be interpreted as mutation of these residues to cysteine somewhat affects the MelB structure which in turn affects the binding of the substrates

The biophysical work was further validated using transport assay measurements showing the reduced [^{3}H] melibiose transport for the mutants R175C, N168C, F177C and no transport for D35C when compared to Cys-less.

7.2 Comparison of experimental results with computational data

MD simulations have become a common tool to investigate structure–activity relationships in biological macromolecules. These simulations help us to interpret the experimental data by providing the element of dynamics at an atomic level and give access to information not directly accessible by experiments [Kandt *et al.,* 2008]. GPCRs have been the subject of some of the longest all-atom simulations reported so far. These simulations have made significant contributions to evaluating the importance of internal waters to GPCR functions [Grossfield *et al.,* 2008; Hurst *et al.,* 2010] which was later confirmed by solid-state NMR [Grossfield *et al.,* 2008].

In this work, the role of some residues in Helix V, Loop5-6 and Helix VI of E. *Coli* MelB transporter was studied using MD simulations. MD simulations of E. *Coli* MelB embedded in a POPE bilayer yielded 10 partially/totally occluded conformations and two replicas showing the largest distance between Tyr-33 and Tyr-256 indicating the substrate free outward open state.

The role of target residues in Helix V, Loop5-6, and Helix VI was studied in open and partially occluded conformation. Some of the residues like Leu164 (Helix V), Pro165 (Helix V), Asn168 (Loop5-6) and Phe177 (Helix VI) showed interactions with other helices in the transporter. These data helped us to explain the possible role of these residues in atomic detail:

Leu164: From the experimental approach it has been shown that mutation of Leu-164 to Cys shows very similar binding affinity as that of the positive control Cys-less. But, the computational approach has shown a tighter hydrophobic lock less than 4 Å between Leu164 TM5 and Leu272 TM8 in the partially/fully occluded state, whereas no interaction (14 Å) was observed in the open state. This is an indication that Leu-164 could be part of a hydrophobic lock the periplasmic side necessary to hold N and C terminal domains together.

Pro165: As mentioned before proline residues are important for the correct folding of membrane proteins. The experimental results showed an intermediate effect on binding of substrates which may be due to the broken nteraction between Pro165 (TM5) and

Pro269 (TM8) when it was mutated to Cys. We can conclude that this hydrophobic patch between Pro165-Pro269 is essential maintain the proper fold of the transporter for substrate binding.

Asn168: The only polar residue in the loop5-6 of the melibiose transporter shows very high effect on Na^+ binding and intermediate effect on melibiose binding when mutated to Cys, which was further confirmed by the simulation analysis where it showed an unstable ionic lock between the triad (Asp-19, Asp-55 and Asp-124) and the triad (Arg-149, Lys-377, Lys-18) respectively. The polar interaction between Tyr31 (TM1) Asn168 (Loop5-6) identified in molecular dynamics study might be important to keep the structure in the occluded conformation.

Phe177: The aromatic stacking interaction between Phe177 (TM6) and Phe95 (TM3) may be important to maintain the occluded conformation of the transporter or play a vital role in the formation of outward open state. This might be the reason for weak binding or reduced transport of radioactive sugar.

CONCLUSIONS

1. The FRET and FTIR results show that the mutants D35C and R175C/D35C cannot bind the substrates, whereas the mutation R175C has an intermediate effect. Therefore, Asp-35 is crucial for substrate binding and hence for the transport mechanism in MelB.

2. L164C and G172C reconstituted into proteoliposomes show reduced affinity for the fluorescent sugar D^2G. However, the same mutants show similar binding affinity to Cys-less in the case of native membrane vesicles. Therefore, the purification and reconstitution processes lead to some alterations of the MelB structure for these particular mutants.

3. Mutation of the polar residue, Asn-168 located at the beginning of the loop 5-6 showed reduced binding affinity and transport for Na^+ suggesting that this is a important residue for Na^+ binding. Further, N168C MD simulations showed an unstable ionic lock between the triad Asp-19, Asp-55 and Asp-124 which further validates that this residue may be important for Na^+ recognition.

4. Mutation of the aromatic residue, Phe-177 located at the beginning of Helix VI showed reduced substrate binding and transport of substrates when compared to Cys-less, indicating that this residue is essential for the correct fold of the native structure.

5. Molecular dynamics simulations of $MelB_{EC}$ show conformational transition from occluded state to an outward open state. Combining these results with the experimental approach, it is clear that the interaction between Helix V/ Helix VIII (L164:L272; P165:P269), loop 5-6/ Helix I (N168:Y31) and Helix VI/Helix III (F177:F95) are important for maintaining the structure in the partially occluded conformation.

A

Abdel-Dayem M, Basquin C, Pourcher T, Cordat E, Leblanc G. Cytoplasmic loop connecting helices IV and V of the melibiose permease from Escherichia coli is involved in the process of Na+-coupled sugar translocation. J Biol Chem.;278(3):1518–24. **(2003).**

Abramson J, Smirnova I, Kasho V, Verner G, Kaback HR, Iwata S. Structure and mechanism of the lactose permease of Escherichia coli. Science.;301(5633):610–5. **(2003).**

Alberts B, Johnson A, Lewis J, et al. Molecular Biology of the Cell. 4th edition. New York: Garland Science;. Membrane Proteins.**(2002)**

Amin.A , Abdul S. Ethayathulla LG. Suppression of Conformation-Compromised Mutants of Salmonella enterica Serovar Typhimurium MelB. J Bacteriol 196(17):3134–9. **(2014).**

Arrondo JL, Goni FM.Strucutre and dynamics of membrane proteins as studied by infrared spectroscopy. Prog Biophys Mol Biol 72(4):367-405**.(1999).**

Arinaminpathy, Y., Khurana, E., Engelman, D. M. & Gerstein, M. B. Computational analysis of membrane proteins: the largest class of drug targets. *Drug Discov. Today* **14,** 1130–1135 **(2009).**

