# Structural and Functional Studies of Melibiose permease of *Escherichia coli*

TESI DOCTORAL

Yibin Lin
CEB-Unitat de Biofísica
UAB 2012

Universitat Autònoma de Barcelona
Departament de Bioquímica i Biologia Molecular
Facultat de Medicina

# Structural and Functional Studies of Melibiose Permease of *Escherichia coli*

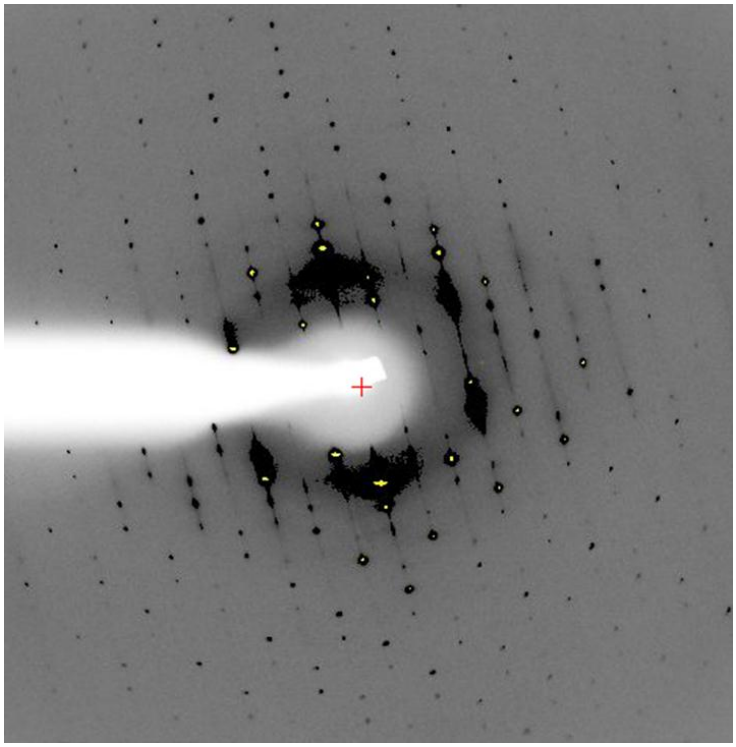
Memòria presentada per Yibin Lin per optar al grau de doctor.

Aquest treball ha estat realitzat en la Unitat de Biofísica del Departament de Bioquímica i de Biologia Molecular, Facultat de Medicina – Centre d'Estudis en Biofísica de la Universitat Autònoma de Barcelona.

Vist i plau del director de la tesi:

Dr. Esteve Padrós i Morell

Bellaterra, Juny, 2012



**X-ray diffraction pattern of a MelB crystal**

业精于勤荒于嬉

The progress of studies are achieved through painstaking endeavors while ends with playing around.

行成于思毁于随

A deed is accomplished through thinking and destroyed by idleness.

——Yu Han, Tang Dynasty, writer, thinker, politician

# Index

# Abbreviations

$\lambda_{em}$,emitted wavelength
$\lambda_{ex}$,excited wavelength
A, Ala, alanine
ABS, absorbance
APS, ammonium persulfate
ATR-FTIR, attenuated total reflection-Fourier transform infrared
BME, β-mercaptoethanol
BSA, bovine serum albumen
C, Cys, cysteine
$C_{12}E_8$, dodecyl octaethylene glycol ether
$C_{12}E_9$, dodecyl nonaethylene glycol ether
$C_8E_4$, tetraethylene glycol monooctyl ether
CFTR, cystic fibrosis transmembrane conductance regulator
CMC, the critical micelle concentration
cryo-EM, Cryo-electron microscopy
Cymal-5, 5-cyclohexyl-1-pentyl-β-D-maltoside
Cymal-6, 6-cyclohexyl-1-hexyl-β-D-maltoside
Cys-less, the mutant proteins without cysteines
D, Asp, aspartic
$D^2G$,$dns^2$-S-Gal,2'-(N-dansyl)-aminoethyl-1-thio-D-galactopyranoside
$dH_2O$, distilled water
$ddH_2O$, double distilled water
DDM, n-Dodecyl β-D-maltoside
DM, Decyl β-D-maltopyranoside
DMSO, dimethyl sulfoxide
DTT, 1, 4-dithio-threitol
E, Glu, glutamic
EDTA, ethylenediaminetetraacetic acid
EM, electron microscopy
ESRF, the European Synchrotron Radiation Facility located in Grenoble - France
F, Phe phenylalanine
FocA, a member of the formate–nitrite transporter family, which transports short-chain acids
Fos-Choline-10, n-Decylphosphocholine
FRET, energy transfer by resonance fluorescence
FTIR, The Fourier transform infrared spectroscopy

GlpT, glycerol-3-phosphate transporter
GltPh, a homolog of the glutamate transporter family from *Pyrococcushorikoshii*
*GPCRs, G protein coupled receptors*
GPH, glycoside-pentoside-hexuronide:cationsymporter family
H, His, histidine
H/D, hydrogen/deuterium
HPLC, High-performance liquid chromatography
IMAC, immobilized metal affinity chromatography
IMPs, integral membrane proteins
IMV, inverted membrane vesicles
IPTG, isopropyltio-β-galactoside
$IR_{diff}$, IR difference spectrum
ISO, Inside-out
K, Lys, lysine
KPi, potassium phosphate buffer
L, Leu, leucine
lactose, 4-O-β-galactopyranosyl-D-glucose
LacY, lactose permease of *Escherichia coli*
LAPAO, (3-lauryl-starch)-N, N'-(dimethylamino) propylamine oxide
LB, Luria Broth
LCP, lipidic cubic phase
LDAO, Lauryldimethylamineoxide
$LeuT_{Aa}$, Leucine transporter from the bacterium *Aquifexaeolicus*
M, Met, methionine
MAD, multi-wavelength anomalous diffraction
MAPEG, membrane-associated proteins in eicosanoid, and gluththione metabolism
Mel, melibiose
MelB, melibiose permease of *Escherichia coli*
$MelB_{EC}$, melibiose permease of *Escherichia coli*
$MelB_{KP}$, melibiose permease from *Klebsiellapneumoniae*
$MelB_{SY}$, melibiose permease of *Salmonella typhimurium*
Melibiose, 6-O-α-D-galactopyranosyl-D-glucose
MFS, major facilitator superfamily
Mhp1, the hydantoin transporter from *Microbacterium liquefaciens*
MIANS, 2 - (4'-maleimidylanilino) naphthalene-6-sulfonic acid
MTSET, 2-(trimethylammonium)ethylmethanethiosulfonate
MW, molecular weight
N, Asn, asparagine
NaPi, sodium phosphate buffer
Native PAGE, polyacrylamide gel electrophoresis in the absence of sodium dodecyl sulfate
NEM, N-ethylmaleimide

NG, n-nonyl-ß-D-glucoside
NM, n-nonyl-*β*-D-maltoside
NMA, N-metilacetamida
NMR, nuclear magnetic resonance
NYCOMPS, New York Consortium on Membrane Protein Structure
$OD_{600}$, the optical density of a sample measured at a wavelength of 600 nm
OG, n-octyl-β-D-glucoside
OM, n-octyl-β-D-maltoside,
PCR, polymerasechainreaction
PDB, Protein Data Bank
PDC, protein-detergent complex
PEG, polyethylene glycol
PL, phospholipids
Q, Gln, glutamine
R, Arg, arginine
$R^2$, the correlation coefficient
RSO, right-side-out
S, Ser, serine
SAD, single-wavelength anomalous diffraction
SDS-PAGE, sodium dodecyl sulfate polyacrylamide gel electrophoresis
SeMet, selenomethionine
SGC, Structural Genomics Consortium
TCDB, transporter classification database
TDG, β-D-galactopyranosyl-1-guy-β-D-galactopiranoside
TEMED, Tetramethylethylenediamine
TEV, Tobacco Etch Virus
Thimerosal, Ethyl mercurithiosalicylate
TM, transmembrane
TMG, methyl-β-D-thiogalactopyranoside
TMRM, Tetramethylrhodamine-5-maleimide
UDM, Undecyl-β-d-maltoside
V, Val, valine
*vSGLT,* the $Na^+$/galactosecotransporter of *Vibrio parahaemolyticus*
W, Trp, tryptophan
Y, Tyr, tyrosine
α-DDM, n-dodecyl-α-D-maltoside
α-NPG, p-nitrophenyl α-D-galactopiranoside
ΔCTMelB R149C, MelB R149C removed the last C-terminal 17 residues and the engineered His-tag.

# 1 Introduction

*"…The vast majority of diseases are caused due to lack of aspecific membrane protein, 80% of the drugs currently on the market target membrane proteins…"---Hartmut Michel in Beijing, 2006*

## 1.1 The importance of membrane proteins

Membrane protein is the general term of a series of proteins that are attached to, or integrated into the membrane of cell or organelle. Nearly 40% of all known proteins are membrane proteins that includeamong others receptors, channel proteins, transporters and signaling molecules, which are necessary for cell communication, nutrition uptake, signal transduction, etc. Importantly, more than half of all proteins in the cell interact with the membrane. Depending on the strength and the membrane-bound mode, membrane proteins can be divided into three main categories: peripheral membrane proteins, integral membrane proteins (IMVs), and lipid-anchored proteins (Figure 1.1). Peripheral membrane proteins, accounting for 20%-30% of the membrane protein are temporarily bound to the membrane via either lipid polar head or via integral proteins by a combination of ionic bonds, hydrogen bonds, and other non-covalent interactions. The integral membrane proteinsaccountfor 70% to 80% of the membrane proteins embedded into the lipid bilayer to a different degree. They can be classified according to their dominating secondary structure: helical transmembrane proteins or betabarrels.Finally, the lipid-anchored membrane proteins areattached to different fatty acid acyl chains on the cytoplasmic side of the cell membraneand covalently bound to the lipid bilayer through lipidated amino acid residues.

Membrane proteins are very important in our bodies, and comprise 20–30% of all proteins in both prokaryotic and eukaryotic organisms (Wallin and von Heijne 1998). This proportion is higher in eukaryotes. For example, 40% of the 6,000 gene products encoded by the genome of baker's yeast are expected to be integral membrane proteins (Ostermeier and Michel 1997). The original human genome sequence project estimated 15%-39% of the total gene count of 31,778 genes to code for membrane proteins

(Lander et al. 2001; Almen et al. 2009). Membrane proteins play very important roles in many important physiological processes, e.g., aquaporins (uptake water), transporters (uptake solublemolecules), ion channels (involved in nerve impulses), G-protein coupled receptors (signal transduction), etc. This enormous family of proteins is of high medical importance indicated by the fact that mutations in membrane proteins are involved in many common diseases, including heart diseases, high blood pressure and angina, related to the malfunctioning of ion channels;cancer, where errors in signaling pathways can lead to cells dividing out of control; migraine, depression and Alzheimer's, linked to problems with transporters and channels; cystic fibrosis, a genetic disorder caused by mutations in the cystic fibrosis transmembrane conductance regulator (CFTR) gene which encodes a chloride ion channel. Since many membrane proteins sit at the surface of cells, over 60% of small molecule drugs target to membrane proteins. In particular, more than 400 membrane transporters annotated in the human genome play a part *in vivo* in drug disposition, therapeutic efficacy and adverse drug reactions(Brunton 2006).

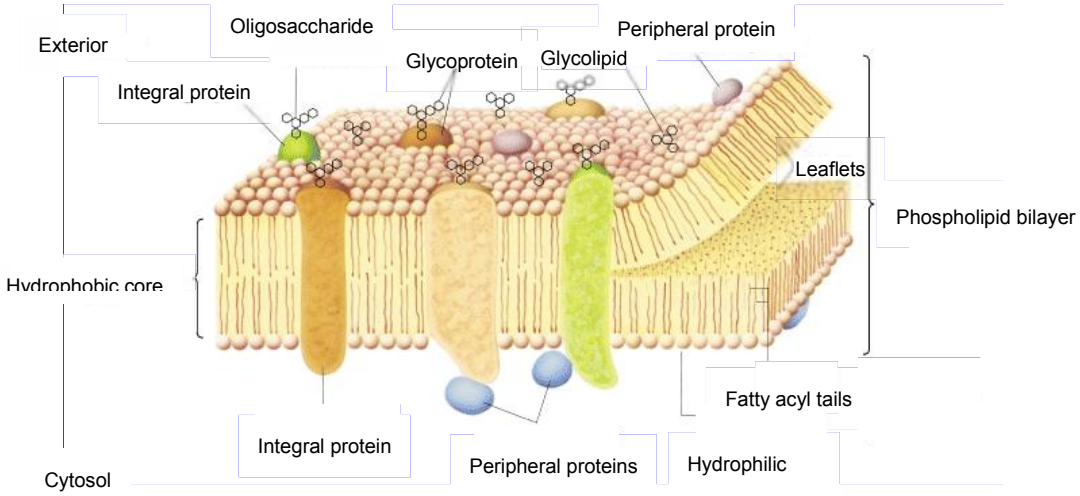



**Figure 1.1**Schematic diagram of typical membrane proteins in a biological membrane. According to (Lodish et al. 2000).

## 1.2 The membrane transport proteins

Fewer molecules enter or leave cells, or cross organellar membranes, without being aided by proteins. Membrane transporter protein is a class of specific membrane proteins which catalyze the movement of a number of substances across the cell membranes. Membrane transport proteins are integral membrane proteins or a complex of transmembrane molecules. The main function of this type of proteins is involved in the passive transport (promoting diffusion) or active transport (transport pumps). The membrane transport proteins involved in promoting the diffusion of molecules do not have enzyme activity, but they share enzymatic characteristics, such as the highest turnover rate, specificity and competitive inhibition. Therefore, the transport protein is also known as an enzyme (permease). Transport proteinsare divided into three categories: carrier (transporter), ion channel, and the ion pump (Figure 1.2A), according to the mode of action.A membrane transport protein exhibits a high degree of specificity for the substance transported. The rate of transport by the three types differs considerably owing to differences in their mechanism of action(Lodish et al. 2000).

The transporters move a wide variety of ions and molecules across cell membraneswith a relatively high specificity by alternately opening their binding site to the one and to the other membrane side. According to the mechanism of substrate transport, the transporter can be divided into three categories (Figure 1.2B). Uniporter mediates the movement of a single molecule across the membrane down a concentration gradient without input of energy. In contrast, antiporters and symporterstransportthe substrates across the membrane against its concentration gradient driven by electrochemical potential gradients. In a transporter classification database (TCDB, http://www.tcdb.org), the transporters are classified in over 600 families incorporating both functional and phylogenetic information (Saier 1999; Saier et al. 2006; Saier et al. 2009).

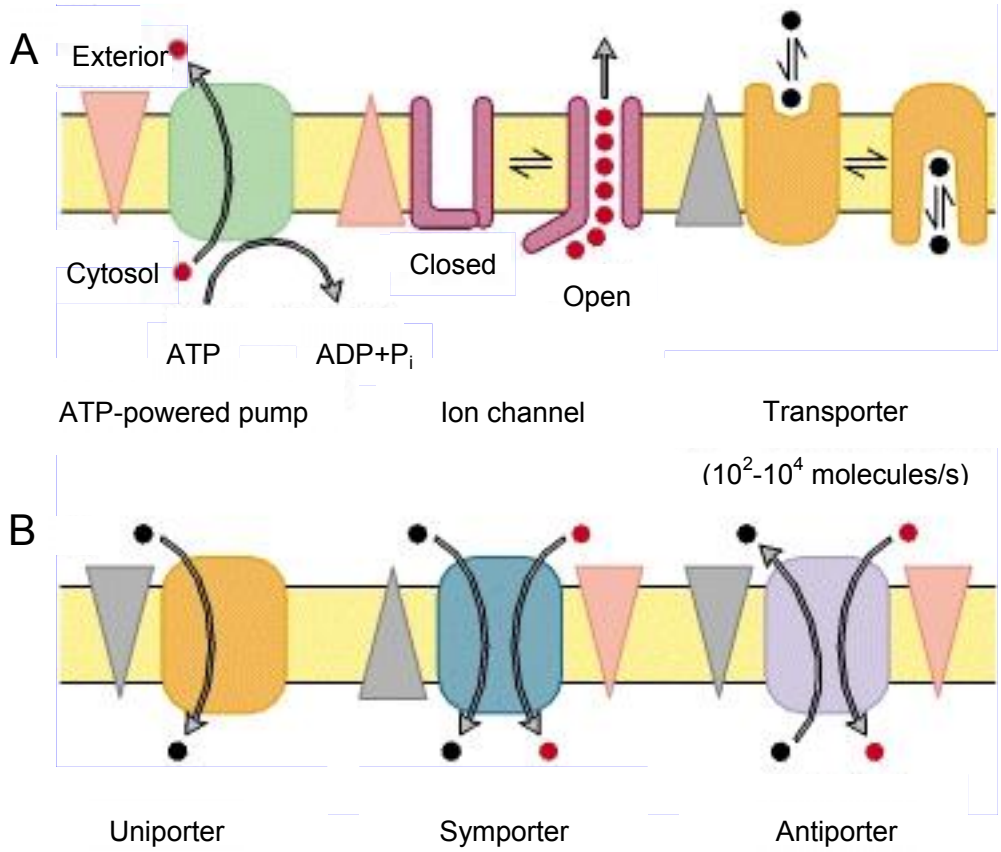



**Figure 1.2** Schematic diagrams illustrating the action of membrane transport proteins. Gradients are indicated by triangles with the tip pointing toward lower concentration, electrical potential, or both. (a) The three major types of transport proteins. Pumps utilize the energy released by ATP hydrolysis to power movement of specific ions (red circles) or small molecules against their electrochemical gradient. Channels catalyze movement of specific ions (or water) down their electrochemical gradient. Transporters, which fall into three groups, facilitate movement of specific small molecules or ions (black circles). (b) The three groups of transporters. Uniporters, transport a single type of molecule down its concentration gradient. Cotransport proteins (symporters and antiporters) catalyze the movement of one molecule against its concentration gradient (black circles), driven by movement of one or more ions down an electrochemical gradient (red circles). According to (Lodish et al. 2000).

Membrane transporters control the influx of essential nutrients and ions and the efflux of cellular waste, environmental toxins, and other xenobiotics. They play a very important role to maintain the normal physiology of the organism stable as well as in

drug safety and efficacy (Figure 1.3. for review about membrane transporters in drug development see Giacomini et al.2010). Consistent with their critical roles in cellular homeostasis, approximately 2000 genes in the human genome (~7% of the total number of genes) code for transporters or transporter-related proteins(Brunton 2006).

The core issues with membrane transporters are how they recognizesubstrates and carry themto pass through the membrane.

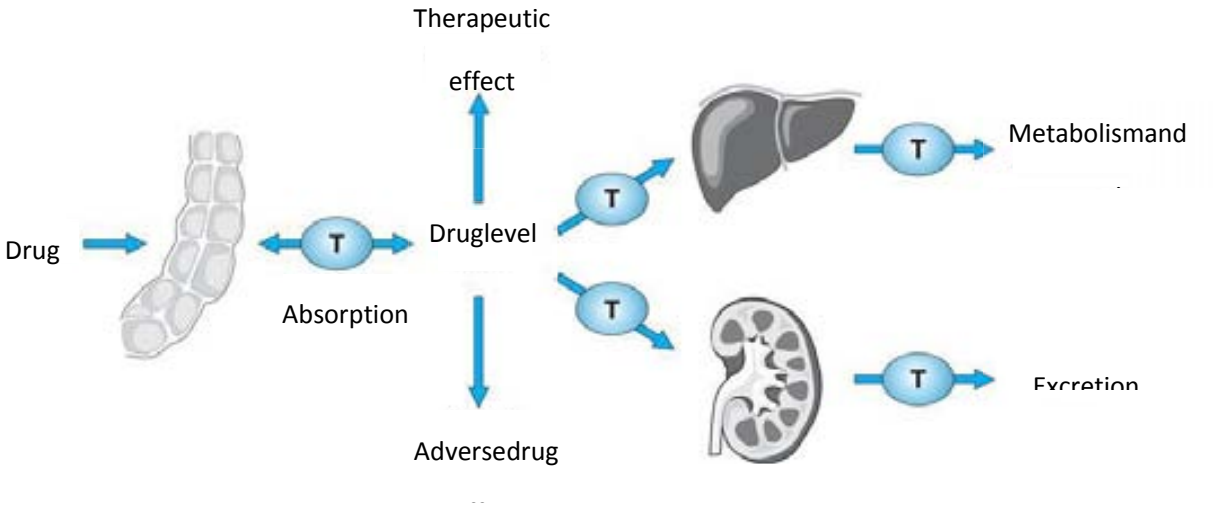



**Figure 1.3**Roles of membrane transporters in pharmacokinetic pathways. Membrane transporters (T) play roles in pharmacokinetic pathways (drug absorption, distribution, metabolism, and excretion), thereby setting systemic drug levels. Drug levels often drive therapeutic and adverse drug effects (According to Brunton 2006).

### 1.2.1 Substrate recognition

In the last years, several atomic structures of membrane transporters have been solved (See http://blanco.biomol.uci.edu/mpstruc/listAll/list for a listing of membrane proteins

of known 3D structure). They share several structural characteristics, however, the nature of the ligands used for substratesrecognition and binding varies. For example, the sodium galactose transporter vSGLT of *Vibrio parahaemolyticus* uses mainly carbonyl oxygens from the main chain(Faham et al. 2008), whereas melibiose permease mainly uses oxygens from Asp side chains (Toyoshima et al. 2000; Leon et al. 2006).Despite this, membrane transporters share several common characteristics of substrate recognition.

i) The substrate binding site is always at protein's center, e.g., LacY(Abramson et al. 2003), GlpT(Huang et al. 2003), LeuT(Yamashita et al. 2005), Mhp1(Shimamura et al. 2010), vSGLT(Faham et al. 2008), etc. This organizational pattern should facilitate to control the substrate pass through the membrane.

ii) For the co-transporters, either antiporter or symporter, two substrates binding sites are always organized together in space. In many cases, they share the same place. This organizational pattern mayfacilitate the cooperation betweenthe bindingof the two substrates.

iii) Substrate-binding sites are surrounded by some other residues from different transmembrane helices, which may offer a suitable charge, polarity, hydrophobic environment, etc for substrate recognition and binding. In general, we call this special environment the "binding pocket".

### 1.2.2 Translocation

The mechanistic basis for solute transport across membranes has been shaped by the alternating access mechanism (Patlak 1957; Jardetzky 1966; Vidaver 1966) indicated by the studies of dynamics and atomic level structure in recent years (for review seeForrestand Rudnick 2009 and Krishnamurthy et al.2009). In this mechanism, transporters bind only one (or some) substrate molecules at a time; after binding the substrate molecules, the transporter undergoes a conformational change such that the bound substrate molecules, and only these molecules, are transported across the membrane (Figure 1.4). In this process, two theories have been widely adopted to

explain how the substrates are transported through the membrane. One is the "rocker switch mechanism"(DeFelice 2004), in which substrate(s) bind in the middle of the protein followed by a conformational change from outward-facing to inward-facing (Figure 1.5A). The other one is the "gated mechanism"(Shimamura et al. 2010), in which substrate(s) binding induces the closing of the outside door and the opening of the inside door (Figure 1.5B).

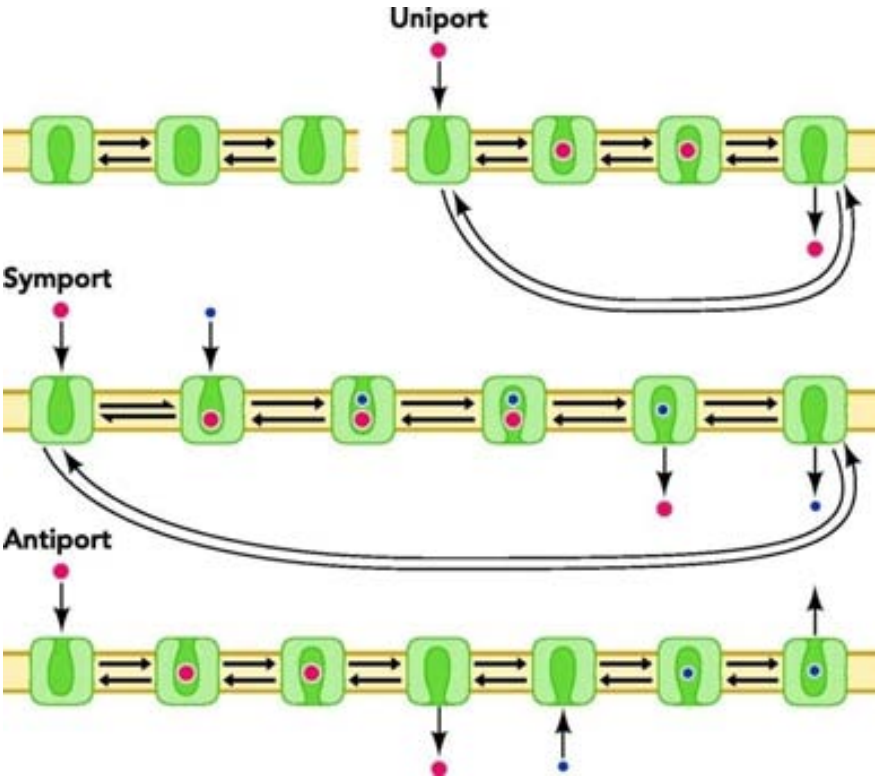



**Figure 1.4** Alternating-access mechanisms. A simple mechanism for alternately exposing a substrate binding site to either side of the membrane involves interconversion through an occluded intermediate (top left). (According to Forrest and Rudnick 2009)

A

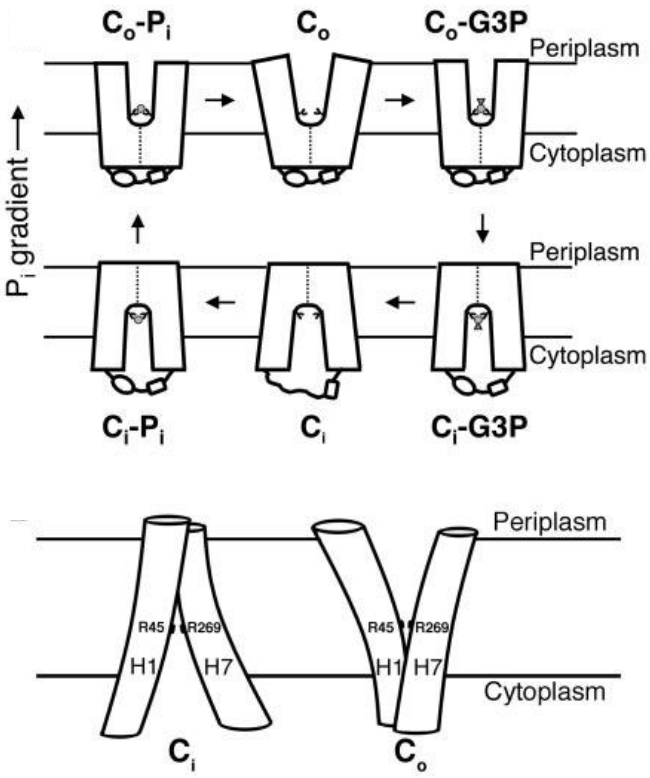
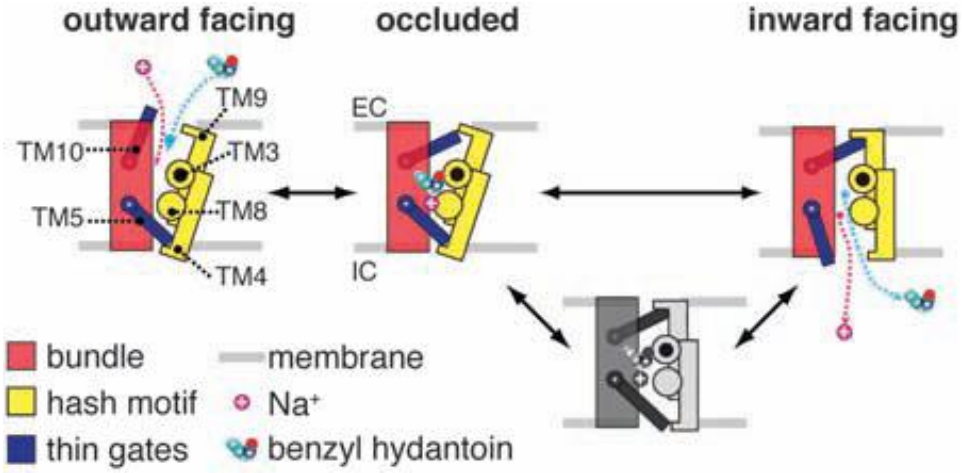



**Figure 1.5** Molecular basis of translocation mechanism.

(A) Rocker-switch mechanism of the glycerol-3-phosphate transporter (According to Huang et al. 2003). (A, top) Reaction cycle of substrate translocation. (A, bottom) Schematic drawing of central helices H1 and H7 in Ci (inward-facingconformation) andCo(outward-facing conformationn). Rockers witch–type movements of the helices that occur upon substrate binding allow the substrate-binding site, which comprises Arg45 and Arg269, to switch between the two sides of the membrane.

B




(B) Gate core mechanism. (According to Shimamura et al. 2010) Upon sodium and substrate binding, the extracellular thin gate (TM 10) closes to form the occluded state. The thick gate then opens with a rigid body rotation of TMs 3, 4, 8, and 9 (the hash motif) relative to TMs 1, 2, 6, and 7 (the bundle). Either independently or concomitantly with this, the intracellular thin gate (TM 5) also opens to allow the substrates to exit toward the cytoplasm.

## 1.3 The melibiose permease

The *Escherichia coli* melibiose permease (MelB) is a member of the glycoside-pentoside-hexuronide:cation symporter family (GPH), which in turn forms part of the major facilitator superfamily (Saier 2000). It is an integral membrane protein that couples the uphill transport of the disaccharide melibiose and other α- or β-galactosides towards the cellular cytosol, to the downhill electrochemical ion gradient of $Na^+$, $Li^+$, or $H^+$. It does so in a 1:1 ratio and the binding of the cations enhances significantly the affinity for the sugar (Pourcher et al. 1990; Pourcher et al. 1995; Poolman et al. 1996).The melibiose permease from *Escherichia coli* is coded by the*melB* gene, and is predicted to consist of 473 amino acid residues, resulting in a protein with a molecular weight of 52,029, 70% of which are apolar(Yazyu et al. 1984).

The GPH family has more than 50 members from bacteria, archaea, and eukaryotes (Saier2000). The studies showed that the homologues of the *E. coli* MelB protein include transporters that facilitate the uptake of galactosides, and use $Na^+$ and $H^+$ and $Li^+$, or only $H^+$ as coupling ions (Table 1.1).These different cation selectivities of the transporters and the varying degrees of primary sequence similarity, ranging from 18 to 85% indentity between pairs of proteins, offer an enormous advantage in indentifying locations/residues with a particular function (Poolman et al. 1996). In contrast, lactose permease (LacY) from *E.coli*, a member of the major facilitator superfamily (MFS) obligatorily couples the transport of galactosides to protons (Hunter and Schulman 2005). The MelB protein can use $Na^+$, $Li^+$ or $H^+$ as coupling ions, implying that there is a different cotransport mechanism between LacY and MelB.In-depth study of MelB would help to reveal the molecular mechanisms of cotransport.

**Table 1.1** Cation selectivity of members of the GPH family

| Substrate | $MelB_{EC}$ | $MelB_{SY}$ | $MelB_{KP}$ | $LacS_{ST}$ | $GusB_{EC}$ |
|---|---|---|---|---|---|
| α-galactosides | | | | | |
| melibiose | $H^+$, $Na^+$, $Li^+$ | $Na^+$, $Li^+$ | $H^+$ | $H^+$ | NT |
| β-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 | $H^+$ |

Melibiose, 6-O-galactopyranosyl-D-glucose; TMG, methyl-β-D-thiogalactopyranoside; lactose, 4-O-β-galactopyranosyl-D-glucose; nd, not determined; NT, not transported. Cations between brackets mean that the activation by them is very poor. $MelB_{EC}$, melibiose permease *Escherichia coli*; $MelB_{SY}$, melibiose permease of Salmonella typhimurium; $MelB_{KP}$, melibiose permease from Klebsiella pneumoniae; $LacS_{ST}$, lactose permease of Streptococcus thermophilus; $GusB_{EC}$, glucuronide transporter of *Escherichia coli*. The table data are described in Poolman et al. 1996.

## 1.4 Structural studies of the melibiose permease

### 1.4.1 The topological structure of the melibiose permease

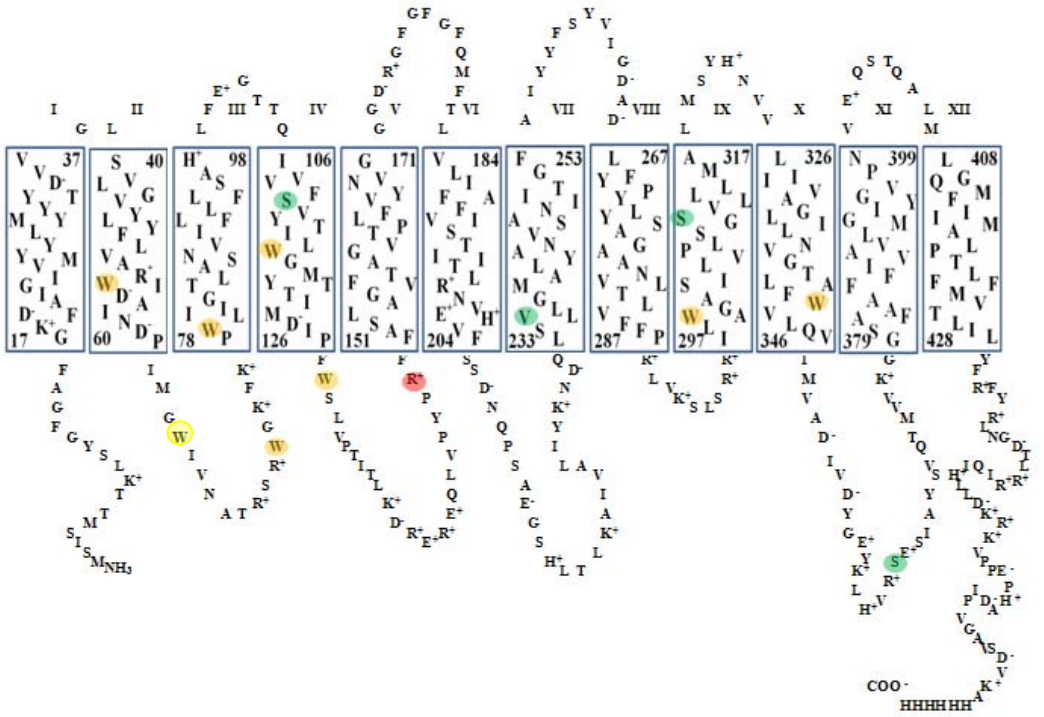


**Figure 1.6** Secondary structure prediction of MelB from *E. coli* where α-helices are represented as rectangles. (According to Pourcher et al. 1996) Four native cysteines were replaced by Ser or Val, colored in green, to obtain Cys-less MelB.

The membrane topological structure is the minimum structural information necessary to formulate and assess mechanistic models of integral membrane protein function (Botfield et al. 1992). The topological structure of the melibiose permease of

*Escherichia coli* was characterized by hydropathy analysis and genetic fusions to the compartment-specific reporter protein alkaline phosphatase. The melibiose permease was predicted to have a cytoplasmic amino terminus, two sets of six transmembrane domains separated by a large cytoplasmic loop ("six-loop-six" arrangement, and a 45-residue cytoplasmic carboxyl tail (Botfield et al. 1992). A more accurate membrane topological structure of melibiose permease was predicted to contain 12 transmembrane domains that traverse the membrane in zigzag fashion connected by hydrophilic loops that are exposed alternatively on the periplasmic or cytoplasmic surfaces of the membrane with the N and C termini on the cytoplasmic face (Figure 1.6, Pourcher 1996).

1.4.2 Toward 3D structure of melibiose permease

Crystallographic methods are the main means to access the protein three-dimensional structure. However, high quality MelB crystals have not been obtained yet. The crystal structures of similar transporters in other organisms are also not enough to understand the peculiarities of the reaction mechanism of MelB, due mainly to the low sequence identity between MelB and these transporters (~ 20% or less), to their use of different substrates when compared to MelB, and finally to the non-conservation of crucial residues. Nevertheless, some useful 3D structural information was obtained by cryoelectron microscopy from 2D crystals and from a 3D model threaded through a crystal structure of lactose permease.

i) Low resolution structure of melibiose permease.

A projection structure at 8Å resolution of the melibiose permease obtained by cryoelectron microscopy from 2D crystals showed that each monomer of MelB had a size of ~49Å×37Å, confirmed the total number of transmembrane helices as well as the feature of 2-helix bundles lining a central cavity (Figure 1.7).

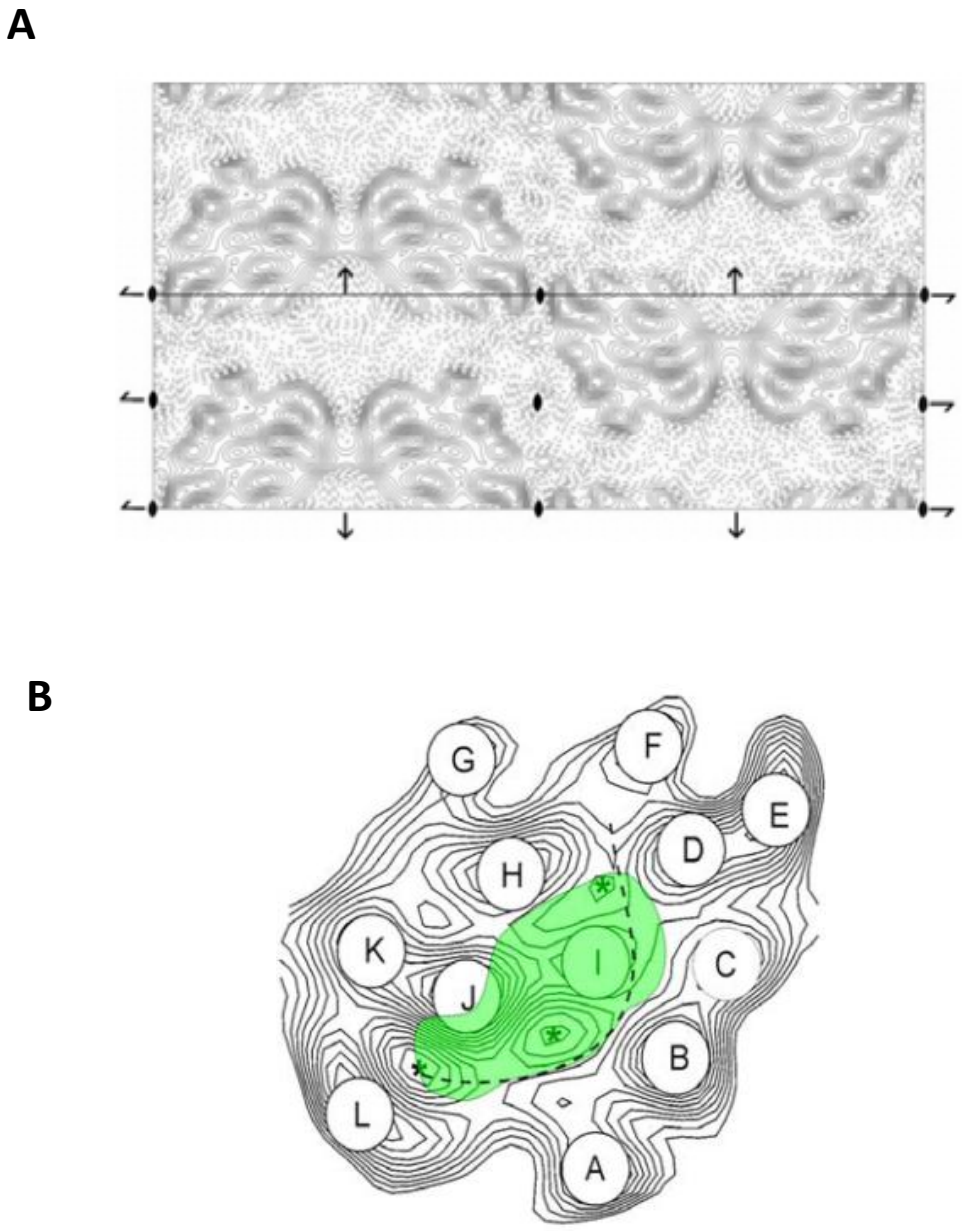



**Figure 1.7 (A)** Projection map at Å resolution of MelB from two-dimensional crystals. Four symmetry-related protein-dense regions are observed in each unit cell corresponding to the projections of the MelB molecules. (B) Protein-dense regions in the MelB map marked with letters (A-L). Asterisks indicate internal density minima in the structure. The dashed line shows protein-deficient regions dividing the projection structure into two parts. A possible aqueous channel corresponds to the elongated green domain enclosing the inner protein-deficient regions in the projection map. (According to Hacksell et al. 2002)

Furthermore, a three-dimensional structure of the melibiose permease was determined by cryo-electron microscopy at 10 Å resolution in the membrane plane from two-dimensional crystals(Purhonen et al. 2005).The three-dimensional map shows a heart-shaped molecule composed of two domains with a large central cavity between them (Figure 1.8),resembling those of lactose permease (Abramson et al. 2003), glycerol-3-phosphate transporter (Huang et al. 2003), and fucose transporter (Dang et al. 2010)of the major facilitator superfamily. However, MelB seems less symmetrical than in those of the major facilitator superfamily (Purhonen et al. 2005), implying that MelB may present a different pattern of how individual helices are organized, particularly around the cation-binding site, which responds to different cation selectivities.

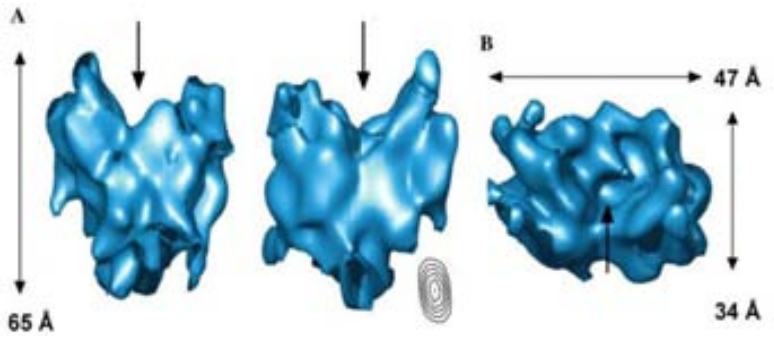



Figure 1.8 The 3D structure of MelB calculated to 10 Å resolution shows a heart-shaped molecule composed of two domains. (According to Purhonen et al. 2005)

ii) The proposals on the organization of the transmembrane helices of melibiose permease

Even in the absence of the high-resolution 3D structure, the organization of the transmembrane helices had been studied by simultaneous mutations of two amino acids, being the second mutation a revertant of the first. That is, the two mutations may be far away each other in the primary structure, but the functional defect introduced by the first mutation will be entirely or partially compensated by the second mutation (Wilson et al. 1995).Disulfide cross-linking of the transmembrane regions have also been used to gain information about the relative position of transmembrane helices (Ding and Wilson 2000a, b). The results indicated that (Figure 1.9):

Helix IV may be close to helix XI (Hastings Wilson and Wilson 1998), helix XI may be close to Helix II (Franco et al. 2001), helix I may be close to helix II, transmembrane domains IV and VII close to helix II in the threedimensional structure of the protein (Franco and Wilson 1999), Helix XI may be close to Helices I, II, and V (Ding and Wilson 2000b).

In addition, the organization of the transmembrane helices of the melibiose permease was implied from the analysis of the protein-dense regions in the MelB map obtained from 2D crystals(Hacksell et al. 2002). In this model, MelB was divided into two parts (Figure 1.9H). The analysis of reagentsreacting with Cys suggests thathelix I is completely surrounded by the aqueous environment(Ding and Wilson 2001a), helices II and XI are exposed about 35% residues to the aqueous channel (Ding 2003) and helix VI is embedded in phospholipid and does not face the aqueous channel through which melibiose passes(Ding and Wilson 2001a).

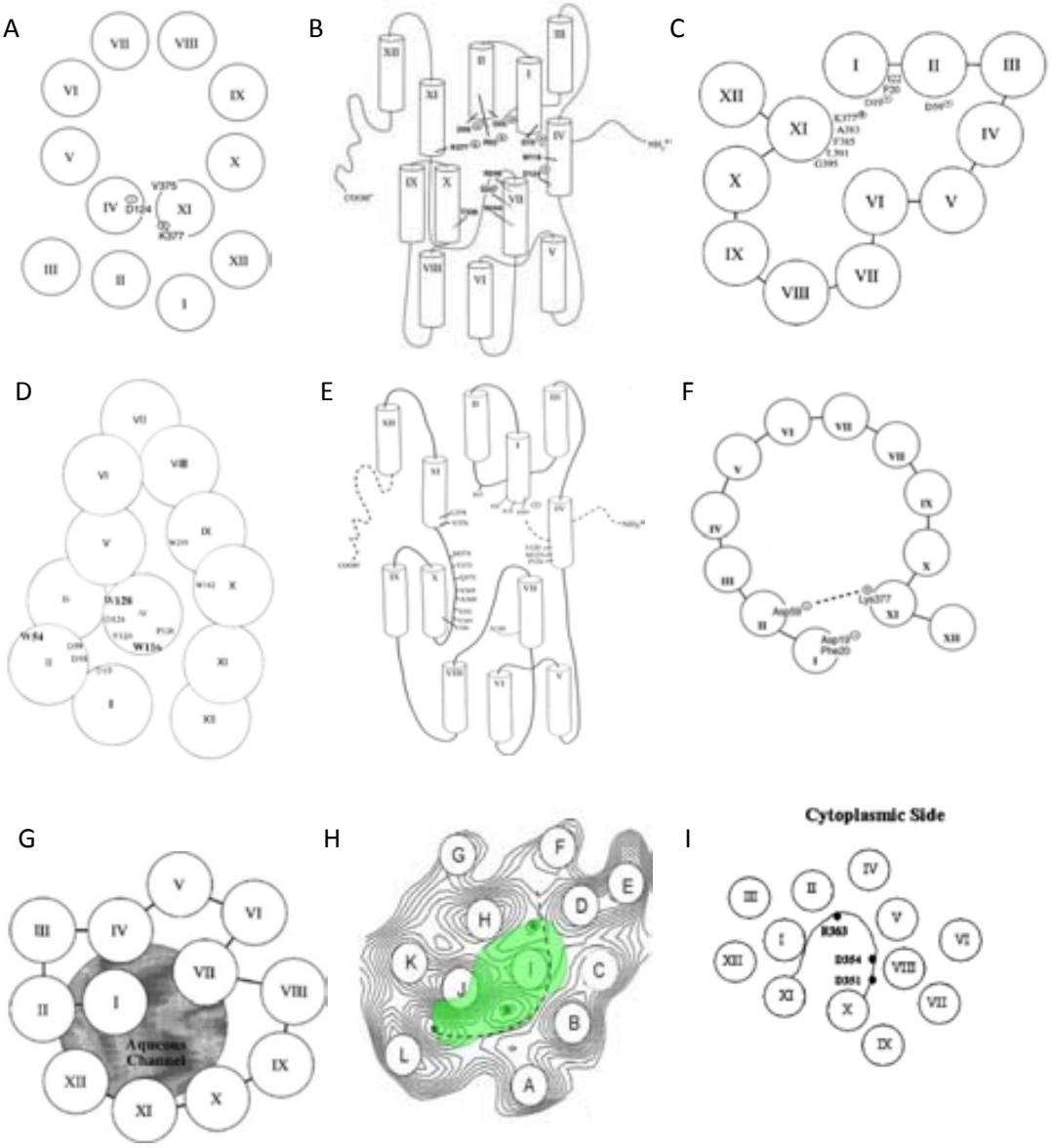



**Figure 1.9** Evolution over time of the proposed distribution of MelB transmembrane helices. According to (A) Hastings Wilson and Wilson 1998. (B) Franco and Wilson 1999. (C) Ding and Wilson 2000b. (D) Cordat et al. 2000. (E) Ding et al. 2000.(F) Franco et al. 2001. (G) Ding and Wilson 2001. (H) Hacksell et al. 2002. (I) Ding 2003.

iii) 3D model.

In recent years, a 3D structure model of MelB threaded through a crystal structure of the lactose permease of E. coli, manually adjusted, and energetically minimized was obtained (Figure 1.10A)(Yousef and Guan 2009). The model contains 442 consecutive

residues (~94% of the polypeptide), includingall 12 transmembrane helices and connecting loops, with no steric clashes and superimposes well with the template structure. The electrostatic surface potential calculated from the model is typical for a membrane protein and exhibits a characteristic ring of positive charges around the periphery of the cytoplasmic side. The 3D model indicates that MelB consists of two pseudosymmetrical two domains with 6 α-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 (Figure 1.10B). The model is consistent with numerous previous mutational, biochemical/biophysical characterizations as well as low-resolution structural data.In the absence of a high-resolution crystal structure of MelB, this theoretical model will help to explain some biochemical and biophysical experimental data (Granell et al. 2010; Ganea et al. 2011),and may play a role for guiding researchers in the design of new experiments(Guan et al. 2012).

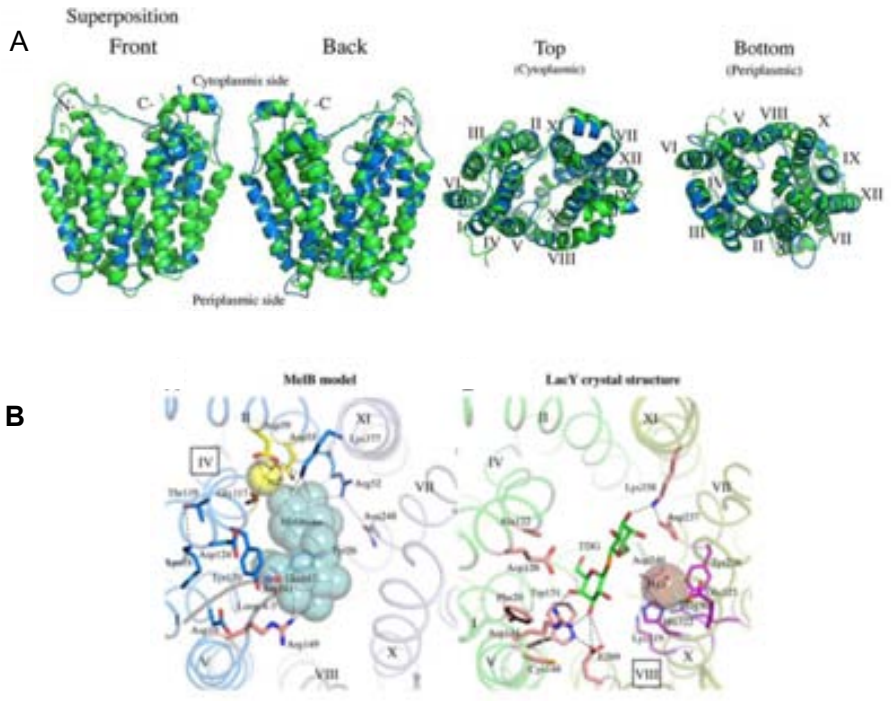



**Figure 1.10** 3D model of melibiose permease. (A) Comparison between the threading model of MelB and the crystal structure of LacY. (B) The central cavity in the MelB model containing putative binding sites for $Na^+$ and sugar. (According to Yousef and Guan 2009)

### 1.4.3 Structural studies of the melibiose permease by IR spectroscopy

Unlike X-ray crystallography, whichis limited by the difficulty of obtaining high-quality crystals, and multidimensional nuclear magnetic resonance (NMR) spectroscopy, which is limited by the molecular weight of the protein, Fourier transform infrared (FTIR) spectroscopy is a well-established technique for the examination of protein secondary structure and structural changes (Surewicz et al. 1993; Goormaghtigh et al. 1994a).In recent years, this technology is getting popularity due to its ease of application to determine the secondary structure of peptides, proteins and also membrane proteins under different physiological states. Estimation of protein secondary structure components and their relative proportion in the overall structure can be derived from an analysis of the protein absorption due to carbonyl stretching vibration of the peptide backbone (appearing in the amide I band), using resolution enhancement and spectral decomposition techniques (Byler and Susi 1986; Surewicz and Mantsch 1988; Fabian et al. 1992; Arrondo et al. 1993; Jackson and Mantsch 1995). The ability to probe structural changes at a molecular level as a function of light, heat, pH or other external factors using difference spectroscopy is making a detailed understanding of the mechanism of action of different proteins possible (Haris and Chapman 1992; Simon-Vazquez et al. 2009; Furutani et al. 2011). Moreover, the use of isotopically labeled molecules helps in both the assignment of spectra and the attainment of greater specificity, and provides anexcellent approach for the study of protein-protein interactions (Haris and Chapman 1992; Ludlam et al. 1996).

i) Secondary structure components and properties of melibiose permease.

The secondary structure componentsof the melibiose permeasefrom *E.coli* had been investigated by Fourier transform infrared spectroscopy, using the purified transporter either in the solubilized state or reconstituted in *E. coli* lipids (Dave et al. 2000). The spectra suggest that the secondary structure of melibiose permease is dominated by α-helical components (up to 50%) and contains β-structure (20%) and additional components assigned to turns, $3_{10}$ helix, and unordered structures (30%) (See Table 1.2). Two distinct and strong absorption bands are recorded at 1660 and 1653 $cm^{-1}$ which is in the usual range of absorption of helices of membrane proteins (Figure 1.11). Furthermore, the secondary structure will be changed by the binding of substrates and ~80% of the backbone amide protons can be exchanged suggesting high conformational flexibilityand/or a large accessibility of the membrane domains to the aqueous solvent.

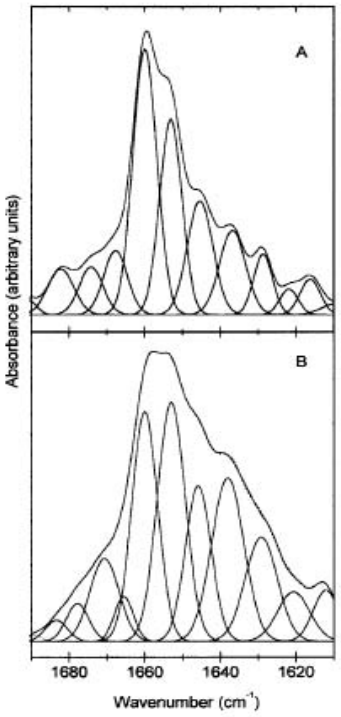



**Figure 1.11** Amide I deconvoluted spectra and the best-fitted componentbands. (A) Reconstituted MelB in $H_2O$ buffer. (B) Reconstituted MelB in$D_2O$ buffer. (According to Dave et al. 2000)

ii) Cation-dependent conformational changes of melibiose permease.

Information about substrate-dependent conformational changes of MelB reconstituted into liposomes was obtained by Fourier transform infrared difference spectroscopy (Leon et al. 2006). The $Na^+$-dependent infrared difference spectroscopy (Figure 1.12A) given rise by the binding of the cosubstrate $Na^+$suggested its binding leads to a series of conformational changesin all types of secondary structures (α-helices, β-sheets, loops), as well as in protonation/deprotonation or in environment of carboxylicgroups (Table 1.3). These conformational changes suggest the presence of a structural

adjustment,which should be need for increasing the affinity for sugar (Abdel-Dayem et al. 2003).

iii) Sugar-dependent conformational changes of melibiose permease.

The difference spectra obtained by subtracting the spectrum cation·MelB from the spectrum complex cation·melibiose·MelB indicated that the binding of melibiose and the ensuing translocation induced a series of conformational changes in all types of secondary structures, includingα-helices, β-sheets, loops, as well aschanges in the protonation state and/or in the environment of given carboxylic residues (Figure 1.12B, Table 1.4), suggesting that a structural adjustment is needed to respond to the binding of sugar and the subsequent translocation.

**Table 1.2** Secondary structure composition and assignments of *E. coli* melibiose permease(according to Dave et al. 2000)

| $H_2O$ | | | $D_2O$ | | |
|---|---|---|---|---|---|
| Wavenumber ($cm^{-1}$) | % area | Assignment | Wavenumber ($cm^{-1}$) | % area | Assignment |
| 1683 | 17 | Rev. turns | 1683 | 16 | Rev. turns |
| 1676 | | | 1678 | | |
| 1669 | | | 1671 | | |
| | | | 1665 | | |
| 1660 | 49 | α, unordered | 1660 | 42 | α |
| 1653 | | | 1653 | 49 | |
| 1647 | 12 | $3_{10}$, open loops, α | 1646 | 13 | unordered |
| 1640 | | | 1638 | 29 | β-sheets,$3_{10}$, open loops, |
| 1634 | 20 | β-sheets | 1629 | | |
| 1628 | | | | | |

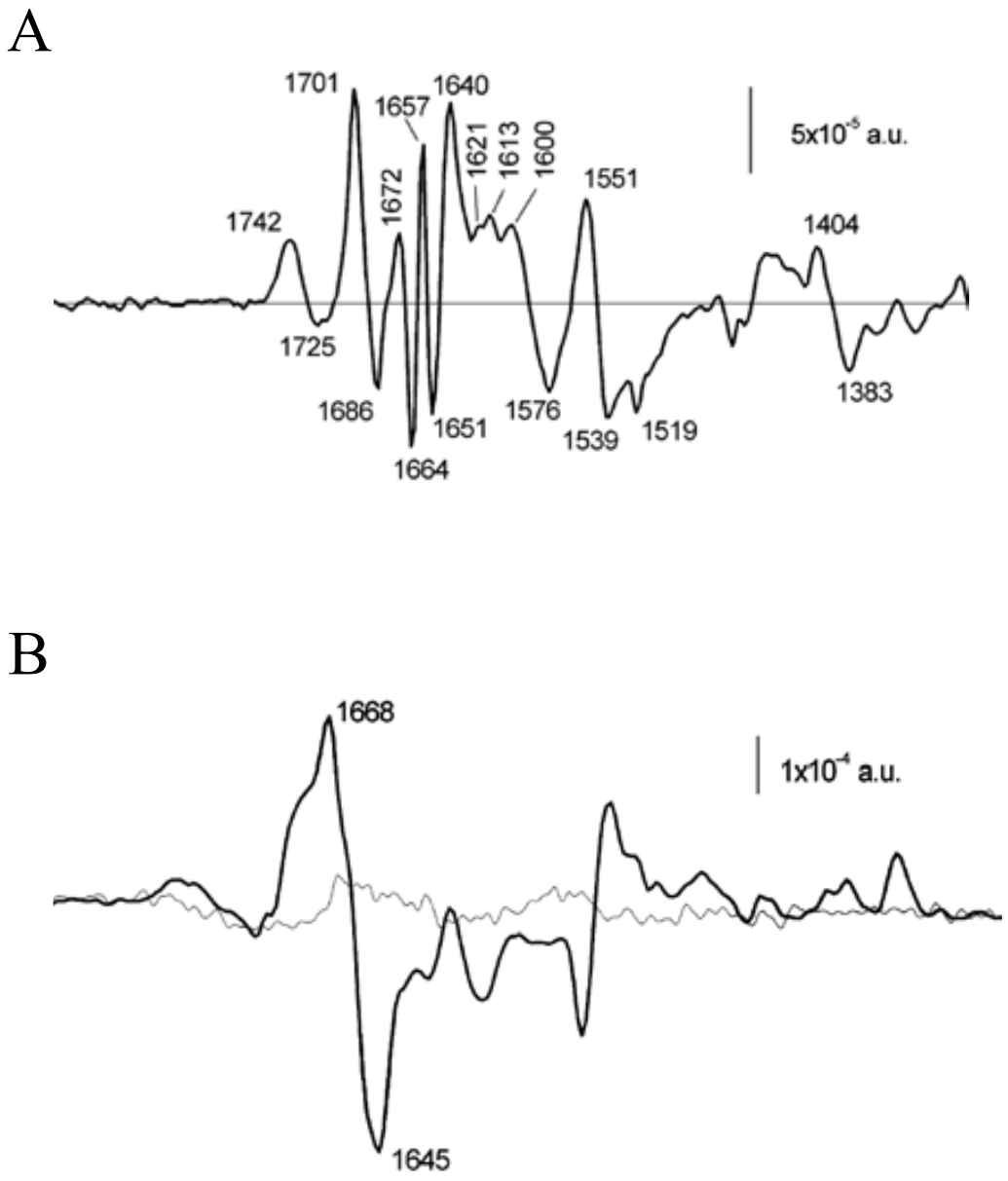



**Figure 1.12** Substrates-dependent secondary structural changes of melibiose permease determined by IR different spectra (A) $Na^+$-dependent IR different spectrum. (B) Melibiose-dependent IR difference spectrum in the presence of $Na^+$. Thin line: saccharose induced IR different spectrum. (According to Leon et al. 2005)

Table 1.3 Tentative assignment of the infrared peaks induced by $Na^+$ binding to MelB (Leon et al. 2006)

| $H_2O$ | $D_2O$ | Assignment |
|---|---|---|
| 1742 | 1741 | Lipid |
| 1725 | 1716 | COOH (with a possible lipid contribution) |
| 1700 | 1690 | Asp/Glu/solvent-accessible reverse turn |
| 1687 | 1687 | β-sheet |
| — | 1678 | Solvent-accessible turn-like structure |
| 1671 | 1670 | Turn/aII-helix |
| 1662 | 1662 | α-helix |
| 1657 | 1657 | α-helix |
| 1651 | 1651 | α-helix |
| — | 1646 | Open loops |
| 1641 | 1639 | β-sheet |
| 1599 | 1599 | $COO^-$ (antisymmetric stretching) |
| 1576 | 1576 | $COO^-$ (antisymmetric stretching) |
| 1551 | 1450 | Amide II vibration |
| 1540 | 1540 | Amide II vibration |
| 1403 | 1403 | $COO^-$ (symmetric stretching) |
| 1384 | 1385 | $COO^-$ (symmetric stretching) |

Table 1.4 Tentative assignments of the infrared peaks induced by melibiose binding to MelB, in the presence of $Na^+$ (Leon et al. 2006)

| $H_2O$ | $D_2O$ | Assignment |
|---|---|---|
| 1703 | 1693 | Asp/Glu/turn |
| 1703 | — | Asn |
| 1693 | 1693 | β-sheet |
| 1693 | 1683 | Solvent-accessible turn |
| 1688 | 1686 | Turn/β-sheet |
| 1680 | 1681 | β-sheet |
| 1674 | 1674 | β-sheet |
| 1668 | 1666 | α-helix/turn |
| 1659 | 1659 | α-helix |
| 1652 | 1653 | α-helix |
| 1643 | 1645 | β-sheet/310-helix/open loops |
| 1631 | 1630 | β-sheet |
| 1620 | 1616 | Secondary structure/aromatic side chain |
| 1611 | 1609 | Secondary structure/aromatic side chain |
| 1595 | 1591 | $COO^-$ (antisymmetric stretching) |
| — | 1567 | $COO^-$ (antisymmetric stretching) |
| 1518/1513 | 1520/1515 | Tyr |
| 1404 | 1402 | $COO^-$ (symmetric stretching) |
| 1390 | 1389 | $COO^-$ (symmetric stretching) |

1.4.4 Structural studies of the melibiose permease by fluorescence spectroscopy.

The effects of sugars and coupling cations ($H^+$, $Na^+$, or $Li^+$) on the conformational properties of melibiose permease can be investigated by tryptophan fluorescence spectroscopy (Mus-Veteau et al. 1995). The melibiose permease emission fluorescence is selectively enhanced by sugar binding. Moreover, the sugar-dependent fluorescence increase is specifically potentiated by cation (Figure 1.13). On the other hand, cation binding produces a small quenching of the fluorescence signal (Figure 1.13).

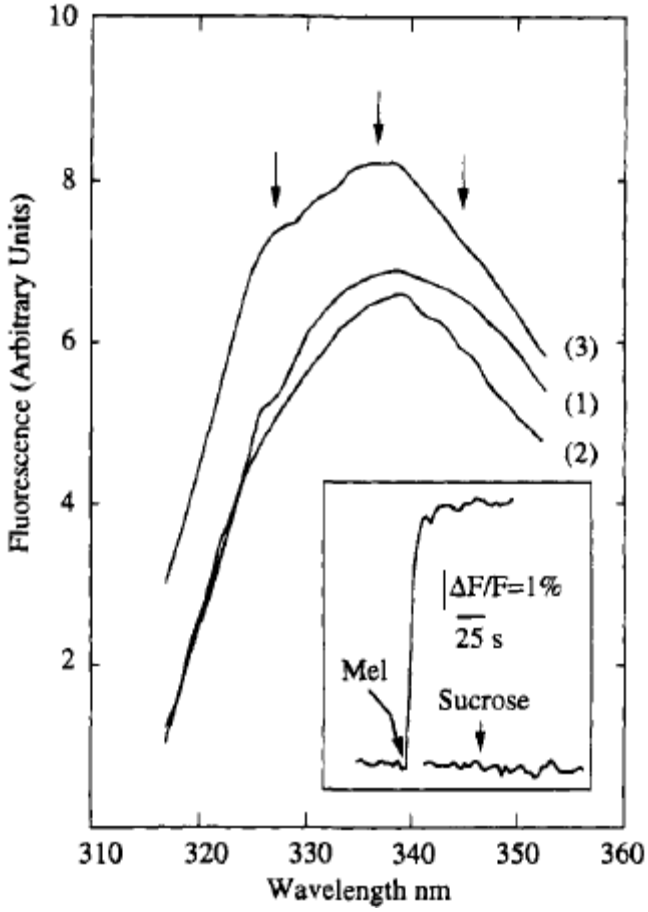



**Figure 1.13** Effects of sodium and melibiose on the intrinsic fluorescence of the proteoliposomes containing melibiose permease. Spectrum 1: fluorescence emission spectrum of melibiose permease obtained by illuminating at 297 nm; specturm 1: fluorescence emission spectrum recorded in the presence of NaCl; spectrum 3: fluorescence emission spectrum recorded in the presence of NaCl and melibiose. Inset: Effect of the addition of sugar on the integrated fluorescence signal of melibiose permease. (According to Mus-Veteau et al. 1995)

Further insight into the cosubstrate-induced structural changes of MelB has been obtainedby a means of Förster resonance energy transfer (FRET) from Trp to a fluorescent sugar analog, 2'-(N-dansyl)-aminoethyl-1-thio-D-galactopyranoside ($D^2G$), which can bind to MelB, but cannot be transported (Maehrel et al. 1998). In the presence of $D^2G$, sample excitation at 297 nm gives rise to a fluorescent signal at around 465 nm, indicating that a tryptophan-mediated fluorescence resonance energy transfer phenomenon is involved in the response (Figure 1.14). The signal of $D^2G$ bound to MelB recorded in sodium-free medium is red shifted by up to 25 nm from that recorded in the presence of NaCl, indicating the presence of a highly hydrophobic environment close to or at the sugar-binding site, and that the interaction of sodium ions with MelB enhances the hydrophobicity of this environment (Figure 1.15).

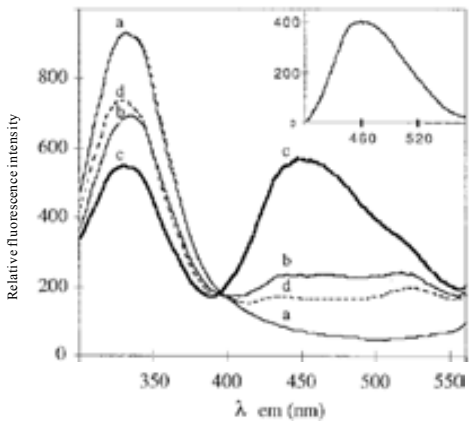



**Figure 1.14** FRET phenomenon between MelB Trp and $D^2G$. (a) MelB reconstituted into proteoliposomes was excited at 297 nm; (b) the same as (a) after the addition of 15 mM $D^2G$; (c) the same as (a) after the addition of 15 mM $D^2G$ and 10 mM $Na^+$; (d) the same as (a) after the addition of 15 mM $D^2G$, 10 mM $Na^+$ and 100 mM melibiose. (According to Maehrel et al. 1998)

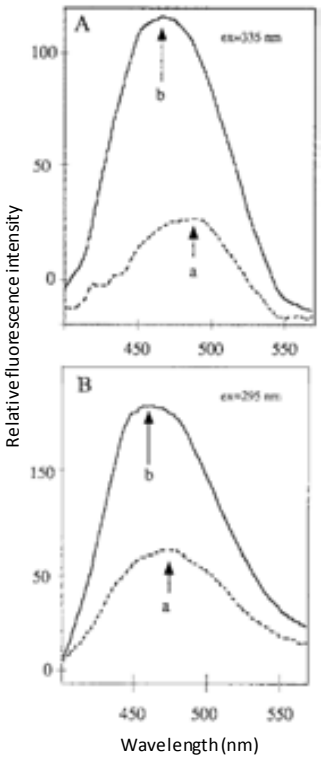



**Figure 1.15** Effect of NaCl on the spectral properties of $D^2G$ bound to MelB. Spectrum a: in the absence of NaCl; spectrum b: after the addition of NaCl. (A) Spectra were obtained by illuminating at 335 nm. (B) Spectra were obtained by illuminating at 295 nm. (According to Maehrel et al. 1998)

These results obtained from fluorescence spectroscopy suggested that MelB was subjected to conformational changes induced by the substrates binding, confirmingthe results obtained from IR spectroscopy.

1.4.5 Towardthe identification of the substrates binding sites of melibiose permease

Although the structure at atomic resolution of melibiose permease is not obtained yet, many efforts towards the identification of substrates binding sites of MelB has been done by combining site-directed mutagenesis with biochemical and biophysical methods.

i) The cation binding sites of melibiose permease

Site-directed mutagenesis studies on several acidic and polar residues located in the N-terminal hydrophobic domains of the permease suggest that three aspartic acid residues, i.e., Asp55, Asp59, and Asp124, distributed on helices I1 and IV, act as a network involved in coordination of the coupling cation(Pourcher et al. 1991; Wilson and Wilson 1992; Pourcher et al. 1993; Zani et al. 1993; Wilson and Wilson 1994; Zani et al. 1994). Construction of chimeric melibiose transporters from various microorganisms (Hama and Wilson 1993)further emphasizes the importance of the permease N-terminal domains for cation recognition. The most important results were summarized in the light of the

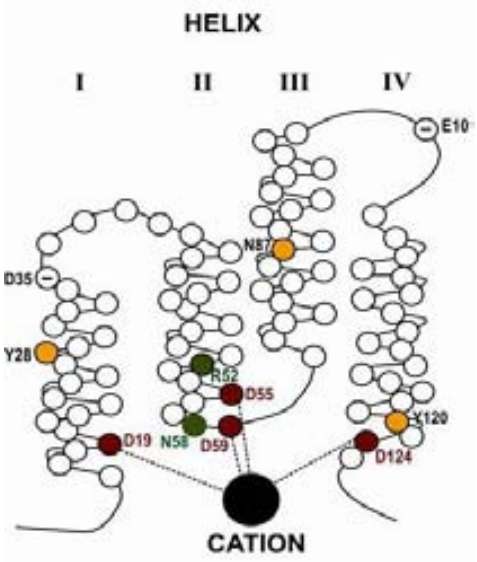



**Figure 1-16** Putative cation binding site. Asp19 (helix I), Asp55 (helix II), Asp59 (helix II), and Asp124 (helix II) were suggested to form a cation coordination site (in red). Asn-58 and Arg-52 are important for $Na^+$ recognition (in green). Polar amino acids suggested to be important for the overall structure are colored yellow.Taken from Poolman et al. 1996.

hypothetical structural model of the amino-terminal region of MelB (Figure 1.16, for mini review see Poolman et al. 1996). In this tertiary structure model, neutral residues at MelB positions 19 (helix I), 55 and 59 (helix II) and 124 (helix IV) were proposed to be involved in the cation binding site. A recent study suggested that Asp55 and Asp59 are essential ligands for $Na^+$ binding, Asp124 is not essential for $Na^+$ binding, but it is affected after $Na^+$ and sugar binding, whereas Asp19 is not involved in $Na^+$ binding(Granell et al. 2010).

ii) Residues involved in sugar binding sites of melbiose permease.

As a cation/sugar cotransporter, the binding of $Na^+$ clearly affect the binding of melibiose by means of adjustment of the spatial structure to enhancing the affinity for sugar (Abdel-Dayem et al. 2003; Leon et al. 2005). In the 3D structures the binding site of the first substrate is always close to the binding site of the second substrate(Yamashita et al. 2005; Faham et al. 2008; Weyand et al. 2008), hence individual point mutation of MelB often results in more than one change in the functional properties of the transporter. In general, the binding of substrates require a special environment, e.g., polarity, charge, hydrophobicity, etc. We may call it “binding pocket”. Several residues from different helices would be involved in the formation of this pocket. Clearly, a single residue mutation might lead to changes in the environment of the binding pocket or its instability. In addition, based on the alternating-access mechanism, which has been widely adopted to explain how membrane transport protein carries substrates pass-through the membrane, the binding of melibiose will drive the protein to reorientateits conformation for translocation. Therefore, it is difficult to get a hypothetical model for the sugar binding site only from site-directed mutagenesis studies. Amino acids or regions involved in sugar binding or lying close to the sugar binding site have, however, been identified.

