

---

**Characterisation of the phosphate transporters  
PitA and PitB from *Escherichia coli***

---

A thesis submitted for the Degree of  
Doctor of Philosophy  
of  
The Australian National University

Robyn Margaret Harris

April, 2002

Membrane Biochemistry Group  
Division of Biochemistry and Molecular Biology  
John Curtin School of Medical Research  
and  
The School of Biochemistry and Molecular Biology  
The Australian National University

## *Errata*

### **Insert the following paragraph after the second paragraph on page 19:**

Experimentally determined topologies for proteins with the above “6 + 6” TM structure usually have both their N- and C-termini located in the cytoplasm (28, 78, 198). However, while membrane proteins with N-termini located in the periplasm are less common, they are not unusual. For example, bacteriorhodopsin has 7 TM helices and a periplasmic N-terminus (201, and references therein).

### **Insert the following paragraph after the second paragraph on page 69:**

The inhibition of cell growth by the addition of IPTG to AN3521 (pAN910 in MC15) cell cultures suggested that over-expression of the *pitA::6xHis* at the levels induced by IPTG from the T5 promoter may be toxic to the cells. This was further tested by transferring *pitA::6xHis* into vector pBR322, so the characteristics of PitA with a 6xHis tag could be compared with wild type PitA in a similar environment. While a protein was produced by *in vitro* transcription/translation from this plasmid, it did not allow growth of strain AN3066 on 500 $\mu$ M Pi minimal media, unlike the PitA control plasmid. Thus the C-terminal 6xHis tag has a detrimental effect on the activity of PitA. However, *pitA::6xHis* placed in the pGex vector system under the control of the *tac* promoter could grow in strain AN3066 upon the addition of IPTG. This strain showed negligible growth without IPTG induction. Thus under certain conditions PitA6xHis protein can allow Pi transport. The varying results in different vector systems suggest that PitA6xHis protein can be toxic to cells, but this depends on the level of expression and other undefined conditions within the cells and cellular environment.



**Replace the first paragraph on page 74 with the following paragraph:**

Measurement of  $^{33}\text{Pi}$  uptake by cells was carried out under conditions in which Pi uptake was linear over time. Examples of the experiments used to determine  $K_m^{\text{app}}$  and  $V_{\text{max}}^{\text{app}}$  are given in Figure 3.17. Two families of curves exist for PitB activity, with significantly different  $V_{\text{max}}^{\text{app}}$  values. The experiments providing the two higher  $V_{\text{max}}^{\text{app}}$  values were carried out within a week of each other. Subsequent experiments could not replicate the high maximum velocities obtained here. All possible variables were checked, including buffers, substrates, equipment and cultures. Statistical analysis was used to combine the  $K_m^{\text{app}}$  and  $V_{\text{max}}^{\text{app}}$  parameters from individual experiments to produce the values listed in Table 3.3.

**Insert the following as the second paragraph on page 75:**

It is possible that the different  $V_{\text{max}}^{\text{app}}$  values obtained for PitB activity in Figure 3.17 could be caused by a complex balance of regulatory systems that have a large effect on PitB expression (discussed in more detail in Chapter 7.)

**Insert the following sentence after the second paragraph on page 108.**

Pi uptake for the PitA amino acid substitutions listed in Table 6.3 which have positive growth on  $500\mu\text{M}$  Pi could be measured to ascertain whether these mutations also affect PitA activity.

## ***Statement***

The work described in this Thesis was performed by the author, except where specifically stated in the text.

A handwritten signature in black ink, appearing to read "Robyn Harris". The signature is written in a cursive style with a large initial 'R' and 'H'.

Robyn Margaret Harris

April, 2002

# ***Acknowledgements***

Learning the intricate art of independent research has been a rite of passage, filled with the challenges of acquiring new skills while balancing the need to keep your vision clear, your imagination wide open, your feet on the ground and your eye on the goal. As in any journey of discovery, developing resilience and learning to work with your own strengths and weaknesses is an integral part of the process. Add a major illness into the mix, and you have a life-changing experience that would test the mettle of any adventurer. I would like to acknowledge and thank the many people who have helped me to navigate my way along this journey.

First and foremost I would like to thank my supervisors, Professor Graeme Cox and Dr Susan Howitt. Thank you Graeme for your positive attitude to life, your enthusiasm for science and your ability maintain a group environment where everyone felt comfortable asking questions and making suggestions. Your guidance has been inspirational. Susan, thank you for so generously and ably taking up the baton Graeme passed you. The clarity and focus you brought to this project was invaluable. My writing skills have certainly improved under your guidance and my appreciation of Freddo moments has grown immensely.

I greatly appreciated the easygoing, enthusiastic environment of the membrane biochemistry students lab, where an experiment always seemed to be in progress, but everyone could down tools and bop for a funky track. Thank you Brett for all your experimental innovations, Louise for your sage advice and good music, and Andy for knowing how to cheer me up after one of those “oh no” moments.

It was wonderful to work within the membrane biochemistry group and I am sad at its demise. Thank you Di, for your help in transforming this protein chemist into a molecular biologist, to Lyndall for all your knowledge and advice (and the excitement of a few rampages), to Gaz for always knowing how to trim an experimental technique down to its essentials, to Russell and Terri for all your technical assistance, and to Frank, Susie Mac, Davy Dogbox, Tim, Greg, Rowena, Judy, Anna, Jules and Phyllis for keeping the place friendly and interesting. A special thanks to Helen Cornwall in JCSMR administration, for your care, concern and skill in guiding me through all the



complicated paperwork associated with suspensions and sick leave. Thank you also to all the JCSMR PhD students who laughed, sympathised and shared confidences with me on the bumpy ride that is student life. The student network we established was an important support for me and I hope it is still thriving.

I would like to thank my 'new lab' for making me feel so welcome. To the legendary BAMBI lab of "long lean women who always eat cake" – Susan, Louise, Megan, Kate, Fiona, Lani, Rhianna – plus the very tolerant Patrick.

I'd like to thank my parents for providing a haven for a bedraggled and very sick daughter without showing too many outward signs of panic!

To my friends, who tried to (and partly succeeded) in keeping my outlook balanced. Linda, Jo, Liz, Peter, Dave and Jo V. – it was a valiant effort, thanks heaps. And Jules, you can stop nagging now.

Michael – you have been such a steady and encouraging support, even though you had very little idea of what you were letting yourself in for when you said "just finish it". I'm so glad you dragged me away on sea kayaking trips and Lizard Island lounges and let me take the time to reach my own decisions. Your ability to do the practical things that really count always amazes me – love you!

Finally, thank you to all those involved in Dru Yoga for showing me how to meet life's challenges with an open heart.

*Om shanti*

## Abstract

*Escherichia coli* contains two major systems for active uptake of inorganic phosphate (Pi). The high-affinity Pst system (*pstSCAB*) is induced by low external Pi concentrations as part of the *pho* regulon and is an ABC transporter, requiring ATP for active transport. The low-affinity Pit system had previously been shown to be a secondary active transporter, dependent on the proton-motive force, expressed constitutively and with all known mutations mapping to one locus. However, the discovery of a second *E. coli* Pit transporter by the Cox laboratory suggested these conclusions may need to be re-examined. These two Pit genes, subsequently named *pitA* and *pitB*, have a putative amino acid sequence identity of 81% and preliminary experiments indicated that the proteins have significantly different  $K_m^{\text{app}}$  values for Pi.

Thus, PitA and PitB have been characterised and compared to investigate the relationship between the structure and function of these Pi transporters. *pitA* and *pitB* were defined as separate *E. coli* gene products and the putative open reading frames were confirmed. Plasmid-borne PitA was found to have an apparent  $K_m$  ( $K_m^{\text{app}}$ ) of approximately 2 $\mu\text{M}$ , which was 14-fold lower than the  $K_m^{\text{app}}$  of 28 $\mu\text{M}$  obtained for PitB.

Strain K-10 lacks Pit activity when the Pst system is mutated and this Pit mutation was identified. The K-10 strain contains a nonfunctional *pitA*, which has a single base change (G to A at nucleotide 658 in the open reading frame), causing an aspartic acid to replace glycine 220. This mutation greatly decreased the amount of PitA protein present in cell membranes, indicating that the aspartic acid substitution disrupts membrane protein assembly. Genomic *pitB* had wild type sequence, even though this strain was not active as a Pi transporter.

Subsequent investigation of *pitB* regulation suggested that *pitB* encodes a functional Pi transporter whose repression at low Pi levels is mediated through the *pho* regulon. While wild type plasmid-borne *pitB* expressed in a *pitA*  $\Delta$ *pstC345* mutant allowed growth on media containing 500 $\mu\text{M}$  Pi as the sole source of phosphate, *E. coli* with a wild type genomic *pitB* (*pitA*  $\Delta$ *pstC345* double mutant) was unable to grow under these

conditions, making it indistinguishable from a *pitA pitB ΔpstC345* triple mutant. The mutation *ΔpstC345* constitutively activates the *pho* regulon, which is normally induced by phosphate starvation. Removal of *pho* regulon activation, by deleting the *phoB-phoR* operon, allowed the *pitB<sup>+</sup> pitA ΔpstC345* strain to utilise Pi and to transport Pi at rates significantly higher than background levels. In addition, post-transcriptional or post-translational regulation of PitB protein may also occur. This was suggested by an unexpected decrease in the  $K_m^{app}$  of plasmid-borne PitB in the pAN1116 plasmid construct that was accompanied by greatly increased levels of protein expression.

The three charged residues His-225, Asp-229 and Lys-232 were shown to be important for PitA structure or function, as mutation of each residue to a polar amino acid greatly reduced the initial rate of Pi uptake, while still allowing assembly of the mutant protein in the membrane. These residues are analogous to charged amino acids important in the function of lactose permease and are identical to residues in PitB, with His-225 and Asp-229 being highly conserved throughout the PiT family of Pi transporters. A possible model for the secondary structure of PitA within the cytoplasmic membrane was revised to move this triad from a ‘periplasmic loop 6’ into a ‘transmembrane  $\alpha$ -helix 6’ position. (The alternative periplasmic location of this loop still remains plausible.) This area of PitA is obviously important for the mechanism. As well as containing the above triad of amino acids, this putative transmembrane helix plus the putative ‘loop 5’ region contains sequence determining the difference in  $K_m^{app}$  between PitA and PitB, as shown by chimera analysis. Nearby single mutations G220D and A213D cause complete loss of Pi uptake and minimal insertion of protein in the membrane. Mutation of amino acids associated with one of the ‘highly conserved regions’ of the PiT family of transporters completely abolished Pi uptake activity, while allowing normal levels of membrane protein assembly, indicating the importance of this sequence in Pit structure or function. More detailed studies are required before this information can be integrated into a mechanistic model for PitA function.



# ***Abbreviations***

A	adenosine
Å	angstrom
AMP	adenosine monophosphate
ANU	Australian National University
Amp	ampicillin
APS	ammonium persulfate
ATP	adenosine triphosphate
bp	base pairs
C	cytosine
°C	degrees Celsius
Cat	chloramphenicol
cpm	counts per minute
DNA	deoxyribonucleic acid
ds	double strand
DTT	DL-dithiothreitol
EDTA	ethylenediamine tetra-acetic acid
G	guanosine
IPTG	isopropyl $\beta$ -d-thiogalactoside
JCSMR	John Curtin School of Medical Research
kb	kilobase(s)
kDa	kilodalton(s)
Km	kanamycin
$K_m^{app}$	apparent maximum $K_m$
LB	Luria Bertani
mA	milliamps
mRNA	messenger ribonucleic acid
MDR	multiple drug resistance
ORF	open reading frame
PAGE	polyacrylamide gel electrophoresis
PBS	phosphate buffered saline
PCR	polymerase chain reaction

Pi	inorganic phosphate
Pi media	minimal media containing 500 $\mu$ M Pi as the sole source of phosphate
PMSF	phenyl-methanesulfonylfluoride
PVDF	polyvinylidene difluoride
RF	replicative form
RNA	ribonucleic acid
rpm	revolutions per minute
SEM	standard error of the mean
SDS	sodium dodecyl sulfate
ss	single strand
T	thymidine
TE	Tris-EDTA
TEMED	N.N.N'N'-tetramethyl-ethylenediamine
TM	transmembrane
Tris	Tris {hydroxymethol amino-methane}
w/v	weight per volume
UV	ultraviolet
V	voltage
$V_{\max}^{\text{app}}$	apparent maximum velocity
v/v	volume per volume

