SODIUM CLORIDE AND SUCROSE RESCUED *Cupriavidus necator* H16 Δ*secB* MUTANT WHEN GROWN IN NUTRIENT RICH MEDIA

SIM XUAN YI

UNIVERSITI SAINS MALAYSIA

MARCH 2016

SODIUM CLORIDE AND SUCROSE RESCUED *Cupriavidus necator* H16 Δ*secB* MUTANT WHEN GROWN IN NUTRIENT RICH MEDIA

By

SIM XUAN YI

Thesis submitted in fulfilment of the requirements for the degree of

Master of Science

MARCH 2016

ACKNOWLEDGEMENT

I would like to extend my deepest appreciation to my supervisor, Prof. Razip Samian for his continuous guidance, advice, encouragement, and support which largely contribute to the completion of my master degree. I also want to thank Prof. Nazalan Najimudin who never fails to give me his advice and support whenever needed.

I would also like to thank my seniors, Dr. Yam Hok Chai, Ooi Gim Luan, Adrian Chek, Tan Yi Fen, Cleo, Kak Aini, Kak Kem, Kak Balqis, Kak Hani, Abg Syafiq, Kak Zaf, Abg Adi, Abg Asraff who have bend over backwards to help me, whenever I need it. Besides, my thanks to my other lab members who have helped me in one way or another, in the completion of this project. To all the members of Lab 414, your support, love and care make lab work a joy, you guys make Lab 414 like a second home.

Not forgetting to thank my family and friends for their endless support. Thank you so much for being there for me, be it virtually or physically.

Most important of all, I thank God for His Grace and Strength in enabling me to complete this thesis.

TABLE OF CONTENTS

CHAPTER 3: MATERIALS AND METHODS

LIST OF FIGURES

LIST OF TABLES

LIST OF SYMBOLS AND ABBREVIATIONS

°C	Celcius
%	percent
µg	microgram
µl	microlitre
µm	micrometer
Amp^r	ampicillin resistance
ATP	adenosine-5'-triphosphate
bp	base pair
cat^R	chloramphenicol resistance gene
Cm^r	chloramphenicol resistance phenotype
CoA	Coenzyme A
dH_2O	distilled water
kb	kilo base pairs
NR	Nutrient rich medium
LA	Luria Bertani agar
LB	Luria Bertani broth
IPTG	isopropyl β-D-1-thiogalactopyranoside
ORF	Open reading frame
sacB	levansucrase gene
mob	mobilization gene
X-gal	5-bromo-4-chloro-indolyl-β -D-galactopyranoside
w/v	weight per volume
OD	Optical density
rpm	Revolutions per minute
v/v	Volume per volume

NATRIUM KLORIDA DAN SUKROSA MENYELAMATKAN MUTAN *Cupriavidus. necator* H16 Δ*secB* YANG DITUMBUHKAN DALAM MEDIA NUTRIEN KAYA

ABSTRAK

SecB adalah salah satu komponen utama dalam sistem translokasi protein Sec terutamanya bagi bakteria Gram-negatif. Dalam kajian ini, fungsi SecB telah dikaji dengan membina mutan *Cupriavidus necator* H16 Δ*secB.* Mutan ini telah dibina melalui pertukaran alel menggunakan vektor swa-hapus pDM4. Untuk mengesahkan perubahan fenotip disebabkan oleh gen yang dimutasi, mutan pelengkap juga dibina, yang mana vektor pBBR1MCS2 yang mengandungi gen SecB telah dikonjugasi ke dalam mutan *C. necator* H16 Δ*secB.* Jenis liar, *C. necator* H16 ΔsecB mutan dan mutan pelengkap telah dikaji mengguna sistem Biolog Phenotype Microarray. Mutan pelengkap menunjukkan fenotip jenis liar, bermakna penglengkapan *secB* telah berjaya. Mutan *Cupriavidus necator* H16 Δ*secB* tidak menunjukkan perbezaan yang signifikan apabila ditumbuh dialam kebanyakan substrat. Hanya satu pengecualian iaitu mutan *Cupriavidus necator* H16 Δ*secB* tidak dapat tumbuh dalam medium nutrien kaya dan medium Bloomfield kecuali jika sukrosa (10%) atau NaCl (0.5-1.0 %) ditambah. Walau bagaimanapun, penambahan KCl tidak dapat mengatasi kecacatan ini. Fenotip daripada mutan yang tidak dapat tumbuh pada medium kaya bukan disebabkan oleh “kesan polar” seperti yang pernah dicadangkan kerana tidak semacam *E.coli*, kodon penamat *secB* dan *gpsA* daripada *C. necator* .tidak bertindih . Kajian ini juga menunjukkan bahawa mutasi pada *secB* tidak mempengaruhi pengeluaran PHB dalam *C. necator.* Kajian dalam *Pesudomonas* sp. menunjukkan peningkatan pengeluaran PHA sebanyak 50% apabila *secB* dihapuskan.

SODIUM CLORIDE AND SUCROSE RESCUED *Cupriavidus necator* H16 Δ*secB* MUTANT WHEN GROWN IN NUTRIENT RICH MEDIA

ABSTRACT

SecB is one of the key players in the Sec protein-translocation system ubiquitous to Gram-negative bacteria. In this study the function of SecB was studied by studying an unmarked *Cupriavidus necator* H16 Δ*secB* mutant. The mutant was constructed through allelic exchange utilizing the pDM4 suicidal vector. In order to confirm that the phenotypic change is due to the mutated gene, a complementation mutant was also constructed, in which vector pBBR1MCS2 harbouring a *secB* gene was introduced into the *C. necator* H16 Δ*secB* mutant via conjugation. The wild type, *C. necator* H16 Δ*secB* mutant and the complementation mutant were characterised using Biolog Phenotype Microarray. The result shows that the complementation mutant exhibits the wild type phenotype, indicating successful complementation of the unmarked *secB* deletion. The *C. necator* H16 Δ*secB* mutant did not show any significant difference when grown on most substrate. The only exception was that the *C. necator* H16 Δ*secB* mutant cannot grow on nutrient rich and Bloomfield medium unless supplemented with either sucrose (10%) or NaCl (0.5 – 1.0%). The result also showed that KCl was unable to rescue growth, unlike sucrose and NaCl. The inability of the mutant to grow in rich media (without sucrose or NaCl supplementation) is unlikely due to polar effect of *secB* deletion in *gspA* as previously proposed, because unlike *E. coli*, the stop codon of *secB* and *gps*A in *C. necator* do not overlap. The result also showed that the deletion of *secB* did not affect the PHB production of *C. necator*. In *Pseudomonas* sp. deletion of *secB* led to a 50% increase in PHA production.