In MelB, Arg-149 (helix V) is important for sugar recognition (Abdel-Dayem et al. 2003), and the S153Cmutant (helix V) shows a strong decrease for α-NPG affinity. Therefore Ser153 seems to board or be close to the sugar binding site (Kim et al. 2001). Asp-124 and Tyr-120 (helix IV) were also suggested to be involved in both sugar

and/or cation binding (Zani et al. 1994). Asp19 was shown to be involved in melibiose binding (Granell et al. 2010), but did not affect $Na^+$ binding. The cysteine-scanning mutagenesis of helix I showed that Gly17, Lys18, Asp19, Tyr32, Thr34, and Asp35 are important residues for keeping native transport activity (Ding and Wilson 2001a). Cysteine substitution studies for helix XI suggested that seven of the residues (Lys377, Gly379, Ala383, Phe385, Leu391, Gly395 and Tyr396) are important residues for keeping transport activity of MelB (Ding and Wilson 2001b). Evidences showed that helix IV plays a role in connecting cation- and sugar-binding of the melibiose permease. In addition, cytoplasmic loops IV/V and X/XI are also involved in the sugar binding and/or translocation (Abdel-Dayem et al. 2003; Ding2003).

In summary, the charged/H-bond network observed between helices IV, I, and V, loop IV/V, loop X/XI as well as in helices II and XI could account for the cooperative binding of the cation and sugar in MelB (Yousef and Guan 2009).

## 1.5 Substrates transport of melibiose permease

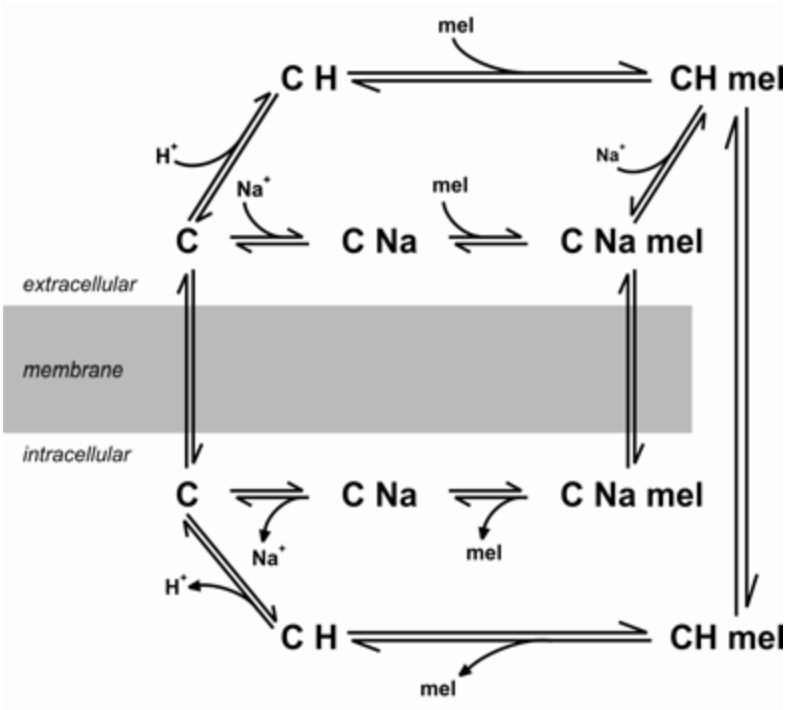
### 1.5.1 Kinetic transport model of melibiose permease

A wealth of kinetic, biochemical, and biophysical studies as well as thermodynamic data (Meyer-Lipp et al. 2004; Meyer-Lipp et al. 2006) suggest that sugar translocation occurs via the alternating-access mechanism, which means that a single polar binding site is accessible to only one side of the membrane at a time: i.e. the existence of the so-called outward and an inward conformation states or states 1 and 6 in Figure 1.17A, respectively. This overly crude model has been extended over the years with the introduction of several intermediates as they proceed through the transport cycle resulting in the 3'-step mechanism shown in Figure 1.17B, corresponding to an occluded transporter. Another empty occluded transporter state is supposed to be present in the cycle between steps 1 and 6.

### 1.5.2 The symport mechanism for melibiose permease

Combining kinetic transport model and alternating-access mechanism, Yousef and Guan postulated the symport mechanism for the efflux mode for melibiose permease (Figure 1.18, Yousef and Guan 2009), which can be explained by a simplified 6-step scheme similar to that proposed for lactose permease (Kim and Barry 2001). In this mechanism, the empty MelB presents a conformational dynamic behavior between inward-facing and outward-facing orientations (step 6). The subsequent binding of cation (step 1) and sugar (step 2) will induce the conformational changes from outward-facing to inward-facing (step 3), then the substrates binding sites are faced to cytoplasmic medium. Sugar (step 4) and cation (step 5) will subsequent leave MelB, and then the empty MelB returns to the initial state (step 6). The turnover between the inward-facing conformation and the outward-facing conformation may be expected to occur by two transition states, i.e., two occluded states, in which cosubstrate binding sites are concurrently occupied or unoccupied (Granell et al. 2010).

A




B

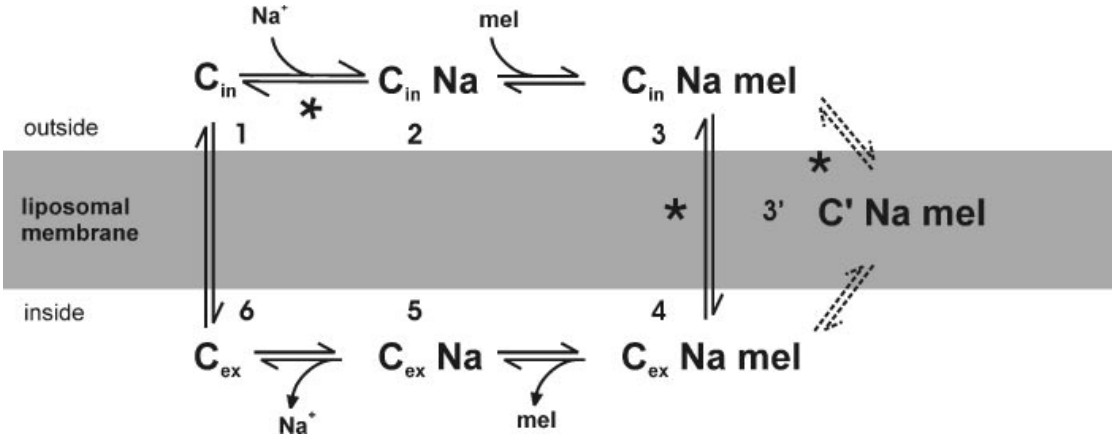



**Figure 1.17** Kinetic transport model of melibiose permease. (A) Kinetic model of MelB. The transporter C, facing the extracellular space, binds $Na^+$ or $H^+$ and sequentially the sugar melibiose. The ternary complex (C-ion-sugar) reorients its binding sites to the intracellular site, releases first the sugar and then the ion. Finally, the empty carrier turns back to the extracellular side. An arrow from C H mel to C Na mel indicates that melibiose can also enhance the carriers' affinity for $Na^+$ (After Meyer-Lipp et al. 2005). (B) An additional intermediate state (3' C' Na mel) was suggested by the study of the R141C mutant (After Meyer-Lipp et al. 2006).

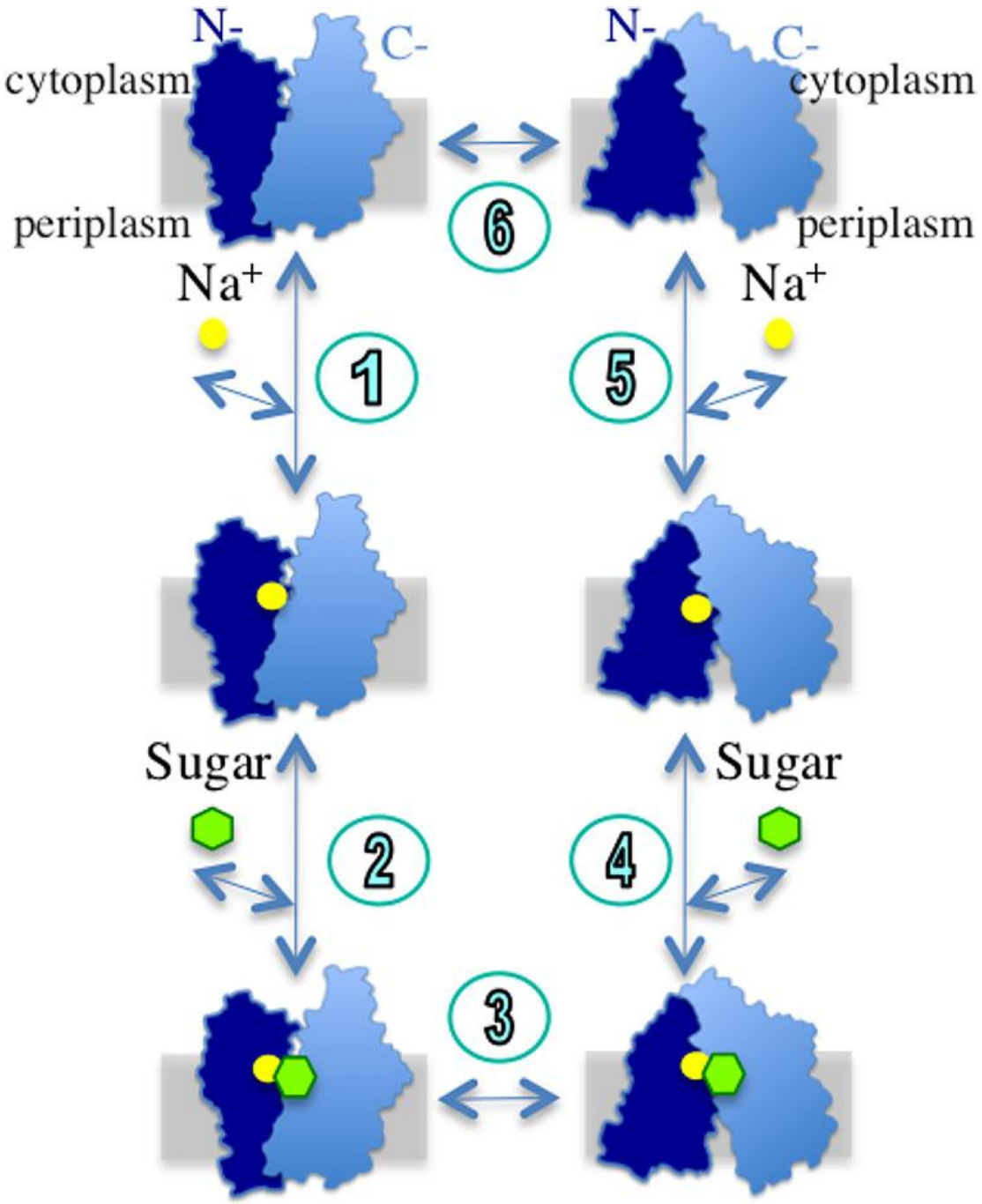



**Figure 1.18** A kinetic scheme of the efflux mode of galactoside/$Na^+$ symport for MelB. A cross section of the membrane is shown as a gray rectangle. (After Yousef and Guan 2009)

## 1.6 Membrane protein crystallization and structure determination

### 1.6.1 X-ray crystallography is the main method for detecting the structure of biological macromolecules

The structure of a protein can be described at different levels, i.e., primary structure (amino acid sequence), secondary structure (α-helix, β-sheet, turn, etc.), tertiary structure, and quaternary structure. The tertiary structure of a protein describes the folding of its secondary structural elements and specifies the positions of each atom in the protein, including those of its side chains. The quaternary structure of a protein is the arrangement of multiple folded protein molecules in a multi-subunit complex. As proteins are performing their function in their 3D form, the three-dimensional structure of the protein in the level of atomic resolution, as well as the structure of the protein and the formation of complexes with their interacting molecules is essential to explain the specific mechanism of the protein. When we know the 3D structure of the protein, we can begin to understand how the protein works. The known protein structures have come to light through X-ray crystallographic, nuclear magnetic resonance (NMR), or Cryo-electron microscopy (cryo-EM). The atomic coordinates of most of these structures are deposited in a database known as the Protein Data Bank (PDB). These data are readily available via the internet (http://www.pbd.bnl.gov), which allows the tertiary structures of a variety of proteins to be analyzed and compared.

X-ray crystallography is one of the most powerful methods for studying macromolecular structures. X-ray crystallography is a method of determining the arrangement of atoms within a crystal, in which a beam of X-rays strikes a crystal and causes the beam of light to spread into many specific directions. From the angles and intensities of these diffracted beams, a crystallographer can calculate a three-dimensional picture of the density of electrons within the crystal. From this electron density, the mean positions of the atoms in the crystal can be determined, as

well as their chemical bonds, their disorder and various other information. Since 1960, more than 80,000 atomic level resolution structures have been obtained by X-ray crystallography. With the quick development of the high-throughput protein expression, purification, crystal screening, and synchrotron radiation, the development of structural biology is very fast (Figure 1.19). Although protein crystallization is always the bottleneck of crystallography, x-ray diffraction method is the most important and most accurate method for the determination of protein three-dimensional structure (Figure 1.19).

In 1971, Jean Jeener, a Belgian physical chemist and physicist, first proposed the concept of 2D NMR. Richard R. Ernst and his colleagues developed a series of homonuclear 2D NMR methods. After 20 years of exploration and research, the multi-dimensional nuclear magnetic resonance spectroscopy technique had been developed into the only effective means of determining the three-dimensional structure of biomolecules such as proteins and nucleic acid in solution. Since 1980s $^{15}N$ and $^{13}C$ stable isotope labeling of protein samples were prepared, which leaded the two-dimensional heteronuclear NMR method to be implemented, this method is widely used in the study of the structure and function of proteins and other biological molecules. In the early 1990s, Ad Bax from United States and others proposed heteronuclear 3D and 4D NMR methods. In recent years, the rapid development of heteronuclear multi-dimensional NMR methods made it possible to determine the three-dimensional structure of a protein whose molecular weight is up to 15 ~ 25kDa, in solution, with an accuracy of up to 2Å resolution. In recent years, the multi-dimensional NMR determination, as the most important means of three-dimensional structure of the protein in solution, has achieved a rapid development. Today, the multi-dimensional NMR techniques and the X-ray crystallography technology, complementing each other, take an important position in structural molecular biology. However, for these methods the maximum size of a protein cannot break through 35kDa. As data collection time is too long (45-60 days), especially in data analysis, which may need 6-12 months,

high-throughput structure determination using NMR is also a great challenge (Figure 1.19).

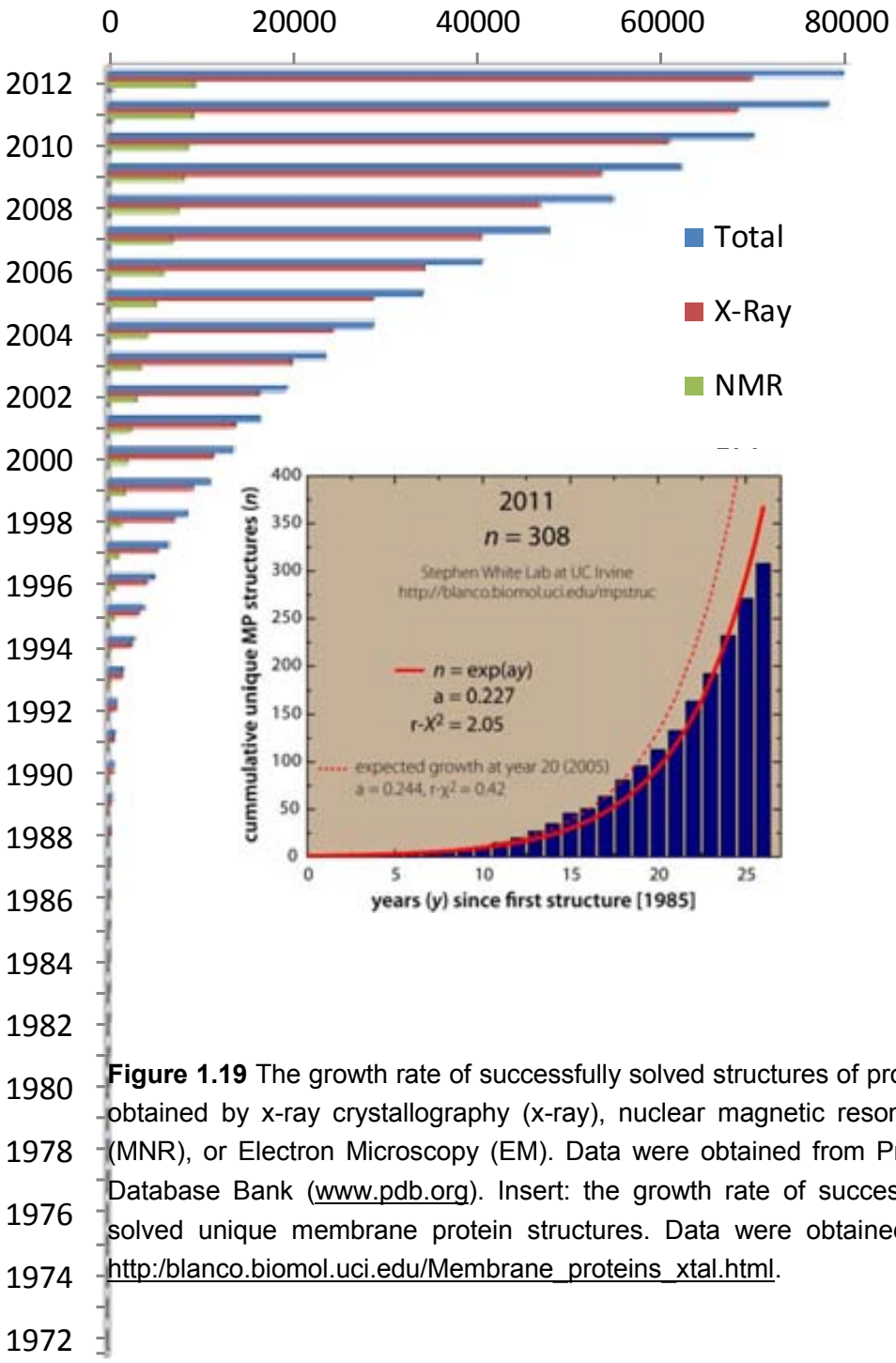



**Figure 1.19** The growth rate of successfully solved structures of proteins obtained by x-ray crystallography (x-ray), nuclear magnetic resonance (MNR), or Electron Microscopy (EM). Data were obtained from Protein Database Bank (www.pdb.org). Insert: the growth rate of successfully solved unique membrane protein structures. Data were obtainedfrom http:/blanco.biomol.uci.edu/Membrane_proteins_xtal.html.

Cryo-electron microscopy (cryo-EM) was developed from the early electron diffraction analysis. Cryo-electron microscopy (cryo-EM) is a form of transmission electron microscopy (EM) where the sample is studied at cryogenic temperatures

(generally liquid nitrogen temperatures). Cryo-EM is gaining popularity in structural biology. The popularity of cryoelectron microscopy stems from the fact that it allows the observation of specimens that have not been stained or fixed in any way, showing them in their native environment, in contrast to X-ray crystallography, which generally requires placing the samples in non-physiological environments, which can occasionally lead to functionally irrelevant conformational changes. In practice, the resolution of cryo-electron microscopy maps is not high enough to allow for unambiguous model construction on the basis of EM maps only, and models obtained by protein crystallography are used to interpret the cryo-EM maps. However, the resolution of cryo-EM maps is improving steadily, and some virus structures obtained by cryo-EM are already at a resolution that can be interpreted in terms of an atomic model. Figure 1.19 compares the amount of the atomic level structures obtaining from different methods. It is clear that more than 90% structures were obtained by X-ray crystallography method.

### 1.6.2 Difficulty to work with membrane proteins

Unlike soluble proteins, membrane proteins are embedded into or intimately associated with the lipid bilayer, which displays a hydrophilic feature on external surfaces, but a hydrophobic feature in the bilayer interior as it is composed of aliphatic chains. Accordingly, membrane proteins embedded into the lipid bilayer have hydrophobic surfaces in association with the lipids and, typically, hydrophilic portions protruding from the membrane surface. Such integral membrane proteins (IMPs) cannot directly be solubilized in aqueous media but require detergents to protect the hydrophobic surfaces against the water when extracted (Garavito et al. 1996). These properties make working with IMPs significantly more complex than working with soluble proteins.

As membrane proteins are strongly hydrophobic, they were considered to be insoluble. 50 years ago, it was thought to be impossible to get crystals of membrane proteins, and this had been included in authoritative textbooks. The breakthrough was the first result in three dimensions of bacteriorhodopsin at 7 Å resolution obtained by

electron microscopy (Henderson and Unwin 1975) and the first atomic-level structure of a membrane protein complex of the photosynthetic reaction center from *Rhodopseudomonas viridis* at 3 Å resolution by X-ray crystallography (Deisenhofer et al. 1984). Taking advantage of the quick development of new technologies, especially synchrotron radiation light sources, which makes it possible to test very small crystals, the pace of the structure determination of membrane proteins is accelerating in recent years, and many membrane proteins have been solved, which render a new scene for this type of proteins. Many of these new structures have had a dramatic impact as in the cases of cytochrome c oxidases (Iwata et al. 1995; Tsukihara et al. 1996), potassium channels (Doyle et al. 1998; Jiang et al. 2003), aquaporins (Murata et al. 2000; Savage et al. 2003), G-protein coupled receptors (for review see Hanson and Stevens 2009), and neurotransmitter transporter (Singh et al. 2008; Krishnamurthy and Gouaux 2012).Nevertheless, the structural output on membrane proteins is a very small fraction of that for soluble macromolecules. Since 1985, when the first atomic-level membrane protein structure was reported, only 308 unique membrane protein structures have been obtained (http://blanco.biomol.uci.edu/Membrane_Proteins_xtal.htmL), whereas there were over 80,000 PDB entries determined by diffraction methods at the same time for soluble proteins (Figure 1.19). Thought membrane proteins are very important in our bodies, and comprise 20–30% of all proteins in both prokaryotic and eukaryotic organisms (Wallin and von Heijne 1998), they comprise less than 0.5 % of all the known structures. For higher eukaryotes the story is even starker, with only 20 structures of human IMPs and less than 50 mammalian IMPs solved. Membrane proteins are therefore one of the most important remaining frontiers of structural biology research. Due to the importance and the difficulties in the studies of membrane proteins，so far, five scientists (Johann Deisenhofer, Robert Huber, Hartmut Michel, 1988; and Peter Agre, Roderick MacKinnon, 2003) won the Nobel Prize due to their outstanding contributions to the detection of membrane protein structure.

Determining membrane protein structure is still a challenge for structural biochemists that arises at almost every stage from protein preparation to structural analysis (for a

review, see Lacapereet al. 2007). Only the initial cloning for the recombinant expression may be expected to offer no more difficulty than for soluble proteins(Love et al. 2010). However, membrane proteins was thought to be toxic for its expression in bacteria, which leads cells growing slowly, and then giving less yield. Due to its strong hydrophobic character, the over-production of proteins in *E. coli* often leads to inclusion body formation. Although in recent years, recombinant bacterial membrane proteins have been well expressed in bacterial expression systems, obtaining an adequate amount of eukaryotic membrane proteins for structural analysis is still a bottleneck. Even though there are structures of important eukaryotic membrane proteins, all but a few have come from natural sources (Love et al. 2010).

In order to understand how a membrane protein works, one has to isolate it from the plasma membrane, either as naturally enriched in the membrane vesicles or reconstituted into lipid liposomes, or in the detergent micelles. Although some atomic-level structures have been obtained by electron crystallography from two-dimensional membrane protein crystals(Raunser and Walz 2009), or solid-state NMR method(McDermott 2009), x-ray crystallography is the principal means of determining membrane protein 3D structure. Membrane proteins solubilized in detergent micelles represent the first step to get crystals to be used for x-ray crystallography. To this end, it is essential to obtain stable and high homogeneous protein. The selection of suitable detergents for extraction and purification of a specific membrane protein is essential in the membrane protein preparation.

Being strongly hydrophobic, membrane proteins can only be present in the form of protein-detergent complexes in aqueous solution. However, the presence of detergents increases the complication of the membrane protein crystallization. In the protein-detergent complex, there are detergent-mediated interactions, which may affect crystallization (Garavito et al. 1996). And the detergent-covered hydrophobic surface varied interaction between two proteins in the crystal and thus unsuitable for lattice contacts(Ostermeier and Michel 1997), which theoretically reduces the probability of crystallization by a high power of the fractional surface area(Kwong et al. 1999).

Therefore, detergents, which play a role in the stability and the aggregation of membrane proteins as well as in the colloidal properties of the protein-detergent complexes, need to be appreciated and controlled before and during the crystallization trials(Garavito et al. 1996). In general, extensive screening of many detergents or mixing detergents, which may help to expose more surfaces to increase protein-protein interaction, is essential for improving resolution limit.

### 1.6.3 General methods used for membrane protein preparation and crystallization

#### 1.6.3.1 General process

A workflow for generating membrane protein crystals is summarized in Figure 1.20. This workflow can be used to design experiments for crystallization of a new membrane protein. Due to the high diversity of the protein properties, some important steps need to be optimized in order to obtain the best results.

#### 1.6.3.2 Cloning, designing, reconstituting, and expressing

Selecting a suitable vector may be directly related to the problem of getting a good expression. In order to obtain the maximum expression it would be necessary to test some vectors. As immobilized metal affinity chromatography (IMAC) is an effective first step in protein purification (Hemdan and Porath 1985a; Hemdan and Porath 1985b), the construct is cloned into a suitable vector with a cleavable poly-histidine tag at N- or C-terminal. In some cases, N-terminal tags should be avoided when there is a signal sequence that directs the transport of a protein because the tag may interfere with proper targeting of the protein or be cleaved during signal sequence processing. The pET system based on the T7 promoter-driven using isopropyl-β-D-thiogalactopyranoside (IPTG) induction is the most powerful system yet

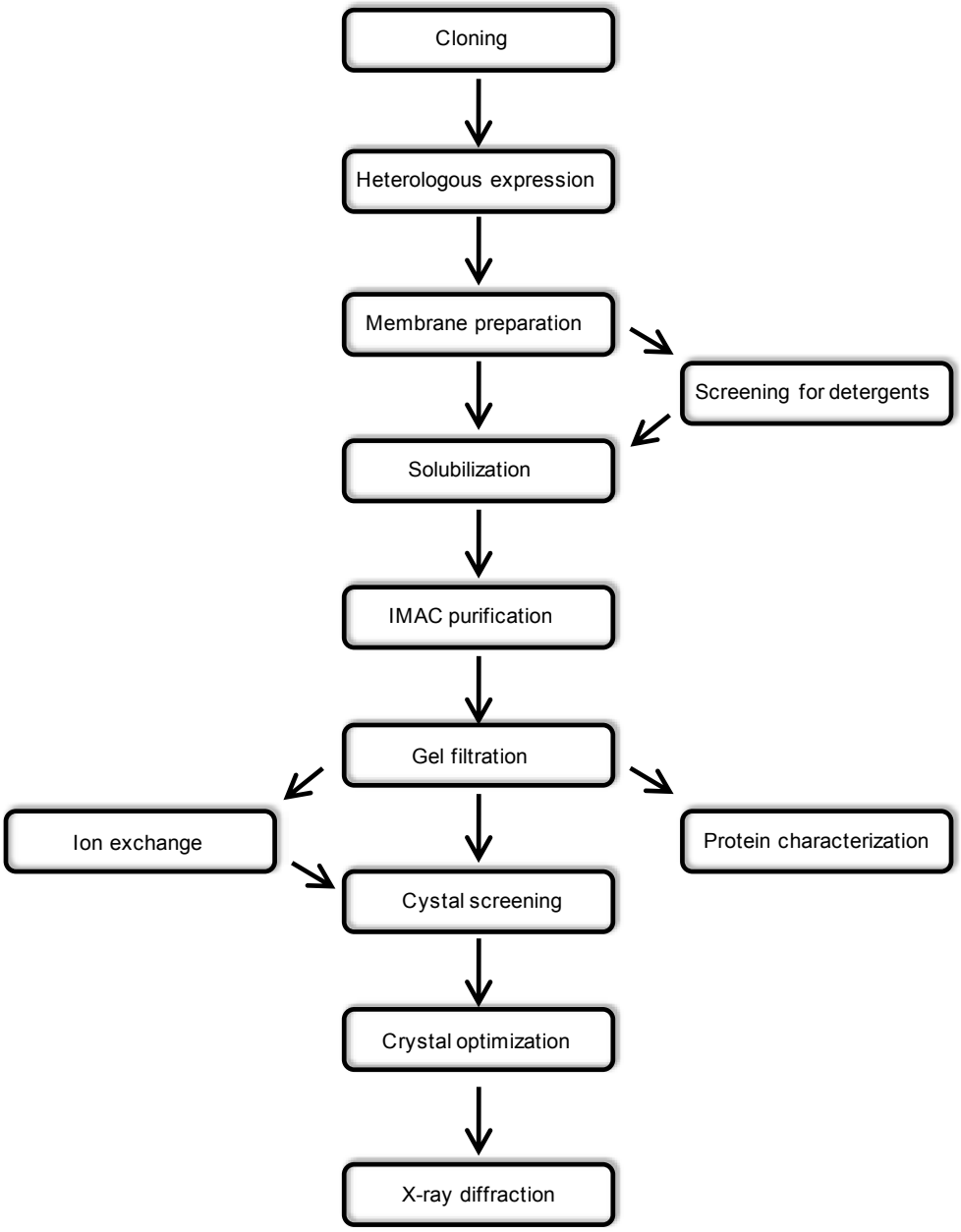



**Figure 1.20** Workflow for generating membrane protein crystals useful for X-ray.

developed for the cloning and expression of recombinant proteins in *E. coli*(Studier et al. 1990). Alternatively, the pBAD vector system containing the arabinose pBAD promoter has been implemented successfully for the production of membrane proteins for X-ray studies (Guzman et al. 1995; Yernool et al. 2004; Zhou et al. 2007). In some cases, the addition of an N-terminal fusion protein may increase the amount of protein

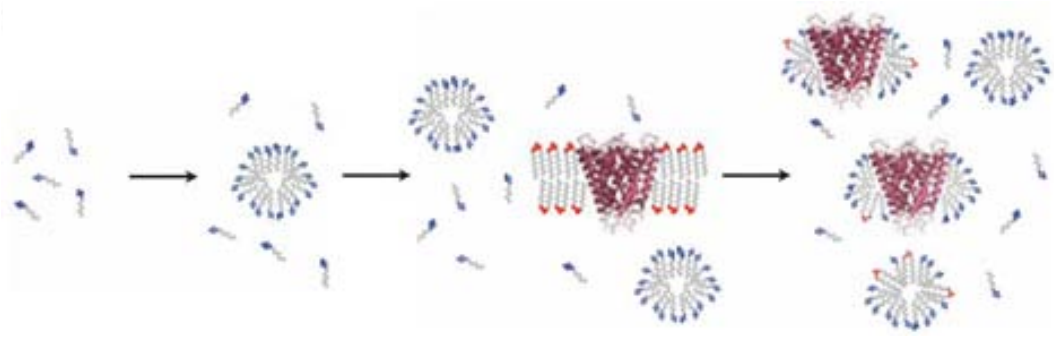
heterologously expressed in the membrane (Hilf and Dutzler 2008). In order to obtain the best quality protein, to try several expression systems are necessary. A new method by tagging green fluorescent protein (GFP) to the C-terminal of membrane proteins facilitates monitoring of both expression and purification of target proteins (Drew et al. 2001; Drew et al. 2005).

Excessive membrane protein expression as well as a non-correctly folded protein will cause an incorrectly-inserted membrane protein, which may affect the expression of endogenous membrane proteins, thus becoming toxic to the host cell. In this case, reducing the speed of the expression of membrane proteins can alleviate this problem. In general, the use of low-temperature expression or reducing the inducer concentration is a good way to decrease the rate of membrane protein expression, thereby improve the quality of the protein.

Cell-free expression system, which does not appear to be restricted by origin, size or topology of the target protein, is a simple and fast way to obtain protein for structural and functional studies. Many commercially available Kits can be used for cell-free protein expression, e.g., MembraneMax™ protein expression Kit (Invitrogen™). In recent years, cell-free expression has emerged as a promising tool for the fast and efficient production of membrane proteins (Schneider et al. 2010). This technology has been used in preparing many important membrane proteins, e.g., G-protein coupled receptors (Haberstock et al. 2012).

1.6.3.3 Membrane preparation, solubilization and purification

Membrane preparation may also be important for obtaining high quality crystals. In some reports, it was pointed out that high-resolution diffraction data was obtained only with well-purified membranes. Clearly, purified membranes containing almost the target protein will facilitate the purification in the next step. Solubilization of the membrane is a critical step for crystallization. Figure 1.21 shows the schematic of the solubilization



**a** **b** **c** **d**

**Figure 1.21** Detergent solubilization of membrane proteins. Schematic of the solubilization process. From left to right: free detergent monomers (a) associate to form detergent micelles (b) at concentrations above the CMC. When added to a membrane preparation (c), the micelles extract membrane proteins from the lipid bilayer yielding a solution containing PDLC complexes, free lipid-detergent micelles and detergent monomers (d). After Newby et al.2009.

process. Extensive screening of a large collection of detergents is essential to obtain stable and highly homogeneous protein required for crystallization. Benefiting from the development of high-technology, this step is performed with a robot in 96-well plates. The quantity of protein can be evaluated by size exclusion chromatography equipped with a small column, which can load samples down to μL (Love et al. 2010). In many cases, DDM (C12-maltoside) is very useful for the solubilization of the membrane. So, it may be a good idea to start from DDM. However, many observations imply that the resolution limit of 8Å cannot be surpassed when using DDM (Ostermeier and Michel 1997; Sonoda et al. 2011). Purification of membrane proteins is not particularly difficult. The Ni-NTA column is an effective first step in protein purification. In many cases, membrane proteins containing a 6-His-tag just need a single pass through Ni-NTA

column to achieve purity higher than 90%, which is suitable for the crystal screening study. If needed, ion exchange chromatography and gel filtration chromatography could be used for further purification. Many observations showed that purity is not the most important parameter for protein crystallization. Some researchers have noted that excessive protein purification may affect subsequent protein crystallization. This may be due to the separation of some native lipids from the membrane protein. These lipids may be important for maintaining the stabilization of the protein.

1.6.3.4 Membrane protein crystallization

Extensive screening for a large set of conditions is essential for obtaining crystals. To this end, a number of commercially available screens, e.g., MemStart, MemSys, MemGold, JCSG-plus, PACT, CP-Custom-IV and the Classics, PEGs, PEGsII, MbClass and MbClass II Suites can be used for screening different pH, precipitants and additives. In general, extensive screening trials are performed by using a crystallisation robot. The CrystalMation™ system from Rigaku is a fully integrated platform for protein crystallisation, automating every step from custom screen making to crystallization trial imaging and analysis. Although the robot can greatly facilitate the crystal screen, optimization is necessary to obtain high resolution crystals. Some common methods, used in soluble proteins can be applied to membrane protein crystallization also, e.g, the vapor diffusion method. In recent years some special methods were developed specially for the crystallization of membrane protein, e.g, crystallization in lipidic cubic phases.

i) The vapor diffusion method

The vapor diffusion method is the most commonly used methods for protein crystallization. These are known as the hanging drop and sitting drop methods. Vapor diffusion method is a simple method. One can control the volume of the drop, the composition of the well solution, and the ratio between the protein sample and the reservoir solution to obtain a large series of conditions, i.e., pH, precipitants, salts, detergents, etc. It is also easy to be performed on automation platform. In many cases,

the vapor diffusion method, either hanging drop or sitting drop, has been shown to be useful in obtaining membrane protein crystals for x-ray diffraction studies.

ii) Crystallization of membrane proteins in lipidic cubic phase (LCP)

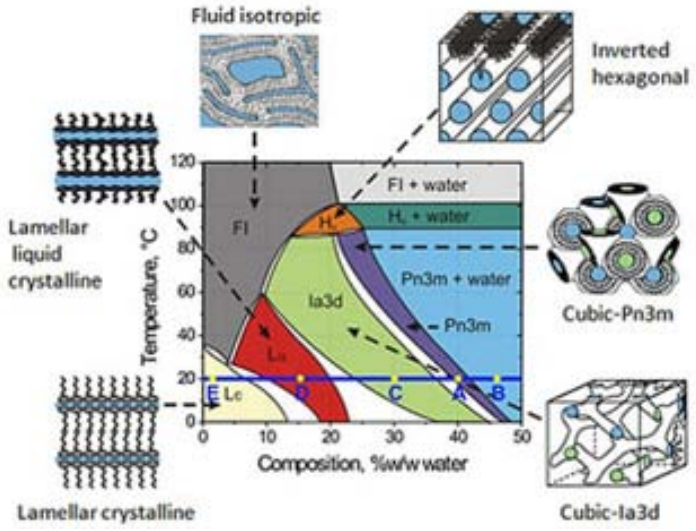



**Figure 1.22** Stability and structure of the phases formed by the monoolein/water system. (After Misquitta et al. 2004)

**Figure 1.23** Cartoon representation of the events proposed to take place during the crystallization of an integral membrane protein from the lipidic cubic mesophase (After Caffrey and Cherezov 2009).(Caffrey and Cherezov 2009)

Under certain conditions, lipids aggregate into different forms, e.g., spheres of lipid molecules (micelles), pairs of layers that face one another (lamellar phase, observed in biological system as a lipid bilayer), a tubular arrangement (hexagonal), or various cubic phases (Fd3m, Im3m, Ia3m, Pn3m, and Pm3m being those discovered so far), etc. LCP can provide an environment that is similar to natural membrane for membrane protein, which can prevent the degeneration of the membrane protein solubilized in detergent for a long time. Monoolein mixed with water is often used to generate the LCP for crystallization studies. This system exhibits a large existence range of two cubic phases (Ia3d and Pn3m) in its temperature-composition phase diagram (Figure 1.22, Briggs 1996). Figure 1.23 proposes the crystallization process of a membrane integral protein from the lipidic cubic mesophase.This approach was firstly used in the study of the crystal structure of bacterial rhodopsin (Rummel et al. 1998). However, due to the complexity of the experimental operation, the LCP method has not been widely used. In recent years, the robot was designed and built first and foremost to perform automated crystallization of membrane proteins using the in meso method (Cherezov et al. 2004), which leads this method as becoming feasible(Cherezov et al. 2007; Hanson et al. 2008; Jaakola et al. 2008). At present, 23 unique proteins and 93 total structures have been obtained by in meso crystallization (including crystallization in lipidic cubic and sponge phases).

iii) Antibody fragment mediated cystallization

The antibody fragment mediated crystallization method was invented by Dr. Hartmut Michel. In the protein-detergent complex crystal, the interactions between adjacent protein molecules are made by the polar surfaces of the protein protruding from the detergent micelle(Michel 1983). Although detergent-mediated protein-protein interactions might stabilize the crystal packing, these interactions do not lead to rigid crystal contacts. The strategy of the antibody fragment mediated crystallization method is to attaching polar domains with specifically bound antibody fragments, in order to increase the probability of getting well-ordered crystals by enlarging the polar surface of the protein (Ostermeier et al. 1995). A recent study shows that the specifically bound

antibody fragments may be useful to obtain different conformational crystal structures(Krishnamurthy and Gouaux 2012).

# 2 Objectives

As membrane transporters play a very important role to maintain the normal physiology of the organism as well as in drug safety and efficacy, it is very important to study the structure and function of these types of membrane proteins. Melibiose permease (MelB), a membrane transporter, couples the uphill transport of the sugar to the downhill electrochemical ion gradient. MelB is of great interest because it can use different sugars (either α- or β-galactosides) and cations ($Na^+$, $Li^+$, and $H^+$), whereas other symporters like lactose permease, a member of the major facilitator superfamily (MFS) uses only $H^+$ (Saier 2000). This implies that MelB should present some unique substrate recognition transport characteristics. The in-depth study of the structure and function of MelB may provide us with key advancements in the understanding of the cotransport mechanism of membrane transporters.

The current biochemical, biophysical, and structural data for MelB fail to give us a clear scene of the substrates recognition mechanism, and explain the structural reorganization that occurs and that ultimately forces the conformational changes needed for transport through the membrane. Although an enormous amount of data on MelB has been published, until now many questions remain unanswered:

- Which residues are involved in cations and sugar binding?

- Why does the binding of $Na^+$ or $Li^+$ to MelB greatly enhance its affinity to sugar?

- What is the mechanism by which the sugar is carried through the membrane?

- Why can MelB use $Na^+$, $Li^+$, $H^+$ as a coupling cation to uptake sugar?

-How can MelB bind α- and β-galactosides?

To answer these questions, more biochemical and biophysical evidences, as well as a high resolution structure are required. Fourier transform infrared difference ($IR_{diff}$) spectroscopy, as well as fluorescene spectroscopy combined with site-directed mutagenesis has been shown tobe a powerful technique to detect conformational changes induced by cation and sugar binding to MelB.

In this thesis, I plan to use these techniques to explore the structure and function of melibiose prermease. **Firstly**, I will focus on the R149C mutant, which cannot bind sugar and cannot transport (Abdel-Dayem et al. 2003). In this work, I am going to reveal how R149C affects sugar binding and translocation. **Then** I will continue to study the role of helix V by cysteine scanning mutagenesis. As previous studies indicate, helix V may be involved in the sugar binding (Basquin 2001) and in a 3D model, helix V was shown to be close to theimportant helix II (Yousef and Guan 2009). **Finally**, I will tryto crystallyze the MelB transporter and apply X-ray diffraction to obtain its 3D structure.

# 3 Materials and Methods

## 3.1 Materials

Tetramethylrhodamine-5-maleimide (TMRM) was obtained from Sigma-Aldrich. (2-(trimethylammonium)ethyl)MethaneThioSulfonate bromide was obtained from Affymetrix, Inc. Trimethylaminoethylmethanethiosulfonate (MTSET) is from Affymetrix, Inc.Synthesis of 2'-(N-dansyl)-aminoethyl-1-thio-D-galactopyranoside (thereafter $D^2G$) was carried out by Dr. B. Rousseau and Y. Ambroise (Département de Biologie Joliot Curie/CEA-Saclay, France) and kindly provided by Dr. Gérard Leblanc. *Escherichia coli* lipid extract for the reconstitution of proteins was from Avanti Polar Lipids, Inc. Restrictions enzymes were purchased from New England Biolabs, cloning kits from Qiagen, detergents from Anatrace, and PEG from Fluka. Crystallization supplies were from Hampton Research or Molecular Dimensions. All other materials were reagent grade and obtained from commercial sources.

## 3.2 Protein preparation

### 3.2.1 Bacterial strains and plasmids

The complementary DNA of full-length MelB encoding the tagged permease (six successive His residues at C-terminal) from *E. coli* strain K-12 was subcloned into the vector pK95ΔAHB, derived from pKK 223-3. pK95ΔAHB, carried an ampicillin resistance and the wild-type or mutated *melB* gene under the control of the tac promoter (Mus-Veteau and Leblanc 1996; Pourcher et al. 1995), was transformed into *E.coli* DW2-R, a *recA*$^{-}$ derivative of strain DW2 (*melA*$^{+}$, Δ*melB*, Δ*lacZY*) (Botfield and Wilson 1988). Extension of the C-terminal end of melibiose permease with six consecutive histidines has no effect on its $H^{+}$- or $Na^{+}$- dependent melibiose transport properties (Pourcher et al. 1995).

### 3.3.2 Site-directed mutagenesis

All of the mutants involved in this study were constructed using the Cys-less MelB, which contains a valine instead of Cys-235 and serines instead of Cys-110, Cys-310, and Cys-364(Weissborn et al. 1997) as a genetic background (Table 3.1). R149C, R149Q, Q149K, R139C, S153C, T163C, and G156C mutants were generous gifts from Dr. Gérard Leblanc. The other mutants were obtained by the polymerase chain reaction (PCR) using two single-primer reactions in parallel as described (Edelheit et al. 2009), whichis a simple and efficient site-directed mutagenesis method avoiding random repeats of primer in cloned DNA. To help the oligo-design, we used an automated web site (http://bioinformatics.org/primerx) (Table 3.2). In this method, the length, complementary, $T_m$ of primers were not severely limited. In general, the mutation can be placed as close as four bases away from the 5'-terminus and at least 6–8 bases from the 3'-terminus, and at least one G or C should be placed at the end of each terminus (Zheng et al. 2004). All of the primers were ordered from Invitrogen™.

**Table 3.1**Site-directed mutagenesis.cDNA encoding cysteine-less *E.coli*MelB was subjected to oligonucleotide-mediated, site-directed mutagenesis, creating a series of mutant cDNAs in which each of the residues was individually changed to cysteine or other amino acids. Residue # refers to the amino acid numbering for MelB. Amino acids are designated by the three-letter code.

| Residue No. | Amino acid change | Codon change |
|---|---|---|
| 139 | Arg→Cys | CGT→TGT |
| | Arg→Cys | CGT→TGT |
| 149 | Arg→Gln | CGT→CAG |
| | Arg→Lys | CGT→AAA |
| 150 | Phe→Cys | TTT→TGT |
| 151 | Phe→Cys | TTT→TGT |
| 152 | Ala→Cys | GCC→TGC |
| 153 | Ser→Cys | AGT→TGT |
| 153 | Ser→Ala | AGT→GCT |
| 154 | Leu→Cys | CTG→TGC |
| 155 | Ala→Cys | GCA→TGC |
| 156 | Gly→Cys | GGC→TGC |
| 157 | Phe→Cys | TTT→TGC |
| 158 | Val →Cys | GTT→TGT |
| 159 | Thr→Cys | ACG→TGC |
| 159 | Thr→Ala | ACG→GCA |
| 160 | Ala→Cys | GCA→TGC |
| 161 | Gly→Cys | GGT→TGT |
| 162 | Val→Cys | GTG→TGC |
| 163 | Thr→Cys | ACG→TGC |
| 262 | Ile→Cys | ATC→TGT |

**Table 3.2** The primers for site-directed mutagensis

| Primer | Sequence* |
|---|---|
| F150C-F | 5' CCTTATCCGCGTT**G**TTTTGCCAGTCTGG 3' |
| F150C-R | 5' CCAGACTGGCAAAA**C**AACGCGGATAAGG 3' |
| F151C-F | 5' CTTATCCGCGTTTTT**G**TGCCAGTCTGGCAG 3' |
| F151C-R | 5' CTGCCAGACTGGCA**C**AAAAACGCGGATAAG 3' |
| A152C-F | 5' CTTATCCGCGTTTTTTT**TG**CAGTCTGGCAGGCTTTG 3' |
| A152C-R | 5' CAAAGCCTGCCAGACTG**CA**AAAAAAACGCGGATAAG 3' |
| S153A-F | 5' CGCGTTTTTTTGCC**GC**TCTGGCAGGCTTTG 3' |
| S153A-R | 5' CAAAGCCTGCCAGA**GC**GGCAAAAAAACGCG 3' |
| L154C-F | 5' CGCGTTTTTTTGCCAGT**TGC**GCAGGCTTTGTTACGG 3' |
| L154C-R | 5' CCGTAACAAAGCCTGC**GCA**ACTGGCAAAAAAACGCG 3' |
| A155C-F | 5' GTTTTTTTGCCAGTCTG**TGC**GGCTTTGTTACGGCAGG 3' |
| A155C-R | 5' CCTGCCGTAACAAAGCC**GCA**CAGACTGGCAAAAAAAC 3' |
| F157C-F | 5' GGCAGGCT**GC**GTTACGGCAGGTGTGACGC 3' |
| F157C-R | 5' CCTGCCGTAAC**GC**AGCCTGCCAGACTGGC 3' |
| V158C-F | 5' GTCTGGCAGGCTTT**TG**TACGGCAGGTGTG 3' |
| V158C-R | 5' CACACCTGCCGTA**CA**AAAGCCTGCCAGAC 3' |
| T159C-F | 5' CTGGCAGGCTTTGTT**TGC**GCAGGTGTGACGCTAC 3' |
| T159C-R | 5' GTAGCGTCACACCTGC**GCA**AACAAAGCCTGCCAG 3' |
| T159A-F | 5' CTGGCAGGCTTTGTT**GCA**GCAGGTGTGACGCTAC 3' |
| T159A-R | 5' GTAGCGTCACACCTGC**TGC**AACAAAGCCTGCCAG 3' |
| A160C-F | 5' GCAGGCTTTGTTACG**TGC**GGTGTGACGCTACC 3' |
| A160C-R | 5' GGTAGCGTCACACC**GCA**CGTAACAAAGCCTGC 3' |
| G161C-F | 5' GCTTTGTTACGGCA**T**GTGTGACGCTACC 3' |
| G161C-R | 5' GGTAGCGTCACAC**A**TGCCGTAACAAAGC 3' |
| V162C-F | 5' GTTACGGCAGGT**TGC**ACGCTACCATTTG 3' |
| V162C-R | 5' GCGT**GCA**ACCTGCCGTAACAAAGCCTGC 3' |
| I262C-F | 5' CATATGTTTGCGGTGATGCGGATTTGTTCCC 3' |
| I262C-R | 5' CGCATCACCGCAAACATATGAGAAATAATAG 3' |

* Mutated nucleotides are written in bold characters.

**Table 3.3** PCR reaction components

| | Reaction 1 | Reaction 2 |
|---|---|---|
| Template plasmid DNA | ~500ng | ~500ng |
| Forward primer | 40 pmol | - |
| Reverse primer | - | 40 pmol |
| $MgCl_2$ | 0.2 mM | 0.2 mM |
| dNTPs | 0.2 mM | 0.2 mM |
| 10X PCR Buffer for KOD Hot Start DNA Polymerase | 1X | 1X |
| PCR Grade Water | | |
| KOD hot start DNA polymerase | 1 U/µl | 1 U/µl |
| Total volume | 25 µl | 25 µl |

**Table 3.4** Denaturation and slow cooling conditions to allow reannealing of PCR products. (After Edelheit et al. 2009)

| Step | Temperature | Time (min) |
|---|---|---|
| 1 | 95 | 5 |
| 2 | 90 | 1 |
| 3 | 80 | 1 |
| 4 | 70 | 0.5 |
| 5 | 60 | 0.5 |
| 6 | 50 | 0.5 |
| 7 | 40 | 0.5 |
| 8 | 37 | Holding |

Two PCR reactions in parallel were performed each using a mismatched primer (forward primer or reverse primer), designed to introduce a mutation at a specific point in a final volume of 25 µl using KOD hot start DNA polymerase (Novagen). After the initial denaturation step at 95 ℃ for 5 min, the PCR was conducted for 30 cycles with denaturation at 95 ℃ for 50 s, primer annealing at 55 ℃ for 50 s and DNA synthesis at 68 ℃ for 7.5 min. When all of the cycles ended, the samples were kept at 68 ℃for 7 min to finish all of DNA synthesis. After the PCR, the two PCR products are combined (giving a total volume of 50 ul) and heated to 95 ℃ for 5 min to separate the PCR product from the plasmid template, then followed by slowly cooling to 37 ℃ as indicated in table 3.4, to promote reannealing of denatured plasmid templates and PCR products (Edelheit et al. 2009).

The PCR products were purified by using the QIAquick PCR Purification Kit (Qiagen). After digestion by adding 30 units of DpnI (New England Biolabs) to the purified PCR products, which digests the methylated parental plasmid strands, and incubating overnight at 37 ℃, the PCR products were transformed to DH5α competent cells by incubating for 30 min on ice, followed by heat-shocking at 42 ℃ for 90 s and then transfering to ice for 5 min. After adding 1 ml LB (Lysogeny Broth), the cells were incubated on a shaker at 37 ℃ for 60 min. Then the cells were precipitated by centrifuging at 3000g for 5 min. Pellets were responded in 200 µl LB, and tracked on LB plate containing 0.1 mg/ml ampicillin. After incubating the plates overnight at 37 ℃, for each transformation we select 3 colonies at random and grow overnight in 12 ml LB+ampicillin medium at 37 ℃. The plasmids were isolated using a QIAprep Spin Miniprep Kit (Qiagen). Mutations were checked by DNA sequencing.

### 3.2.3 Cell culture

The pK95ΔAHB plasmids with a cassette containing wild-type or mutated *melB* gene were transformed into *E.coli* DW2-R strain and were stored in glycerol storage solution at -80 ℃. When used, the cells were streaked on Macconkey plates containing 0.1 mg/ml ampicillin and incubated at 30 ℃ over night. A single colony was picked, and mini cultured in 12 ml LB containing 0.1 mg/ml ampicillin at 30 ℃ overnight.

Thereafter, the 12 ml mini culture was poured into 200 ml M9 medium (6 g/L $Na_2HPO_4$, 3 g/L $KH_2PO_4$, 0.5 g/L NaCl, 1 g/L $NH_4Cl$)supplemented with 0.5% (v/v) glycerol, 0.2% (w/v) casaminoacids, 10 mM thiamine, and 0.1 mg/ml ampicillin, and incubated at 30 ℃ for 7-9 h (midi culture). For the big culture, 50 ml of the midi culture was poured into 1 L of M9 medium supplemented with 0.5% (v/v) glycerol, 0.2% (w/v) casaminoacids, 10 mM thiamine, and 0.1 mg/ml ampicillin, until an $OD_{600}$ of 1.6-1.8 was reached. Cells were harvested by centrifuging at 13000 g for 10 min. Pellets were resuspended in 50 mM Tris pH8.0, 50 mM NaCl, 5 mM BME (β-mercaptoethanol), and frozen at -80 ℃.

3.2.4 Protein purification

The protein purification was carried out mainly as described (Pourcher et al. 1995).In short, frozen cells were thawed and washed with50 mM Tris-HCl pH8.0, 50 mM NaCl, and 5 mM 2-mercaptoethanol. Cells were homogenized in the same buffer supplemented with lysozyme (dissolved with water) and EDTA (dissolved with 3 M NaOH) at a final concentration of 1 mg/ml and 15 mM respectively, and incubated at 4 ℃ for 30 min with slight shaking. Before the membranes were disrupted by passing through a microfluidizer (Model 110S, Microfluidics) at 20,000 psi, DNAase, RNAase, and $MgSO_4$ at final concentration 20 μg/ml, 20 μg/ml, and 15 mM respectively, were added. Cell debris was removed by low-speed centrifugation for 10 minutes. The supernatant was collected and ultracentrifuged at 310,000 *g* for 30 minutes. The membrane vesicle pellets were washed with 10% (v/v) glycerol, 600 mM NaCl, 20 mM Tris, 5mM BME, pH8.0.

The membrane fraction was incubated with 1% (w/v) 3-(laurylamido)-N, N′-dimethylaminopropylamine oxide (LAPAO, Anatrace) for 30 minutes at 4 ℃. Following another ultracentrifugation step at 310,000 *g* for 15 minutes, the supernatant was collected and incubated with $Ni^{2+}$-NTA affinity resin (Sigma-Aldrich) at 4 ℃ for 1 h in an orbital shaker. $Ni^{2+}$-NTA affinity resin was washed previously with 20 mM Tris, 600 mM NaCl, 10% (v/v) Glycerol, 10 mM melibiose, 10 mM imidazole, 0.1 % (w/v) LAPAO, and 5 mM BME at pH 8.0. The resin was pelleted by centrifuging at 4000g for

5 min, and the pellet was again washed with the same buffer. $Ni^{2+}$-NTA affinity resin was then loaded into the column (Rio-Rad), and washed with one column volume of 20 mM Tris, 600 mM NaCl, 10% (v/v) Glycerol, 10 mM melibiose, 10 mM imidazole, 0.1 % (w/v) LAPAO, and 5 mM BME at pH 8.0. Thereafter, the detergent was exchanged by extensive wash with 20 mM Tris buffer, 100 mM NaCl, 10% (v/v) Glycerol, 10 mM melibiose, 10 mM imidazole, 0.1 % (w/v) dodecyl-β-D-maltopyranoside (DDM, Anatrace), and 5 mM BME at pH 8.0. The concentration of salt was reduced by washing with three column volumes of the same buffer, expect 100mM NaCl. The protein was eluted at pH 8.0 with 100 mM imidazole and 0.1% (w/v) DDM.

If necessary, the protein was further purified by ion-exchange chromatography and size-exclusion chromatography. The concentration of purified protein was evaluated by measuring the UV absorbance at 280nm using quartz cuvette. The purity of protein was evaluated by sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) (Laemmli 1970). In general, following this method, about 8-10 mg purified MelB WT or mutants would be obtained from 16 L cells. The pure protein was pooled and frozen at -80 ℃ until used.

3.2.5 Reconstitution

Reconstitution was carried out by a procedure involving detergent adsorption on polystyrene beads (Meyer-Lipp et al. 2006). *E.coli* lipids (Avanti Polar Lipids) were solubilized with chloroform. The chloroform was eliminated using a rotavapor. The lipids were resuspended in 20 mM MES buffer, 100 mM KCl, pH6.6 at 10 mg/ml by vortexing, then subjected to 3-5 cycles of freeze/defreeze procedure with liquid nitrogen and hot water to make big unilamelar liposomes. Before using, the liposomes were homogenized by extruding with 0.1 μm filter (Millipore). Bio-Beads SM-2 (Bio-Rad) were washed 5-7 times with distilled water, and once with 20 mM MES, 100 mM KCl, pH6.6. Bio-Beads SM-2weredried by vacm pump, distributed into three aliquots based on 120 mg Bio-Beads every 1 ml solution.

The lipids were directly added to the solution containing the purified protein in 2-fold quality protein to lipids, and then incubated at 4 ℃ for 10 min. Three Bio-Beads

aliquots were added progressively at 10 min interval and the sample was left overnight at 4 ℃ with slight shaking. Bio-Beads were eliminated by filtering the solution in a column (Bio-Rad). Proteoliposomes were pelleted by centrifuging at 310,000g at 4 ℃ for 30 min. Pellets were washed twice with 20 mM MES buffer, 100 mM KCl, pH6.6.Proteoliposomes were finally resuspended in the same buffer at ~10 mg/ml, and kept at -80 ℃. The concentration of proteoliposomes was determined using the Lowry method (Lowry et al. 1951) by using serum albumin as standard.

## 3.3 Preparation of membrane vesicles

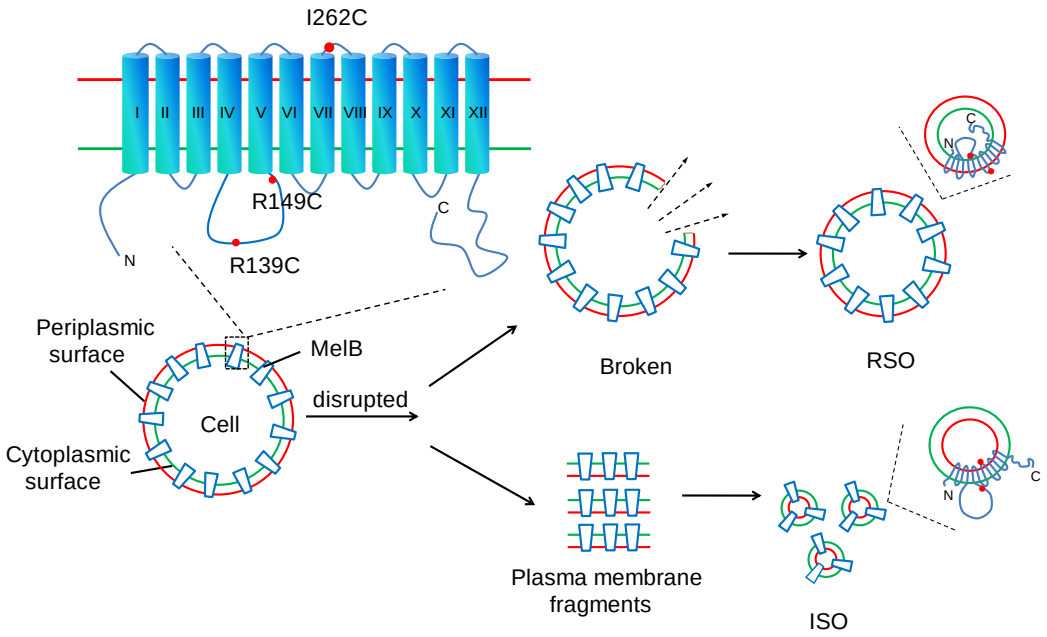



**Figure 3.1** Scheme representing the formation of ISO and RSO membrane vesicles from disrupted cells (see text).

Right-site-out (RSO) and inside-out (ISO) membrane vesicles were prepared from the same culture to guarantee the same concentration of MelB in the membranes (Figure 3.1). RSO membrane vesicles were prepared by lysozyme-ethylenediaminetetraacetic acid treatment and osmotic lysis (Kaback 1971). Approximatedly 3g (wet weight) of cells were rapidly thawed at 30 ℃ and washed with 10 mM Tris-HCl (pH 8.0) at 4 ℃. The cells were resuspended (1g, wet weight, per 80 ml) at room temperature in 50 mM Tris-HCl, pH 8.0, containing 30% (w/v) sucrose (270 ml). The suspension was swirled by means of a magnetic stirrer. Potassium EDTA (pH 7.0) and lysozyme were added to a final concentration of 10 mM and 0.5 mg/ml, respectively, and the suspensions were incubated for 30 min at room temperature, thereafter subjected to centrifuge at approximately 16,000g for 10 min. The pellet was homogenized in the smallest possible volume of 0.1 M KPi, pH 6.6, containing 30% (w/v) sucrose and 20 mM $MgSO_4$ using a syringe. 5 mg RNAase and DNAase each were added to the homogenate. The homogenized spheroplast suspension was poured directly into 500 ml of 50 mM KPi buffer, pH 6.6, which had been pre-equilibrated at37 ℃. The lysate was incubated for 15 min at 37 ℃ with vigorous stirring. Thereafter, potassium EDTA (pH 7.0) was added to 10 mM final concentration, and incubated for another 15 min. $MgSO_4$was added to 15 mM final concentration and incubated for another 15 min. The lysate was centrifuged at 16,000g for 30 min. The pellet was homogenized in 0.1 M KPi buffer, pH 6.6, 10 mM EDTA, following by a centrifugation at 900g for 1h. The supernatant was carefully collected and centrifuged at 146,000 *g* for 30 minutes and the pellet was washed twice with 0.1 M KPi, pH 6.6, 10 mM EDTA. In order to remove whole cells and partially lysed forms from the membrane preparation, the RSO membrane vesicles pellet was resuspended in 0.1 M KPi (pH 6.6), containing 10 mM $MgSO_4$ and 20% (w/v) sucrose. The suspension was carefully layered on top of 60% sucrose (w/v) containing 0.1 MKPi (pH 6.6), and 10 mM $MgSO_4$ and centrifuged at 64,000*g* overnight. The thick layer of membrane remaining at the interface was carefully aspirated, diluted with 0.1 MKPi (pH 6.6), 10 mM $MgSO_4$ and centrifuged at 146,000*g* until clear. The purified RSO vesicles were washed several times with the same buffer, and resuspended to a

protein concentration of 10-20 mg/ml in 0.1 mM KPi (pH 6.6), 10 mM $MgSO_4$, frozen in liquid nitrogen, and stored at -80 ℃ until use.

The ISO membrane vesicles were prepared by microfluidizer pressure as described (Pourcher et al. 1995), which is similar to a French press. In short, approximately 4 g (wet weight) of cells were rapidly thawed at 30 ℃and washed twice with a buffer containing 10 mM Tris-HCl (pH 8.0). Cells were resuspended in 100 ml of the same buffer supplemented with 15 mM EDTA and 100 mg/ml lysozyme and incubated at 4 ℃ for 30 min with slight shaking. Before applying a pressure of 20000 psi, 20 μg/ml of each DNAase and RNAase as well as 15 mM $MgSO_4$ were added to the cell suspensions. The cell suspensions were passed through the microfluidizer three times and unbroken cells and large debris were eliminated by centrifugation at 13,000g for 30 min. The supernatant was collected and centrifuged at 146,000 *g* for 30 minutes. The ISO membrane vesicles were washed 3-4 times with 0.1 M KPi (pH 6.6), and resuspended at about 10-20 mg of protein/ml in 0.1 M KPi, 10 mM $MgSO_4$, frozen in liquid nitrogen and stored at −80 ℃ until use.

The total amount of protein in membrane vesicles was determined using the Lowry method (Lowry et al. 1951) by using serum albumin as standard.

## 3.4 Orientation assay

The orientation of MelB in the membrane vesicles and liposomes was determined by means of site-directed alkylation of cysteine as described (Guan and Kaback 2007; Nie et al. 2007; Nie et al. 2008) with some modifications. Figure 3.2 shows the principle of orientation assay. TMRM is a membrane-permeant fluorescent alkylating agent, which can label cysteine residues facing to the inside or the outside medium. MTSET is a quaternary ammonium ion and is expected to be membrane-impermeant, labeling only the cysteines facing to the outside medium. We firstly used MTSET to protect all of the cysteines facing to the outside medium, and thereafter TMRM was used to determine

the residual cysteines. There are two cases: i) TMRM signal can be determined, suggesting that the labeled cysteines should be facing to the inside medium; ii) No TMRM signal can be detected, suggesting that the cysteines should be facing to the outside medium.

TMRM was dissolved in dimethyl sulfoxide (DMSO) and its concentration determined by measuring the absorbance in methanol at 541 nm (extinction coefficient, 95,000 $cm^{-1}$·$M^{-1}$). RSO or ISO membrane vesicles (about 0.1 mg of total protein in 200 µL of 0.1 M KPi pH 7.5) were incubated with 0.1 mM MTSET, (or reference samples without incubating with MTSET) on ice for 5 minutes. Reactions were stopped by adding 1.4 ml of ice-cooled 0.1 M KPi (pH 7.5), and centrifuged. Pellets were washed 3 times with the same buffer, and resuspended in 200 µL of 0.1 M KPi (pH 7.5). TMRM was added to a final concentration of 40 µM, and incubated in ice for 30 min. Reactions were terminated at the indicated time by adding 10 mM dithiothreitol (DTT). The membranes were then incubated with 2% (w/v) DDM at 4 ℃ for 30 minutes on a rotating platform. The solubilized protein was incubated with 25 µL Ni-NTA affinity resin on a rotating platform at 4 ℃ for 1 h and centrifuged at 1,000 *g* for 1 min. The pellet was washed 6 times with 0.1 M KPi (pH 7.5), 100 mM NaCl, 10 mM imidazole and the protein was eluted with 25 µl of the same buffer containing 300 mM imidazole. The samples were then subjected to sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE). The wet gels sandwiched between two glasses plates were then imaged directly on an Imager VersaDoc MP 4000 System (Bio-Rad). The Cy3 laser setting (excitation at 532 nm) with a 580 nm filter (emission wavelength) was used to capture the image. The SDS-PAGE gels were then coomassie stained and scanned. The TMRM signal and the amount of protein were estimated by measuring the density of each band by using Quantity One software (Bio-Rad).

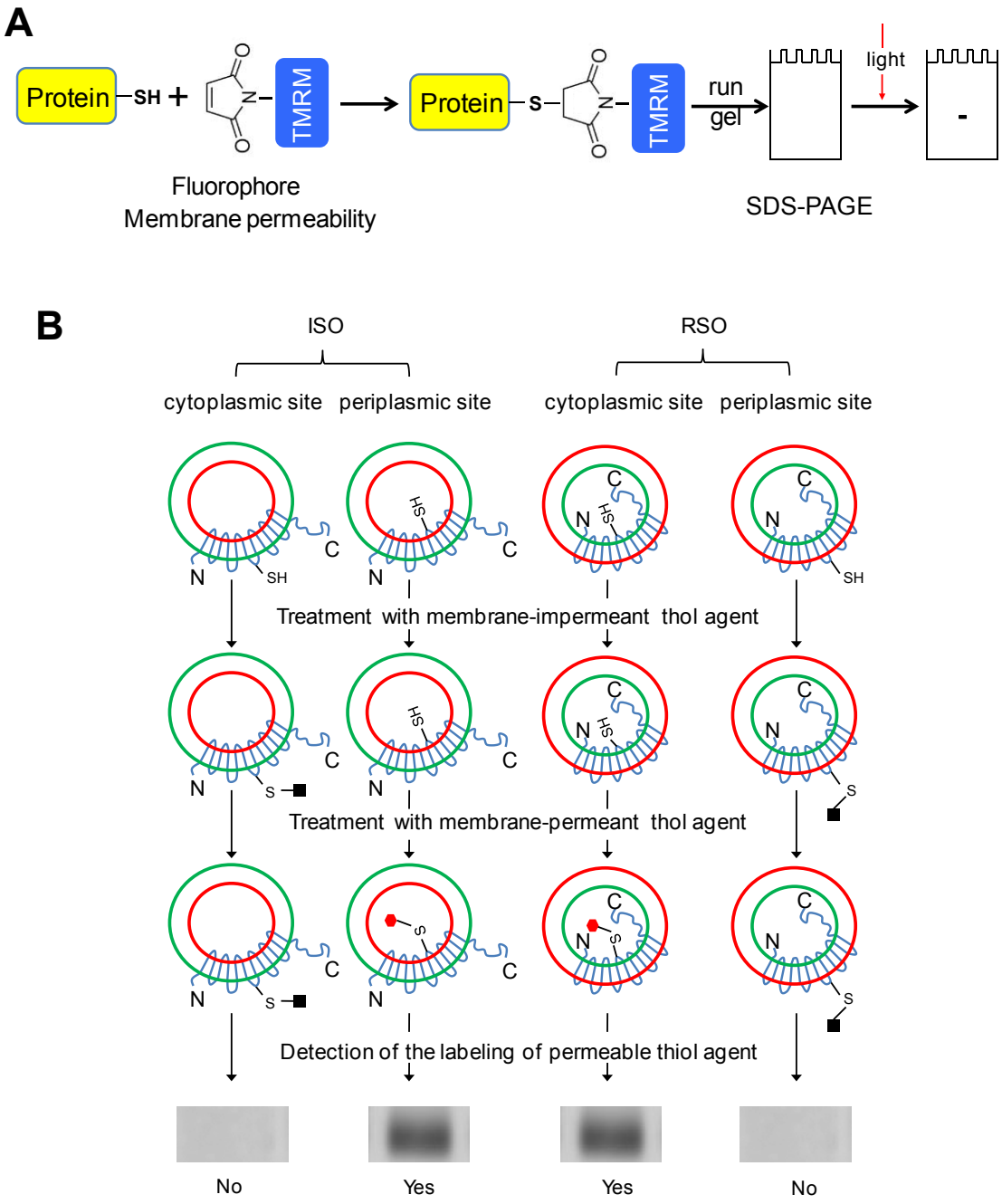



**Figure 3.2** Illustration of protein orientation assay. (A) Scheme representing the sit-directed alkylation of cysteine with TMRM and the detectment of TMRM labeling. (B) The strategy of labeling experiments to distinguish cysteines location (see text).

## 3.5 Infrared spectroscopy

The infrared spectroscopy (FTIR) is a valuable tool for the investigation of protein structure (Arrondo et al. 1993; Goormaghtigh et al. 1994a, b, c; Jackson and Mantsch 1995; Siebert 1995; Arrondo and Goni 1999) and the molecular mechanism of protein reactions (reviewed by (Siebert et al. 1990; Rothschild et al. 1992; Mantele 1993; Maeda 1995; Siebert 1995; Slayton and Anfinrud 1997; Gerwert 1999; Fahmy et al. 2000; Jung 2000; Vogel and Siebert 2000; Breton 2001; Fahmy 2001; Kim and Barry 2001; Barth and Zscherp 2002). One advantage of this technique over other spectroscopies is that due to its high wavelength, no problem appears due to light scattering that may occur by other techniques in the study of proteoliposomes. Unlike other spectroscopic techniques, it does not require incorporation of any external probes. FTIR can be used to study the secondary structure of proteins in different environments, e.g., proteoliposome in aqueous suspensions (Kabsch 1993; Dave et al. 2000), in organic solutions (Torres and Padros 1993), solubilized (Dave et al. 2000; Lorenz-Fonfria et al. 2001) or in dry films (Lorenz-Fonfria et al. 2003).

### 3.5.1 Vibration

At temperatures higher than 0K, all molecules vibrate. The simple two-atomic oscillator illustrates well some of the fundamental principles that govern the relationship between the vibrational spectrum of a molecule and its structure and environment. The frequency, υ, of a two-atomic oscillator is given by

$$\upsilon=\frac{1}{2\pi}\sqrt{k/m_r} \quad (1),$$

where k is the force constant between the two atoms, and $m_r$ the reduced mass ($1/m_r=1/m_1+1/m_2$). The value of kis related to the type of the chemical bond and the surrounding environment and reflects the electron density in the bond between the two atoms. Clearly, any inter- or intra-molecular factor that alters the electron density in the bonds will affect the vibrational spectrum(Barth and Zscherp 2002). Therefore, valuable information can be derived from the vibrational spectra.