# Table of contents

<b>CHAPTER ONE</b>	<b>1</b>
<b>GENERAL INTRODUCTION</b>	
<b>1.1 Phosphorus transport in <i>Escherichia coli</i></b>	<b>1</b>
1.2.1 The outer membrane	1
1.2.2 The periplasm	2
1.2.3 The cytoplasmic membrane	3
<b>1.3 The phosphate transporters of <i>Escherichia coli</i> K-12</b>	<b>4</b>
1.3.1 PhoE	4
1.3.2 Pit	5
1.3.3 Pst system	7
1.3.4 GlpT	8
1.3.5 UhpT	9
1.3.6 Ugp system	10
<b>1.4 Regulation of inorganic phosphate transport in <i>Escherichia coli</i></b>	<b>10</b>
<b>1.5 Mechanisms of membrane transporters</b>	<b>13</b>
1.5.1 Classification of membrane proteins by their mechanism of transport	13
1.5.2 Secondary active transporters	14
<b>1.6 The PiT family</b>	<b>15</b>
<b>1.7 The structure of membrane proteins</b>	<b>16</b>
1.7.1 The basic structural motifs of membrane proteins	16
1.7.2 The general structure of outer membrane $\beta$ -barrel proteins	17
1.7.3 The general structure of cytoplasmic membrane proteins	18
<b>1.8 Methods to determine membrane protein structure</b>	<b>21</b>
1.8.1 Direct methods for the determination of membrane protein structure	21
1.8.2 Indirect methods for the determination of membrane protein structure	25
<b>1.9 The functional analysis of membrane proteins</b>	<b>27</b>
1.9.1 Methods for the functional analysis of membrane proteins	27
1.9.2 Structural and functional analysis of lactose permease	28
1.9.3 The success of indirect techniques in structure and function predictions	33
<b>1.10 Aims of the project</b>	<b>37</b>
<b>CHAPTER TWO</b>	<b>38</b>
<b>EXPERIMENTAL PROCEDURES</b>	
<b>2.1 Bacterial strains and plasmids</b>	<b>38</b>
<b>2.2 Oligonucleotide sequences</b>	<b>38</b>
<b>2.3 Growth media</b>	<b>38</b>
<b>2.4 General molecular biology techniques</b>	<b>40</b>
2.4.1 Small scale purification of plasmid DNA or replicative form M13 DNA	40
2.4.2 Large scale purification of plasmid DNA by caesium chloride (CsCl) density gradient	41
2.4.3 Purification of chromosomal DNA from <i>E. coli</i>	42



2.4.4	Preparation of single-stranded M13 DNA	42
2.4.5	Spectrophotometric analysis of DNA	43
2.4.6	Agarose gel electrophoresis	43
2.4.7	“Rapid cracking” of phage or colonies	43
2.4.8	Synthesis and preparation of oligonucleotides	44
2.4.9	Polymerase chain reaction (PCR)	44
2.4.10	Purification of PCR products	45
2.4.11	Site-directed mutagenesis	45
2.4.12	Manual sequencing of DNA	46
2.4.13	Automated sequencing of DNA	48
<b>2.5</b>	<b>Recombinant DNA techniques</b>	<b>48</b>
2.5.1	Restriction endonuclease digestion of DNA	48
2.5.2	Elimination of 5' and 3' protruding ends	49
2.5.3	Removal of terminal phosphates from linearised DNA	49
2.5.4	Phosphorylation of oligonucleotides	49
2.5.5	Purification of DNA fragments from agarose gels	50
2.5.6	Ligation of DNA	50
2.5.7	Preparation of competent <i>E. coli</i>	50
2.5.8	Transformation of competent <i>E. coli</i> with plasmid DNA	52
<b>2.6</b>	<b>Genomic recombinant DNA techniques</b>	<b>53</b>
2.6.1	Genetic recombination on the <i>E. coli</i> genome	53
2.6.2	Bacteriophage P1 <sub>kc</sub> transduction	53
<b>2.7</b>	<b>Protein biochemistry techniques</b>	<b>55</b>
2.7.1	Gel electrophoresis of proteins	55
2.7.2	Coomassie staining of polyacrylamide gels	55
2.7.3	Measurement of protein concentration	56
2.7.4	Detection of protein by Western blot	56
2.7.5	Preparation of cytoplasmic membranes from <i>E. coli</i> by centrifugation	57
2.7.6	Preparation of cytoplasmic membranes from <i>E. coli</i> by ammonium sulfate precipitation	58
2.7.7	Solubilisation of cytoplasmic membranes	58
<b>2.8</b>	<b>Over-expression and purification of protein using the 6xHistidine affinity tag</b>	<b>58</b>
<b>2.9</b>	<b><i>In vitro</i> transcription/translation of plasmid DNA</b>	<b>60</b>
<b>2.10</b>	<b>Identification of <i>E. coli</i> cells able to transport inorganic phosphate (Pi)</b>	<b>60</b>
<b>2.11</b>	<b>Inorganic phosphate (Pi) transport assay</b>	<b>61</b>
<b>2.12</b>	<b><i>phoA</i> alkaline phosphatase assay</b>	<b>62</b>
<b>2.13</b>	<b>Production of antipeptide polyclonal antibodies</b>	<b>62</b>
2.13.1	Peptides used for raising antibodies	62
2.13.2	Immunisation of rabbits with Multi-Antigenic Peptides	63
2.13.3	Detection of antipeptide polyclonal antibodies in rabbit serum	63
2.13.4	Purification of polyclonal antipeptide antibodies from serum by immunoaffinity chromatography	64
<b>2.14</b>	<b>Analysis of DNA sequences using MacTargsearch 2.0</b>	<b>64</b>
<b>CHAPTER THREE</b>		<b>66</b>
<b><i>ESCHERICHIA COLI</i> HAS TWO <i>PIT</i> GENES WITH DISTINCT CHARACTERISTICS</b>		
<b>3.1</b>	<b>Introduction</b>	<b>66</b>

3.2 Comparison of the nucleotide sequences of <i>pitA</i> and <i>pitB</i>	67
3.3 Are <i>pitA</i> and <i>pitB</i> genes or alleles?	67
3.4 Over-expression and purification of PitA	68
3.5 Identification of the translation starts of the <i>pitA</i> and <i>pitB</i> open reading frames using site-directed mutagenesis	71
3.6 Comparison of the deduced amino acid sequences of PitA and PitB	72
3.7 Kinetic parameters of PitA and PitB	73
3.8 Discussion	74
<b>CHAPTER FOUR</b>	<b>76</b>
<b>IDENTIFICATION OF THE <i>E. COLI</i> K-10 PIT MUTATION</b>	
4.1 Introduction	76
4.2 Identification of the <i>E. coli</i> K-10 <i>pit</i> mutation/s	77
4.3 Functional analysis of <i>pitA1</i> from K-10 strain AN3066	78
4.3.1 <i>pitA1</i> Pi transport	78
4.3.2 Preparation of <i>pitA</i> (G220D) by site-directed mutagenesis	78
4.3.3 Analysis of PitA expression using polyclonal PitA antipeptide antisera	79
4.3.4 Comparison of PitA assembly in strains containing either wild type <i>pitA</i> or <i>pitA1</i>	80
4.4 Discussion	81
<b>CHAPTER FIVE</b>	<b>84</b>
<b>EXPRESSION AND REGULATION OF <i>PIT A</i> AND <i>PIT B</i></b>	
5.1 Introduction	84
5.2 Preparation of a Pi transport triple mutant	86
5.3 Expression of wild type <i>pitA</i>	87
5.3.1 Expression of wild type <i>pitA</i> from the K-12 genome	87
5.3.2 Expression of plasmid-borne <i>pitA</i> or <i>pitB</i> in a Pi transport triple mutant strain	88
5.4 Expression and activity of wild type <i>pitB</i>	88
5.4.1 Variability of <i>pitB</i> kinetic parameters $K_m^{app}$ and $V_{max}^{app}$	88
5.4.2 PitB antipeptide antibody production	89
5.4.3 Investigation of PitB protein expression	90
5.5 Regulation of <i>pitB</i>	91
5.5.1 Deletion of the <i>phoB-phoR</i> operon from <i>pitA1</i> $\Delta$ <i>pstC345</i> strains with/without genomic <i>pitB</i>	91
5.5.2 Effect of the <i>pho</i> regulon on <i>pitB</i> activity	92
5.6 Discussion	93

<b>CHAPTER SIX</b>	<b>102</b>
<b>INVESTIGATIONS INTO THE STRUCTURE AND FUNCTION OF PITA</b>	
6.1 Introduction	102
6.2 Putative topological models for PitA and PitB	104
6.3 Identification of important amino acids in PitA	106
6.3.1 Investigation of charged or polar amino acids located within the putative transmembrane helices of PitA	106
6.3.2 Investigation of the role of other PitA amino acids	107
6.4 Investigation of substrate binding by PitA and PitB through the construction and analysis of chimeras	108
6.4.1 Preparation of <i>pitA</i> and <i>pitB</i> with silent restriction endonuclease sites	108
6.4.2 Production and analysis of chimeras	109
6.5 Sequence comparison and domain analysis of PitA and PitB with other members of the PiT family	110
6.6 Discussion	112
<b>CHAPTER SEVEN</b>	<b>118</b>
<b>OVERVIEW AND FUTURE DIRECTIONS</b>	
7.1 Introduction	118
7.2 Regulation of the Pit system	118
7.3 Structure and function studies on PitA and PitB	124
<b>BIBLIOGRAPHY</b>	<b>131</b>



## ***Publications***

Harris, R. M., Webb, D. C., Howitt, S. M. and Cox, G. B. 2001. Characterisation of PitA and PitB from *Escherichia coli* J. Bacteriol. 183:5008-5014.

---

# ***Chapter one***

## ***General Introduction***

---

# **General Introduction**

## **1.1 Phosphorus transport in *Escherichia coli***

Phosphorus plays a key role in the metabolism of all cells. It is intrinsic for DNA and RNA synthesis, is a key to the lipid component of membranes, is essential for cell energy and membrane transport and is involved in many cell signaling and regulatory functions through phosphorylation and dephosphorylation. Therefore it is not surprising that many bacteria have evolved multiple systems for transporting phosphorus. *Escherichia coli* can utilise inorganic phosphate (Pi), organophosphates or phosphonate as phosphorus sources for growth, although Pi is the preferred source. Many phosphorylated compounds which cannot be metabolised are hydrolysed to release Pi into the periplasm. When Pi is plentiful transport is carried out by general porins and the inner membrane Pit transporter. The level of external Pi regulates expression of genes in 0.5-1% of the *E. coli* chromosome (204). When Pi concentrations become limiting, many proteins involved in Pi transport and metabolism are induced. These include an outer membrane porin, periplasmic enzymes which break down a variety of phosphorylated compounds and inner membrane transporters which have a high affinity for Pi or transport additional sources of phosphate, such as organophosphates or phosphonate (Figure 1.1).

### **1.2.1 The outer membrane**

Gram negative bacteria have two membrane structures that nutrients have to cross before reaching the cytoplasm. The major function of the outer membrane is to act as a coarse sieve to exclude noxious molecules present in the external environment (184). It consists of an asymmetric lipid bilayer, with a lipopolysaccharide outer leaflet and a phospholipid inner leaflet. The hydrophobic core of this membrane prevents the entry of amphiphilic and hydrophilic compounds, while the outer leaflet of lipopolysaccharide is thought to reduce spontaneous diffusion by uncharged hydrophobic molecules by creating a membrane interior with low fluidity. Thus the outer membrane is 50-100 times less permeable to steroids than the symmetrical phospholipid bilayer of a typical inner (cytoplasmic) membrane (184).

## Figure 1.1

### Proteins involved in phosphate transport under high and low external concentrations of inorganic phosphate ( $P_i$ ).

The transport systems and proteins involved are:

#### Outer membrane

OmpF/other general porins

Transport small molecules (less than 600 Daltons).

PhoE (phosphoporin)

Transports small molecules (less than 600 Daltons), with a slight preference for negatively charged substrates. Induced by  $P_i$  limitation.

#### Periplasm

AP (alkaline phosphatase)

Breaks down a variety of phosphate compounds, releasing  $P_i$ . Induced by  $P_i$  limitation.

Binding proteins from various binding protein dependent ABC transporters.

#### Inner membrane

Pit (low-affinity inorganic phosphate transporter)

Transports inorganic phosphate. PitA seems to be constitutive. Single protein.

Pst (high-affinity inorganic phosphate transporter)

Transports inorganic phosphate. Induced by  $P_i$  limitation. A binding protein dependent ABC transporter consisting of PstS (binding protein), PstA and PstC (membrane components) and PstB (contains the nucleotide binding domain).

GlpT (*sn*-glycerol 3-phosphate transporter)

Transports *sn*-glycerol 3-phosphate (G3P) as a  $P_i$ -linked antiporter, and can provide a net uptake of  $P_i$  by importing 2 monovalent G3P in exchange for one divalent G3P. Induced by G3P or glycerol. Undergoes catabolite repression.

UhpT (hexose phosphate transporter)

Transports various hexose phosphates, such as glucose 6-phosphate (G6P) as a  $P_i$ -linked antiporter. May provide a net uptake of  $P_i$  by a similar mechanism to GlpT, but the available evidence is conflicting. Activated by external G6P. Undergoes catabolite repression.