1.0 INTRODCTION

All nascent polypeptides in bacterial cells are produced in cytosolic compartments and later sorted to a specific subcellular location according to their specific biological duties. Transportation of non-cytosolic primary proteins is mediated by a group of translocases that is vital to living cells. A typical translocation system/machinery comprises a chaperone that maintains the nascent polypeptide in unfolded stage and a protein-conducting channel to translocate the competent polypeptide across lipid bilayer membrane. Sec secretion machinery represents the general translocase in Gram-negative bacteria that is responsible for transporting a wide range of periplasmic located proteins. Sec translocase consists of a protein-conducting channel formed by SecYEG, a chaperone SecB for maintaining the target protein in translocation competent stage and an ATPase protein SecA for targeting the SecB-bounded polypeptide to the channel.

Using *E. coli* as a model, the *secB* gene was discovered by Kumamoto and Beckwith (1983) in an attempt to isolate mutants that were defective in protein secretion. The null mutant of *secB* can only grow on minimal medium but not rich medium. However, none of the other sec mutants was found to have this phenotype despite working in concert to form a functional Sec-dependent protein translocation machinery. This observation leads to a speculation that *secB* mutant may cause a polar effect on the immediate downstream gene *gpsA*. This issue had been addressed by Shimizu and they found that the growth cessation of *secB* mutant on rich medium can be rescued by introducing a foreign copy of *gpsA* but not *secB* itself (Shimizu et al., 1997). Conceivably, the growth-defective phenotype of *secB* mutant was due to the polar effect that caused the lost-of-function of *gpsA* since the stop codon of *secB* ORF overlaps with the start codon of gpsA ORF (Sofia et al., 1994). It was later

found that missense mutation on *secB* were able to grow on LB medium with or without the addition of foreign *gpsA* gene (Ullers et al., 2007). The discrepancies of the resultant phenotypes of *secB* mutation may be partly due to different mutagenesis approaches and variation in bacterial strains used as study models. The physiological purpose of *secB* in bacterial cells is thus perplexing.

It was reported that mutation on *secB* had increased the production of polyhydroxyalkanoate (PHA) in *Pseudomonas* sp. USM 4-55(Hassan, 2013). Hassan's work had showed that when the transposon jumped into the *secB* region, PHA production of the mutant showed an increase of 50%. Therefore, in this study, the function of sec gene is investigated in a different Gram-negative bacterium namely *Cupravidus necator. C. necator* is a soil-living bacterium under the class of Betaproteobacteria; not only it is able to oxidize hydrogen, live under aerobic or anaerobic condition, but also capable of utilising organic compounds and hydrogen as energy sources. It is the model organism for polyhdroxyalkanoate (PHA) biosynthesis studies as it can produce very high PHA mass over cell dry weight under growth condition of limited nitrogen and excess carbon source. This trait makes it a very important industrial strain, especially in bioplastic production industry as bioplastic is biodegradable and able to be used as raw materials in many applications.

Based on the complete genome sequence, the *secB* ORF in *C. necator* does not overlap with the downstream *gpsA* ORF in which mutation on *secB* will not cause polar effect. Moreover, this strain is easy to be genetically manipulated which makes it a suitable non-*E. coli* strain for *secB* functional elucidation.

The objective of this study is to functionally characterise the *secB* gene in *C. necator* by scrutinizing the phenotypic differences between an unmarked *secB*

deletion mutant and the wild type strain and also the effect of *secB* gene in PHA production. The *secB* gene was removed from the wild-type genome using homologous recombination without insertion of antibiotic gene. The resultant phenotypes of the mutant are also cross compared with a *secB*-complemented strain. The results presented herein strengthen the knowledge of the *secB* gene in *C. necator*.

2.0 LITERATURE REVIEW

2.1 Protein transportation in Gram-negative bacteria

Thousands of different proteins produced by a bacterial cell are teither transported out or localized to different cellular compartments of a cell to carry out specific duties. Up to 50% of the proteins which are synthesized in the cytosol at the ribosomes do not stay there, only the inner membrane is directly accessible to the newly synthesised protein (Dalbey and Kuhn, 2012). They are either inserted into a membrane (insertion process) or transported across the membrane to the outer layer (periplasm or the extracellular milieu). Transporter proteins, energy conversion related channels, sensory transducers and etc. are usually embedded on the membrane while proteins like hydrolytic enzymes, toxins, small molecule scavengers etc are normally secreted out of cytosol. Secretion typically is the process of a protein from the cell to the extracellular milieu. Secretion studies in bacteria were done by observing the protein from the ribosome to membrane targeting to crossing over or integration into the membrane (translocation). Secretion of protein components is often essential for the pathogenicity of the bacteria. Not only that, other functions such as biofilm formation, modulation of host, energy conversion and nutrient uptake are closely related to protein secretion (Kusters and Driessen, 2011). The proteins can either be transported directly from cytoplasm out of the cell by one step or two steps. The protein is first exported to the periplasm before entering the outer membrane. There are many challenges for the proteins to undergo this translocation process. Bacterial cells have evolved many different secretion systems to handle the secretion of different types of proteins and to various destinations (Papanikou et al., 2007). Bacteria usually have two or more secretion systems.

The surface of Gram-negative bacteria contains two membranes, inner (cytoplasmic) and outer membranes. There is a periplasm between the two membranes which gives cell its shape and rigidity. There are many integral proteins that are needed to be inserted into the inner membrane. The main pathway will be through the Sec pathway (Xie and Dalbey, 2008). The Sec translocase facilitates the proteins across the membrane and insert their apolar segments into the membrane with the help of YidC insertase. YidC plays a role in insertion of both sec-dependent and sec-independent membrane protein. YidC can work alone as true membrane insertase in a separate pathway (Samuelson et al., 2000, Serek et al., 2004).

For the transportation of protein into the membrane or the periplasmic space or outer membrane in the Gram-negative bacteria (*E.coli*), two systems have been identified; the twin-arginine translocation (TAT) system and the general secretory (Sec) system. The major difference between these two systems is that the Tat system deals with the proteins that are completely folded in the cytosol and a cofactoris needed to be present, while the Sec system deals with the preproteins that are not folded prior to the transportation to happen. All preproteins that undergo the sec system, must be transported before they acquire a stable folded structure (Randall and Hardy, 2002).