B

Basquin C. Etude de la contribution relative des domaines intra et extra-membranaires au transport Na^+/sucre par la Melibiose Permease d' Escherichia coli. DEA. Universite de Nice Sophia-Antipolis **(2001).**

Botfield MC and **Wilson** TH. Mutations that simultaneously alter both sugar and cation specificity in the melibiose carrier of Escherichia coli. J Biol Chem. 263(26):12909–15. **(1988).**

Botfield MC and Wilson TH. Carboxyl-terminal truncations of the melibiose carrier of Escherichia coli. J Biol Chem.;264(20):11643–8. **(1989).**

Botfield MC, Naguchi K, Tsuchiya T, and Wilson TH. Membrane topology of the melibiose carrier of Escherichia coli. J. Biol. Chem. 267, 1818-22 **(1992).**

Borhani DW, Shaw DE. The future of molecular dynamics simulations in drug discovery. J Comput Aided Mol Des.;26(1):15–26. **(2012).**

C

Cordat E, Mus-veteau I, Maehrel C. Structural Studies of the Melibiose Permease of Escherichia coli by Flourescence Energy Transfer. II. Identification of the trytophan residues acting as energy donors.J. Biol.Chem.;273(50):33198–202. **(1998).**
Cordat E, Leblanc G, Mus-Veteau I. Evidence for a role of helix IV in connecting cation- and sugar-binding sites of Escherichia coli melibiose permease. Biochemistry.;39(15):4493–9. **(2000).**

D

Dave N, Troullier A, Mus-Veteau I, Duñach M, Leblanc G, Padrós E. Secondary structure components and properties of the melibiose permease from Escherichia coli: a fourier transform infrared spectroscopy analysis. Biophys J.79(2):747–55. (**2000).**
Ding, P.Z., Botfield M.C and Wilson T.H. Sugar recognitio mutants of the melibiose carrier of Escherichia coli: possible structural information concerning the arrangement of membrane-bound helices and sugar/cation recognition site." Biochim Biophys Acta.1509(1-2):123-130 **(2000)**
Ding PZ and **Wilson TH**. Physiological evidence for an interaction between helices XI and helices I, II, and V in the melibiose carrier of Escherichia coli.. Biochem Biophys Acta.;268(2):409–13. **(2000a).**
Ding PZ , Wilson T.H . The melibiose carrier of Escherichia coli: cysteine substitutions for individual residues in helix XI. J Membr Biol.**(2000b).**
Ding PZ, Weissborn AC, and Wilson TH. Cysteine substitutions for individual residues in helix IV of the melibiose carrier of Escherichia coli. J. Membr. Biol.183, 33-8 **(2001).**
Ding PZ and Wilson TH. Cysteine mutagenesis of the amino acid residues of transmembrane helix I in the melibiose carrier of Escherichia coli. Biochemistry. 40(18):5506–10.(**2001a**).
Ding PZ and **Wilson TH**. The effect of modifications of the charged residues in the transmembrane helices on the transport activity of the melibiose carrier of Escherichia coli. Biochem Biophys Res Commun.;285(2):348–54. (**2001b).**
Ding PZ and Wilson TH. The proximity between helix I and helix XI in the melibiose carrier of Escherichia coli as determined by crosslinking. *Biochim. Biophys. Acta* 1514, 230-8 **(2001c).**
Ding PZ. An investigation of cysteine mutants on the cytoplasmic loop X/XI in the melibiose transporter of Escherichia coli by using thiol reagents: Implication of structural conservation of charged residues. Biochem Biophys Res Commun.;307(4):864–9. **(2003).**

Ding PZ. Loop X/XI, the largest cytoplasmic loop in the membrane-bound melibiose carrier of Escherichia coli, is a functional re-entrant loop. Biochim Biophys Acta - Biomembr.;1660(1):106–17.**(2004).**

E

Edelheit, O., Hanukoglu, A. & Hanukoglu, I. Simple and efficient site-directed mutagenesis using two single-primer reactions in parallel to generate mutants for protein structure-function studies. *BMC Biotechnol.* **9,** 61 **(2009).**

Ethayathulla AS, Yousef MS, Amin A, Leblanc G, Kaback HR, Guan L. Structure-based mechanism for Na(+)/melibiose symport by MelB. Nat Commun [Internet]. 5:3009. (**2014).**

F

Faham.S, Akira.W , Gabriel M.B, Duilio C.A, Specht, Bruce A. Hirayama , Ernest M. Wright. The crystal structure of a sodium galactose transporter reveals mechanistic insights into Na+/sugar symport. Science (80-)**;**(6):997–1003.(**2008)**

Franco PJ and **Wilson TH**. Arg-52 in the melibiose carrier of Escherichia coli is important for cation-coupled sugar transport and participates in an intrahelical salt bridge. J Bacteriol.;181(20):6377–86. **(1999).**

Franco PJ, Jena AB and Wilson.Physiological evidence for an interaction between helices II and XI in the melibiose carrier of Escherichia coli. Biochim. Biophys. Acta 1510, 231-42 **(2001).**

Fringeli, U.P., GuÈnthard, Hs.H. Infrared Membrane Spectroscopy. In: Grell, E. (Ed.), Membr. Spectrosc. Springer, Berlin, pp. 270±332.**(1981).**

Fuerst O, Lin Y, Granell M, Leblanc G, Padrós E, Víctor A. Melibiose Transporter of Escherichia coli : Critical Contribution of Lys377 to the Structural Organization of the Interacting Substrate Binding Sites. JBiol chema. (**2015).**

G

Gafvelin, G., and von Heijne, G. Topological “frustration” in multispanning E. *Coli* membrane proteins. Cell 77,401-12. **(1994).**

Ganea C, Meyer-Lipp K, Lemonnier R, Krah A, Leblanc G, Fendler K. G117C MelB, a mutant melibiose permease with a changed conformational equilibrium. Biochim Biophys Acta - Biomembrane.Elsevier B.V.;;1808(10):2508–16. (**2011).**