Ugp (G3P/glycerophosphoryl diester transporter)

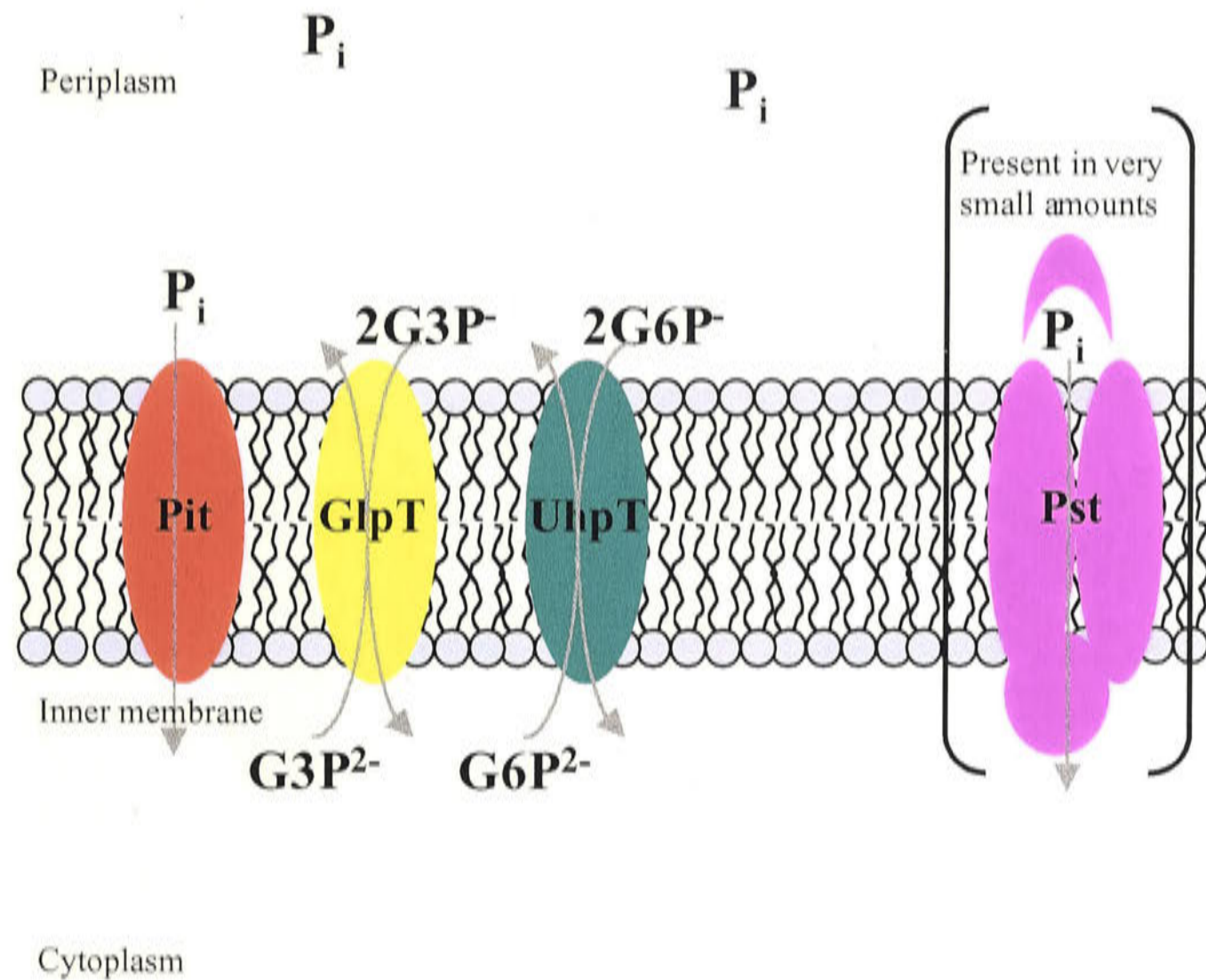
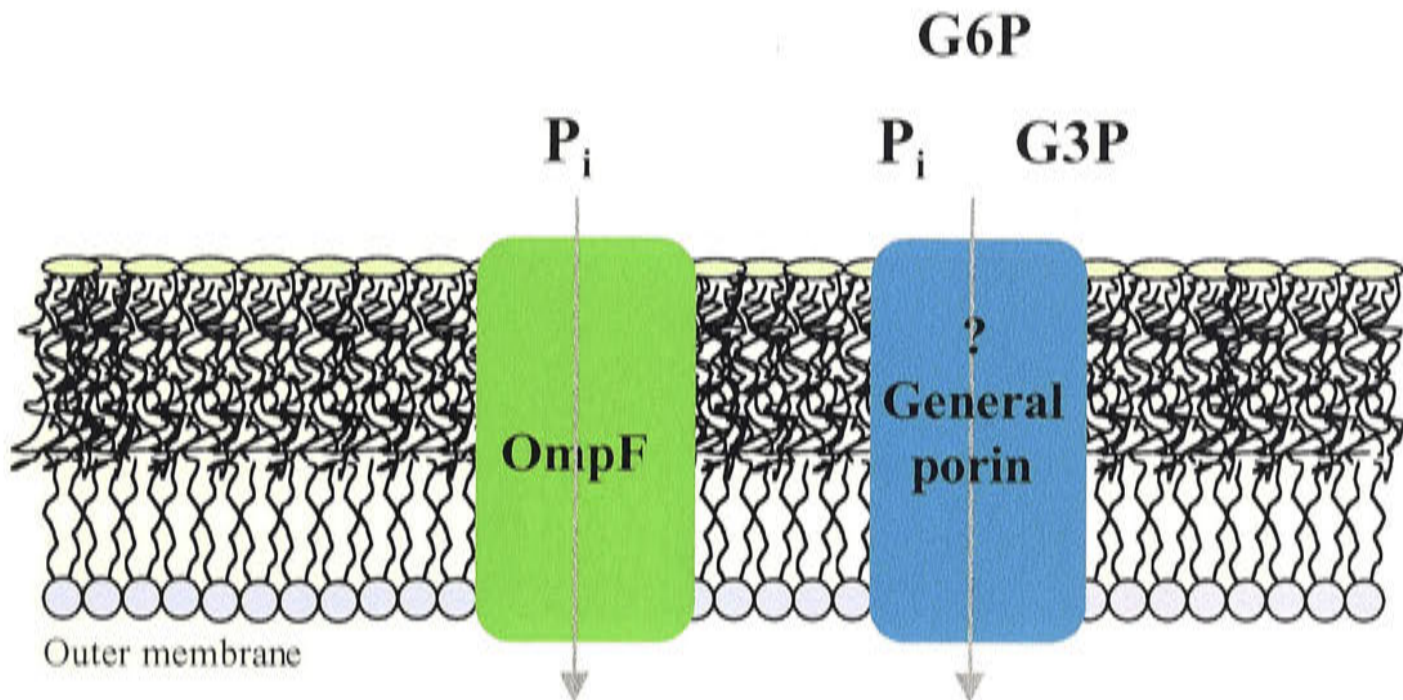
Transports G3P and glycerophosphoryl diesters. Induced by  $P_i$  limitation. G3P can provide the sole source of  $P_i$  through this transporter. Undergoes catabolite repression. This is a binding protein dependent ABC transporter.

Phn (phosphonate transporter)

Transports phosphonates ( $P_n$ ), which have a direct phosphate/carbon bond. This binding protein dependent ABC transporter is induced by  $P_i$  limitation, but is cryptic in K-12 and K-10 strains.

**Figure 1.1A**

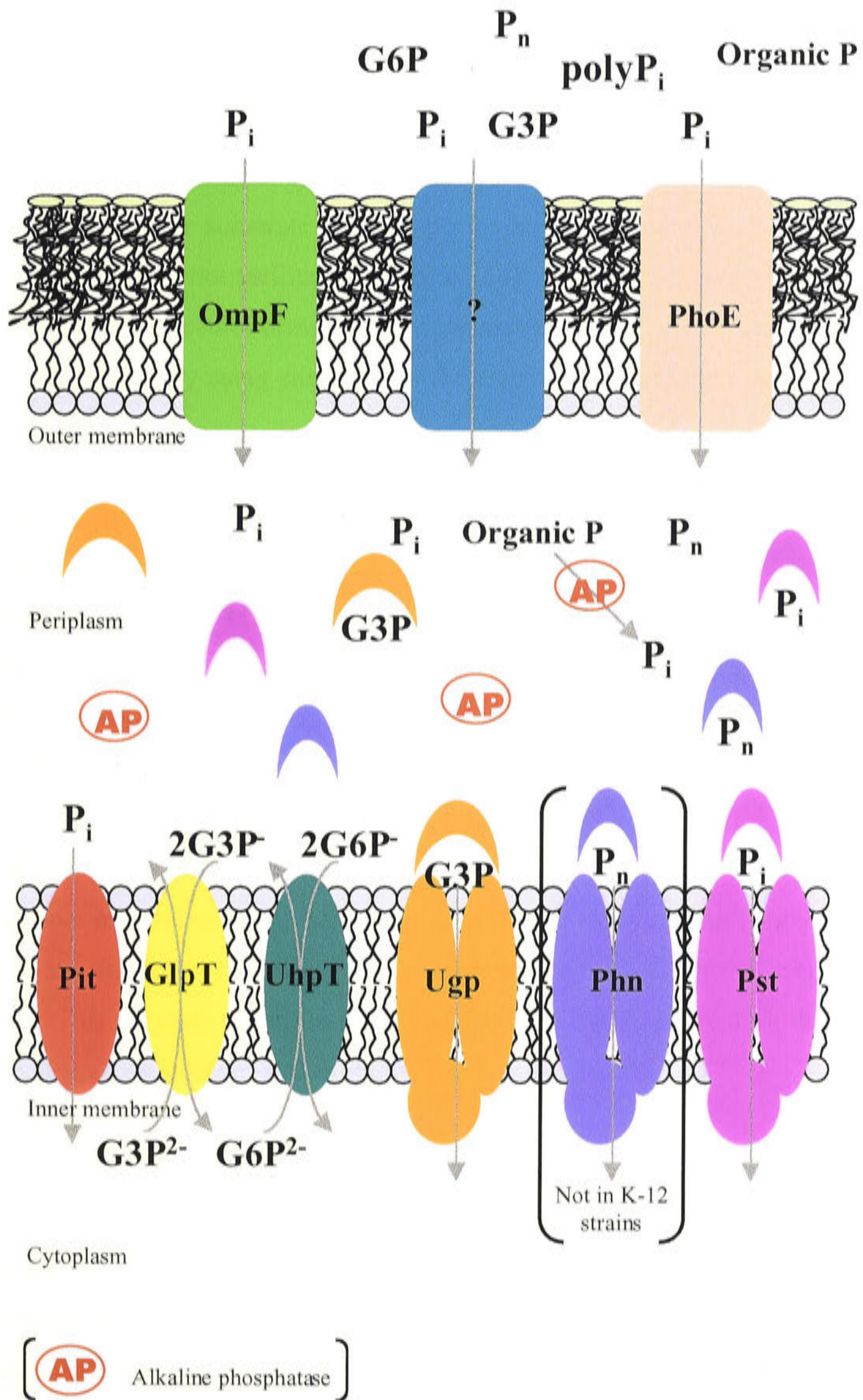
**Phosphate transport at high inorganic phosphate ( $P_i$ ) concentrations**





**Figure 1.1B**

Phosphate transport at low inorganic phosphate ( $P_i$ ) concentrations





Proteins inserted in the membrane assist in the passage of nutrients, wastes and information and those exposed at the cell surface have also become receptors for pathogenic agents such as bacteriophages and bacteriocins (130).

About 50% of the outer membrane consists of protein, either as lipoproteins anchored to the membrane by N-terminally attached lipids, or in the form of integral membrane proteins. In particular, this membrane contains a large number of pore forming proteins called porins, which allow the diffusion of ions and small hydrophilic molecules. Several porins have little solute specificity and behave as general diffusion channels (184), while other substrate specific porins have binding sites for nutrients such as maltose and carry out facilitated diffusion (14). TonB-dependent receptors concentrate large nutrient molecules present at low concentrations in the external environment, such as vitamin B12, by using energy from the electrochemical potential of the cytoplasmic membrane.

Inorganic phosphate (Pi) can diffuse through many of the general porins in the outer membrane. Under conditions of limiting Pi the general porin PhoE is induced. PhoE has a slight preference for negative ions but transports only molecules with a molecular mass below 600 daltons. Therefore larger polyphosphates and phosphorylated compounds diffuse through the outer membrane through proteins that have yet to be determined (204, 205).

### **1.2.2 The periplasm**

The outer membrane is attached by proteins to an underlying peptidoglycan layer which provides rigidity to the cell while being permeable to all nutrients and wastes (129, 248). This polymer of amino sugars acts as a cell wall, preventing the rupture of the cytoplasmic membrane in environments of low osmolarity.

The space between the outer and inner membranes, called the periplasm, contains the peptidoglycan layer and a gel composed of proteins and polysaccharides (97). The polysaccharides help to buffer the cell against the variable ionic and osmotic environments it may be exposed to. Periplasmic proteins are involved in the growth and maintenance of the outer membrane, the inactivation of toxins, chemotaxis and the processing of nutrients. The inner cytoplasmic membrane is impermeable to many

nutrients that can cross the outer membrane. Therefore many compounds, such as polyphosphates and organophosphates (other than the *sn*-glycerol phosphates and hexose phosphates), must be broken down by periplasmic enzymes into transportable nutrients. The periplasm contains a number of phosphatases that release Pi from phosphorylated compounds. Some, such as acid phosphatase (*appA*) and 2'-3' cyclic phosphodiesterase (*cdpB*) are present in the periplasm under all conditions, while others like alkaline phosphatase (*phoA*) are induced under conditions of Pi limitation. PhoA is important in Pi metabolism as it has a wide substrate specificity. The periplasm also contains specific binding proteins that efficiently deliver substrates present at low concentrations to their relevant transporters present in the inner membrane. Low concentrations of Pi induce the expression of substrate binding proteins for several high affinity transporters such as Pst (Pi transporter) and Ugp (G3P transporter) (Figure 1.1B) (204).

### **1.2.3 The cytoplasmic membrane**

The inner cytoplasmic membrane consists of a double layer of phospholipid, rather than the asymmetric lipopolysaccharide/phospholipid bilayer of the outer membrane. It is much less permeable than the outer membrane. This allows nutrients to be concentrated across the inner lipid bilayer (184). The membrane is permeable only to water, oxygen, carbon dioxide and small hydrophobic molecules. This impermeability is maintained by specific integral membrane proteins which are required to transport ions and nutrients into the cytoplasm (46). Approximately 70% of the mass of the cytoplasmic membrane is protein, and these proteins have roles in nutrient and waste transport, energy generation, transmembrane signaling, chemotaxis and the translocation of envelope macromolecules (182). The impermeability of this membrane to protons and other ions allows the development of an electrochemical potential across this barrier which can be used directly by proteins to energise the concentration of nutrients in the cell, or it may be used to store chemical energy via the synthesis of ATP. Membrane proteins do form ion channels, but these are gated to preserve the integrity of the membrane potential (203). The cytoplasmic membrane is much more fluid than the outer membrane, due to the high phospholipid content of the lipid bilayer, allowing proteins to freely diffuse sideways within the plane of the membrane (243).