After the proteins are translocated across the inner membrane via Sec system or TAT system, two pathways are involved in sorting the protein to the outer membrane which is the Lol pathway and β-barrel assembly machinery (BAM) pathway. Lol pathway is essential for the transport of lipoproteins from the inner membrane to the outer membrane. While BAM pathway is able to recruit the unfolded β-barrel protein in the periplasmic space and transport them to the outer membrane.

Not only that, there are several bacterial proteins that need to be exported across the outer membrane by different kinds of secretion systems which produce surface structures like type I pili, curli and type IV pili. The subunit proteins for all the piliare transported via sec translocase into the periplasm. After that, it is assembled into the growing pilus. The pili are critical in adhering on to the surface and the host cells.

2.1.1 Sec System

Of all the bacterial secretion systems, the general secretion system (Sec system) mediates the exports of most secretory proteins and some membrane proteins. The sec system plays a prominent role in both insertions of protein into the inner membrane and also secretion of proteins to the periplasm. The membrane-embedded protein Sec channel (SecYEG) is conserved in archaea, bacteria and eukaryotes (Berg et al., 2004). It is also by far the most understood pathway. The Sec system also shows a good general view of challenges that needed to be tackled by the bacteria secretion pathway which is shown in Figure 2.1 The challenges include sorting process whereby how the proteins are separated out from the cytoplasmic resident, targeting how the protein find its way to the right location, why and how the protein remain unfolded, translocation - where and how the protein cross the membrane, and energetics, what is the cost of the secretion process (Chatzi et al., 2013).

A secretory protein is synthesized as a preprotein with a signal sequence at the N-terminal that will be removed in the post-translocation step. Signal peptide is the tag to differentiate the secretory protein and to direct the protein to the final destination (ribosome to membrane). Once the preprotein is bound at the membrane,

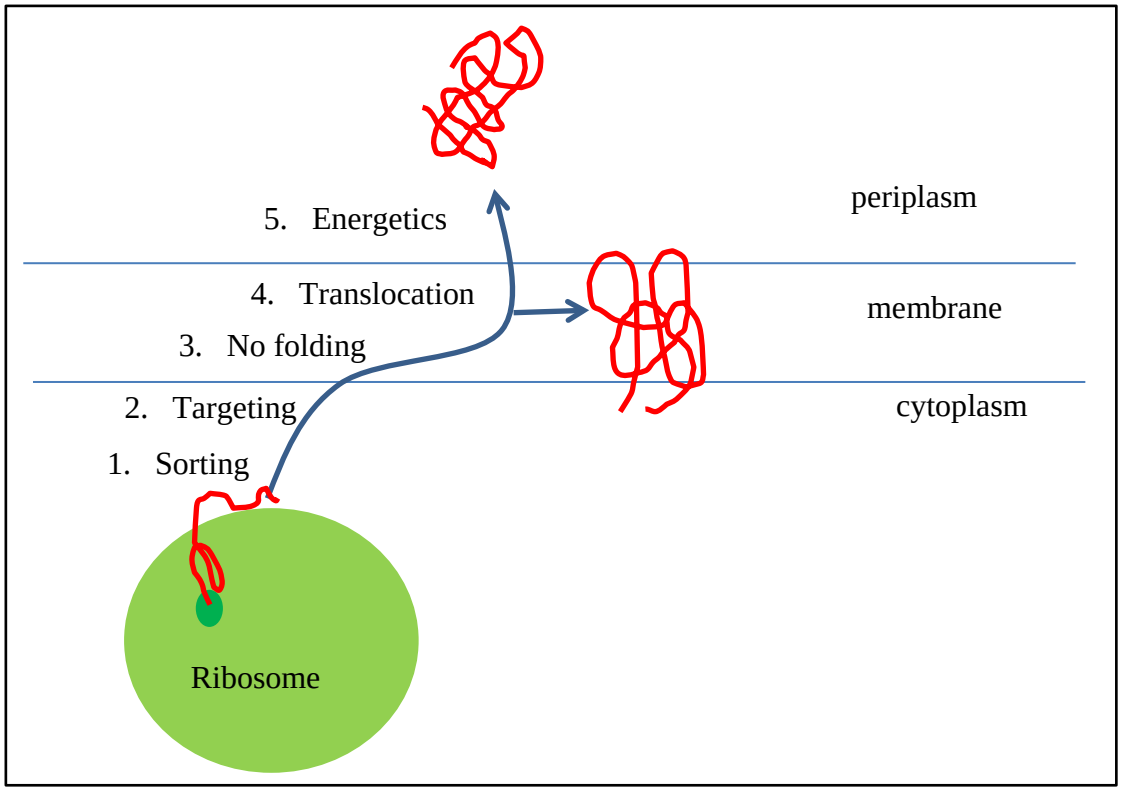


Figure 2.1: Challenges faced by the protein before it can be secreted to perform their duty (Adapted from Chatzi et al., 2013) .

1. How are proteins determined to be sent out or remain in cytoplasm (sorting)
2. How do the proteins find the membrane and the substrates? (targeting)
3. How to prevent the proteins from folding? (No folding)
4. How do the proteins cross through the membrane? (translocation)
5. What are the costs or energy needed? (energetics)

the signal peptide also acts as the allosteric key to activate the Sec channel. Preprotein that undergo sec pathway should remain unfolded. Almost all the unfolded preproteins remain in a non-native state through the help of specialized chaperones. Sec system will either incorporate the preprotein into the plasma membrane or transport across it. Metabolic energy input is needed for all the transportation. Besides, preprotein translocation can happen in two modes: co-translation and post-translation. SecYEG channel is utilized in both.Inner membrane proteins are normally co-translationally targeted. However in bacteria, additional post-translational mode is usually followed. . Figure 2.2 shows the overall sec secretion system in bacteria (Chatzi et al., 2013).

2.1.2 Sec system co-translation mode

Membrane protein insertion undergoes co-translation mode. Signal recognition particle (SRP) is made up of Ffh protein together with the short 4.5S RNA in *E. coli* co-translation system (Grudnik et al., 2009). SRP docks on the L23 ribosomal subunits and binds the nascent preprotein chain as it emerges from the ribosomes. It targets the nascent preprotein chain to a membrane receptor, FtsY. FtsY is attached to SecY. When GTP is hydrolysed, the dissociation of the SRP-Fts complex happens, the protein is delivered to the SecY channel and the protein is inserted into the membrane (Luirink et al., 1994). This SRP pathway is very important for the bacteria (Phillips and Silhavy, 1992). For some membrane proteins, YidC protein is involved in the translocation (Samuelson et al., 2000).