Vibrations can be classified into two groups: Stretching vibration, where the atoms vibrate in the direction of the bond, and bending vibration, where the atoms move perpendicular to the direction of the bond.

3.5.2 Molecular vibration of proteins

i) Amide vibrations

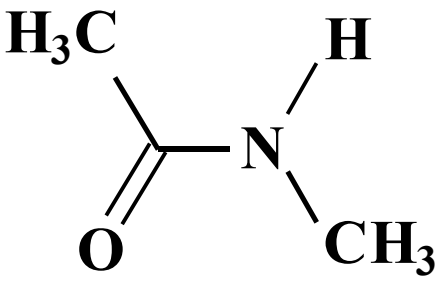



**Figure 3.3** Structure of N-methylacetamide (NMA)

The peptide group, the structural repeat unit of proteins, gives up to 9 characteristic bands named amide A, B, I, II ... VII. N- N-methylacetamide (NMA) (Figure 3.3), the smallest molecule containing the peptide bond (Krimm and Bandekar 1986), 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, specially Amide I, and Amide II, are major bands of the protein infrared spectrum.

**Table 3.5** The amino acid side chain vibrations (according to Venyaminov and Kalnin 1990).

| Amino Acid | Vibration | $cm^{-1}$ |
|---|---|---|
| Asp | -COO st as | 1574 |
| | -COOH st | 1716 |
| Glu | -COO st as | 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 |
| | | 1498 |
| His | ring | 1596 |
| Phe | ring | 1494 |
| terminal | | |
| | -COO st as | 1598 |
| | -COOH st | 1740 |
| | $-NH_3^+$bd as | 1631 |
| | bd s | 1515 |
| | $-NH_2$bd | 1560 |

*st=stretching vibration*

*bd=bending*

*s=symetrical*

*as=asymetrical*

Amide A: absorbs between 3310 and 3270 $cm^{-1}$. This vibration is exclusively formed by the vibration of the bond NH (stretching) and therefore is very sensitive to the secondary structure of the protein.

Amide I: absorbs between 1700 and 1600 $cm^{-1}$ and comes mainly from the stretching vibration of the C=O (70-85%) and C-N groups (10-20%), and is directly related to the backbone conformation and the hydrogen bonding pattern.

Amide II: absorbs at approximately 1550 $cm^{-1}$ and corresponds to the combination of N-H bending vibration (40-60%) and C-N stretching vibration (18-40%). This band is also conformational sensitive.

ii) Amino acid side chain vibrations

Amino acid side chains are involved in the protein reactions, so the presence of bands arising from amino acid side chains can provide valuable information about the protein reaction mechanism, for example, on the protonation state, the coordination of cations, hydrogen bonding, etc.

By thoroughly investigating the contribution of the side chain vibrations in the region between 1800 and 1400 $cm^{-1}$, Venyaminov and Kalnin had established that among the 20 proteinogenous amino acids, only arginine, asparagine, glutamine, aspartic and glutamic acids, lysine, tyrosine, histidine, and phenylalanine have intense absorption in this spectral region(Venyaminov and Kalnin 1990).The amino acid side chain vibrations are summarized in Table 3.5.

3.5.3 Design of experiments to make the IR difference spectra

The conformational changes responding to the binding of ligands can be obtained by IR difference spectrum. This technique has been shown to be useful to obtain valuable information of the substrate binding of melibiose permease (Leon et al. 2005; Leon et al. 2006). Furthermore, this technique has been used to explore the important residues which are involved in the binding of substrates to the protein (Granell et al. 2010).

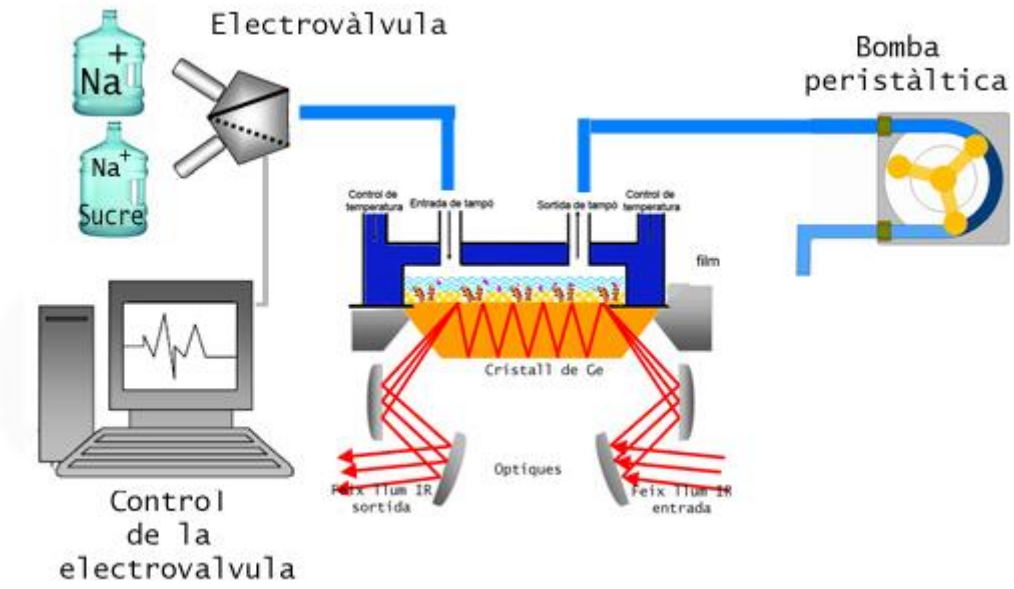



**Figure 3.4** Outlining the assembly used to get substrates dependent IR different spectra. (According to GranellDoctoral thesis)

The experimental setup was the same as that described in a previous study (Figure 3.4, Leon et al. 2005). In summary, 10 µl MelB containing proteoliposome suspensions (~100 µg of protein) was spread homogeneously on a germanium ATR crystal (Harrick, Ossining, NY; 50 × 10 × 2 mm, yielding 12 internal reflections at the sample side) and dried under a stream of nitrogen. The substrate-containing buffer and the reference buffer were alternatively perfused over the proteoliposome film at a rate of ~1.5 ml/min. The switch of buffers was carried out by a computer-controlled electro valve. The film was firstly exposed to the substrate-containing buffer, and then washed with the reference buffer. Table 5.6 summarized the conditions used for obtaining the IR difference spectra for MelB.

In order to increase the signal-to-noise ratio, each difference spectrum was obtained from the mean of about 25,000scans. Spectra were recorded with an FTS6000 Bio-Rad spectrometer equipped with a Mercury-Cadmium-Telluride detector at a resolution of 4 $cm^{-1}$.

Table 5.6 The conditions used for obtaining IR difference spectra for MelB

| Difference spectrum | Substrates containing buffer | Reference buffer | Scans | Cycles |
|---|---|---|---|---|
| 10 mM NaCl vs. H | 2 min for 20 mM MES (pH6.6), 100 mM KCl, 10 mM NaCl | 10 min for 20 mM MES (pH6.6), 100 mM KCl | 500 | 50 |
| 10 mM NaCl,10 mM melibiose vs. 10 mM NaCl | 4 min for 20 mM MES (pH6.6), 100 mM KCl, 10 mM melibiose, 10 mM NaCl | 10 min for 20 mM MES (pH6.6), 100 mM KCl, 10 mM NaCl | 1000 | 25 |
| 50 mM Melibiose vs. H | 4 min for 20 mM MES (pH6.6), 100 mM KCl, 50 mM melibiose | 30 min for 20 mM MES (pH6.6), 100 mM KCl | 1000 | 25 |
| 50 mM melibiose, 10 mM NaCl vs 10 mM NaCl | 4 min for 20 mM MES (pH6.6), 50 mM melibiose, 10 mM NaCl | 30 min for 20 mM MES (pH6.6), 100 mM KCl | 1000 | 25 |

3.5.4 Correction of the difference spectra

The experimental difference spectrum contains four possible contributions:

i) Difference spectrum of the sample (protein and lipid) induced specifically for the binding of substrates.

ii) Difference spectrum of water induced by the presence of substrates, which modify the structure of water (Venyaminov and Kalnin 1990).

iii) Absorbance of the substrate (in our case melibiose absorbs at 1080 $cm^{-1}$; cations do not absorb).

iv) Change in the swelling of the film, with an apparent gain/loss of water with a concomitant apparent loss/gain of sample (protein and lipid).

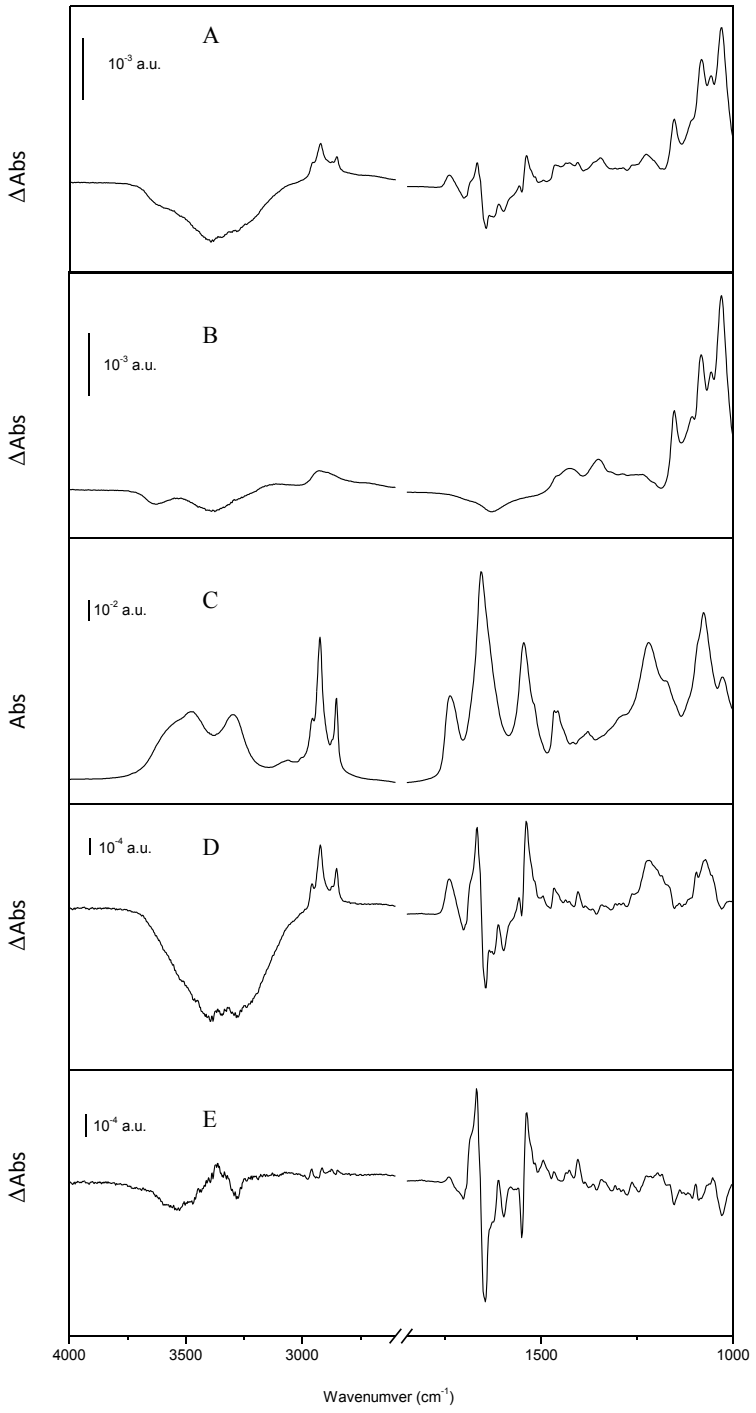



**Figure 3.5** The collection of IR different spectrum. A. Melibiose-dependent IR different spectrum of MelB Cys-less mutant in the presence of $Na^+$, which was obtained from the mean of about 25,000 scans. B. IR different spectrum of buffer. C. absorbance spectrum of a rehydrated film of proteoliposomes in the substrate containing buffer. D. IR different spectrum substrated the contribution of buffer. E. IR different spectrum substrated the contribution of swelling.

Contributions ii and iii were corrected by subtracting the difference spectra obtained in the substrate-containing buffer and the reference buffer (Figure 3.5D). The subtraction factor was adjusted to flat water absorption between 3700 and 2800 $cm^{-1}$ and the absorption bands due to melibiose at 1080 $cm^{-1}$ (Figure 3.5A).

The last remaining contribution was corrected by subtracting the difference specturm from an absorbance spectrum of a rehydrated film of proteoliposomes in the substrate containing buffer (see Figure 3.5C, E). The subtraction factor was adjusted to remove completely the lipid bands (corresponding to the vibration of $CH_2$) at 2920 $cm^{-1}$ and 2850 $cm^{-1}$ (see Figure 3.5C).

### 3.5.5 Spectra deconvolution

Spectra deconvolution was performed using the maximum entropy method as described (Lorenz-Fonfria and Padros 2005). Deconvolution was performed using a Lorentzian band of 7 $cm^{-1}$ width, which value was determined from the spectra as described (Saarinen et al. 1995), and a regularization parameter of $10^{-11}$, which was derived in accordance with the spectral noise content (Lorenz-Fonfria and Padros 2005).

### 3.5.6 Quantitative comparison of intensity and similarity of FTIR difference spectra

Quantitative comparison of intensity and similarity of FTIR difference (thereafter $IR_{diff}$) spectra were performed as described (Granell et al. 2010). In brief, all the difference spectra were normalized for the amount of protein contributing to the IR signal in the 1,700-1,500 $cm^{-1}$ interval, which includes the protein conformation-sensitive amide I and amide II bands from the peptide bond. Quantitative comparison of two normalized difference spectra (an input and a reference spectrum) was performed by a linear regression in the 1,710-1,500 $cm^{-1}$ interval on the first derivative of the difference spectra (Figure 3.6) The correlation analysis provides two relevant parameters: a) the correlation coefficient ($R^2$), which shows the spectral similarity of the input with respect to the reference spectrum in response to the added substrate; b) the slope, which quantifies the relative intensity of common features in the

input with respect to the reference spectrum arising in response to the effect of substrates (Figure 3.6).

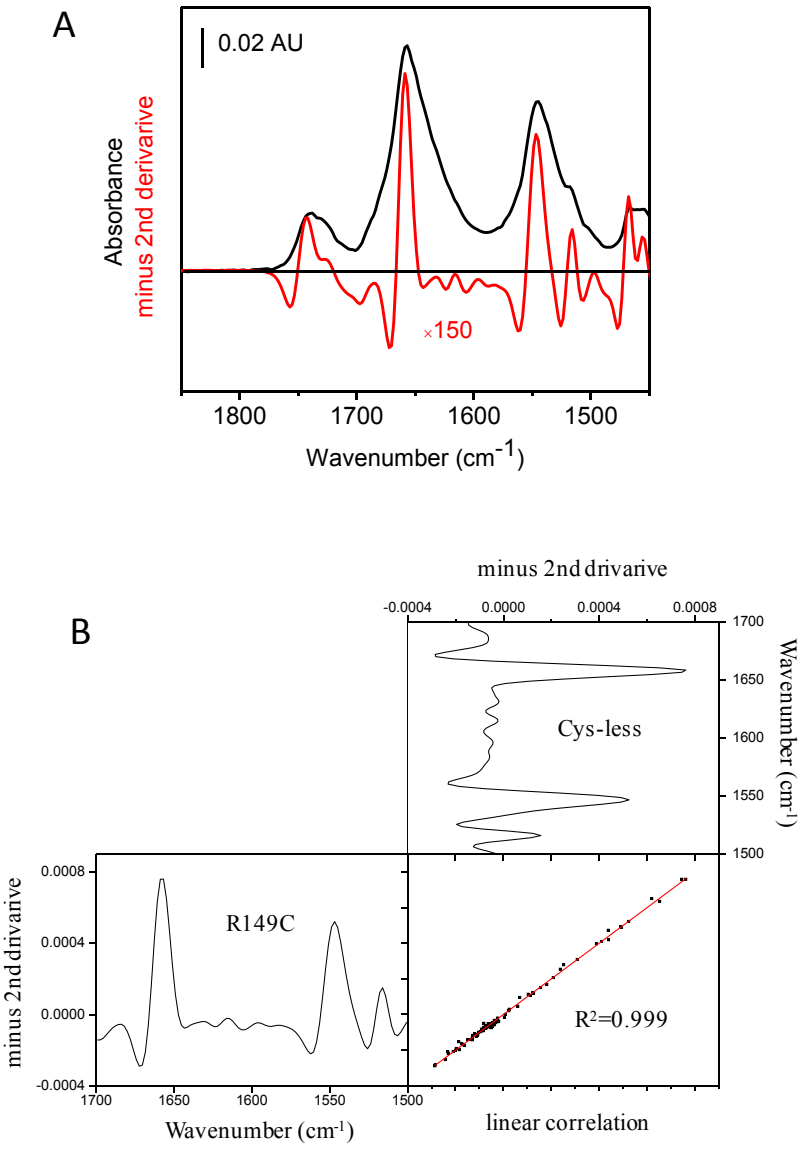



**Figure 3.6** Illustration how to compare the structure of different form MelB using IR spectroscopy of hydrated samples. (A) Absorbance IR spectrum at 4 $cm^{-1}$ resolution of the Cys-less MelB mutant in a hydrated film with the buffer contribution subtracted (black). The spectrum includes the structure-sensitive amide I and II regions. The second derivative at 8 $cm^{-1}$ resolution (red) provides more resolved details, robust to small differences in the buffer subtraction, and therefore more suitable for the structural comparison performed below. (B) The correlation analysis of two absorbance spectra using their second derivative spectra. The parameter m was used to normalize difference spectra of different form MelB mutants to Cys-less, and $R^2$ corresponds to the similarity of the Structural components between different form MelB mutants.

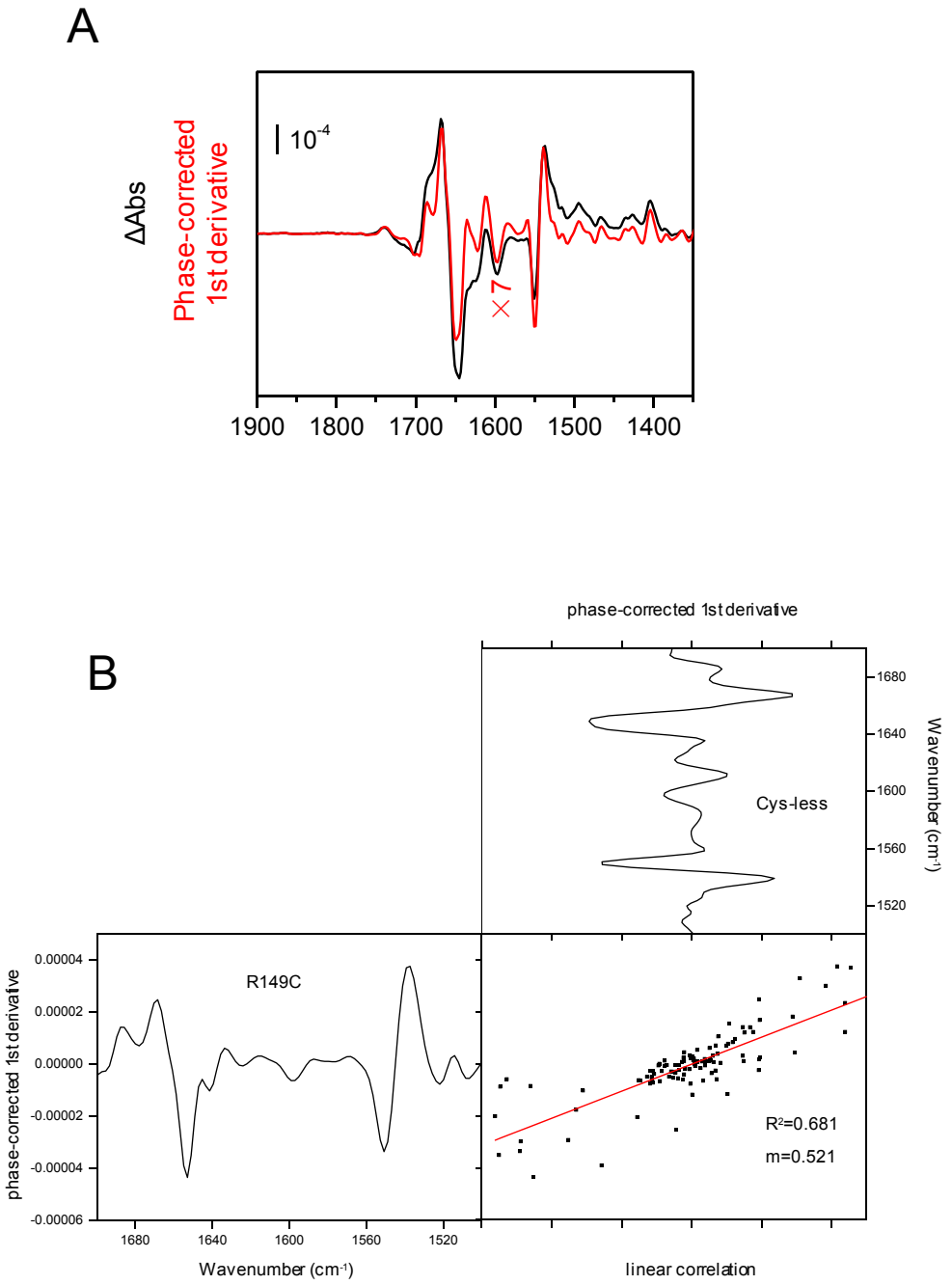



**Figure 3.6** Illustration of how the quantitative comparison of the intensity and similarity of the substrate-dependent IR difference spectra. (A) melibiose-induced IR difference spectrum of Cys-less MelB in the presence of $Na^+$ at 4 $cm^{-1}$ resolution (black), and its phase-corrected first derivative at 6 $cm^{-1}$ resolution (red). (B) The correlation analysis of two difference spectra using their first derivative spectra. The two parameters, $R^2$ and m, correspond to similarity and intensity of the IR different spectra between mutants and Cys-less.

## 3.6 Fluorescence spectroscopy

Fluorescence measurements were performed at 20 ℃in a 1cm quartz cuvette (Hellma Analytics) with a UV-visible QuantaMaster$^{TM}$ spectrofluorometer and processed with the Felix 32 software (Photon Technology International).

### 3.6.1 Trp fluorescence

Trp fluorescence spectra were obtained by setting the excitation wavelength at 290 nm, with a half-bandwidth of 5 nm and collecting the emission spectrum in the290-400 nm interval. The samples were prepared in 0.1 M KPi, pH 7.0 containing about 23 μg/ml protein. The time course of Trp fluorescence spectrum was obtained by exciting at sample at 290 nm and recording the emission intensity at 325 nm as function of time. Before starting the measures, MelB reconstituted in liposomes was subjected to a 30 s sonication in a sonic ultrasonic bath (Ultrasonic cleaner for Fungilab U.S. 1.6).

Water with identical volume (for $Na^+$-dependent fluorescence change) or sucrose with identical volume and concentration (for melibiose-dependent fluorescence change) was applied for the control to correct for the dilution effect.

### 3.6.2 Fluorescence resonance energy transfer (FRET)

FRET signal from Trp to $Dns^2$-S-Gal (thereafter $D^2G$) was obtained by setting the excitation wavelength at 290 nm, with a half-bandwidth of 5 nm and collecting the emission spectrum in the interval310-570 nm, in 0.1 M KPi, 0.1 M KCl, pH7.0 with the addition of substrates as described. The emission spectrum of $D^2G$by direct excitation was obtained by setting the excitation wavelength at 335 nm and collecting the emission spectrum in the interval400-570 nm, in 0.1 M KPi, 0.1 M KCl, pH7.0 with the addition of substrates as described.

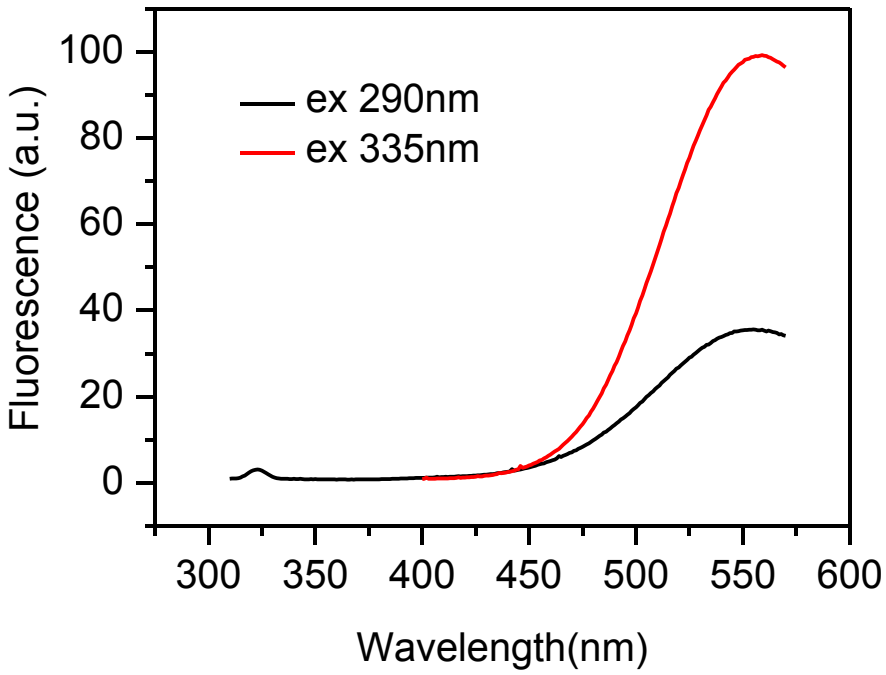



**Figure 3.8** Emission spectrum of $D^2G$ in the buffer containing 0.1 M KPi, 0.1 M KCl, pH 7.0 by illuminating at 290 nm (black) or 335 nm (red).

The FRET signal arising from membrane vesicles was collected in the interval400-570 nm due to the intense fluorescence of the total protein content below 400 nm. 100 µg of protein/ml of membrane vesicles or 30 µg of protein/ml of proteoliposomes was used for the FRET experiments. Before starting the measures, the ISO membrane vesicles and proteoliposome samples were subjected to 30 s sonication in a sonic ultrasonic bath (Ultrasonic cleaner for Fungilab U.S. 1.6).As sonication will lead that a part of RSO vesicles converts into ISO orientation, the RSO membrane vesicles preparations cannot be sonicated. The concentration of protein in liposome or vesicles was determined using the Lowry method (Lowry et al. 1951) by using serum albumin as standard.

When illuminated at either 290 nm or 335 nm, $D^2G$ dissolved in buffer emits fluorescence with a maximum intensity at 558 nm (Figure 3.8). Contributions due to the

emission of free $D^2G$were corrected by subtracting the emission spectrum obtained in the buffer in the absence of protein from the protein-containing buffer. The subtraction factor was adjusted to remove completely the $D^2G$ band at 558 nm.

3.6.3 Measurements of the $Na^+$-activation constant and melibiose inhibitory constant using the $D^2G$ fluorescence assay

Experiments were carried out either on RSO or ISO membrane vesicles (100 μgof protein/ml) in nominally $Na^+$-free, 0.1 M KPi buffer, 0.1 M KCl, pH 7.0. Variations of the $D^2G$ FRET signal intensity at 460 nm (bandwidth, 5 nm) upon excitation at 290 nm (bandwidth, 5 nm) were recorded as a function of time as $Na^+$ or melibiose were added stepwise as described (Guan et al. 2011).

i) $Na^+$-activation constant for $D^2G$ FRET

All samples were initially incubated in $Na^+$-free buffer containing $D^2G$ at 10μM, and the dependence of the intensity of the $D^2G$ FRET signal on $Na^+$ concentration was analyzed following stepwise addition of NaCl to the medium at 60-sec interval. Water with identical volume was applied for the control to correct for the dilution effect.The increase of the fluorescence intensity relative to the signal recorded in $Na^+$-free buffer was plotted as a function of cation concentration present in the solution, corrected for the dilution effect and expressed as % of the maximal increase in fluorescence signal. The $K_{0.5}$ $Na^+$ was determined by fitting the data with a hyperbolic equation.

ii) $K_{0.5}$ for melibiose displacement of bound $D^2G$

All samples were initially incubated in a medium containing 20 mM NaCl and 10 μM $D^2G$. Next, stepwise addition of melibiose (every 60sec) was undertaken to measure the extent of $D^2G$ fluorescent signal extinction resulting from the displacement of bound $D^2G$ by the competing melibiose. Sucrose with identical volume and concentration was applied for the control to correct for the dilution effect. The $K_{0.5}$ for melibiose was calculated by fitting the data expressed as % of the maximal initial $D^2G$ intensity with a hyperbolic equation.

3.6.4 Kinetic determination of MINAS labeling

2 - (4'-maleimidylanilino)-naphthalene-6-sulfonic acid (MIANS, Molecular Probes, Inc., Eugene, OR) is an environment sensitive fluorophore. To get fluorescence, MIANS must be reacted with thiols that are located in hydrophobic sites. Therefore, MIANS can be used as probe to explore the environment of Cys residues.

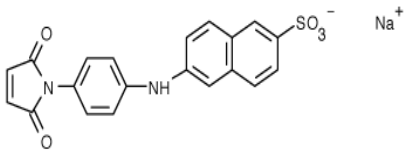



Figure 3.9Structure of MIANS

The concentration of MIANS was determined in dimethyl sulfoxide (DMSO) by measuring the absorbance at 322 nm using an extinction coefficient of 17000 $cm^{-1}$·$M^{-1}$(Haugland 1994). To record the MIANS emission spectrum, the protein (~20 μg/ml) was first preincubated for 10 min in 0.1M KPi, 0.1 M KCl, pH 7.0 with the addition of substrates as indicated. MIANS (5-fold molar excess of MIANS; 0.07% (v/v) DMSO final concentration), was added, the sample was illuminated at 325 nm(bandwidth, 5 nm), and the emission spectrum was collected in the interval 310-570 nm. To obtain the time course of MIANS labeling, the protein sample was excited at 325 nm(bandwidth, 5 nm), MIANS was added to the sample, and the emission intensity was recorded at 415nm(bandwidth, 5 nm) as function of time (Figure 3.10). Before starting the measurements, MelB reconstituted in liposomes was subjected to a 30 s sonication in a sonic ultrasonic bath (Ultrasonic cleaner for Fungilab U.S. 1.6). The half-time of the MIANS reaction was calculated by fitting the data with an exponential equation (Origin 8.0, OriginLab).

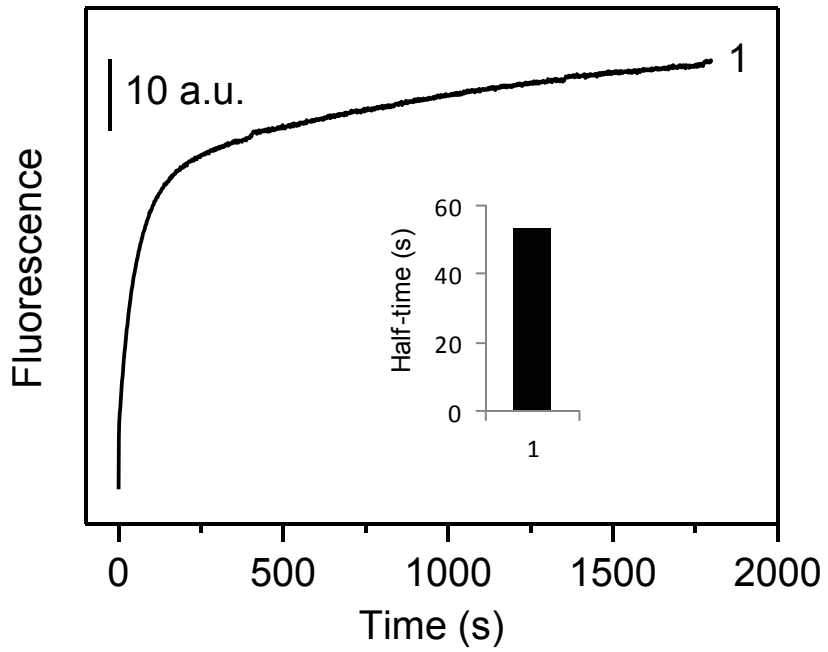



**Figure 3.10** the time course of MIANS labeling of R139C mutant reconstituted in lipid liposomes. The sample containing μg 20 g/ml R139C in 0.1 M KPi, 0.1 M KCl, pH 7.0 was illumated at 325 nm (bandwidth, 5 nm), the emission spectrum was recorded at 415 nm (bandwidth, 5 nm) as function of time. Insert: half-time of MIANS labeling. (See text)

## 3.7 X-ray crystallographic method

The MelB purification and crystallization procedure has been summarized in Figure 3.11.

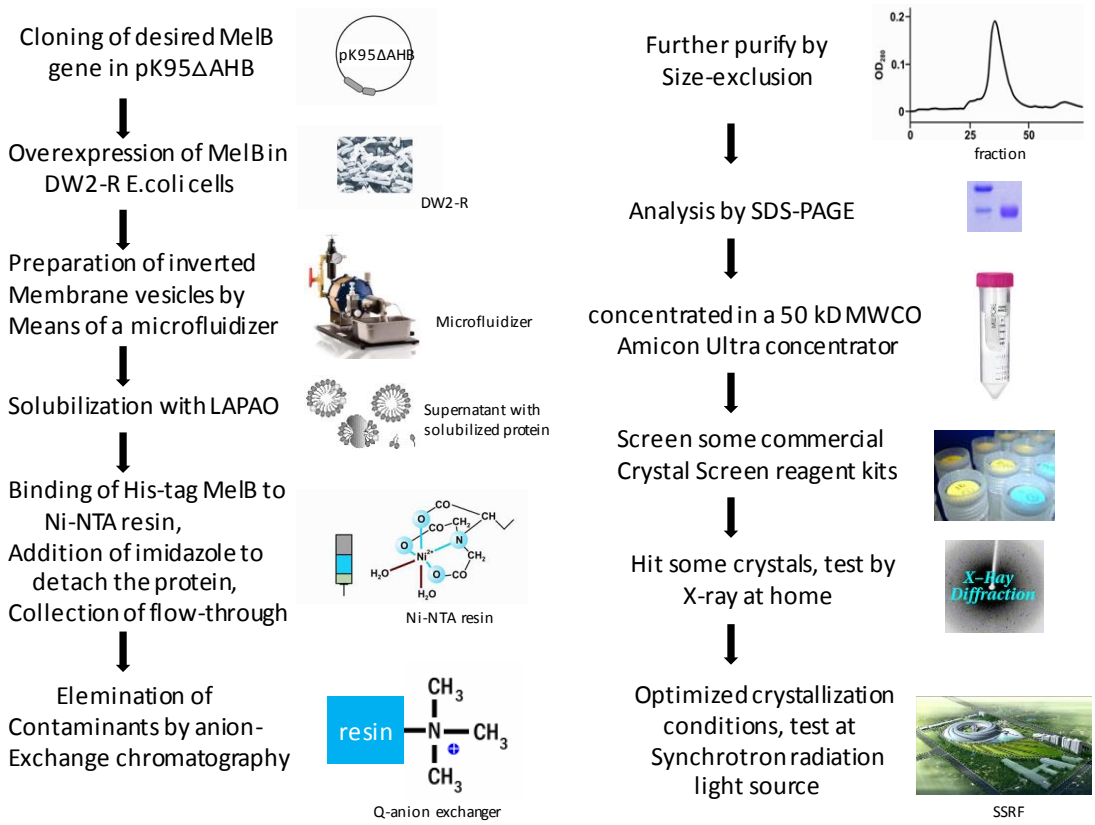



**Figure 3.11** Flow chart of the protein purification and crystallization procedure

### 3.7.1 Protein preparation

MelB and the mutants were cloned, expressed and purified similarly as described in **3.2**. The protein eluted from the Ni-NTA affinity resin column was further purified on a Superdex 200 GL size exclusion column (GE Healthcare) equilibrated with 20 mM Tris, 100 mM NaCl, 5% (v/v) Glycerol, 10 mM melibiose, 0.017 % (w/v) DDM, and 5 mM 2-mercaptoethanol at pH 8.0 using an ÄKTA purifier system. The peak fraction was collected for crystallization. The Purified MelB was concentrated using a 50kDaAmicon Ultra-4 (Millipore) filter and washed against the crystallization buffer composed of 20mMTris pH7.5, 100 mM NaCl, 10 mM melibiose, 0.017% (w/v) DDM.

3.7.2 Silver staining

A protocol from www.protana.com/services/protocols/default.asp was used in this study. Table 3.7 lists the buffers used.

Table 3.7 Buffers for silver staining.

| Buffer | Composition |
|---|---|
| Fixer | 50% (v/v) methanol, 12% (v/v) acetic acid, 0.05% (v/v) formalin |
| Wash | 35% (v/v) ethanol |
| Sensitizing | 0.02% (w/v) $Na_2S_2O_3$ |
| Silver nitrate | 0.2% (w/v) $AgNO_3$, 0.076% (v/v) formalin |
| Developer | 6% (w/v) $Na_2CO_3$, 0.05% (v/v) formalin, 0.0004% (w/v) $Na_2S_2O_3$ |
| Stop solution | 50% (v/v) methanol, 12% (v/v) acetic acid |

Experiment procedure:

1)Fix gel: fixing buffer for 2 hrs or overnight

2)Wash gel: washing buffer for 20 min.

3)Wash gel: washing buffer for 20 min.

4)Wash gel: washing buffer for 20 min.

5)Sensitize gel: sensitizing buffer for 2 min.

6)Wash gel: $H_2O$ for 5 min.

7)Wash gel: $H_2O$ for 5 min.

8)Wash gel: $H_2O$ for 5 min.

9)Stain gel: staining buffer for 20 min.

10) Wash gel: $H_2O$ for 1 min.

11) Wash gel: $H_2O$ for 1 min.

12) Develop gel: developing buffer; may need 1-5 min.

13) Stop staining: stopping buffer for 5 min.

14) Leave the gel at 4℃ in 1% (v/v)acetic acid; Scanning gel.

3.7.3 Native PAGE

Native polyacrylamide gel electrophoresis (Native PAGE) is made without SDS and mercaptoethanol, therefore allowing the separation of proteins in their native conformation. The nonreducing and nondenaturing environment of native PAGE allows the detection of protein homogeneity.

Table 3.8 Work buffers for native PAGE

| No. | Work buffers |
|---|---|
| 1 | 40% acrylamide/bis-acrylamide solution 29:1 (Bio-RAD) |
| 2 | 4×separating buffer (1.5 M Tris-HCl, pH 8.8) |
| 3 | 4×stacking buffer (0.5 M Tris-HCl, pH 6.8) |
| 4 | 10×running buffer (pH8.8 Tris-Gly) |
| 5 | 2×sample buffer: 1.25 ml pH 6.8, 0.5 M Tris-HCl+3.0 ml glycerol, 0.2 ml 0.5% (w/v) bromophenol blue+5.5 ml $ddH_2O$ |
| 6 | 10% ammonium persulfate (APS) |
| 7 | 0.25% (w/v) coomassie blue R250: Coomassie blue R-250 2.5g, methanol 450 ml, acetic acid 100 ml, $dH_2O$ 450 ml |

Table 3.9 Native gel preparation

| | 17% separating gel (10 ml) | 4% stacking gel (4 ml) |
|---|---|---|
| 40% acrylamide/bis-acrylamide | 4.25 ml | 0.5 ml |
| 4×separating buffer | 2.5 ml | |
| 4×stacking buffer | | 1.25 ml |
| Water | 3.25 ml | 2.7 ml |
| 10 % APS | 35 µl | 14 µl |
| TEMED | 15 µl | 6 µl |
| 10 % (w/v) DDM | 17 µl | 6.8 µl |

Table 3.8 shows the work buffer for native PAGE. Before used, DDM was added to running buffer to 0.1% (w/v).Native gel was prepared as shown in Table 3.9. Native PAGE was performed at 4℃ and the gels were run at 100 V. When the indicator went into the separating gel, at the voltage was increased to 160 V until the indicator ran out of the gel. As MelB is an alkaline protein (PI>9), the electrodes were reversed to let the protein come down.

3.7.4 Measurement of tryptophan fluorescence and Trp→$D^2G$ FRET in detergent containing solution

Fluorescence measurements were performed at 20℃with a UV-visible QuantaMaster™ spectrofluorometer and processed with the Felix 32 software (Photon Technology International). Trp fluorescence spectra were obtained by setting the excitation wavelength at 290 nm (half-bandwidth of 5 nm) and collecting the emission spectrum in the interval 290-400 nm, in 20mM Tris-HCl pH 7.0, 5% (v/v) Glycerol, 0.017% (w/v) DDM with the addition of 10 mM NaCl and 10 mM melibiose as indicated. Trp→$D^2G$ FRET spectra were obtained by setting the excitation wavelength at 290 nm (half-bandwidth of 5 nm) and collecting the emission spectrum in the interval 290-570 nm, in 20m M Tris-HCl pH 7.0, 5% (v/v) Glycerol, 0.017% (w/v) DDM with the addition of 15 µm $D^2G$ and 15 mM NaCl as indicated.

3.7.5 Preparation of selenomethionine (Se-Met) labeled R149C MelB

The incorporation of selenomethionine into proteins in place of methionine helps solve the problem of phase determination in X-ray crystallography.

For the protein labeled with seleno-L-methionine, transformed bacteria (B834 DE3) were grown in LeMaster medium, and the protein was purified as above with 10 mM DTT added to all buffers. LeMaster medium was prepared as described (LeMaster and Richards 1985)

3.7.5.1 LeMaster medium preparation:

1) Make pre-medium (4 L):

The mixture of:alanine 1.6g, arginine HCl 1.86g, aspartic acid 1.28g, cysteine 0.1g, glutamic acid 2.14g, glutamine 1.06g, glycine 1.7g, histidine 0.2g, isoleucine 0.736g, leucine 0.736g, lysine HCl 1.34g, phenylalanine 0.42g, proline 0.32g, serine 6.66g, threonine 0.736g, tyrosine 0.54g, valine 0.736g, adenine 1.6g, guanisine 2.14g, thymine 0.54g, uracil 1.6g, Na acetate 4.8g, succinic acid 4.8g, were mixed with 400 mlof 10×M9 (75 g/l $Na_2HPO_4$, 30 g/l $KH_2PO_4$, 10 g/l $NH_4Cl$, 5 g/l NaCl). Adjust volume to 4 Lwith water.

2) Autoclave this “pre-medium”.

3) Cool it down at room temperature. Check that the medium is at pH 7.5. Adjust if necessary.

4) Distribute the pre-medium to four 1L.

5) Autoclave them.

6) Cool to room temperature. Make the non-autoclavable portion below:

Nonautoclavable portion:

1) Prepare 80 mlof 20 mg/ml glucose, 20 ml 0.5M $MgSO_4$, 20 ml 0.05M $CaCl_2$, and 100 ml 34mg/mlthiamine.

2) Filter using a 0.2 μm filter.

3) Add 20 mlglucose, 2 ml $MgSO_4$, 2 ml $CaCl_2$, and 10 mlthiamine to each liter of the autoclaved pre-medium.

4) Before inoculation, add 50 mg DL-SeMet to each liter to complete the master medium.

3.7.5.2 The protocol for large scale expression (1 L):

1) Fresh transform of the expression vector into B834. Incubate the plates overnight at 37 ℃.

2) Inoculate 3 ml culture overnight (LeMaster medium) with one single colony. Grow at 37 ℃with 240 rpm shaking.

3) Inoculate 50 ml of LeMaster medium with the 3 ml overnight culture. Grow at 37 ℃with 240 rpm shaking.

4) Inoculate 1 L Master medium with 50 ml of the overnight culture. Grow at 37 ℃with 240 rpm shaking for about 20 h.

3.7.6 Crystallization and X-ray diffraction

MelB wild type and other mutated forms of MelB purified in different detergents were subjected to extensive crystallization screens consisting of PEG400 and PEG200 at intervals of 1.0 pH unit (from pH3.5 to pH9.0), as well as four major commercial crystal screening kits: MemGold, MemSys&MemStart, JCSG-plus and PACT (Molecular Dimension). Additives such as organics, detergents, and mono- and divalent salts (Hampton Research) were screened to improve crystal order and packing. Crystallization drops were assembled in a hanging drop fashion in 24-well plate or sitting drop fashion in 96-well plate.

MelB crystals were frozen in liquid nitrogen, or using a cold nitrogen stream. Crystals were examined by X-ray diffraction home light source or, more often, with a synchrotron source at one of the following beamlines: BM16, ID23-1, and ID14-4 at the European Synchrotron Radiation Facility (ESRF), located in Grenoble, France.

Diffraction data were processed by using the HKL2000 program (Otwinowski and Minor 1997) or Mosflm (Leslie and Powell 2007).

# PART I The key role of Arg149 in function of melibiose permease

# 4 Results and Discussions of Part I

In a search for important amino acids for sugar transport in MelB, it was demonstrated the importance of Arg149, located presumably in the cytoplasmic end of TM5.The substitution of Arg149 with Cys was found to be completely inactive for sugar transport and sugar binding in right-side-out (RSO) membrane vesicles, while the replacement of Arg149 by Lys and Gln retained some transport activity (Abdel-Dayem et al. 2003). Furthermore, neither positive charge nor negative charge, labeled to Cys149, cannot recover transport activity (Abdel-Dayem et al. 2003).

With the data at hand, it could not be decided whether R149C inactivation results from an overall alteration of the transporter structure or from a defect restricted to a given step of the transport cycle. To decipher the structural and/or catalytic defect responsible for the transport inactivation in R149C, we present in this work spectroscopic and biochemical data reporting on the substrate-induced conformational changes and protein accessibility in vesicles and in proteoliposomes. Fourier transform infrared difference ($IR_{diff}$) spectroscopy (Granell et al. 2010; Leon et al. 2005) has been used to bring information on the global transporter structure and on the substrate-induced conformational changes triggered by the binding of $Na^+$ and melibiose to the Arg149 mutants. On the other hand, fluorescence spectroscopy has been used to complement this structural analysis, as well as to determine the protein orientation and the binding site accessibility in different vesicles using a fluorescent probe for Cys residues and a fluorescent sugar analog (Abdel-Dayem et al. 2003; Mus-Veteau et al. 1995; Maehrel et al. 1998).

Like in previous works, R149C, R149K, and R149Q, mutants involved in this study were constructed using the Cys-less MelB as a genetic background (Abdel-Dayem et al. 2003). Therefore, in the following the behavior of these mutants will be compared to Cys-less, which displays properties very similar to those of WT (Weissborn et al. 1997).

## 4.1 R149C reconstituted in proteoliposomes

To find out which is the defect introduced by the mutation, we first analyzed whether the mutant can bind the substrates once reconstituted into liposomes, where the MelB orientation is inside-out (see Meyer-Lipp et al. 2006 and below).

4.1.1 Analysis of structural components of mutants

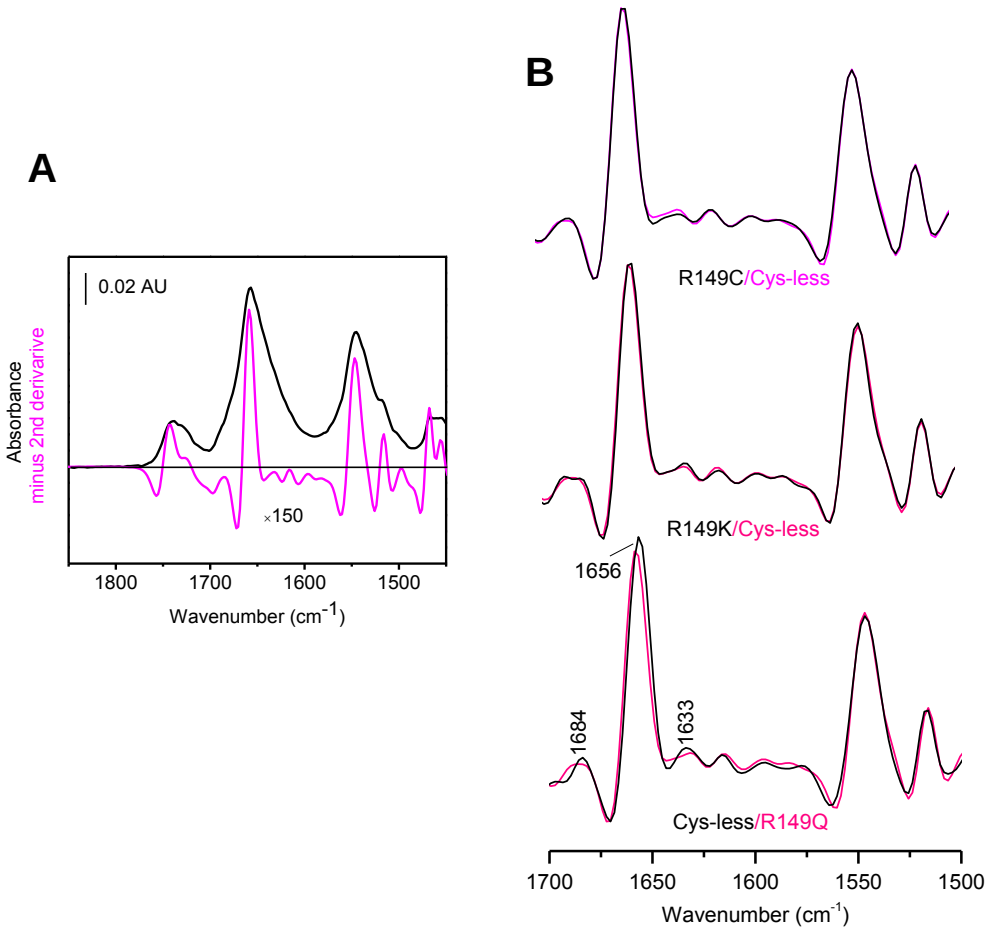



**Figure4.1**Structural comparison of MelB mutants using IR spectroscopy of hydrated samples. (A) Absorbance IR spectrum at 4 $cm^{-1}$ resolution of the Cys-less MelB mutant in a hydrated film with the buffer contribution subtracted (black). The second derivative at 8 $cm^{-1}$ resolution (red) provides more resolved details. (B) Second derivative in the amide I/II region for R149C, R149K, and R149Q compared to Cys-less (shown in red).

Prior to analyze the substrate-induced conformational changes by infrared difference spectroscopy, we looked for any structural change of the mutants carrying a Cys, Lys or Gln in place of Arg149. This information can be extracted from of the second derivative of the IR absorbance spectra in the structure-sensitive amide I and II regions measured in the absence of substrates (Figure 4.1). It is observed that Cys-less, R149C and R149K show almost identical spectra, and thus presumably a highly conserved structure, while R149Q shows few small band-shifts in comparison to Cys-less, suggesting some minor local structural alterations. These results discard global conformational alteration, like protein denaturation or aggregation, as a possible origin of the absence of transport activity for R149C.

4.1.2 Infrared difference spectra

It has been previously shown that measuring the substrate-induced $IR_{diff}$ spectraon purified WT or Cys-less transporters reconstituted in liposomes provides a means to assess whether or not the substrates bind to the transporters and trigger conformational changes (Leon et al. 2005; Leon et al. 2009). We first compare the $Na^+$-induced $IR_{diff}$ spectra ($Na^+$-MelB versus $H^+$-MelB) recorded from R149C or from Cys-less. Figure 4.2A shows that a very significant signal is recorded from R149C. Application of a spectral correlation analysis (Granell et al. 2010)to the signal of the two permeases shows that R149C has an intensity of~90% and a spectral similarity of ~44% with respect to Cys-less (Figure4.2C, left side).At this stage, it is important to recall that no measurable signal is detected in MelB mutants with impaired capacity to bind $Na^+$(Granell et al. 2010). These data proof that $Na^+$ not only binds to the R149C mutant, but its binding induces conformational changes comparable to those occurring in the Cys-less. One can however note that the R149C and Cys-less spectra display a few spectral differences, especially in the intensity of some peaks located in the amide I region (see Figure4.2A and Figure 4.3A). Namely, the peaks at 1657 $cm^{-1}$ and 1663 $cm^{-1}$, proposed to arise from transmembrane α-helices (Leon et al. 2006) are absent and strongly reduced in intensity in R149C.In contrast, the peak at 1651 $cm^{-1}$ arising from another α-helix structure (Leon et al. 2006), is enhanced in R149C and slightly shifted.

Peaks assigned to environment changes in deprotonated Asp side chains, at around 1404 $cm^{-1}$, appear also altered in the mutant. Therefore, $Na^+$ binding to R149C gives rise to slightly modified conformational changes as compared to those of the active Cys-less permease.

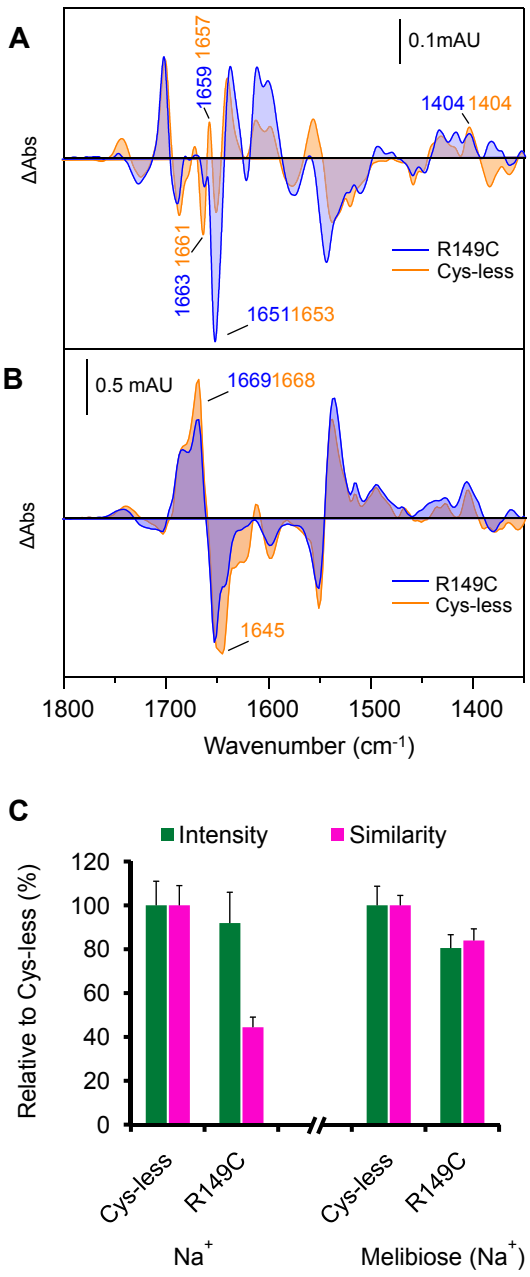


**Figure4.2** Substrate-induced conformational changes. Substrate-induced $IR_{diff}$ spectra at 4 $cm^{-1}$ resolution of MelB Cys-less (orange) and R149C (blue) at 25° C and pH 6.6 (20 mM MES and 100 mM KCl), normalized to the amount of probed protein (see Figure 4.1). The buffer exchange protocol and data acquisition are described in *Materials and Methods*. (**A**) 10 mM $Na^+$-induced$IR_{diff}$ spectra.(**B**) 50 mM melibiose-induced $IR_{diff}$ spectra in the presence of 10 mM$Na^+$. (**C)** Spectral similarity and intensity of the substrate-induced $IR_{diff}$ spectra. The error bar corresponds to one standard error of the mean of three independent spectra.

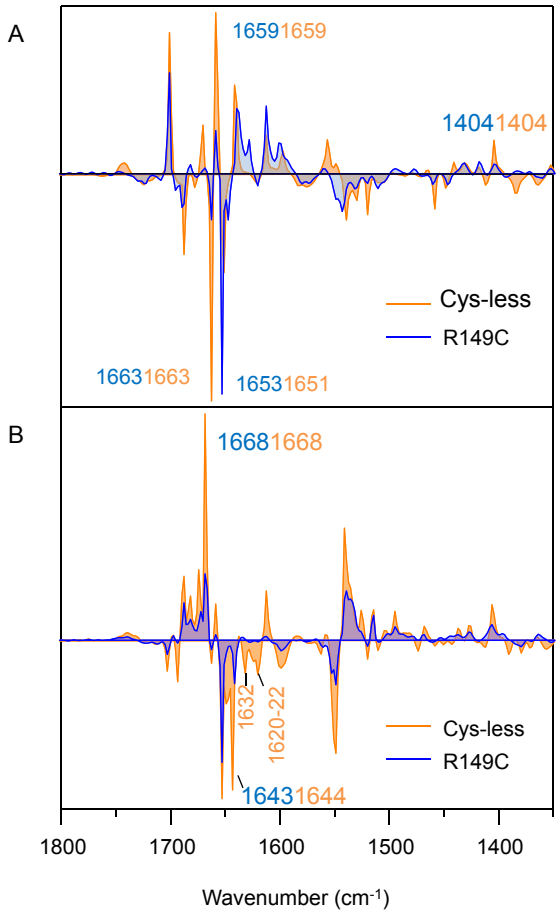



**Figure4.3**(A) Deconvoluted spectra of the $Na^+$-induced MelB difference spectra of Figure 2A. (B) Deconvoluted spectra of the melibiose-induced MelB difference spectra of Figure 2B. Deconvolution (Lorenz-Fonfria and Padros 2005) was performed using a Lorentzian band of 7 $cm^{-1}$ width.

The R149C spectral response to addition of 10 mM melibiose in the presence of $Na^+$ was next investigated (Figure4.2B). The spectrum exhibits an overall shape comparable

to that of Cys-less, but with decreased intensity (82% of Cys-less), suggesting a reduced affinity for the sugar. This interpretation is likely, since performing equivalent experiments with 50 mM melibiose, gives rise to a difference spectrum showing about 80% intensity of that of Cys-less and 84% similarity (Figure 4.2B and C, right side). This observation implies that, like sodium, melibiose can bind to the non-transporting R149C mutant. There are some variations in the amide I region, including the decrease of the peak at 1668-1669 $cm^{-1}$, that may correspond to α-helix or turn structure (Leon et al. 2006) and the decreased intensity of the negative peak at 1645 $cm^{-1}$, assigned to β-sheet, $3_{10}$ helices or open loops (Leon et al 2006). These and other changes are better resolved after band narrowing by maximum entropy deconvolution (Figure 4.3B).

To evaluate the effect of the charge of Arg149, we analyzed two additional mutants, namely R149K and R149Q. A previous report showed that both remain functional, although with a decreased transport efficiency (Abdel-Dayem et al. 2003). The $Na^+$-induced $IR_{diff}$ spectra of these 2 mutants are comparable to that of Cys-less, with a similarity of 57% for R149K and 69% for R149Q (Figure 4.4A-C), indicating that both mutants bind $Na^+$. A more thoughtful comparison shows that $Na^+$-induced $IR_{diff}$ spectra are very similar between R149K and R149Q (>90% similarity), and slightly more similar to Cys-less than to R149C (Figure 4.4D). It suggests that mutation of Arg149 to Lys or Gln is more conservative than its mutation to Cys from the point of view of $Na^+$ binding. The sugar-induced difference spectra in the presence of $Na^+$ recorded from R149K and R149Q show also a high similarity to Cys-less(between 70 and 80%, Figure 4.5C), confirming these two mutants retain native-like melibiose-induced conformational changes. However, the spectral intensity is significantly lower than that of Cys-less and R149C, suggesting that the affinity for melibiose might be reduced in both R149K and R149Q. In agreement, a spectral correlation matrix suggests that for melibiose binding, mutation of Arg149 to Cys is more conservative than its mutation to Lys or Gln (Figure 4.5D).

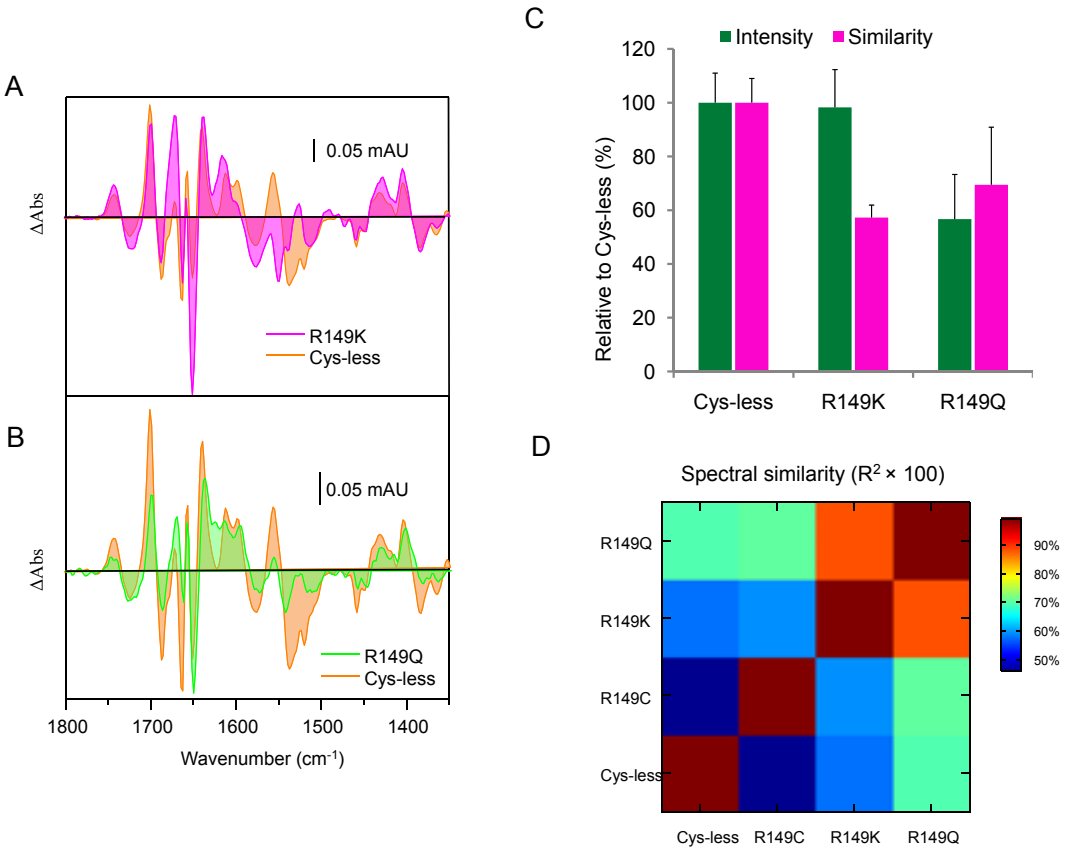



**Figure 4.4** $Na^+$-induced $IR_{diff}$ spectra at 4 $cm^{-1}$ resolution of MelB R149K (A), and R149Q mutants (B), compared to Cys-less. The difference spectra were obtained by replacing $Na^+$-free buffer by a buffer medium containing 10 mM $Na^+$ at 25ºC. The buffer exchange protocol and data acquisition are described in ***Materials and Methods***. The average difference spectrum from three independent experiments is shown. (C) Spectral similarity and intensity of the $IR_{diff}$ spectra of R149K and R149Q compared to the Cys-less (see ***Materials and Methods***). The error bar corresponds to one standard error of the mean. (D) Representation of all possible comparisons between pair of $Na^+$-induced $IR_{diff}$ spectra, with a color code indicating their degree of similarity, measured as $R^2$x100.

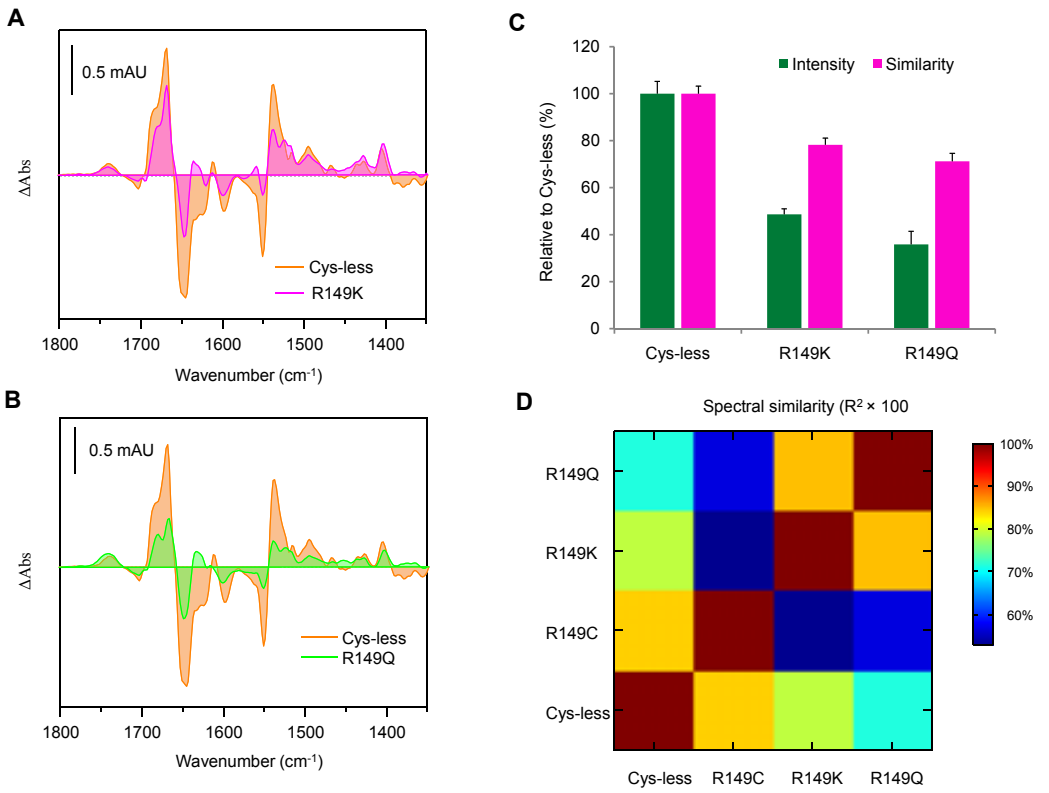



**Figure 4.5**Melibiose-induced $IR_{diff}$ spectra at 4 $cm^{-1}$ resolution of MelB R149K (A) and R149Q (B) in the present of 10 mM $Na^+$. The difference spectra were obtained by replacing melibiose-free buffer by a buffer medium containing 50 mM melibiose at 25 ºC in the present of 10 mM $Na^+$. The buffer exchange protocol and data acquisition are shown in ***Materials and Methods***. The average difference spectrum from three independent experiments is shown. (C) Spectral similarity and intensity of the $IR_{diff}$ spectra of R149K and R149Q compared to Cys-less (see ***Materials and Methods***). The error bar corresponds to one standard error of the mean. (D) Representation of all possible comparisons between pair of melibiose-induced $IR_{diff}$ spectra, with a color code indicating their degree of similarity, measured as $R^2$x100.

4.1.3 Intrinsic fluorescence spectra and fluorescence resonance energy transfer in R149C

To give more strength to the conclusion drawn by $IR_{diff}$ spectroscopy that purified R149C in proteoliposomes does retain the capacity to bind $Na^+$ and melibiose and triggers associated conformational changes, we undertook complementary studies with the same proteoliposomes. We used intrinsic fluorescence and Fluorescence Resonance Energy Transfer (FRET) spectroscopy, since these techniques were previously shown to bring significant information on these MelB properties (Mus-Veteau et al. 1995; Maehrel et al. 1998; Cordat et al. 2000).

In the absence of $Na^+$, incubation of R149C with melibiose (10 mM) leads to an intrinsic fluorescence increase (Figure 4.5A). This signal is further enhanced after $Na^+$ addition because of the induced affinity increase for melibiose upon $Na^+$ binding (Mus-Veteau et al. 1995). These substrate-induced changes are about half the intensity observed in Cys-less, in concordance with the reduced melibiose affinity of R149C deduced from $IR_{diff}$ experiments. The results not only confirm that R149C retains melibiose and $Na^+$ binding capacity, but also indicates that the reciprocal activation of the binding of one substrate by the other is maintained in R149C proteoliposomes, even if it looks somewhat altered, in keeping with the $IR_{diff}$ spectroscopy results.

Substrate binding to proteoliposomes was also assessed using FRET from Trp residues to the fluorescent sugar analog $D^2G$ (Maehrel et al. 1998). Figure 4.6 shows that, as $D^2G$ is added, the Trp fluorescence signal (band around 330 nm) decreased in part because of FRET from Trp side chains to $D^2G$ bound to the sugar-binding site (Maehrel et al. 1998). Concomitantly, the fluorescence emission from the protein-bound fluorescent sugar appears with a maximum at 460 nm. Subsequent addition of $Na^+$ enhances the affinity for the sugar analog and increases its bound fraction, leading to a further increase of FRET (Maehrel et al. 1998). In Figure 4.6B, C it is also seen that addition of excess melibiose decreases the FRET signal, indicating that the fluorescent sugar and melibiose compete for the same binding site. The $Na^+$-dependent $D^2G$ emitted fluorescence can be obtained from the difference between the spectra recorded after and before addition of NaCl (Figure 4.6D). It is important to note that the $\lambda_{max}$ of $D^2G$

molecules bound to R149C is similar to that bound to the Cys-less MelB, suggesting the probe experiences an equivalent hydrophobic environment when bound to the two permeases. Overall, these data show that the fluorescent sugar analog can bind to the R149C mutant, and that $Na^+$ retains the capacity to increase the affinity of the mutant for the sugar.

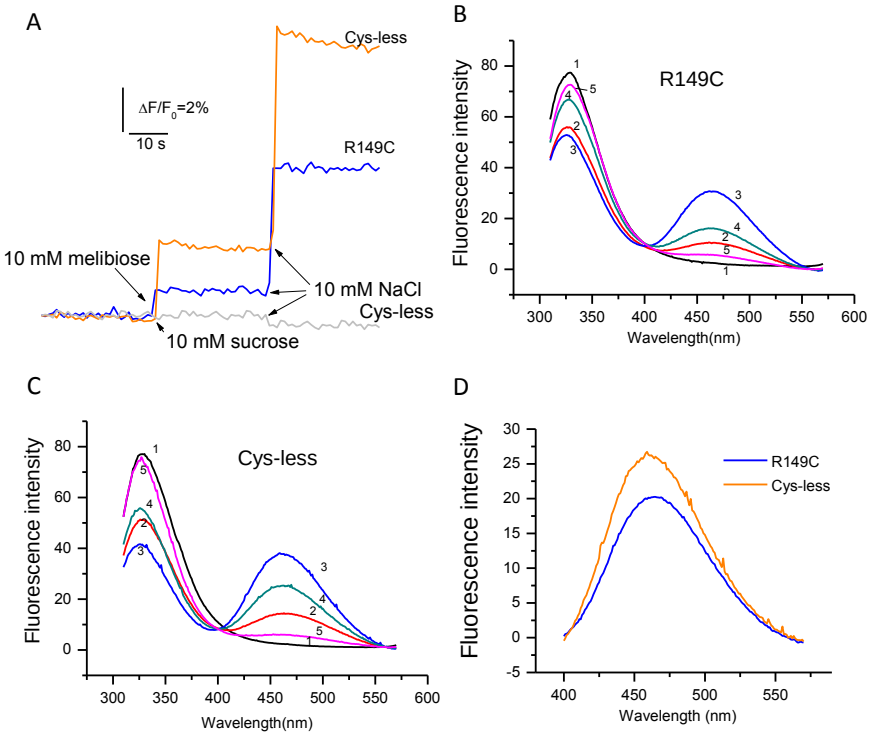



**Figure 4.6** Substrate-induced fluorescence changes of MelB in proteoliposomes and FRET changes. (A) Tryptophan fluorescence changes in proteoliposomes ($\lambda_{ex}$ = 290nm; half-bandwidth = 5 nm, and $\lambda_{em}$ =325 nm; half-bandwidth = 5 nm) containing purified R149C (or Cys-less) at 20 µg/mL in 100 mM KPi after the addition of sugar and $Na^+$ to a final concentration of 10 mM (see arrows). Sucrose addition is included as a negative control. (B) FRET signal between R149C tryptophans and the fluorescent sugar analog $D^2G$ ($\lambda_{ex}$ = 290 nm)in nominally $Na^+$-free 100 mM potassium phosphate and 100 mMKCl. The emission fluorescence was recorded before (trace 1) and after the consecutive additions of 16 µM $D^2G$ (trace 2), 10 mM NaCl (trace 3), 10 mM melibiose (trace 4) and150 mM melibiose (trace 5). Each spectrum is the mean of three scans. (C) FRET signal for Cys-less, under the same conditions and substrate additions as for R149C. (D) $Na^+$-induced FRET signal changes, calculated from the

difference between the fluorescence spectra recorded after (trace 3 in (B) or (C)) and before addition of NaCl (trace 2 in (B) or (C)).