*E. coli* has four major systems which can transport Pi through its cytoplasmic membrane, with two of these being highly specific for Pi (211). The high affinity phosphate-specific transporter (Pst) is induced when Pi is scarce (170, 289), while the lower affinity inorganic phosphate transporter (Pit) is the general phosphate transporter (211). The *sn*-glycerol 3-phosphate transporter (GlpT) and the hexose phosphate transporter (UhpT) accept Pi as a low-affinity analog of G3P (91) and glucose-6P (199, 295) respectively. In the absence of Pst and Pit neither of these transporters can support cell growth when supplied with Pi. Under physiological conditions these two systems are induced by the presence of their respective organophosphates in the extracellular environment, and can use these to fill the cell's carbon requirements. While GlpT can reportedly use G3P as the sole source of Pi (237, 249), there are conflicting reports on the ability of UhpT to do the same when supplied with glucose-6P Pi (101, 249). A second *sn*-glycerol-3-phosphate transporter (Ugp) induced by the *pho* regulon under conditions of Pi limitation, is probably the preferred mechanism for supplying Pi from G3P (Figure 1.1B) (302).

Many bacteria can also use phosphonates (Pn) as the sole source of phosphate under conditions of Pi limitation. Phosphonates have a direct carbon-phosphorus bond in place of the carbon-oxygen-phosphorus ester linkage found in organophosphates. *E. coli* contains a binding protein dependent Pn transporter that is under *pho* regulon control ((280) and references therein). However Pn utilisation is cryptic in *E. coli* K12, due to a frameshift mutation in the *phnE* (EcoK) gene, so Pn cannot serve as a source of phosphate in this strain (160, 282).

## ***1.3 The phosphate transporters of Escherichia coli K-12***

### ***1.3.1 PhoE***

Phosphoporin (PhoE) is an outer membrane porin which allows the diffusion of small hydrophilic molecules, (less than 600 daltons), through a water filled channel. The X-ray crystal structure of PhoE has been solved, showing it to form stable trimers, each consisting of a 16-stranded anti-parallel  $\beta$ -barrel with one strand looping into the



internal channel to form a constricting eyelet (42). This porin is weakly anion selective (13), and this charge specificity is provided by positively charged residues on the eyelet, within the channel and near the channel mouth (11, 15, 267). PhoE expression is induced by the *pho* regulon under conditions of phosphate limitation (See Section 1.4 for details), increasing the diffusion rate of anions, such as small polyphosphates, through the outer membrane. However these molecules, plus various organophosphates, also enter the periplasm via other proteins which have yet to be determined (205). While voltage gated closing of this porin can be induced (235), it is not certain that this occurs under physiological conditions (130, 183, 239).

### **1.3.2 Pit**

Studies on Pit in wild type cells (170, 290), Pst deficient mutants (212), spheroplasts (211) and membrane vesicles (131) suggest that Pit is a constitutive system which catalyses an electrogenic symport of Pi and protons. Transport of these substrates is coupled exclusively to the proton motive force and is completely abolished by the uncoupler CCCP (211). Divalent cations, such as  $Mg^{2+}$  or  $Ca^{2+}$ , were shown to be essential for Pit activity (218) and experiments by van Veen *et al* (268) with membrane vesicles indicate that Pi is transported as a neutral metal phosphate ( $MeHPO_4$ ) chelate formed by complexation of divalent metal ions, such as  $Mg^{2+}$ ,  $Ca^{2+}$ ,  $Mn^{2+}$ , or  $Co^{2+}$  with  $HPO_4^{2-}$ . The complex is symported with a proton stoichiometry close to one (268). This cotransport of divalent metal ions is supported by the recent identification of a *pitA* mutant that accumulates reduced amounts of zinc(II), conferring resistance to toxic external concentrations of zinc (12).

Efflux and homologous exchange of metal phosphate can occur under particular conditions, but there is no mixed exchange of metal phosphate for Pi, glycerol-3-phosphate or glucose-6-phosphate. As expected, efflux was inhibited by the proton motive force, with the deprotonation of the carrier on its outer surface the limiting step (268). Interestingly, Beard *et al* (12) suggest that PitA may also play a role in  $Zn^{2+}$  efflux when the ion reaches toxic external concentrations ( $>2mM$ ). The strictly aerobic bacteria, *Acinetobacter johnsonii*, which accumulates polyphosphate, becomes enriched in wastewater treatment plants with alternating aerobic and anaerobic conditions by generating a proton-motive force through the coupled efflux of  $MeHPO_4$  and protons

(269). During exchange the Pit carrier is recycled in a protonated form. Exchange was inhibited by the membrane potential only, presumably by preventing the outward movement of this positively charged ternary complex (268).

Pi transport in *E. coli* has also been shown to be increased by  $K^+$  (170). Russell and Rosenberg (217, 218) showed that  $K^+$  transport indirectly stimulated Pit. Electrogenic  $K^+$  uptake coupled to  $H^+$  expulsion helps cells to maintain an alkaline pH in the cytosol, preventing the internal deprotonation of Pit being inhibited by high  $H^+$  concentrations.

Pit is the low affinity transporter of Pi in *E. coli*, having an apparent  $K_m$  ( $K_m^{app}$ ) for Pi which is between 40- to 200-fold **higher** than the Pst system ( $K_m^{app} \sim 0.2 \mu M$  (211, 290)). The kinetic constants for Pit have been measured under a variety of conditions producing a  $K_m^{app}$  of between 12 and  $38 \mu M$ . Originally the  $K_m^{app}$ s of Pit and Pst were measured simultaneously in wild type cells, giving Pit a  $K_m^{app}$  of  $9.2 \mu M$  (pH 7.0,  $400 \mu M$   $MgSO_4$ ) (170). When Pit was genetically isolated from Pst a  $K_m^{app}$  of  $25 \mu M$  (pH 6.9,  $1 mM$   $MgSO_4$ ), or  $38.2 \mu M$  (pH 7.5,  $8 mM$   $MgSO_4$ ) was observed in whole cell assays (211, 290). The  $K_m^{app}$  obtained for Pit in membrane vesicles was  $11.9 \mu M$  (pH 7.0,  $10 mM$   $MgSO_4$ ). However at pH 7.0 only 36% of  $Mg^{2+}$  is complexed with  $HPO_4^{2-}$  giving  $MgHPO_4$  a  $K_m^{app}$  of  $8.8 \mu M$  (268).

Elvin *et al* (59) isolated a Pit gene that confers Pi uptake on a Pi auxotrophic *pit pst* mutant in a dosage dependent manner. The single copy plasmid pCE26 expressing Pit could restore wild type activity, showing *pit*<sup>+</sup> to be dominant over *pit*, while the multicopy plasmid pCE27 could increase uptake 8- to 10-fold. While there is no direct proof that the Pit system has a single membrane component, Elvin *et al* (59) suggest that this amplification would not be expected if Pit was not the sole component of this system (60).

There are now over 50 members of the PiT transporter family, which is defined by homology to domain PD1131 (ProDom database – (37)). All characterised PiT members transport Pi, with the exception of *cysP*, an inorganic sulfate transporter (167). This family is described in more detail in Section 1.6.

### 1.3.3 *Pst* system

If the external Pi concentration falls below around 0.1mM, the high-affinity phosphate-specific transport (*Pst*) system is induced by the *pho* regulon (291). This has a  $K_m^{app}$  for Pi of around 0.2  $\mu$ M (211, 290). The *Pst* system is a periplasmic binding protein-dependent transport system composed of four proteins. *PstS* is the periplasmic binding protein, while *PstA* and *PstC* are integral membrane proteins that are proposed to have 6 transmembrane (TM)  $\alpha$ -helices each. *PstB* is a hydrophilic protein associated with *PstA* and *PstC* on the cytoplasmic side of the membrane that contains an ATP binding fold and is thought to be important in energy generation for the system (31, 74, 284). The *pst* operon contains genes for these four proteins plus *phoU*, which encodes a protein required for repression of the *pho* regulon but which is apparently not required for Pi transport through the *Pst* system (2, 254, 259). An intact *Pst* system is also thought to be required for repression of the *pho* regulon, and may play a role in detecting the external Pi concentration. Many *Pst* mutants derepress the *pho* regulon as well as prevent Pi uptake. However, there are single amino acid substitutions within *PstA* and *PstC* membrane proteins which separate these functions, inactivating Pi transport without affecting *pho* regulon control (43, 44). This is discussed further in Section 1.4.

Site-directed mutagenesis studies indicate R-220 in *PstA* and R-237 and E-241 in *PstC* are required for Pi transport and that these residues interact with each other, possibly by forming a Pi relay system. Paired proline residues may also play an important role in the configuration of the protein. A double proline to alanine *PstC* mutant caused a permanently 'open' conformation in *Pst* and the equivalent double proline *PstA* mutant represented a permanently 'closed' configuration (284).

The X-ray crystal structure of the *PstS* binding protein shows that Pi is tightly held in place by 12 hydrogen bonds and is totally devoid of water of hydration. Eleven of these hydrogen bonds are donor groups (mainly from peptide backbone NH and side chain OH groups, plus one arginine side chain ) and one aspartic carboxylate side chain is an acceptor group, recognising a proton on either the monobasic or dibasic Pi.



Therefore there are no strong charge interactions (155). Whether other Pi transporters bind Pi in a similar manner is not known.

### **1.3.4 *GlpT***

*sn*-glycerol 3-phosphate is a precursor for membrane phospholipid biosynthesis, and can also be used by the cell as the sole source of carbon or phosphate. When used as a carbon source G3P is taken up exclusively by the GlpT transport system, whose genes are induced by the presence of glycerol or G3P as part of the *glp* regulon. Repression occurs via the *glp* repressor GlpR in the absence of glycerol or G3P, and through catabolite repression mediated by CAP-cAMP (56).

GlpT is a Pi-linked antiporter which catalyses the uptake of G3P by electroneutral exchange of G3P and Pi, although it is also able to exchange Pi:Pi or G3P:G3P. Other substrates identified include glycerol-2-phosphate, arsenate and phosphonomycin (61). It is a chemiosmotic transporter, concentrating substrate by use of the electrochemical potential of the cell. GlpT may be able use G3P as the sole source of phosphate, as well as carbon (237). Maloney *et al* (164) proposed a mechanism whereby GlpT can have a net import of Pi. Antiporters recognise both the monovalent and divalent forms of their substrates. As the cytoplasm is usually more alkaline than the periplasm, two monovalent G3P can be imported, while one divalent G3P is exported, giving a net import of a divalent G3P and two protons (164). However, under conditions of Pi limitation Ugp, a G3P transporter induced by the *pho* regulon, would be expected to be the normal route of G3P import (Figure 1.1B) (302).

GlpT is a member of the Major Facilitator Superfamily (227) and has a typical topology of 12 TM  $\alpha$ -helices, which has been supported by PhoA and LacZ protein fusion experiments (78). The protein has been over-expressed, purified and actively reconstituted into liposomes (6). While this protein has been found to operate as a monomer, the native protein has been identified as a homo-oligomer (141).

Several *E. coli* strains that were originally selected for wild type Pit activity by their sensitivity to arsenate, such as strain C600, underwent rapid Pi exchange when supplied with carbon sources other than glucose. These were found to have a *glpR* mutation that

causes arsenate sensitivity as well as rapid Pi exchange through constitutive expression of GlpT (61). This occurs because GlpT can substitute arsenate for Pi in its catalysis of G3P:Pi antiport or Pi:Pi exchange. Glucose inhibits this response by causing catabolite repression of the *glpT* gene. Although wild type Pit is sensitive to arsenate, constitutive GlpT confers arsenate sensitivity on *pit* mutant strains that would otherwise be resistant, such as strain HR159 (*pit zhg::TN10 glpR*). Thus, isolation of Pit mutants using the auxotrophic requirement for G3P has proven to be a less ambiguous selection criteria than arsenate resistance (60, 249).

### **1.3.5 UhpT**

The hexose phosphate transporter has a high degree of sequence similarity to GlpT, and is also an obligate anion antiporter whose main function is assumed to be the import of sugar phosphates for use as a carbon source (164). UhpT catalyses the uptake of a number of organophosphate compounds including the phosphate esters of hexoses, pentoses, heptoses, amino sugars and sugar alcohols by electroneutral exchange with Pi (119). Pi:Pi exchange also occurs, as in the similar GlpT system (247). Thus active transport is directly coupled to the downhill movement of Pi (119). There are four genes in the *uhp* operon, which are all essential for *uhp* expression. Induction through the *uhp* regulatory system (UhpA, UhpB) is activated by external glucose 6-P. A third protein, UhpC, is also involved in signaling the presence of glucose 6-phosphate (104). The promoter contains multiple sites for UhpA and also contains a single binding site for the cAMP-dependent transcriptional activator CAP (172). UhpT is a member of the Major Facilitator Superfamily (227) and contains 12 TM segments, based on hydropathy profiles and a number of *phoA* gene fusions (149).

It has previously been reported that UhpT supplied with organophosphate can provide the cell's sole source of phosphate (249). This would be possible through a mechanism similar to that described for GlpT, using glucose 6-P instead of G3P (164). However Hoffer *et al* (101) have been unable to grow *pit pst* strains with glucose 6-P as the phosphate source and either glucose or glycerol as the carbon source. (This strain does grow with G3P as the phosphate source, which could be transported through either GlpT or Ugp.)