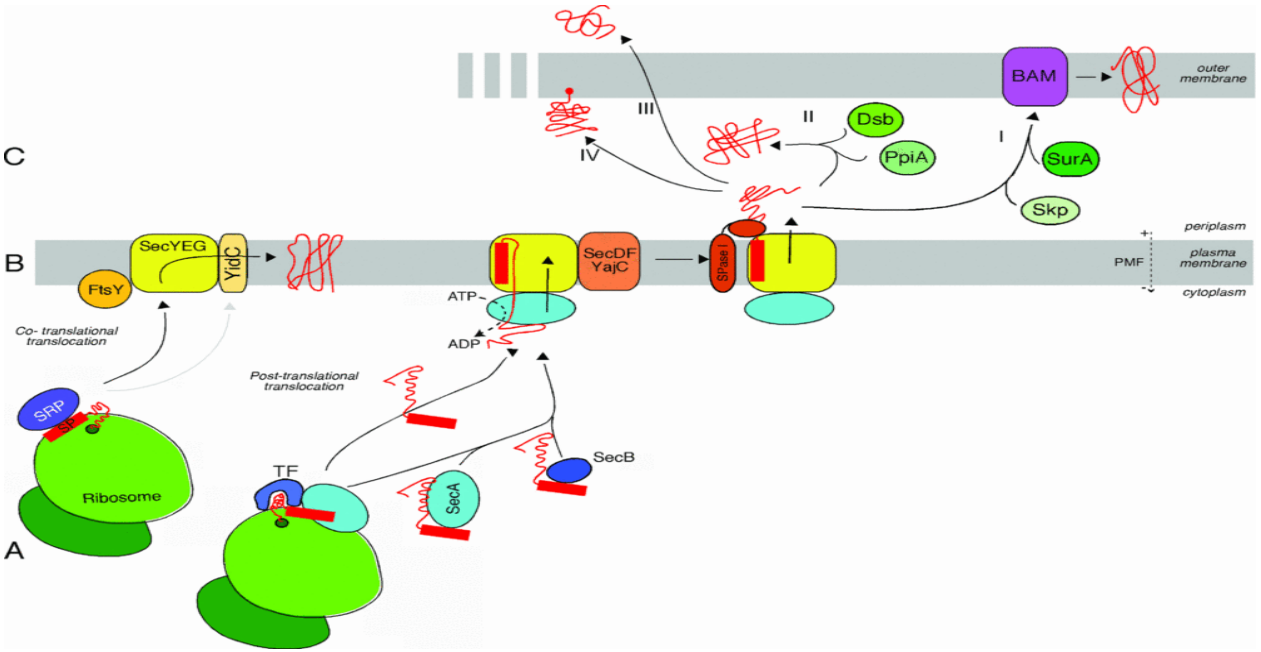


Figure 2.2: Schematic view of the overall sec secretion system in bacteria (Chatzi et al., 2013).

(**A**) For membrane-embedded proteins, the nascent polypeptide chain emerging from the ribosome is bound by SRP (co-translational translocation). The complex is transferred to SecYEG assist by FtsY which is the SRP receptor. Some proteins may directly use the YidC route (greyed arrow). Secreted proteins will be scanned by TF and through a process that might involve chaperones, such as SecB, will be delivered to soluble or SecYEG-bound SecA (post-translational translocation). (**B**) Once the complex reaches SecYEG, the preprotein, as it is being synthesized, enters the membrane through the SecYEG lateral gate, in some cases with the help of YidC. Following preprotein docking on SecA bound to SecYEG. SecDFYajC might help in preprotein release. SPase I (or II for lipoproteins, not shown) cleaves the signal peptide, hence releasing the mature domain to the periplasm. (**C**) Proteins newly released in the periplasm can interact with chaperones (Skp and SurA) and get transferred to machineries such as the BAM for insertion into the outer membrane (I), fold spontaneously or with the help of periplasmic chaperones [e.g. Dsb proteins and PPiase (peptidylprolyl isomerase)] (II), get secreted using other specialized systems (III), or become membrane-anchored by specialized lipoprotein sorting systems (IV) (Chatzi et al., 2013).

2.1.3 Sec system post-translation mode

Secretory protein normally follows the post-translation secretion pathway. This mode can be put into three key stages:

(i) **The preprotein is targeted to the translocase (SecYEG-SecA)**

Preproteins are prevented from folding activity by the chaperones that recognized them either during their synthesis or after they are released from the ribosome before targeting them to the correct membrane location. The chaperones and the factors that recognize and bind the newly synthesized proteins include SRP, SecA, trigger factor (TF) and SecB. TF, SRP and SecA bind directly to the L23 ribosomal subunit but SecB are not known to bind directly with the ribosome (Ullers et al., 2003). SecB attaches to preproteins that are larger than 150 amino acids and lead them to SecA (Randall et al., 1997). In some bacteria, the preproteins are able to maintain their non-native form even without any chaperones (Gouridis et al., 2009).

(ii) Assembly of SecY-SecA-preprotein complex and the preprotein translocation

The preprotein is translocated through the SecYEG channel when the ternary complex of SecYEG-SecA-preprotein is obtained. This translocation process needs ATP and proton motive force. Although SecDFYajC is not needed in translocation, it has been involved in several steps of the reaction. Therefore, SecYEG can work with this membrane protein complex (Duong and Wickner, 1997).

(iii) Release of preprotein and periplasmic folding or further trafficking

Spase I will cleave the signal peptide to release the protein into the periplasm after it is successfully translocated. The protein will fold with the help of periplasmic chaperones (Skp, prolyl isomerases etc) and cysteine oxidase (Dsb protein) (Allen et al., 2009). The fate of the protein will either be further trafficked to the outer membrane or even out of the cell through other secretion system.