Goormaghtigh, E., V. Cabiaux and J. M. Ruysscahert. “Determination of soluble and membrane protein structure by Fourier transform infrared spectroscopy. I. Assignments and model compounds.” Subcell Biochem. 23:329-362 **(1994a).**

Granell M. Estudi del lloc d'unio del sodi de la permeasa de melibiosa d'Escherichia coli. Doctoral Thesis **(2009)**

Granell M, León X, Leblanc G, Padrós E, Lórenz-Fonfría V a. Structural insights into the activation mechanism of melibiose permease by sodium binding. Proc Natl Acad Sci U S A;107(51):22078–83. **(2010).**

Grossfield A., Pitman M.C., Feller S.E., Soubias O., Gawrisch K. Internal hydration increases during activation of the G-protein-coupled receptor rhodopsin. J. Mol. Biol; 381:478–486. doi: 10.1016/j.jmb.2008.05.036**. (2008)**

Guan, L., Nurva, S. & Ankeshwarapu, S. P. Mechanism of melibiose/cation symport of the melibiose permease of Salmonella typhimurium. *J. Biol. Chem.* **286,** 6367–74 **(2011)**.

Gwizdek C, Leblanc G, Bassilana M. Proteolytic mapping and substrate protection of the escherichia coli melibiose permease. Biochemistry.;36(28):8522–9. (**1997).**

H

Hama H and Wilson TH. Cation-coupling in Chimeric Melibiose Carriers Derived from Escherichia coli and Klebsiella pneumoniae. J Biol Chem. 268:10060–5. (**1993).**

Hama H, Wilson TH. Replacement of Alanine-58 by Asparagine Enables the Melibiose Carrier of Klebsiella-pneumoniae to Couple Sugar Transport to Na+. J Biol Chem. 269(2):1063–7. **(1994).**

Hacksell I, Rigaud J, Purhonen P, Pourcher T, Hebert H. Projection structure at 8 A melibiose permease , an Na ± sugar co-transporter from Escherichia coli.;21(14):3569–74. (**2002).**

Huang Y, Lemieux MJ, Song J, Auer M, Wang D-N. Structure and mechanism of the glycerol-3-phosphate transporter from Escherichia coli. Science.;301(5633):616–20. **(2003).**

Humphrey W.F., Dalke A, Schulten K. VMD- Visual molecular dynamics. J. Mol. Graphics, 14, p. 33. **(1996)**

Hurst DP, Grossfield A, Lynch DL, Feller S, Romo TD, Gawrisch K, Pitman MC, Reggio PH. A lipid pathway for ligand binding is necessary for a cannabinoid G protein-coupled receptor. J Biol Chem; 285:17954–64. (**2010).**

J

Jackson, M and Mantsch H.H. The use and misuse of FTIR spectroscopy in determination of protein structure. Crit Rev Biochem Mol Biol. 30(2) 95-120 **(1995).**

Jardetzky.O. PROTEIN DYNAMICS AND CONFORMATIONAL TRANSITIONS IN ALLOSTERIC PROTEINS. Proteins.;65(3):171–219. (**1996).**

Jakkula SV, Guan L. Reduced Na+ affinity increases turnover of Salmonella enterica serovar typhimurium MelB. J Bacteriol.;194(20):5538–44. **(2012).**

Jo, S., Kim, T., Iyer, V. G. and Im. CHARMM-GUI: A web-based graphical user interface for CHARMM. J. Comput. Chem., 29: 1859–1865. **(2008).**

K

Kale L, Skeel R & Bhandarkar M. Namd2: greater scalability for parallel molecular dynamics. J comput Phys 151: 283-312 **(1999).**

Kandt C, Matyus E, Tieleman DP. Protein Lipid Interactions from a Molecular Dynamics Simulation Point of View In: Nag K, editor. , editor. Structure & Dynamics of Membranous Interfaces: Hoboken, NJ; pp. 267–282. **(2008).**

Kawakami T, Akizawa Y, Ishikawa T, Shimamoto T, Tsuda and Tsuchiya T. Amino acid substitutions and alteration in cation specificity in the melibiose carrier of *Escherichia Coli. J. Biol. Chem.* 263, 14276-80 **(1988).**

Knol J., Veenhoff, L.M., Liang,W.J., Henderson,P.J.F., Leblanc,G. and Poolman,B. Unidirectional reconstitution into detergent-destabilized liposomes of the purified lactose transport system of *Streptococcus thermophilus*. J. Biol. Chem., 271, 15358–15366. **(1996)**

Krimm S, Bandekar J. Vibrational spectroscopy and conformation of peptides, polypeptides, and proteins. *Adv Protein Chem.* 38:181–364. **(1986).**

Krulwich T. Bacterial Energetics: A Treatise on Structure and Function. (**1990).**

L

Law CJ, Maloney PC, Wang D. Ins and Outs of Major Facilitator Superfamily Antiporters. (67):289–305. **(2008)**

Leen, W. G., J. Klepper, M. M. Verbeek, M. Leferink, T. Hofste, B. G. Van Engelen, R. A. Wevers, T. Arthur, N. Bahi-Buisson, D. Ballhausen, J. Bekhof, P. Van Bogaert, I. Carrilho, B. Chabrol, M. P. Champion, J. Coldwell, P. Clayton, E. Donner, A. Evangeliou, F. Ebinger, K. Farrell, R. J. Forsyth, C. G. E. L. De Goede, S. Gross, S. Grunewald, H. Holthausen, S. Jayawant, K. Lachlan, V. Laugel, K. Leppig, M. J. Lim, G. Mancini, A. D. Marina, L. Martorell, J. Mcmenamin, M. E. C. Meuwissen, H. Mundy, N. O. Nilsson, A. Panzer, B. T. Poll-The, C. Rauscher, C. M. R. Rouselle, I. Sandvig, T. Scheffner, E. Sheridan, N. Simpson, P. Sykora, R. Tomlinson, J. Trounce, D. Webb, B. Weschke, H. Scheffer, and M. A. Willemsen. "Glucose Transporter-1 Deficiency Syndrome: The Expanding Clinical and Genetic Spectrum of a Treatable Disorder."*Brain* 133.3: 655-70. **(2010).**