### **1.3.6 *Ugp* system**

The Ugp transporter catalyses the uptake of G3P and glycerophosphoryl diesters (the diacylation products of phospholipids) and is induced by the *pho* regulon under conditions of Pi limitation (4, 237). When G3P is taken up exclusively by Ugp it can be used as the sole source of Pi but not as the sole source of carbon (26). However Kasahara *et al* (120) found that the *ugp* operon contains a second promoter with a cAMP-CRP binding site that is induced by carbon starvation, suggesting Ugp plays a role in carbon metabolism. Ugp activity increases the internal Pi concentration upon degradation of transported G3P, and is inhibited by high intracellular Pi concentrations (302).

Ugp is a periplasmic binding protein-dependent transport system. It consists of four proteins, which are a substrate binding protein (UgpB) two integral membrane components (UgpA and UgpE), and an ATP-binding subunit (UgpC) (4, 302).

## **1.4 Regulation of inorganic phosphate transport in *Escherichia coli***

When inorganic phosphate (Pi), the preferred phosphorus source for *E. coli*, is in excess the low affinity Pit transport system imports the phosphorus requirements of the cell. Genes involved in the high affinity transport of Pi and the transport and assimilation of other phosphate sources are only activated when the Pi concentration in the culture medium becomes low, falling below 0.1-1mM (291). The induction of many of these genes is mediated through the *pho* regulon (280). Recent 2D protein experiments, which examined only proteins with pIs less than seven, have shown that the *pho* regulon induces about 118 proteins and represses around 19 proteins under conditions of Pi limitation (271). Over 38 of these genes have been identified, with most being involved in phosphate transport and/or assimilation (86, 281).

The *pho* regulon is controlled by the *phoB-phoR* operon which has identity to a number of two-component regulatory systems that sense and respond to changing environmental stimuli in bacteria (210). Pi-dependent activation requires the histidine kinase PhoR to autophosphorylate and then phosphorylate the transcriptional regulator PhoB. A phosphorylated PhoB dimer then activates genes in the *pho* regulon by



binding to a *pho* box consensus sequence found within the regulatory regions of these genes, which is also present in the *phoBR* operon, allowing auto-amplification of the system ((161) plus references therein, 83, 191, 242, 280). When the Pi concentration is restored to normal, interactions between the Pst system, PhoU and PhoR catalyse dephosphorylation of PhoB, and transcription is suppressed (Figure 1.2) (181, 282). While the mechanism of this inhibition has yet to be determined, it has been shown that the majority of mutations in either PhoU or any of the four Pst proteins result in high constitutive *pho* regulon gene expression. It is believed that the Pst system plays a role in the detection of the external Pi concentration to initiate inhibition of the *pho* regulon, and that this repression is independent of Pi transport, as specific mutations in PstA or PstC abolished Pi transport without affecting repression (43, 44). PhoU is a product of the *pstSCAB-phoU* operon, and is loosely associated with the inner membrane but has no role in Pi transport (250). An extremely deleterious growth defect caused by a *phoU* deletion was found to be a consequence of high-level Pst synthesis (84). External Pi was proposed to provide the signal for *pho* regulon activation and inhibition because derepression of the *pho* regulon by low external Pi concentrations is not accompanied by a lowering of the cell's internal Pi concentration (280). Recently, this idea has been challenged by evidence that overproduction of either PitA or PitB in the absence of the PstS phosphate binding protein can restore regulation of the *pho* regulon. Hoffer *et al* (100) propose that PhoU associated with the Pst system may detect Pi concentration internally by direct access to Pi that enters the cell via this system. PitA or PitB transport Pi as a neutral complex associated with divalent cations (268), which normally may not be available to PhoU. However upon overproduction of Pit in the PstS mutant, this Pi may become accessible to PhoU (Figure 1.3).

PhoB consists of an N-terminal phosphorylation domain and a C-terminal DNA-binding/transactivation domain, which interacts with the RNA polymerase  $\sigma^{70}$  subunit and the DNA binding sequence TGTCA (159). There is evidence that these domains interact in the inactive, unphosphorylated PhoB (58). PhoB is activated by phosphorylation, which disrupts this interaction and allows the dimerisation of PhoB through the N-terminal domains (65). The crystal structure of the N-terminal domain has recently been determined, showing it is similar to other response regulators from the two component regulator family (246). The three dimensional nuclear magnetic

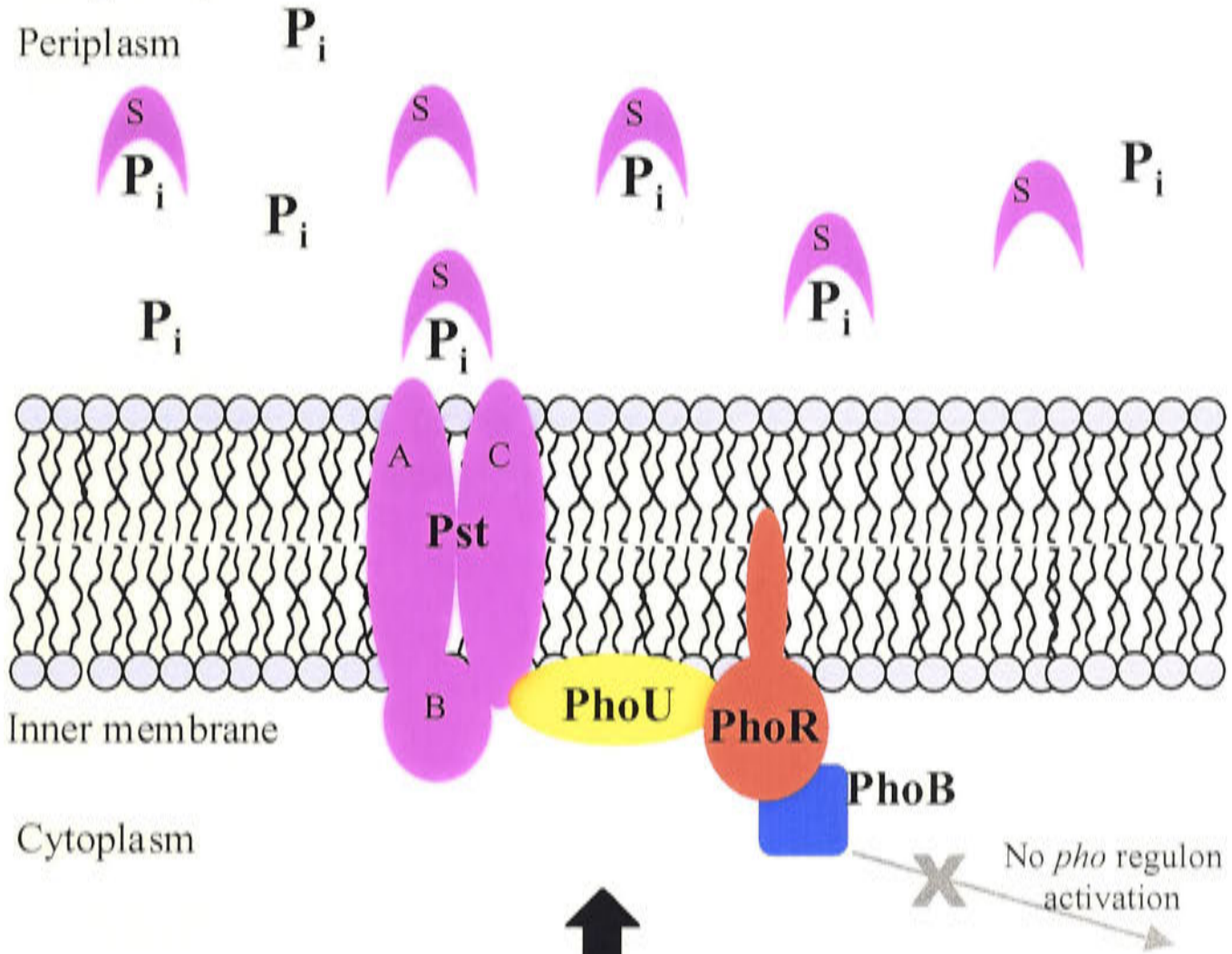
## Figure 1.2

### Postulated mechanism for the involvement of the Pst system in the repression of the *pho* regulon.

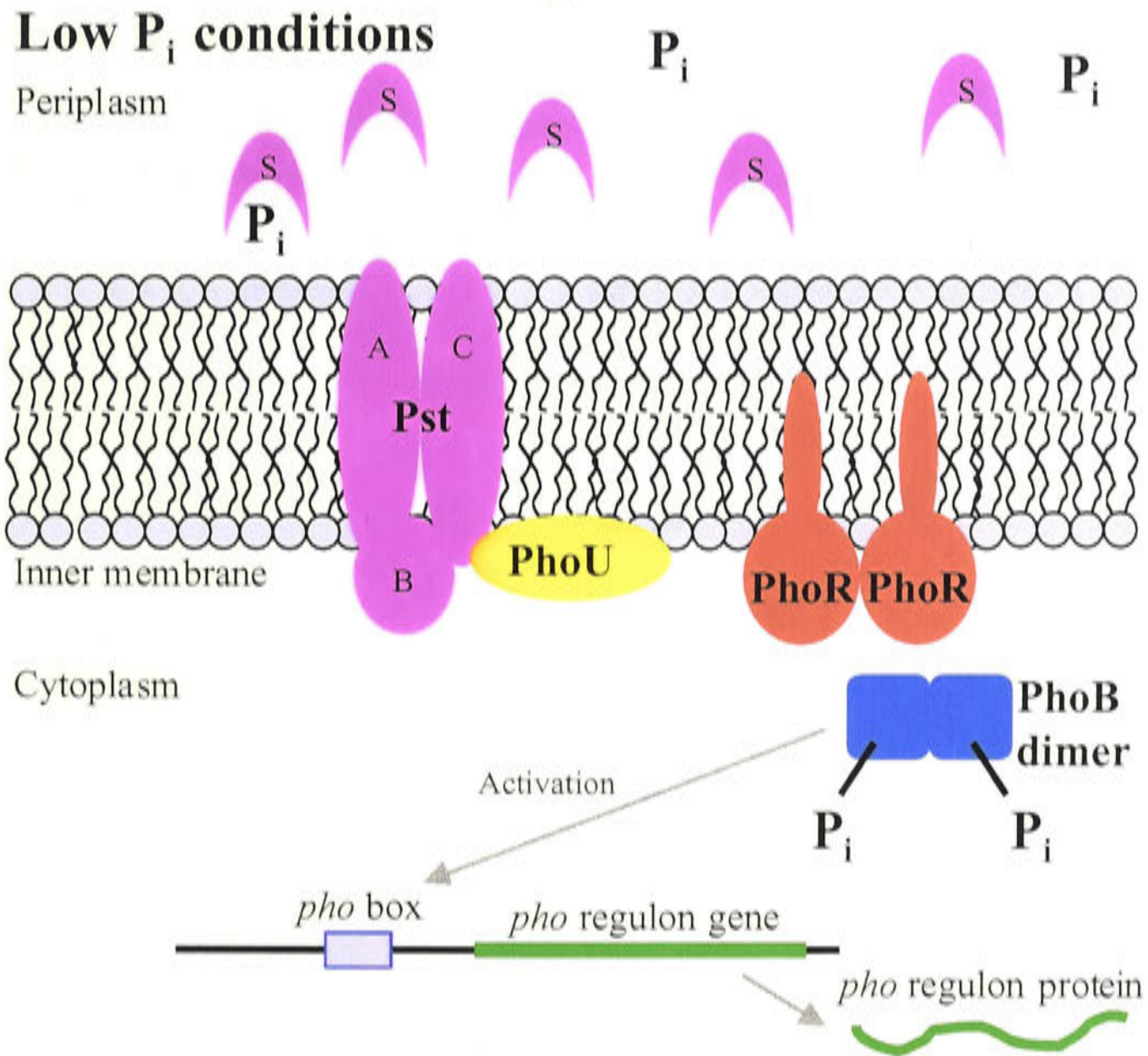
This diagram is modified from Wanner, B. L. (280).

Under low Pi conditions Pi-dependent activation requires the histidine kinase PhoR to autophosphorylate and then phosphorylate the transcriptional regulator PhoB. A phosphorylated PhoB dimer then activates genes in the *pho* regulon by binding to a *pho* box consensus sequence found within the regulatory regions of these genes. When the Pi concentration is restored to normal, interactions between the Pst system, PhoU and PhoR catalyse dephosphorylation of PhoB, and transcription is suppressed.