2.1.4 Model of secretory protein translocation

The secretory protein translocation can be separated into six steps.. SecA attaches itself to SecYEG channel forming a holoenzyme, a biochemically active compound formed by the combination of an enzyme with a coenzyme (step I). Chaperones bound or free preprotein are transferred to the translocase, SecA-SecYEG. The signal peptide binds on the specific site on SecA which is bond with SecY (step II). When the signal peptide is bound, it will lower the activation energy of the translocase. It will become “triggered” (step III). The mature domains irreversibly trapped within the channels (requires the physical presence of the signal peptides) after the translocase is triggered (step IV) (Gouridis et al., 2009) A structural conserved salt bridge at the base of ATPase motor breaks, the preprotein translocation begins in defined segments (20-30 residues) driven by the binding energy of ATP (Nouwen et al., 1996, Schiebel et al., 1991) . This cycle is repeated a few times for the translocation to be completed. The preprotein will be released to the other side of the membrane. When the chain is not engaged by SecA, proton motive force will drive the chain to move forward and also prevent the chain from back-slipping (step V) (Arkowitz et al., 1993, Yamada et al., 1989). The preprotein is released once it is

cleaved by Spase I into the periplasm to be folded or to continue further trafficking (step VI) (Chatzi et al., 2013).

2.2 SecB as a chaperone

SecB is a molecular chaperone in Gram-negative bacteria which is dedicated to the translocation of proteins across the cytoplasmic membrane by bacteria utilizing the Sec-system. A chaperone in modern molecular biology is defined as a protein which acts as a supporter in non-covalent folding and assembly of proteins and proteins complexes. It is needed to prevent protein aggregation or denaturation into a non-functional structure (Bechtluft et al., 2010).

2.2.1 History of SecB Studies

In the 1980s, mechanism of protein secretion and translocation machinery had been studied widely in eukaryotic and prokaryotic system through genetic approaches. Sec system is one of the machinery that has been widely looked into. Therefore, the genes that constitute this system are essential for cell viability. Most Sec mutants were identified by screening for conditional lethal mutations. The first reference of (*secB*) was in 1983 by Kumamoto and Beckwith who found that *E. coli* with a mutated gene (*secB*), eliminated the export of a subset of periplasmic proteins (Kumamoto and Beckwith, 1983). *E. coli* cells that had a disrupted *secB* gene were unable to grow on rich media, Luria broth media. However, this phenotype was not seen on minimal media. Thus, SecB is not an essential protein for cell viability. In *secB* mutants, the preMBP which was the signal sequence of maltose binding protein was found to accumulate in the cytosol and to become protease resistant state. The retardation or blockage of the preMBP observed in pulse-chase experiment suggested that SecB plays a role in stabilizing MBP in a translocation competent

state (Gannon and Kumamoto, 1993). The first indication of SecB anti-folding activity was shown in 1988 where the folding rate of *in vitro* synthesis preMBP was slowed down by SecB (Collier et al., 1988). Phil Bassford's laboratory was the first to purify SecB. Using this purified SecB through an *in vitro* translocation assay, SecB was shown to be involved in the stimulation of preMBP translocation into inside-out membrane vesicle of *E. coli* (Weiss and Bassford, 1990, Weiss et al., 1988*)*.

In the early nineties, Hardy and Randall suggested the "Kinetic partitioning model" for SecB function, the rate of the folding will differentiate the secretory and cytosolic protein, using the purified SecB, preMBP and mature MBP. This model suggested that secretory protein folds slower and will buy some time for the secretory to bind with SecB partially due to the presence of signal sequence (Hardy and Randall, 1991). However, at the same period, Topping and Randall who determined the binding frame of SecB disagree that the signal sequence is the binding partner (Topping and Randall, 1994). During the early stage of protein synthesis, SecB that are already present, are readily associate with the ribosome bound nascent secretory protein (Randall et al., 1997).

The *secB* gene does not only function solely as an anti-folding agent but also as a preprotein targeting factor. It binds directly to the C-domain of a motor domain of Sec-translocase, SecA. The C-domain consists of a zinc ion which is essential for the binding to happen. In the year 2000, Xu and coworkers elucidated the three dimensional structure of SecB, although it is in complex with a peptide. Recent proteomics studies show that the preprotein substrates are dependent of SecB. The preprotein substrates are discussed in section 2.2.3. The interaction sites of SecB with the preprotein were identified using the site directed spin-labelling techniques.

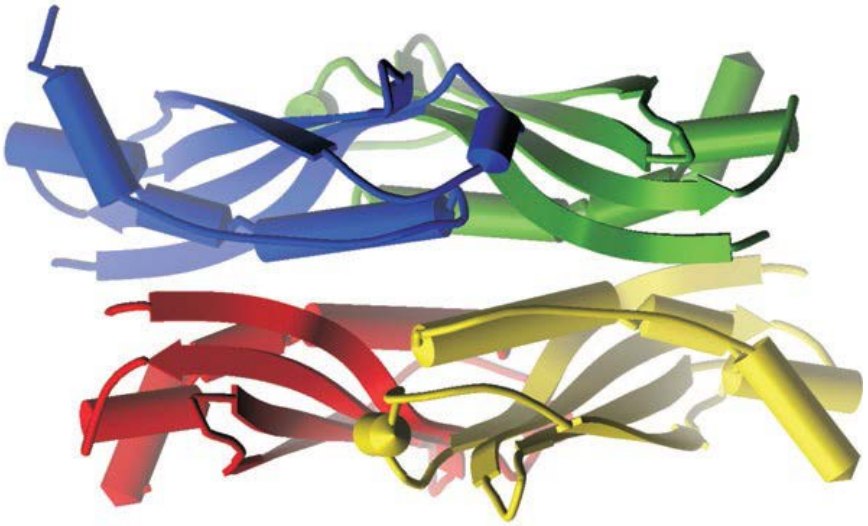
Evidence of SecB in anti-folding and anti-aggregation activity was also shown in the study of the impact of SecB on the secretory protein pathway.

2 2.2 Structure of SecB

In the year 2000, SecB chaperone was crystallized from *Haemophilus influenza* (Xu et al., 2000). Three years later, another SecB chaperone was crystallized from *E. coli (*Dekker *et al., 2003).* After some comparisons, these two proteins showed 55% sequence identity and they also exhibited a similar structure (Muren et al., 1999, Topping and Randall, 1994). SecB isorganized as a tetramer.It is rectangular in shape with a molecular mass of 17,200 dalton for each of the monomer (Smith et al., 1996). SecB consists of a simple α/β fold which is made up of four β-strands and two α-helices forming a stable and solid tetramer. The four β-strands are arranged into an antiparallel β-sheet that lies on the surface of the molecule above the α-helix. The two monomers are paired into a dimer through β-strand and α helix of each monomer while the tetramer is formed by the interaction of the dimers via their α-helix (Figure 2.4). The extreme C-terminal helix of each subunit does not pack against the tetramer after the third turn and this is consistent with the proton nuclear magnetic resonance studies done by Volkert and his colleague. They have shown that the C-terminal (13 residues) in *E. coli* SecB is highly mobile (Volkert et al., 1999).