## 4.2 Orientation of the protein

The observations described above proof that the R149C permease does bind sugar in proteoliposomes. Puzzlingly, it was described that this same mutant does not display binding activity in RSO membrane vesicles (Abdel-Dayem et al. 2003).An attractive interpretation is to assume that the difference in binding capacity observed in the two membrane types may be related to an opposite orientation of the MelB transporters in these membranes. In RSO vesicles, as in cells, the periplasmic side of MelB faces the outside medium (Figure 3.1); in proteoliposomes, as in ISO membrane vesicles, MelB has an inverted orientation as the cytoplasmic side faces the outside medium (Meyer-Lipp et al. 2006). To test this interpretation, we undertook a comparative determination of the orientation of R149C permease, either in proteoliposomes or in the more natural environment of RSO and ISO vesicles. In addition, accessibility of the substrate-binding sites in RSO and ISO vesicles was probed by using the fluorescence sugar analog $D^2G$.

To confirm that the orientation of R149C was the one actually expected in the vesicles and liposomes, we used mutants with a single cysteine located either in the cytoplasmic side, or in the periplasmic side. We detected the accessibility of the single cysteine using tetramethylrhodamine-5-maleimide (TMRM), a fluorescent reagent specific for cysteines which can permeate the plasma membrane under the conditions of concentration and time of labeling used here (Nie et al. 2007). It was combined with the previous blocking of all the cysteines exposed to the external medium by the use of a membrane-impermeant thiol reagent, (2-(trimethylammonium)ethyl)methanethiosu-lfonate (MTSET). We performed three control experiments, namely:(i) the mutant I262C, with the Cys located within loop 7-8 in the periplasmic side; (ii) the mutant R139C, with the Cys located within loop 4-5 in the cytoplasmic side and functional properties similar to those of Cys-less (Abdel-Dayem et al. 2003); (iii) Cys-less MelB,

which was used as a negative control since it should not react with TMRM. Finally, we also performed similar experiments forI262C, R139C, and R149C reconstituted into proteoliposomes.

As shown in Figure 4.7A, R149C RSO membrane vesicles pretreated with MTSET show a clear fluorescence signal from TMRM at the expected molecular weight of MelB, implying full reaction of TMRM with Cys149. The same was observed for RSO vesicles not previously treated with MTSET. Since MTSET reacts only with external Cys residues, it is obvious that the R149C single cysteine faces to the inside in RSO vesicles. For the ISO vesicles, no TMRM signal was observed after pretreatment with MTSET (Figure 4.7A), demonstrating that Cys149 faces the external medium in ISO vesicles. Similar results were obtained from the other mutant, R139C, for which the single cysteine is located in the cytoplasmic loop 4-5. On the contrary, the single cysteine of the mutant I262C, which is located at the periplasmic side, cannot be labeled by TMRM in RSO vesicles, whereas it is labeled in ISO vesicles (Figure 4.7A). Therefore, in the RSO vesicles this single cysteine should face to the outside, whereas in ISO vesicles it faces to the inside.

As a negative control, we did not observe any fluorescence signal from TMRM for ISO or RSO vesicles containing the Cys-less MelB (Figure 4.7B). Based on these data we confirmed that the orientation of the protein in the ISO and RSO vesicles is the one expected: MelB in the ISO is inside-out, and in the RSO vesicles it is right-side-out. We performed equivalent experiments to determine the orientation of MelB reconstituted into proteoliposomes (Figure 4.7C). There is no TMRM reaction for R149C or R139C pre-treated with MTSET, whereas for the I262C mutant a clear TMRM signal was observed both before and after MTSET addition. These results indicate that the protein orientation in these proteoliposomes is inside-out, in agreement with a previous report (Meyer-Lipp et al. 2006).

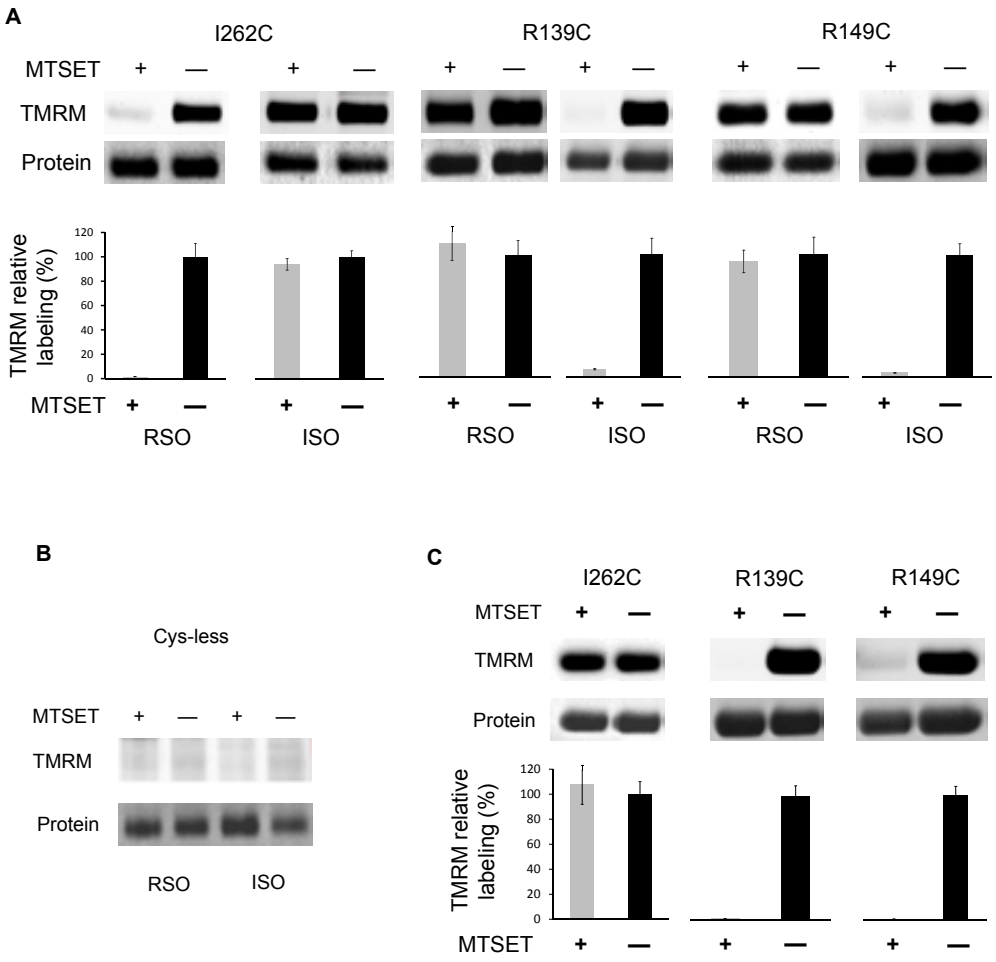



**Figure 4.7**Orientation of R149C in vesicles and proteoliposomes. (**A**) RSO or ISO membrane vesicles containing I262C, R139C, and R149C cysteine mutants were reacted with 40 μM TMRM directly, or after treatment with 0.1 mM MTSET. MelB was purified from the solubilized vesicles and subjected to SDS-PAGE. The coomassie-stained bands report the amount of the loaded MelB protein, and the TMRM reacted with MelB wasimaged on the gel using the TMRM intrinsic fluorescence as described in *Materials and Methods*. Histograms represent the ratio between the intensity of TMRM and protein bands area, expressed relative to the vesicles without treating with MTEST. Each value is the mean of three independent experiments (±S.E.). (**B**) TMRM labeling of MelB Cys-less mutant in ISO and RSO vesicles, following the same protocol as in (A). (**C**) Orientation of I262C, R139C, and R149C in proteoliposomes. The experiments were performed as in (A).

## 4.3 Accessibility of the sugar-binding sites: $Na^+$-induced change of the FRET signal

To study the accessibility of the sugar-binding site in vesicles, we recorded the FRET signal changes of the fluorescent sugar analog $D^2G$ induced by sodium (Maehrel et al. 1998), in both ISO and RSO membrane vesicles (Figure 4.8A to G). No significant or at most a very weak $Na^+$-induced FRET signal was recorded for R149C in RSO vesicles (Figure 4.8E, black line), in complete agreement with previous results (Abdel-Dayem et al. 2003). In that work, R149C mutant not only showed lack of binding of $D^2G$ in RSO as we show here, but also of the labeled sugar analog [3H]p-nitrophenyl-α-D-galacto -pyranoside (NPG) (Abdel-Dayem et al. 2003). Furthermore, we found no significant differences between the fluorescence signal in R149C RSO vesicles and in RSO or ISO vesicles prepared from DW2 cells that do not express MelB, which were used as controls (Figure 4.9). It is therefore evident that sugars cannot bind to R149C RSO vesicles. In contrast to these results, a clear FRET signal was observed in R149C ISO vesicles (Figure 4.8E, red line), similar to that observed for proteoliposomes in Figure 4.6. We also analyzed the responses of RSO and ISO vesicles from cells expressing R139C, I262C, WT, Cys-less, R149K, and R149Q MelB (Figure 4.8). A final control was done by corroborating that the FRET signal indeed decreased and eventually disappeared by adding excess of melibiose.

The ratio between the $Na^+$-dependent FRET signals in ISO and RSO vesicles (ISO/RSO) carrying any single mutant was used as a quantitative index of the relative accessibility of $D^2G$ to the sugar binding site from the periplasmic and cytoplasmic sides in each transporter (Figure 4.8H). For R139C, I262C, WT, Cys-less, and R149K permeases the ratio of ∼1 indicates equal accessibility from either sides. With a ratio of ~2.5, R149Q seems to have a slightly easier access from the periplasmic side than from the cytoplasmic one. Remarkably, the ratio for R149C reaches a value as high as 25. As a whole, these results suggest that the sugar analog cannot reach the binding site

reflecting a complete or almost complete blocking access of the fluorescent sugar to the binding site in R149C RSO membrane vesicles.

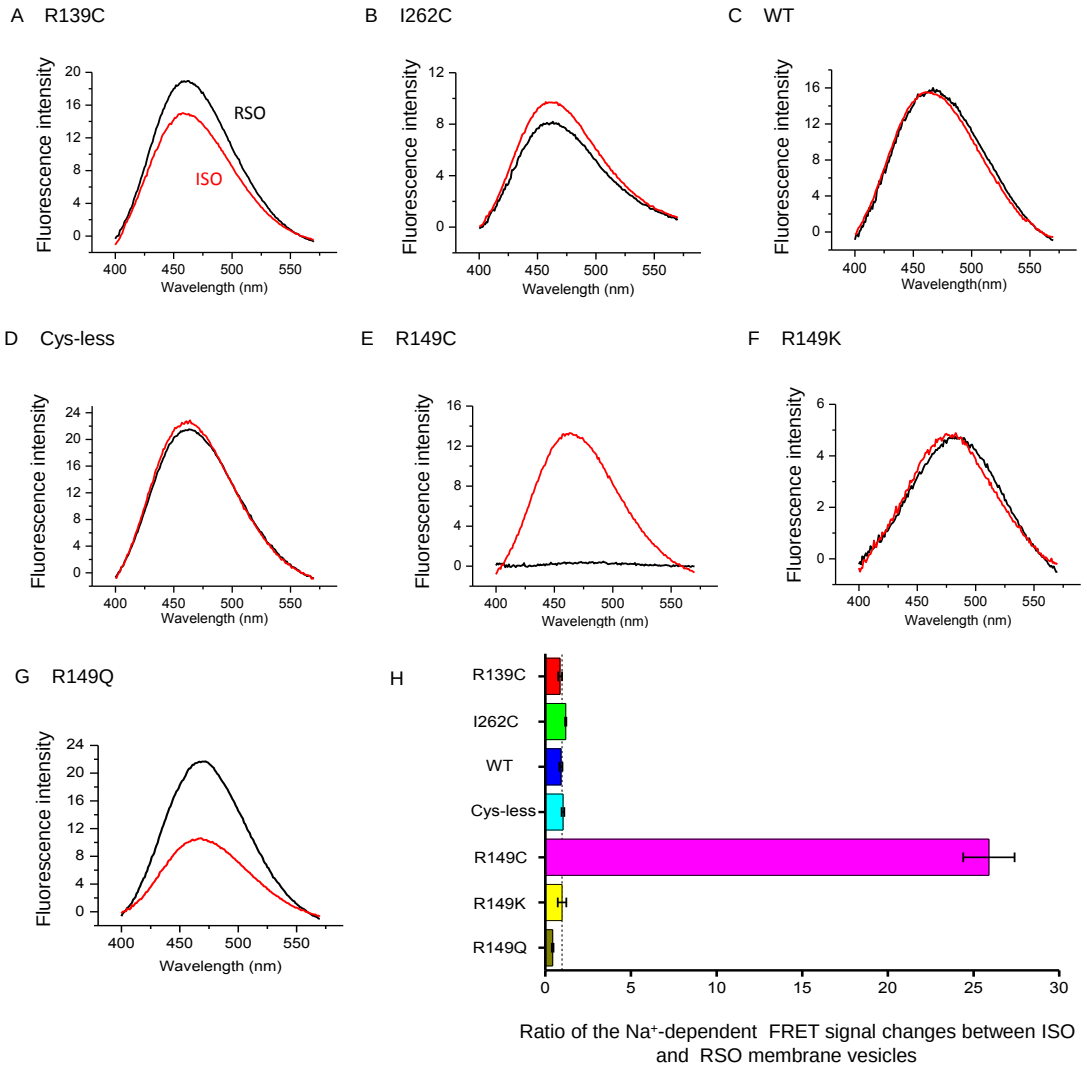



**Figure 4.8** $Na^+$-dependent variation of the FRET signal recorded from ISO and RSO membrane vesicles. ISO (red trace) or RSO membrane vesicles (black trace) were incubated with the sugar analog $D^2G$ (10 µM) and excited at 290 nm (half-bandwidth = 5 nm). The FRET signal from Trp residues to $Dns^2$-S-Gal recorded before the addition of $Na^+$ was subtracted from the FRET signal recorded after the addition of $Na^+$. The $Na^+$-dependent variation of the FRET signal presented in (A)-(G) is the average of at least three independent experiments. (H) Histograms showing the ratio of the $Na^+$-dependent FRET signal changes integrated between 400 and 570 nm, between ISO and RSO membrane vesicles. Values are the mean of three independent experiments (±S.E.).

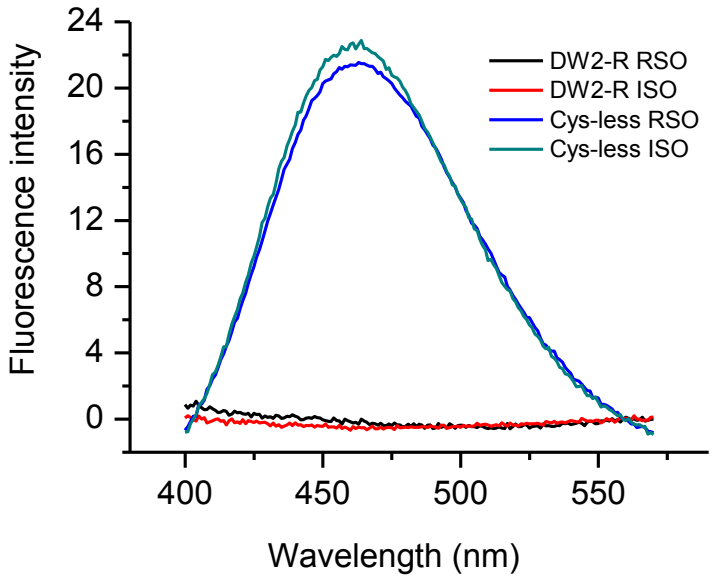



**Figure 4.9** $Na^+$-dependent variation of the FRET signal recorded from Cys-less MelB and *E. coli* DW2-R, which was used to express MelB, membrane vesicles incubated in the presence of the fluorescent sugar analog $D^2G$. The same method was used as presented in Figure 4.6.

To achieve a more complete characterization of the co-substrate binding properties of R149C and their comparison with those of the Cys-less, we used $D^2G$ as a reporter of sugar binding and determined the constants of activation by $Na^+$ ($K_{0.5}$ [$Na^+$]) and of the inhibition by melibiose ($K_{0.5}$ [mel]) of the fluorescent probe binding in RSO and ISO vesicles (Table 4.1). In Cys-less ISO and RSO vesicles, the $K_{0.5}$ [$Na^+$] values are nearly identical and the same value was found in R149C ISO vesicles. This indicates that the $Na^+$ affinity and the $Na^+$-sugar coupling are not modified in R149C. The insignificant $D^2G$ FRET signal R149C RSO vesicles (Figure 4.8E) precluded the determination of the constants in these membranes.

**Table 4.1** Affinity constants for $Na^+$ activation of $D^2G$ binding ($K_{0.5}$ [$Na^+$]) and melibiose inhibition of $D^2G$ binding ($K_{0.5}$ [mel]) to ISO and RSO vesicles

| | Cys-less | | R149C | |
|---|---|---|---|---|
| | ISO | RSO | ISO | RSO |
| $K_{0.5}$ [$Na^+$] (mM) | 0.29 | 0.3 | 0.3 | ND[a] |
| $K_{0.5}$ [mel] (mM) | 3.35 | 2.01 | 4.68 | ND[a] |

a: not determined due to weak signal.

Regarding the inhibition of $D^2G$ binding by melibiose ($K_{0.5}$ [mel]), the values of 3.35 and 2.01 mM in Cys-less ISO and RSO vesicles (Table 4.1) suggest a slightly better binding affinity for melibiose originating from the outer medium than for melibiose originating from the cytoplasmic medium. Interestingly, the $K_{0.5}$ [mel] values found in Cys-less (3.35 mM) and R149C (4.68 mM) ISO vesicles have the same order of magnitude, and suggests that the mutation of Arg into Cys causes a limited reduction of affinity for the sugar. As a whole, these results strongly suggest that the co substrate binding process is only moderately affected by the mutation, in agreement with $IR_{diff}$ and fluorescence data in proteoliposomes. In contrast, the mutation appears to induce a major defect in the reorientation mechanism required for MelB to adopt an outward-facing conformation.

## 4.4 MIANS reactivity

Ligand protection against MIANS (2-(4'-maleimidylanilino)naphthalene-6-sulfonic acid) labeling of MelB provides a more convenient binding assay than flow dialysis, as

it is more sensitive, requires less protein, and is much less dependent upon the affinity of MelB for ligand. Since MIANS is a sulfhydryl-specific probe that is not fluorescent until the maleimide group undergoes chemical reaction, the time course of the reaction can be studied readily. Addition of MIANS to a reaction mixture containing purified MelB reconstituted into liposomes results in a rapid, almost linear increase in fluorescence emission intensity at the beginning (Figure 4.10A, B).

Previous studies showed that replacing Arg149 by Cys leads to completely inactive transporter. Introduction of an additional mutation in the C-terminal helix X (Gly for Val343) of R149C restored sugar transport (Abdel-Dayem et al. 2003). R149C and R149C/V343G mutants contain a single cysteine at the same position. We can detect the MIANS reactivity of these single cysteines to monitor the conformation changes at this position or around this position, which would be affected by substrates. Figure 4.10A and B show the time course of the reaction of R149C and R149C/V343G in the presence of different substrates condition. The half-time of the MIANS reaction for these two mutants, obtained under different substrates condition, are presented in Figure 4.10C, D. As these MIANS signals came from the same Cys, i.e. Cys149, these data reflect the different conformations at the same position in the two different mutants.

The high half time values of R149C indicates that MIANS has lower affinity to this mutant compared to R149C/V343G. This suggests that the additional mutation at Val343 affects the conformation at Arg149. In spite that we don’t know how V343G affects the conformation of R149C, the facts: i) R149C is inactive for sugar transport and sugar binding in right-side-out (RSO) membrane vesicles; ii) R149C/V343G restored sugar transport to some extent; iii) R149C binds sugars in inside-out (ISO) membrane vesicles, suggesting that the conformational changes at Arg149 or around this position are very important for the transport.

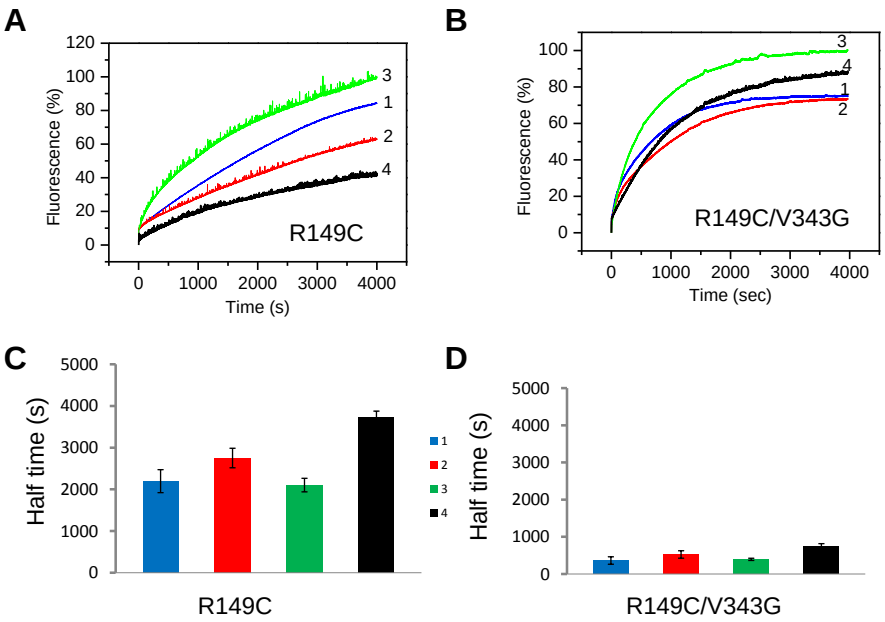



**Figure 4.10** Reactivity of R149C (A), and R149C/V343G (B) MelB reconstituted in liposomes with MIANS and the effect of substrates. Samples (30 µg protein/ml) were illuminated at 325 nm in the absence of any substrates (trace 1), or in the presence of 50 mM NaCl and 50 mM melibiose (trace 2), or 50 mM NaCl (trace 3), or 50 mM melibiose (trace 4) as described when 16 µM MIANS was added to supernatant. The emission spectrum at 415 nm was collected as the function of time. All traces were normalized to the same maximum fluorescence intensity after subtraction of the fluorescence observed without MIANS, and fitted with an exponential equation (Origin 8, ExpAssoc). Values are the mean of three independent experiments (±S.E.).

# 5 General Discussions of Part I

The FTIR and fluorescence results presented in this work show: (i) that R149C binds $Na^+$ and melibiose in proteoliposomes, where the protein has an inverted orientation with respect to that in the cell i.e., with the cytoplasmic side oriented to the exterior medium; (ii) that the conformational changes induced by $Na^+$ or sugar binding to R149C are comparable to the Cys-less protein albeit with a few significant differences in α-helical structures, indicating that the mutation affects the interaction of $Na^+$ and the sugar with the transporter; (iii) that $Na^+$ and the fluorescent sugar analog $D^2G$, as well as melibiose are capable of accessing the R149C substrate-binding sites in inside-out membrane vesicles but not in the right-side-out vesicles. (iv) Furthermore, the MIANS labeling analysis suggests that the replacement of Arg149 by Cys interferes the accessibility of MIANS to 149 position, implying this replacement may modify the environment around Arg149, which would be important for efficient transport. Therefore, the conclusion emerges that the R149C mutant remains largely in an inward-facing conformation and is not capable of reorienting its binding sites to the periplasmic side. This defect explains the absence of transport in R149C, even if it is capable of binding the substrates in ISO vesicles.

According to current hypothesis on the transport mechanism by the alternate access model(Jardetzky 1966; Zhou et al. 2008; Forrest and Rudnick 2009; Krishnamurthy et al. 2009; Shimamura et al. 2010; Radestock and Forrest 2011), a critical feature of substrate transport is the equilibrium established between the outward- and the inward-facing orientations, allowing for an efficient transport turnover (Figure 4.11). Thus, the substrate-binding sites of the empty transporter is assumed to be alternatively accessible from both sides of the membrane according to an equilibrium established between both conformations, with a free energy barrier for this conformational change in the range of the thermal energy. For MelB, this is substantiated by the similar accessibility of the substrates to the substrate-binding sites from both sides of the vesicles (Figure 4.8), even if the frequency of opening to the periplasmic space might be much lower than opening to the cytoplasmic space(Smirnova et al. 2011). Under normal conditions, the binding of the substrates from the periplasmic side is followed by the

reorientation of the outward-facing to the inward-facing conformation (Forrest and Rudnick 2009; Krishnamurthy et al. 2009). The binding process itself might also contribute to speed up the reorientation process by decreasing its free energy barrier, as it occurs in enzymatic reactions. On molecular terms, this reorientation and change in accessibility would take place by means of the so-called rocking bundle mechanism, in which some transmembrane helices or part of helices tilt as rigid bodies, in concerted movements involving also conformational changes in the helical ends and loops(Law et al. 2008; Forrest et al. 2011). Changes of helical tilting of MelB upon substrate binding have been described by using polarized $IR_{diff}$ spectroscopy (Lorenz-Fonfria et al. 2009), giving support to this notion. The results presented in this thesis demonstrate that Arg149 is a required side chain for this switching mechanism to occur in MelB.

The positive charge at position 149 does not seem to be an absolute requirement for the reorientation mechanism, in keeping with previous suggestions (Abdel-Dayem et al. 2003). Our data demonstrate that, although R149Q shows easier access from the extracellular side, it is accessible to the fluorescent sugar analog from both sides (Figure 4.8). On the other hand, R149K shows the same accessibility as Cys-less from both sides. The $IR_{diff}$ spectra for these two mutants are altered in shape and are reduced in intensity, especially for sugar binding (Figure 4.4 and Figure 4.5), indicating that both mutations affect substrate binding and moderately reduce the affinity for melibiose. These results are in agreement with a previous work reporting that these mutants retain 25-30% of transport efficiency (Abdel-Dayem et al. 2003). Therefore, it seems likely that the reorientation of the carrier requires at least a polar residue at position 149, and that the positive charge of Arg149 would provide the optimal arrangement and interactions for an effective change of orientation and efficient transport.

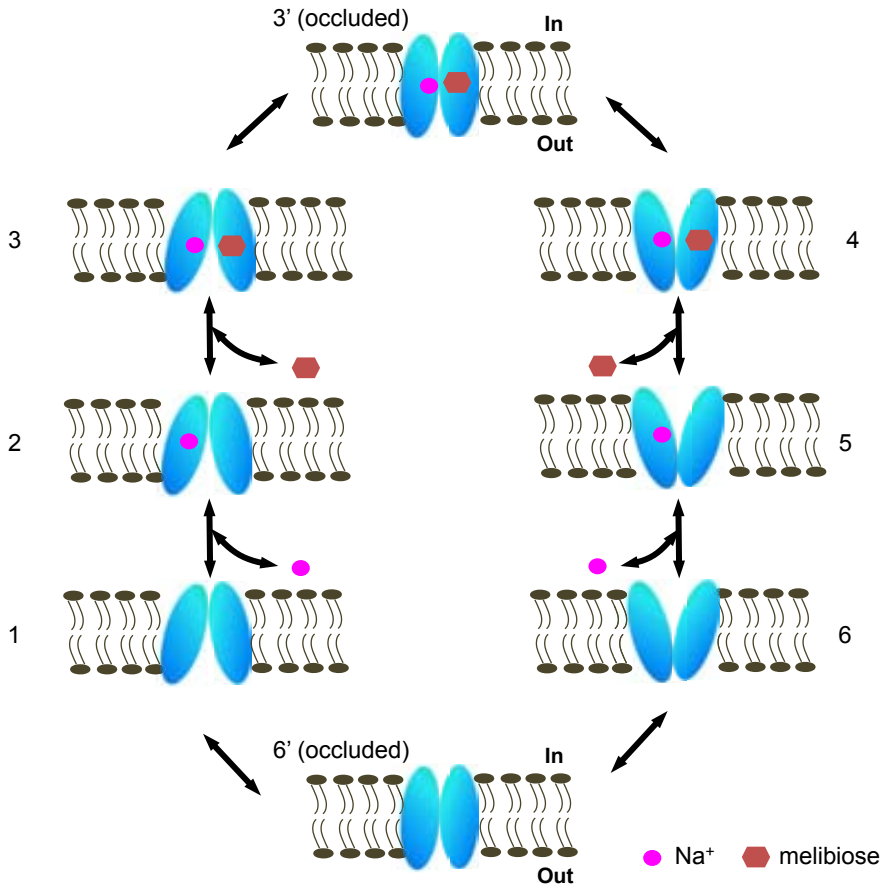



**Figure 4.11** Extended kinetic model for MelB. In this model, MelB presents outward-facing (states 1-3), inward-facing (states 4-6), and occluded conformations (states 3' and 6'). The transition between outward- and inward-facing conformations is expected to occur through the occluded states. (Adapted from Granell et al. 2010)

In the absence of an atomic 3D structure for MelB, it is challenging to presume how Arg149 performs its role. It could be hypothesized that Arg149, most probably located in the inner (cytoplasmic) half of TM5, could interact with a side chain of TM4. It could form a hydrogen bond with Tyr120 or a salt bridge with Asp124, both located favorably in TM4 in a recent MelB model (Yousef and Guan 2009). One or both of these interactions could be crucial for stabilizing the helix tilt changes (see Figure 4.12), and without them the energy barrier for the reorientation of the binding sites might largely

increase, explaining the effect of R149C mutation. In this regard, it has been proposed that TM4 in MelB could behave as a hinge for the required helical movements during transport (Abdel-Dayem 2003; Cordat et al. 2000). In TM4, Gly117 has also been implicated in the reorientation mechanism (Ganea et al. 2011). In this case, the mutant was proposed to be in an initial outward-facing conformation, contrary to R149C. It could be envisaged that the contiguous TM4 (where Gly117 is located) and TM5 (containing Arg149) could undergo coordinated movements during transport by a mechanism in which both side chains are critical components. Comparison with the *E.coli* lactose permease (LacY), another member of the Major Facilitator Superfamily, can help to set up a reasonable hypothesis. In LacY, Arg144, located in TM5, is irreplaceable for active lactose transport by establishing alternating salt bridges with Glu126 and Glu269 side chains during transport (Mirza et al. 2006). It is remarkable that the two available predictions about the MelB structure propose a similar location of Arg149 compared to Arg144 of LacY in TM5 and of Asp124 compared to Glu126 of LacY in TM4 (Yousef and Guan 2009) (see Figure 4.12B). In addition, it is worth noting that the putative TM5 of MelB, like the corresponding region of LacY, contains two Gly side chains that may facilitate the reorientation movements.

Importantly, a mutation also in TM5 of LacY (C154G) has been described to stabilize the transporter in one conformation that allows sugar binding from both sides, but cannot undergo the conformational changes on the periplasmic side required for translocation (Smirnova and Kaback 2003; Guan et al. 2007; Kaback et al. 2011). It is therefore tempting to speculate that key side chains in TM5 participate in interactions important for the reorientation mechanism and that Arg149 of MelB could perform a similar role as Arg144 of LacY. In this scenario, Arg149 could be ion paired with Asp124 and/or hydrogen bonded to Tyr120, both located in TM4, inaddition to directly participating in sugar binding. These interactions could be essential for the reorientation mechanism or, at least, for the opening of the periplasmic side. The observation that in all the studied Arg149 mutants both $Na^+$ and melibiose-induced structural changes are

in some way alter favors an interaction of Arg149 with Asp124, a residue found to be essential in coupling the $Na^+$ and melibiose binding site (Granell et al. 2010).

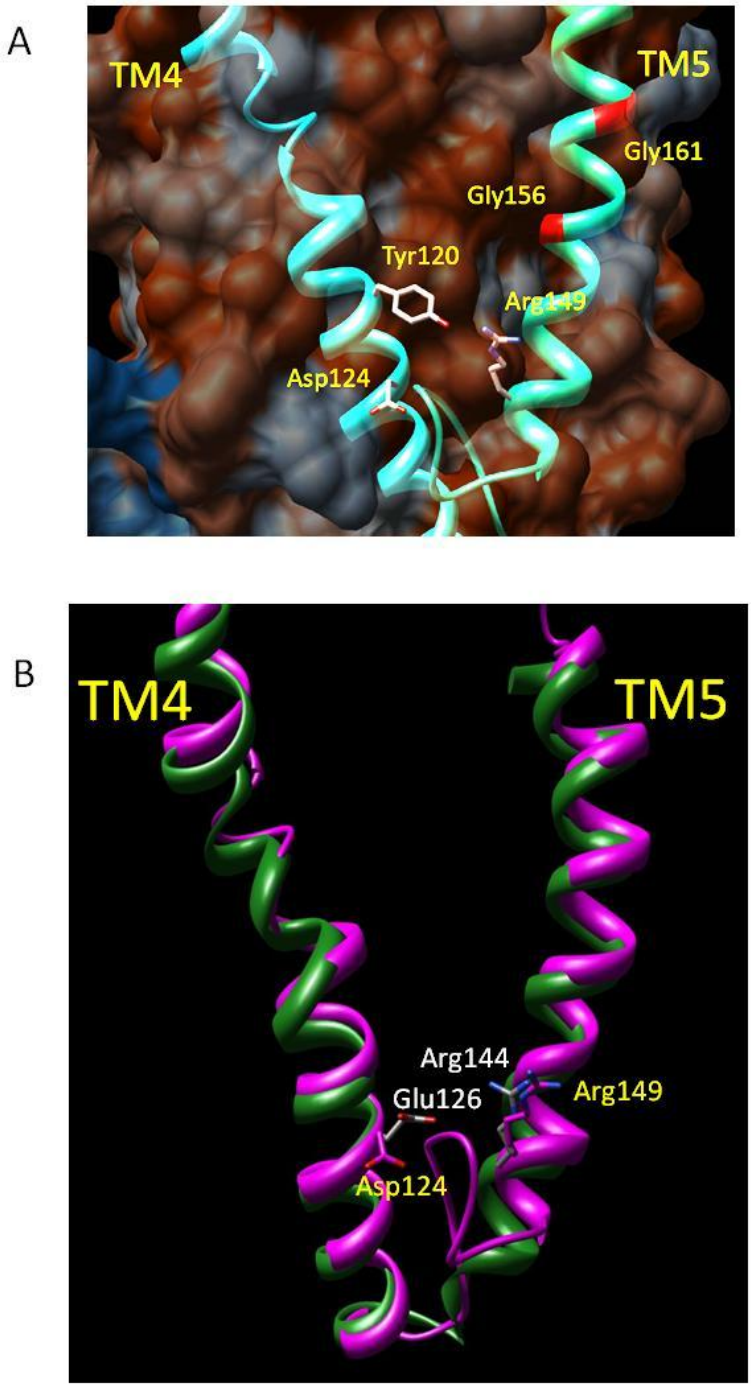



**Figure 4.12** (A) Model of helices TM4 and TM5 of MelB. The location of Gly residues in TM5 is marked in red color. The MelB model was obtained from Yousef and Guan 2009. (B) Comparison of helices TM4 and TM5 of the MelB model (pink) with the homologous helices of the LacY structure (green) corresponding to the TDG-bound protein (pdb 1PV7). The model of MelB outlined in Granell et al. 2010 discloses equivalent locations of these residues.

# 6 Conclusions of Part I

1) The R149C mutant can bind $Na^+$ as well as melibiose in proteoliposomes and in ISO membrane vesicles. It cannot bind sugars in RSO membrane vesicles.

2) The R149C mutation fixes the MelB in an inward-facing conformation. Therefore, Arg149, located probably in the cytoplasmic half of transmembrane helix 5, is a crucial side chain for the reorientation mechanism of MelB.

# Part II Cysteine-scanning mutagenesis of the helix 5: evidences for the role of helix 5 in the substrates binding

# 4 Results and Discussions of Part II

In general, secondary active transporters organize their substrates binding sites and the amino acid residues that take part in the conformational changes, in a way that facilitates the cooperation of the substrates binding and the conformational changes. In this particular structure, the substrates binding site and the amino acid residues, participating in the substrates-induced conformational changes, are always close to each other in space. For example $LeuT_{Aa}$, a bacterial homologue of the $Na^+/Cl^-$-dependent neurotransmitter transporters from *Aquifex aeolicus* has a single substrate-binding site at its center, surrounded by the interior helices TM1, TM3, TM6 and TM8.The binding of $Na^+$ and Leu induce the movement of TM1 and TM6, which leads to the conformational changes from outward-facing to inward-facing (Yamashita et al. 2005). This phenomenon is commonly found in the membrane transport proteins (Toyoshima et al. 2000; Dutzler et al. 2002; Hunte et al. 2005; Faham et al. 2008; Weyand et al. 2008). Our previous work presented in **PART I**, showed that Arg149 located at the cytoplasmic end of helix 5fixes MelB in an inward-facing conformation when it is mutated to Cys, which suggests Arg149 may be a key residue for the reorientation of MelB or take part in this mechanism. Therefore, it may be reasonable to think that Arg149 should be close to or participate in the substrates binding site.

As a secondary active transporter, the binding of the first substrate to MelB will control the binding of the second substrate. One may suggest that these two substrate binding sites should be close to each other in three-dimensional structure (in many cases they share the same site) because this is the simplest way for a cooperation between these two substrates. This feature can be seen in many cotransporters that have been crystallized and their structure resolved to high-resolution, e.g., LeuT (Yamashita et al. 2005), vSGLT (Faham et al. 2008), Mhp1 (Shimamura et al. 2010), etc. For lactose permease, Glu269 is not only at the sugar binding site and is a ligand for the sugar in one transport step, but it is also at the proton binding site (Abramson et al. 2003). So it is reasonable to suppose that for MelB the cation binding site and the sugar binding site are close to each other or even they overlap.

Previous studies showed that the $Na^+$ coordination ligands are essentially located in the N-terminal domain of MelB (Poolman et al. 1996). A recent study showed that two aspartates located at helix II, i.e. Asp55 and Asp59, are essential for $Na^+$ binding and Asp19 is involved in sugar binding (Toyoshima et al. 2000). In contrast, sugar binding ligands seem to be distributed in both the N-terminal and C-terminal domains of MelB. Some evidences showed that helix IV may play a role in connecting $Na^+$ binding and sugar binding (Poolman et al. 1996). A cysteine scanning mutagenesis study for helix I showed that several residues were involved in substrates binding and translocation (Ding and Wilson 2001a). In a 3D threading model of MelB (Yousef and Guan 2009), Yousef and Guan docked the sugar in the central cavity that is lined by the central part of several helices from both the N-terminal (helices I, II, IV, V) and C-terminal (helices VII, VIII, X, XI) domains. In this 3D model, Helix 5 is expected to be close to TM 4 (containing important residues, e.g., Tyr120, Asp124, Gly117) and helix I (containing some important residues, e.g., Gly17, Lys18, Asp19, Tyr32, Thr34, and Asp35). A study suggested that cytoplasmic loop connecting helices IV and V of the melibiose permease is involved in the process of $Na^+$-coupled sugar translocation (Abdel-Dayem et al. 2003) and conformational changes after sugar binding (Meyer-Lipp et al. 2006).

All of these evidences imply that helix 5 is an important structure of the function of MelB. We may suggest that some residues from helix 5or whole helix 5 take part in the reorientation of MelB, and that the substrates binding site may directly be located at helix 5 or is close to substrates binding sites in the 3D structure. Then we performed cysteine-scanning mutagenesis for helix 5 (residues 150-163). Using a functional MelB permease devoid of Cys residues (Cys-less), each amino acid residue in the sequence $^{149}$RFFASLAGFVTAGVT$^{163}$ in helix 5 was replaced individually with Cys, and determine the important residues involved in substrates binding by FTIR difference spectroscopy and fluorescence spectroscopy techniques. In order to have a more overall view of the effects of the different mutations in helix 5, some mutants of Arg149, located at the cytoplasmic terminal of helix 5, which have already been studied in the PART I, were included in this Section.

## 4.1 Infrared spectroscopy analysis

The $IR_{diff}$ technique is useful to determine the interaction between the substrates and the protein (Leon et al. 2005). Some previous works have demonstrated that $IR_{diff}$ is able to disclose the substrates-dependent conformational changes of MelB (Leon et al. 2005; Leon et al. 2006; Leon et al. 2009; Granell et al. 2010).

Here, the same technique was used for detecting and characterizing substrate-induced conformational changes on mutants of cysteine-scanning mutagenesis for helix 5. After verification of the sequence, each single Cys mutant as well as the Cys-less mutant used as a control was transformed into *E.coli* DW-2R, and then expressed, purified, reconstituted in *E.coli* lipids, deposited over an ATR crystal, and the $IR_{diff}$ was obtained as shown in ***Materials and Methods***. For more a convenient comparison with other mutants, two additional mutants, i.e., S153A and T159A, which are further discussed later on, are also considered in this section.

### 4.1.1 The effect of mutations on the MelB structure

The effect of mutations on the secondary structure of MelB can be determined by quantitative comparison of the shape of the structure-sensitive amide I and amide II protein bands of the second derivative of the absorbance spectra between 1500-1700 $cm^{-1}$ (Figure 3.6). The difference of the structure between these single-site mutants is lower than 2.5% (Figure 5.1A), suggesting that none of the single residue mutations will disorder the whole structure of the protein. This observation reasonably discards that the observed defective substrate-binding phenotypes could be due to a loss of the protein native structure.

In spite of the small difference of the structure between the different mutants, all of them were clustered into three groups (Figure 5.1A), i.e., F151C, S153C, L154C, G161C, T159C, A155C, and T159A (named Group 1); F150C, F157C, A160C, T163C, V162C, A152C, S153A, and V158C (named Group 2); Cys-less, R149C, and G156C (named Group 3), indicating that these mutations affect somewhat the structure of MelB.

Figure 5.1B shows the second derivative of the absorbance spectra in the amide I/II region for the MelB mutants considered in this study. In Group 1 (Figure 5.1B), these mutations lead to a shift and an increase in intensity of a band centered at 1660 $cm^{-1}$, corresponding mainly to α-helices and to an increase in intensity of a band centered at 1630 $cm^{-1}$ specially for A155C, assigned to a mixture of β-sheets, $3_{10}$ helices, and open loops (Fabian et al. 1992; Arrondo et al. 1993; Jackson and Mantsch 1995). As the difference of the structure between these mutants is very small, the more reasonable possibility is that these mutations affect the structural surroundings of the mutated side chain, special those structures that may interact with these side chains. Among the mutants in the Group 1, it is seen that the MelB structure is subject to a relatively significant effect, special those of β-sheets, $3_{10}$ helices, and open loop in the A155C mutant. Based on previous studies of the topology of MelB, Ala155 should be located at the middle of the helix 5. In general, the main structural component of the transmembrane part of the membrane transport proteins is α-helix, except for the discontinue structures, which always play key role in the substrates binding or/ and substrate-dependent conformational changes (Krishnamurthy et al. 2009). Therefore, it may be reasonable to suppose that Ala155 is located at or close to discontinue structures.

T159A is a mutant which presents a quite similar structure as that of A155C. The T159A mutation also affects α-helix (band at 1660 $cm^{-1}$) and β-sheets, $3_{10}$ helices, and open loops (band at 1630 $cm^{-1}$), only with less intensity of the band at 1630 $cm^{-1}$as that of A155C. In the Group 2, all of the mutants present very similar absorbance spectra between 1500-1700 $cm^{-1}$, except for band at 1660 $cm^{-1}$, displaying a little difference between mutants and Cys-less. These results suggest that the mutations in the Group 2 affect principallyα-helices, although these effects are very limited. In the Group 3, it is seen that the differences in the secondary structure of Cys-less, R149C, and G156C are not significant (Figure 5.1B).

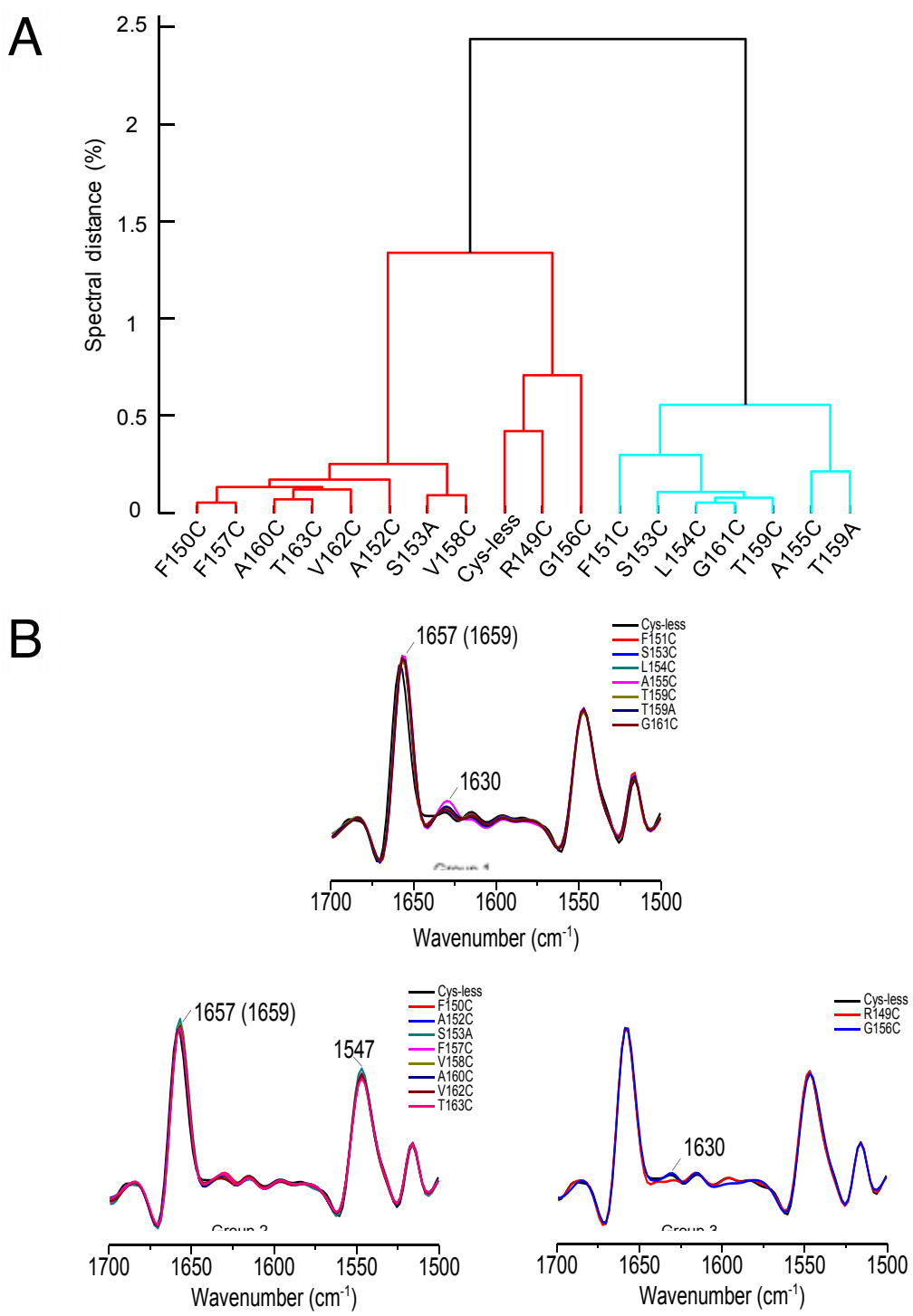



Figure 5.1 Structural comparison of MelB mutants using IR spectroscopy of hydrated samples. (A) Dendrogram representing the spectral distance (related to the structural distance) between different proteins forms, measured as $(1-R^2) \times 100$. (B) Second derivative in the amide I/II region for the MelB mutants considered in this study. All spectra intensities were normalized to the Cys-less. For detail, see ***Materials and Methods***.

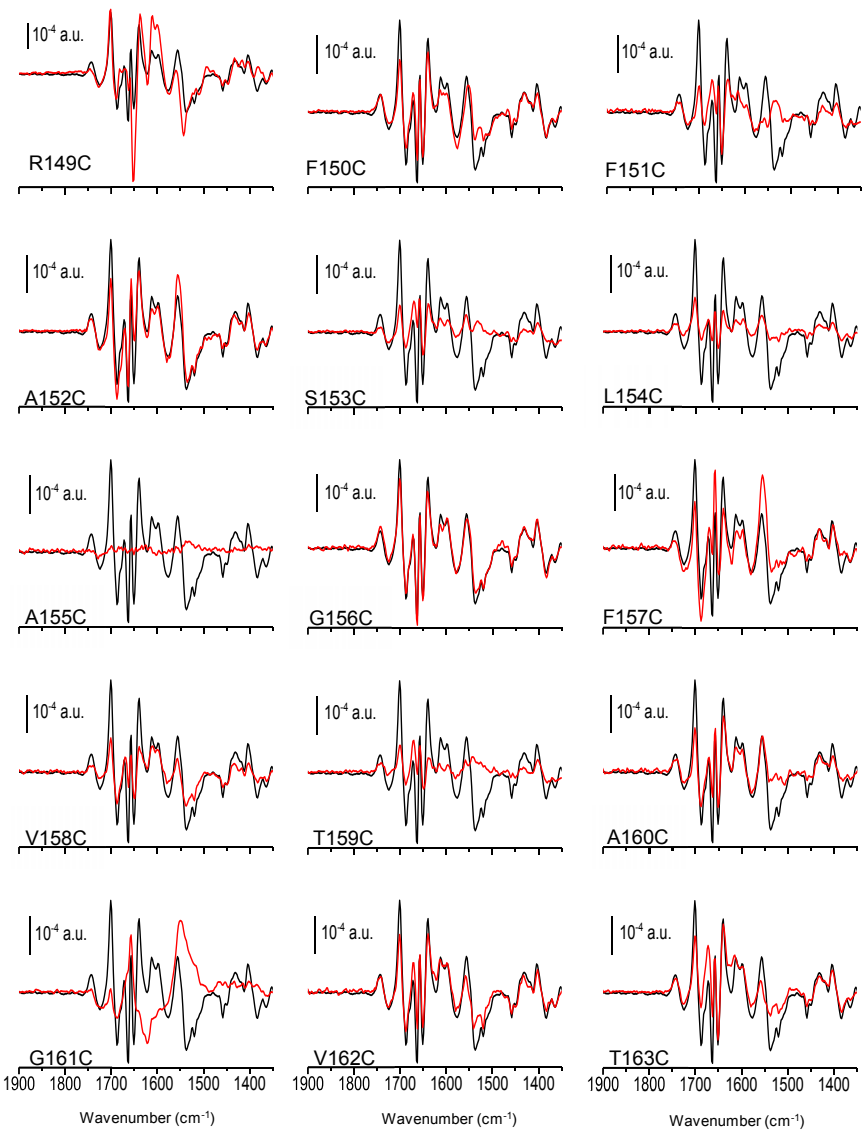
For the comparison of the similitude of the secondary structure of these different form MelB mutants, see Appendix 1.

4.1.2 Effect of MelB mutations on $Na^+$ binding

To determine the effect of MelB mutations on $Na^+$ binding, we obtained the $Na^+$-dependent $IR_{diff}$ spectra for these MelB mutants. As described in Granell et al. 2010, the addition of 10 mM $Na^+$ (17-fold above the affinity constant) (Abdel-Dayem et al. 2003) generates a reproducible difference spectrum, formally originated from the substitution of a proton ($H_3O^+$) with a $Na^+$ in the cation-binding site, reflecting not only interaction(s) of $Na^+$ with the cationic-binding site ligands and associated local structural adjustments, but also the induced protein structural changes responsible for the well-established increase in sugar affinity following $Na^+$ binding.

Figure5.2A shows that many mutants displayed difference spectra similar to Cys-less, but some others showed strongly decreased spectra. These $Na^+$-dependent $IR_{diff}$spectra of the MelB mutants can be quantitatively compared as unbiased as possible by means of a linear regression analysis encompassing the structure-sensitive 1,700-1,500 $cm^{-1}$ region from their difference spectra(see ***Materials and methods***, and Granell et al. 2010). First, the linear correlation parameter, $R^2$, quantifies the spectral similarity of a mutant response relative to the Cys-less, i.e., the percentage of spectral features in common with the control Cys-less (Figure 5.2B red columns). A high spectral similarity for a mutant is expected to correlate with structural changes in response to the substrate, highly similar to those to the Cys-less. Second, the slope of the linear correlation gives the relative intensity of the spectral features in common with the Cys-less intensity (Figure 5.2B blue columns). A relative intensity lower than 100% for any given mutant implies either a reduced affinity for the added substrate or smaller structural changes in response to substrate binding than for the Cys-less.

A


10-4 a.u.
R149C
F150C
F151C
A152C
S153C
L154C
A155C
G156C
F157C
V158C
T159C
A160C
G161C
V162C
T163C
1900 1800 1700 1600 1500 1400
Wavenumber (cm-1)

B

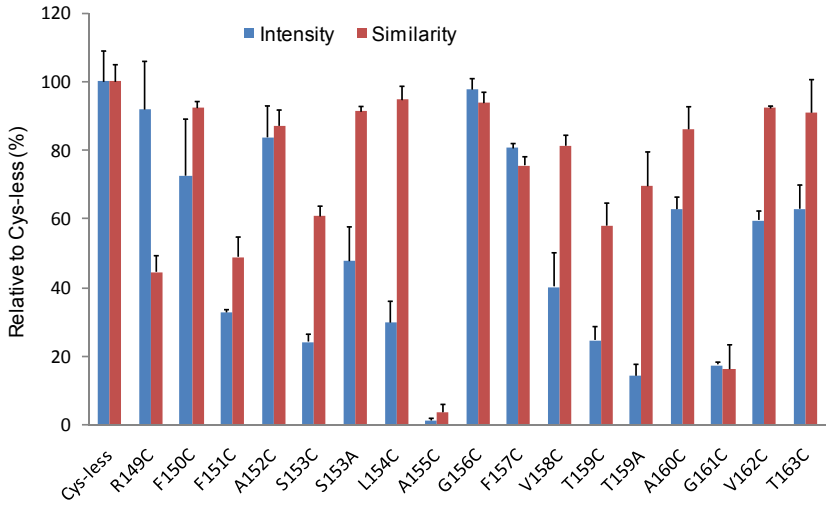

Intensity
Similarity
Relative to Cys-less (%)
120
100
80
60
40
20
0
Cys-less
R149C
F150C
F151C
A152C
S153C
S153A
L154C
A155C
G156C
F157C
V158C
T159C
T159A
A160C
G161C
V162C
T163C

**Figure 5.2** $Na^+$-induced IR difference spectra at 4 $cm^{-1}$resolution of different MelB mutants (red curves) compared to Cys-less (black curve). Difference spectra were obtained by replacing $Na^+$-free buffer by a buffer media containing $Na^+$ at the concentration of 10 mM (*See **Materials and Methods***). All the difference spectra were normalized to the amount of probed protein. The average difference spectrum from two or more independent experiments is shown in (A). (B) Spectral similarity and intensity of the $Na^+$-induced IR difference spectra in the mutants compared to the Cys-less. Spectral similarity and intensity for S153A and T159A were obtained by calculating the $IR_{diff}$ spectra presented in Figure 5.14. The error bar corresponds to one standard error of the mean.

As shown in Figure 5.1B of the 16 single Cys mutants, four mutants show quite similar intensity compared to Cys-less (G156C, R149C, A152C and F157C), four mutants display 60%-80% intensity to Cys-less (For F150C, A160C, V162C and T163C), five mutants present an intensity between 20-40% (F151C, S153C, L154C, V158C and T159C). G161C displays a low intensity of 18%. A155C shows a quite low intensity compared to Cys-less of 1.3%, indicating it is inactive with respect to $Na^+$ binding. The similarity of $Na^+$-induced $IR_{diff}$ spectra shows that eight mutants are highly similar to Cys-less with values between 80 and 100% (L154C, G156C, F150C, V162C, T163C,A152C, A160C, and V158C), three mutants display a similarity between 50-80%, i.e., F157C, S153C and T159C, two mutants display similarities between 30-50%, i.e., R149C and F151C. G161C displays a low similarity of 16%, whereas A155C shows an insignificant spectral similarity with respect Cys-less.

As seen in Figure 5.2A, A155C did not show any clear peak assignable to the protein in its difference spectrum (flat signal below ~ 1,720 $cm^{-1}$).Not only the intensity is very low, but also the similarity, when comparing intensity and similarity for A155C and Cys-less. Therefore, it does not produce conformational changes typical for $Na^+$ binding. These results establish that Ala155 is an essential residue for $Na^+$ binding. The behavior of F151Cand G161C deserves some special comments. These two mutants display a

$Na^+$-induced $IR_{diff}$ spectrum, indicating that they retain the ability to bind $Na^+$ (Figure 5.2A), implying that Phe151and Gly161 are not strictly essential for $Na^+$ binding to MelB. However, the resulting difference spectra show that both the similarity and the intensity are reduced with respect to the Cys-less difference spectrum, suggesting that these two residues may participate in the configuration of the $Na^+$ pocket. The behavior of S153C and T159C, which contain a hydroxyl group, deserve some special comments also. These mutants display a $Na^+$-induced $IR_{diff}$ spectrum, indicating that they retain the ability to bind $Na^+$ (Figure 5.2A), implying that Ser153 and Thr159 are not essential for $Na^+$ binding to MelB. The resulting difference spectra show a moderate similarity (~55-65%) and a reduced intensity (~25-35%) with respect to the Cys-less difference spectrum (Figure 5.2B). The moderate similarity suggests that the $Na^+$-induced structural changes of S153C and T159C may be less complete that those occurring in Cys-less. This conclusion is supported by the fact that both of these two mutants lacks some intense peaks present in Cys-less (e.g., at 1662 $cm^{-1}$, 1657 $cm^{-1}$, and 1651 $cm^{-1}$ which were assigned to α-helix changes), whereas all peaks in these two mutants $IR_{diff}$ spectrum are also seen in Cys-less. Consequently, not only the structural changes induced by $Na^+$ in S153C and T159C are less complete than in Cys-less, but are also of smaller amplitude, i.e., Ser153 and Thr159 are required for full and native-like structural protein changes in response to $Na^+$ binding.

L154C and V158C display $Na^+$-induced $IR_{diff}$ spectra with a reduced intensity and a slightly reduced similarity. These results suggest that they retain the ability to bind $Na^+$ (Figure 5.2B). The resulting difference spectra show a good similarity between these two mutants and Cys-less. This implies that $Na^+$ can bind to these two mutants with a decreased affinity, inducing quite similar conformational changes. F150C, A160C, V162C, and T163C show $Na^+$-induced $IR_{diff}$ spectra with a reduced intensity between 63-73% and a slightly reduced similarity between 86-92% (Figure 5.2B). These data indicate that these mutants retain the ability to bind $Na^+$ with a slight reduced affinity. It is clear that Phe150, Ala160, Val162, and Thr163 are not essential residues for $Na^+$ binding.A152C and G156C display almost a Cys-less-type signal in terms of intensity

and similarity (Figure 5.2B), meaning that these two residues are not involved in $Na^+$ binding or in the configuration of the $Na^+$-binding pocket.

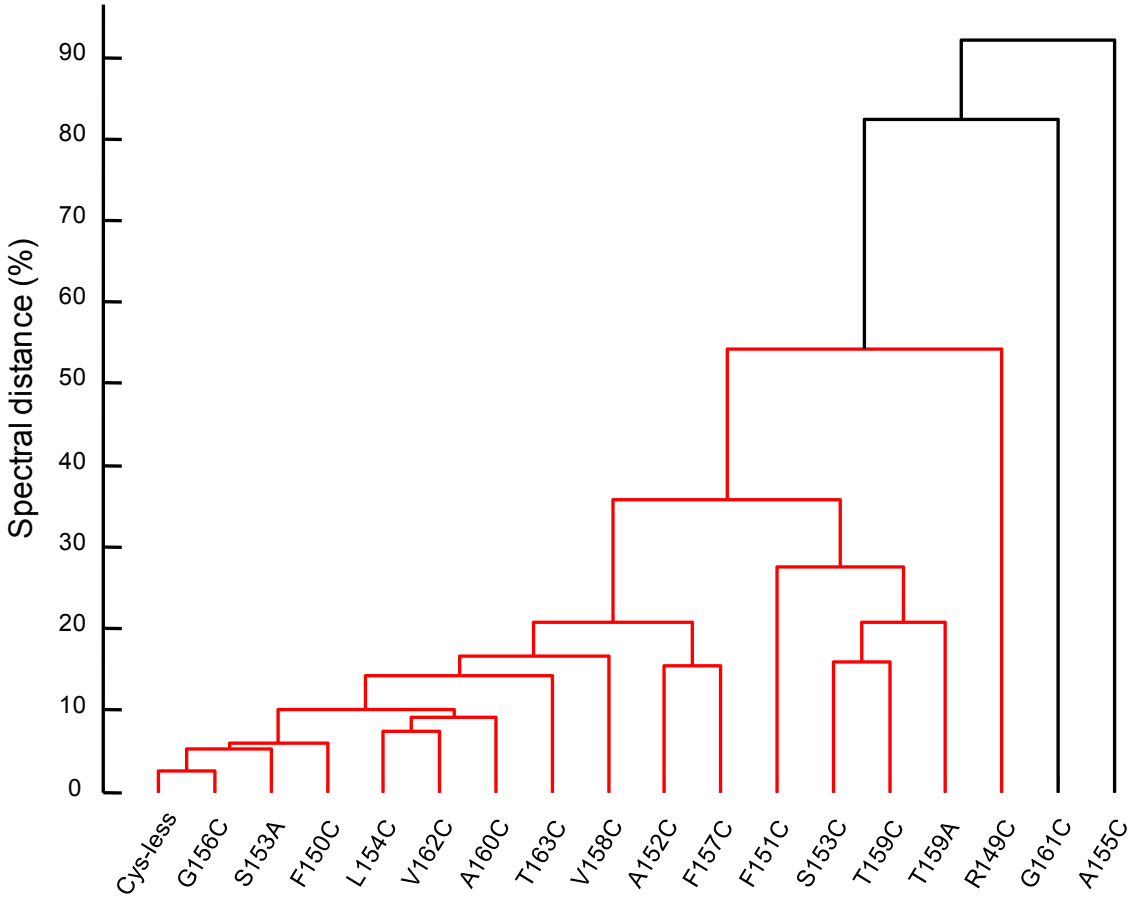



**Figure 5.3**Dendrogram clustering samples according to their spectral similitude in response to $Na^+$ binding (shown in Figure 5.2), measured as $(1 – R^2) \times 100$. (See ***Materials and Methods***)

To further complete the characterization of the studied mutants, we performed a comparison of the similitude of the $Na^+$-induced $IR_{diff}$ spectra across all mutants, collected in a correlation matrix (see Appendix2) used to construct a dendrogram (conceptually similar to a phylogenetic tree) (Figure 5.3). The response of the A155C mutant appears unclustered, since the distance between this mutant and the other mutants is quite high (~90%), as expected for a protein without any specific interaction with $Na^+$. The response of the G161C mutant also appears unclustered, corresponding to a protein with a mutation that affects the specific interaction with $Na^+$. The responses of the F151C, S153C, T159A, and T159C appear to cluster together with a specific distance of about 40% to Cys-less, as expected for residues that are required for full and native-like structural protein changes in response to $Na^+$ binding. The responses of the Cys-less, G156C, F150C, S153A, L154C, A160C, A152C, T163C, V158C, V162C, and F157C are clustered together with specific distance lower than 20%, suggesting that all of these mutations almost preserve the native conformational changes responding to $Na^+$ binding. Since L154C and V158C showed a strongly reduced intensity, they can be removed from this cluster. Therefore we can conclude that Gly156, Phe150, Ala160, Thr163, Ala152, Val162 are not essential residues involved in $Na^+$ binding.

4.1.3 The effect of the MelB mutations on the binding of melibiose in the absence of $Na^+$

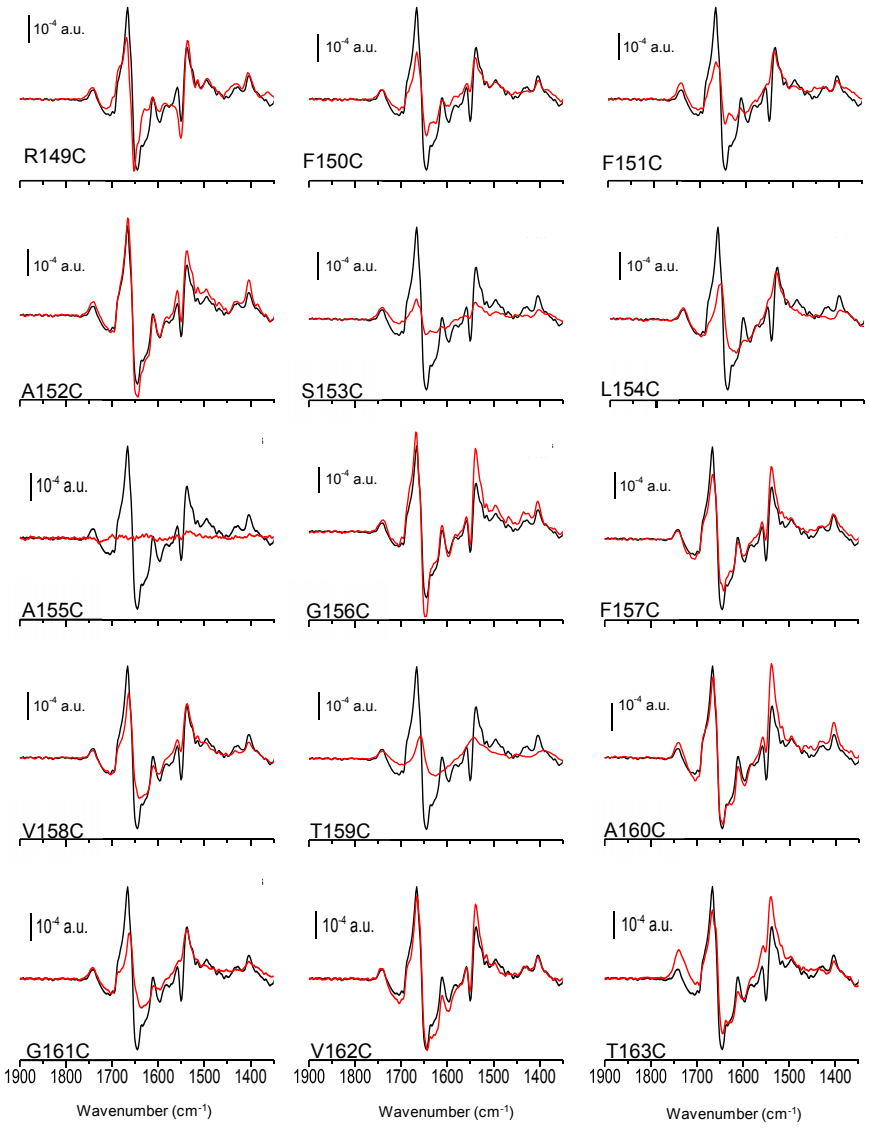
We next explored how each mutation affects the MelB ability to bind melibiose in the absence of $Na^+$, i.e., when the only possible coupling ion is $H^+$. In these experiments the sugar was added at a concentration of 50 mM, a value close to the half-saturating concentration of Cys-less for melibiose (Meyer-Lipp et al. 2006).

As shown in Figure 5.4B of the 16 single Cys mutants, five mutants show quite similar intensity to Cys-less (A152C, G156C, F157C, A160C and V162C), three mutants display 60%-80% intensity to Cys-less (For R149C, V158C, and T163C), F150C present 48% intensity to Cys-less, and three mutants present at lower intensity between 30-40% (F151C, L154C and G161C).Two containing hydroxyl amino acids, S153C and T159C,displayquitelow intensities of about 14% for S153C and about 8% for T159C respectively. A155C shows a very small signal (1.8% in intensity as that of

Cys-less), indicating its inactivity for melibiose binding. The similarity of the $IR_{diff}$ spectra induced by melibiose in the absence of $Na^+$ shows that eight mutants are similar to Cys-less between 80 and 100% (F150C, A152C, G156C, F157C, V158C, A160C, V162C and T163C), four mutants display similarities between 60-80%, i.e., R149C, F151C, S153Cand G161C, and two mutants display a quite low similarity, i.e., L154C of 38% and T159C of 17%. The A155C mutant shows an insignificant spectral similarity with respect Cys-less (about 6%).

As seen in Figure 5.4A, A155C did not show any clear peak assignable to the protein in its difference spectrum induced by melibiose. These results prove that Ala155 is an essential residue for melibiose binding. The behaviors of S153C and T159C are similar with quite low intensity (Figure 5.4B), indicating their importance for melibiose binding. However, S153C has higher similarity and intensity than T159C. These data suggest that Ser153 may affect the binding of melibiose to MelB, but Thr159 may be an important residue in the binding of melibiose. L154C displays a melibiose-induced $IR_{diff}$ spectrum with similarly reduced intensity and similarity, indicating that it retains the ability to bind melibiose (Figure 5.4B). This implies that Leu154 is not essential for melibiose binding to MelB, but its mutation to Cys impedes the binding of melibiose to some extent. F151C and G161C display melibiose-induced $IR_{diff}$ spectra with low intensity (~37-45%) and a moderate similarity (~64-73%) (Figure 5.4B). Therefore, Phe151 and Gly161 are not essential for melibiose binding to MelB, although the melibiose-induced structural changes of F151C and G161C may be less complete that those occurring in Cys-less. This conclusion is supported by the fact that both of these two mutants lack some intense peaks present in Cys-less (e.g., at 1668 $cm^{-1}$, and (-)1645 $cm^{-1}$), whereas all peaks in their $IR_{diff}$ spectra are also seen in Cys-less (Figure 5.4A). Consequently, Phe151 and Gly161 are required for full and native-like structural protein changes in response to melibiose binding.

A


10-4 a.u.
R149C
F150C
F151C
A152C
S153C
L154C
A155C
G156C
F157C
V158C
T159C
A160C
G161C
V162C
T163C
1900 1800 1700 1600 1500 1400
Wavenumber (cm-1)

B

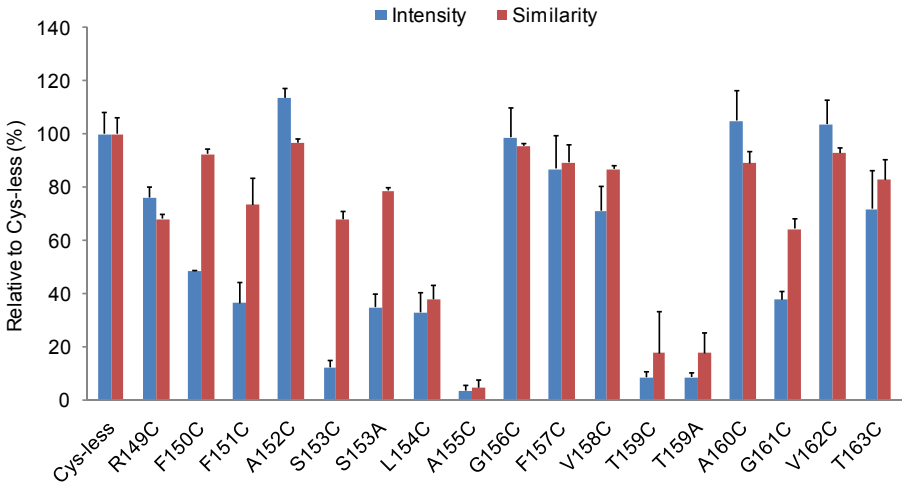

Intensity
Similarity
Relative to Cys-less (%)
0
20
40
60
80
100
120
140
Cys-less
R149C
F150C
F151C
A152C
S153C
S153A
L154C
A155C
G156C
F157C
V158C
T159C
T159A
A160C
G161C
V162C
T163C

**Figure 5.4** Melibiose-induced IR difference spectra at 4 $cm^{-1}$ resolution of different MelB mutants (red curves) in the absence of $Na^+$ compared to Cys-less (black curve). Difference spectra were obtained by replacing melibiose-free buffer by a buffer containing melibiose at the concentration of 50 mM. (See ***Materials and Methods***) All the difference spectra were normalized to the amount of probed protein. The average difference spectrum from two or more independent experiments is shown in (A). (B) Spectral similarity and intensity of the melibiose-induced IR difference spectra in the mutants compared to the Cys-less in the absence of $Na^+$. Spectral similarity and intensity for S153A and T159A were obtained by calculating the $IR_{diff}$ spectra presented in Figure 5.14. The error bar corresponds to one standard error of the mean.