## High $P_i$ conditions



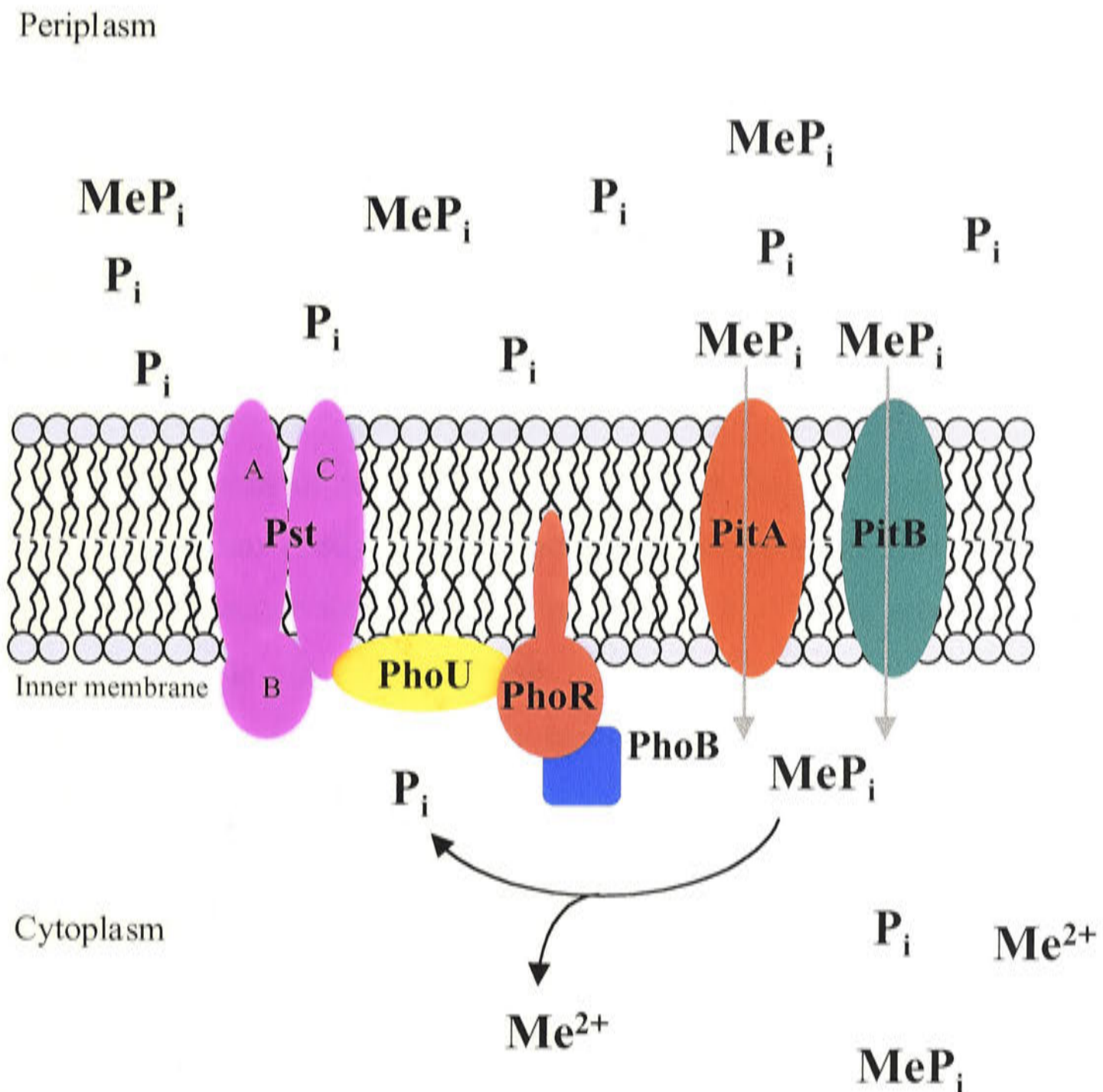
## Low $P_i$ conditions





**Figure 1.3**  
**Proposed control of the *pho* regulon by internal inorganic phosphate ( $P_i$ ) concentrations.**

Overproduction of either PitA or PitB in the absence of the PstS phosphate binding protein can restore regulation of the *pho* regulon. Hoffer *et al* (100) propose that PhoU associated with the Pst system may normally detect  $P_i$  concentration internally by direct access to  $P_i$  that enters the cell via this system. PitA or PitB transport  $P_i$  as a neutral complex associated with divalent cations, which usually may not be available to PhoU. However upon overproduction of Pit in the PstS mutant,  $P_i$  transported through either Pit protein may become accessible to PhoU.



resonance (NMR) structure of the C-terminal domain has also recently been solved, allowing interactions of this domain with the *pho* box to be proposed and modeled (191).

Genes activated by the *pho* regulon contain an 18 base consensus *pho* box normally located 10 bases upstream from the -10 region of the promoter which effectively replaces the -35 promoter region ((191) and references therein). The generally accepted sequence is 5'-CTGTCATA(A/T)A(T/A)CTGTCA(C/T)-3', which consists of two binding sites for PhoB separated by an AT-rich region. NMR studies using the PhoB binding/transactivation domain and a variety of DNA molecules confirmed that two PhoB molecules bind in tandem on the *pho* box. PhoB recognises the TGTCA sequence in the major groove of the DNA with an  $\alpha$ -helical region and contacts the 3' site out of this TGTCA sequence within the minor groove in a non-specific but stabilising manner (191). Each PhoB within the dimer binds to the DNA sequence about 11 nucleotides apart, which is slightly greater than one turn in the standard B-form DNA conformation, enhancing DNA bending in this region. It is proposed that phosphorylated PhoB and the  $\sigma^{70}$  subunit from RNA polymerase holoenzyme are positioned close to each other (159).

The promoter regions of several genes induced by the *pho* regulon contain multiple copies of the *pho* box (280). The *psiE* promoter contains two *pho* boxes, one over the -35 region and a second four bases upstream of the first. Both sequences are protected from DNaseI in the presence of PhoB. A binding site for cAMP-CRP covers the putative -35 region and overlaps the downstream *pho* box. In low glucose, low phosphate conditions, when both phosphorylated PhoB and cAMP-CRP would be present in high concentrations, *in vivo* data suggests PhoB binding predominates (124). The *phoE* promoter also contains two *pho* boxes in a similar arrangement to the *psiE* promoter. Both are required for full expression of PhoE during phosphate starvation. Optimisation of the -10 region created a mutant with high basal levels of expression under high phosphate conditions. This expression was partly dependent on the presence of the *phoB* gene. Only the downstream *pho* box was necessary for full *phoE* induction of this mutant under low phosphate conditions. Thus regulation of this promoter is carefully balanced by deviations from the optimal -10 region and by the

presence of several copies of the *pho* box, which enhance expression under phosphate starvation (236).

Most research has concentrated on the genes activated by the *pho* regulon. The mechanism of repression for *pho* regulon genes has yet to be determined. Significant evidence for an inhibitory role for PhoB has only recently been published and will be discussed in detail in Chapter 5.

## **1.5 Mechanisms of membrane transporters**

### **1.5.1 Classification of membrane proteins by their mechanism of transport**

Many transporters in the cytoplasmic membrane couple transport with the input of energy, as they often function to concentrate compounds within the cell. There are four basic classifications for these transporters based on the presence or nature of the energy source. These are:

- 1 facilitated diffusion
- 2 secondary active transport
- 3 group translocation
- 4 primary active transport

(See Figure 1.4)

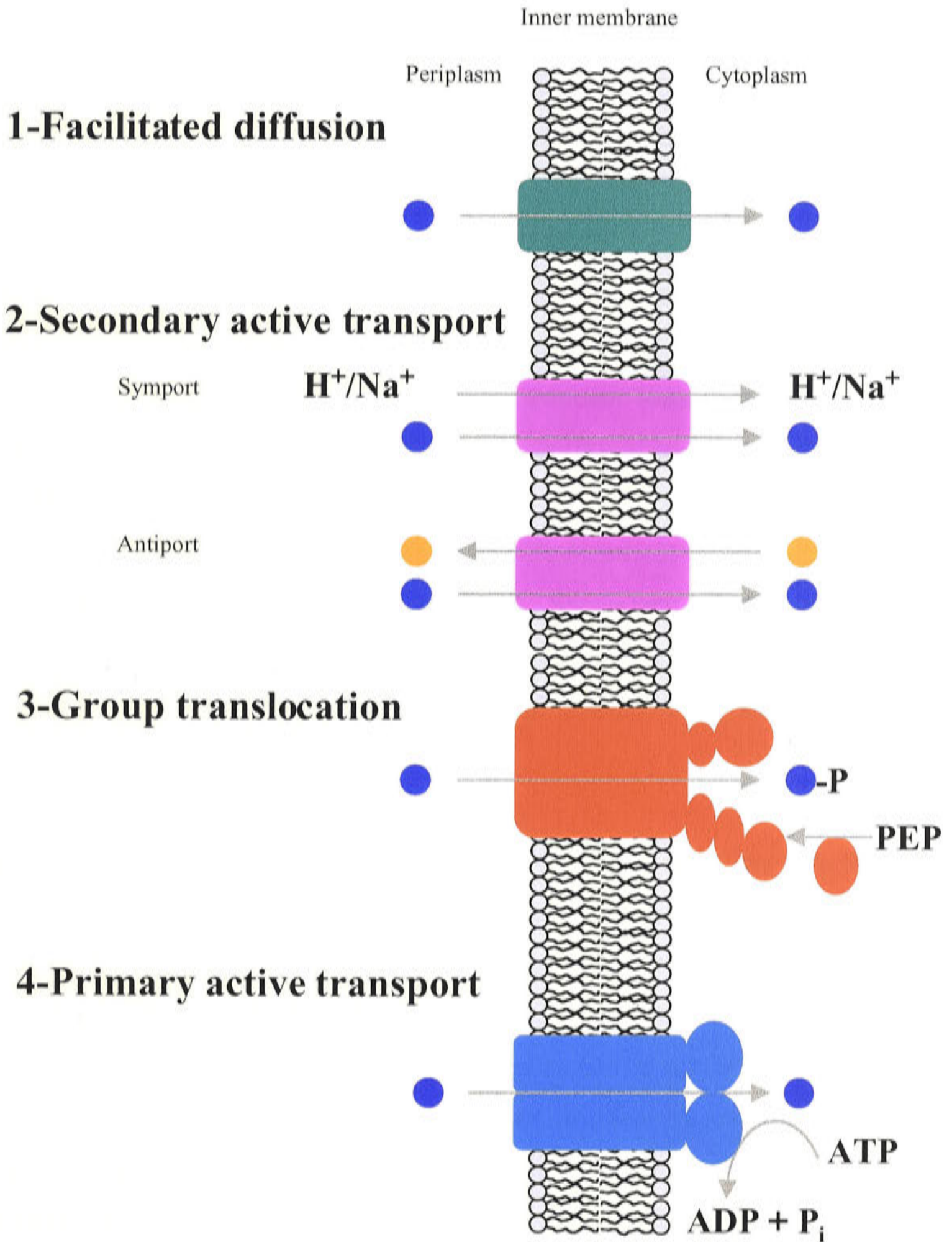
Facilitated diffusion allows a compound to diffuse down its concentration gradient until equilibrium concentrations are achieved. There are very few of these transporters in *E. coli*, with the glycerol facilitator GlpF being one example (93, 187). These are sometimes classified as the uniporter subgroup within secondary transporters. Many secondary active transporters couple the uphill transport of a substrate to the electrochemical potential of the membrane by the downhill transport of an ion. This may occur by symport (e.g. the influx of H<sup>+</sup> or Na<sup>+</sup> with substrate) or antiport (e.g. the export of Pi coupled to the import of substrate). Group translocation systems tightly couple the import and phosphorylation of a sugar to the breakdown of intracellular phosphoenolpyruvate (PEP). The only family in this category is the



**Figure 1.4**

**Mechanisms of cytoplasmic membrane transporters.**

Facilitated diffusion is a passive process, while the remaining mechanisms of transport are active, requiring energy input from either ATP hydrolysis or a substrate moving down its electrochemical gradient.



phosphoenolpyruvate (PEP) phosphoryl transfer-driven group translocators. Primary active transport uses ATP hydrolysis to energise transport. One example is the periplasmic binding protein dependent systems. These translocators have high affinity for their substrates, particularly when compared to secondary transport systems specific for the same substrate. For example, Pst is an active transporter ( $K_m \sim 2\mu\text{M}$ ), while Pit is a secondary transporter ( $K_m \sim 25\mu\text{M}$ ). Substrates are bound by a periplasmic binding protein and delivered to a membrane transporter that contains four modules. These have two membrane spanning regions and two nucleotide binding domains which may be fused as one, two or four peptides. Other types of primary active systems include pumps such as the  $\text{Na}^+/\text{ATPase}$  and ABC transporters without binding proteins.

### **1.5.2 Secondary active transporters**

Secondary transporters use an electrochemical potential to drive the uptake or efflux of substrates across the membrane. Bacteria develop an electrochemical potential by using chemical, light or redox energy to translocate ions across the cytoplasmic membrane. Usually protons are exported, but a number of bacteria can also use sodium ions. The difference in concentration of these ions on each side of the membrane creates an electrochemical gradient, which has an electrical potential, and a chemical potential (pH for protons, pNa for sodium ions). Secondary transport systems convert the electrochemical energy of one solute into the electrochemical energy of another solute. This may be done by coupling the downhill transport of protons or sodium ions with the uphill transport of solute (symport).

There are three groups of secondary transport proteins. Uniporters catalyse the transport of only one solute down an electrochemical gradient and these solutes are usually neutral or positively charged. Symporters couple the transport of two or more solutes (with one usually being a proton or sodium ion) in the same direction. Antiporters couple the import of one solute to the export of another. Usually one solute is a proton or sodium ion, but other ions such as phosphate are also used (e.g. the GlpT Pi-linked antiporter). Under certain conditions secondary transport systems can reverse their direction of transport, building up the proton (or sodium) motive force by letting the substrate move down its concentration gradient.

Secondary transporters often import nutrients such as sugars, amino acids and ions, and account for approximately 40% of the active transport in *E. coli* and *S. typhimurium* (163). Currently the Transport Commission Database, set up by the Saier Laboratory Bioinformatics Group (University of California, San Diego), has 80 different families listed in the porter subclass of secondary transporters. The largest family is the Major Facilitator Superfamily, which currently contains over 1000 sequenced members, including lactose permease. Another family is the Inorganic Phosphate Transporter (PiT) family, to which PitA and PitB belong.