The tetramer is stable as the equilibrium constant of the dissociation from tetramer to dimer is about 20 nM at pH7.6 (Muren et al., 1999). However, the high pH and high salt concentration will shift the equilibrium to the dimeric state (Smith et al., 1996). Three mutant forms of *sec*B, each having a single aminoacyl substitution at Cys76, Val78, and Gln80, which is positioned in the surface β-sheets with the side chain pointing to the inside of the molecule and not part of the

(a)



(b)

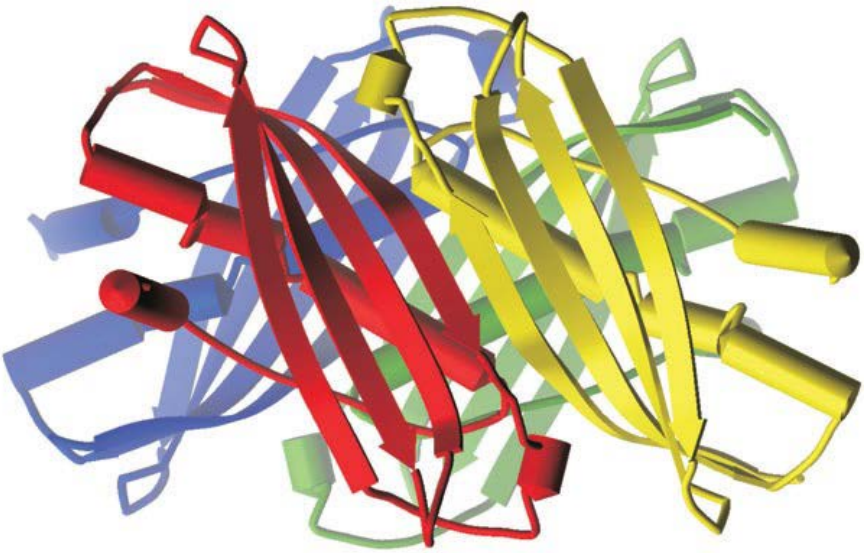


Figure 2.3: The crystal structure of SecB tetramer. The ribbon drawing is based on the Protein Data Base where (a) is the front view and (b) is the side view. Each subunit is represent by different colour (Bechtluft et al., 2010).

peptide-binding groove, caused the tetramer to be destabilized and yielded to form dimers under physiological conditions (Muren et al., 1999). SecB is unable to bind preproteins in this dimeric state. It is believed that the dimer-dimer interface is distorted due to the conformational changes caused by the mutation.

SecB is a highly acidic protein. Although SecB was crystalized without a peptide substrate in the binding site, a long groove is noticeable (Figure 2.5 dashed yellow circle) in the polarity surface profile of the tetramer, was suggested to be the peptide binding channel. There are two subsites in this long groove. Subsite one, is a deep cleft located close to the outside of the channel. It might recognize hydrophobic and aromatic regions of peptide substrates as it is formed by mostly conserved aromatic residues. The length of this groove is enough to accommodate peptides of the size of the SecB-binding motif (Knoblauch et al., 1999). Subsite two, is a shallow, open groove that is positioned in the middle with a hydrophobic surface. This subsite is proposed to be involved in the binding of extended region of polypeptides through the forming of hydrogen bonds due to its lack of aromatic residues. Around the groove, there are many negatively charged residues which could explain the selectivity of SecB for peptides with basic residues (Kim and Kendall, 1998).

2.2.3 SecB and other preproteins

Proteins that are produced by the ribosome are exposed to the molecular chaperones to aid them in the translocation process across the cytoplasmic membrane by folding/unfolding the proteins. Unfolded proteins are inclined to fold or aggregate in the cytosol in the absence of chaperones. Even though there are some general chaperones that can associate with the unfolded secretory proteins,