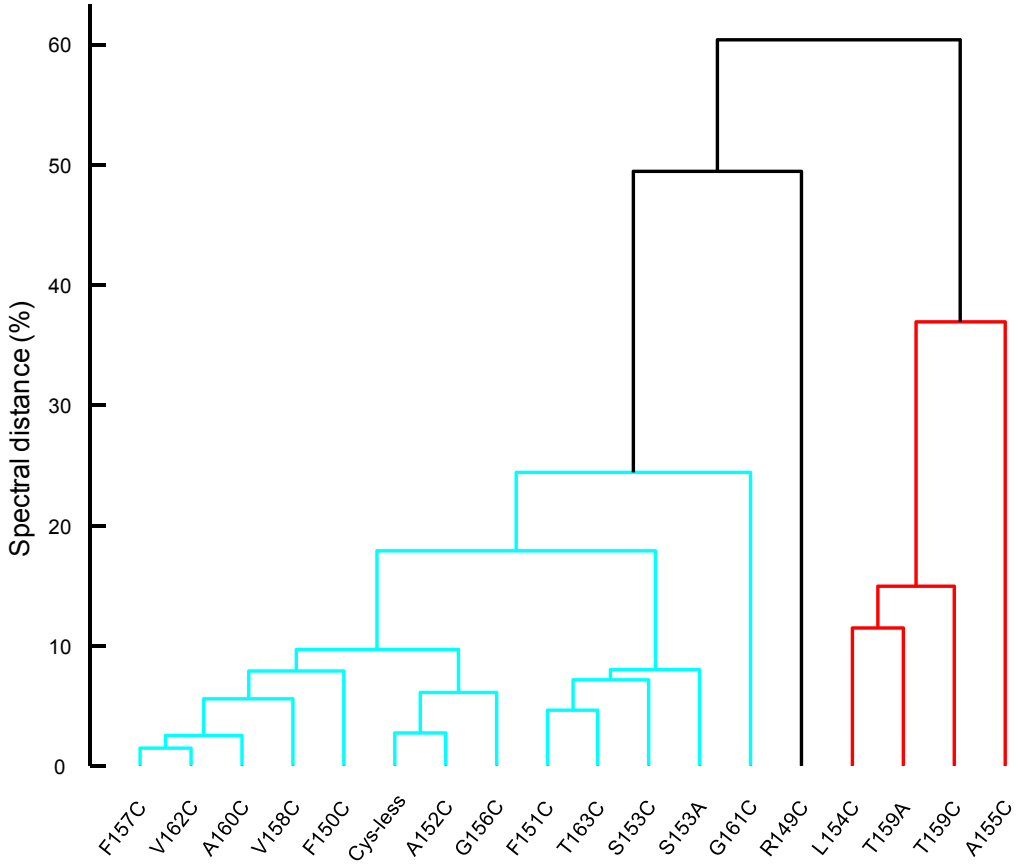



**Figure 5.5** Dendrogram clustering samples according to their spectral similitude in response to melibiose binding in the absence of $Na^+$, (shown in Figure 5.4), measured as $(1-R^2) \times 100$. (See ***Materials and Methods***)

F150C, V158C and T163C display melibiose-induced $IR_{diff}$spectra with a moderate intensity and a slightly reduced similarity (Figure 5.4B). These results suggest that they retain the ability to bind melibioseand undergo quite similar conformational changes.Thisimplies that Phe150, Val158 and Thr163 do not form part of the sugar binding pocket.

The behaviors of A152C, G156C, F157C, A160C, and V162C, especially of A152C and G156Care quite similar to Cys-less with very few differences in the intensity and similarly (Figure 5.4B).These data indicate that these mutants retain the ability to bind melibiose, and undergoconformational changes similar to Cys-less. Therefore, it is clear that Ala152, Gly156, Phe157, Ala160, and Val162 are not essential residues for melibiose binding.

4.1.4 The effect of the MelBmutations on the binding of melibiose in the presence of $Na^+$

We next explored how each mutation affectson the MelB'sability to bind melibiose in the presence of $Na^+$. In these experiments, the sugar was added at a concentration of 10 mM in the presence of 10 mM $Na^+$in the medium (double above the affinity constant, Meyer-Lipp et al. 2006). The melibiose-induced $IR_{diff}$ spectrum in the presence of $Na^+$reflects the coupling between the cation- and sugar-binding sites, and corresponds to the second step of the transport cycle after binding of $Na^+$.

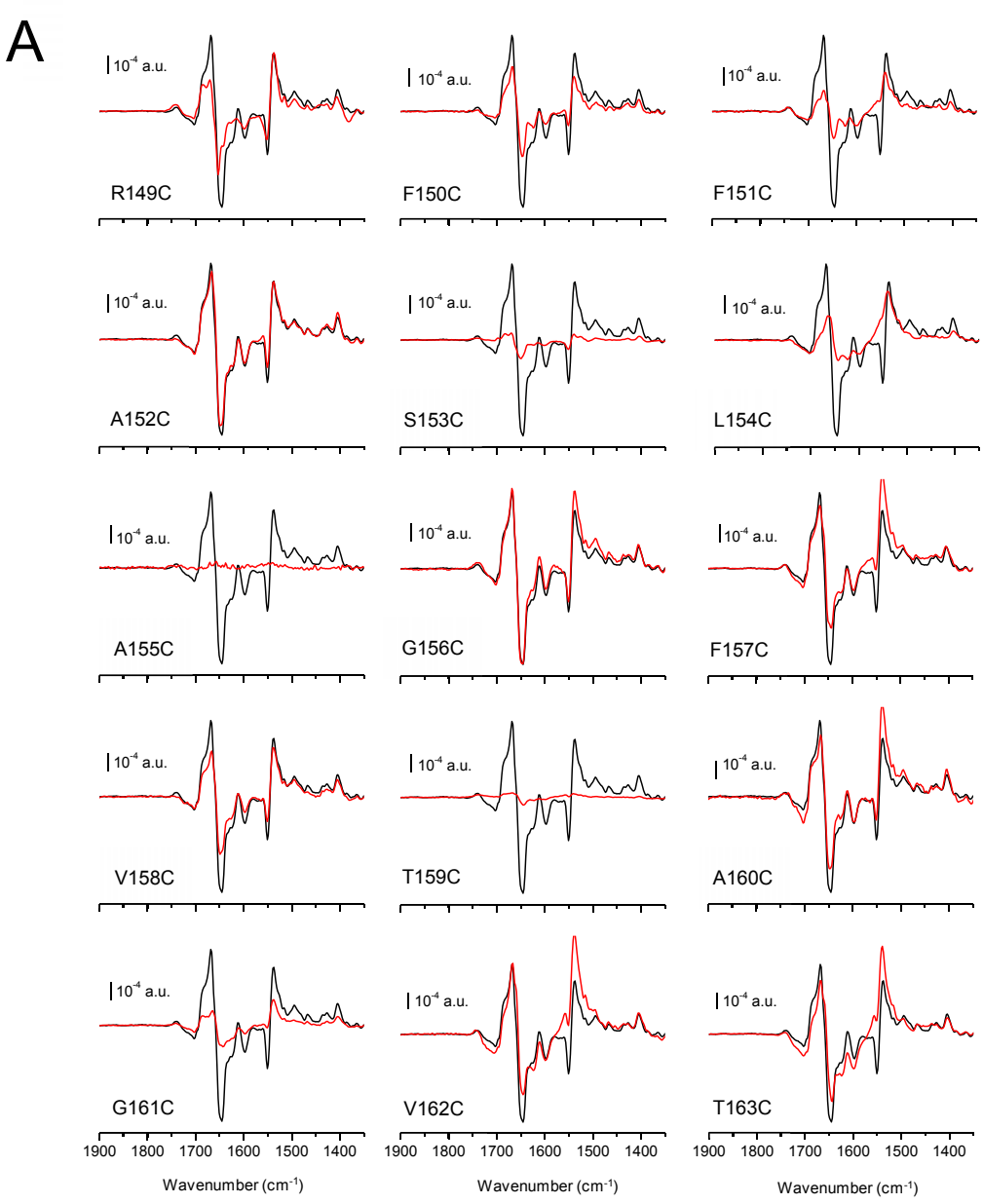
As shown in Figure 5.6A, of the 16 single Cys mutants, four mutants show quite similar intensity to Cys-less (A152C, G156C, A160C and V162C), three mutants display 60%-80% intensity to Cys-less (F157C, V158C and T163C), R149C and F150C present intensities of 54% and 51% respectively, and two mutants present intensities between 20-30% (F151C and G161C). Two containing hydroxyl amino acids, S153C and T159C, display intensities between 12% and 5% and A155C shows a very low intensity. Similarities of $IR_{diff}$ induced by melibiose in the presence of $Na^+$ shows that eight mutants give similar difference spectra to Cys-less, between 80 and 100% (F150C, A152C, G156C, F157C, V158C, A160C, G161C and V162C), five mutants display similarity values between 60-80% (R149C, F151C, S153C, T159C, and T163C), and

L154C display a low similarity of 20%. Finally, the A155C mutant shows an insignificant spectral similarity with respect to Cys-less.

As seen in Figure 5.6A, A155C did not show any clear peak assignable to the protein in the difference spectrum induced by the binding of melibiosein the presence of $Na^+$. When comparing intensity and similarity of A155C with Cys-less, both parameters are near zero, suggesting a loss in the capacity for melibiose binding. When consideringthe $IR_{diff}$spectra induced by the binding of $Na^+$, the binding of melibiose in the presence of $H^+$ and the binding of melibiose in the presence of $Na^+$, the A155C mutant leads to complete disappearance of the difference spectra at any conditions. This suggests that Ala155 playsan essential role in both $Na^+$ and melibiose binding.

The behaviors of S153C and T159C are similar showing $IR_{diff}$ spectra of low intensity, especially for T159C, suggestingthat they areimportant for melibiose binding. When comparing themelibiose-induced $IR_{diff}$spectra in the presence or absence $Na^+$, the similarity of melibiose-induced $IR_{diff}$ of S153C and T159C increase from 63% to 75%, and 17% to 66%, respectively, in the presence of $Na^+$, although there is not a clear effect on the intensity (Figure 5.4B and Figure 5.6B). Thisimplies that the binding of $Na^+$ may be required for full and native-like structural protein changes in response to melibiose binding.

L154C display melibiose-induced $IR_{diff}$ spectrum with a decreased intensity and similarity compared to Cys-less, indicating that it retains a certainability to bind melibiose (Figure 5.6B).The spectra suggest that this residue may not only participate in the melibiose binding, but also in the subsequent structural changes. It is important for the melibiose binding.


A
10-4 a.u.
R149C
F150C
F151C
A152C
S153C
L154C
A155C
G156C
F157C
V158C
T159C
A160C
G161C
V162C
T163C
1900 1800 1700 1600 1500 1400
Wavenumber (cm-1)

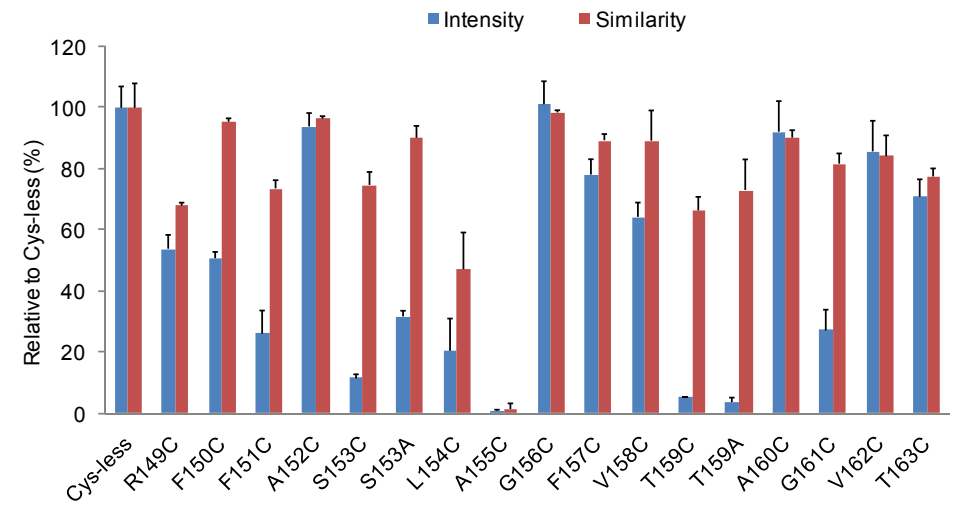

B
Intensity
Similarity
Relative to Cys-less (%)
120
100
80
60
40
20
0
Cys-less
R149C
F150C
F151C
A152C
S153C
S153A
L154C
A155C
G156C
F157C
V158C
T159C
T159A
A160C
G161C
V162C
T163C

Figure 5.6 melibiose-induced IR difference spectra at 4 $cm^{-1}$resolution of different form MelB mutants (red curves) in the presence of $Na^+$ compared to Cys-less (black curve). Difference spectra were obtained by replacing melibiose-free buffer by a buffer containing melibiose at the concentration of 10 mM. (*See* ***Materials and Methods***) All the difference spectra were normalized to the amount of probed protein. The average difference spectrum from two or more independent experiments is shown in (A).(B) Spectral similarity and intensity of the melibiose-induced IR difference spectra in the mutants compared to the Cys-less in the presence of $Na^+$. Spectral similarity and intensity for S153A and T159A were obtained by calculating the $IR_{diff}$ spectra presented in Figure 5.14.The error bar corresponds to one standard error of the mean.

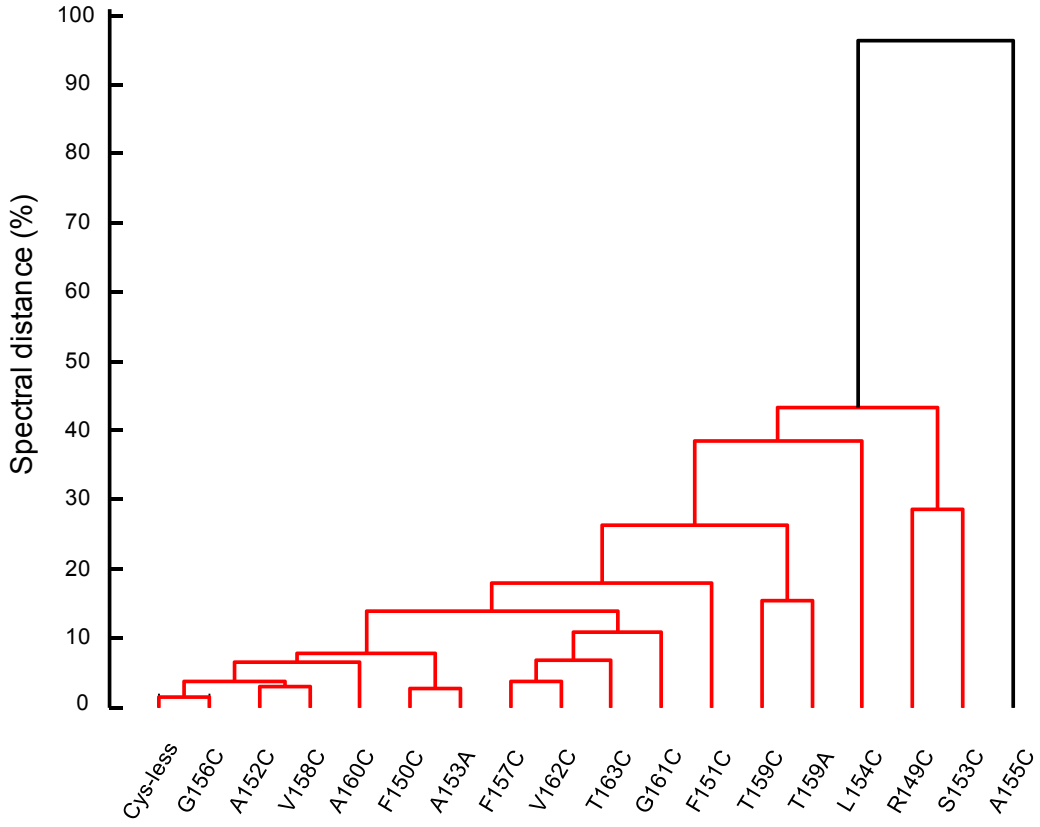



**Figure 5.7** Dendrogram clustering samples according to their spectral similitude in response to melibiose binding in the presence of $Na^+$,(shown in Figure 5.6), measured as $(1-R^2) \times 100$. (See ***Materials and Methods***)

The mutants F151C and G161C display melibiose-induced $IR_{diff}$ spectra, indicating that they retain the ability to bind melibiose (Figure 5.6B). Therefore,Phe151 and Gly161 are not essential for melibiose binding to MelB. However, the resulting difference spectra show a moderate similarity (~74-82%) and a reduced intensity (~26-27%) with respect to the Cys-less difference spectrum (Figure 5.6B). The moderate similarity suggests that the melibiose-induced structural changes of F151C and G161C may be less complete that those occurring in Cys-less. This conclusion is supported by the fact that the difference spectra lack some intense peaks present in Cys-less (e.g., at 1668 $cm^{-1}$, and (-)1645 $cm^{-1}$), whereas all peaks in the mutants $IR_{diff}$ spectrum are also seen in Cys-less (Figure 5.6A). Consequently, not only the structural changes induced by melibiose in F151C and G161C are less complete than in Cys-less, but are also of smaller amplitude, i.e., Phe151 and Gly161 are required for full and native-like structural protein changes in response to melibiose binding.

The F150C mutant displays a melibiose-induced $IR_{diff}$ spectrum with a moderate intensity (~50%), and a slightly reduced similarity (Figure 5.6B). Therefore, F150C binds melibiosewith a reduced affinity and similar conformational changes compared to Cys-less. The mutants V158C, A160C and T163C also display melibiose-induced $IR_{diff}$spectra with a moderate decrease of intensity and similarity. Thus, these mutants retain the ability to bind melibiose with a little effect on the affinity or on theconformational changes induced by the binding of melibiose. It implies that Val158, Ala160 and Thr163 are not involved in sugar binding.

The behaviors of A152C, G156C, F157C, and V162C are quite similar to Cys-less with very few differences in the intensity and similarly, retaining the ability to bind melibiose with conformational changes similar to Cys-less. It is clear that Ala152, Gly156, Phe157, and Val162 are not essential residues for melibiose binding.

## 4.2 Fluorescence spectroscopy analysis.

To give more strength to the conclusion drawn by $IR_{diff}$ spectroscopy that purified mutants in proteoliposomes do retain the capacity to bind $Na^+$ and melibiose and trigger associated conformational changes, we undertook complementary studies with the same proteoliposomes. We used intrinsic fluorescence and Fluorescence Resonance Energy Transfer (FRET) spectroscopy, since these techniques were previouslyshown to bring significant information on the MelB properties (Maehrel et al. 1998).

4.2.1 Intrinsic fluorescence spectra

As a result of the substrate binding tomelB, the tryptophan fluorescence intensity increases upon binding of melibiose, reflecting conformational changes due to its binding. The posterior addition of $Na^+$leads to further tryptophan fluorescence intensity increase due to the binding of additionalmelibiose as the melibiose affinity increases(Mus-Veteau et al. 1995). In these studies, 10 mMmelibiose and 10 mM $Na^+$(17-fold for $Na^+$and double for melibiosein the presence of$Na^+$ above the affinity constant) were added to the medium consecutively.

As shown in Figure 5.8G, the A155C mutant does not display any significant intensity variation of the Trp fluorescence upon incubation with melibiose or with $Na^+$ (in the presence of melibiose). This result is in agreement with the previous section concluding that this mutant loses all capacity to bind substrates.The mutants S153C (Figure 5.8E) and T159C (Figure 5.8H) showsmall Trp fluorescence changesformelibiose or $Na^+$addition, implying that these two mutations strongly disturb substrates binding. These results are in agreement with the studies of substrates-dependent infrared spectra changes in the previous section. Three mutants, F151C (Figure 5.8C), L154C (Figure 5.8F), and G161C (Figure 5.8N), display strongly reduced Trpfluorescence changes compared to Cys-less whether for melibiose or for $Na^+$addition.This suggeststhat these three mutations affect substrates binding, in partial agreement with the

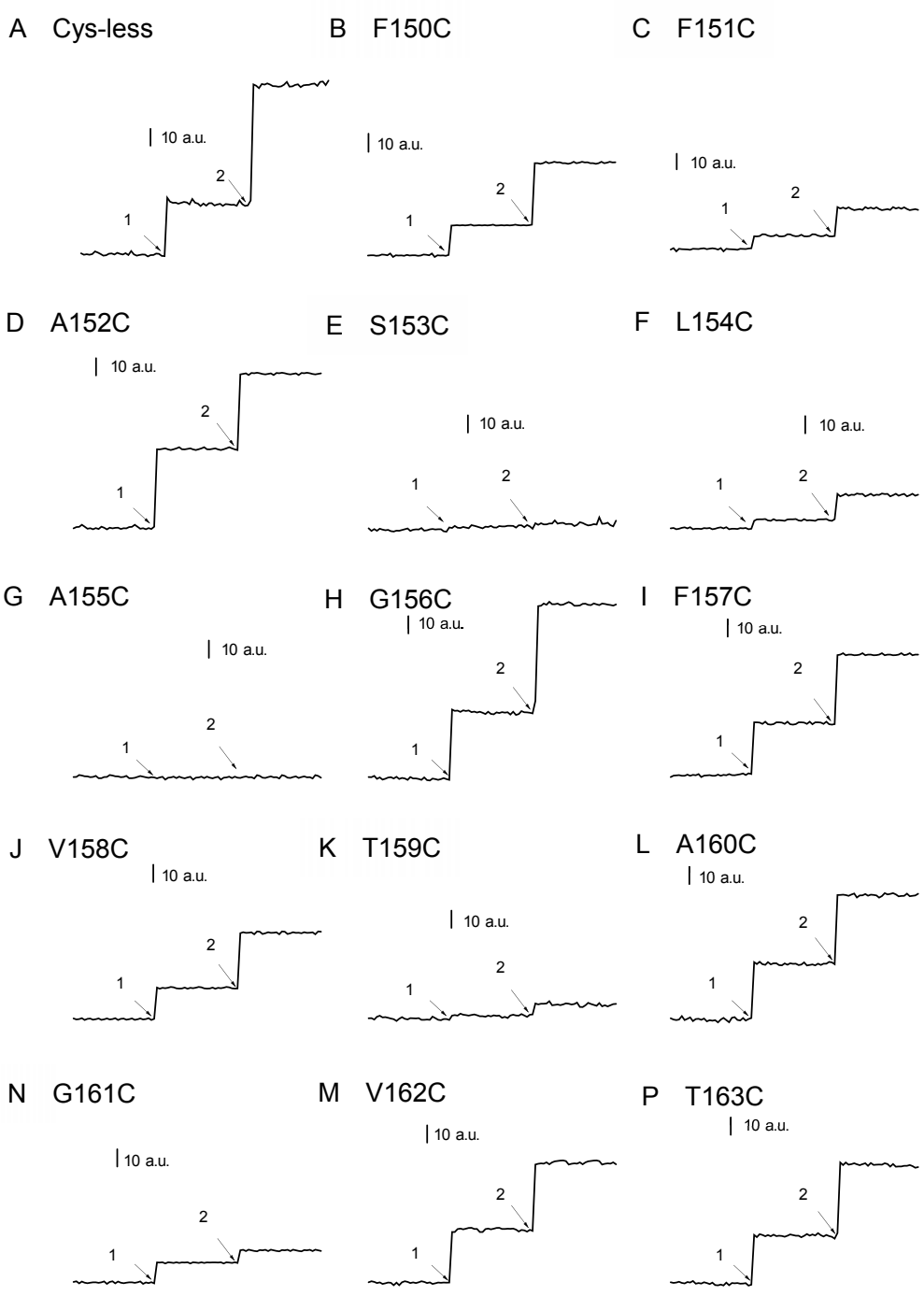



**Figure 5.8** Substrates-induced Trp-fluorescence changes of MelB mutants in proteoliposomes. (A) - (P) Representative plot of fluorescence variation of different MelB mutants induced by the addition of melibiose (1, 10 mM final concentration), followed by the addition of NaCl (2, 10 mM final concentration) where indicated (see arrows). The protein samples reconstituted in liposomes at 23 µg/ml in 100 mM KPi pH7.0 were illuminated at 290 nm (half-bandwidth=5 nm), and fluorescence was collected as a function of time. Abscissa, the fluorescence intensity variation (a.u.); ordinate, time (s).

infrared difference studies, since for G161C these studies indicate a relatively low affectation of melibiose binding. It is, however, worth noting that the intrinsic fluorescence changes are responding only to changes localized in the immediate vicinity of Trp residues; other changes not affecting Trp residues are not detected.

In the absence of $Na^+$, incubation of F150C (Figure 5.8B) and V158C (Figure 5.8J) with 10 mMmelibiose leads to an intrinsic fluorescence increase. This signal is further enhanced after 10 mM $Na^+$ addition.These substrate-induced changesare about half the intensity observed inCys-less, in concordance with the reduced melibiose affinity of F150C and V158C deduced from $IR_{diff}$experiments.The results not only confirm that F150C and V158Cretainmelibiose and $Na^+$ binding capacity, but also indicates that the reciprocal activation of the binding of one substrate by the other is maintained in these mutants inproteoliposomes, even if they look somewhat altered, in keeping with the $IR_{diff}$spectroscopyresults.The moderate effect of these mutations on the Trpfluorescence implies that Phe150 and Val158 may take part in the binding of substrates, but are not essential residues for substrates binding.

In the absence of $Na^+$, incubation of A152C (Figure 5.8D), G156C (Figure 5.8H), F157C (Figure 5.8I), A160C (Figure 5.8L), V162C (Figure 5.8M), and T163C (Figure 5.8P) with 10 mMmelibiose leads to an intrinsic fluorescence increase, which is further enhanced after 10 mM $Na^+$ addition. The intensity of these substrate-induced fluorescence changesarequite similar asthoseobserved inCys-less (Figure 5.8A), especially for A152C and G156C, in concordance with the high similarityof the $IR_{diff}$ spectra, whether induced by $Na^+$ or melibiose/$Na^+$.These resultsimply that Ala152, Gly156, Phe157, Ala160, Val162, and Thr163 are not essential residues for substrates binding.

4.2.2 Fluorescence resonance energy transfer (FRET) analysis in proteoliposomes

Substrate binding to proteoliposomescan be assessed by using FRET from Trpresidues to the bound fluorescent sugar analog $D^2G$ (excited at 290 nm) or

determining direct fluorescence emission signals from bound $D^2G$ (excited at 335 nm)(Maehrel et. al. 1998).Figure5.9A shows that, as 16 μM $D^2G$ is added, the Trp fluorescence signal (band around 330 nm) decreased in part because of FRET from Trp side chains to $D^2G$ bound to the sugar-binding site (Maehrel et al. 1998). Concomitantly, the fluorescence emission from the protein-bound fluorescent sugar appears with a maximum at 460 nm (Figure 5.9A). Subsequent addition of $Na^+$ enhances the affinity for the sugar analog and increases its bound fraction, leading to a further increase of FRET (Maehrel et al. 1998). It is also seen that addition of excess of melibiose decreases the FRET signal, indicating that the fluorescent sugar and melibiose compete for the same binding site. Figure 5.10 shows specific fluorescence emission signals from bound $D^2G$ by direct excitation at 335 nm. In this case, the protein-bound $D^2G$ emitsfluorescence with a maximum at 475nm (Figure 5.10A).As seen in FRET, and for the same reasons, subsequent addition of $Na^+$leads to a further increase of fluorescence andthe addition of excess melibiose decreases the fluorescence emission signals from bound $D^2G$. This method can be used to determine the capability of substrates binding for specific mutants (Cordat et al. 1998; Abdel-Dayem et al. 2003; Guan et al. 2011; Guan et al. 2012).

As shown in Figure 5.9, there is no clear FRET signal from Trp side chains to $D^2G$ bound to the sugar-binding site obtained from A155C (Figure 5.9G), T159C (Figure 5.9K), and G161C (Figure5.9M). The fluorescence emission signals from the protein-bound fluorescent sugar excited at 335 nmfor A155C is weak and didn't show substrates-dependent feature (Figure 5.11G). There are small fluorescence emission from the protein-bound fluorescent sugar for T159C (Figure 5.11K) and G161C (Figure 5.11M). It is clear that the fluorescent sugar analog does not bind to the A155C mutant either in the presence or in the absence of $Na^+$.T159C and G161C show a smallbinding capability for the fluorescent sugar and $Na^+$(Figure 5.9 and Figure 5.10). So, from these experiments we can safely conclude that the mutations on Ala155, Thr159, and Gly161 affect the $D^2G/Na^+$ coupled binding in proteoliposomes.

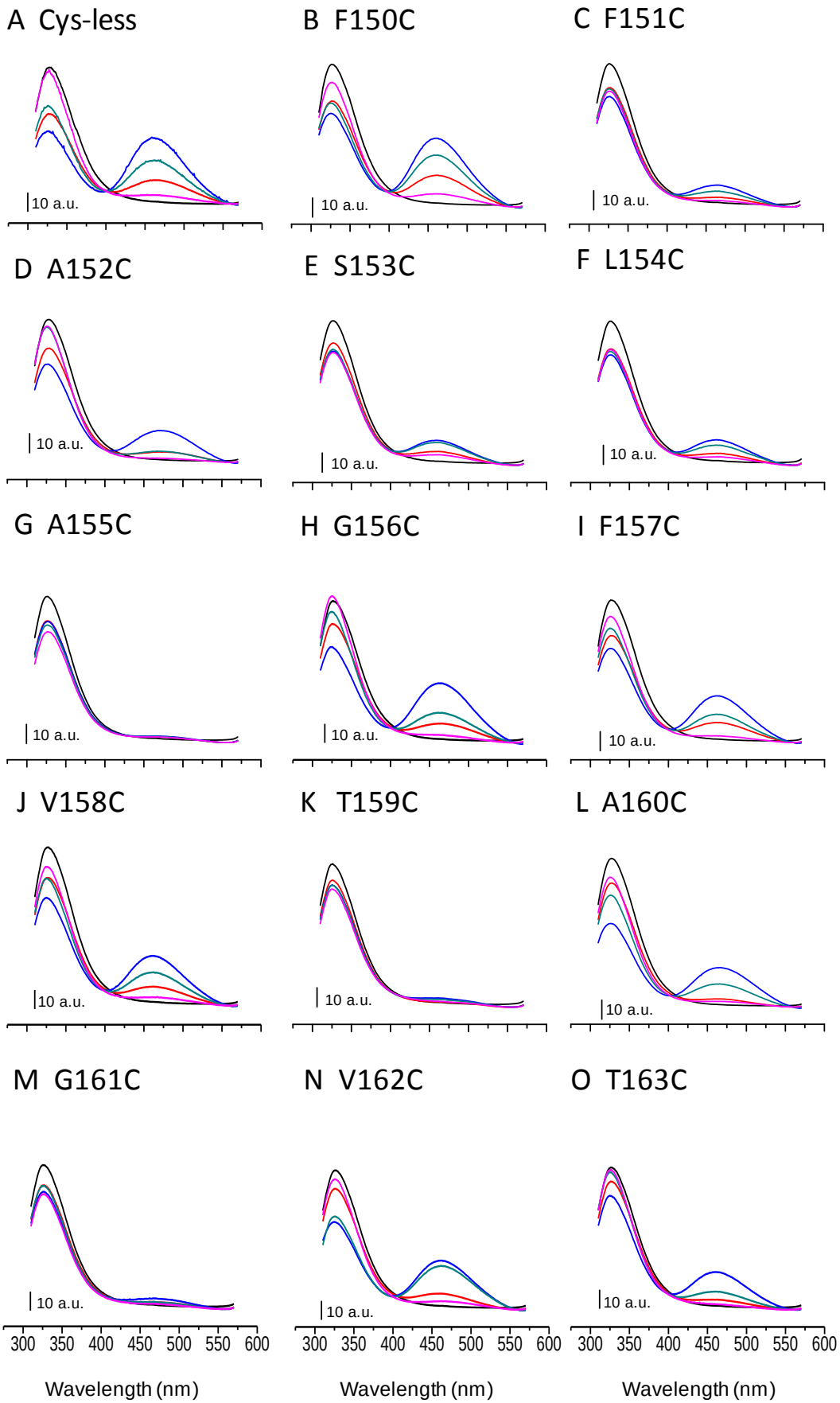

A Cys-less
B F150C
C F151C
D A152C
E S153C
F L154C
G A155C
H G156C
I F157C
J V158C
K T159C
L A160C
M G161C
N V162C
O T163C
10 a.u.
300 350 400 450 500 550 600
Wavelength (nm)

**Figure 5.9** Substrates-induced Trp→$D^2G$ FRET signal changes of MelB mutants in proteoliposomes. (A) - (P) Representative plot of fluorescence changes of different MelB mutants induced by the sequential addition of 30 μg/ml protein (black), 16 μ M $D^2G$ (red), 10 mM NaCl (blue), 10 mM melibiose (dark cyan), 150 mM melibiose (magenta). The protein samples reconstituted in liposomes at 30 μg/ml in 100 mM KPi and 100 mM KCl pH7.0 with different substrates where indicated, were illuminated at 290 nm (half-bandwidth 5), and the emission spectra were collected at 310-570 nm. Abscissa, the fluorescence intensity (a.u.); ordinate, wavelength (nm).

Three mutants, i.e., F151C (Figure 5.9C), S153C (Figure 5.9E) and L154C (Figure 5.9F), present a similar behavior of FRET,responding to the consecutive additions of $D^2G$, $Na^+$, and excess melibiose, indicating that all of them preserve the capability of sugar/$Na^+$ binding with a little reduction in the intensity, suggesting that the mutations may affect the affinity of the substrates.This conclusion was further confirmed by directly determining bound $D^2G$ emission (Figure 5.10C, E, F). For F150C, A152C, G156C, F157C, V158C, A160C, V162C, and T163C, the data show that the fluorescent sugar analog can bind tothese mutants, and that $Na^+$ retains the capacity to increase the affinity of the mutant for the sugar (Figure 5.9, Figure 5.10). Their behaviors are quite similar to Cys-less (Figure 5.10A). The mutations may affect the binding of substrates in different degree, but these side chains should not play a key role in the process of substrates binding to MelB. It is safe to conclude that Phe150, Ala152, G156C, F157C, V158C, Ala160, Val162, and Thr163 are not essential residues for substrates binding.

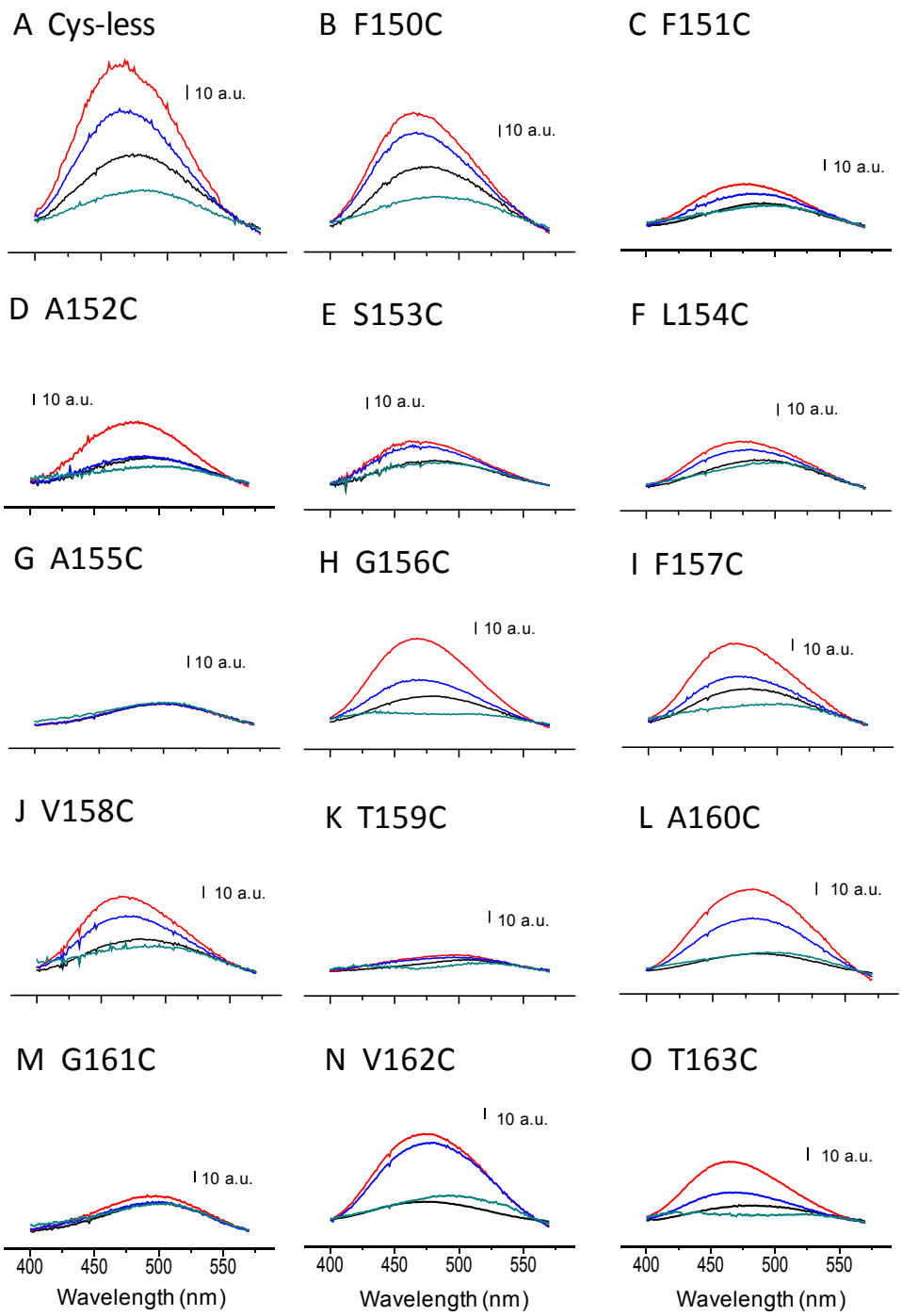



**Figure 5.10** Substrates-induced $D^2G$ emission fluorescence variations of MelB mutants in proteoliposomes. (A) - (O) Representative plot of fluorescence changes of different MelB mutants induced by the sequential addition of 16 μM $D^2G$ (black), 10 mM NaCl (red), 10 mM melibiose (blue), 150 mM melibiose (dark cyan). The protein samples reconstituted in liposomes at 30 μg/ml in 100 mM KPi and 100 mM KCl pH7.0 were illuminated at 335 nm (half-bandwidth 5), emission spectra were collected at 400-570 nm. Abscissa, the fluorescence intensity (a.u.); ordinate, wavelength (nm).

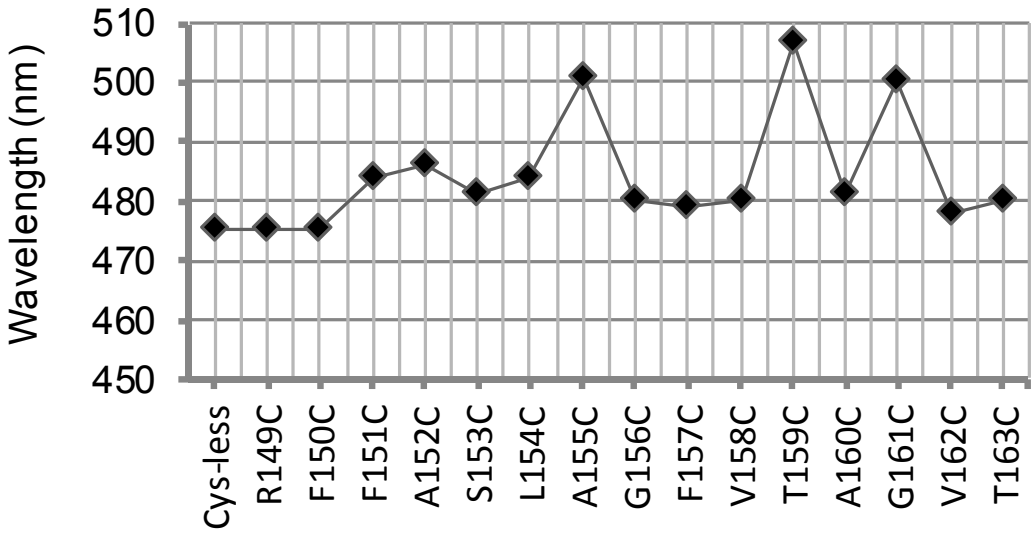



**Figure5.11** The wavelength of the maximum emission fluorescence of $D^2G$ molecules bound to MelB mutants by illuminating at 335 nm, calculated from Figure 5.10.

Figure 5.11 shows the $\lambda_{max}$ of $D^2G$ molecules bound to these mutants reconstituted into proteoliposomes by illuminating at 335 nm. Surprisingly, the $\lambda_{max}$ of $D^2G$ bound to A155C, T159C, and G161C is red shifted by up to about 25 nm from that recorded for Cys-less or the other mutants in the same condition, suggesting the probe experiences a less hydrophobic environment when bound to the these permeases, although the bound $D^2G$ emission signals from these mutants are very weak. These results suggest that A155C, T159C, and G161C mutations reduce the highly hydrophobic environment close to or at the sugar-binding site. Therefore, Ala155, Thr159, and Gly161 should be important residues in maintaining a suitable environment for sugar binding. Another explanation would be that in these mutants, $D^2G$ is not bound in the functional sugar binding site, but in a secondary site near the surface of the transporter. This would explain why there is no $Na^+$ effect and also the polar nature of the $D^2G$ environment.The

$\lambda_{max}$ of $D^2G$ molecules bound to the rest of mutants are similar to that bound to the Cys-less MelB(Figure 5.11), suggesting the probe experiences an equivalent hydrophobic environment when bound to these permeases.

4.2.3 Accessibility of the substrates-binding sites by analyzing $Na^+$-induced change of the FRET signal in vesicles

The features of the substrate binding for the cysteine replacement mutants described above have been studied in the proteoliposomes. Before the purified proteins were reconstituted into the proteoliposomes, the native protein was firstly solubilized by the use of detergents, and then subjected to a series of purification steps. In these processes, some native conformations may be lost. So the features obtained from the proteoliposomes may not really reflect the features of the native protein. In this section, we will study the substrates binding features in the native membrane, i.e. in the right-site-out (RSO) membrane vesicles, which were prepared from the cells by an osmotic process and in the inside-out (ISO) membrane vesicles, which were prepared from the cells by a Microfluoidizer treatment, following a method as described in PART I (see **PART I4.3** and ***Materials and Methods*** also).

As shown in Figure 5.12A, for Cys-less, the $Na^+$-induced FRET signal obtained from RSO and ISO vesicles after normalization for the total protein contentare nearly equal, implying that the accessibility of the substrate-binding sites from both sides of the plasma membrane are quite similar. For I262C and R139C mutations that are found in an exposed location in the periplasmic and cytoplasmic loops, respectively, the $Na^+$-induced FRET signal is similar to Cys-less, with a ratio near to 1, indicating a very similar accessibility from both sides.

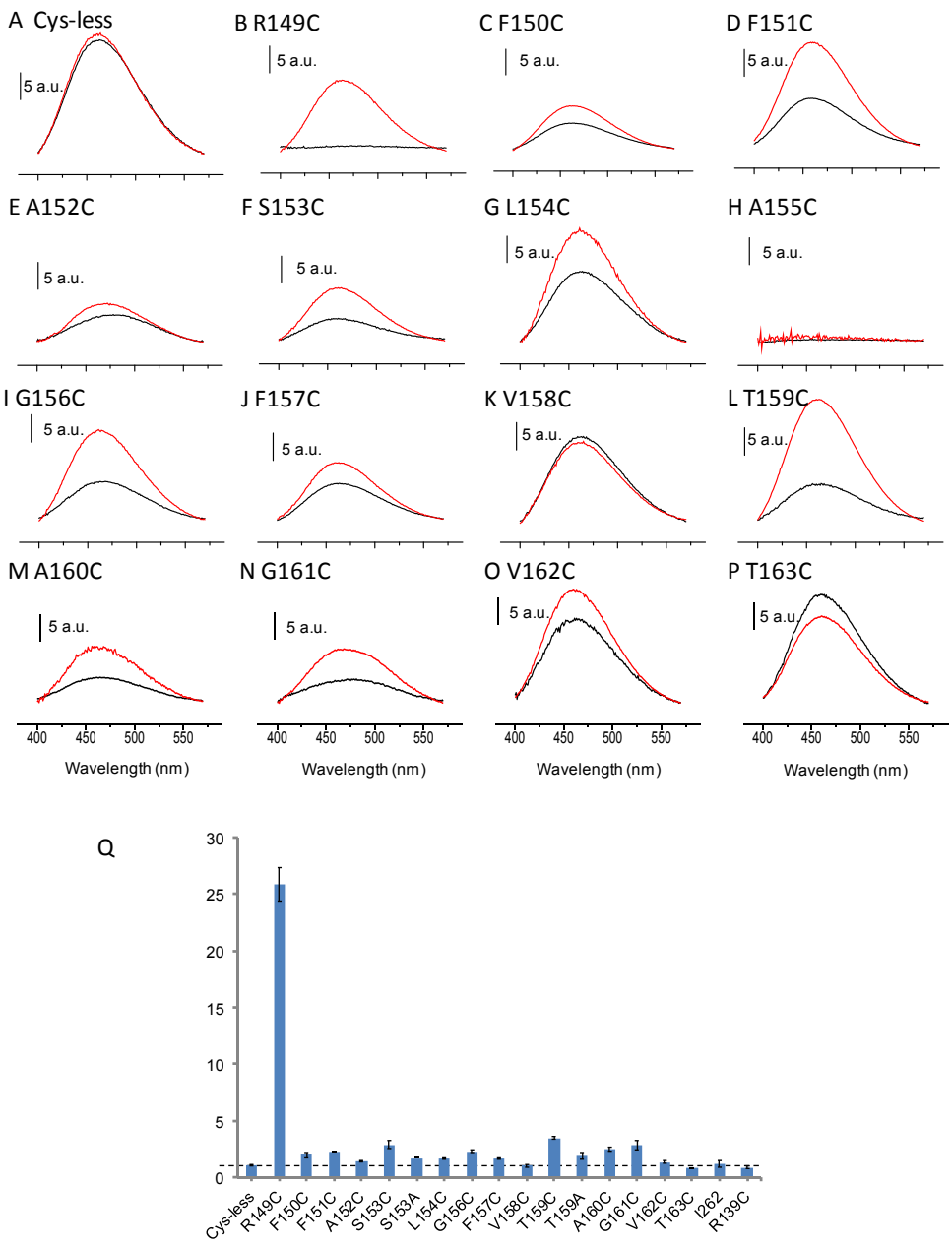



**Figure 5.12** $Na^+$-dependent variation of the FRET signal recorded from ISO and RSO membrane vesicles. ISO (red trace) or RSO membrane vesicles (black trace) were incubated with the sugar analog $D^2G$ (10 µM) and excited at 290 nm (half-bandwidth = 5 nm). The FRET signal from Trp residues to $D^2G$ recorded before the addition of $Na^+$ was subtracted from the FRET signal recorded after the addition of $Na^+$ (20 mM final concentration). The $Na^+$-dependent variation of the FRET signal presented in (A)-(P) is the average of at least three independent experiments. Abscissa, the fluorescence intensity variation (a.u.); ordinate, wavelength (nm).(Q) Histogram showing the ratio of the $Na^+$-dependent FRET signal changes integrated between 400 and 570 nm, between ISO and RSO membrane vesicles. Values are the mean of three independent experiments (±S.E.).

No significant $Na^+$-induced FRET signal was recorded for A155C in both RSO and ISO membrane vesicles (Figure 5.12H), suggesting that this mutant cannot bind any substrates in the native membrane,in agreement with results obtained by infrared spectroscopy and fluorescence spectroscopy from the samples in proteoliposomes (shownabove).

T159C displays a $Na^+$-induced FRET signal in the ISO vesicles similar to Cys-less (Figure 5.12L red line) anda decreased signal in RSO vesicles (Figure 5.12L black line). Figure 5.12 shows the ratio of the $Na^+$-dependent FRET signal between ISO and RSO membrane vesicles. For T159C this ratio is 3.6. So this mutant keeps near-native capability for the binding of the substrates from the cytoplasmic side, but shows a decreased access of the substrates from extracellular side.

The results obtained from $D^2G$ binding to T159C vesicles are in contradiction to those obtained from proteoliposomes, where almost no signal were obtained either from infrared or fluorescence spectroscopy (see **PART II4.1** and **4.2**). A similar case occurs with G161C, which displays a measurable intensity of $Na^+$-induced FRET signal (Figure 5.12N), but a negligible signal in proteoliposomes. This different 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 substrates binding.

S153C, another hydroxyl-containing residue in helix 5, displays aclear $Na^+$-induced FRET signal, but it shows a reduced intensity compared to Cys-less (Figure 5.12F). The intensity of the FRET signal induced by $Na^+$ in the two typesof vesicles isdifferent,with a ratio of 2.9. Therefore, this mutant presents a capability of substratesbinding, but with a somewhat reduced affinity, in keeping with the results obtained from proteoliposomes. The F150C mutant displays a decreased intensity of $Na^+$-induced FRET signal (Figure 5.12N) in both RSO and ISO vesicles, with an increased access in the ISO vesicles. F151C (Figure 5.12D) and L154C (Figure 5.12G) display reduced intensity of the FRET signals in proteoliposome (inside-out orientation). However, quite similar $Na^+$-induced FRET signal changes were obtained from both of these two mutants in the ISO membrane vesicles (with the same orientation in proteoliposomes).

A152C (Figure 5.12E), F157C (Figure 5.12J), and A160C (Figure 5.12M) display moderate reduced intensity of $Na^+$-induced FRET signal changes, whereas G156C (Figure 5.12I), V158C (Figure 5.12K), V162C (Figure 5.12O), and T163C (Figure 5.12P) present Cys-less like $Na^+$-induced FRET signal changes in ISO membrane vesicles, which is in agreement with the studies in the proteoliposomes.

## 4.3 Complementary analysis of two hydroxyl residues: Ser153 and Thr159

Ser153, Thr159, and Thr163 are the only polar residues of the highly hydrophobic helix 5, suggestingthat they may play important roles in the function of MelB. The studies of infrared and fluorescence spectroscopiesshow that Thr163 maynot be an essential residue for the function of MelB, since it shows similar spectra to Cys-less. The infrared spectroscopy shows that replacement of Ser153 or Thr159 by Cys reduces the intensity of the $Na^+$-induced $IR_{diff}$ spectra, and that the melibiose-induced $IR_{diff}$spectraeither in the presence of $Na^+$ or in its absence are of very low intensity and low similarity. Furthermore,intrinsic fluorescence spectra indicated that purified T159C reconstituted in liposomesis not capable of bindingsubstrates. A very important feature of Ser or Thr is that they contain a hydroxyl group, which plays important role in the formation of H-bonds.

The replacement of Seror Thrby Cys may still form H-bond because of the thiol group. To completely eliminate the possibility of the formation of hydrogen bonds, we decided to mutate these two hydroxyl-containing amino acids to Ala, i.e., S153A and T159A, which avoid all the hydrogen-bonding interactions.

4.3.1 The determinations of substrates binding

4.3.1.1Infrared difference spectra analysis

As shown in Figure 5.13A, the replacement of Ser153 by Ala displays a $Na^+$-induced $IR_{diff}$ spectrum, indicating that it retains the ability to bind $Na^+$, once reconstituted into liposomes. The resulting difference spectrum shows a high similarity (~92%) and a moderate intensity (~47%) with respect to the Cys-less difference spectrum (Figure 5.13B). Therefore, S153A preserves similarconformational changes induced by $Na^+$ binding, buta decrease of $Na^+$affinity. Comparedto the replacement of Ser153 by Cys, S153A leads to an increase in the similarity and the intensity, suggesting that Ala is more suitable for this position than Cys.

Figure 5.13E shows the melibiose-induced $IR_{diff}$ spectra in the absence of $Na^+$. S153A retains melibiose-induced $IR_{diff}$ spectra in the absence of $Na^+$ with a decreasedintensity(~35%) but similar shape (~79%) compared toCys-less (see also Figure 5.13F). Therefore, S153A preserves similarconformational changes induced by melibiose binding, with a decreased affinity for melibiose binding. Comparedwith the replacement of Ser153 by Cys, the replacement with Alaleads to an increase of both the similarity and the intensity, suggesting again that Ala is more suitable for this position than Cys.

Figure 5.13C shows the melibiose-induced $IR_{diff}$ spectra in the presence of $Na^+$ andFigure 5.13D compares the intensity and similarity of S153A and Cys-less. In the same way as the $Na^+$-induced $IR_{diff}$ spectrum or the melibiose-induced spectra in the presence of $H^+$, the replacement of Ser153 by Ala displays a melibiose-induced $IR_{diff}$ spectrum with increased intensity and similarity compared to S153C. This suggests thatAla is more appropriatefor this position than Cys for binding of melibiose in the presence of $Na^+$.

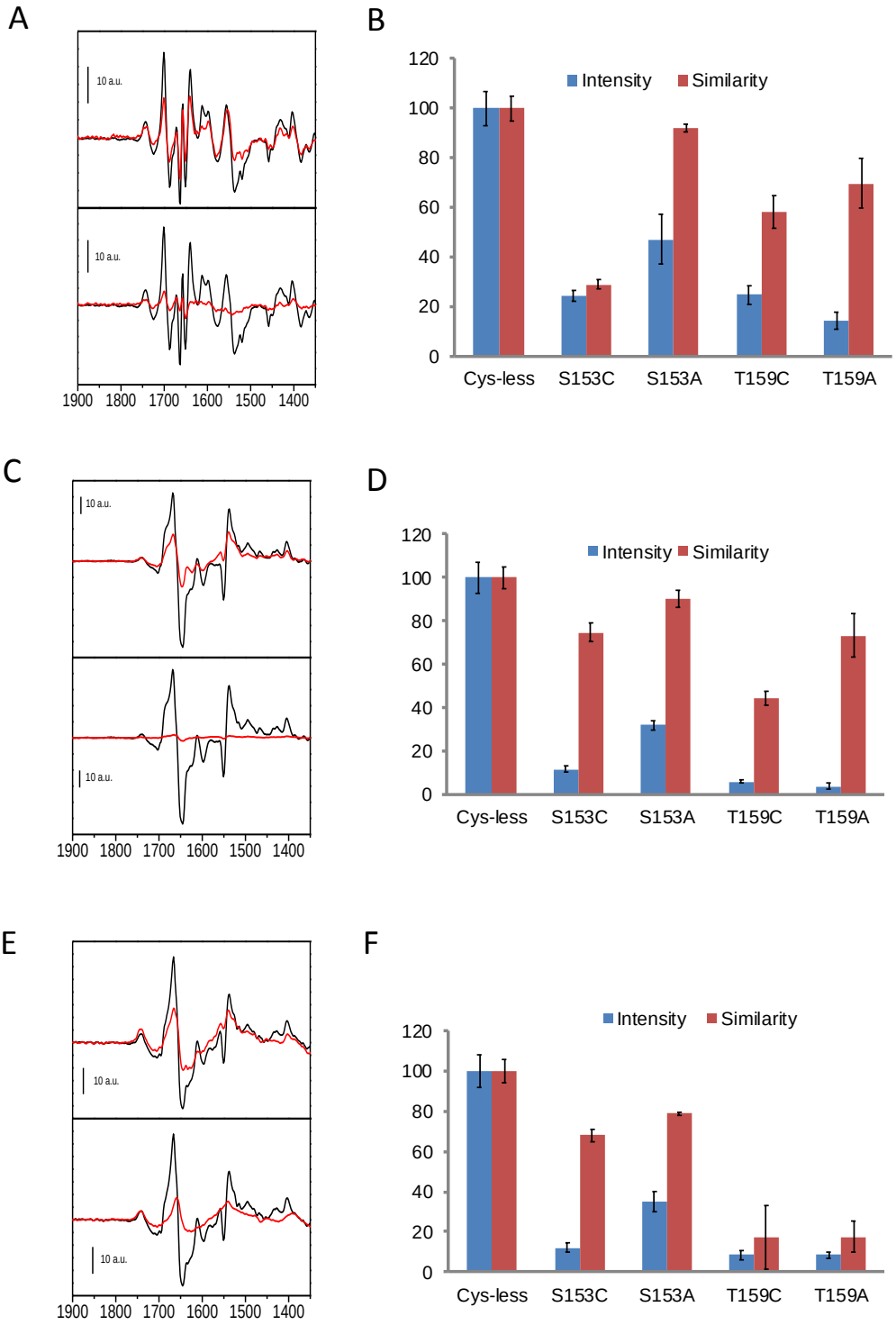

A
10 a.u.
10 a.u.
1900 1800 1700 1600 1500 1400
B
120
100
80
60
40
20
0
Intensity
Similarity
Cys-less
S153C
S153A
T159C
T159A
C
10 a.u.
10 a.u.
1900 1800 1700 1600 1500 1400
D
120
100
80
60
40
20
0
Intensity
Similarity
Cys-less
S153C
S153A
T159C
T159A
E
10 a.u.
10 a.u.
1900 1800 1700 1600 1500 1400
F
120
100
80
60
40
20
0
Intensity
Similarity
Cys-less
S153C
S153A
T159C
T159A

**Figure 5.13** Substrates-induced conformational changes of S153A and T159A. Substrates-induced $IR_{diff}$ spectra at 4 $cm^{-1}$ resolution of MelB Cys-less (black curve), S153A (red curve) and T159A (red curve) at 25 °C and pH 6.6 (20 mM MES and 100 mM KCl), were normalized to the amount of probed protein as described in ***Materials and Methods***(see Figure 3.6). The buffer exchange protocol and data acquisition are described in ***Materials and Methods***. (A) 10 mM $Na^+$-induced $IR_{diff}$ spectra of S153A (up panel) or T159A (bellow panel); (C) 10 mMmelibiose-induced $IR_{diff}$ spectra in the presence of 10 mM $Na^+$ of S153A (up panel) or T159A (bellow panel); (E) 50 mM melibiose-induced $IR_{diff}$ spectra in the absence of $Na^+$ of S153A (up panel) or T159A (bellow panel). (B) Spectral similarity and intensity of the $Na^+$-induced $IR_{diff}$ spectra. (D) Spectral similarity and intensity of the melibiose-induced $IR_{diff}$ spectra in the presence of $Na^+$. (F) Spectral similarity and intensity of the melibiose-induced $IR_{diff}$ spectra in the absence of $Na^+$. The error bar corresponds to one standard error of the mean of three independent spectra.

As shown in Figure 5.13A, the T159A mutant displays a $Na^+$-induced $IR_{diff}$ spectrum with a high similarity (~70%) and a reduced intensity (~14%) with respect to the Cys-less difference spectrum (Figure 5.13B). The high similarity suggests that the T159A mutation preserves almost all conformational changes induced by $Na^+$ binding. However, a ~30% similarity difference between T159A and Cys-less suggests that the $Na^+$-induced structural changes of T159A may be less complete that those occurring in Cys-less. This conclusion is supported by the fact that T159A lacks some intense peaks present in Cys-less (e.g., at 1,640, 1,575, and 1,550 $cm^{-1}$), whereas all peaks in T159A $IR_{diff}$spectrum are also seen in Cys-less. Comparedto the replacement of Thr159 by Cys, T159A leads to a decrease in thesimilarity and an increase of the intensity, implying that the substitution of Thr159 with Ala decreases the capability of $Na^+$ binding toMelBcomparedto the substitution of Thr159 with Cys. This suggests that the hydroxyl in Thr159 may be an important factor for the $Na^+$ binding.

Figure 5.13Eshows the melibiose-induced $IR_{diff}$spectrumin the absence of $Na^+$ andFigure 5.13F compares the intensity and similarity for T159C and T159A mutants.

Both display a small intensity and similarity with respect to the Cys-less difference spectrum. Thissuggeststhat Thr159 is required for full and native-like structural protein changes in response to melibiose binding in the absence of $Na^+$.

Figure 5.13C shows the melibiose-induced $IR_{diff}$ spectra of T159A in the presence of $Na^+$ and compares the intensity and similarity of T159C with Cys-less. Comparedto the replacement of Thr159 by Cys, T159A leads to a decrease of the intensity from ~5.7% to ~3.6%, and to asimilarity increase from ~44% to ~70%, suggesting Thr159 is very important for melibiose binding.

4.3.1.2 Fluorescence analysis

As shown in Figure 5.14A, the replacement of Ser153 with Ala preserves some capabilities of Trp fluorescence changes responding to the binding of melibiose and $Na^+$, inagreement with $IR_{diff}$ studies. Compared to S153C, the substitution of Ser153 with Ala leads to a partial recovery of the capability of substrates binding, suggesting thatthe possible hydrogen bonding interactions of residue 153 are not the key factor for substrates binding.

As shown in Figure 5.14B, the replacement of Thr159 by Ala doesn't show any clear Trp fluorescence changes responding to melibioseor $Na^+$. In agreement with $IR_{diff}$ studies, the replacement of Thr159 by AlaorCys does not recover any capabilities of substrates binding.

As in previous studies, substrate binding to proteoliposomes was also assessed using FRET from Trp residues to the fluorescent sugar analog $D^2G$ and determining directly the bound $D^2G$ emission (Maehrel et al. 1998). As shown in Figure 5.15A, up panel, there is FRET signal from Trp side chains to $D^2G$ bound to the sugar-binding site, obtained from S153A (Figure 5.15A). The addition of 10 mM NaCl produces an increase of the FRET signal at 460 nm due to more $D^2G$ binding to MelB. And the addition of 10 mM melibiose to the medium leads to a decrease of the FRET signal at 460 nm due to the replacement of $D^2G$ by melibiose. Higher concentration of melibiose addition leads nearly disappearance of the FRET signal. Specific fluorescence emission

signals from bound $D^2G$ by exiting at 335 nm also show substrate-dependent features (Figure 5.15B, top panel), and the $\lambda_{max}$ of $D^2G$ molecules bound to S153A is 475 nm, which is the same as that of Cys-less. Therefore, the substitution of Ser153 with Alapreserves the capability of substrates binding.

As shown in Figure 5.15A, bottom panel, there is no clear FRET from Trp side chains to $D^2G$ bound to the sugar-binding site, obtained from T159A. There is very little fluorescence emission from the protein-bound fluorescent sugar forT159C (Figure 5.9K) or for T159A (Figure 5.15A, bellow panel). Specific fluorescence emission signals from $D^2G$ bound to T159A by exiting at 335 nm also is weak and didn't show substrate-dependent feature (Figure 5.15B, bellow panel). Figure 5.15B shows that the $\lambda_{max}$ of $D^2G$ molecules bound to T159A is 500 nm, implying that the mutation of T159A changes the hydrophobic characteristics of sugar binding sites. Clearly, the replacement of Thr159 by Ala leads to the loosingof almost any capability of substrates binding.

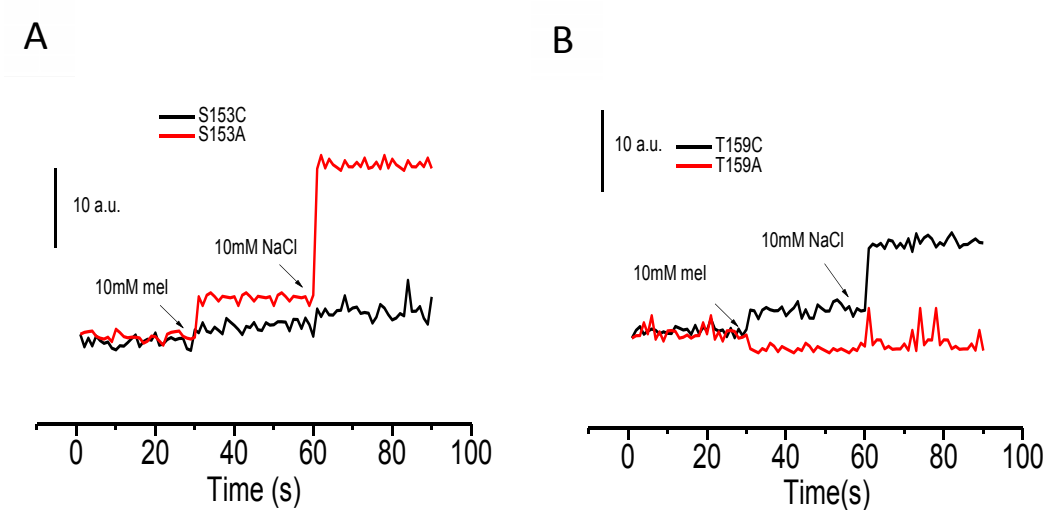



**Figure 5.14** Substrate-induced Trp fluorescence changes of MelB in proteoliposomes. (A) Tryptophan fluorescence changes in proteoliposomes ($\lambda_{ex}$ = 290 nm; half-bandwidth = 5 nm; $\lambda_{em}$ = 325 nm, half-bandwidth = 5 nm) containing purified S153C (black) or S153A (red) at 20 µg/mL in 100 mM KPi after the addition of melibiose and $Na^+$ to a final concentration of 10 mM (see arrows). (B) Tryptophan fluorescence changes in proteoliposomes ($\lambda_{ex}$ = 290 nm; half-bandwidth = 5 nm; $\lambda_{em}$ = 325 nm, half-bandwidth = 5 nm) containing purified T159C (black) or T159A (red) at 20 µg/mL in 100 mM KPi after the addition of melibiose and $Na^+$ to a final concentration of 10 mM (see arrows).

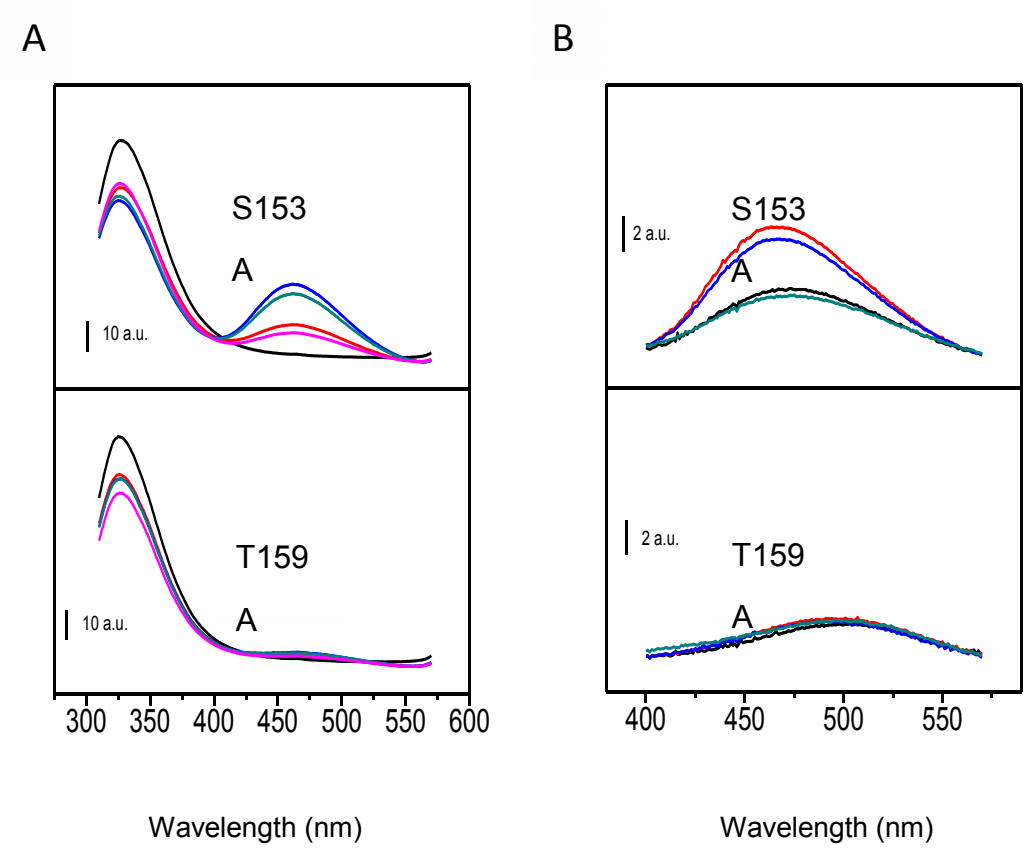



**Figure 5.15** Trp→$D^2G$ FRET signal changes of S153A (A, top panel) or T159A (A, bellow panel) in proteoliposomes induced by the sequential addition of 30 µg/ml protein (black), 16 µ M $D^2G$ (red), 10 mM NaCl (blue), 10 mM melibiose (dark cyan), 150 mM melibiose (magenta). Sample was illuminated at 290 nm (half-bandwidth 5 nm), emission spectra were collected at 310-570 nm.

$D^2G$ emission spectra of S153A (B, up panel) or T159A (B, bottom panel) in proteoliposomes induced by the sequential addition of 30 µg/ml protein (not shown), 16 µ M $D^2G$ (black), 10 mM NaCl (red), 10 mM melibiose (blue), 150 mM melibiose (dark cyan). Sample was illuminated at 335 nm (half-bandwidth 5 nm), emission spectra were collected at 400-570 nm.

Abscissa, the fluorescence intensity (a.u.); ordinate, wavelength (nm).

4.3.1.3 Accessibility of the sugar-binding sites in vesicles

Figure 5.16 shows the $Na^+$-induced FRET changes of $D^2G$ bound to S153A in ISO and RSO membrane vesicles. In the ISO vesicles the signal is quite similar to Cys-less (Figure 5.12A red line and Figure 5.16A, right), suggesting a similar accessibility of the

sugar-binding sites from the cytoplasmic side between S153A and Cys-less.Thisis in agreement with the FRET feature obtained for the proteoliposomes in ISO orientation. However, the $Na^+$-induced FRET signal changes appear somewhat decreased in the intensity in RSO vesicles(Figure 5.16B, right), suggesting that the accessibility of the sugar-binding sites from the extracellular side is somewhat impeded. In contrast, the replacement of Ser153 by Cysgives rise to ahighly decreased$Na^+$-induced FRET signal comparedwithCys-less (Figure 5.12F). Clearly, Ala is a more appropriateresidue than Cysfor the 153 position for keeping the native function ofthe substrates binding.

Figure 5.17 shows the $Na^+$-induced FRET changes of T159A in ISO and RSO membrane vesicles. In both vesicle types, T159A displays a weak $Na^+$-induced FRET signal, indicating that this mutant lose the capability of the binding of fluorescent sugar, $D^2G$, in agreement with the results obtained by infrared spectroscopy and fluorescence spectroscopy from the samples in proteoliposomes (shown above). We can conclude that the substitution of Thr159 with Ala, which avoids the hydrogen-bonding interactions, lose almost anycapability of substrates binding.