León X, Lórenz-Fonfría V a., Lemonnier R, Leblanc G, Padrós E. Substrate-induced conformational changes of melibiose permease from Escherichia coli studied by infrared difference spectroscopy. Biochemistry.;44(9):3506–14. (**2005).**

León X, Lemonnier R, Leblanc G, Padrós E. Changes in secondary structures and acidic side chains of melibiose permease upon cosubstrates binding. Biophys J.;91(12):4440–9. (**2006).**

León X, Leblanc G, Padrós E. Alteration of sugar-induced conformational changes of the melibiose permease by mutating Arg141 in loop 4-5. Biophys J. 96(12):4877–86. **(2009).**

Lin Y. Structural and Functional Studies of Melibiose permease of Escherichia coli. Thesis. (**2012).**

Lin Y, Fuerst O, Granell M, Leblanc G, Lórenz-Fonfría, V, Padrós, E. The substitution of Arg149 with Cys fixes the melibiose transporter in an inward-open conformation. Biochim Biophys Acta. 182(8):1690-9 **(2013).**

Lórenz-Fonfría, V. A., León, X. & Padrós, E. Studying substrate binding to reconstituted secondary transporters by attenuated total reflection infrared difference spectroscopy. *Methods Mol. Biol.* **914,** 107–26 **(2012).**

M

Maehrel.C, Emmanuelle Cordat‡ IM-V and GL. Structural Studies of the Melibiose Permease of Escherichia coli by Fluorescence Resonance Energy Transfer I. EVIDENCE FOR ION-INDUCED CONFORMATIONAL CHANGE. Biochemistry.;273(50):33198–202. **(1998).**

Matsuzaki S, Weissborn AC, Tamai E, Tsuchiya T, Wilson TH. Melibiose carrier of Escherichia coli: Use of cysteine mutagenesis to identify the amino acids on the hydrophilic face of transmembrane helix 2. Biochim Biophys Acta - Biomembr.;1420(1-2):63–72. (**1999).**

Menguy, T., F. Corre, L. Bouneau, S. Deschamps, J. V. Moller, P. Champeil, M. le Maire, and P. Falson. The cytoplasmic loop located between transmembrane segments 6 and 7 controls activation by Ca^{2+} of sarcoplasmic reticulum Ca^{2+}-ATPase. J. Biol. Chem. 273:20134–20143. **(1998).**

Meyer-Lipp K. Time-Resolved Measurements of Sugar-Binding-Induced Conformational Changes in the Melibiose Permease from Escherichia Coli. Thesis. 1-174 p.(**2005).**

Meyer-Lipp K, Séry N, Ganea C, Basquin C, Fendler K, Leblanc G. The inner interhelix loop 4-5 of the melibiose permease from Escherichia coli takes part in conformational changes after sugar binding. J Biol Chem. 281(36):25882–92. **(2006).**

Mitchell P. A General Theory of Membrane Transport From Studies of Bacteria. Nature. 180(4577):134–6. (**1957).**

Mizushima K, Awakihara S, Kuroda M, Ishikawa T, Tsuda M, Tsuchiya T. Cloning and sequencing of the melB gene encoding the melibiose permease of Salmonella typhimurium LT2. 234(1):74–80. **(1992)**.

Mus-Veteau & Leblanc, G. Melibiose permease of Escherichia coli: structural organization of cosubstrate binding sites as deduced from tryptophan fluorescence analyses. *Biochemistry* **35,** 12053–60 **(1996)**.

N

Newstead S, Drew D, Cameron AD, Postis VLG, Xia X, Fowler PW, Ingram JC, Carpenter EP, Sansom MSP, Mcpherson MJ, Baldwin SA, Iwata S. Crystal structure of a prokaryotic homologue of the mammalian oligopeptide-proton symporters, PepT1 and PepT2. EMBO J 30: 417–426. **(2011).**

O

Okazaki N, Kuroda M, Shimamoto T, Shimamoto T, Tsuchiya T. Characteristics of the melibiose transporter and its primary structure in Enterobacter aerogenes. Biochim Biophys Acta - Biomembr. 1326(1):83–91. (**1997).**

P

Pao S., Paulsen IT, Saier MH. Major facilitator superfamily. Microbiol Mol Biol Rev.;62(1):1–34. (**1998)**

Poolman B, Knol J, van der Does C, Henderson PJ, Liang WJ, Leblanc G, et al. Cation and sugar selectivity determinants in a novel family of transport proteins. Mol Microbiol.;19(5):911–22. (**1996)**

Pourcher T, Bassilana M, Sarkar HK, Kaback HR LG. Melibiose permease and alpha-galactosidase of Escherichia coli: identification by selective labeling using a T7 RNA polymerase/promoter expression system.Biochemistry ,29,690-696.(**1990).**

Pourcher T, Zani ML, Leblanc G. Mutagenesis of Acidic Residues in Putative Membrane-spanning segments of the Melibiose Permease of Escherichia coli. J Biol Chem.;268(32). (**1993).**

Pourcher.T, Sophie.L, Gerard.B GL. Melibiose permease of Escherichia coli: Large scale purification and evidence that H+, Na+, and Li+ sugar symport is catalyzed by a single polypeptide. (**1995).**

Pourcher T, Bibi E, Kaback HR, Leblanc G. Membrane topology of the melibiose permease of Escherichia coli studied by melB-phoA fusion analysis. Biochemistry.;35(13):4161–8.**(1996).**

Poulsen SB, Fenton RA, Reig T. Sodium- glucose cotransport. *Curr Opin Nephrol* Hypertens. (**2015)**

Purhonen P, Lundbäck AK, Lemonnier R, Leblanc G, Hebert H. Three-dimensional structure of the sugar symporter melibiose permease from cryo-electron microscopy. J Struct Biol.;152(1):76–83. **(2005).**

R

Reddy V.S, Maksim A. Shlykov, Rostislav Castillo EIS and MHSJ. The major facilitator superfamily (MFS) revisited. FEBS J;279(11):2022–35. (**2012)**

Rigaud JL, Pitard B, Levy D. Reconstitution of membrane proteins into liposomes: application to energy-transducing membrane proteins. Biochimica et Biophysica Acta.1231. 223-246 **(1995)**

S

Saier M.H.J, JR . A functional-phylogenetic classification system for transmembrane solute transporters. 64(2):354–411.(**2000a)**.