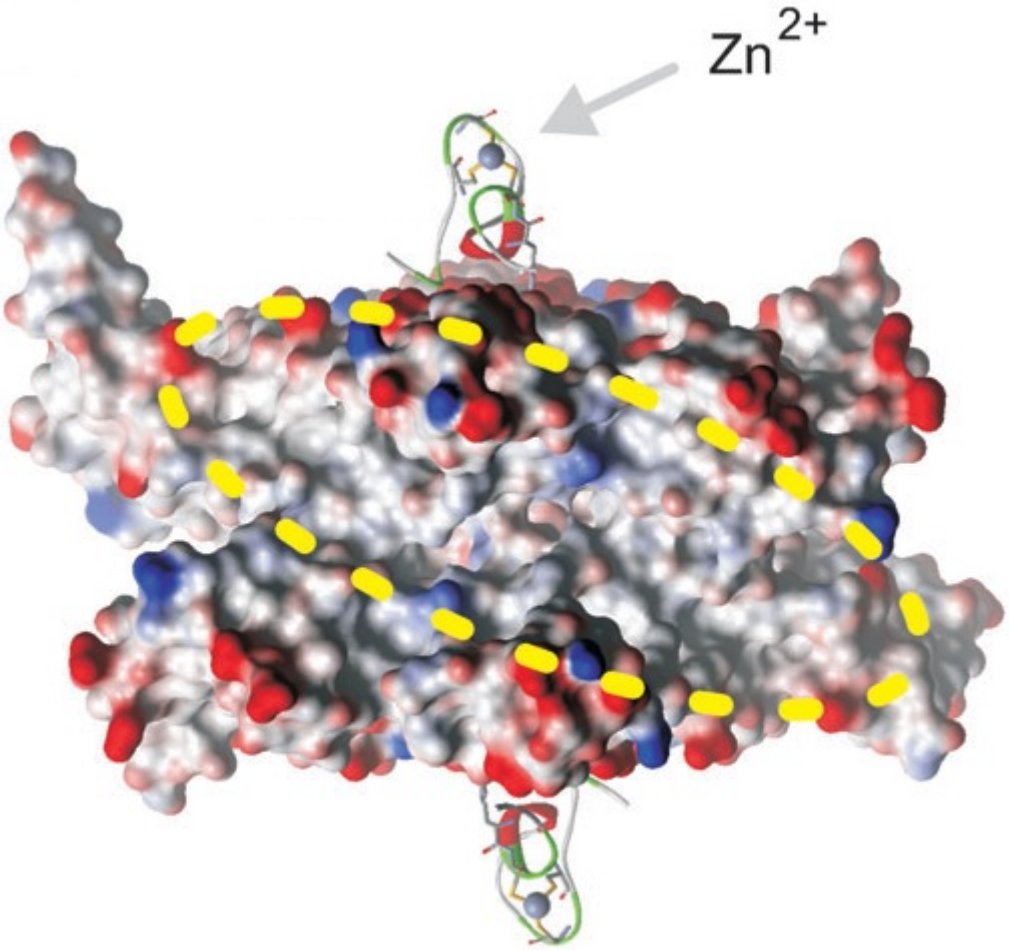


Figure 2.4: The side view of tetrameric SecB. The red colour indicates negative electrostatic potential while the blue colour represents the positive potential electrostatic. The yellow dash circle indicates the long groove where the polypeptides may be bound (Bechtluft et al., 2010).

only SecB seems to play a specific role in protein export. SecB has the ability to target secretory proteins to SecA besides preventing the proteins to fold and aggregate.

It has been shown that SecB binds to some nascent secretory protein only *in vivo.* Baars *et al*. has shown the direct evidence of SecB dependence for the transport of eighteen proteins in *E.coli*, namely MBP, GBP, PhoE, LamB, OmpF, OpA, DegP, FhuA, FkpA, OmpT, OmpX, OppA, TolB, TolC, YbgF, YgiW and YncF, using pulse labelling protein secretion studies and comparative proteomics (Baars et al., 2006). SecB associates only with the internal polypeptide in the critical cores that play a role in protein folding but does not recognize or bind to their signal sequences although all these proteins have a signal sequence that is located at the N-terminal, which directs them to the Sec-system (Knoblauch et al., 1999). SecB has weak consensus binding motif, a length of nine amino acids. It prefers unstructured peptides which are aromatic and basic residues; acid residues are disfavoured. It can bind to a few small ligands simultaneously (Bechtluft et al., 2010).

Naturally, SecB binds strongly with naturally unfolded preprotein with a one to one ratio. This was supported by the small dissociation constant, which is in the nanomolar range (Randall et al., 1998, Fekkes et al., 1995). However, SecB has been found to associate with nascent polypeptide chains of preMBP (150 residues in length) despite of its role in post-translational translocation of preprotiens. SecB also binds to stabilized molten globular-like proteins, eg. barstar, bovine pancrease trypsin inhibitor (BPTI), α-lactalbumin or ribonuclease A (RNaseA). SecB appears to be less selective and binds with many different proteins in a four to one ratio as long as they are unfolded *in vitro* (Panse et al., 2000, Panse et al., 1998).

2.2.4 Binding with SecA

SecB is one of the important counterparts for SecA. SecB is able to direct the preprotein substrates to SecA at the membrane translocase as it binds directly to SecA specifically. SecA is the motor domain of the translocase which is arranged as a homodimer. It is a 102 kDa ATPase motor. In the cell, SecA is either membrane-bound, soluble in cytosol or bound to the ribosome (Driessen and Nouwen, 2008). Soluble SecA binds loosely with SecB with a K_d of micromolar range and the binding is 50 times poorer compared to the membrane associated SecA (den Blaauwen et al., 1997). SecB binds strongly with the SecA attached to the protein conducting pore, SECYEG with a K_d of 30 nM. The binding affinity increased even greater to K_d (10 nM) when SecB is attached with a preprotein (SecB-preprotein complex) (Fekkes et al., 1998, Hartl et al., 1990). The signal sequence of the preprotein has contributed in this increased binding affinity. The signal sequence itself binds tightly to SecA forming a ternary SecA-SecB-preprotein complex. The high affinity binding site of SecA with SecB is highly conserved in bacteria even in those that lack the SecB homolog. The site was determined by truncation analysis and it only consists of the 22 amino acid in C-terminal of SecA (Breukink et al., 1995). High abundance of lysyl and arginyl residues have caused the presence of a net positive charge at the binding site. Besides, a zinc atom (chelated by three cysteines and a histidine) which is present at the C-terminus of SecA, is needed for the strong SecB binding as it stabilizes the binding site. This was proven in NMR studies. (Fekkes et al., 1999, Dempsey et al., 2004).

The C-terminal peptide of SecA bound to SecB has been crystallized in 2003. The structure suggests that the binding of SecA to SecB is in a two to one ratio at the acidic and negatively charged eight β-strand SecB surface. Electrostatic interaction,

hydrogen bonding and intermolecular salt bridges are likely found on the surface of the interaction. Four conserved residues in SecA that are required for the high bonding affinity are the neutral asparagine 882 and positively charged arginine 881, lysine 891 and lysine 893 and these were found by site-directed mutagenesis. When these four residues were mutated to alanine, SecA was still able to bind with the zinc atom but not SecB (Zhou and Xu, 2003).

There are three important sites for a tight binding of SecA and SecB to happen. The first is a stabilizing point at the zinc atom position at the extreme terminus of SecA and the negatively charged SecB (Fekkes et al., 1999). The second is the interaction between the C-terminal of SecB and the interfacial area of the SecA homodimer (Randall et al., 2005, Randall et al., 2004). The third contact point is not as clearly defined as the other two as it involves residues on the negatively charged β-sheet of SecB and the interface of the dimer of dimers (Woodbury et al., 2000). The NMR studies also suggested that the interaction of SecA-SecB complex is asymmetric although the structures of SecB and SecA showed that they have a two-fold symmetry (Randall et al., 2005).

There is also an anionic binding cluster in SecB that consists of some conserved residues that are pivotal for SecA binding especially in the functional binding of SecB to the SecYEG–bound SecA protein (Woodbury et al., 2000). However, in solution, the loose- binding of SecB-SecA does not depend on this cluster nor the C-terminal end of SecA. This cluster (Asp20, Glu24, Leu75 and Glu77) are positioned on the outer β-strand surface of *E. coli* SecB. Mutation of these residues does not affect the binding of preprotein with SecB but greatly reduced the bonding affinity with SecA (Fekkes et al., 1998). The mutation might cause a defect in the

transferring of preprotein from SecB to SecA, hence causing the translocation process to break down (Bechtluft et al., 2010).