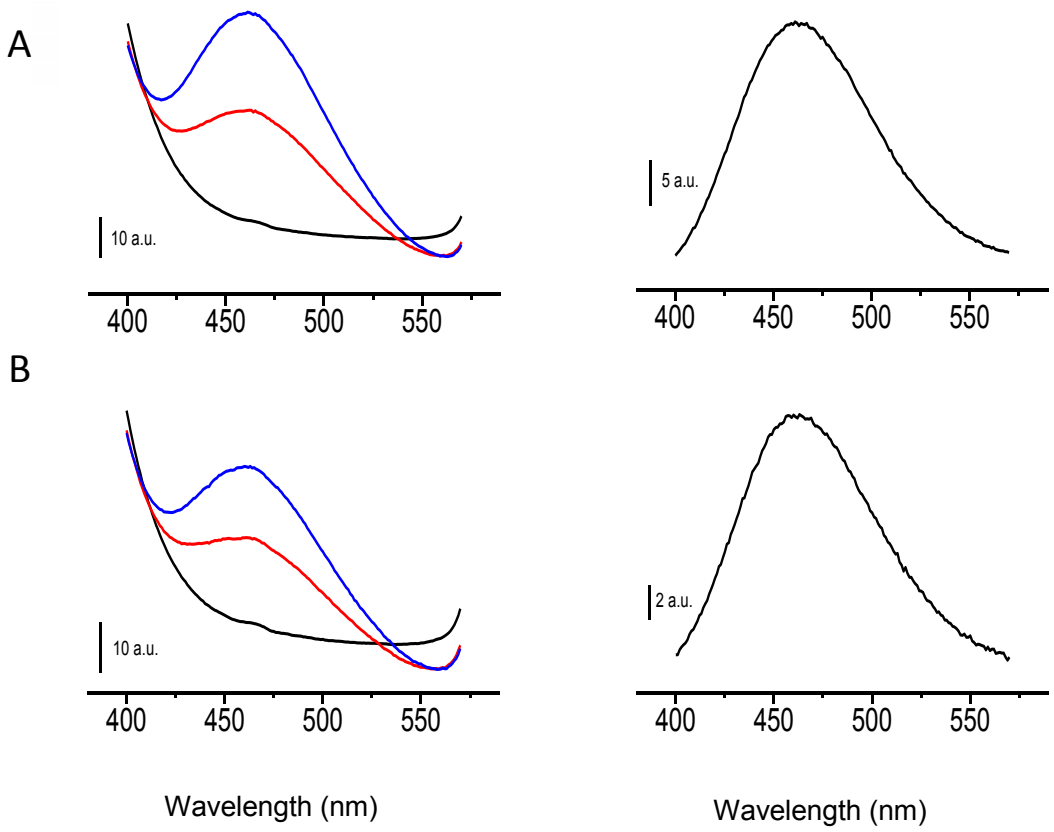



Figure 5.16 $Na^+$-dependent variation of the FRET signal recorded from S153A in ISO (A) and RSO (B) membrane vesicles. Sample was illuminated at 290 nm and emission spectra were collected at 400-570 nm. (A, left) Representative plot of fluorescence changes of S153A in ISO membrane vesicles induced by the sequential addition of 100 µg of protein/ml (black traces), 10 µM $D^2G$ (red traces), and 20 mM NaCl (blue traces).Abscissa, the fluorescence intensity (a.u.).(A, right panel) $Na^+$-dependent variation of the FRET signal of S153A in RSO membrane vesicles obtained by subtracting the FRET signal from Trp residues to $D^2G$ recorded before the addition of $Na^+$ from the FRET signal recorded after the addition of $Na^+$. Abscissa, the fluorescence intensity variation (a.u.).(B, left panel) Representative plot of fluorescence changes of S153A in RSO membrane vesicles induced by the sequential addition of 100 µg of protein/ml (black traces), 10 µM $D^2G$ (red traces), and 20 mM NaCl (blue traces). (B, right panel) $Na^+$-dependent variation of the FRET signal of S153A in RSO membrane vesicles obtained by subtracting the FRET signal from Trp residues to $D^2G$ recorded before the addition of $Na^+$ from the

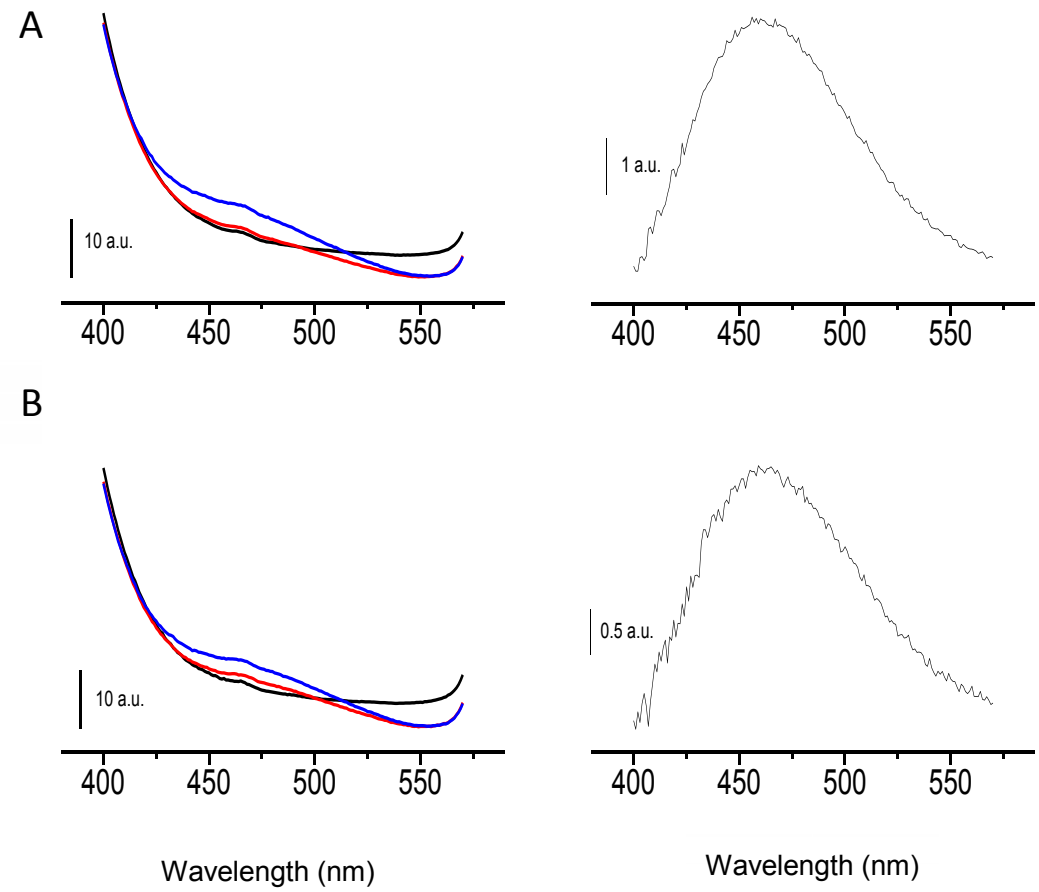



**Figure 5.17** $Na^+$-dependent variation of the FRET signal recorded from T159A in ISO (A) and RSO (B) membrane vesicles. Sample was illuminated at 290 nm and emission spectra were collected at 400-570 nm. (A, left) Representative plot of fluorescence changes of T159A in ISO membrane vesicles induced by the sequential addition of 100 µg of protein/ml (black traces), 10 µM $D^2G$ (red traces), and 20 mM NaCl (blue traces).Abscissa, the fluorescence intensity (a.u.).(A, right panel) $Na^+$-dependent variation of the FRET signal of S153A in RSO membrane vesicles obtained by subtracting the FRET signal from Trp residues to $D^2G$ recorded before the addition of $Na^+$ from the FRET signal recorded after the addition of $Na^+$. Abscissa, the fluorescence intensity variation (a.u.). (B, left panel) Representative plot of fluorescence changes of T159A in RSO membrane vesicles induced by the sequential addition of 100 µg of protein/ml (black traces), 10 µM $D^2G$ (red traces), and 20 mM NaCl (blue traces). (B, right panel) $Na^+$-dependent variation of the FRET signal of T159A in RSO membrane vesicles obtained by subtracting the FRET signal from Trp residues to $D^2G$ recorded before the addition of $Na^+$ from the FRET signal

4.3.2 Structural analysis for the mutations at Ser153 and Thr159

As mentioned in **PART II4.1.1**,the effect of the mutations on the secondary structure of MelB can be determined by a quantitative comparison of the shape of the structure-sensitive amide I and amide II protein bands of the second derivative of the absorbance spectra between1500 and 1700 $cm^{-1}$. As shown in Figure 5.1A, T159C was clustered together with F151C, S153C, L154C, and G161C, which affected strongly the function of MelB. However, substitution of Cys159 with Ala move T159A to cluster with A155C, a mutant showing a complete lack ofsubstrates binding, suggesting that this replacement further disordersthe MelB structure. More detailed structural differences can be seen in Figure 5.1B. A155C, T159A, and T159Cmutations affect a band at 1630 $cm^{-1}$, assigned to a mixture of β-sheets, $3_{10}$ helices, and open loops (Fabian et al. 1992; Arrondo et al. 1993; Jackson and Mantsch 1995).

As shown in Figure 5.1A, S153C was clustered together with F151C, T159C, L154C, and G161C, indicating that the replacement of Ser153 by Cys affects theMelBstructure.The substitution of Cys153 with Ala move S153A to cluster with F150C, F157C, A160C, T163C, V162C, A152C, and V158C, which showed insignificant effects on substrates binding of MelB, suggesting that this replacement recovers the MelB structure.

# 5 General Discussions of Part II

Previous studies showed that MelB cytoplasmic loop 4-5 is important for the binding of substrates (Abdel-Dayem et al. 2003). In the **PART I**, we discussed the role of Arg149, which is an important residue in the reorientation mechanism of MelB. A 3D model (Yoursef and Guan 2009) proposed that Helix 5 is close to the other two important transmembrane helices, i.e., helix IV, which may connectthe $Na^+$ binding site and the melibiose binding site (Poolman et al. 1996), and helix I, which is important for melibiose binding (Ding and Wilson 2001a). All of these evidences imply that helix 5 might be an important structure for the function of MelB. Thereforewe did cysteine-scanning mutagenesis of helix 5 using a Cys-less MelB, which was obtained from WT MelB avoiding all of the four native cysteines, i.e., Cys110, 310 and 364 were changed to Ser, and Cys235 to Val and displays properties very similar to those of WT. By using Cys-less as a genetic background,the effects of the one by one replacementby Cys were studied.

Taking into account the high sensitivity of infrared spectroscopy, any change in the protein structure or in the protonation state of side chains should be detected in the substrate-induced $IR_{diff}$ spectra. As in previous studies (Leon et al. 2005; Leon et al. 2006; Leon et al. 2009), the infrared difference spectra show multiple and discrete absorbance changes (in the range 1740-1400 $cm^{-1}$) of purified MelB transporter, reconstituted in liposomes, associated with binding of the cosubstrates. Some of the sharp peaks observed at specific wavenumbers were assigned to conformational changes occurring at the level of the major structural components of MelB (α-helices, β-sheets, and turns). In this study, we compared the intensity and similarity of the substrate-induced $IR_{diff}$ spectra of each mutant to Cys-less, as parameters for reflecting the capability of the substrates binding to each mutant.

Fluorescence spectroscopy was used as a complementary technique to report substrate-induced conformational changes, reflecting the capability of the substrates for binding to the transporter. As shown in previous reports, the binding of melibioseor $Na^+$to MelB leads to Trp fluorescence changes, which are further enhanced afterthe additionof the second substrate, because of the induced affinity increase upon binding of

the first substrate, reflecting the cooperational binding of melibiose and $Na^+$. This method was used to reflect the capability for substrates binding to the mutants. Substrate binding to purified MelB reconstituted in liposomes was also assessed using FRET from Trp residues to the fluorescent sugar analog $D^2G$, which can bind to MelB, but cannot be transported. As shown in previous studies of MelB reconstituted in liposomes, when$D^2G$ is added, the Trp fluorescence signal (band around 330 nm) decreases in part because of FRET from Trp side chains to $D^2G$bound to the sugar-binding site. Concomitantly, the fluorescence emission from the protein-bound fluorescent sugar appears with a maximum at 460 nm. Subsequent addition of $Na^+$enhances the affinity for the sugar analog and increases its bound fraction, leading to a further increase of FRET. In this study, we compare these FRET features responding to the subsequent addition of different substrates to report the capability of the binding of substrates to the different mutants.

In general, before a protein was reconstituted into liposomes, it was firstly solubilized from plasma membrane by detergent, and then subjected to a series of steps for purification. Some native structure may be lost in these procedures until reconstitution. So the substrates binding features reported by the $IR_{diff}$ spectra and fluorescence spectroscopy obtained from proteoliposomesmay not correspond to a totally native transporter. Therefore, we studied the binding of the fluorescent analog $D^2G$ in ISO membrane vesicles, which reflectsthe accessibility from the cytoplasmic side, and in RSO membrane vesicles, which reflects the accessibility from the extracellular side. The proteins in vesicles preserve the native structure. Therefore, the substrate binding features detected in the plasma membrane vesicles should correspond to native transporters.

## 5.1 Ala155 is an essential residue for substrates binding

Our studies suggest that Ala155, located in helix 5is an essential residue to either $Na^+$ or melibiose binding because the mutant A155Cdoesnot exhibit any structural variations upon incubation with the $Na^+$ or melibiose in the presence or absence of $Na^+$, as is deduced from the $IR_{diff}$ spectra. The results obtained from the fluorescence spectroscopy also confirm this conclusion.Nosubstrate-induced Trp fluorescence changes are detected,whereas no FRET signal was obtained under any situations with different substrates from the reconstituted transporters into proteoliposomes or from ISO or RSO membrane vesicles important for substrate binding.

From a quantitative comparison of the shape of the structure-sensitive amide I and amide II protein bands of the infrared absorbance spectra in the absence of any substrates, A155C was clustered with T159A, another important mutant thatstrongly affects melibiose binding. This indicates that each of these two mutations would disorder somewhat the MelB native structure in a similar way. It is reasonable to deduce that these structural changes are important for keeping the native function for substrates binding.

## 5.2 The role of three polar residues in helix 5, i.e., Ser153, Thr159 and Thr163

According to the Results section, the MelB mutant S153C displays a $Na^+$-induced $IR_{diff}$spectrum of low intensity, a very weak fluorescence signal and weak signals of the fluorescent sugar $D^2G$ in proteoliposomes and vesicles. Similar results were obtained for the melibiose-induced infrared and fluorescence spectra either in the presence or absence of $Na^+$, suggesting that Ser153 plays a role in the binding of substrates. To completely eliminate any possibilities of the formation of hydrogen bonds, we mutated Cys to Ala. The S153A mutant increases the ability of the binding of substrates compared to S153C, in both proteoliposomes and in vesicles. Therefore, the

replacement of Cys153 by Ala gives rise to a mutant that recovers the capability of substrates binding and of conformational changes.

We then continued to studysubstrate binding featurein either ISO or RSO membrane vesicles, which excludethe possibility that purification and reconstitution affect membrane protein native structure. The T159C mutant has a low capability of substrates binding in proteoliposomes, suggesting that Thr159 is required for full and native-like structural protein changes in response to substrates binding. However, T159C displays a strong $Na^+$-induced FRET signal changes in the ISO vesicles, which is quite similar to Cys-less (Figure 5.12L red line), indicatingT159C preserves a similar capability of substrates binding compared to Cys-less. The different results obtained from proteoliposomes and vesicles can be explained by assuming that the purification and reconstitution processes may lead to a disordering of some importantstructures, which are key in the binding of substrates. Unlike S153C, for which the substitution with Ala improves substrates binding, the results show that Ala is an unsuitable residue at position 159. The substitution of Cys159 with Ala clearly further disorders the sugar binding site or pocket. The analysis of the secondary structure (See **PART II 4.1.1**) also confirms that the substitution of Thr159 with Ala disorders some types of secondary structures of MelB. All of these data suggest that Thr159 should be an important residue for keeping the native capability of substrates binding, acting probably through the hydroxyl group.Unlike the other two polar residues, Thr163 should not be an essential residue for the $Na^+$ and melibiose binding. This is proved by the studies of FTIR spectroscopy and fluorescence spectroscopy: i) T163C mutant shows Cys-less-like similarity of $Na^+$-induced $IR_{diff}$ spectrum. ii) Melibiose-induced $IR_{diff}$ spectrum shows a Cys-less-like feature similarity and a small reduction in the intensity either in the presence or absence of $Na^+$; iii) The T163C mutant preserves a Cys-less-like behavior of the Trp fluorescence changes and the FRET signal changes induced by addition of the substrates.

## 5.3 The role of the two glycine residues, i.e., Gly156 and Gly161

Glycine is important in transmembrane helix packing. Since its side chain consists of a single H atom, glycine allows adjacent helices to approach more closely than any other residue to form stabilizing Van-der-Waals interactions between surrounding residues (MacKenzie et al. 1997). The presence of a glycine residue also exposes the polar backboneatoms of its own chain, facilitating the formation of hydrogen bonds and dipolar interactions (Javadpour et al. 1999). Based on computational analysis of several membrane protein crystal structures, it has been observed that glycine residues are frequently involved in the formation of $C^{\alpha}$-hydrogen bonds (Senes et al. 2001). The hydrogen atoms of glycine residues can form electrostatic interactions to backbone carbonyl groups of amino acids on the adjacent helix. Experimental work has very recently shown that $C^{\alpha}$-hydrogen bonds can indeed contribute to stabilize a transmembrane helix interaction (Arbely and Arkin 2004). However, there are conflicting reports on this issue in the literature (Yohannan et al. 2004), and based on computational studies it has been suggested that the formation of a stabilizing $C^{\alpha}$-hydrogen bond highly depends on the sequence context (Mottamal and Lazaridis 2005). Although this kind of interaction cannot contribute strongly to the stabilization of transmembrane helix oligomers, it could be one important force. For MelB, previous studies showed that a carboxyl group at position 117 can partially compensate for the loss of the carboxyl group at position 55 (Wilson et al. 1995), indicating there may be some kind of interaction between Gly117 in helix IV andAsp55 in helix II.Recent spectroscopic and electrical studies imply Gly117 in the reorientation mechanism (Ganea et al. 2011).

There are two glycines in helix 5. In our studies, only mutation of Gly161 with Cysaffects the capability of substrates binding. The other one, G156C preserves a Cys-less-like substrates binding feature. This conclusion is proved by the following evidences: i) For G156C, both$Na^{+}$-induced andmelibiose-induced $IR_{diff}$spectrashow Cys-less-like features not only for intensity but also for similarity, indicating that the

G156C mutant preservesthe same affinity forthe binding of the substrates and all of the conformational changesinduced by their binding; ii) The G156C mutant preserves the Cys-less-like Trp fluorescence changes induced by addition of different substrates; iii) G156C shows Cys-less like FRET signal feature in proteoliposomes. However, when we compared the accessibility of the sugar-binding sites, the G156C mutant preserves a Cys-less-like $Na^+$-dependent FRET signal only in ISO membrane vesicles, where the protein is disposed in the same orientation as in proteoliposomes. G156C shows a decrease in the intensity of $Na^+$-dependent FRET signal in RSO membrane, indicating that the substitution of Gly156 with Cys disturbs the accessibility of the substrates-binding sites from the extracellular side. As mentioned before, glycine allows the adjacent helices to approach more closely than any other residue to form stabilizing Van-der-Waals interactions between surrounding residues or exposes the polar backboneatoms of its own chain, facilitating the formation of $C^{\alpha}$-hydrogen bonds. The replacement of Gly156 by Cys may lead to a less stable transmembrane helix packing, which then disturbs the accessof the substrates to their binding sites from extracellular side. But Gly156 does not seem to participate in the reorientation mechanism of MelBbecause the ratio of the $Na^+$-dependent FRET signal changes between ISO and RSO membrane vesicles is 2.3, far lower than the value of R149C mutant (~25).

Unlike G156C, the G161C mutant shows that the capability of the binding of $Na^+$ is affected. This conclusion is proved by the followingresults: i) the $Na^+$-induced $IR_{diff}$ spectrum shows a decrease not only inthe intensity but also in the similarity, indicating that the G161C mutant affects the binding of the substrates and impairthe conformational changes induced by the binding of $Na^+$; ii) the G161C mutant shows a decreased intensity of the $Na^+$-dependent Trp fluorescence changes.

These data suggest that Gly161 playsan important role in the binding of $Na^+$ and the subsequent conformational changes. However, the G161C mutant preserves almost the complete capability of the conformational changes induced by the binding of melibiose as detected by infrared spectroscopy, implying that Gly161 is only marginally involved in the binding of sugar. The somewhat decreased intensity of melibiose-dependent

conformational changes may be due to the reducedaffinity of melibiose binding. The G161C mutant displaysa measurableaccessibility of the sugar-binding site in ISO membrane vesicles, where the protein is disposed the same orientation to proteoliposomes, with a decreased intensity. This isnot in full agreement with the substrates-dependent FRET signal features obtained from proteoliposomes, where the signal is of lower intensity. However, it is in agreement with the infrared data for melibiose binding. This suggests that the conformation of the G161C mutant in proteoliposomes suffered from a subtle conformational change that lowered the $D^2G$ affinity more than the melibiose affinity. It is possible that the different configuration of both sugars (melibiose with an α linkage and $D^2G$ with a β linkage)could explain this effect. In summary, the FTIR and fluorescence results presented in this work showthat Gly161 should play important role in the binding of $Na^+$, but should only participate marginally in the sugar binding site.

## 5.4 Phe151 and Leu154 mutants behave differently in proteoliposomes and in vesicles

Phe151 and Leu154 behave similarly with regard to their mutation by Cys. They seem to be relatively important residues for the $Na^+$ and as well as for melibiose binding in proteoliposomes, which was suggested by the studies of FTIR spectroscopy and fluorescence spectroscopy: i) they show reduced intensity and similarityof $Na^+$-dependent $IR_{diff}$ spectra; ii) theydisplaya reduced intensity and similarity of melibiose-dependent $IR_{diff}$ spectrum either in the presence or absence of $Na^+$; iii) F151C and L154Cpreserve similar substrates-dependent FRET feature and substrates-dependent Trp fluorescence feature as that of Cys-less but areduced amplitude. The intensity of the emission fluorescence of $D^2G$ upon direct excitation shows a relatively large decrease, suggesting adecreased affinity of $D^2G$ to MelB.

However, when we detected the accessibility of the sugar-binding sites in vesicles, F151C and L154Cdisplayeda strong $Na^+$-induced FRET signal in the ISO vesicles, similar to Cys-less (Figure 5.12D and G, red line), implyingasimilar accessibility of the sugar-binding sites from cytoplasmic site. The different conclusions between the results obtained from proteoliposomes (strong effect upon substrates binding) and vesicles (similar to native feature in ISO) imply that the purification and reconstitutionprocesses may lead to the loss of some important conformations, which are key in the binding of substrates. Therefore, Phe151 and Leu154 do not seem to be important residues for substrates binding of MelB.

## 5.5 Phe150, Ala152, Phe157, Val158, Ala160, V162, and Thr163are not essential residues for the substrates binding

Mutation of Phe150, Ala152, Phe157, Val158, Ala160, V162, and Thr163by Cys gives rise to proteins that retain similar spectral features as Cys-less in all assays. Therefore, these side chains are not involved in the binding of substrates or in the ensuing conformational changes.

## 5.6 Helix 5was involved in substrates binding

The interspectral distance between different mutants reconstituted in liposomes under different situations reflects the effect of the mutations on the substrate-dependent conformational changes. For $Na^+$-dependent $IR_{diff}$ spectra, R149C,F151C, S153C, A155C, T159C, T159A, and G161C were clustered together with spectral distances higher than 37% against Cys-less, whereas F150C, A152C,S153A,L154C,G156C, F157C, V158C, A160C,A162C, and T163Cpresent interspectral distances lower than 22% (Figure 5.3).This suggests that the mutations R149C,F151C, S153C, A155C,T159C,

T159A, and G161Cwould affect $Na^+$ binding. Therefore, Arg149, Phe151, Ser153, Ala155, Thr159, and Gly161 may be involved in $Na^+$ binding or in the architecture of the $Na^+$-binding site.

For melibiose-dependent $IR_{diff}$spectra either in the presence or in the absence of $Na^+$, R149C, L154C, A155C, and T159C were clustered together with spectral distances higher than 40% against other mutants, which present interspectral distances lower than 20% (Figures 5.5 and 5.7).T159A and S153C show also high interspectral distances for melibiose binding in the absence or the presence of $Na^+$, respectively.Therefore, from the analysis of MelB mutants reconstituted into liposomes, we conclude that Arg149, Ser153, Leu154, Ala155, and Thr159 may be involved in melibiose binding or in the structural design of the binding site.

On the other hand, the analysis of the accessibility of the substrates-binding sites in vesicles showed that F151C, L154C, and T159C preserve the Cys-less like fluorescent sugar analog binding features (See Figure 5.12). We may conclude that Arg149, Ser153, Ala155 andThr159 may be involved in substrates binding or in the building of the binding site,and Gly161 may be additionally involved in $Na^+$ binding. As commented in other sections in this work, it is again noticed that the purification and reconstitution processes can be deleterious to some mutants, since these mutations have adverse effects on the substrates-binding ability in proteoliposomes, not seen in vesicles.

## 5.7 The relation of structure and function

From a quantitative comparison of the shape of the structure-sensitive amide I and amide II protein bands of the infrared absorbance spectra, F151C, S153C, A155C,T159C, T159A and G161C, on one side, and R149C, F150C, A152C, S153A,G156C, F157C, V158C,A160C, T163C, V162C, and Cys-less on the other side were clustered together, indicating that the secondary structures of these two groups of mutants display high similarity among them. In addition, the studies of

substrate-induced FTIR difference spectroscopy and fluorescence spectroscopy have shown that Phe151, Ser153, Gly161, Thr159, and Ala155 are important residues for keeping the native function of MelB. On the other hand, A152C, S153A,F157C, V158C, A160C, V162C, and T163C, preserve very similar substrates binding features as those of Cys-less.

It is important to recall that these results were obtained with purified MelB incorporated into proteoliposomes and that the behavior of some MelB mutants in vesicles may be different, as commented before. One explanation for this different behavior is that the purification and reconstitution procedures alter is some manner the mutant structure (see comments in **PART II 5.2** and **5.4**). Therefore, the results obtained from a few reconstituted samples come from partially altered proteins, compared to the same native proteins in vesicles. Even this, these features can serve to gain information about MelB behavior. In this line, the results imply that the changes of the secondary structures may be the main reason that contributes to the observed functional changes. Furthermore, we have shown that the substitution of Cys153 with Ala restored some of the native-like substrates binding features. From the quantitative comparison of the shape of the structure-sensitive IR bands, we can see that S153C and S153A were separated into different clusters, indicating that this replacement affects somewhat the MelB structure and that this affects in turn the binding of substrates. On the other hand, the replacement of Cys159 by Ala that prevents all polar interactions, further affects substrates binding. When we analyzed the structures of these two mutants, they were again separated into two different clusters. T159A was clustered together with A155C, which lost all of the capability of substrates binding whereas T159C was clustered together with L154C and G161C, which preserve somewhat substrates binding. Clearly, the substitution of Cys159 with Ala further disorders the structure of MelB, which then prevents the binding of substrates.

As the structural differences between these mutants are very small (less than 2.5%, see Figure 5.1), it may be reasonable to imagine that these changes of the secondary structures may happen in the functional domains only. Then we may conclude that

Phe151, Ser153, Leu154, Ala155, Thr159and Gly161 are located in the (or close to) functional domains of MelB.

## 5.8 Helix 5 was involved in the reorientation of MelB

A method, developed in PART I, has been shown to be useful for determining the key residues involved in the reorientation of the protein. In this part, we used this method to determine the involvement of helix 5 in the reorientation of MelB. Fig12Q shows the ratio of the $D^2G$ accessibility to the sugar-binding site for ISO and RSO membrane vesicles for all of mutants. Besides R149C, which may be a key residue for the reorientation of MelB, all the other mutants show a ratio larger than 1 except V158C and T163C. Although not as prominent as R149C (ratio ~25), all these mutations from helix 5 give rise to a decrease of the accessibility of the sugar analog from the extracellular matrix. These interesting results may imply that helix 5 or part of it is involved in the reorientation mechanism of MelB.

# 6 Conclusions of Part II

1) Ala155, located probably in the middle of helix 5 is an essential residue for either $Na^+$ or melibiose binding, since the mutant A155C absolutely loses the capability to bind substrates.

2) Some mutants reconstituted into proteoliposomes show hindered substrates binding, like T159C or G161C. However, the same mutants show interactions with substrates in membrane vesicles, which are relatively akin to Cys-less. Therefore, the purification and reconstitution processes lead to some alterations of the MelB structure for these particular mutants.

3) Cysteine-scanning mutagenesis combined with the determination of the sugar accessibility to binding sites shows that many cysteine mutations disturb the sugar accessibility to binding sites from the extracellular matrix. Therefore, the helix 5 should be involved in the reorientation mechanism of MelB from the outward-facing to the inward-facing conformation.

# Part III Crystallization and preliminary X-ray diffraction studies of the melibiose permease

# 4 Results and Discussions of Part III

Because the highly conformational dynamics of secondary transporters and their condition of membrane proteins, it is very difficult to crystallize them and get a good resolution crystal. In general, finding a mutant that may cause a more compact structure and decreased conformational flexibility, is a good way to facilitate the crystallization of transporters. A good example corresponds to lactose permease (LacY), which catalyzes the coupled symport of a galactopyranoside and an $H^+$ and is a paradigm for the major facilitator superfamily (MFS) of membrane transport proteins. LacYwild type couldn't be crystallized at the beginning. However, they found one mutant, C154G that causes a more compact structure and decreased conformational flexibility, an alteration that specifically blocks the structural changes necessary for substrate translocation with little or no effect on ligand binding(Smirnova and Kaback 2003). Kaback and his colleagues were successful to crystallize LacY using this mutant. In spite of that, they finally crystallized LacYwild type after additional time and hard work. Clearly, finding a good mutant will facilitate the crystallization.

The model of alternating access is a basic mechanistic explanation for the transport function of the secondary active transport, and has been supported by numerouskinetic, biochemical and biophysical studies. According to this model, the transporter exposes its substratebinding site(s) to one side of the membrane or the other during transport catalysis, requiring a substantialconformational change of the carrier protein. Typical states include outward-facing, occluded, or inward-facing conformations. Clearly, to understand how a transporter protein translocates its substrates through the membrane, it is important to get different states structures. Therefore, it is important to crystallize mutants, which trap the protein in a state.

Our previous studies (**PART I**) and results from others identify some mutants suitable for crystallization trials. One of the most interesting isthe R149C mutant, which preserves the capability of substrates binding, but traps the protein in an inward-facing conformation. It seems reasonable to suppose that this mutant will decrease the conformational flexibility of the protein, and then facilitate the crystallization of MelB. Therefore, we decided to use this mutant for crystallization. Previous studies have

shown that the R141C mutant may display an occluded conformation, in which MelB can bind substrates but cannot translocatethem (Abdel-Dayem et al. 2003 and Leon et al. 2009). So, R141C is another mutant of interest for crystallization that may present an occluded conformation. R149Q may disturb the accessibility of the sugar binding site from the extracellular side, but not from the cytoplasmic side. This suggests that R149Q may present a higher facilityfor an outward-facing conformation. The K377C mutant lost all of the capabilities of substrates binding (unbublished work). Using these mutants for crystallization, one may expect to get different states MelB structure. Combining all the structures of these mutants, we may explain the translocation mechanism of MelB.

## 4.1 Crystallization of MelB R149C mutant

### 4.1.1 Protein preparation

MelB and the mutants were cloned, expressed and purified similarly as previously reported in ***Materials and Methods***. Following this method, 8 L of *E.coli* cell culture typically produced 15 g of wet cells, resulting in 5 g of membranes after cell fractionation. From this, 10-15 mg of the MelB R149C mutant could be eluted from a $Ni^{2+}$-NTA affinity column, and 8-12 mg of pure, stable and monomeric R149C was recovered from the size-exclusion column. We observed that the concentration of the detergent used for solubilizing protein from plasma membrane is essential for improving yield. For example, when we increased the concentration of LAPAO from 0.1% to 0.2% (w/v), the yield of protein will increase to double. The purified MelB was concentrated using a 50kDaAmicon Ultra-4 filter (Millipore) and subjected to a buffer exchange process to control the concentration of detergent in protein solution (see **PART III4.1.4.1)**. In this step, about 30% protein can be lost.

For protein labeled with seleno-L-methionine, transformed bacteria (B834 DE3) were grown in LeMastermedium (L-methionine replaced by seleno-L-methionine), and the protein was purified as abovewith 10 mM DTT added to all buffers. In general, 8 L

of cell culture typically produced 10 g of cells. From this, 3 mg of SeMet-R149C MelB could be eluted from a $Ni^{2+}$-NTA affinity.

4.1.2 Characterization of oligomericstates

Monodispersity, purity, and identity are characterized by gel filtration chromatography during the last step of the purification in an ÄKTA Purifier (GE Healthcare), by coomassie blue staining SDS-PAGE Gels, silver staining SDS-PAGE gel, and Native PAGE gel.

Gel filtration chromatography separates proteins on the basis of size. Molecules move through a bed of porous beads, diffusing into the beads to greater or lesser degrees. Both molecular weight and three-dimensional shape contribute to the degree of retention. In this study, size-exclusion chromatography was used to separate other contaminants. For good separation during size-exclusion chromatographic analysis, the protein was applied to the column at ~5 mg/ml concentration. MelB fractions (5 mL) were collected. In Figure 6.1 we can see the elution profile of MelB in 0.1 % (w/v)DDM. The symmetry and shape of this profile points to a hydrodynamic behavior ideal for crystallographic studies. Figure 6.2 show coomassie blue staining SDS-PAGE Gel and silver staining SDS-PAGE Gel for the sample which was ready for crystallization studies. The protein purity was estimated at 95-98%.

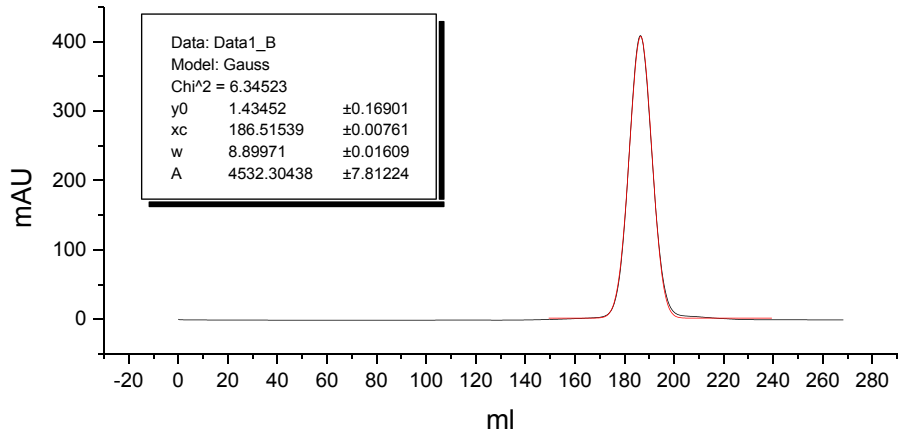



**Figure 6.1** Chromatogram of a MelB *E. coli* sample purified in 0.1 % DDM, measured at 280 nm in a gel filtration column (Superdex200). The curve fits a Gaussian function with a $chi^2$ of 6.3 and the maximum is at 186.5 ml.

Previous studies have shown that when the protein purity is higher than 90%, this should be suitable for crystallization studies. However, the homogeneity of the protein is a very important indicator for protein crystallization. Only very high homogeneous proteins can be crystallized. Native PAGE electrophoresis is run in non-denaturing conditions. It can be used to determine oligomeric states of the test protein. Figure 6.3 shows the native PAGE for purified R149C MelBin the solubilized state or reconstituted in *E. coli* lipids. Only one band was detectedin this native PAGE, indicating that purified R149C MelB just presents single state. As all of the samples used in this study were stocked at -80ºC. Based on the result obtained from native PAGE, we may conclude that in spite that protein was subjected to freeze-thaw-freeze cycles, the aggregation state of the protein was not changed (Figure 6.3).

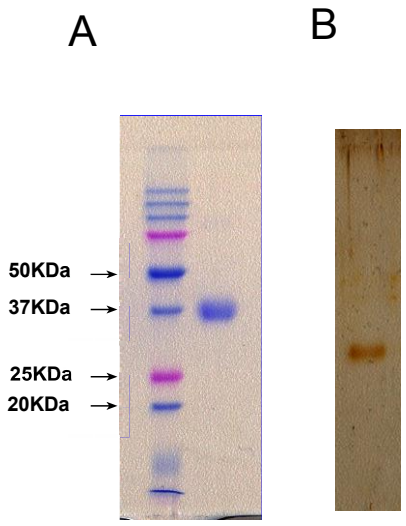



**Figure 6.2** SDS-PAGE gels.
(A) Coomassie blue staining.
(B) Silver staining.

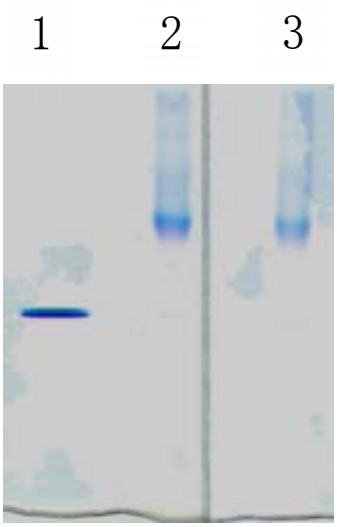



**Figure 6.3** Native PAGE.
1) 10 µg lysozyme.
2) 10 µg fresh purified R149C in 0.017% DDM.
3) 10 µg frozen R149C in 0.017% DDM.

4.1.3 Trp fluorescence and Trp→$D^2G$ FRET

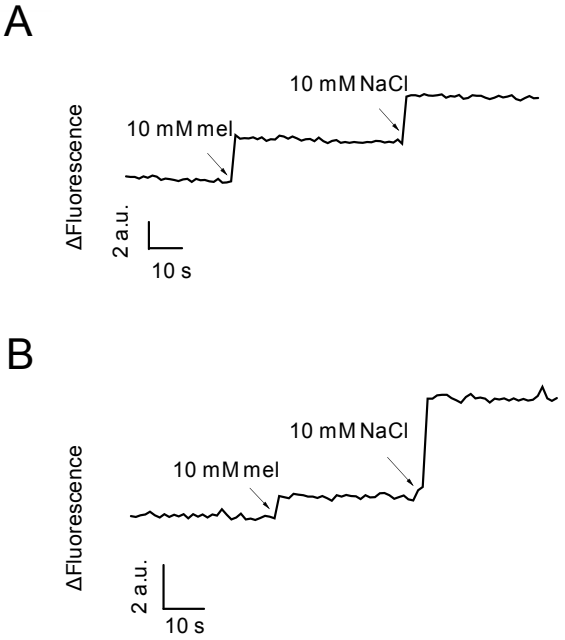



**Figure 6.4** Substrates-dependent Trp fluorescence changes of R149C.
(A) Tryptophan fluorescence changes of purified R149C in the buffer containing 0.017% DDM at 20 μg protein/ml ($\lambda_{ex}$ = 290 nm and $\lambda_{em}$ = 325 nm) after the addition of sugar and $Na^+$ to a final concentration of 10 mM
(B) Tryptophan fluorescence changes of purified R149C reconsitituted in lipid liposome.

To assay the integrityof the R149C MelB in detergent solution, its tryptophan fluorescence changes upon substrate binding was measured as described forproteoliposomes (Figure 6.4). R149C MelB in detergent solution displays similar substrates-dependent Trp fluorescence changes feature as that of R149C reconstituted in liposome, indicating that R149C MelB in detergent solution preserves the capability of substrates binding and substrates-dependent conformational changes. However, the amplitude of Trp fluorescence changes induced by the presence of melibiose along or both melibiose and $Na^+$ appears somewhat decreased, which may be due toa looser structure in the protein-detergent complex. Another important probe of integrity is to

detect the accessibility of the fluorescent sugar analog $D^2G$ to sugar-binding site. Figure 6.5 shows substrates-dependent Trp→$D^2G$ FRET feature of the R149C MelB in detergent solution. Unlike in proteoliposome, R149C MelB only displays a small FRET signal with a clear red shift. In agreement with the results obtained from Trp fluorescence, the R149C MelB in detergent solution displays a relatively loose structure although it preserves a certain capability of $Na^+$ and sugar binding.

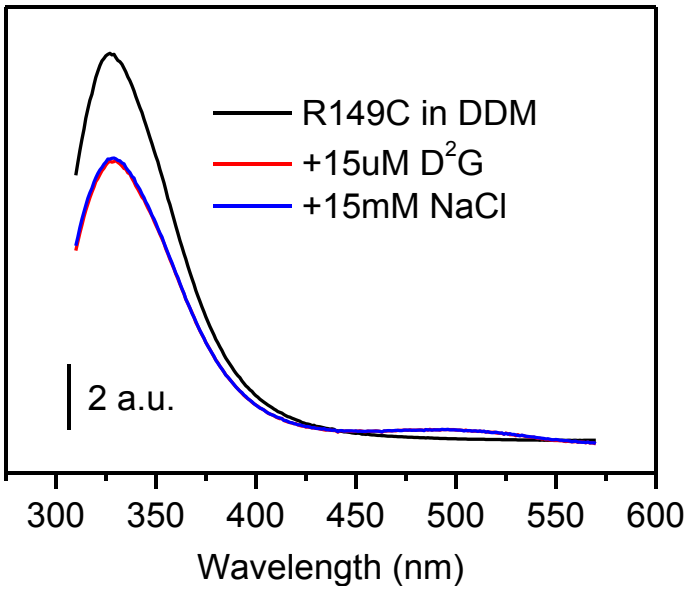



**Figure 6.5** FRET signal between R149C tryptophans and the fluorescent sugar analog $D^2G$ ($\lambda_{ex}$ = 290 nm)recorded in the buffer containing 0.017% (w/v) DDM.

4.1.4 Crystallization and preliminary X-ray diffraction

4.1.4.1 The effect of the concentration of detergents

Detergentsare quite easily co-concentrated during sample protein preparation when a Molecular Weight Cutoff Filter is used, particularly for those detergents with high

micellar average molecular weight. A too high concentration of a detergent will lead to phase separation. Therefore, caution should be taken to not allowing high detergent concentrations.Before being subjected to crystal screening, purified R149C MelB was concentrated using an AmiconUltra 15 centrifugal concentrator (Millipore) with 50kDa or 100 KDa molecular weight cutoff. When comparing these two type filters, we noted that more protein passed though the 100 KDa molecular weight cutoff, which leadsto the loss of protein.Therefore, we went to use the 50 KDa molecular weight cutoff. To get an exact concentration of detergent and to avoid co-concentration of the detergent, the protein sample was buffer exchanged by the batch affinity method by using a small volume of Ni-NTA resin (in a ratio of 15mg protein/1 ml Ni-NTA resin) and centrifuging the sample in the smallest amount possible ofthe elution buffer which contained 0.017% (w/v)DDM, 100 mMNaCl, 5% (v/v) Glycerol, 20 mMTris-HCl pH8.0, 5 mM DTT, 10 mMmelibiose, and 300 mM imidazole (final protein concentration of 5-10 mg/ml). Then the sample was dialyzed against the crystallization buffer containing 0.017% (w/v)DDM, 100 mMNaCl, 5% (v/v) Glycerol, 20 mMTris-HCl pH 7.5, 2 mM DTT, 10 mMmelibiose, and 0.5 mM EDTA,for 3h to remove the extra imidazole using a MINI Dialysis Devices (20K MWCO, Thermo SCIENTIFIC).Following this process, the concentration of the detergent can be controlledtothe expected value.

4.1.4.2 Crystal preparation

The protein solution was clarified by centrifugation (100,000g, 4℃, 30 min) before starting the crystallization trials. Purified R149C MelB (5-10 mg/ml in a buffer containing 0.017% (w/v)DDM, 100 mMNaCl, 5% (v/v) Glycerol, 20 mMTris-HCl pH 7.5, 2 mM DTT, 10 mMmelibiose, and 0.5 mM EDTA) was initially screened usinga Phoenix RE robot, which sets up the crystallization experiment in 96 well plates using volumes down to 100 nL, against the commercial protein crystal growth screensMemGold, MemSys, MemStart, JCSG-plus and PACT (Molecular Dimensions)at both 18 ℃ and 4 ℃.After several attempts of crystallizing the protein during a relatively long time, initialneedle-sharp crystals of R149C MelB in DDMwere

observed in condition 30 of the JCSG-plus crystal screen (Figure 6.6) at 18ºC. We then repeated this condition (changing both the concentration as well as the molecular weight of the PEG used) using the hanging drop vapor diffusion techniqueby mixing 0.7 μL of protein and 0.7μL reservoir solution at 18ºC.Crystals of tetragonal rod appeared from 46-55% (w/v) PEG200 and 29-35% (w/v) PEG400 as precipitant in about one week and grew to full size in about one month (Figure 6.7). They were frozen directly in liquid nitrogen using PEG as cryoprotectant.

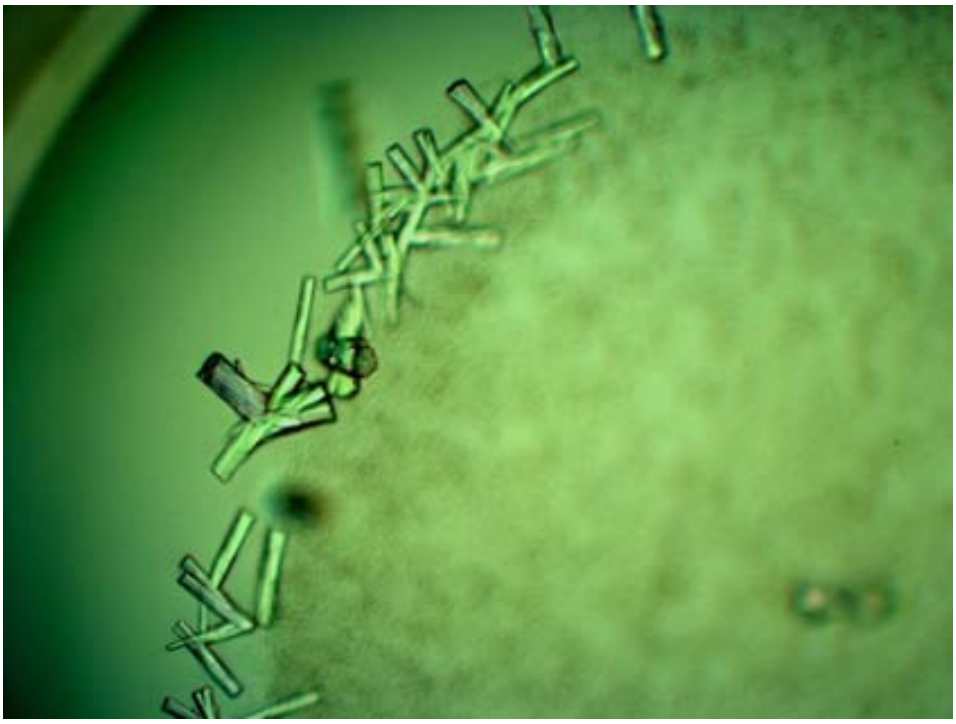


**Figure 6.6** Initial needle-sharp crystals of R149C MelB in DDM observed in condition 30 of the JCSG-plus crystal screen.

We recognized at that time that the main reason why we did not succeed to crystallize MelB was that we suffered from excessive co-concentration of the detergent when concentrating the protein right after the gel filtration procedure. Due to the large dilution effect of the Superdex200 column, we had to concentrate the sample 50-100 times, which in turn also concentrated the detergent to a final concentration of 0.5-2% from its

initial 0.01-0.02%. All the commercial crystallization screens evaluated with several forms of MelB at this high detergent concentration only yielded spherulite like pseudo-crystalline forms that did not diffract. By using the protein sample containing the desireddetergent concentrationobtained following the protocol described above, we were able to obtain diffracting crystals of R149C MelB. The condition used for crystallization, which is similar to those used to crystallize membrane transport protein (see Table 6.1) is: To form the drops, 1 ul of protein solution in 0.017% (w/v)DDM, 100 mMNaCl, 5% (v/v) Glycerol, 20 mMTris-HCl pH 7.5, 2 mM DTT, 10 mMmelibiose, and 0.5 mM EDTA at a concentration of about 7-10 mg/ml was mixed with an equal volume of reservoir solution containing 30-50% (w/v) PEG 200-400 and 100 mMphosphate/citrate pH 4.2 by the hanging drop method at 18 ºC. Hexagonal rods appear after 3 days to that range from 10 to 100 μm.

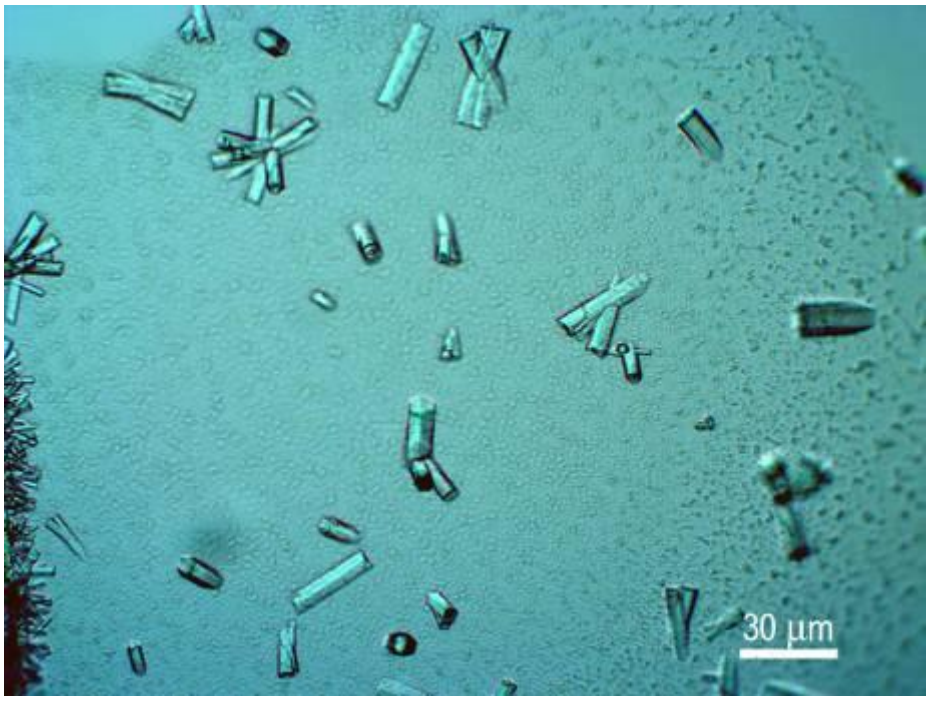



**Figure 6.7** Crystals of R149C mutant form of *E.coli* MelB in the PEG 200 condition.



A B

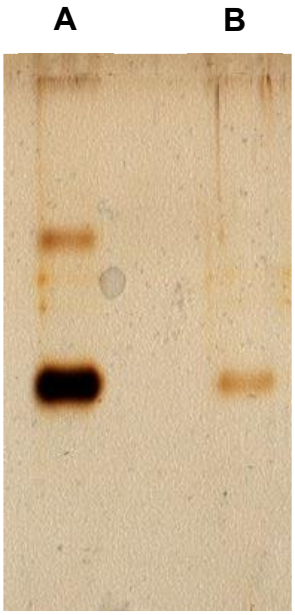



**Figure 6.8** Silver staining SDS-PAGE gel of R149C crystals. A) Crystals; B) control: R149C in detergent containing solution.

Crystals were washed, dissolved, and their identity verified by silver staining SDS-PAGE Gels (Figure 6.8A). Crystals lead a band at same position as that of MelB in detergent-containing solution, indicating that the crystals obtained in this study contained the MelB protein. Some bigger crystals were then mounted in cryo-loops and diffracted at 100K in the Plataforma de Cristal.lografia (ParcCientífic de Barcelona, Barcelona). Using the rotating Cu anode X-ray source at several exposure times and mar345 image plate detector, we obtained the best macromolecular diffraction to ~17 Å (Figure 6.9). The very approximate cell parameters obtained from these images range from 100 - 200 Å which would agree with one molecule of MelB in the asymmetric unit (considering that the MW of the MelB-detergent complex is ~ 100-120 KDa) and if one considers the following cell parameter/space group combinations: a=b=100 Å c=200 Å in P6 or a=b=200 Å c=100 Å in P622. This crystal was flash-frozen in liquid $N_2$ and tested at a synchrotron (beamline BM16, ESRF). Itdiffracted to about 16Å, which was not better than that obtained at home light source.

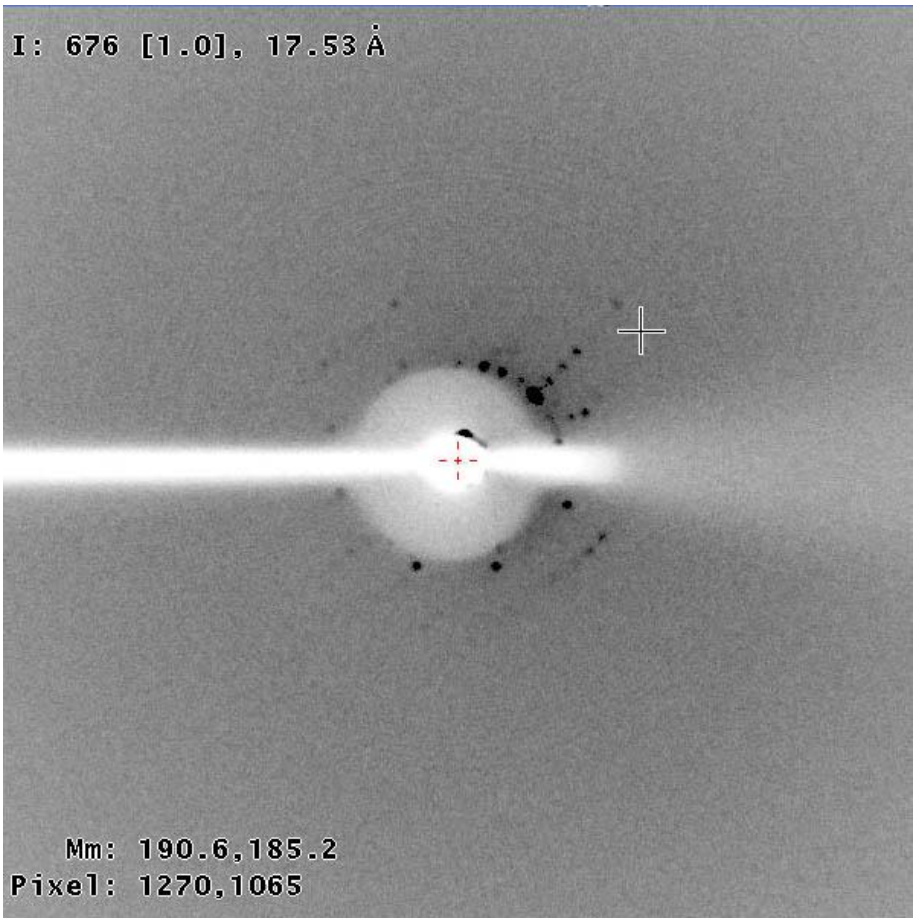



**Figure 6.9** X-ray diffraction pattern obtained using a Cu-Kα radiation from a rotating anode and from crystals similar to those shown in Figure 6.7. The crystals were flash frozen in liquid nitrogen in their own mother liquor (~40% PEG 400) which served as cryoprotectant and mounted in a Molecular Dimensions cryoloop/caps. This frame was obtained from a 1 º-1200 s X-ray exposure. The dimensions of this image are 345 mm x 345 mm.

**Table 6.1** Summary of the crystallization of some membrane transport proteins

| Organism/Protein | Function | Expression system | Vector | Protein modification | Induction | [prot.] (mg/ml) | Detergent Sol. | Detergent Cry. | Cleaved tag | Crystallization method | Protein:reservoir | [substrate] | T (℃) | Precipitant | Buffer | Salt | Additive(s) |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| *E. coli* LacY[1] | symporter $H^+$:lactose | *E.coli* XL1-Blue | | C154G | IPTG | 2 | 2% β-DDM | 0.01% β-DDM | No | hanging drop vapor diffusion | 1µl:1µl | 5 mM TDG | 20 | 27-30 % PEG400 | 100 mM HEPES pH 7.0 | 200 mM $CaCl_2$ | 3% 1.6-hexanediol 0.8 mM CHAPS |
| *E. coli* NhaA[2] | antiporter $H^+$:$Na^+$ | *E. coli* RK20 | | | IPTG | 4-6 | 1% β-DDM | 0.03% β-DDM | No | hanging drop vapor diffusion | 1µl:1µl | | 6 | 28-34 % PEG 400 | 25 mM Na citrate pH 4.0 | 200-450 mM $MgCl_2$, 100 mM KCl | 0.5 % ethanol 1% β-OG |
| *E. coli* GlpT[3] | antiporter $H^+$:glycerol-3-phosphate | LMG194 strain of E. coli | | Leu2 changed to Gly2 and sequence Arg449-Gly452 to Leu-Val-Pro-Arg | 0.1% arabinose | 6 | 1% β-DDM for 30 min at 4℃ | 0.1-0.5 % β-DDM 0.04-0.1 % $C_{12}E_9$ | thrombin digestion | hanging drop vapor diffusion | | | 15-20 | 25-27 % PEG 2000 MME | 100 mM Tris pH 8.5-8.9 | 5-100 mM NaCl | 5mM $SrCl_2$/$MgCl_2$ 25% MPD, 20% glycerol |
| *A. aeolicus* LeuT[4] | symporter $Na^+$:leucine | *E. coli* C41 | | | 0.1 mM IPTG for 20 h at 20℃ | 3-6 | 40 mM β-DDM | 40 mM β-OG | thrombin digestion | hanging drop vapor diffusion | | 50 mM tryptophan | 20 | 18 % PEG 550 MME | 100 mM HEPES pH 6.8 | 200 mM NaCl | |
| *E. coli* EmrD[5] | antiporter $H^+$:$Na^+$ | *E coli*. BL 21 (DE3) star | | | | 10-26 | 1% β-DDM at 4℃ | 0.1 % β-DDM | | sitting drop technique | | | 22 | 18-25 % PEG 400 | 50 mM Na citrate pH 3.8-5.2 | 25-100 mM KCl/NaCl | |
| *E. coli* FucP[6] | symporter $H^+$:L-fucose | E. coli BL21(DE3) | | | 0.2 mM IPTG at 37℃ | | 1.5% β-DDM for 1 h at 4℃ | 0.4 % β-NG | | hanging drop vapor diffusion | | | 18 | 30% PEG400 | 0.1 M MES pH 6.0 | 0.1 M $MgCl_2$ | 2 mM L-fucose, 50 mM NaF |
| *E. coli* vSGLT[7] | symporter $Na^+$:Galactose | *E.coli* XL1-Blue | | pVNH6A,411A,423C and p3C423 | 0.66mM L-Arabinose | ~16.5 | 2% β-DM | 0.174% β-DM | | hanging drop vapor diffusion | 1µl:1µl | 20 mM galactose | 4 | 20-25% PEG400 | 0.1M HEPES pH7.0 | 0.2 M $CaAc_2$ | 1x cmc Anzergent 3-12 |
| *M.liquefaciens* Mhp1[8] | symporter $Na^+$:Hydantoin | *E.coli* BLR | pSHP11 | | | 20-28 | 1% β-DDM at 4℃ | 0.7% β-NM | | | 1µl:1µl | | 18 | 31-33 % PEG 300 | 0.1 M Na-phosphate pH 7.0 | 100 mM NaCl | |
| *E. coli* AdiC9 | antiporter arginine:agmatine | *E. coli* BL21(DE3) | pET15b | | 0.2 mM IPTG at 37℃ for 4 h | 10 | 1.5% β-DDM for 3 h at 4℃ | 0.4% NG | thrombin digestion | hanging drop vapor diffusion | 1µl:1µl | | 18 or 4 | 22% PEG 400 | 0.1 mM Tris, pH 7.0 | | |
| *P. Horikoshii* GltPh[10] | ion:glutamate transporter | *E.coli* Top10 cells | pBAD24 | | 0.05 mM IPTG at 22℃ for 2 h | 7-10 | 40 mM DM | 10 mM DM | tobacco etch virus (TEV) protease | | | | 4 | 14–18% PEG 1000 | 50mM $Na_2HPO_4$, 50mM citric acid, pH4.2 | 0.1 M $Li_2SO_4$ | |
| *E. coli* YiiP[11] | $Zn^{2+}$/$H^+$ exchange | *E. coli* BL21(DE3) pLysS cells | pET15b | | auto-induced | | | | | hanging drop vapor diffusion | | 5 mM $ZnSO_4$ | 20 | 10% PEG400, 15-20% PEG2000 | 0.1 M mM Na-citrate (pH 6.0) | 100 mM NaCl, 200 mM $(NH_4)_2SO_4$ | 3 mM Fos-Choline-12, 4% (w/v) benzamidine, 10% (w/v) glycerol and 4% (v/v) 1,3-propanediol |
| *S. aureus* Sav1866[12] | Multidrug transporters | *E.coli* BL21-codon plus (DE3) RIPL | pET19b | | 0.4 mM IPTG for 2 h | 15 | 0.1% DDM+f% C12E8 for 1.5 h | $C_{12}E_8$ | | sitting drop technique | 2µl:1µl | 1 mM ADP | 10 | 20 % PEG 6000 | 50 mM $Li_3$ citrate, 150 mM $K_3$ citrate, 100 mM $Na_2HPO_4$ | 3 mM $MgCl_2$ | |
| *E.coli* $MalFGK_2$[13] | catalytic intermediate of the maltose transporter | | pFG26 | | | 10-15 | 0.3% DDM | 0.06% UDM | | sitting drop technique | 1µl:1µl | 0.5 mM ATP, 0.2 mM maltose | 20 | 27% PEG 400 | 0.1 M HEPES pH7.5 | 0.5 M NaCl | 10 mM betaine hydrochloride |
| *E.coli* FocA[14] | formate–nitrite transporter | *E. coli* BL21(DE3) | pET21b | Truncation (residues 22–285) | 0.2 mM IPTG for 16 h at 22 ℃ | - | 1.5% OG for 3 h at 4℃ | 0.8% β-OG,0.046% LDAO | | hanging drop vapor diffusion | | | 18 | 36% PEG400 | 0.1 M MOPS pH7.5 | 200 mM NaCl or sodium formate | 0.2% Cysmal-2 |
| *M.jannaschii* ApcT[15] | $Na^+$-Independent Amino Acid Transporter | *E. coli* C41 | GFP fusion vector pCGFP-BC and pNGFP-BC | | IPTG | 5-10 | DDM | 20 mM C8SG | | paraffin oil | 1:1 | | 20 | 39-45% PEG 550 MME | 100 mM Bicine (pH 9.0) | 100 mM NaCl | mercury acetate, PCMB, or thimerosal or holmium salts ($HoCl_3$) |
| *V.cholerae* NorM[16] | cation-bound multidrug and toxic compound extrusion transporter | *E. coli* BL21(DE3) | pET19b | | 0.4 g/l IPTG at 30 ℃ for 4 h | 6-12 | 1% DDM | 0.23% β-NG | enterokinase | | 2µl:2µl | | 22 | 16–24% PEG250 DME | 50 mM Tris-HCl pH7.2–8.6 | 87 mM $(NH_4)_2SO_4$ | 0.02% FA-231 |

Note: 1. (Abramson et al. 2003); 2. (Hunte et al. 2005); 3. (Huang et al. 2003); 4. (Yamashita et al. 2005); 5. (Yin et al. 2006); 6. (Dang et al. 2010); 7. (Watanabe et al. 2010); 8. (Shimamura et al. 2010); 9. (Gao et al. 2009); 10. (Yernool et al. 2004); 11. (Lu and Fu 2007); 12. (Dawson and Locher 2006); 13. (Khare et al. 2009); 14. (Wang et al. 2009); 15. (Shaffer et al. 2009); 16.(He et al. 2010).

Clearly, to gain structural information from the crystals showed in the previous section they need to be improved, that is, their resolution has to be increased from the current 17 Å to at least 4 or 3 Å. Although we expected to improve the crystal diffraction by using synchrotron light, especially when using the microfocus undulator

beamline, we were aware that this would need careful and possibly lengthy crystal optimization. These methodologies are explained in the following points.

4.1.4.3 Crystal optimization by modifying the crystallization conditions

To obtain high-quality crystals, screening of a large number of conditions was shown to be essential, specially for membrane proteins. Faham et al. reported that crystal optimization and structure determination of a Sodium Galactose Transporter required ~50,000 crystallization trials and 25 synchrotron trips where more than 2,500 crystals were screened and nearly 120 data sets collected. This shows how difficult it is to get high suitable crystals for structure determination (Faham et al. 2008). Avery important feature of membrane proteins is their lack of solubility in aqueous solution, unless with the help of detergents. However, detergents may interfere with crystal order. It is clear that detergents are the core of membrane protein crystal optimization process. Detergent solubilization of proteins entails the formation of protein:detergent complexes or mixed micelles, where easily 50% or more of the total mass of the complex belongs to the detergent. However, historically the interactions between the soluble parts are the first ones to be optimized during protein crystallization. This is a method that has been used for many years in crystallization of soluble proteins. However, such a high percentage of detergent in the protein:detergent complex must have an important effect during crystallization. In addition, the type of detergent is also important to get a stable protein for crystallization. As a matter of fact, it has been observed that many membrane proteins crystallize at detergent concentrations near its cloud point. At this detergent concentration, the detergent micelles optimally interact with each other without reaching phase separation. The cloud-point concentration of a specific detergent varies considerably depending on the temperature, pH, precipitant, and concentration of the salts/additives present in solution(Berger et al. 2006). All of these give us the clues for crystal optimization.

In order to improve the condition that produces the crystal form obtained in **PART III4.1.4.2** or other possible crystal forms that we may encounter during this work, the first logical and easiest step is to systematically vary in just a few percent all the

variables of the conditions, and setting up a multi-dimensional matrix of conditions in 24-well plates (preferably in hanging drop). Lists of several variables that will have to be modified are:

- ❖ Altering the precipitant concentration: changing both the concentration as well as the molecular weight of the PEG used;
- ❖ Changing the pH of the crystallization condition: before and after crystallization;
- ❖ Varying the drop ratio (protein solution vs. reservoir solution), that is instead of 1:1 µl:µl try 0.5:1, 1:1.5, 0.75:1, 1:0.75, etc.;
- ❖ Testing different temperatures;
- ❖ Modifying the protein concentration used for crystallization (variations of 20% above and below the current protein concentration);
- ❖ Extensive screening for small molecule additives;
- ❖ Trying to extract the protein from the membrane with different detergents;
- ❖ Trying to purify the protein with different detergents;

4.1.4.3.1 Effect of the precipitants

Figure 6.10 summarizes the different precipitants and the concentrations of the polyethylene glycols (PEGs) used in the successful crystallization of the membrane proteins. Unlike soluble proteins, small MW (molecular weight) PEGs, in particular PEG 400, have been more successful for membrane proteins (Figure 6.10A). Figure 6.10B shows that the optimum concentration range for low MW PEGs lies between 20% (w/v) and 30% (w/v), which is relatively narrow compared to the large MW PEGs. As other transporters, we obtained the MelB crystals from the condition using PEG200-400 as precipitant. To find a suitable precipitant for R149C MelB crystallization, we fixed the buffer and changed the MW and the concentration of PEG. We tested PEG200 at 30-60% (w/v), PEG300 at 20-50% (w/v), PEG400 at 10-40% (w/v), and PEG500 at 5-25% (w/v). Crystals appeared with PEG200 at 46-55% (w/v), PEG300 at 40-50% (w/v), and PEG400 at a narrow range, i.e., 29-35% (w/v) (Figure 6.11). No crystal were

obtained with the PEG of MW higher than 400. These results show that only small molecular weight PEGs are suitable for MelB

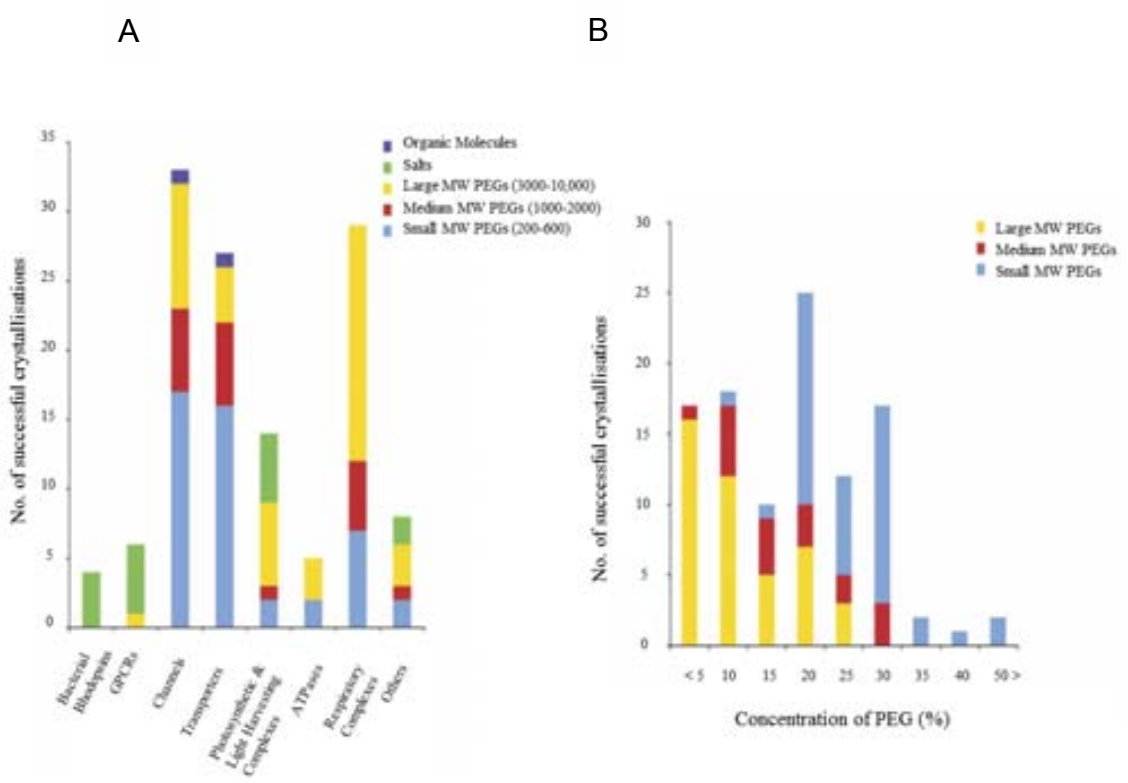



**Figure 6.10** (A) the different precipitants used in the successful crystallization of the membrane protein families are shown. (A) Precipitants. (B) The concentrations of the polyethylene glycols used for successful crystallization of the membrane protein families are shown. (According to Newstead et al. 2008)

crystallization. In contrast, we tested some traditional precipitants, e.g. $(NH_4)_2SO_4$, which were used for protein crystallization. Only a brown precipitation was obtained at any concentrations, which indicates that only low MW PEGs are appropriate for stabilization of the MelB-DDM complex. The best crystal diffracted to 11Å at ESRF BM16.

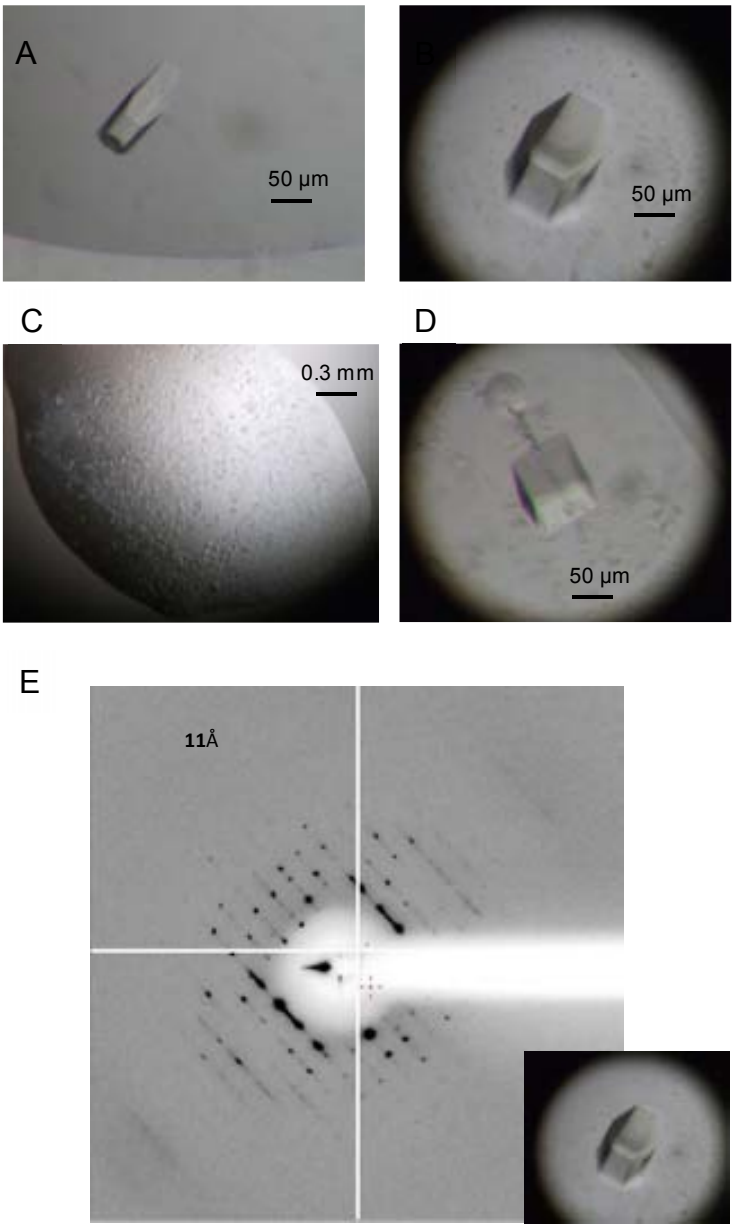



**Figure 6.11** Optimization in R149C MelB crystallization by changing precipitants. (A)-(D) photos of R149C MelB crystals obtained in the 50 mM phosphate/citrate pH4.2 with the addition of 51% (w/v) PEG200 (A), 46% (w/v) PEG200 (B), 29% PEG400 (C), and 42% (w/v) PEG300 (D) as precipitant. (E) X-ray diffraction pattern obtained at synchrotron ESRF BM16 beamline. Insert figure in E shows the photo of crystal which was used for x-ray diffraction analysis. The crystal belongs to space group P622, with unit cell dimensions of a =b = 212 Å, c = 131 Å, and α = β= 90°, γ = 120°. The best R149C MelB crystals diffracted to a resolution of 11Å.

4.1.4.3.2 Effect of the pH in the reservoir solution

pH is an effective variable in crystals screening experiments because most proteins demonstrate pH dependent solubility minima and will solubilize, precipitate, or crystallize at particular pH values. Fine screening of pH as a single variable has been successfully used for optimization(McPherson et al. 1995).To test the effect of the pH in MelB crystallization, we fixed the precipitant, i.e., 48% (w/v) PEG200, and then varied the pH by 0.1 pH unit from 3.5 to 9.0. Table 6.2 lists the entire buffers used in this study. Similar rod-like crystals were obtained with sodium citrate buffer at pH 4.0 to 4.5. This buffer change neither clearly improves the size of the crystals nor improves the resolution of the crystals. Only needle-like crystals were obtained with MES buffer at pH 6.5. Some bigger crystals were only diffracting to 18-20 Å at synchrotron (ESRF BM16 beamline).The study of buffer screening suggested that phosphate/citrate pH 4.2 may be the best buffer for MelB crystallization.

4.1.4.3.3 The effect of the concentration of the protein

Initially, we obtained R149C MelB crystals at the protein concentration of 9.2 mg/ml. Then we test series of the concentration of the protein from 4 mg/ml to 11 mg/ml by fixing the buffer (50 mM phosphate/citrate pH 4.2) and the precipitant (48% (w/v) PEG200) in hanging drops at 18 ℃ by equilibrating a 1:1 mixture of protein and reservoir solutions against the reservoir. It seems that R149C MelB prefers to crystallize at a concentration from 4.5 to 9 mg/ml. Another easy way to test the effect of the concentration of the protein is by varying the drop ratio (protein solution vs. reservoir solution), that is, instead of 1:1 μl:μl, trying 0.5:1, 1:1.5, 0.75:1, 1:0.75, etc. The result showed that using the ratio of 0.7μl protein to 1.4μl reservoir solution would facilitate the R149C MelB crystallization. At this ratio, crystals will appear in 3-7 days. In contrast, at the ratio of 1:1, crystals will need more time to appear, typically two weeks. However, with the ratio of 1:1, it would be much easy to obtain large crystals with fewer crystals in one drop. Although with the 1:2 ratio it is much easy to obtain crystals, they are small and there are many crystals in one drop.