SaierM.H.J. Families of transmembrane sugar transport proteins. Mol Microbiol..35(4):699–710. **(2000b)**

Scalise, M., Pochini, L., Giangregorio, N., Tonazzi, A. & Indiveri, C. Proteoliposomes as tool for assaying membrane transporter functions and interactions with xenobiotics. *Pharmaceutics* **5,** 472–497 **(2013).**

Seidner G, Alvarez MG, Yeh JI, O'Driscoll KR, Klepper J, Stump TS, Wang D, Spinner NB, Birnbaum MJ, De Vivo DC. GLUT-1 deficiency syndrome caused by haploinsufficiency of the blood-brain barrier hexose carrier. Nat Genet.18:188–91 **(1998)**

Shimamura T, Weyand S, Beckstein O, G. N, Rutherford, Jonathan M. Hadden, David Sharples, Mark S.P. Sansom5 SI, Peter J. F. Henderson6,† and ADC. Molecular Basis of Alternating Access Membrane Transort by the Sodium-Hydantoin Transporter, Mhp1. October.;328(5977):470–3, **(2010)**

Simonescu CM. Application of FTIR Spectroscopy in Environmental Studies. Adv Asp Spectrosc. (1):49–84. **(2012).**

Stone J.E, Saam J, Hardy D.J.,Vandivort K.L., Wen-mei W. Hwu, and Klaus Schulten. High performance computation and interactive display of molecular orbitals on GPUs and multi-core CPUs. In Proceedings of the 2nd Workshop on General-Purpose Processing on Graphics Processing Units, ACM International Conference Proceeding Series, volume 383. **(2009)**

Surewicz WK, Mantsch HH CD. Determination of protein secondary structure by Fourier transform infrared spectroscopy: a critical assessment. Biochemistry (**1993).**

T

Tastan O, Klein-Seetharaman J, Meirovitch H. The effect of loops on the structural organization of alpha-helical membrane proteins. Biophys. J. 96(6) 2299-312 **(2009).**

Tatulian SA. Attenuated total reflection Fourier transform infrared spectroscopy: a mothod of choice for studying membrane proteins and lipids. Biochemistry. 42(41):11898-907. **(2003).**

Thevenin D and Lazarova T. Identifying and measuring transmembrane helix-helix interactions by FRET **914,** 87-106 **(2012).**

Tsuchiya T, Takeda K WT. H + -substrate cotransport by the melibiose membrane carrier in Escherichia coli. Membr Biochem. (**1980).**

V

Venyaminov S, Kalnin NN Quantitative IR spectrophotometry of peptide compounds in water (H2O) solutions. I. Spectral parameters of amino acid residue absorption bands. Biopolymers 30: 1243–1257. **(1990)**

W

Wallin E and Heijne G. Genome-wide analysis of integral membrane proteins from eubacterial, archaean, and eukaryotic organisms. Protein Sci.;7(4):1029–38.**(1998)**
Weissborn AC, Botfield MC, Kuroda M, Tsuchiya T, Wilson TH. The construction of a cysteine-less melibiose carrier from E. coli. Biochim Biophys Acta - Biomembr. 1329(2):237–44. **(1997)**
Weng J, Wang W. Molecualr dynamics simulation of membrane proteins. 805:305-29. 805:305-29 **(2014).**
Weinglass, A.B., and H.R. Kaback. The central cytoplasmic loop of the major facilitator superfamily of transport proteins governs efficient membrane insertion. Proc. Natl. Acad. Sci. USA.97:8938–8943**(2000).**
Whitley P, Zander T, Ehrmann M, Haardt M, Bremer E, von Heijne G. SecA-independent translocation of the periplasmic N-terminal tail of an Escherichia coli inner membrane protein. Position-specific effects on translocation of positively charged residues and construction of a protein with a C-terminal translocation signal. J Biol Chem.Dec 15;270(50):29831-5**. (1995).**
Winkelmann G. Microbial Transport Systems**(2008).**
Wilson DM and **Wilson TH.** Cation specificity for sugar substrates of the melibiose carrier in Escherichia coli. Biochim Biophys Acta. 904(2):191–200. **(1987).**

Wilson H and Wilson T.H. Primary Structure and Characteristics of the Melibiose Carrier of Klebsiella pneumoniae*. (26):18371–6. **(1992)**.
Wilson DM, Hama H, and Wilson TH**.** Gly113-Asp can restore activity to the Asp-Ser Mutant in the Melibiose carrier of *Escherichia Coli. Biophys. Res. Commun.* 2009,242-9 **(1995)**
Wilson TH and **Wilson DM**. Evidence for a close association between helix IV and helix XI in the melibiose carrier of Escherichia coli. Biochimica et Biophysica Acta. (**1998).**

X

Xie, H. Activity assay of membrane transport proteins. *Acta Biochim. Biophys. Sin. (Shanghai).* **40,** 269–277 **(2008).**