The subunit stoichiometry for the SecA-SecB-preprotein complex to function is 2:4:1. After the SecB-preprotein is attached with SecA that is bound with SecYEG, the preprotein is transferred to SecA by SecB. SecB bind more tightly with SecYEG-bound SecA when it habours a preprotein. This is due to the fact that SecA is able to attach the signal sequence of the preprotein. This fits in the picture of the transfer mechanism of the preprotein. The preprotein most likely is attached to SecA instead of SecB in the stable ternary SecB-preprotein-SecA complex. The binding of the SecB to SecA will partially kick start the SecA ATPase activity. As soon as ATP is attached to SecA, SecB is released from SecA to cytosol after the preprotein translocation is activated. In the cytosol, SecB is free to bind with new preproteins (Fekkes et al., 1997). It has been proposed that the dissociation of SecA dimer to a monomer at SecYEG complex causes the release of SecB.

2.4.5 Other functions of SecB

Besides the roles in transportation, SecB may also act as a general buffer (holdase) for the unfolded polypeptides in the cytosol. This was supported by the depletion of heat shock chaperones (DnaKJ and GroEL/GroES) when SecB is overexpressed (Muller, 1996). It has been shown that SecB is able to help the role of DnaK in luciferease protein folding (Knoblauch et al., 1999).

SecB also takes part in the secretion of hemophore, HasA although HasA does not utilize the sec system as their translocation machinery. HasA is secreted by a specific ABC transport. HasA is a protein that rapidly folds and SecB has the ability to stop the folding process, thusmaintaining HasA in a translocation competent state

when HasA is bound to it. Therefore, HasA is strictly dependent on SecB (Delepelaire and Wandersman, 1998, Wolff et al., 2003).

SecB does not only have the ability to prevent the preprotein from folding, it also can dissolve aggregates of misfolded protein. This function was observed by several heat shocked chaperones. However, Bechtluft and team found that SecB was unable to dissolve MBP aggregates in a single molecule measurement, but study from Panse and team shows that SecB dissolves insulin B chain aggregates (Bechtluft et al., 2007, Panse et al., 2000).

2.3 Polyhydroxyalkanoates (PHA)

Polyhydroxyalkanoates (PHAs) are a group of biologically synthesized polyesters by a wide range of prokaryotic microorganisms, from kingdom of eubacteria to members of family Halobacteriacea of the Achaea (Steinbuchel et al., 1995). These microorganisms have the ability to form an intracellular and spherical organic inclusion under nutrient limited conditions. This biopolymer is accumulated as a unique water-insoluble inclusion in the cytoplasm. This polyester is stored as intracellular carbon and energy storage material when there is imbalance growth condition due to excess of carbon sources and limited amounts of essential nutrients such as nitrogen, phosphate, magnesium and etc (Anderson and Dawes, 1990, Doi, 1990). The PHAs which are usually stored as granules can be accumulated by bacteria intracellularly up to 90% of the cell dry weight (Madison and Huisman, 1999). The accumulation of PHAs helps in bacteria's survival under extreme environment or starvation (Kadouri et al., 2005) . When starvation happens, the bacteria will secrete intracellular or extracellular PHA depolymerase to degrade the PHAs and utilize the decomposed compounds as carbon and energy source (Philip et al., 2007). PHA synthesis that undergoes a process of polymerizing soluble

intermediates into insoluble intermediates will also maintain the osmotic state of the cell and prevent essential compounds leak out of the cell (Madison and Huisman, 1999). Besides being carbon and energy source, the biopolymer has an impact on bio-film formation and is able to function as an electron sink (Pham et al., 2004). PHAs' unique features have made them popular as substitutes for many petrochemical products. PHAs are used in a wide range of applications which include packaging, medical, pharmaceutical, food, manufacturing, agricultural and fisheries industries (Rehm and Steinbuchel, 1999). Although PHAs have many applications in various industries and are environmental friendly, they are not widely produced on a commercial scale. The production and status of cheap conventional plastics such as polypropylene and polystyrene, is still not challenged by PHAs until today.

2.4 Allelic Exchange

As the sequencing technology is more and more advanced these years, the number of uncharacterized Open Reading Frame (ORFs) has increased with the increased number of sequenced genome. The fundamental studies to understand the bacterial metabolism is through the characterization of the gene functions. The best way to study gene function is by gene disruption. Gene mutation can be categorized into random mutagenesis or site specific mutagenesis. While random mutagenesis can be done by using chemical or transposons, site-specific mutation can be done through allelic exchange on the bacteria chromosome. The method to achieved allelic exchange is through homologous recombination.

Homologous recombination is the exchange of nucleotide sequences between two similar DNA molecules. It can be done through two methods, either by linear DNA fragment transfer or with the use of a suicide vector. Both methods are done by

delivering a DNA flanking segment of the altered targeted site either in linear form or carried by a suicide vector into the recipient by electroporation or conjugation. Suicide vector is a vector that is replication incompetent and is unable to be replicated during cell differentiation in the host cell. This method was first introduced by Scherer and Davis 1979 and has been widely used to obtain unmarked deletion of bacteria and yeasts (Scherer and Davis, 1979). The housekeeping mechanism will swap the mutated version DNA fragment on the plasmid vector with the wild type allele on the bacteria chromosome. Figure 2.6 will show the graphic presentation of the allelic exchange.

Besides that, an antibiotic resistance gene or a counter selection marker is needed to select the cells which had received the vector and remove the vector after mutagenesis has taken place. Therefore this selection marker is incorporated in the vector before it carries the DNA fragment. A few selection markers has been developed and used. The use of selection marker is largely dependent on the bacteria, the gene that is targeted and the vector used.

The sucrose-sensitivity system, which is based on the *Bacillus subtilis sacB,* which encodes levansucrase, causes lethality in most Gram-negative bacteria when grown in sucrose containing medium.

With the use of a suicide plasmid, allelic exchange happens via merodiploid formation (essentially a haploid organism which carries a second copy of a part of its genome). The merodiploid, which is formed by a single crossover event, can be selected by a marker on the vector. The resolved merodiploids are a mixture of the wild type and the mutant in the absence of the antibiotic used for selection. Therefore, the deletion mutant can be easily distinguished from the merodiploids by their phenotypes.