4.1.4.3.4 The effect of the substrates

Table 6.2 Conditions for buffer screening

| No. | buffer | pH | No. | buffer | pH |
|---|---|---|---|---|---|
| 1 | 0.1M NaPi | 6 | 49 | 0.1M Tris | 7.7 |
| 2 | 0.1M NaPi | 6.1 | 50 | 0.1M Tris | 7.8 |
| 3 | 0.1M NaPi | 6.2 | 51 | 0.1M Tris | 7.9 |
| 4 | 0.1M NaPi | 6.3 | 52 | 0.1M Tris | 8 |
| 5 | 0.1M NaPi | 6.4 | 53 | 0.1M Tris | 8.1 |
| 6 | 0.1M NaPi | 6.5 | 54 | 0.1M Tris | 8.2 |
| 7 | 0.1M NaPi | 6.6 | 55 | 0.1M Tris | 8.3 |
| 8 | 0.1M NaPi | 6.7 | 56 | 0.1M Tris | 8.4 |
| 9 | 0.1M NaPi | 6.8 | 57 | 0.1M Tris | 8.5 |
| 10 | 0.1M NaPi | 6.9 | 58 | 0.1M Tris | 8.6 |
| 11 | 0.1M NaPi | 7 | 59 | 0.1M HEPES | 6.9 |
| 12 | 0.1M NaPi | 7.1 | 60 | 0.1M HEPES | 7 |
| 13 | 0.1M NaPi | 7.2 | 61 | 0.1M HEPES | 7.1 |
| 14 | 0.1M NaPi | 7.3 | 62 | 0.1M HEPES | 7.2 |
| 15 | 0.1M NaPi | 7.4 | 63 | 0.1M HEPES | 7.3 |
| 16 | 0.1M KPi | 6 | 64 | 0.1M HEPES | 7.4 |
| 17 | 0.1M KPi | 6.2 | 65 | 0.1M HEPES | 7.5 |
| 18 | 0.1M KPi | 6.4 | 66 | 0.1M HEPES | 7.6 |
| 19 | 0.1M KPi | 6.6 | 67 | 0.1M HEPES | 7.7 |
| 20 | 0.1M KPi | 6.8 | 68 | 0.1M HEPES | 7.8 |
| 21 | 0.1M KPi | 7 | 69 | 0.1M tri-Na Citrate | 3.8 |
| 22 | 0.1M KPi | 7.2 | 70 | 0.1M tri-Na Citrate | 3.9 |
| 23 | 0.1M KPi | 7.4 | 71 | 0.1M tri-Na Citrate | 4 |
| 24 | 0.1M MES | 5.9 | 72 | 0.1M tri-Na Citrate | 4.1 |
| 25 | 0.1M MES | 6 | 73 | 0.1M tri-Na Citrate | 4.2 |
| 26 | 0.1M MES | 6.1 | 74 | 0.1M tri-Na Citrate | 4.3 |
| 27 | 0.1M MES | 6.2 | 75 | 0.1M tri-Na Citrate | 4.4 |
| 28 | 0.1M MES | 6.3 | 76 | 0.1M tri-Na Citrate | 4.5 |
| 29 | 0.1M MES | 6.4 | 77 | 0.1M tri-Na Citrate | 4.6 |

| | | | | | |
|---|---|---|---|---|---|
| 30 | 0.1M MES | 6.5 | 78 | 0.1M tri-Na Citrate | 4.7 |
| 31 | 0.1M MES | 6.6 | 79 | 0.1M tri-Na Citrate | 4.8 |
| 32 | 0.1M Na Acetate | 3.6 | 80 | 0.1M tri-Na Citrate | 4.9 |
| 33 | 0.1M Na Acetate | 3.8 | 81 | 0.1M tri-Na Citrate | 5 |
| 34 | 0.1M Na Acetate | 4 | 82 | 0.1M tri-Na Citrate | 5.1 |
| 35 | 0.1M Na Acetate | 4.2 | 83 | 0.1M tri-Na Citrate | 5.2 |
| 36 | 0.1M Na Acetate | 4.4 | 84 | 0.1M tri-Na Citrate | 5.3 |
| 37 | 0.1M Na Acetate | 4.6 | 85 | 0.1M tri-Na Citrate | 5.4 |
| 38 | 0.1M Na Acetate | 4.8 | 86 | 0.1M tri-Na Citrate | 5.5 |
| 39 | 0.1M Na Acetate | 5 | 87 | 0.1M tri-Na Citrate | 5.6 |
| 40 | 0.1M Na Acetate | 5.2 | 88 | 0.1M tri-Na Citrate | 5.7 |
| 41 | 0.1M Na Acetate | 5.4 | 89 | 0.1M tri-Na Citrate | 5.8 |
| 42 | 0.1M Na Acetate | 5.6 | 90 | 0.1M tri-Na Citrate | 5.9 |
| 43 | 0.1M Tris | 7.1 | 91 | 0.1M tri-Na Citrate | 6 |
| 44 | 0.1M Tris | 7.2 | 92 | 0.1M Na Cacodylic acid | 6.1 |
| 45 | 0.1M Tris | 7.3 | 93 | 0.1M Na Cacodylic acid | 6.2 |
| 46 | 0.1M Tris | 7.4 | 94 | 0.1M Na Cacodylic acid | 6.3 |
| 47 | 0.1M Tris | 7.5 | 95 | 0.1M Na Cacodylic acid | 6.4 |
| 48 | 0.1M Tris | 7.6 | 96 | 0.1M Na Cacodylic acid | 6.5 |

It is clear that to get high quality crystals, it is essential to work with a stable protein. Membrane transporters are characterized by their high dynamic conformation. The binding of substrates may help to stabilize protein. However, it has been reported that high substrates concentrations in the drop lead to a decrease of the resolution (Kowalczyk et al. 2011). For MelB, there are two substrates, i.e., $Na^+$ or $Li^+$ and melibiose. $Na^+$ is one of the substrates, but it also plays an important role as salt in crystallization. From the reported structures of sodium-coupled transporters, a NaCl concentration in the range of 50 to 100 mM was used. And the organic substrates, e.g., sugar, amino acid, nucleic acid, etc. were used in a lower concentration, e.g., in the 2-5 mM range, always close to their binding constants. To test the effect of the substrates, we tried to change their concentration in the protein solution or added them to the

reservoir solution as additives. To change the concentration of substrates in the protein solution, we modified the dialysis buffer by changing the concentration of $Na^+$ or melibiose. We test different concentration of NaCl from 20 mM to 150 mM and melibiose from 2 mM to 20 mM. The results show that crystals appear in all of these conditions. However, with higher concentration of the substrates it is easy to obtain small, but many crystals in a drop. With lower concentrations of the substrates it is much difficult to obtain crystals under the same condition. So, finally we choose 100 mM NaCl and 10 mM melibiose for R149C MelB crystallization.

4.1.4.3.5The effect of temperature

From the already reported structures of this type transporter, the major part of the proteins was successfully crystallized at 18-20 ℃ or 4 ℃. In this study, we tested two different temperatures, i.e., 18 ℃ and 4 ℃ under the same conditions. The results show that R149C MelB prefers to crystallize at 18-20 ℃. We also obtained crystals at 4 ℃ (Figure 6.12). However, crystals grew slower than at higher temperature (more than one month) and were always small. When we diffracted these crystals the best resolution was about 17 Å at synchrotronline (ESRF ID 23-1 beamline). These results suggest that low temperature may not be suitable for the crystallization of R149C MelB. Then we chose 18 ℃ for further crystal optimization.


**Figure 6.12** Crystals growing in 50 mM phosphate/ citrate, pH4.2, and 26% PEG400 at 4 ºCby the hanging-drop vapour-diffusion method.

4.1.4.3.6 Effect of phospholipids (PL)

Phospholipids can also have a positive impact in the crystallization of membrane proteins. Of the transporters that are most similar to MelB and described in Table 6.1, two of them: *E. coli* GlpT and LacY require phospholipids for its crystallization, or their presence during crystallization increased significantly the quality of the obtained crystals (Lemieux et al. 2003; Guan et al. 2006). During the purification of GlpT, it was observed that endogenous phospholipids co-purified with the protein in molar ratios that ranged from 1:20 to 1:40 (protein:phospholipid) (Lemieux et al. 2003). Guan and her colleagues reported three different crystal forms that diffract to increasingly better resolution in a manner that correlates with the concentration of copurified phospholipids, i.e., mol PL/ mol LacY <8→5Å; mol PL/ mol LacY 18-25→3Å; mol PL/ mol LacY 9-16→2.6Å (Guan et al. 2006). In the case of LacY, and similar to GlpT, the addition of *E. coli* phospholipids to purified protein was a requirement to obtain crystals of sufficient quality of the wild-type protein (Mirza et al. 2006).

To test the effect of the phospholipids in MelB crystallization, we used the same method employed in the cited work. To this end, we tried to directly add *E.coli* phospholipids (Avanti Polar Lipids, Inc.) to the reservoir solution as additive, or tried to purify the protein with the addition of *E.coli* phospholipids. We tested the effect of phospholipids with the protein in molar ratios that ranged from 1:5 to 1:40 (protein:phospholipid). It seems that the addition of phospholipids in the reservoir solution may not help to facilitate MelB crystallization. The problem is that the phospholipids cannot dissolve in water, but in detergent containing buffer. In general, phospholipids were dissolved in 0.5% (w/v) DDM. However, in the reservoir solution, there is not detergent or just 0.017% (w/v) (double to CMC). When we added phospholipids to the reservoir solution, almost all lipids precipitated. Then we went to purify protein with the addition of phospholipids to the washing buffer and eluting buffer. Our results showed that with the help of lipids, the crystals were grown within one month to optimal size ($800 \times 60 \times 60 \mu m^3$). The best crystal diffracted to about 8 Å at synchrotron ESRF ID 23-1 beamline (Figure 6.13A) and to about 9 Å at synchrotron ESRF ID 14-4 beamline (Figure 6.13B).

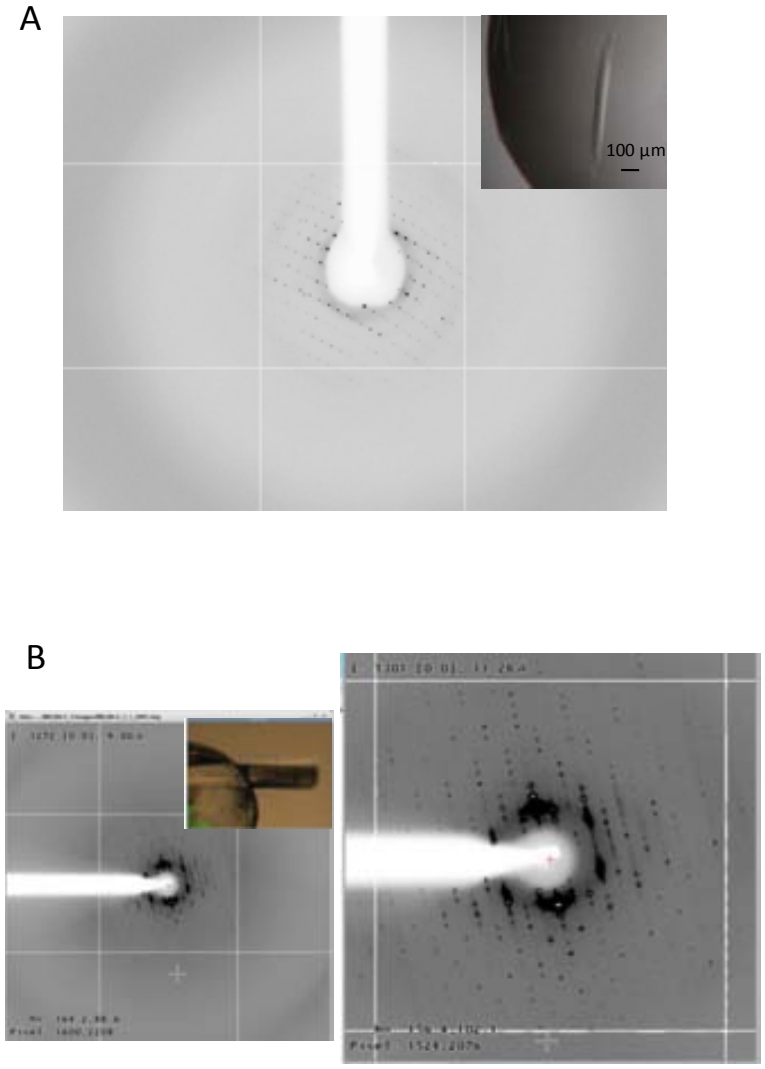



**Figure 6.13** X-ray diffraction pattern obtained at synchrotron ESRF ID 23-1 beamline (A) and at synchrotron ESRF ID 14-4 beamline (B). The crystal were flash frozen in liquid nitrogen in its own mother liquor (~48% PEG 200) which served as cryoprotectant and mounted in a Molecular Dimensions cryoloop/caps. Insert figure in A shows the photo of crystal using for X-ray diffraction study. (B, right) shows enlarged image of (B, left) to show detail diffraction spots. Insert figure in B shows the photo of crystal which was mounted on the cryoloop.

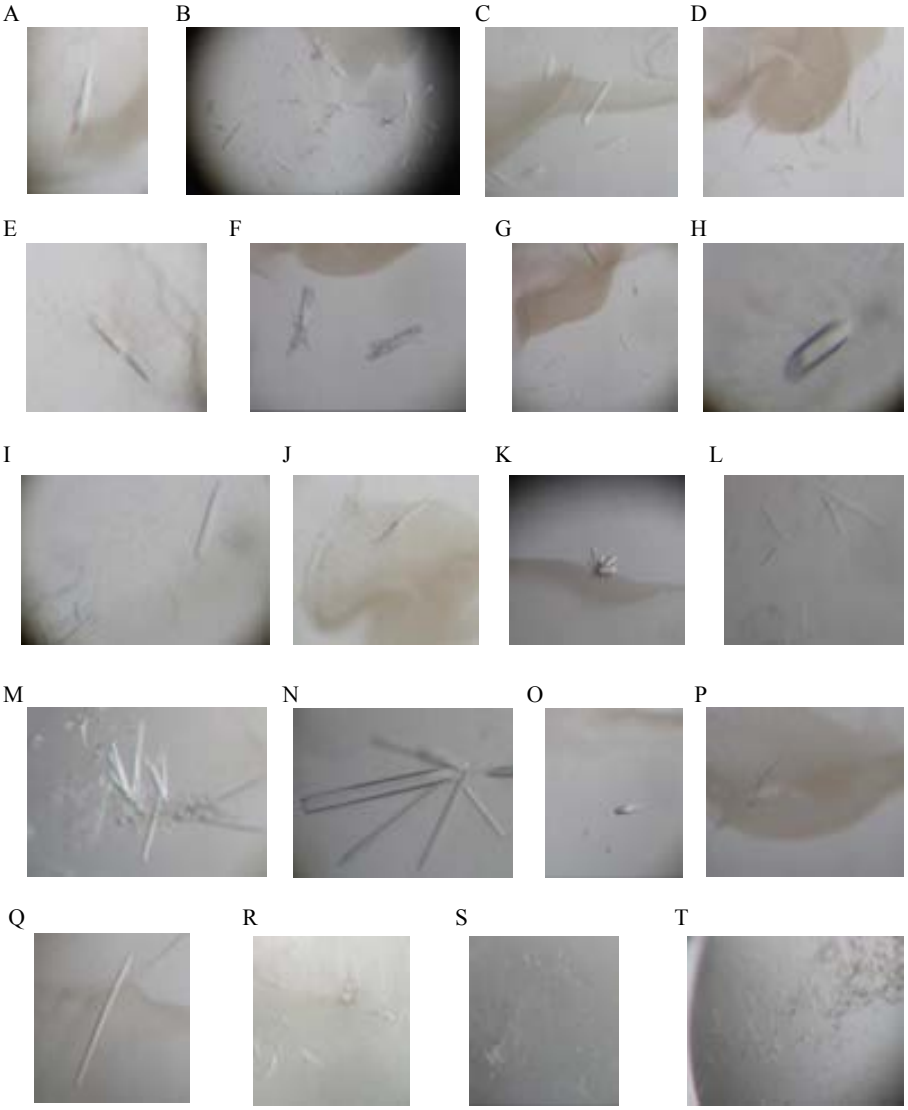



**Figure 6.14** photos of the crystals obtained from additive screening. R149C MelB crystals were obtained under the condition of 48% (w/v) PEG200, 50 mM phosphate/citrate pH 4.2 with the addition of different additives: (A) 10 mM 0.1 M barium chloride, (B) 10 mM 0.1 M cadmium chloride, (C) 10 mM cobalt(II) chloride, (D) 10mM zinc chloride, (E) 100 mM potassium chloride, (F) 100 mM lithium chloride, (G) 10 mM L-proline, (H) 10 mM phenol, (I) 3% (v/v) dimethyl sulfoxide, (J) 10 mM sodium bromide, (K) 30 mMglycyl-glycyl-glycine, (L) 10 mMsperminetetrahydrochloride, (M) 10 mMhexammine cobalt(III) chloride, (N) 10 mM guanidine hydrochloride, (O) 10 mM urea, (P) 10 mM b-Nicotinamide adenine dinucleotide hydrate, (Q) 10 mMEthylenediaminetetraacetic acid disodium salt dihydrate, (R) 0.5% (w/v) Polyvinylpyrrolidone K15,(S) 3.0% (w/v) Dextran sulfate sodium salt (Mr 5,000), (T) (v/v) Pentaerythritolethoxylate (3/4 EO/OH).

4.1.4.3.7Additive screening

Previous studies showed that small additives may help to improve diffraction limit. For example, glycerol may stop nucleation and may give fewer, larger crystals, and has the advantage of acting as a cryo-protectant. Ethanol or dioxane has the effect of poisoning the crystals and stopping too much nucleation. To conserve sample, we performed the additive screen at the nanoliter scale using a Phoenix RE robot in 96-well format from Hampton Research (HR2-138). Once a suitable additive is identified, a focused grid screen on the microliter scale is performed to find the optimal well condition in the presence of the additive. Figure 6.14 shows the pictures of the R149C MelB crystals, obtained under the condition of PEG200, 50 mM phosphate/citrate pH 4.2 with the presence of different additives from the additive screening in sitting drop at the ratio of 200 nl protein solution to 200 nl reservoir solution at 18 ℃. Then we performed grid screen on the microliter scale to find the optimal well condition in the presence of the additive.

4.1.4.3.8Detergent screening

Since the detergents are responsible for the solubilization of membrane proteins, they have a paramount role in establishing or hampering crystal contacts. Previous studies have shown that finding a suitable detergent is essential for getting high quality crystals and good resolution. In some instance, it may be necessary to screen a second or a third detergent as additive. For example, Wang Y. and her colleagues reported that a third detergent, as an additive, was essential for improving the diffracting resolution of FocA crystal from 3.5 Å to 2.2 Å (Wang et al. 2009). In many cases, performing extensive crystallization trials for the proteins purified in more than 20 different detergents is the essential process for getting better resolution crystals. Therefore, modifying the detergent in which MelB has been solubilized is a crucial point to consider when trying to improve the crystals or obtain better-diffracting crystal forms. To this end, we tried to purify R149C MelB in different detergents and performed a detergent screen for R149C MelB on the nanoliter scale using a Phoenix RE robot in the 96-well format from Hampton Research (HR2-406).

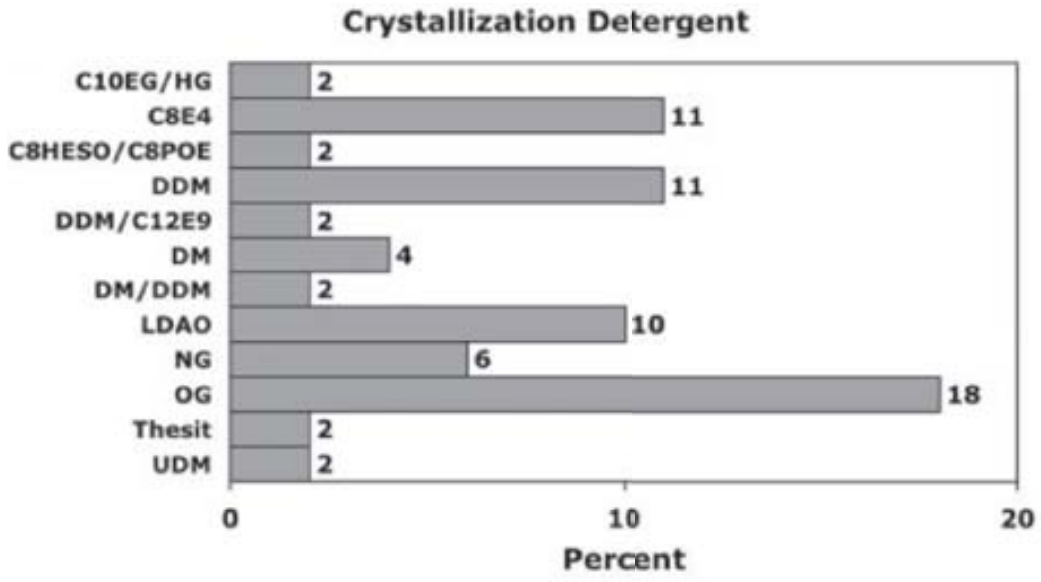



**Figure 6.15** Detergents used in the successful crystallization of membrane proteins. (According to Willis and Koth 2008).

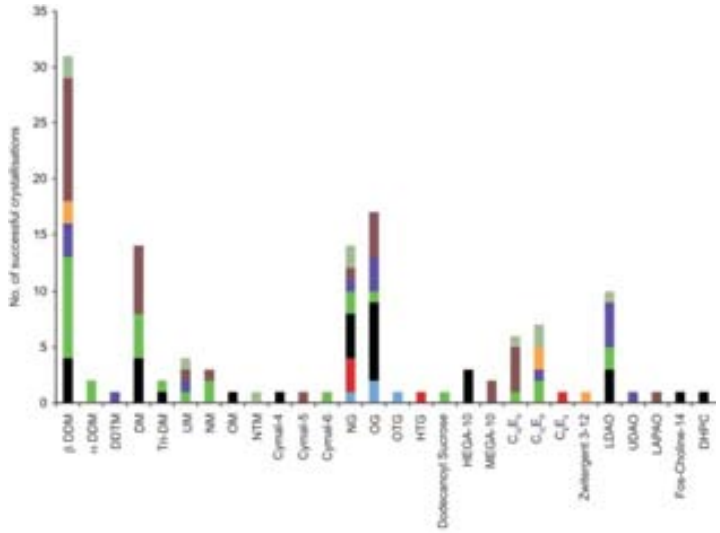


**Figure 6.16** The different detergents that have been used to successfully crystallize a-helical membrane proteins. Color code: Bacterial rhodopsins (blue), GPCR (red), channels (black), transporters (green), photosynthetic and light harvesting complexes (purple), ATPases (orange), respiratory complexes (brown), others (DsbB-DsbA oxidase, intramembrane proteases, membrane-associated proteins in eicosanoid, and gluththione metabolism [MAPEG]) (olive). (According to Newstead et al. 2008)

In order to have the highest probability of success, one has to look at the list of detergents that have been the most successful in crystallizing membrane proteins. The current literature has extensive data on this, and the statistics show that OG, DDM, $C_8E_4$,

LDAO, and NG (Figure 6.15) are the most commonly used detergents to solubilize and crystallize membrane proteins (Willis and Koth 2008). In another study, focused on α-helical membrane proteins (Newstead et al. 2008), it was determined that DDM is the detergent that has given the best results in crystallization trials, followed by OG, DM, NG, LDAO, $C_{12}E_8$, $C_{12}E_9$, etc. (Figure 6.16). The rank of the detergents used in the crystallization of α-helical membrane transport proteins is DDM, DM, α-DDM, NG, NM, $C_{12}E_8$, LDAO (Table 1).

**Table 6.3** the results of the protein preparation in different detergents

| Detergents | [detergent] | [protein] | status | Crystal |
|---|---|---|---|---|
| DM | 0.17% (w/v) | 8.4 mg/ml | stable | No |
| LDAO | 0.046% (w/v) | 6.7 mg/ml | stable | No |
| Cymal-6 | 0.056% (w/v) | 5.0 mg/ml | stable | NO |
| NM | 0.56% (w/v) | 7.1 mg/ml | stable | NO |
| UDM | 0.06% (w/v) | 6.7 mg/ml | stable | Yes |
| DDM | 0.017% (w/v) | 8.5 mg/ml | stable | Yes |
| OG | 1.06% (w/v) | ND | precipitation | ND |
| NG | 0.4% (w/v) | 6.5 mg/ml | unstable | ND |
| $C_8E_4$ | 0.5% (w/v) | 7.3 mg/ml | stable | NO |

According to the results of statistical analysis previously reported, we performed protein purification with several common detergents, e.g., DDM, UDM, DM, NM, OM, OG, NG, Cymal-6, Cymal-5, LDAO, $C_{12}E_8$, $C_{12}E_9$, $C_8E_4$. For purification, membrane fractions containing R149C MelB were collected and incubated with 1.0% (w/v) dodecyl-β-D-maltopyranoside (DDM, Anatrace) for 3 h at 4 ºC, then subjected to buffer exchange with a buffer containing the detergent to be tested. Table 6.3 shows the results of the protein preparation in different detergents. Our results show that R149C MelB is not stable in the buffer containing glucopyranoside-type detergent, e.g., OG, NG. Almost all the protein precipitated during buffer exchange. In contrast, R149C MelB is quite stable in the buffer containing maltoside-type detergent, e.g., DDM, UDM, DM, Cymal-5 or Cymal-6. These results may suggest that a two-pyran structure is essential for the stabilization of the R149C-detergent complex. Then we performed detergent

**Table 6.4** lists of 24 common salts, small organic molecules, and detergents

| No. | Conditions |
|---|---|
| 1 | 0.029% (w/v)UDM, 48% (w/v) PEG200, 50 mM phosphate/citrate pH4.2 |
| 2 | 0.087% (w/v)DM, 48% (w/v) PEG200, 50 mM phosphate/citrate pH4.2 |
| 3 | 0.1M LiCl, 48% (w/v) PEG200, 50 mM phosphate/citrate pH4.2 |
| 4 | 3% (v/v) (+/-)-2-Methyl-2,4-pentanediol, 48% (w/v) PEG200, 50 mM phosphate/citrate pH4.2 |
| 5 | 3% (v/v) 1,6-hexanediex, 48% (w/v) PEG200, 50 mM phosphate/citrate pH4.2 |
| 6 | 0.023% (w/v) LDAO, 48% (w/v) PEG200, 50 mM phosphate/citrate pH4.2 |
| 7 | 0.53% (w/v)OG, 48% (w/v) PEG200, 50 mM phosphate/citrate pH4.2 |
| 8 | 0.1M Guanidinium chloride, 48% (w/v) PEG200, 50 mM phosphate/citrate pH4.2 |
| 9 | 5% (w/v) PEG400, 48% (w/v) PEG200, 50 mM phosphate/citrate pH4.2 |
| 10 | 0.0085% (w/v)α-DDM, 48% (w/v) PEG200, 50 mM phosphate/citrate pH4.2 |
| 11 | 0.25% (w/v) $C_8E_4$, 48% (w/v) PEG200, 50 mM phosphate/citrate pH4.2 |
| 12 | 3% (v/v) 2-propanol, 48% (w/v) PEG200, 50 mM phosphate/citrate pH4.2 |
| 13 | 0.35% (w/v) Fos-choline 10, 48% (w/v) PEG200, 50 mM phosphate/citrate pH4.2 |
| 14 | 0.028% (w/v) Cymal-6, 48% (w/v) PEG200, 50 mM phosphate/citrate pH4.2 |
| 15 | 10mM Phenol, 48% (w/v) PEG200, 50 mM phosphate/citrate pH4.2 |
| 16 | 0.28% (w/v)NM, 48% (w/v) PEG200, 50 mM phosphate/citrate pH4.2 |
| 17 | 0.2% (w/v)NG, 48% (w/v) PEG200, 50 mM phosphate/citrate pH4.2 |
| 18 | 0.89% (w/v)β-OM, 48% (w/v) PEG200, 50 mM phosphate/citrate pH4.2 |
| 19 | 0.12% (w/v) Cymal-5, 48% (w/v) PEG200, 50 mM phosphate/citrate pH4.2 |
| 20 | 0.003% (w/v) $C_{12}E_9$, 48% (w/v) PEG200, 50 mM phosphate/citrate pH4.2 |
| 21 | 0.0048% (w/v) $C_{12}E_8$, 48% (w/v) PEG200, 50 mM phosphate/citrate pH4.2 |
| 22 | 0.5% (v/v) ethanol, 48% (w/v) PEG200, 50 mM phosphate/citrate pH4.2 |
| 23 | 10 mM argentine, 48% (w/v) PEG200, 50 mM phosphate/citrate pH4.2 |
| 24 | 0.0085% (w/v) 2,2-didecylpropane-1,3-bis-β-D-maltopyranoside, 48% (w/v) PEG200, 50 mM phosphate/citrate pH4.2 |

screen using the robot with the kits purchased from Hampton Research, at 18 ℃. Detergents were directly added to the well solution, which containing 48% (w/v) PEG 200 and 50 mM phosphate/citrate pH 4.2. To make the drops, 0.7μlR149C MelB, purified in 0.017% (w/v) DDM were mixed with 0.7μl reservoir solution.

To avoid detergent co-concentration during sample preparation, especially for those with higher micellar average molecular weight, e.g., $C_{12}E_8$ (66KDa), $C_{12}E_9$ (83KDa), we followed a method, described in **PART III 4.1.1** to control detergent concentration in the protein sample. For those with lower micellar average molecular weight, e.g., NM (11.7KDa), DM (33.3 KDa), LDAO (17 KDa), etc., the protein samples were directly concentrated to the expected value using 50 KDa molecular weight cutoff, and then subjected to crystallization studies.

Based on statistical analysis of previously reported membrane protein crystallization conditions, and our experimental results, we developed a simplified grid screen method, which is appropriate to rapid detection of newly purified MelB or other mutants to see if they are suitable for crystallization. Our studies showed that 2 L *E. coli* cultures will yield about 4-5 mg R149C MelB, which is quite similar to that of MelB Cys-less or to wt. Clearly, this amount of the protein should be enough for crystal screening. Then the membrane protein was extracted by a standard detergent, i.e., DDM (1.5 % (w/v) for 1h), followed by purification using a Ni-NTA column. We compared the purity of the protein purified just by Ni-NTA column or size exclusion, and no clear difference was found. Actually, by just passing through the Ni-NTA column, the sample is pure enough for the crystallization study. We compared the crystals obtained from the samples passed through two columns (Ni-NTA + size exclusion) or just one column (Ni-NTA), and the resolution was not essentially different. So, here we simplified the protein purification steps. Then we went to detergent exchange by using centrifugal concentrators with 50 kDa or buffer exchanged by the batch affinity method by using a small volume of Ni-NTA resin, measured the concentration of the sample and then subjected them to crystal screen. Here we just did a small grid screen, i.e., PEG200 45%-56% (w/v) (step 1%) and PEG400 26%-37% (w/v), 50 mM phosphate/citrate pH4.2. These conditions

were summarized from the experiments of the MelB crystallization. PEG200 and PEG400 were also quite normally used as precipitants in other membrane transporters of known structures. For R149C MelB, crystals appeared in 3-7 days when we did small grid screen for the new purified protein. When the crystals appeared, we then went to do additive screen. Table 6.4 lists 24 common salts, small organic molecules, and detergents, which are normally used as an additive for improving crystal quality and resolution. Figure 6.17 shows the results corresponding to the conditions listed in Table 6.4. In general, following this protocol, we obtained crystals which can be used for X-ray diffraction studies. It should be noted that, as there are many parameters, repeatability of the membrane protein crystallization is relatively poor. Therefore, repeating the experiments several times is essential for obtaining high quantity crystals.

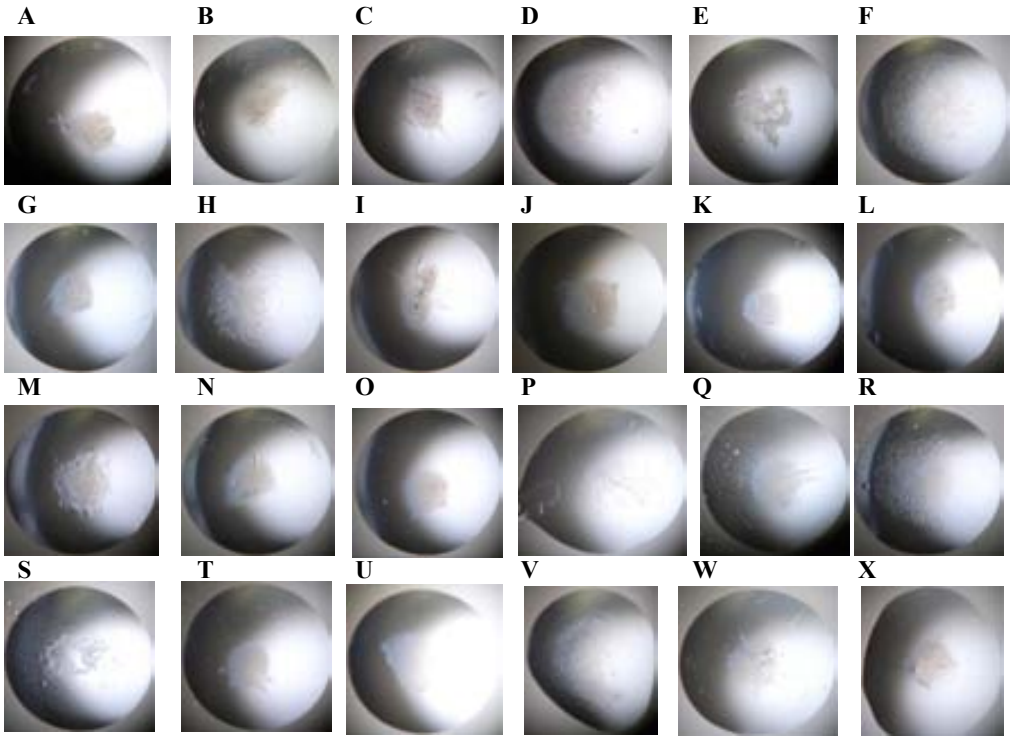



**Figure 6.17** Results of crystal screening using the conditions listed in Table 6.4. Experients were performed in the conditions of 48% PEG200, 50 mM phosphate/citrate pH4.2 with the addition of salts, small organic molecules, or detergents as additives. (A)-(X) correspond to the condition (1)-(24) listed in Table 6.4, respectively.

4.1.4.3.9 Crystal optimization by expression constructs reengineering

The MelB R149C construct may have to be modified to obtain crystals of sufficient quality:

i) Modifications of the His-tag: the His-tag used to purify the protein is usually disordered in the majority of protein crystal structures. This is due to the fact that this sequence tail is very flexible. Although the His-tag end should not affect the crystallization process, the absence of crystals or poor diffraction may often be related to the presence of this tail. These tags can be eliminated by using controlled protease digestion (as used to obtain ΔCTMelB R149C in point 2) or modifying the constructs by adding protease cleavage sites between the protein and His-tag. The most commonly used cleavage sites belong to thrombin or TEV (Tobacco Etch Virus) protease. These cleavage sites may be introduced by using pET vectors and molecular biology techniques. In this way, they can be eliminated after they have been used for the affinity chromatography step during protein purification. However, without the His-tag, we met one problem, that is, how to control the concentration of detergent? It is clear that the concentration of DDM in the protein solution is essential for getting R149C MelB crystals. We had tried to concentrate the R149C MelB using centrifugal concentrators with 100 kDa cutoff. But none protein crystal was obtained from these samples. We also tried to use a MINI Dialysis Device (20K MWCO, Thermo SCIENTIFIC) to dialyze overnight against the crystallization buffer to remove the excess of detergents, which is due to co-concentration during sample preparation. No crystal was obtained from these samples.

ii) Trimming or fixing the protein in a specific conformation: another strategy that can be employed consists in finding regions in the sequence that are thought to be disordered, and eliminate them by designing a new expression construct or by introducing mutations that successfully fix the transporter in a specific conformation. For example, the crystallized Glycerol-3-Phosphate Transporter (GlpT) construct contains 451 amino acids (aa. 2-452), with Leu2 changed to Gly2 and the

sequence Arg449-Gly452 to Leu-Val-Pro-Arg (Huang et al. 2003). This modification is essential for improving the diffracting resolution (Lemieux et al. 2003).

*Limited proteolysis*: as it was suggested in the case of the His-tag, the presence of regions at the N- or C-termini of the protein that are too flexible may hinder the formation of crystals. To identify these regions, limited proteolysis assays can be carried out where small quantities of the protein are treated at 37 ℃ with a gradient of protease concentration (usually from 1:1000 to 1:100 molar ratio protease:protein). The most used proteases are trypsin, chymotrypsin, subtilisin, elastase, or thermolysin. The results are evaluated with SDS/PAGE gels and the bands that may correspond to protein segments are identified by mass spectroscopy. Identification of these segments will indicate what region of the protein will have to be used to design new constructs.

Gwizdek and colleagues reported the proteolytic mapping of the *Escherichia coli* melibiose permease (Gwizdek et al. 1997), which provides us with important information. There are only two trypsin-sensitive sites, one located at cytoplasmic loop 4-5, the other one at C-termini tail (Arg442). It may be possible to cleave the polypeptide at position Arg-442, thus removing the last C-terminal 17 residues of MelB and the engineered His-tag. The detailed experimental methods will be discussed later.

4.1.4.3.10Seeding

Seeding has been shown to be a useful method to obtain better crystals (Bergfors 2003).In this way, an already existing crystal which was added to a new crystallized drop as a seed provides a template on which further molecules can assemble. Seeding allows controlling the nucleation and growing crystals of the same lattice and symmetry because it will inherit many of the characteristics of the seed from which it is originated. At given proper environment, the seed will enlarge into a crystal. We had tried some commonly used seeding method, e.g., microseeding, macroseeding, streaking seeding. Unfortunately, no crystal was obtained from any microseeding or streaking seeding drops. When we tried macroseeding, it seemed that seeded crystals were stable in the new drops, but no new or better crystals was obtained from them.

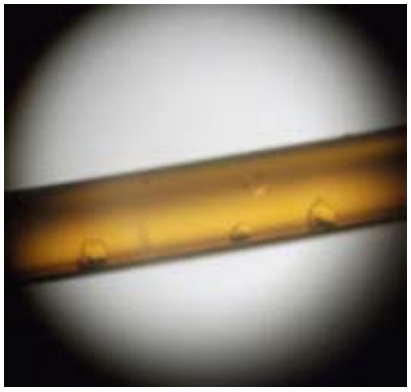
4.1.4.3.11 Crystal optimization by less standard crystallization techniques

24-well and 96-well plates are commonly used to crystallize proteins by the vapor diffusion method. Sometimes crystal diffraction improvement is only obtained by employing radically different crystallization methods, a few examples are listed below.

- Growth in a capillary by the counter-diffusion method (Garcia-Ruiz et al. 2002): Granada crystallization box (Triana Science & Technology) or CrystalHarp™ (Molecular Dimensions).
- Batch method under oil (Rayment 2002).
- Lipidic Cubic Phase (Willis and Koth 2008): Lipidic Cubic Phase Kit (Molecular Dimensions).

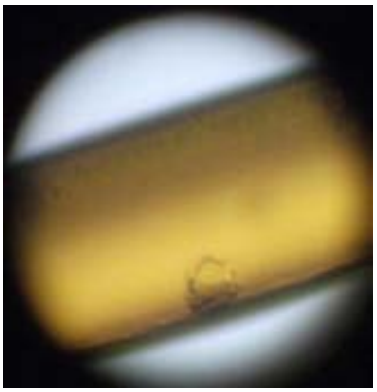
A

B



23#

C



33#

**Figure 6.18** Crystal optimization using CrystalHarp. (A)CrystalHarp$^{TM}$ plate (Molecular Dimensions). (B) Photo of the crystals obtained from the condition of 40% (w/v) PEG400, 0.1 M phosphate/citrate pH4.2, 0.73% (w/v)NG. (B) Photo of the crystals obtained from the condition of 40% (w/v) PEG400, 0.1 M phosphate/citrate pH4.2, 6% (v/v) MPD.

Capillary diffusion achieves a much broader screening of variables in one single experiment. Here we chose CrystalHarp$^{TM}$ plate (Molecular Dimensions), designed for crystallization based on capillary diffusion for crystallization screening and optimization (Figure 6.18A). Table 6.5 listed all of the 24 conditions, which were used in this experiment. A protein solution containing 8.5 mg/ml R149C MelB in 20 mM Tris (pH 7.5), 0.017 % (w/v) DDM, 2mMDDT, 100 mM NaCl, 5% (v/v) Glycerol, 10 mM melibiose, 0.5 mM EDTA was injected to the capillary, then per crystallization condition was pipetted into the precipitation wells as described in the User Guide. Figure 6.18 B, C shows the photos of some crystalline forms. These crystals (crystal-like) are not similar to those obtained using normal methods. Some of these crystals were tested at synchrotron light source. Unfortunately, there was no diffraction.

4.1.4.3.12 Improving diffraction by humidity control

The controlled dehydration of protein crystals has been shown to improve the diffraction qualities of many proteins (Russi et al. 2011). The Humidity Control device (HC1) from the diffraction instrumentation group at the EMBL provides an easy to use dehydration setup that can be used in conjunction with the world-class MX beamlines at the ESRF. Crystals are mounted on meshes (from Mitegen or MDL) and kept in a stream of humidified air from a modified cryostream nozzle. Once crystals have been conditioned, they can be immediately cryocooled by simply unmounting them with the sample changer (Sanchez-Weatherby et al. 2009). We used the dehumidifier for the MelB crystals during our visit. Unfortunately, MelB crystals cracked soon when kept in a stream of humidified air.

4.1.4.4 Preparation of heavy atom derivative

Despite the development in recent times of a range of techniques for phasing macromolecules, the conventional heavy-atom derivatization method still plays a significant role in protein structure determination. Table 6.6 summarized the methods used in the structure determination of the membrane transport proteins.

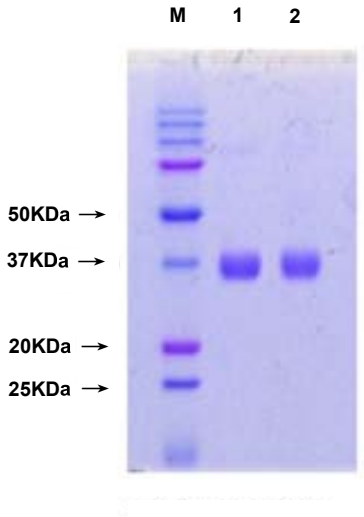



**Figure 6.19** SDS/PAGE gel of R149C MelB (1) and selenomethionine-labelled R149C MelB (2) at concentrations that range from 5-7 mg/ml in 0.017 % DDM after gel purification step and ready for crystallization assays. The gel is loaded with 2 µl of sample.

4.1.4.4.1 Preparation of seleno-L-methionine -labelled R149C MelB

Incorporation of selenomethionine into proteins in place of methionine aids the structure elucidation by X-ray crystallography using single- or multi-wavelength anomalous diffraction (SAD or MAD). The incorporation of heavy atoms such as selenium helps to solve the phase problem in X-ray crystallography. Selenomethionine (Se-Met) labeled R149C MelB has been expressed in *E. coli* B834 (DE3) grown in enriched LeMaster medium (Hendrickson et al. 1990) containing L-selenomethionine. In this case, enough protein for crystallization assays (about 3mg SeMet-R149C from 8 L culture) has been obtained by following the same purification steps described previously for the native protein (see **PART III 4.1.1)**. The purity of Se-Met R149C MelB was checked by SDS-PAGE (Figure 6.19). To conserve sample, we performed a small grid screen in 24-well plate as described at **PART III 4.1.4.3.8**. After several attempts, no SeMet-R149C crystal has been obtained yet.

4.1.4.4.2 Soaking with heavy atom

As it is difficult to get SeMet-R149C crystals, then we went to try to soak crystals with heavy atoms. Mercury derivatives were prepared by soaking crystals in 2 mM

Ethyl mercurithiosalicylate (Thimerosal, Hampton Research) solution for 1min, 5 min, 10 min, 1 h, and 2 h at 19 ºC, and then were fleshly frozen in liquid nitrogen. Crystals were stable in the solution containing heavy atoms.

**Table 6.5** Lists of the 48 condition for crystal screening using crystal harp

| No. | salt | precipitant | buffer | pH | additive (detergent/lipids) | additive (non-detergent) |
|---|---|---|---|---|---|---|
| 1 | 200mM $CaCl_2$ | 30% (w/v) PEG 400 | 100 mM HEPES | 7 | 0.8 mM CHAPS MMA | 3% (v/v) 1,6-hexanediol |
| 2 | 450 mM $MgCl_2$ 100mM KCl | 34% (w/v) PEG 400 | 25 mM NaCitrate | 4 | 1% (w/v)OG | 0.5% (v/v) ethanol |
| 3 | 100 mM NaCl | 27% (w/v) PEG 1500 | 100 mM Tris | 8.5 | 0.1% (w/v)DDM + 0.04% (w/v) $C_{12}E_9$ | 5mM $MgCl_2$, 25% (v/v) MPD, 20% (v/v) glycerol |
| 4 | 200 mM NaCl | 18% (w/v) PEG 500 | 100 mM HEPES | 7 | 1.2 % (w/v)OG | None |
| 5 | 100 mM KCl/NaCl | 25% (w/v) PEG 400 | 50 mM Na Citrate | 4 | 0.1 % (w/v)DDM | None |
| 6 | 100 mM $MgCl_2$ | 30% (w/v) PEG 400 | 100 mM MES | 6 | 0.4 % (w/v)NG | 50mM NaF |
| 7 | None | 40% (w/v) PEG 400 | 100 mM Na Citrate | 4 | None | None |
| 8 | None | 40% (w/v) PEG 400 | 100 mM CAPS | 10 | None | None |
| 9 | 100 mM NaCl | 40% (w/v) PEG 400 | 100 mM Na Citrate | 4 | None | None |
| 10 | 100 mM NaCl | 40% (w/v) PEG 400 | 100 mM CAPS | 10 | None | None |
| 11 | 100 mM $MgCl_2$ | 40% (w/v) PEG 400 | 100 mM Na Citrate | 4 | None | None |
| 12 | 100 mM $MgCl_2$ | 40% (w/v) PEG 400 | 100 mM CAPS | 10 | None | None |
| 13 | 100 mM $CaCl_2$ | 40% (w/v) PEG 400 | 100 mM Na Citrate | 4 | None | None |
| 14 | 100 mM $CaCl_2$ | 40% (w/v) PEG 400 | 100 mM CAPS | 10 | None | None |
| 15 | None | 40% (w/v) PEG 400 | 100mM phosphate/citrate | 4.2 | None | None |
| 16 | 100 mM Hg-Acetate | 40% (w/v) PEG 400 | 100 mM Na Citrate | 4.2 | None | None |
| 17 | 100mM $ZnSO_4$ | 40% (w/v) PEG 400 | 100 mM Na Citrate | 4.2 | None | None |
| 18 | 100mM LiCl | 40% (w/v) PEG 400 | 100mM phosphate/citrate | 4.2 | None | None |
| 19 | 100mM NaCl | 40% (w/v) PEG 400 | 100mM phosphate/citrate | 4.2 | None | None |
| 20 | 100mM KCl | 40% (w/v) PEG 400 | 100mM phosphate/citrate | 4.2 | None | None |
| 21 | 100mM NaI | 40% (w/v) PEG 400 | 100mM phosphate/citrate | 4.2 | None | None |
| 22 | None | 40% (w/v) PEG 400 | 100mM phosphate/citrate | 4.2 | 0.35 % (w/v)DM (4x CMC) | None |
| 23 | None | 40% (w/v) PEG 400 | 100mM phosphate/citrate | 4.2 | 0.73 % (w/v)NG | None |
| 24 | None | 40% (w/v) PEG 400 | 100mM phosphate/citrate | 4.2 | 1.12% (w/v)NM | None |
| 25 | None | 40% (w/v) PEG 400 | 100mM phosphate/citrate | 4.2 | 0.024% (w/v) $C_{12}E_8$ | None |
| 26 | None | 40% (w/v) PEG 400 | 100mM phosphate/citrate | 4.2 | 0.018% (w/v) LDAO | None |
| 27 | None | 40% (w/v) PEG 400 | 100mM phosphate/citrate | 4.2 | 3.44% (w/v)OG | None |
| 28 | None | 40% (w/v) PEG 400 | 100mM phosphate/citrate | 4.2 | 0.032 % (w/v) $C_{12}E_9$ | None |
| 29 | None | 40% (w/v) PEG 400 | 100mM phosphate/citrate | 4.2 | 1 % (w/v) $C_8E_4$ | None |
| 30 | None | 40% (w/v) PEG 400 | 100mM phosphate/citrate | 4.2 | E. coli lipids 18:1 | None |
| 31 | None | 40% (w/v) PEG 400 | 100mM phosphate/citrate | 4.2 | None | 6% (v/v) 1,6-hexanediol |
| 32 | None | 40% (w/v) PEG 400 | 100mM phosphate/citrate | 4.2 | None | 6% (v/v) MPD |
| 33 | None | 40% (w/v) PEG 400 | 100mM phosphate/citrate | 4.2 | None | 8% (w/v) 1,3-butanediol |
| 34 | None | 40% (w/v) PEG 400 | 100mM phosphate/citrate | 4.2 | None | 6% (w/v) Trimethylamine N-oxide dihydrate |
| 35 | None | 40% (w/v) PEG 400 | 100mM phosphate/citrate | 4.2 | None | 6% (v/v) glycerol |
| 36 | None | 40% (w/v) PEG 400 | 100mM phosphate/citrate | 4.2 | None | 6% (w/v) ethyleneglycol |
| 37 | None | 50% (w/v) PEG 200 | 100mM phosphate/citrate | 4.2 | None | None |
| 38 | 100 mM $MgCl_2$ | 50% (w/v) PEG 200 | 100 mM Na Citrate | 4.2 | None | None |
| 39 | 100 mM $CaCl_2$ | 50% (w/v) PEG 200 | 100 mM Na Citrate | 4.2 | None | None |
| 40 | 100 mM NaCl | 50% (w/v) PEG 200 | 100mM phosphate/citrate | 4.2 | None | None |

| | | | | | | |
|---|---|---|---|---|---|---|
| **41** | None | 25% (w/v) PEG 4000 | 100mM phosphate/citrate | 4.2 | None | None |
| **42** | 100 mM $MgCl_2$ | 25% (w/v) PEG 4000 | 100 mM Na Citrate | 4.2 | None | None |
| **43** | 100 mM $CaCl_2$ | 25% (w/v) PEG 4000 | 100 mM Na Citrate | 4.2 | None | None |
| **44** | 100 mM NaCl | 25% (w/v) PEG 4000 | 100mM phosphate/citrate | 4.2 | None | None |
| **45** | 100 mM $CaCl_2$ | 32% (w/v) PEG 400 | 100 mM HEPES | 7.5 | None | None |
| **46** | 100 mM $MgCl_2$, 100 mM NaCl | 45% (w/v) PEG 400 | 100 mM Na Citrate | 5 | None | None |
| **47** | 100 mM NaCl, 100 mM $LiSO_4$ | 20% (w/v) PEG 4000 | 100 mM Na Acetate | 4.5 | None | None |
| **48** | 100 mM $MgCl_2$, 100 mM NaCl | 45% (w/v) PEG 400 | 100 mM cacodylate | 6.5 | None | None |

**Table 6.6** Summary of the methods were used in membrane transport protein structure determination

| Protein[a] | Method |
|---|---|
| C154G LacY | Soaking with methylmercuric acetate (MMA, 3-fold molar excess over protein) for 3 h at 20 °C |
| *E. coli* NhaA | SeMet-labelled protein |
| *E. coli* GlpT | Soaking crystals in $Na_9[P_2Nb_3W_{15}O_{62}].nH_2O$; SeMet-GlpT |
| *A. aeolicus* LeuT | Selenomethionine-labelled protein |
| *E. coli* EmrD | Determined by a two-wavelength anomalous scattering experiment at the Gold LIII edge using a gold thiomalate derivative |
| *E. coli* FucP | Soaking crystals for 10 min in mother liquor containing 1mg/ml $HgCN_2$ |
| *E. coli* vSGLT | Incubated with methylmercuric acetate (MMA) in a 2-5 fold molar excessfor 2 hours at 20 °C prior to crystallization; Selenomethionine-labelled protein |
| *M.liquefaciens* Mhp1 | Soaked for 1 min in 2.5mM methyl mercury acetate, soaked in 10mM $K_2P_t(NO_2)_4$ overnight, or soaked in 0.5mM ethyl mercury thiosalicylate (EMTS) overnight. |
| *E. coli* AdiC | Soaking crystals for 4 hours in mother liquor containing 10 mM $OsCl_3$ followed by back-soaking for 1 min in well buffer plus 0.5% (w/v) NG. |
| *P. Horikoshii* GltPh | Soaking crystals in a solution containing 50mM $K_2P_t(NO_2)_4$ for 6 h followed by a 1 h back-soak. |
| *E. coli* YiiP | Incubating crystals with 2 to 50 mM heavy metal compounds for 2 to 24 hours at 20 °C. |
| *S.aureus* Sav1866 | Multiple isomorphous replacement with anomalous scattering using data from xenon (collected at a wavelength of 1.54004 Å), $Ta_6Br_{14}$ (collected at a wavelength of 1.25447 Å), selenomethionine (collected at a wavelength of 0.97894 Å), 2´-iodo-ADP (collected at a wavelength of 1.06994 Å), and ethyl mercury phosphate (EMP, collected at a wavelength of 1.00799 Å) derivative crystals |
| *E.coli* $MalFGK_2$ | Selenomethionine-labelled protein |
| *E.coli* FocA | Soaking crystals for 24 h in mother liquor containing 2 mM $K_2P_tCl4$ followed by back-soaking for 3 min in well buffer plus 1.0% (w/v)OG. |
| *M.jannaschii* ApcT | Selenomethionine-labelled protein |
| *V.cholerae* NorM | Soaking with either 5mM mercury(II) acetate ($Hg(OAc)_2$), 5mM Mercury(II) cyanide |
| | ($Hg(CN)_2$) or 5mM Bakers dimercurial (BDM) |

[a]: for references, see Table 6.1

4.1.4.5 Initial X-ray diffraction experiments

After extensive crystallization trials, several different form MelB R149C mutant crystals were obtained. Figure 6.20A shows the hexagonal rod MelB crystal, which appear after three days to one week and continue to grow for several more weeks to sizes that range from 100-200 µm. These crystals are mounted in cryo-loops and flash-frozen in liquid nitrogen using their own mother liquor as cryoprotectant (30-50% PEG 200). We were able to collect diffraction data at ID23-1 (ESRF, Grenoble) to ~ 8 Å resolution (Figure 6.20B). The data were processed with XDS (Kabsch 1993), and SCALA (CCP4 1994) and the space group was determined manually and with the help of POINTLESS (CCP4 1994) (Table 6.7). Assuming 1-2 molecules per asymmetric unit, we obtain a Matthew's coefficient of ~8.6-4.8 Å3/Da or a solvent content of 85-70% which resemble the values observed in other transporters (Abramson et al. 2003; Huang et al. 2003).

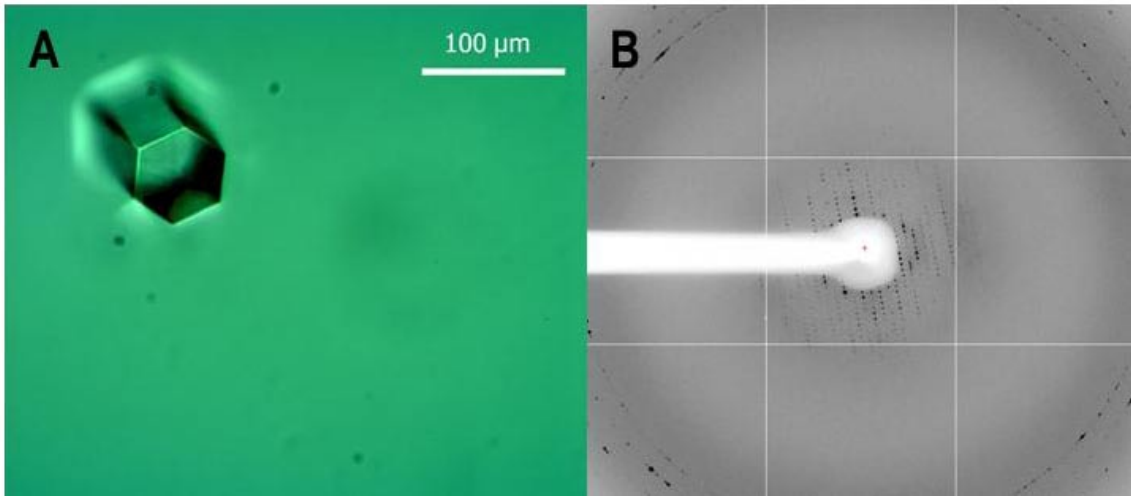



**Figure 6.20**(A) Crystal of R149C mutant form of E. coli MelB in the PEG200 condition. (B) X-ray diffraction pattern obtained at ID23-1 from crystals shown in (A). The crystal was flash frozen in liquid nitrogen in its own mother liquor (~40% PEG200) which served as cryoprotectant and mounted in a Molecular Dimensions cryoloop/caps. This frame was obtained from a 1 º-1 s X-ray exposure.

Clearly, if we intend to gain structural information from the crystals just shown, they have to be improved. Their resolution has to be increased from the current 8 Å to at least 4 or 3 Å, so the crystals will need careful and lengthy optimization.

**Table 6.7** Summary of X-ray and structural statistics of MelB R149C crystal[a]

| **Crystal Parameters:** | |
|---|---|
| Space group | P $6_1$ 2 2 |
| Cell dimensions: | |
| a, b, c (Å) | 216.3, 216.3, 132.4 |
| α, β, γ (º) | 90, 90, 120 |
| Matthews coefficient (Å$^3$/Da) | 4.8 |
| Solvent content (%) | 70 |
| Mw Protein (Da) | 52,200 |
| **Data Quality:** | |
| Wavelength (Å) | 0.9724 |
| Resolution (Å) | 50.0 - 8.0 (8.43 - 8.00) |
| $R_{sym}$ (%) [b] | 3.7 (53) |
| No. of reflections (unique) | 10,096(2,137) |
| Mean Redundancy | 4.7 (5.6) |
| Completeness, overall (%) | 99.5 (100.0) |
| Mean I/$\sigma_I$ | 25.1 (1.9) |

[a] Standard definitions were used for all parameters(Drenth 1994). Entries in parentheses report data from the limiting resolution shell. Data collection and refinement statistics come from SCALEPACK (Otwinowski and Minor 1997)and CNS (Brunger et al. 1998) respectively.[b] All observations with I≥-3$\sigma_I$ were merged and included in the calculation of $R_{sym}$.

## 4.2 Crystallization of MelB wt

Despite years of research into the crystallization of MelB wild-type, no crystal was obtained. The breakthrough was made from R149C MelB mutant, as described in **PART III 4.1.** Then we tried to apply the crystallization experience, summarized from R149C MelB.

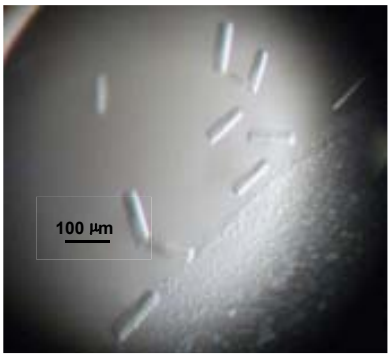



**Figure 6.21**MelB WT crystals obtained from the condition of PEG200 49%, 50 mM phosphate/citrate pH4.2.

MelB wild-type (below described as MelB) has been cloned, expressed and purified similarly as previously reported (**PART III4.1.1)**. In short, the proteins were expressed in *E.coli* DW2-R grown in M9 at 30 ºC. Membranes were isolated from disrupted cells and solubilized with 1.5% (w/v) DDM (Anatrace). MelB was affinity purified on a Ni-NTA affinity column (Bio-Rad). The purified MelB was then subjected to buffer exchange by the batch affinity method by using a small volume of Ni-NTA resin to control the concentration of the detergent in protein solution. MelB in 0.017% (w/v) DDM, 10 mM Tris-HCl (pH 7.5), 5% (v/v) Glycerol, 10 mM melibiose, 100 mM NaCl, and 0.5 mM EDTA was subjected to a small grid screen as described in **PART III 4.1.4.3.8** at 18 ºC. Some crystals were obtained from the similar conditions as those of R149C MelB (Figure 6.21). Then we went to do additive screening and detergent screening as described in **PART III4.1.4.3.7** and**4.1.4.3.8**. Several different sharp crystals were obtained from the conditions with different additives (Figure 6.22). These crystals diffracted poorly at synchrotron (ESRF, ID14-4), usually at 17-20 Å. However, R149C MelB crystals obtained from similar conditions, diffracted to about 8 Å at synchrotron (ESRF, ID23-1). These facts implied how difficult it is to work with wild-type membrane transport proteins.

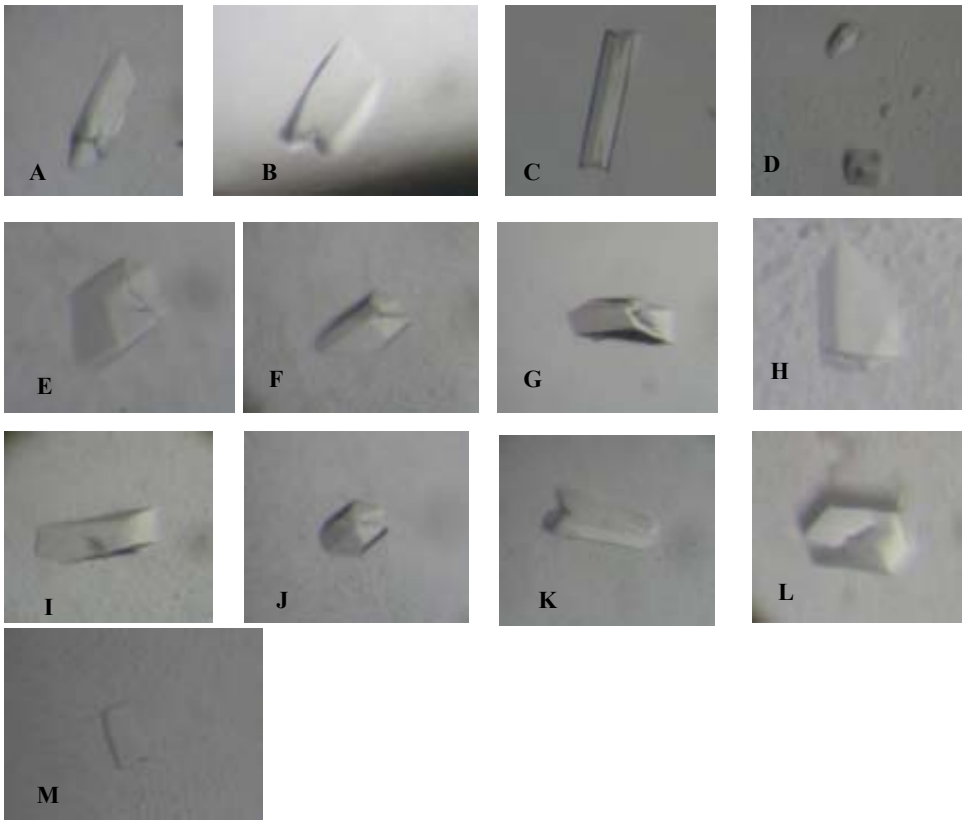



**Figure 6.22**MelB WT crystals obtained from detergent screen. MelB WT crystals grew at the condition of 48% (w/v) PEG200, 50 mM phosphate/citrate pH4.2 with the addition of 0.56 mM 6-Cyclohexyl-1-hexyl-β-D-maltoside (A), 0.59 mM n-Undecyl-β-D-maltopyranoside (B), 1.02 mM (1H, 1H, 2H, 2H-Perfluorooctyl)-β-D-maltopyranoside (C), 1.4 mM N,N-bis-(3-D-Gluconamidopropyl)deoxycholamide (D), 1.8 mM n-Decyl-β-D-maltopyranoside (E), 2.9 mM n-Nonyl-β-D-thioglucopyranoside (F), 3.2 mM n-Nonyl-β-D-thiomaltopyranoside (G), 6.0 mM n-Nonyl-β-D-maltopyranoside (H), 6.5 mM n-Nonyl-β-D-glucopyranoside (I), 7.6 mM 4-Cyclohexyl-1-butyl-β-D-maltoside (J), 29.5% n-Octyl-β-D-galactoside (K), 0.5 M Dimethyl(2-hydroxyethyl)ammonium propane sulfonate (L), 0.5 M 3-(1-Methylpiperidinium)-1-propane sulfonate (M).

## 4.3 Cystallization of other mutants

The R141C, R149Q, K377C and other mutant forms of MelB are obtained by PCR using the cys-less MelB-carrying plasmid as template, expressing and purifying them in

the same strain as described in **PART III4.1.1**. The ΔCtR149C or C-terminally truncated form of R149CMelB was obtained by controlled trypsin digestion. In this case protein samples of the R149C mutant obtained after Ni-NTA step at ~ 1 mg/ml concentration in 0.1% (w/v)DDM, 100 mM imidazole, 0.1 mM AEBSF, 10 mM melibiose, 100 mM NaCl, 10 % (v/v) glycerol, and 5mM beta-mercaptoethanol were incubated with trypsin, at a ratio of 1:20 (w/w) at 37 ºC for 2 hours. Samples were analyzed by Coomassie Blue-stained SDS-PAGE (Figure 6.23). This procedure yields a stable truncated form of R149C MelB, where the endopeptidase cuts the polypeptide at position Arg-442, thus removing the last C-terminal 17 residues of MelB and the engineered His-tag. ΔCtR149C were further purified to separate other proteins by Size-Exclusion HPLC (SE-HPLC, ÄKTA™ Purifier, GE). Purified ΔCtR149C were concentrated to 9 mg/ml by AmiconUltra 15 centrifugal concentrators (Millipore) 100 KDa molecular weight cutoff.

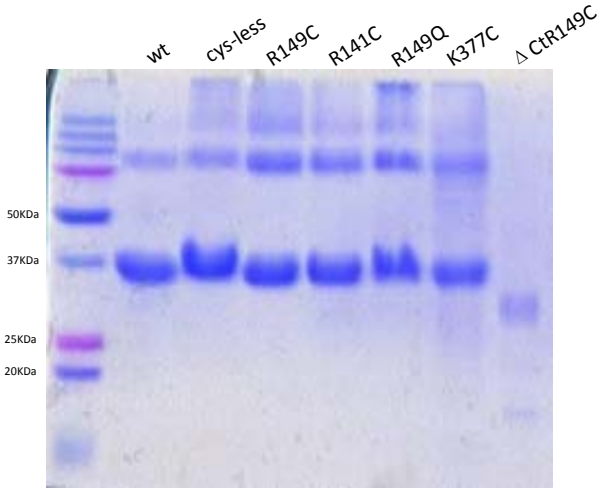



**Figure 6.23** SDS-PAGE gel of several protein samples at concentrations that range ~10-15 mg/ml in 0.1 % DDM.

These modifications are supposed to arrest the transporter in a defined conformation, affect ligand binding, or just be more suitable for crystallization. They are expressed and purified in a similar manner as shown for the original MelB form with the only exception that the protein yield is a bit lower.

Some of these mutant forms have been purified in DDM and LAPAO by gel filtration (using the methodology explained in the previous paragraphs) and show single elution profiles that are consistent with a monodisperse protein sample, suitable for crystallization assays.

All of these mutants and ΔCtR149C were subjected to extensive crystallization trials against some commercially obtained sparse matrix screens or a grid screen summarized from the experiments of the crystallization of R149C mutant. Despite several attempts, no crystal was obtained from any mutant forms MelB or ΔCtR149C. C-terminally truncated form of R149C MelB was difficult to crystallize, possibly because of an excess of detergent in the protein solution. As we discussed in **PART III4.1.1**, the concentration of detergent in protein solution is essential for obtaining MelB crystals. Without His-tag, we still didn't find an effective way to control the concentration of detergent in the protein solution. One way may be useful to reach this aim. Before digestion with Trypsin, ΔCtR149C may be firstly detergent-exchanged to other detergent with smaller micellar average molecular weight, e.g., Undecyl-β-D-Maltopyranoside (UDM, 35.2 KDa). Then it would be possible to concentrate the sample by using a centrifugal concentrator. We had performed same trials with R149C MelB. The results showed that R149C in the solution containing UDM can be crystallized, suggesting that the concentration of UDM should not be co-concentrated by 50 KDa centrifugal concentrator. In contrast to R149C MelB, when we performed crystal screening for R149Q MelB, only a brown precipitate was obtained, implying R149Q mutant may be not stable in the same condition as that for R149C mutant. In coherence with previous studies as described in **PART I**, the replacement of Arg149 by Gln may cause the MelB to become unstable.

# 5 General Discussions of Part III

Inspired by the successful structure determination of about 40 membrane transport protein (see Table 6.1), we took a similar approach to crystallize MelB. The crystallization of the MelB is summarized in Figure 6.24. The key elements in our experiments were: i) identifying a rigid protein core by modifying the protein construct; ii) prescreening the detergents for their ability of maintaining protein monodispersity; iii) performing extensive crystallization screens for every detergent identified in the previous step and repeating any condition several times to improve reproducibility; iv) improving protein contact in the crystals by using a detergent mixture. Hence, the emphasis was as much on the protein itself, as it was on detergent type or efficient screening and refinement of the crystallization conditions.

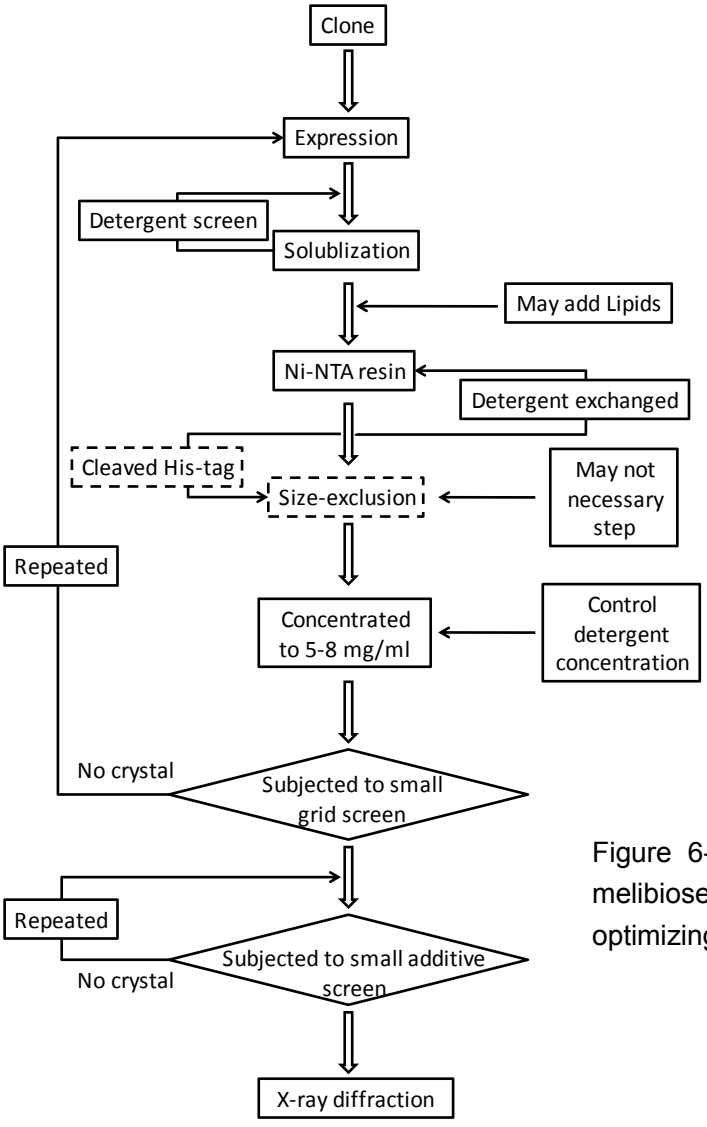



Figure 6-24 Workflow for generating melibiose permease crystallization and optimizing crystals.