Y

Yamashita A, Singh SK, Kawate T, Jin Y, Gouaux E. Crystal structure of a bacterial homologue of Na+/Cl--dependent neurotransmitter transporters. Nature.;437(7056):215–23. (**2005).**
Yazyu .H , Sumiko S.N, Tadashi S, Hiroshi K. TsuchiyaS and T. Nucleotide Sequenceof the melB Gene and Characteristicscof Deduced Amino Acid Sequence of the Melibiose Carrier in Escherichia coli*;261(4):1782–5. **(1984)**
Yazyu H, Shiota S, Futai M, Tsuchiya T. Alteration in cation specificity of the melibiose transport carrier of Escherichia coli due to replacement of proline 122 with serine. J Bacteriol.;162(3):933–7. **(1985).**

Yousef MS, Guan L. A 3D structure model of the melibiose permease of Escherichia coli represents a distinctive fold for Na+ symporters. Proc Natl Acad Sci U S A.;106(36):15291–6.(**2009).**

Z

Zani ML, Fourcher T, Leblanc G. Mutation of polar and charged residues in the hydrophobic NH2-terminal domains of the melibiose permease of Escherichia coli. J Biol Chem.;269(40):24883–9.(**1994).**

Zeng H, Parthasarathy R, Rampal a L, Jung CY. Proposed structure of putative glucose channel in GLUT1 facilitative glucose transporter. Biophys ; 70(1):14–21. (**1996).**

Zhang, Y., R.H. Fillingame. Subunits coupling H+ transport and ATP synthesis in the Escherichia coli ATP synthase. Cys-Cys cross-linking of F1 subunit epsilon to the polar loop of F0 subunit c. J.Biol.Chem.(**1995).**

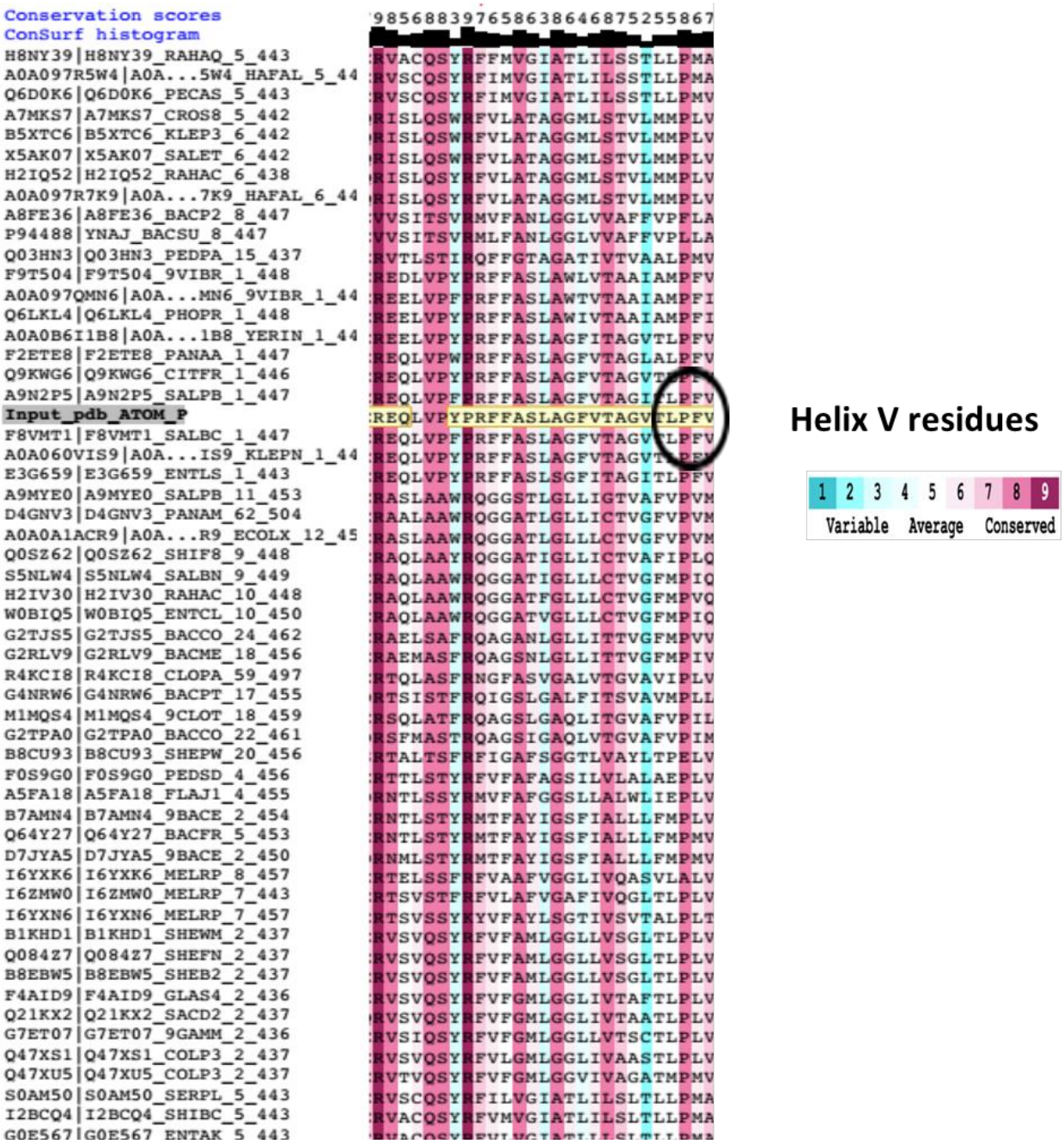
Appendix I:

Amino acid sequence of E. *coli* MelB

MSISMTTKLSYGFGAFGKDFAIGIVYMYLMYYYTDVVGLSVGLVGTLFLVARIWDAINDPIMGWIVN
ATRSRWGKFKPWILIGTLANSVILFLLFSAHLFEGTTQIVFVSVTYILWGMTYTIMDIPFWSLVPTITLD
KREREQLVPYPRFFASLAGFVTAGVTLPFVNYVGGGDRGFGFQMFTLVLIAFFIVSTIITLRNVHEVFSS
DNQPSAEGSHLTLKAIVALIYKNDQLSVLLGMALAYNVASNIITGFAIYYFSYVIGDADLFPYYLSYAGA
ANLVTLVFFPRLVKSLSRRILWAGASILPVLSSGVLLLMALMSYHNVVLIVIAGILLNVGTALFWVLQVI
MVADIVDYGEYKLHVRSESIAYSVQTMVVKGGSAFAAFFIAVVLGMIGYVPNVEQSTQALLGMQFI
MIALPTLFFMVTLILYFRFYRLNGDTLRRIQIHLLDKYRKVPPEPVHADIPVGAVSD VKAHHHHHH