## **1.6 The PiT family**

The PiT family contains transporters belonging to a wide variety of organisms such as Gram-positive and Gram-negative bacteria, archaea, yeast, fungi, plants and animals (121, 174, 226). The Transport Commission has defined the PiT family as containing a signature sequence (226) and this sequence forms part of domain PD1131 (ProDom database – (37)) which is usually duplicated within each protein (229). Most PiT family members which have been functionally characterised catalyse inorganic phosphate (Pi) uptake energised by either H<sup>+</sup> or Na<sup>+</sup> symport – depending on which ion can be used to develop an electrochemical membrane potential in the particular organism (21, 122, 168, 192, 211, 270, 289, 293). PitA and PitB from *Escherichia coli* have been shown to import Pi as a neutral divalent metal ion complex (268). Currently, only *cysP* from *Bacillus subtilis*, which is a sulfate permease, has a different substrate profile. Inorganic sulfate is structurally quite similar to Pi and the ability of *cysP* to also transport Pi has yet to be determined (167). While the majority of active transport systems in bacteria, fungi and plants use the proton-motive force, several Pit proteins in these organisms can be coupled to the sodium-motive force, sometimes exclusively (21, 168). The sodium-motive force is used by animal PiT genes (122, 192).

Several PiT family phosphate transporters from varied organisms seem to serve as the general Pi importer for Pi used to sustain normal cellular function. These include PitA from *Escherichia coli*, type III NaPi transporters (PiT-1 and PiT-2) in mammalian tissues and Pht2;1 in the vascular plant *Arabidopsis thaliana* (211, 122, 47, 192, 293). Other members are more specialised, such as Pho89 from *Saccharomyces cerevisiae*, and Pho4 from *Neurospora crassa* which require alkaline conditions to function (168,



274). Thus Pit regulation can be quite varied. Pit may be constitutive (289), or expression may be stimulated (21, 34, 121, 168) or repressed (9) by phosphate starvation. Modulation of activity may also be caused by pH (168, 274), by molecules such as insulin-like growth factor 1 (194) or by the binding of retroviral envelope glycoprotein (133, 174, 190). The mammalian transporter PiT-2, which seems to exist in active or inactive conformations, associates with the actin cytoskeleton network and forms high molecular-weight complexes when Pi concentrations are low. This suggests post-translational changes may be important for function (207).

Some species, such as humans, rats, the nematode *Caenorhabditis elegans* and a growing number of prokaryotes such as *Pseudomonas aeruginosa*, and *Listeria monocytogenes* have several different but related Pit transporters (122, 272) (NCBI Blast analysis performed at SIB (1)). In rats PiT-1 and PiT-2 have distinctive, but overlapping, expression in a wide variety of tissues, and over-expression of PiT-2 in rat fibroblasts severely repressed PiT-1 synthesis transporters (122) (121). Genetically unrelated Pi transporters are also present in these organisms, so analyses of organisms with multiple Pit genes and multiple Pi transporters are revealing complex and subtle interactions.

## **1.7 The structure of membrane proteins**

### **1.7.1 The basic structural motifs of membrane proteins**

Membrane proteins need to traverse a lipid bilayer that has a strongly hydrophobic core, (i.e. a very low dielectric constant), where it is energetically costly to place charged and polar residues. Polypeptide chains, which can hydrogen bond with water in an aqueous environment, need alternate sources of hydrogen bonds in a lipid environment. There are two conformations of polypeptide chains that maximise the number of hydrogen bonds -  $\alpha$ -helices, which bond internally, and  $\beta$ -sheets, where the bonds form between different strands (214). Sequence analysis of cytoplasmic membrane proteins has found that many have regions containing hydrophobic residues which also have a high propensity for forming  $\alpha$ -helices that are long enough to cross the cytoplasmic membrane. Recently determined atomic structures of the prokaryotic

potassium channel, KcsA (50), and the *E. coli* lipid flippase, MsbA (32), are just two examples of structural studies which support a consensus protein topology of hydrophobic and amphipathic  $\alpha$ -helices connected by hydrophilic loops. In contrast, crystallographic studies on outer membrane proteins suggest that most of these proteins form anti-parallel  $\beta$ -sheets which are arranged in a barrel conformation (130). This  $\beta$ -sheet structure allows outer membrane proteins to have, in general, a low overall hydrophobic content, as only alternate side chains need to be exposed to the lipid bilayer. It has been proposed that these characteristics allow outer membrane proteins to be translocated through the cytoplasmic membrane into the periplasm, whereas the more hydrophobic  $\alpha$ -helical sequences of inner membrane proteins may be retained by the cytoplasmic membrane (240).

### **1.7.2 The general structure of outer membrane $\beta$ -barrel proteins**

High resolution structures of 6 outer membrane protein families show all form  $\beta$ -barrels with an even number of  $\beta$ -strands. Each  $\beta$ -strand lies adjacent to the next  $\beta$ -strand in the sequence, and these motifs are connected to each other by loops. Despite these similarities in secondary structure, the cross-sectional shapes of these  $\beta$ -barrels vary considerably. Analysis of 11 outer membrane proteins whose atomic structures have been determined shows the length of these  $\beta$ -strands vary between 6 (the minimum needed to span the membrane) and 25 amino acids, with an average length of 12.3 residues (130). Interestingly, while these  $\beta$ -strands start at the periplasmic boundary of the outer membrane, they often extend out into the external environment. The loops exposed to the outer environment are often long (2-46 residues) and hydrophilic, and display considerable sequence variability within families. This variability is most probably an adaptation to the role these regions play as receptors for phages or bacteriocins (130). On the periplasmic side of the membrane  $\beta$ -strands are connected by short (1-12 residues) peptide loops. The general porins are stable as homotrimers, while other outer membrane proteins may be monomeric. In general, the outside of a  $\beta$ -barrel, contacting the lipid bilayer, is covered by non-polar

side chains (with a clear preference for large hydrophobic residues such as phenylalanine, leucine, isoleucine and valine), while the interior cavity is much more hydrophilic (263). The residues which form the transmembrane (TM)  $\beta$ -strands are highly conserved within families, particularly those lining the interior of the  $\beta$ -barrel (130).

### **1.7.3 The general structure of cytoplasmic membrane proteins**

Cytoplasmic membrane proteins may consist of more than one subunit and may have associated proteins or subunits with large periplasmic or cytoplasmic domains. The integral membrane components of most cytoplasmic membrane proteins are predicted to contain an arrangement of  $\alpha$ -helices connected by hydrophilic loops. Experimental studies using biophysical techniques such as circular dichroism (67), and fusion proteins (54, 70, 166, 185) together with the recently determined 3D crystal structures (32, 50, 53) support this hypothesis. The minimum length required for an  $\alpha$ -helix to cross the cytoplasmic membrane is 19 residues, but TM helices are often longer as they may be tilted away from the vertical or extend beyond the membrane, as illustrated in the structure of the prokaryotic potassium channel, KcsA (50). Alternatively, some  $\alpha$ -helices may not fully traverse the membrane, as seen in the crystal structure of the ClC chloride channel from *E. coli* (53) (Figure 1.7). As mentioned previously, TM helices usually contain hydrophobic amino acids for stability within the hydrophobic core of the lipid bilayer. However,  $\alpha$ -helices may be amphipathic, exhibiting more polar residues along a face of the helix that may contact other TM helices within the protein complex (77) or be adjacent to an aqueous pocket, such as in the prokaryotic KcsA potassium channel (50, 77). Occasionally, charged residues that are functionally or structurally important may also be located within these membrane spanning  $\alpha$ -helices (43, 44, 223, 296, 307).

X-ray crystal structures and topological predictions suggest that these TM  $\alpha$ -helical regions are connected by hydrophilic loops. von Heijne *et al* (276) discovered that in *E. coli* cytoplasmic loops with less than 60 residues have four times the number of



positively charged lysine and arginine side chains than periplasmic loops of similar size, while the ratio of negatively charge side chains remained similar in each set of loops ('positive inside' rule). Large loops (greater than 60 residues) however, do not retain this charge ratio. Many of these extra-membranous loops are not required for folding or function. Lactose permease is just one example of a membrane protein that can be expressed *in vivo* as two segments to create a functional transporter (17, 201).

The membrane component of many transport proteins is predicted to contain 12 TM  $\alpha$ -helices. These can consist of two proteins containing 6 predicted  $\alpha$ -helical regions that may be different subunits or a homodimer. Alternatively, one protein may contain 12 predicted  $\alpha$ -helices which have sequence similarity between the N-terminal and C-terminal domains, suggesting that these proteins arose from the duplication of an ancestral gene to give a "6 + 6" topology. Several of these proteins, such as the prokaryotic ClC channel, the mammalian PiT2 phosphate transporter and the *Arabidopsis* Pht2;1 phosphate transporter, have been shown to have an anti-parallel structure where the second domain has the opposite orientation in the membrane (53) (47, 229).

While individual  $\alpha$ -helices are stabilised by internal hydrogen bonding and interactions with the lipid bilayer, it is the interactions between different TM helices and loops which are believed to be important to the overall tertiary structure and mechanism of membrane proteins (144, 200). Popot *et al* (201) suggest that the interactions favouring specific helix associations include the coordination of favorable Van der Waals contacts over complementary surfaces and the binding of prosthetic groups. Single strong interactions, via hydrogen bonds, covalent links or ion pairs are rarely seen between such helical arrangements. Studies derived from the crystal structures of 7 integral membrane proteins and 37 soluble proteins show that, in general, membrane protein  $\alpha$ -helices pack together more tightly (higher packing values) than in soluble proteins (57). The smaller amino acids such as glycine, alanine, serine and threonine seem to stabilise the tight association of these TM helices. Interhelical regions that are more loosely packed often contain polar residues and provide the binding sites for substrates and ligands (57). Analysis of membrane proteins with known structures has also found

that: 1) helices almost always pack against neighbors in the sequence, 2) helix packing in an anti-parallel orientation is more prevalent than a parallel orientation, 3) there is no preference for the positioning of the point of closest contact along the length of the helices, 4) most helices are tilted with respect to the bilayer (by an average of 21 degrees in one study) (19).

The increasing number of membrane proteins with experimentally determined topologies or structures has allowed various analyses of more specific amino acid motifs and distributions within the  $\alpha$ -helical domains. Russ *et al* (216) used a randomised sequence library based on the glycophorin A structure to identify a GxxxG TM helix packing motif that promotes dimerisation in this system and is present in the TM domains of many membrane proteins. While glycine residues are noted helix breakers in soluble proteins, they are moderately common in TM  $\alpha$ -helices, and do not tend to disrupt the  $\alpha$ -helical backbone dihedral angles. Glycine residues are predominantly angled towards helix-helix interfaces, and are highly predominant at helix crossing points, serving as molecular 'notches' for orienting multiple helices in the tertiary structure (105).

Proline residues are also known to be helix breakers in soluble proteins. Prolines can cause kinks in helical secondary structure (24). This characteristic has led to suggestions that prolines may function as molecular hinges that enable conformational transitions, such as opening the internal cavities in rhodopsin and bacteriorhodopsin for the binding of the retinal prosthetic group (48, 252, 287). The involvement of prolines in conformational changes within the *E. coli* Pst phosphate transport system has been demonstrated by mutagenesis experiments. The PstC and PstA membrane components contain paired proline residues that occur at about the same position within the membrane in adjacent helices. A double proline to alanine PstC mutant (P123A, P183A) caused a permanently 'open' conformation in the Pst system while the equivalent double proline mutation in PstA (P132A, P166A) represented a permanently 'closed' configuration (284). In membrane proteins with multiple helices, kinks seem to be more predominant when sequences contain more than one proline or a combination of proline and glycine spaced four residues apart, such as in cytochrome c

oxidase (105). However, studies on model peptides show that the structural effects of proline are strongly dependent on its environment, and that this residue can even stabilise helical structures under certain conditions (57, 146).

Other amino acid patterns have been observed with the subsequent structural determination of membrane proteins. The amphipathic aromatic residues tyrosine and histidine are preferentially found at the lipid/water interfaces, where they are proposed to anchor the proteins into the membrane through an interaction of their aromatic rings with the lipid head groups (263). Charged residues are preferentially found at the interfacial regions and water accessible regions of proteins. Those charged residues inserted within the membrane are usually in the form of salt bridges and/or are involved in function, such as Glu-325 in lactose permease, which is proposed to form part of the proton translocation pathway (223).

## ***1.8 Methods to determine membrane protein structure***

### ***1.8.1 Direct methods for the determination of membrane protein structure***

When studying the mechanism of a protein it is helpful to know the three dimensional (3D) protein structure, so that residues likely to be involved in the mechanism can be identified and examined experimentally. With soluble proteins this structure can often be determined at atomic resolution through the purification, crystallisation and analysis by X-ray diffraction of the protein. By contrast, the structural determination of membrane proteins is notoriously difficult at all stages of this process. Membrane proteins have regions of high hydrophobicity and require the presence of detergent and/or lipids to remain soluble and to retain their 3D structure through the purification and crystallisation process. Crystallisation is hindered by the disorder created by the presence of this detergent or lipid and the subsequent decrease in protein-protein interactions these compounds cause. Replacement of native lipids by detergent may also disturb the natural conformation of the protein. The inherent mobility of many membrane proteins, which have evolved to function in a fluid membrane environment, also works against crystallisation. As a consequence, even when crystals are formed



these can often lack the order needed for resolution to be at the level where the polypeptide backbone and side chains can be precisely mapped (below 4Å, preferably around 2Å) (32, 305).

Despite these challenges, the structures of a number of membrane proteins have recently been solved using X-ray crystallography at resolutions approaching atomic levels. Many bacterial outer membrane porins with stable trimeric homo-oligomer structures, such as OmpF, OmpA and PhoE, have proven amenable to this technique (42, 195, 201). The 3D structures of bacterial cytoplasmic membrane proteins, such as photosynthetic reaction centres, chloride and potassium channels, and cytochrome *c* oxidase have also been resolved at less than 3.5Å (49, 50, 53, 193, 201).