## 5.1 Important to identify a rigid protein core for crystallization studies

Secondary transporters use energy to transport molecules across a membrane by a series of conformational changes. In general, the translocation is expected to take place by means of an alternating access mechanism, where the structurally coupled substrates binding sites present access to the extracellular medium in the binding process and to the cytoplasmatic medium in the release process (Krishnamurthy et al. 2009). This mechanism has been supported by numerous kinetic, biochemical, biophysical studies and atomic-resolution structural evidences (Abramson et al. 2003; Huang et al. 2003; Singh et al. 2008; Krishnamurthy and Gouaux 2012). A major characteristic of these proteins is a dynamical behaviour, needed for translocation of substrates, but making them difficult to crystallize. In many cases, finding a rigid protein core by mutating or trimming the original protein will facilitate the crystallization and help to improve the resolution. For example, lactose permease was firstly crystallized using a mutant, i.e., C154G, which blocks the protein in an inward-facing conformation. Higher resolution crystals of GlpT were obtained by trimming original protein. In our case, we attempted several times to crystallize MelB wild-type, without success. The breakthrough was that we found one mutant, i.e., R149C, which can bind substrates, but fixes MelB in single conformation. We were successful to obtain R149C MelB crystals from some conditions as that of other secondary transporters. We summarized R149C MelB crystallization experiences, and developed a simplified grid screen. MelB wild type was then subjected to this grid screen. We obtained MelB crystals from some condition as that of R149C MelB. In contrast, no crystal was obtained from R149Q MelB, R141CMelB, and K377C MelB when we performed the same grid screen. These interesting results indicate that MelB wild type and R149C MelB should share similar structural features. In agreement with the studies of crystallization, previous studies, based on analysis of infrared spectroscopy and fluorescence spectroscopy have shown that all of R149Q (see PART I), R141C (Leon et al. 2009), and K377C (our unpublished work) mutants may change MelB native structure to some degree. Although the WT MelB crystals could be

obtained from the same conditions as those of R149C MelB, the resolution is clearly lower, which may imply differences of protein-protein interactions in crystal packing. One reasonable assumption already commented above may be that the R149C mutation fixes the MelB in a relatively stable conformation, which reduces the conformational dynamics, with the benefit of forming a stable crystal cell. Our previous studies as described in PART I, showed that R149C MelB trapped MelB in inward-facing conformation. In short, to identify a stable structure that reduces the conformational dynamics is essential for obtaining high-resolution crystals.

## 5.2 The importance of the concentration of detergent

We observed that a critical factor for obtaining membrane protein crystals is the concentration of detergent in the protein solution. Due to the high hydrophobicity, membrane proteins should be prepared with the help of detergents. However, excess detergent will disturb the formation of protein crystals, which may be due to a reduction of protein-protein interactions, and then reduce crystallizability. In general, protein crystallization depends on random protein-protein interactions (Carugo and Argos 1997). If a membrane protein is surrounded by an excess of detergent micelles, it will expose less protein surface areas for protein-protein contacts. On the other hand, too much detergent can denature the protein or impede crystallization by phase separation (Wiener 2004). We observed that the main reason which causes excess detergent is that detergent is co-concentrated during sample protein preparation by a means of Molecular Weight Cutoff Filter. For MelB, this effect should be quite important. No crystals were obtained when the concentration of the detergent was not controlled. We had tried to use 50 kDa or 100 KDa molecular weight cutoff to concentrate protein sample. In any case, no crystal was obtained, implying that the 50 KDa or 100 KDa molecular weight cutoff filter co-concentrated the detergent during the protein preparation. Actually, many researchers reported that they used 50 kDa or 100 KDa molecular weight cutoff to

concentrate protein in the presence of DDM. Therefore, MelB should be quite sensitive to the concentration of detergent. We may suppose that the protein interactions of MelB in DDM containing solution are very weak, which may be the cause of the difficulty of the crystallization of MelB. MelB seems quite stable in DM and NM, since the detergent can be concentrated to high values. However, when we performed crystal screens for these samples, no crystal was obtained, suggesting that there are only detergent-mediated interactions. Our observation implies that DDM, DM and NM, which significantly protect hydrophobic proteins may be suitable for the stability of MelB. However this type of detergent reduces the protein-protein interactions, and then reduces crystallizability. A mixture of detergents may be effective to improve the crystallizability and the resolution.

## 5.3 Screening for crystal nucleation of MelB

Extensive crystallization screens were essential to obtain crystals. Sparse matrix screens allow sampling of a broad range of conditions screens that are highly systematic with respect to pH as well as precipitant type and concentration. We used a number of commercially available screens: MemGold, MemSys&MemStart, JCSG-plus and PACT (Molecular Dimensions). MemGold, designed for membrane proteins, is based on the conditions that have successfully generated α-helical membrane protein crystals used to solve X-ray structures (Newstead et al. 2008) and contains 96 conditions covering a range of pH, PEGs and salt additives. MemStart is a sparse matrix screen, whereas MemSys is a systematic exploration of pH, salt concentration/type and precipitant concentration/type. MemSys&MemStartwere designed for screening and optimizing crystallization conditions for transmembrane proteins using vapour diffusion methods. JCSG-plus is a 96 reagent, optimized sparse-matrix screen of classic and modern conditions. PACT has been developed to systematically test the effect of pH, anions and cations, using PEG as the precipitant.

A survey of the published literature reveals that small MW PEGs, in particular PEG 400 have been more successful for α-helical membrane protein crystallization, as the precipitant (Newstead et al. 2008). Organic solvents tend to disturb the detergent micelles and, at high concentrations, denature membrane proteins. Salt, on the other hand, reduces the solubility of the detergent micelles (Zulauf 1991) and often precipitates the membrane protein embedded in the detergent micelle before crystallization occurs(Lemieux et al. 2003). For MelB, crystals were obtained from the conditions using small MW PEGs as precipitant. PEGs are therefore the best choice for membrane protein crystallization. Another important factor is the pH(McPherson 1990). We surveyed the entire buffer which was used for the membrane transporter protein of known structure. Thus, we used screening kits with 4 small MW PEGs over the pH range of 3.5 to 9.0 (McPherson 1990) based on the initial crystallization results of the protein in DDM. Indeed, their application to MelB purified in several detergents produced numerous hits, which were subsequently pursued for optimization. Furthermore, the utilization of these screening kits significantly improved the efficiency and reproducibility of crystallization experiments, which was particularly beneficial when a dozen of detergents were screened for crystallization. Finally, the 4 PEG/pH screening kits used to search for nucleation of R149C MelB crystals may serve as a starting point for other mutant form MelB.

## 5.4 The importance of identifying suitable detergents

The membrane protein molecule in solution exists as a protein-detergent-lipid complex. Specific lattice contacts in any protein crystal are made exclusively via protein-protein interactions, and too large a detergent micelle can be an obstacle for protein crystallization(Michel 1983; Lemieux et al. 2003). By reducing the heterogeneity of the protein surface and by optimizing the detergent micellar size and shape, we essentially increase the area available for the formation of lattice contacts,

thereby improving the protein crystallizability (Kwong et al. 1999; Lemieux et al. 2003).In many cases, a second detergent or a third detergent has shown to be essential for improving the diffracting resolution. For example, Wang and colleagues reported that for the formate transporter FocA, Cymal-2 was shown to be essential for improving the diffracting resolution from 3.5 Å to 2.2 Å (Wang et al. 2009). In another case, Lemieux and colleagues reported that a detergent mixture of DDM and $C_{12}E_9$ is a requirement to give crystals that diffracted to 3.3 Å resolution, but DDM alone resulted in crystalline order to 7 Å, while complete detergent exchange to $C_{12}E_9$ gave no crystals (Lemieux et al. 2003). Table 6.1 listed detergents used in the transporters of known structure. Based on the data obtained from the transporters of known structure, we found DDM, UDM, DM, OG, and lipid-like detergents, e.g., Cymal-6 were quite common used to extract the protein from the plasma membrane. And $C_{12}E_8$, $C_{12}E_9$ and LDAO might use as second detergent for protein preparation(Huang et al. 2003; Wang et al. 2009; He et al. 2010). In order to improve efficiency and simplify the experiment, we combined published experimental data and our experimental data of R149C MelB, to list some of the most commonly used detergents i.e., UDM, DM, LDAO, OG, NM, NG, β-OM, α-DDM, $C_8E_4$, $C_{12}E_8$, $C_{12}E_9$, Cymal-5, Cymal-6, Fos-choline-12. Then we went to prepare MelB or other mutant forms of MelB with DDM, UDM, DM, OG, and Cymal-6 and performed detergent screen as described above. Following this method, we were successful to obtain MelB wt crystals, and many different sharp crystals of R149C MelB were obtained by this screening.

## 5.5 Important to repeat

In our studies, we note that unlike soluble proteins, in spite that we were strictly following the same experimental procedures, it was quite difficult to reproduce MelB crystals. We believe that the problem lies in the protein preparation. In fact, our observation showed that only a small percentage of new prepared protein could be

crystallized. It may be due to the unstability of protein-detergent complex (PDC). It is clear that the detergent micelles are not as stable as lipid liposomes or plasma membranes. Several factors as temperature, mechanical force, the amount of plasma lipids associated with protein, etc. would affect the stability of PDC, and increasing the disorder of the protein conformation. Obviously, we cannot guarantee that each experiment can be conducted in an absolutely consistent situation. The effect of external factors on the PDC might lead to two possible outcomes: i) the protein-detergent complex is completely undermined, so the detergent cannot continue to protect the host protein. In this case, the membrane proteins expose their hydrophobic surface to hydrophilic environment, then will precipitate; ii) the protein-detergent complex is disordered that lead to some of protein hydrophobic surface would expose to hydrophilic environment. More detergent should be needed to cover the exposure of hydrophobic surfaces. Therefore, host protein is covered by more detergents. In this case, there is only detergent-mediated protein-protein interaction that will reduce the crystallizability. Apparently the reason why the crystallizability experiments resulted in lower reproducibility corresponds to the latter case. In fact, it is inevitable to change the structure of PDC during protein preparation for crystallization studies. The biggest difficulty is that there is still not any effective method used to detect the types of protein-protein interaction. So we needed to perform extensive crystallization trials for the purified proteins. To improve the efficiency and reproducibility of the crystallization experiments, it is necessary to establish a high-speed and convenient crystal screening methods. Based on the data obtained from some of other transporters of known structure and the experience of the crystallization of MelB, we developed a simple crystal screening method combining sparse matrix screen, detergent screen, and additive screen.

## 5.6 MelB and other transporters of known structure

MelB crystals were obtained from the conditions similar to other transporters of known structure. This interesting phenomenon implies that the crystallization of such proteins may follow a common law. In agreement with this phenomenon, all secondary transporters share some common structural features, e.g., transmembraneα-helix secondary structure, heart-shaped with an internal cavity and an approximate two-fold symmetry, higher conformational dynamics, etc. For example, almost all transporters were crystallized using PEG400 as precipitant. The protein and crystallization buffers for the different proteins are quite similar. Almost all crystals were obtained at 18-20 ℃. The difference of the crystallization conditions between different proteins may be because of the differences of surface properties, dimension, stability, etc. The core work of the crystallization of this type of proteins is to find a rigid protein to reduce dynamics and a suitable detergent or detergents mixture to improve the stability of the protein and the protein-protein interactions in the crystal pack. Mutagenesis is an extremely useful method to prepare structure-specific proteins. In many cases, more stable structure of the protein can be obtained by site-directed mutagenesis or trimming the original protein. On the other hand, size exclusion chromatography (Lemieux et al. 2003), infrared spectroscopy and fluorescence spectroscopy have been used in detecting the stability of the modified protein or detergent-protein complex. Another feature of membrane protein crystallization is a very low reproducibility. As the detergent micelle is not as strong as plasma membranes or lipids liposomes, the stability of the membrane protein structure will be strongly affected by temperature, mechanical force, and other physical and chemical parameters. All of these factors explain that repeating several times the whole purification and crystallization processes are essential to obtain good crystals and improved resolution.

# 6 Conclusions of Part III

1) The work presented here show that MelB can be crystallized in conditions similar to those of other membrane transport proteins.

2) MelB-MelB interactions in the DDM-containing solutions are affected by the protein preparation method, which may lead to a reduction in the reproducibility of the crystallization experiments.

3) R149C MelB crystals can be obtained in DDM, but these crystals only diffracted to about 8 Å resolution, whereas MelB WT crystals obtained with the same conditions only diffracted to 17 Å. It is advisable to check other mild detergents for the solubilization step.

# Conclusions

1) The R149C mutant can bind $Na^+$ as well as melibiose in proteoliposomes and in ISO membrane vesicles. It cannot bind sugars in RSO membrane vesicles.

2) The R149C mutation fixes the MelB in an inward-facing conformation. Therefore, Arg149, located probably in the cytoplasmic half of transmembrane helix 5, is a crucial side chain for the reorientation mechanism of MelB.

3) Ala155, located probably in the middle of helix 5 is an essential residue for either $Na^+$ or melibiose binding, since the mutant A155C absolutely loses the capability to bind substrates.

4) Some mutants reconstituted into proteoliposomes show hindered substrates binding, like T159C or G161C. However, the same mutants show interactions with substrates in membrane vesicles, which are relatively akin to Cys-less. Therefore, the purification and reconstitution processes lead to some alterations of the MelB structure for these particular mutants.

5) Cysteine-scanning mutagenesis combined with the determination of the sugar accessibility to binding sites shows that many cysteine mutations disturb the sugar accessibility to binding sites from the extracellular matrix. Therefore, the helix 5 should be involved in the reorientation mechanism of MelB from the outward-facing to the inward-facing conformation.

6) The work presented here show that MelB can be crystallized in conditions similar to those of other membrane transport proteins.

7) MelB-MelB interactions in the DDM-containing solutions are affected by the protein preparation method, which may lead to a reduction in the reproducibility of the crystallization experiments.

8) R149C MelB crystals can be obtained in DDM, but these crystals only diffracted to about 8 Å resolution, whereas MelB WT crystals obtained with the same conditions only diffracted to 17 Å. It is advisable to check other mild detergents for the solubilization step.

# References

# A

Abdel-Dayem, M., C. Basquin, T. Pourcher, E. Cordat and G. Leblanc (2003). "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): 1518-1524.

Abramson, J., I. Smirnova, V. Kasho, G. Verner, H. R. Kaback and S. Iwata (2003). "Structure and mechanism of the lactose permease of Escherichia coli." Science, **301**(5633): 610-615.

Almen, M. S., K. J. Nordstrom, R. Fredriksson and H. B. Schioth (2009). "Mapping the human membrane proteome: a majority of the human membrane proteins can be classified according to function and evolutionary origin." BMC Biol, **7**: 50.

Arbely, E. and I. T. Arkin (2004). "Experimental measurement of the strength of a C alpha-H...O bond in a lipid bilayer." J Am Chem Soc, **126**(17): 5362-5363.

Arrondo, J. L. and F. M. Goni (1999). "Structure and dynamics of membrane proteins as studied by infrared spectroscopy." Prog Biophys Mol Biol, **72**(4): 367-405.

Arrondo, J. L., A. Muga, J. Castresana and F. M. Goni (1993). "Quantitative studies of the structure of proteins in solution by Fourier-transform infrared spectroscopy." Prog Biophys Mol Biol, **59**(1): 23-56.

# B

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

Barth, A. and C. Zscherp (2002). "What vibrations tell us about proteins." Q Rev Biophys, **35**(4): 369-430.

Berger, B. W., C. M. Gendron, A. M. Lenhoff and E. W. Kaler (2006). "Effects of additives on surfactant phase behavior relevant to bacteriorhodopsin crystallization." Protein Sci, **15**(12): 2682-2696.

Bergfors, T. (2003). "Seeds to crystals." J Struct Biol, **142**(1): 66-76.

Botfield, M. C., K. Naguchi, T. Tsuchiya and T. H. Wilson (1992). "Membrane topology of the melibiose carrier of Escherichia coli." J Biol Chem, **267**(3): 1818-1822.

Botfield, M. C, Wilson, T. H. (1988). "Mutations that simultaneously alter both sugar and cation specificity in the melibiose carrier of Escherichia coli." J Biol Chem, **263**: 12909-12915.

Breton, J. (2001). "Fourier transform infrared spectroscopy of primary electron donors in type I photosynthetic reaction centers." Biochim Biophys Acta, **1507**(1-3): 180-193.

Briggs, J., H. Chung, and M. Caffrey (1996). "The temperature-composition phase diagram and mesophase structure characterization of the monoolein/water system." J. Phys. II France, **6**: 723-751.

Brunger, A. T., et al. (1998). "Crystallography & NMR system: A new software suite for macromolecular structure determination." Acta Crystallogr D Biol Crystallogr, **54**(Pt 5): 905-921.

Brunton, L. L., Ed. (2006). Goodman & Gilman's The Pharmacological Basis of Therapeutics. New York, NY McGraw-Hill, Medical Publishing Division.

Byler, D. M. and H. Susi (1986). "Examination of the secondary structure of proteins by deconvolved FTIR spectra." Biopolymers, **25**(3): 469-487.

## C

Caffrey, M. and V. Cherezov (2009). "Crystallizing membrane proteins using lipidic mesophases." Nat Protoc, **4**(5): 706-731.

Carugo, O. and P. Argos (1997). "Protein-protein crystal-packing contacts." Protein Sci, **6**(10): 2261-2263.

Cherezov, V., A. Peddi, L. Muthusubramaniam, Y. F. Zheng and M. Caffrey (2004). "A robotic system for crystallizing membrane and soluble proteins in lipidic mesophases." Acta Crystallogr D Biol Crystallogr, **60**(Pt 10): 1795-1807.

Cherezov, V., et al. (2007). "High-resolution crystal structure of an engineered human beta2-adrenergic G protein-coupled receptor." Science, **318**(5854): 1258-1265.

Collaborative Computational Project Number 4 (1994). "The CCP4 suite: program for protein crystallography." Acta Crystallographica Section D Biological Crystallography, **D50**: 760-763.

Cordat, E., G. Leblanc and I. Mus-Veteau (2000). "Evidence for a role of helix IV in connecting cation- and sugar-binding sites of Escherichia coli melibiose permease." Biochemistry, **39**(15): 4493-4499.

Cordat, E., I. Mus-Veteau and G. Leblanc (1998). "Structural studies of the melibiose permease of Escherichia coli by fluorescence resonance energy transfer. II. Identification of the tryptophan residues acting as energy donors." J Biol Chem, **273**(50): 33198-33202.

# D

Dang, S., L. Sun, Y. Huang, F. Lu, Y. Liu, H. Gong, J. Wang and N. Yan (2010). "Structure of a fucose transporter in an outward-open conformation." Nature, **467**(7316): 734-738.

Dave, N., A. Troullier, I. Mus-Veteau, M. Dunach, G. Leblanc and E. Padros (2000). "Secondary structure components and properties of the melibiose permease from Escherichia coli: a fourier transform infrared spectroscopy analysis." Biophys J, **79**(2): 747-755.

Dawson, R. J. and K. P. Locher (2006). "Structure of a bacterial multidrug ABC transporter." Nature, **443**(7108): 180-185.

DeFelice, L. J. (2004). "Transporter structure and mechanism." Trends Neurosci, **27**(6): 352-359.

Deisenhofer, J., O. Epp, K. Miki, R. Huber and H. Michel (1984). "X-ray structure analysis of a membrane protein complex. Electron density map at 3 A resolution and a model of the chromophores of the photosynthetic reaction center from Rhodopseudomonas viridis." J Mol Biol, **180**(2): 385-398.

Ding, P. Z. (2003). "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): 864-869.

Ding, P. Z., M. C. Botfield and T. H. Wilson (2000). "Sugar recognition 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.

Ding, P. Z. and T. H. Wilson (2000a). "The melibiose carrier of Escherichia coli: cysteine substitutions for individual residues in helix XI." J Membr Biol, **174**(2): 135-140.

Ding, P. Z. and T. H. Wilson (2000b). "Physiological evidence for an interaction between helix XI and helices I, II, and V in the melibiose carrier of Escherichia coli." Biochem Biophys Res Commun, **268**(2): 409-413.

Ding, P. Z. and T. H. Wilson (2001a). "Cysteine mutagenesis of the amino acid residues of transmembrane helix I in the melibiose carrier of Escherichia coli." Biochemistry, **40**(18): 5506-5510.

Ding, P. Z. and T. H. Wilson (2001b). "The proximity between helix I and helix XI in the melibiose carrier of Escherichia coli as determined by cross-linking." Biochim Biophys Acta, **1514**(2): 230-238.

Doyle, D. A., J. Morais Cabral, R. A. Pfuetzner, A. Kuo, J. M. Gulbis, S. L. Cohen, B. T. Chait and R. MacKinnon (1998). "The structure of the potassium channel: molecular basis of K+ conduction and selectivity." Science, **280**(5360): 69-77.

Drenth, J. (1994). "Principles of protein X-ray crystallography." New York, Springer-Verlag.

Drew, D., et al. (2005). "A scalable, GFP-based pipeline for membrane protein overexpression screening and purification." Protein Sci, **14**(8): 2011-2017.

Drew, D. E., G. von Heijne, P. Nordlund and J. W. de Gier (2001). "Green fluorescent protein as an indicator to monitor membrane protein overexpression in Escherichia coli." FEBS Lett, **507**(2): 220-224.

Dutzler, R., E. B. Campbell, M. Cadene, B. T. Chait and R. MacKinnon (2002). "X-ray structure of a ClC chloride channel at 3.0 A reveals the molecular basis of anion selectivity." Nature, **415**(6869): 287-294.

# E

Edelheit, O., A. Hanukoglu and I. Hanukoglu (2009). "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.

# F

Fabian, H., D. Naumann, R. Misselwitz, O. Ristau, D. Gerlach and H. Welfle (1992). "Secondary structure of streptokinase in aqueous solution: a Fourier transform infrared spectroscopic study." Biochemistry, **31**(28): 6532-6538.

Faham, S., A. Watanabe, G. M. Besserer, D. Cascio, A. Specht, B. A. Hirayama, E. M. Wright and J. Abramson (2008). "The crystal structure of a sodium galactose transporter reveals mechanistic insights into Na+/sugar symport." Science, **321**(5890): 810-814.

Fahmy, K. (2001). "Application of ATR-FTIR spectroscopy for studies of biomolecular interactions." Recent Res devel Chem, **2**: 1-17.

Fahmy, K., T. P. Sakmar and F. Siebert (2000). "Structural determinants of active state conformation of rhodopsin: molecular biophysics approaches." Methods Enzymol, **315**: 178-196.

Forrest, L. R., R. Kramer and C. Ziegler (2011). "The structural basis of secondary active transport mechanisms." Biochim Biophys Acta, **1807**(2): 167-188.

Forrest, L. R. and G. Rudnick (2009). "The rocking bundle: a mechanism for ion-coupled solute flux by symmetrical transporters." Physiology (Bethesda), **24**: 377-386.

Franco, P. J., A. B. Jena and T. H. Wilson (2001). "Physiological evidence for an interaction between helices II and XI in the melibiose carrier of Escherichia coli." Biochim Biophys Acta, **1510**(1-2): 231-242.

Franco, P. J. and T. H. Wilson (1999). "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): 6377-6386.

Furutani, Y., T. Murata and H. Kandori (2011). "Sodium or lithium ion-binding-induced structural changes in the K-ring of V-ATPase from Enterococcus hirae revealed by ATR-FTIR spectroscopy." J Am Chem Soc, **133**(9): 2860-2863.

## G

Ganea, C., K. Meyer-Lipp, R. Lemonnier, A. Krah, G. Leblanc and K. Fendler (2011). "G117C MelB, a mutant melibiose permease with a changed conformational equilibrium." Biochim Biophys Acta, **1808**(10): 2508-2516.

Gao, X., F. Lu, L. Zhou, S. Dang, L. Sun, X. Li, J. Wang and Y. Shi (2009). "Structure and mechanism of an amino acid antiporter." Science, **324**(5934): 1565-1568.

Garavito, R. M., D. Picot and P. J. Loll (1996). "Strategies for crystallizing membrane proteins." J Bioenerg Biomembr, **28**(1): 13-27.

Garcia-Ruiz, J. M., L. A. Gonzalez-Ramirez, J. A. Gavira and F. Otalora (2002). "Granada Crystallisation Box: a new device for protein crystallisation by counter-diffusion techniques." Acta Crystallogr D Biol Crystallogr, **58**(Pt 10 Pt 1): 1638-1642.

Gerwert, K. (1999). "Molecular reaction mechanisms of proteins monitored by time-resolved FTIR-spectroscopy." Biol Chem, **380**(7-8): 931-935.

Giacomini, K. M., et al. (2010). "Membrane transporters in drug development." Nat Rev Drug Discov, **9**(3): 215-236.

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

Goormaghtigh, E., V. Cabiaux and J. M. Ruysschaert (1994b). "Determination of soluble and membrane protein structure by Fourier transform infrared spectroscopy. II. Experimental aspects, side chain structure, and H/D exchange." Subcell Biochem, **23**: 363-403.

Goormaghtigh, E., V. Cabiaux and J. M. Ruysschaert (1994c). "Determination of soluble and membrane protein structure by Fourier transform infrared spectroscopy. III. Secondary structures." Subcell Biochem, **23**: 405-450.

Granell, M. (2009). "Estudi del lloc d'unió del sodi de la permeasa de melibiosa d'Escherichia coli." Doctoral thesis.

Granell, M., X. Leon, G. Leblanc, E. Padrós and V. A. Lorenz-Fonfria (2010). "Structural insights into the activation mechanism of melibiose permease by sodium binding." Proc Natl Acad Sci U S A, **107**(51): 8322078-8322083.

Guan, L., S. V. Jakkula, A. A. Hodkoff and Y. Su (2012). "Role of Gly117 in the cation/melibiose symport of MelB of Salmonella typhimurium." Biochemistry, **51**(13): 2950-2957.

Guan, L. and H. R. Kaback (2007). "Site-directed alkylation of cysteine to test solvent accessibility of membrane proteins." Nat Protoc, **2**(8): 2012-2017.

Guan, L., O. Mirza, G. Verner, S. Iwata and H. R. Kaback (2007). "Structural determination of wild-type lactose permease." Proc Natl Acad Sci U S A, **104**(39): 15294-15298.

Guan, L., S. Nurva and S. P. Ankeshwarapu (2011). "Mechanism of melibiose/cation symport of the melibiose permease of Salmonella typhimurium." J Biol Chem**286**(8): 6367-6374.

Guan, L., I. N. Smirnova, G. Verner, S. Nagamori and H. R. Kaback (2006). "Manipulating phospholipids for crystallization of a membrane transport protein." Proc Natl Acad Sci U S A, **103**(6): 1723-1726.

Guzman, L. M., D. Belin, M. J. Carson and J. Beckwith (1995). "Tight regulation, modulation, and high-level expression by vectors containing the arabinose PBAD promoter." J Bacteriol, **177**(14): 4121-4130.

Gwizdek, C., G. Leblanc and M. Bassilana (1997). "Proteolytic mapping and substrate protection of the Escherichia coli melibiose permease." Biochemistry, **36**(28): 8522-8529.

# H

Haberstock, S., C. Roos, Y. Hoevels, V. Dotsch, G. Schnapp, A. Pautsch and F. Bernhard (2012). "A systematic approach to increase the efficiency of membrane protein production in cell-free expression systems." Protein Expr Purif, **82**(2): 308-316.

Hacksell, I., J. L. Rigaud, P. Purhonen, T. Pourcher, H. Hebert and G. Leblanc (2002). "Projection structure at 8 Å resolution of the melibiose permease, an Na-sugar co-transporter from *Escherichia coli*." Embo J, **21**(14): 3569-3574.

Hama, H. and T. H. Wilson (1993). "Cation-coupling in chimeric melibiose carriers derived from Escherichia coli and Klebsiella pneumoniae. The amino-terminal portion is crucial for Na+ recognition in melibiose transport." J Biol Chem, **268**(14): 10060-10065.

Hanson, M. A., V. Cherezov, M. T. Griffith, C. B. Roth, V. P. Jaakola, E. Y. Chien, J. Velasquez, P. Kuhn and R. C. Stevens (2008). "A specific cholesterol binding site is established by the 2.8 Å structure of the human beta2-adrenergic receptor." Structure, **16**(6): 897-905.

Hanson, M. A. and R. C. Stevens (2009). "Discovery of new GPCR biology: one receptor structure at a time." Structure, **17**(1): 8-14.

Haris, P. I. and D. Chapman (1992). "Does Fourier-transform infrared spectroscopy provide useful information on protein structures?" Trends in biochemical sciences, **17**(9): 328-333.

Hastings Wilson, T. and D. M. Wilson (1998). "Evidence for a close association between helix IV and helix XI in the melibiose carrier of *Escherichia coli*." Biochim Biophys Acta, **1374**(1-2): 77-82.

Haugland, R. P. (1994). "Spectra of fluorescent dyes used in flow cytometry." Methods in cell biology, **42 Pt B**: 641-663.

He, X., P. Szewczyk, A. Karyakin, M. Evin, W. X. Hong, Q. Zhang and G. Chang (2010). "Structure of a cation-bound multidrug and toxic compound extrusion transporter." Nature, **467**(7318): 991-994.

Hemdan, E. S. and J. Porath (1985a). "Development of immobilized metal affinity chromatography II. Interaction of amino acids with immobilized nickel iminodiacetate." J Chromatogr, **323**: 255–264.

Hemdan, E. S. and J. Porath (1985b). "Development of immobilized metal affinity chromatography III. Interaction of oligopeptides with immobilized nickel iminodiacetate." J Chromatogr, **323**: 255–264.

Henderson, R. and P. N. Unwin (1975). "Three-dimensional model of purple membrane obtained by electron microscopy." Nature, **257**(5521): 28-32.

Hendrickson, W. A., J. R. Horton and D. M. LeMaster (1990). "Selenomethionyl proteins produced for analysis by multiwavelength anomalous diffraction (MAD): a vehicle for direct determination of three-dimensional structure." Embo J, **9**(5): 1665-1672.

Hilf, R. J. and R. Dutzler (2008). "X-ray structure of a prokaryotic pentameric ligand-gated ion channel." Nature, **452**(7185): 375-379.

Huang, Y., M. J. Lemieux, J. Song, M. Auer and D. N. Wang (2003). "Structure and mechanism of the glycerol-3-phosphate transporter from Escherichia coli." Science, **301**(5633): 616-620.

Hunte, C., E. Screpanti, M. Venturi, A. Rimon, E. Padan and H. Michel (2005). "Structure of a Na+/H+ antiporter and insights into mechanism of action and regulation by pH." Nature, **435**(7046): 1197-1202.

Hunter, T. and H. Schulman (2005). "CaMKII structure--an elegant design." Cell, **123**(5): 765-767.

## I

Iwata, S., C. Ostermeier, B. Ludwig and H. Michel (1995). "Structure at 2.8 A resolution of cytochrome c oxidase from Paracoccus denitrificans." Nature, **376**(6542): 660-669.

## J

Jaakola, V. P., M. T. Griffith, M. A. Hanson, V. Cherezov, E. Y. Chien, J. R. Lane, A. P. Ijzerman and R. C. Stevens (2008). "The 2.6 angstrom crystal structure of a human A2A adenosine receptor bound to an antagonist." Science, **322**(5905): 1211-1217.

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

Jardetzky, O. (1966). "Simple allosteric model for membrane pumps." Nature, **211**(5052): 969-970.

Javadpour, M. M., M. Eilers, M. Groesbeek and S. O. Smith (1999). "Helix packing in polytopic membrane proteins: role of glycine in transmembrane helix association." Biophys J, **77**(3): 1609-1618.

Jiang, Y., A. Lee, J. Chen, V. Ruta, M. Cadene, B. T. Chait and R. MacKinnon (2003). "X-ray structure of a voltage-dependent $K^+$ channel." Nature, **423**(6935): 33-41.

Jung, C. (2000). "Insight into protein structure and protein-ligand recognition by Fourier transform infrared spectroscopy." J Mol Recognit, **13**(6): 325-351.

# K

Kaback, H. R. (1971). "Preparation and characterization of bacterial membranes." Methods in Enzymology, **XXII**: 99-120.

Kaback, H. R., I. Smirnova, V. Kasho, Y. Nie and Y. Zhou (2011). "The alternating access transport mechanism in LacY." J Membr Biol, **239**(1-2): 85-93.

Kabsch, W. (1993). "Automatic processing of rotation diffraction data from crystals of initially unknown symmetry and cell constants." J Appl Cryst , **26**: 795-800.

Khare, D., M. L. Oldham, C. Orelle, A. L. Davidson and J. Chen (2009). "Alternating access in maltose transporter mediated by rigid-body rotations." Mol Cell, **33**(4): 528-536.

Kim, S. and B. A. Barry (2001). "Reaction-induced FT-IR spectroscopic studies of biological energy conversion in oxygenic photosynthesis and transport." J Phys Chem, **105**: 4072-4083.

Kim, S., C. A. Sacksteder, K. A. Bixby and B. A. Barry (2001). "A reaction-induced FT-IR study of cyanobacterial photosystem I." Biochemistry, **40**(50): 15384-15395.

Kowalczyk, L., et al. (2011). "Molecular basis of substrate-induced permeation by an amino acid antiporter." Proc Natl Acad Sci U S A, **108**(10): 3935-3940.

Krimm, S. and J. Bandekar (1986). "Vibrational spectroscopy and conformation of peptides, polypeptides, and proteins." Advances in protein chemistry, **38**: 181-364.

Krishnamurthy, H. and E. Gouaux (2012). "X-ray structures of LeuT in substrate-free outward-open and apo inward-open states." Nature, **481**(7382): 469-474.

Krishnamurthy, H., C. L. Piscitelli and E. Gouaux (2009). "Unlocking the molecular secrets of sodium-coupled transporters." Nature, **459**(7245): 347-355.

Kwong, P. D., R. Wyatt, E. Desjardins, J. Robinson, J. S. Culp, B. D. Hellmig, R. W. Sweet, J. Sodroski and W. A. Hendrickson (1999). "Probability analysis of variational crystallization and its application to gp120, the exterior envelope

glycoprotein of type 1 human immunodeficiency virus (HIV-1)." J Biol Chem, **274**(7): 4115-4123.

# L

Lacapere, J. J., E. Pebay-Peyroula, J. M. Neumann and C. Etchebest (2007). "Determining membrane protein structures: still a challenge!" Trends Biochem Sci, **32**(6): 259-270.

Laemmli, U. K. (1970). "Cleavage of structural proteins during the assembly of the head of bacteriophage T4." Nature, **227**(5259): 680-685.

Lander, E. S., et al. (2001). "Initial sequencing and analysis of the human genome." Nature, **409**(6822): 860-921.

Law, R. J., K. Munson, G. Sachs and F. C. Lightstone (2008). "An ion gating mechanism of gastric H,K-ATPase based on molecular dynamics simulations." Biophys J, **95**(6): 2739-2749.

LeMaster, D. M. and F. M. Richards (1985). "1H-15N heteronuclear NMR studies of Escherichia coli thioredoxin in samples isotopically labeled by residue type." Biochemistry, **24**(25): 7263-7268.

Lemieux, M. J., J. Song, M. J. Kim, Y. Huang, A. Villa, M. Auer, X. D. Li and D. N. Wang (2003). "Three-dimensional crystallization of the Escherichia coli glycerol-3-phosphate transporter: a member of the major facilitator superfamily." Protein Sci, **12**(12): 2748-2756.

Leon, X., G. Leblanc and E. Padros (2009). "Alteration of sugar-induced conformational changes of the melibiose permease by mutating Arg141 in loop 4-5." Biophys J, **96**(12): 4877-4886.

Leon, X., R. Lemonnier, G. Leblanc and E. Padros (2006). "Changes in secondary structures and acidic side chains of melibiose permease upon cosubstrates binding." Biophys J, **91**(12): 4440-4449.

Leslie, A. G. W. and H. R. Powell (2007). "Evolving methods for macromolecular." Crystallography, **245**: 41-51.

Lodish, H., A. Berk, S. L. Zipursky, P. Matsudaira, D. Baltimore and J. Darnell, Eds. (2000). Molecular Cell Biology, New York: W. H. Freeman.

Lorenz-Fonfria, V. A., M. Granell, X. Leon, G. Leblanc and E. Padros (2009). "In-plane and out-of-plane infrared difference spectroscopy unravels tilting of helices and structural changes in a membrane protein upon substrate binding." J Am Chem Soc, **131**(42): 15094-15095.

Lorenz-Fonfria, V. A. and E. Padros (2005). "Maximum entropy deconvolution of infrared spectra: use of a novel entropy expression without sign restriction." Appl Spectrosc, **59**(4): 474-486.

Lorenz-Fonfria, V. A., J. Villaverde, V. Trezeguet, G. J. Lauquin, G. Brandolin and E. Padros (2003). "Structural and functional implications of the instability of the ADP/ATP transporter purified from mitochondria as revealed by FTIR spectroscopy." Biophys J, **85**(1): 255-266.

Love, J., et al. (2010). "The New York Consortium on Membrane Protein Structure (NYCOMPS): a high-throughput platform for structural genomics of integral membrane proteins." J Struct Funct Genomics, **11**(3): 191-199.

Lowry, O. H., N. J. Rosebrough, A. L. Farr and R. J. Randall (1951). "Protein measurement with the folin phenol reagent." J Biol Chem, **193**: 265-275.

Lu, M. and D. Fu (2007). "Structure of the zinc transporter YiiP." Science, **317**(5845): 1746-1748.

Ludlam, C. F., I. T. Arkin, X. M. Liu, M. S. Rothman, P. Rath, S. Aimoto, S. O. Smith, D. M. Engelman and K. J. Rothschild (1996). "Fourier transform infrared spectroscopy and site-directed isotope labeling as a probe of local secondary structure in the transmembrane domain of phospholamban." Biophys J, **70**(4): 1728-1736.

# M

MacKenzie, K. R., J. H. Prestegard and D. M. Engelman (1997). "A transmembrane helix dimer: structure and implications." Science, **276**(5309): 131-133.

Maeda, A. (1995). "Application of FTIR spectroscopy to the structural study on the function of bacteriorhodopsin." Israel J Chem, **35**: 387-400.

Maehrel, C., E. Cordat, I. Mus-Veteau and G. Leblanc (1998). "Structural studies of the melibiose permease of Escherichia coli by fluorescence resonance energy transfer. I. Evidence for ion-induced conformational change." J Biol Chem, **273**(50): 33192-33197.

Mantele, W. (1993). "Reaction-induced infrared difference spectroscopy for the study of protein function and reaction mechanisms." Trends Biochem Sci, **18**(6): 197-202.

McDermott, A. (2009). "Structure and dynamics of membrane proteins by magic angle spinning solid-state NMR." Annu Rev Biophys, **38**: 385-403.

McPherson, A. (1990). "Current approaches to macromolecular crystallization." Eur J Biochem, **189**(1): 1-23.

McPherson, A., A. J. Malkin and Y. G. Kuznetsov (1995). "The science of macromolecular crystallization." Structure, **3**(8): 759-768.

Meyer-Lipp, K. (2005). "Time-Resolved Measurements of Sugar-Binding-Induced Conformational Changes in the Melibiose Permease from Escherichia coli." Doctoral thesis.

Meyer-Lipp, K., C. Ganea, T. Pourcher, G. Leblanc and K. Fendler (2004). "Sugar binding induced charge translocation in the melibiose permease from Escherichia coli." Biochemistry, **43**(39): 12606-12613.

Meyer-Lipp, K., N. Sery, C. Ganea, C. Basquin, K. Fendler and G. Leblanc (2006). "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): 25882-25892.

Michel, H. (1983). "Crystallisation of membrane proteins." Trends Biochem Sci, **8**: 56-59.

Mirza, O., L. Guan, G. Verner, S. Iwata and H. R. Kaback (2006). "Structural evidence for induced fit and a mechanism for sugar/H+ symport in LacY." Embo J, **25**(6): 1177-1183.

Misquitta, Y., V. Cherezov, F. Havas, S. Patterson, J. M. Mohan, A. J. Wells, D. J. Hart and M. Caffrey (2004). "Rational design of lipid for membrane protein crystallization." J Struct Biol, **148**(2): 169-175.

Mottamal, M. and T. Lazaridis (2005). "The contribution of C alpha-H...O hydrogen bonds to membrane protein stability depends on the position of the amide." Biochemistry, **44**(5): 1607-1613.

Murata, K., K. Mitsuoka, T. Hirai, T. Walz, P. Agre, J. B. Heymann, A. Engel and Y. Fujiyoshi (2000). "Structural determinants of water permeation through aquaporin-1." Nature, **407**(6804): 599-605.

Mus-Veteau, I., T. Pourcher and G. Leblanc (1995). "Melibiose permease of Escherichia coli: substrate-induced conformational changes monitored by tryptophan fluorescence spectroscopy." Biochemistry, **34**(20): 6775-6783.

## N

Newby, Z. E., et al. (2009). "A general protocol for the crystallization of membrane proteins for X-ray structural investigation." Nat Protoc, **4**(5): 619-637.

Newstead, S., S. Ferrandon and S. Iwata (2008). "Rationalizing alpha-helical membrane protein crystallization." Protein Sci, **17**(3): 466-472.

Nie, Y., N. Ermolova and H. R. Kaback (2007). "Site-directed alkylation of LacY: effect of the proton electrochemical gradient." J Mol Biol, **374**(2): 356-364.

Nie, Y., F. E. Sabetfard and H. R. Kaback (2008). "The Cys154-->Gly mutation in LacY causes constitutive opening of the hydrophilic periplasmic pathway." J Mol Biol, **379**(4): 695-703.

# O

Ostermeier, C., S. Iwata, B. Ludwig and H. Michel (1995). "Fv fragment-mediated crystallization of the membrane protein bacterial cytochrome c oxidase." Nat Struct Biol, **2**(10): 842-846.

Ostermeier, C. and H. Michel (1997). "Crystallization of membrane proteins." Curr Opin Struct Biol, **7**(5): 697-701.

Otwinowski, Z. and W. Minor (1997). "Processing of X-ray data diffraction data collected in oscillation mode." Meth Enzymol, **276**: 307-326.

# P

Patlak, C. S. (1957). "Contributions to the theory of active transport: II. The gate type non-carrier mechanism and generalizations concerning tracer flow, efficiency, and measurement of energy expenditure." Bull Math Biophys, **19**: 209–235.

Poolman, B., J. Knol, C. van der Does, P. J. Henderson, W. J. Liang, G. Leblanc, T. Pourcher and I. Mus-Veteau (1996). "Cation and sugar selectivity determinants in a novel family of transport proteins." Mol microbiol, **19**(5): 911-922.

Pourcher, T., M. Bassilana, H. K. Sarkar, H. R. Kaback and G. Leblanc (1990). "The melibiose/Na+ symporter of Escherichia coli: kinetic and molecular properties." Philos Trans R Soc Lond B Biol Sci, **326**(1236): 411-423.

Pourcher, T., E. Bibi, H. R. Kaback and G. Leblanc (1996). "Membrane topology of the melibiose permease of Escherichia coli studied by melB-phoA fusion analysis." Biochemistry, **35**: 4161-4168.

Pourcher, T., M. Deckert, M. Bassilana and G. Leblanc (1991). "Melibiose permease of Escherichia coli: mutation of aspartic acid 55 in putative helix II abolishes activation of sugar binding by Na+ ions." Biochem Biophys Res Commun, **178**(3): 1176-1181.

Pourcher, T., S. Leclercq, G. Brandolin and G. Leblanc (1995). "Melibiose permease of Escherichia coli: large scale purification and evidence that H+, Na+, and Li+ sugar symport is catalyzed by a single polypeptide." Biochemistry, **34**(13): 4412-4420.

Pourcher, T., M. L. Zani and G. Leblanc (1993). "Mutagenesis of acidic residues in putative membrane-spanning segments of the melibiose permease of Escherichia coli. I. Effect on Na(+)-dependent transport and binding properties." J Biol Chem, **268**(5): 3209-3215.

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

# R

Radestock, S. and L. R. Forrest (2011). "The alternating-access mechanism of MFS transporters arises from inverted-topology repeats." J Mol Biol, **407**(5): 698-715.

Raunser, S. and T. Walz (2009). "Electron crystallography as a technique to study the structure on membrane proteins in a lipidic environment." Annu Rev Biophys, **38**: 89-105.

Rayment, I. (2002). "Small-scale batch crystallization of proteins revisited: an underutilized way to grow large protein crystals." Structure, **10**(2): 147-151.

Rodriguez-Casado, A., M. Molina and P. Carmona (2006). "New accessory for studies of isotopic 1H/2H exchange and biomolecular interactions using transmission infrared spectroscopy." Anal Bioanal Chem, **385**(1): 134-138.

Rothschild, K. J., Y. W. He, S. Sonar, T. Marti and H. G. Khorana (1992). "Vibrational spectroscopy of bacteriorhodopsin mutants. Evidence that Thr-46 and Thr-89 form part of a transient network of hydrogen bonds." J Biol Chem, **267**(3): 1615-1622.

Rummel, G., A. Hardmeyer, C. Widmer, M. L. Chiu, P. Nollert, K. P. Locher, I. I. Pedruzzi, E. M. Landau and J. P. Rosenbusch (1998). "Lipidic Cubic Phases: New Matrices for the Three-Dimensional Crystallization of Membrane Proteins." J Struct Biol, **121**(2): 82-91.

Russi, S., et al. (2011). "Inducing phase changes in crystals of macromolecules: status and perspectives for controlled crystal dehydration." J Struct Biol, **175**(2): 236-243.

# S

Saarinen, P. E., J. K. Kauppinen and J. O. Partanen (1995). "New method for spectral line shape fitting and critique on the voigt line shape model." Applied Spectroscopy, **49**(10): 1438-1453.

Saier, M. H. (1999). "Genome archeology leading to the characterization and classification of transport proteins." Curr Opin Microbiol**2**(5): 555-561.

Saier, M. H., Jr. (2000). "Families of transmembrane sugar transport proteins." Mol microbiol, **35**(4): 699-710.

Saier, M. H., Jr., C. V. Tran and R. D. Barabote (2006). "TCDB: the Transporter Classification Database for membrane transport protein analyses and information." Nucleic Acids Res, **34**(Database issue): D181-186.

Saier, M. H., Jr., M. R. Yen, K. Noto, D. G. Tamang and C. Elkan (2009). "The Transporter Classification Database: recent advances." Nucleic Acids Res, **37**(Database issue): D274-278.

Sanchez-Weatherby, J., et al. (2009). "Improving diffraction by humidity control: a novel device compatible with X-ray beamlines." Acta Crystallogr D Biol Crystallogr, **65**(Pt 12): 1237-1246.

Savage, D. F., P. F. Egea, Y. Robles-Colmenares, J. D. O'Connell, 3rd and R. M. Stroud (2003). "Architecture and selectivity in aquaporins: 2.5 Å X-ray structure of aquaporin Z." PLoS Biol, **1**(3): E72.

Schneider, B., F. Junge, V. A. Shirokov, F. Durst, D. Schwarz, V. Dotsch and F. Bernhard (2010). "Membrane protein expression in cell-free systems." Methods Mol Biol, **601**: 165-186.

Senes, A., I. Ubarretxena-Belandia and D. M. Engelman (2001). "The Calpha ---H...O hydrogen bond: a determinant of stability and specificity in transmembrane helix interactions." Proc Natl Acad Sci U S A, **98**(16): 9056-9061.

Shaffer, P. L., A. Goehring, A. Shankaranarayanan and E. Gouaux (2009). "Structure and mechanism of a Na+-independent amino acid transporter." Science, **325**(5943): 1010-1014.

Shimamura, T., et al. (2010). "Molecular Basis of Alternating Access Membrane Transport by the Sodium-Hydantoin Transporter Mhp1." Science, **328**(5977): 470-473.

Siebert, F. (1995). "Infrared spectroscopy applied to biochemical and biological problems." Methods Enzymol, **246**: 501-526.

Siebert, F., R. Grimm, W. Rudiger, G. Schmidt and H. Scheer (1990). "Infrared spectroscopy of phytochrome and model pigments." Eur J Biochem, **194**(3): 921-928.

Simon-Vazquez, R., T. Lazarova, A. Peralvarez-Marin, J. L. Bourdelande and E. Padros (2009). "Cross-linking of transmembrane helices reveals a rigid-body mechanism in bacteriorhodopsin transport." Angew Chem Int Ed Engl, **48**(45): 8523-8525.

Singh, S. K., C. L. Piscitelli, A. Yamashita and E. Gouaux (2008). "A Competitive Inhibitor Traps LeuT in an Open-to-Out Conformation." Science, **322**(5908): 1655-1661.

Slayton, R. M. and P. A. Anfinrud (1997). "Time-resolved mid-infrared spectroscopy: methods and biological applications." Curr Opin Struct Biol, **7**(5): 717-721.

Smirnova, I., V. Kasho and H. R. Kaback (2011). "Lactose permease and the alternating access mechanism." Biochemistry, **50**(45): 9684-9693.

Smirnova, I. N. and H. R. Kaback (2003). "A mutation in the lactose permease of Escherichia coli that decreases conformational flexibility and increases protein stability." Biochemistry, **42**(10): 3025-3031.

Smirnova, I. N., Kaback, H. R. (2003). "A mutation in the lactose permease of Escherichia coli that decreases conformational flexibility and increases protein stability." Biochemistry, **42**(10): 3025-3031.

Sonoda, Y., et al. (2011). "Benchmarking membrane protein detergent stability for improving throughput of high-resolution X-ray structures." Structure, **19**(1): 17-25.

Studier, F. W., A. H. Rosenberg, J. J. Dunn and J. W. Dubendorff (1990). "Use of T7 RNA polymerase to direct expression of cloned genes." Methods Enzymol, **185**: 60-89.

Surewicz, W. K. and H. H. Mantsch (1988). "New insight into protein secondary structure from resolution-enhanced infrared spectra." Biochim Biophys Acta, **952**(2): 115-130.

Surewicz, W. K., H. H. Mantsch and D. Chapman (1993). "Determination of protein secondary structure by Fourier transform infrared spectroscopy: a critical assessment." Biochemistry, **32**(2): 389-394.

# T

Torres, J., and Padros, E. (1993) "The secondary structure of bacteriorhodopsin in organic solution. A Fourier transform infrared study." FEBS lett, **318**, 77-79.

Toyoshima, C., M. Nakasako, H. Nomura and H. Ogawa (2000). "Crystal structure of the calcium pump of sarcoplasmic reticulum at 2.6 A resolution." Nature, **405**(6787): 647-655.

Tsukihara, T., H. Aoyama, E. Yamashita, T. Tomizaki, H. Yamaguchi, K. Shinzawa-Itoh, R. Nakashima, R. Yaono and S. Yoshikawa (1996). "The whole structure of the 13-subunit oxidized cytochrome c oxidase at 2.8 A." Science, **272**(5265): 1136-1144.

# V

Venyaminov, S. and N. N. Kalnin (1990). "Quantitative IR spectrophotometry of peptide compounds in water (H2O) solutions. I. Spectral parameters of amino acid residue absorption bands." Biopolymers, **30**(13-14): 1243-1257.

Vidaver, G. A. (1966). "Inhibition of parallel flux and augmentation of counter flux shown by transport models not involving a mobile carrier." J Theor Biol, **10**(2): 301-306.

Vogel, R. and F. Siebert (2000). "Vibrational spectroscopy as a tool for probing protein function." Curr Opin Chem Biol, **4**(5): 518-523.

# W

Wallin, E. and G. von Heijne (1998). "Genome-wide analysis of integral membrane proteins from eubacterial, archaean, and eukaryotic organisms." Protein Sci, **7**(4): 1029-1038.

Wang, Y., et al. (2009). "Structure of the formate transporter FocA reveals a pentameric aquaporin-like channel." Nature, **462**(7272): 467-472.

Watanabe, A., S. Choe, V. Chaptal, J. M. Rosenberg, E. M. Wright, M. Grabe and J. Abramson (2010). "The mechanism of sodium and substrate release from the binding pocket of vSGLT." Nature, **468**(7326): 988-991.

Weissborn, A. C., M. C. Botfield, M. Kuroda, T. Tsuchiya and T. H. Wilson (1997). "The construction of a cysteine-less melibiose carrier from E. coli." Biochim Biophys Acta, **1329**(2): 237-244.

Weyand, S., et al. (2008). "Structure and molecular mechanism of a nucleobase-cation-symport-1 family transporter." Science, **322**(5902): 709-713.

Wiener, M. C. (2004). "A pedestrian guide to membrane protein crystallization." Methods, **34**(3): 364-372.

Willis, M. S. and C. M. Koth (2008). "Structural proteomics of membrane proteins: a survey of published techniques and design of a rational high throughput strategy." Methods Mol Biol, **426**: 277-295.

Wilson, D. M., H. Hama and T. H. Wilson (1995). "GLY113-->ASP can restore activity to the ASP51-->SER mutant in the melibiose carrier of Escherichia coli." Biochem Biophys Res Commun, **209**(1): 242-249.

Wilson, D. M. and T. H. Wilson (1992). "Asp-51 and Asp-120 are important for the transport function of the Escherichia coli melibiose carrier." J Bacteriol**174**(9): 3083-3086.

Wilson, D. M. and T. H. Wilson (1994). "Transport properties of Asp-51-->Glu and Asp-120-->Glu mutants of the melibiose carrier of Escherichia coli." Biochim Biophys Acta, **1190**(2): 225-230.

# Y

Yamashita, A., S. K. Singh, T. Kawate, Y. Jin and E. Gouaux (2005). "Crystal structure of a bacterial homologue of Na+/Cl--dependent neurotransmitter transporters." Nature, **437**(7056): 215-223.

Yazyu, H., S. Shiota-Niiya, T. Shimamoto, H. Kanazawa, M. Futai and T. Tsuchiya (1984). "Nucleotide sequence of the melB gene and characteristics of deduced amino acid sequence of the melibiose carrier in Escherichia coli." J Biol Chem, **259**(7): 4320-4326.

Yernool, D., O. Boudker, Y. Jin and E. Gouaux (2004). "Structure of a glutamate transporter homologue from Pyrococcus horikoshii." Nature, **431**(7010): 811-818.

Yin, Y., X. He, P. Szewczyk, T. Nguyen and G. Chang (2006). "Structure of the multidrug transporter EmrD from Escherichia coli." Science, **312**(5774): 741-744.

Yohannan, S., S. Faham, D. Yang, D. Grosfeld, A. K. Chamberlain and J. U. Bowie (2004). "A C alpha-H...O hydrogen bond in a membrane protein is not stabilizing." J Am Chem Soc, **126**(8): 2284-2285.

Yousef, M. S. and L. Guan (2009). "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): 15291-15296.

# Z

Zani, M. L., T. Pourcher and G. Leblanc (1993). "Mutagenesis of acidic residues in putative membrane-spanning segments of the melibiose permease of Escherichia coli. II. Effect on cationic selectivity and coupling properties." J Biol Chem, **268**(5): 3216-3221.

Zani, M. L., T. Pourcher and G. Leblanc (1994). "Mutation of polar and charged residues in the hydrophobic NH2-terminal domains of the melibiose permease of Escherichia coli." J Biol Chem, **269**(40): 24883-24889.

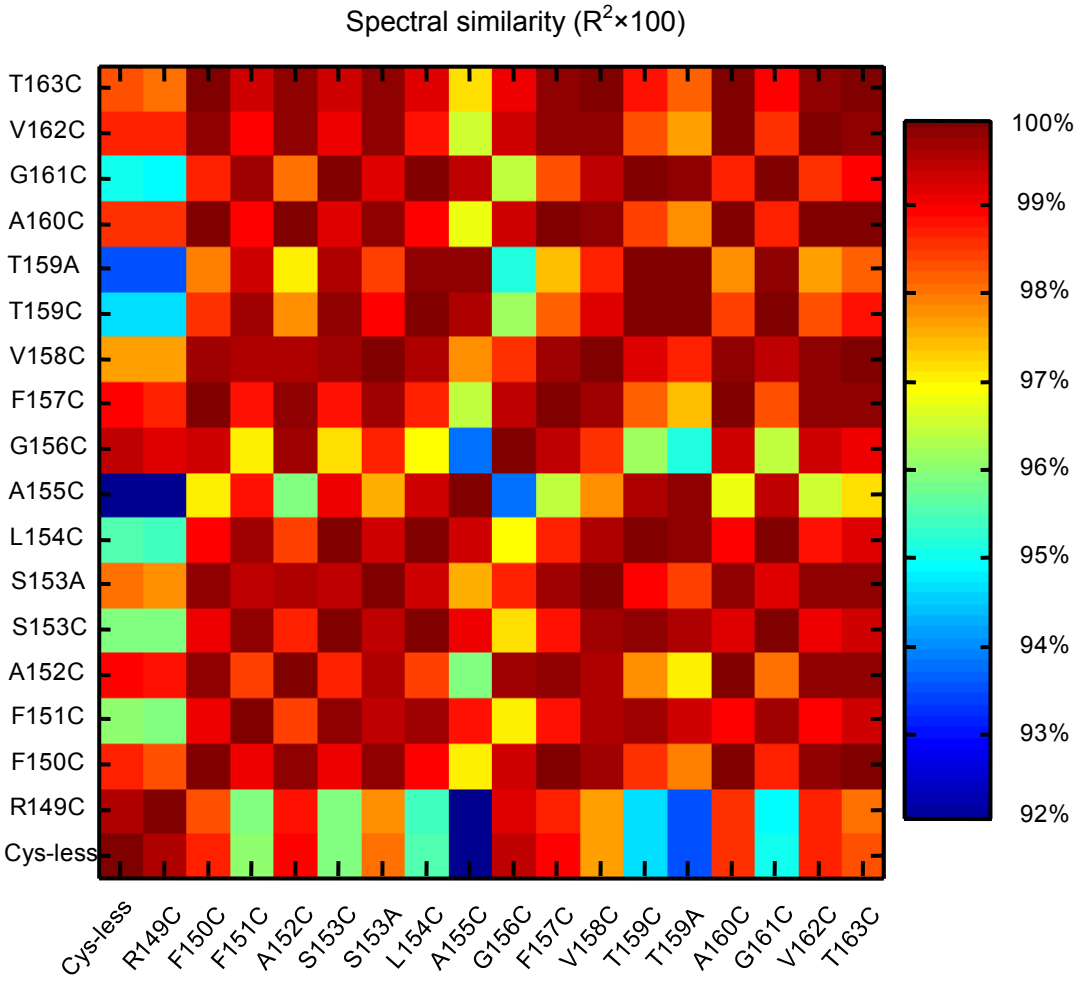
Zheng, L., U. Baumann and J. L. Reymond (2004). "An efficient one-step site-directed and site-saturation mutagenesis protocol." Nucleic Acids Res, **32**(14): e115.

Zhou, Y., L. Guan, J. A. Freites and H. R. Kaback (2008). "Opening and closing of the periplasmic gate in lactose permease." Proc Natl Acad Sci U S A, **105**(10): 3774-3778.

Zhou, Z., J. Zhen, N. K. Karpowich, R. M. Goetz, C. J. Law, M. E. Reith and D. N. Wang (2007). "LeuT-desipramine structure reveals how antidepressants block neurotransmitter reuptake." Science, **317**(5843): 1390-1393.

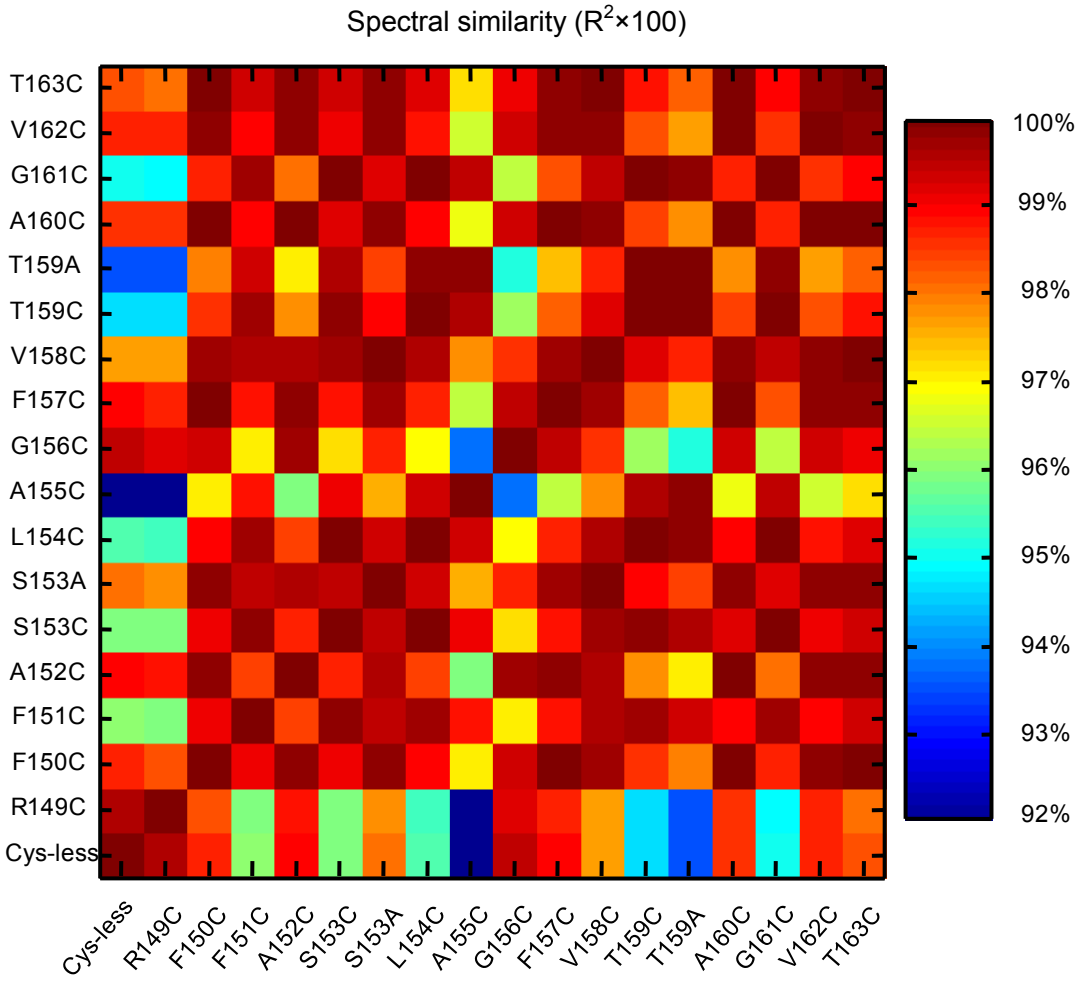
Zulauf, M. (1991). "Detergent phenomena in membrane protein crystallization." In Crystallization of membrane proteins (ed. H. Michel): pp. 53–72. CRC Press, Boca Raton, FL.

# Appendix I



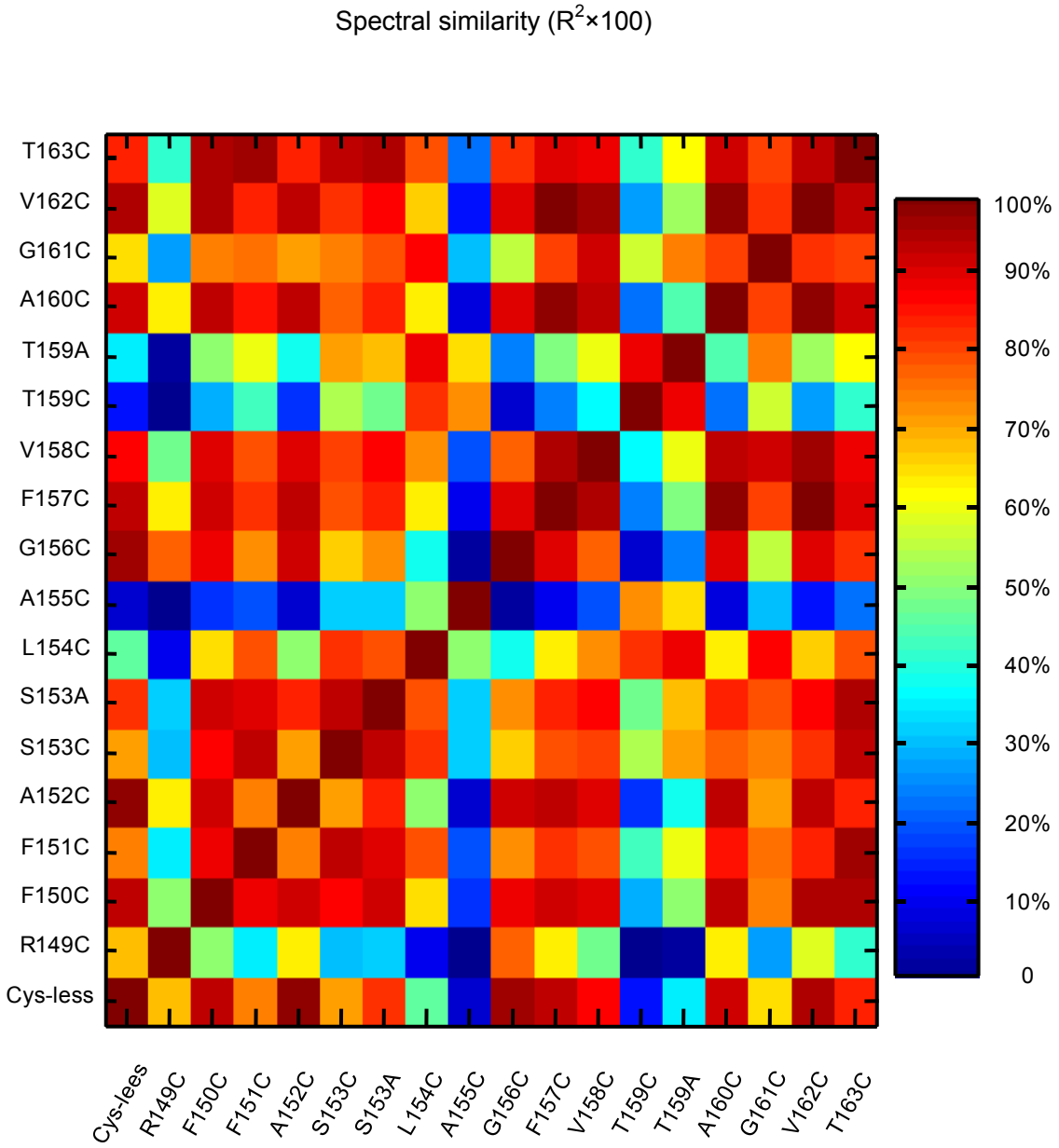

Structural comparison of MelB mutants using IR spectroscopy of hydrated samples (shown in Figure 5.1). Representation of all possible comparisons between pair of protein, with a color code indicating their degree of similarity, measured as $R^2$ × 100. For detail, see ***Materials and Methods***.

# Appendix II



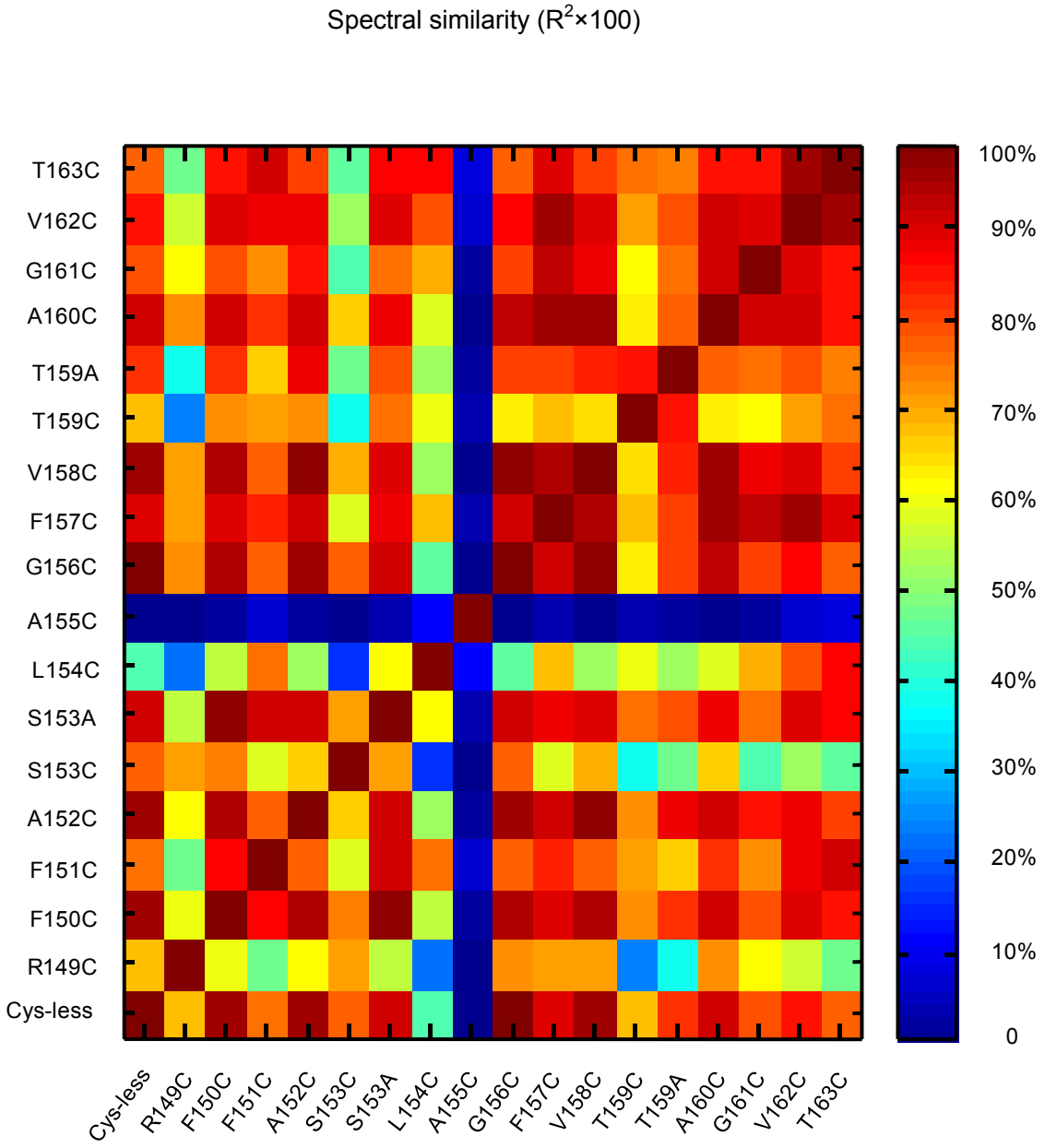

Representation of all possible comparisons between pair of protein, with a color code indicating their degree of similarity in response to $Na^+$ binding, measured as $R^2$ × 100.This figure was obtained by using data presented in Figure 5.2. For detail, see ***Materials and Methods***.

# Appendix III




Representation of all possible comparisons between pair of protein, with a color code indicating their degree of similarity in response to melibiose binding in the absence of $Na^+$, measured as $R^2$ × 100.This figure was obtained by calculating data presented in Figure 5.4. For detail, see ***Materials and Methods***.

# Appendix IV




Representation of all possible comparisons between pair of protein, with a color code indicating their degree of similarity in response to melibiose binding in the absence of $Na^+$, measured as $R^2$ × 100.This figure was obtained by calculating data presented in Figure 5.6. For detail, see ***Materials and Methods***.

Many thanks to all the kind persons around me. Without your help and support, it would not possible to finish this doctoral thesis.

First of all, I would like to thank my wife **Aiping Qian** for standing beside me throughout my four years doctoral study, and **my parents**, **my sister** for their encouragement. And **my daughter**, well she has just been born several days ago.

I would like to express my sincere gratitude to my advisor, Prof. **Esteve Padrós**, for continuous support for my doctoral study and research, for his patience, motivation, enthusiasm, and immense knowledge. Without his guidance it would be impossible to finish this thesis.

Besides my advisor, I would like to thank all my thesis committee members, Prof. **Pere Garriga**, Dr. **José Lusis**, Dr. **Natalia Dave**, for their willingness to serve as my thesis committee. Your encouragement, insightful comments, and hard questions will help me to improve my work in the future.

I would like to thanks Dr. **Víctor Lórenz** and Dr. **Meritxell Granell** to help me to start my Ph.D work, and for their kind discussion.

I would like to thank Prof. **Gérard Leblanc** for his valuable collaboration.

Many thanks for Dr. **Jordi Benach** and Dr. **David Reverter** for their valuable collaboration and kind discussion for melibiose permease crystallization.

I thank especially Dr. **Antoni Morros Carulla**, Prof. Dra. **Maria Josefa SABRIÁ**, Dr. **Josep Cladera Cerdà**, Dr. **Xavier León** for encouragement, insightful comments, and guidance of the thesis preparation.

I would like to thank **Elodia**, **Mateu**, and **Neus** for their kind help and their serious and responsible work.

I would like to thank my classmate, **Oliver**, for his encouragement and kind help in these four years.

At the end of my thesis I would like to thank all the people in Centre d'Estudis en

Biofísica, **Tzvetana**, **Joan**, **Mireia**, **David**, **Manuel**, **Nuria**, **Beatriz**, **Montserrat**, **Arash**, **Silvia**, **Oksana**, **Ero**, **Asrar**, **Guillem**, **Mikhail**, **Glòria**, **Roger**, **Gabriela**, **Meritxell**, **Fanli** … for their friendly help and kind support during these four years study.