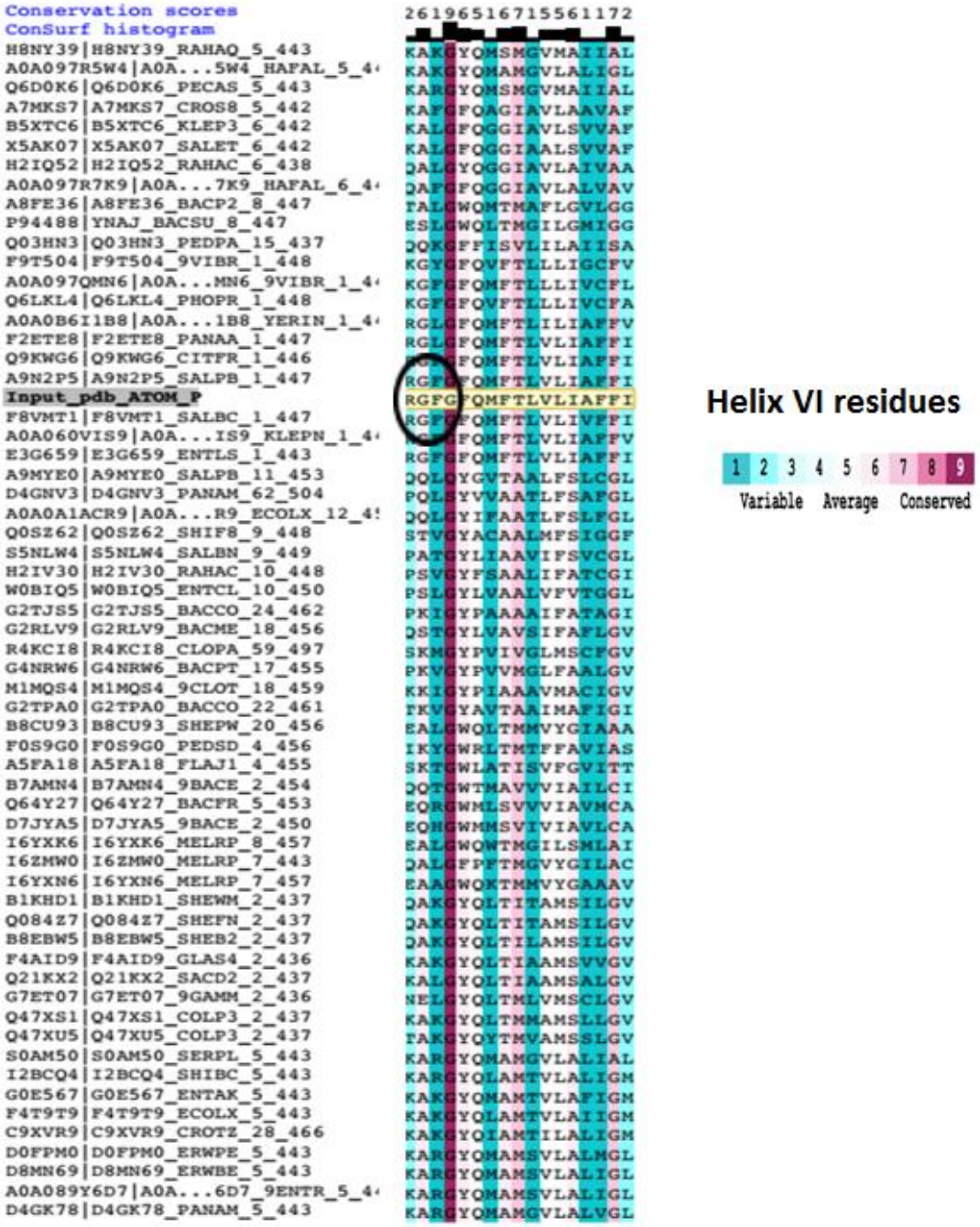
(This has been used as a template for all the site-directed mutagenesis)

Appendix II:



Multiple Sequence alignment of MelB with different MFS transporters. The yellow highlighted bar shows the Input sequence of MelB from the region 141 to 167 of Helix V in MelB and the number shown above represents the level of conservation. Alignment was done using CONSURF web-based program.

Appendix III:



Multiple Sequence alignment of MelB with different MFS transporters. The yellow highlighted bar shows the Input sequence of MelB from the region 175 to 210 of Helix VI in MelB and the number shown above represents the level of conservation. Alignment was done using CONSURF web-based program.

Appendix IV

Command for RMSF analysis in MD simulations

```
#load trajectory
mol load psf protein.psf
mol addfile protein.dcd

#select CA atoms
set ca [atomselect top "protein and name CA"]
set rmsf [measure rmsf $ca]
set outfile [open "rmsfCA.txt" w]
foreach x $rmsf {
  puts $outfile $x
}
close $outfile
```

Appendix III

1. Configuration files for running minimization using NAMD

 # Periodic conditions. 100 ps.

```
Coordinates          ionized.pdb
structure            ionized.psf
parameters           par_all27_prot_lipid_na.inp
paratypecharmm       on
set output           minimized
outputname           $output
dcdfile              ${output}.dcd
xstFile              ${output}.xst
dcdfreq               500
xstFreq               500
binaryoutput         no
binaryrestart         no
outputEnergies       200
restartfreq           1000
rigidbonds            all
rigidTolerance        0.0005
exclude               scaled1-4
1-4scaling            1
COMmotion            no
dielectric           1.0
switching            on
switchdist           8
cutoff               10
```

```
pairlistdist              12
firsttimestep              0
stepspercycle             20
timestep                  1.0
nonbondedFreq             1
fullElectFrequency        2
set temperature           298
temperature               $temperature
cellBasisVector1    x     0     0  ; # according to cell dimensions.
cellBasisVector2    0     y     0
cellBasisVector3    0     0     z
cellOrigin          x   y   z
wrapWater                  on
wrapAll                   on
wrapNearest               off
PME                       yes
PMEGridSpacing            1.0
minimize             100000
```

2. Configuration files for running equilibration using NAMD

100 ps. Heating from 0K in 2K increments.

 # input

```
set input                 minimized
coordinates                ${input}.coor
extendedSystem            ${input}.xsc
structure                 ionized.psf
parameters                par_all27_prot_lipid_na.inp
paratypecharmm             on
```

 # output

```
set output                equilibrated
outputname                $output
dcdfile                   ${output}.dcd
xstFile                    ${output}.xst
dcdfreq                   200
xstFreq                   500
binaryoutput               no
binaryrestart             no
outputEnergies            200
restartfreq               1000
```

 # mobile atom selection:

```
fixedAtoms                 off
```

 # Basic dynamics

```
exclude                   scaled1-4
1-4scaling                1
COMmotion                  no
dielectric                 1.0
```

 # Simulation space partitioning

```
switching                 on
switchdist                8
```

```
cutoff                  10
pairlistdist            12
# Multiple timestepping
firsttimestep           0
stepspercycle           20
timestep                1.0  ;# 1fs/step
rigidBonds              all  ;# needed if 2fs steps
nonbondedFreq           1    ;# nonbonded forces every step
fullElectFrequency      2    ;# PME only every other step
wrapWater               on  ;# wrap water to central cell
wrapAll                 on ;# wrap other molecules too
wrapNearest             off ;# use for non-rectangular cells

#PME (for full-system periodic electrostatics)
PME                     yes
# let NAMD determine grid
PMEGridSpacing          1.0
# Scripting
temperature             0     # starting at 0 K
reassignFreq            500
reassignIncr            2
reassignHold            298      # target T: 298 K
numsteps                100000
```

2. Configuration files for running dynamics using ACEMD

Production in linux with ACEMD (200 ns) after equilibration with NAMD in Win.

```
# input

set input               equilibrated

coordinates             ${input}.coor
extendedsystem          ${input}.xsc
velocities              ${input}.vel
structure               ionized.psf
parameters              par_all27_prot_lipid.prm
# output
restart                 on
set output              Melbec-200
outputname              $output
dcdfile                 ${output}.dcd
dcdfreq                 25000
restartfreq             1000000
restartname             ${output}
restartsave             on
energyfreq              10000
# Basic dynamics
exclude                 scaled1-4
1-4scaling              1
dielectric              1.0
# Simulation space partitioning
```

```
switching                   on
switchdist                  8
cutoff                      10
pairlistdist                12

# Multiple timestepping

timestep                    4.0  ;# 4fs/step
rigidbonds                  all  ;# needed for 4fs steps
fullelectfrequency          2    ;# PME only every other step
hydrogenscale               4    ;# needed for 4fs steps
# Constant Temperature Control
langevin                    on           ;# langevin dynamics
langevindamping             2            ;# damping coefficient of 2/ps
langevintemp                298          ;# random noise at this level
berendsenpressure                 on
berendsenpressuretarget           1.01325
berendsenpressurerelaxationtime   400
useconstantratio                  on
#PME (for full-system periodic electrostatics)
pme             on
wrap            on           ;# wrap all to central cell
pmegridspacing  1.0
# Scripting
run             50000000       ;# 200 ns (because 4 fs/step)
```

Firstly, I would like to express my eternal gratitude to my supervisor, **Prof. Josep Cladera**. I have been always glad to have a supervisor who gave me the freedom to explore on my own, and at the same time the guidance to recover when my steps faltered. His great patience and guidance helped me finish this dissertation without any stress.

A very special thanks goes to my co-supervisor, **Prof. Esteve Padros.** His kindness in teaching encouraged me to learn a lot of new things. I once again thank my supervisors for giving me an opportunity to start my PhD in Barcelona, without them it would have been difficult for me to receive the fellowship from FI-DGR AGAUR. Thanks to AGAUR.

Dr. Jordi Benach, Dr. Fernando Gil deserves a special thanks, for their guidance and support in handling the crystals in synchrotron. I really enjoyed the time in ALBA. Our team was great!!

I would like to thank Dr. Carlos Gil, for assisting me in doing the transport assays in MelB. His extreme patience and guidance helped me to set up the transport assay experiment for the first time.

Thanks as well to my yearly committee members Dr. Ramon Barnadas, Dr. Jose Luis. Your guidance and ideas after my presentation helped me to do lot of things.

My special thanks to Dr. Tzvetana Lazarova and Elodia Serrano**,** their kind words always made me feel better.

Next, I would like to thank Dr. Alex Peralvarez, for his suggestions and ideas.

Thanks to Prof. Joan Manyosa and Prof.Antoni Morros for your cava, wine and cheese suggestions during our department get together. ☺

My special thanks to Pablo, Dr. Elena, Guillem, Carla, Claus, Anita and Ping. It was always nice to spend time with you. I will miss you!!!

Thanks to all the members of my lab Prof. Mireia Dunach, Josu, Nuria, Fanli, Asrar, Arash, LiYing, Albert, Dr. Beatriz, Meribel, Meritxell, Gloria, Roger, Eduvardo, Atieh, Erika, Mateu and Marco, it was always a great pleasure and fun working with you.

My special thanks to you Kevin, I cannot imagine how I would be if I haven't met you. Your love, support, humor, encouragements etc., always make me feel great. <3 <3 ☺

Lastly, I would like to dedicate this thesis to my parents who provided their support through my entire life and encouraging me with their best wishes. And also I thank my entire family for their wishes.