A number of specific approaches have assisted in the preparation of well ordered crystals. Formation of the protein-protein lattice contacts needed during crystallisation is strongly influenced by variations in the surface residues of a protein. Site-directed mutagenesis has been used to make proteins more conducive to crystal formation. For example, the long surface-exposed loops of the outer membrane protein OmpA were predicted to form unfavorable protein-protein contacts that could limit crystal growth. When these loops were reduced in size by mutagenesis well ordered crystals were formed (195). The prokaryotic KcsA potassium channel was also crystallised without the C-terminus (Residues 126 to 158) following limited proteolysis – a method often employed in crystallography trials (50). An alternative method for optimizing lattice interactions is to purify and attempt to crystallise a number of natural homologues of the desired protein. Chang *et al* (32, 33) have successfully used this approach to determine the crystal structure of the mechanosensitive ion channel from *Mycobacterium tuberculosis*, Tb-MscL, to 3.5Å resolution and the *E. coli* lipid flippase, MsbA, to 4.5Å resolution.

Site-directed mutagenesis has also been used to assist in protein purification by the addition of purification tags, such as six histidine residues. However, even the availability of large amounts of pure protein does not guarantee success with 3D X-ray crystallography, so a number of innovative alternatives have been used. Membrane proteins are more amenable to the formation of two-dimensional (2D) crystalline

arrays. While these are too thin to diffract X-rays, electron microscopy can be used to obtain a diffraction pattern. 2D protein-lipid crystals are exposed to a low intensity beam of electrons and images are formed through the summation of conserved patterns. 3D pictures are obtained by combining images taken at different tilt angles. To obtain high resolution images, specimens need to be frozen in amorphous ice and analysed under these conditions, in a process called cryo-electron microscopy (cryo-EM). This preserves the integrity of the protein in its lipid bilayer, reduces radiation damage to the protein and reduces movement or heating of the sample while it is exposed to the electron beam. Although it is easier to form these 2D crystalline arrays the resolution of cryo-EM is usually lower than that of X-ray crystallography. This lower resolution is due mainly to diffraction patterns becoming difficult to obtain once the sample is tilted beyond 60 degrees. Consequently the lack of data from the remaining 30 degrees reduces the final overall resolution (anisotropic data). High tilt data also has a loss of clarity caused by membrane crystals often lacking uniform thickness (251). The resolution of this technique is currently around 7 to 9Å, such as for the *E. coli* Na<sup>+</sup>/H<sup>+</sup> antiporter, NhaA. At this resolution the identification and basic arrangement of  $\alpha$ -helices can be determined, but no information can be derived regarding the peptide backbone and side chain arrangements (286). The structure of the eukaryotic light-harvesting complex II has been resolved to 3.4Å using anisotropic data from well ordered 2D crystals formed from detergent solubilised protein (136, 201). Cryo-EM has also solved the structure of bacteriorhodopsin to 3.0Å resolution (201). The key to this protein's success has been that it forms pseudo 3D crystals, where 2D arrays of purple membranes stack on top of one another, thus giving an additional level of information.

Cryo-EM can also be carried out on membrane proteins that form tubular rather than flat crystals. The advantage of this tubular arrangement is that it supplies a complete set of 3D images of the protein by the very nature of the tubular packing, hence no tilting of the sample is necessary. The main problems associated with tubular arrays arise from distortion and variation within the tubes (251). Cryo-EM of tubular arrays has been used to determine the structure of the nicotinic acetylcholine receptor at 4.6Å resolution in the closed state and 9Å in the open state, and also the calcium ATPase at 8Å resolution (175, 264, 306). Single particle analysis, where the protein is simply



solubilised and then analysed by cryo-EM, is a promising technique for large complexes. Thousands of images need to be collected from particles in different orientations, thus building up the 3D image. The inherent flexibility of many macromolecules and the heterogeneity of the purified sample are some of the major limitations to this technique (251). Cryo-EM is also suited to the study of conformation changes, as the determined structure is from complexes trapped in solution under conditions similar to those used for biochemical and kinetic characterisations (225). For example, the nicotinic acetylcholine receptor has been imaged in both the open and closed states and with the continued improvement in data collection and processing a higher resolution of data will be forthcoming. Thus insight will be gained into how this ion channel transmits a binding signal into the opening of a pore through the membrane resolution (175, 264).

While solution nuclear magnetic resonance (NMR) has been successfully used to determine the structure of the 40 amino acid Glycophorin A TM region, this technique is usually limited to soluble proteins with molecular weights of less than 50000 Daltons (158). Solid-state NMR of proteins in hydrated but anisotropic environments is becoming a promising technique now that more membrane proteins are able to be purified in high quantities (5-10 mg), particularly if the sensitivity of the instruments continues to improve (72).

Proteins that have been analysed by both X-ray crystallography and cryo-EM, such as bacteriorhodopsin, show each technique produces consistent structures (225). This has set the stage to allow the combination of low resolution cryo-EM maps of large complexes that do not form good 3D crystals with the atomic structures of smaller domains or subunits determined by X-ray crystallography or NMR. This combination of techniques provides a powerful tool. Structural determinations for eukaryotic potassium and chloride channels and members of the multi-drug resistance family of transporters have been greatly enhanced by the X-ray crystal structures of their bacterial homologues (50, 53, 32). These bacterial proteins have simpler architecture (e.g. prokaryotic KcsA has 2 TM regions rather than the 6 TM regions found in the *Drosophila* (*Shaker*) and vertebrate eukaryotic potassium channels (50)). As a consequence, these are also easier to over-express and crystallise (132).



While over 40 membrane protein structures with atomic resolution better than 4Å are now available, including bacterial chloride and potassium channels, outer membrane porins and photosynthetic reaction centres (50, 53, 263), none is an ion-coupled transporter. In the absence of a high-resolution structure researchers have had to rely on indirect methods to obtain the structural constraints needed for an educated guess at the structure/function relationships of ion-coupled transporters. (An overview of the methods used to analyse lactose permease is presented in Section 1.9.2.)

### **1.8.2 Indirect methods for the determination of membrane protein structure**

Consensus prediction programs are used to propose a secondary structure for putative membrane proteins from the amino acid sequence. While these began as simple determinations of regions of hydrophobicity, the increase in the number of membrane proteins with experimentally determined topologies has led to increasingly sophisticated and varied algorithms. Nilsson *et al* (185) tested five prediction methods on 60 *E. coli* inner membrane proteins with experimentally determined topologies. They found a greater than 90% success rate if at least four of the five methods agreed, which occurred for around 50% of inner membrane proteins. Thus these programs are commonly used to develop one or more topological models.

Using such a topological model the putative secondary structure of a membrane protein can then be tested, for example, by the introduction of tags into proposed cytoplasmic or periplasmic loops. These tags can include epitopes modified by reagents that cannot penetrate the membrane, such as monoclonal antibodies, protease cleavage sites, and N-linked glycosylation sites. Such methods have their limitations as membrane proteins can be quite resistant to proteolysis and antibody binding studies can be difficult to interpret (277). The preparation of protein fusions with reporter domains that function differently on each side of the membrane have been very successful with prokaryotic proteins. The most commonly used reporter proteins are alkaline phosphatase (PhoA),  $\beta$ -lactamase (Bla) and  $\beta$ -galactosidase (LacZ). *phoA* and *bla* are only active in the periplasm of *E. coli*, while LacZ is only active in the cytoplasm (54, 70, 166). PhoA or Bla genes devoid of their N-terminal signal sequences can be fused to predicted

periplasmic or cytoplasmic domains, replacing the C-terminal portion of the membrane protein being studied. It has been shown that approximately half of an 'outward' TM domain is needed to translocate the PhoA moiety through the membrane to the periplasm (118). LacZ fused to cytoplasmic domains forms a functional protein, while periplasmic domain fusions cause it to become trapped within the inner membrane (70). Complementary studies combining the use of PhoA and LacZ fusions have been carried out (165). A number of methods have been designed to create these fusions, including a method to insert alkaline phosphatase within a protein so that the C-terminus of the protein is retained. This preserves any topological determinants in the C-terminal sequence (260), providing a more accurate representation of membrane protein topology than that supplied by simple fusion proteins (55). Proteins analysed by these methods include lactose permease (28), *sn*-glycerol-3 phosphate permease (78), the phenylalanine-specific permease (198) and the penicillin binding protein 1B (54).

Site-directed mutagenesis and second site suppressor analysis of amino acids identified as important from the topological model (e.g. charged or polar residues buried in the membrane) or from chemical modification studies can reveal potential interactions between residues. These relationships lead to the refinement of this secondary structure and to proposed interactions that may be necessary for determining the tertiary structure. However more precise structural information is needed. The use of site-directed mutagenesis to engineer membrane proteins for biophysical measurements is emerging as a powerful means of determining structure and structural dynamics. A protein mutated to contain a single Cys residue can be chemically modified with alkylating agents (such as N-ethylmaleimide) or labeled for fluorescence spectroscopy to measure the exposure of this position to the external environment or the lipid phase of the membrane (300). A number of fluorophores of different sizes and hydrophobicity, plus a variety of quenching agents can be used for these determinations. These single Cys residues can also be modified with a nitroxide spin label that is analysed in conventional electron paramagnetic resonance experiments to determine solvent exposure, electrostatic potentials and (depending on experimental design) the orientation and movement of individual segments of membrane protein (102).

Proteins engineered to contain two Cys residues can be used to establish the proximity of two residues through thiol cross-linking, excited state dimer fluorescence or spin labeling experiments. Double spin labeling allows the distance, and in some cases the relative orientation between the two spin labels to be defined. Engineered metal binding sites created by the introduction of two His residues have also been used to determine the proximity of two side chains (92). Defining spatial relationships, such as adjacent  $\alpha$ -helical surfaces, is necessary for the establishment of a proteins' tertiary structure. Spin labeling can also be used to study protein dynamics, such as conformational changes caused by ligand binding (102).

## ***1.9 The functional analysis of membrane proteins***

### ***1.9.1 Methods for the functional analysis of membrane proteins***

The 3D structure of a polypeptide chain plays an important role in defining a protein's mechanism. Therefore many of the indirect methods used for structure analysis, described above, are also used for the determination of a protein's mechanism of action when its atomic structure has yet to be elucidated.

An additional technique used to locate functional motifs within a protein is the analysis of chimeras formed between two closely related proteins. A region of one protein is replaced with the corresponding segment of another similar protein. Analysis of the resulting chimeras may allow particular characteristics of a transporter, such as substrate range or substrate affinity, to be located within a particular amino acid sequence. This method has successfully characterised the functional domains of outer membrane porins such as PhoE, OmpF and OmpC (176, 186, 258, 265, 303), and the dopamine and the norepinephrine transporters of mammalian cells (27). A region in the amino acid transporter AroP which can confer the ability to transport tryptophan on a predominantly PheP (high-affinity phenylalanine transporter) protein has also been located, allowing the identification of an AroP residue, Tyr-103, that is important in tryptophan transport (38). However, results from indirect methods must be carefully



interpreted. Substrate recognition sites may involve different parts of the transporter and chimeric proteins may distort these or other structures. For example, while chimeras of the melibiose carrier from *E. coli* and *Klebsiella pneumoniae* led to the identification of an A58N mutation in the *K. pneumoniae* protein that could change its ion selectivity from protons to sodium ions, at least two regions of the *E. coli* carrier seem to be involved in sodium ion recognition (85).

### **1.9.2 Structural and functional analysis of lactose permease**

Lactose permease from *E. coli* is encoded by the *lacY* gene, and catalyses the coupled stoichiometric translocation of protons and galactosides. This was the first membrane protein to be cloned, sequenced, over-expressed and purified, with the purified protein shown to be functional as a monomer (224, 275). It has become a paradigm for secondary transport proteins and has been extensively studied using site-directed mutagenesis and various biophysical and chemical modification techniques. While attempts are being made to produce well-ordered crystals for X-ray crystallography, this is a notoriously difficult task, and it is through indirect methods that the structure and function of lactose permease is slowly being revealed.

Hydropathy analysis predicted lactose permease to be a membrane protein containing 12  $\alpha$ -helices with both the N- and C-termini in the cytoplasm. Circular dichroism measurements indicating the protein is 75-80% helical are consistent with this proposal. A number of techniques, such as laser Raman and Fourier transform infrared spectroscopy, immunological studies using monoclonal antibodies, limited proteolysis and chemical modification, were used to show the N- and C-termini as well as the putative second and third cytoplasmic loops are exposed to the cytoplasm. However it was not until the analysis of this protein with lactose permease-alkaline phosphatase fusions that the 12 TM region topology was confirmed (118).

This model for secondary structure gave a framework for the analysis of amino acid residues important for structure or function. Charged and polar side chains proposed to be within the membrane environment were the first to be targeted by chemical modification. Site-directed mutagenesis was then used on side chains whose

modification disrupted activity, replacing them with various charged, polar and neutral amino acids. These mutants were characterised by assays for active transport, facilitated diffusion (with either coupled or uncoupled translocation of protons), efflux, exchange or counterflow. Thus different functional roles were defined for the mutated amino acids. Charged or polar amino acids that the model suggested might be within the vicinity of disruptive mutants were also modified. Initial mutagenesis studies led to the proposal that His-322 and Glu-325 may be components of a charge relay system coupling lactose and proton transport (117). Second site suppressor experiments were also important for identifying interactions between charged and polar residues (127, 142, 143). However, mutational analysis is not necessarily clear cut. King *et al* (128) found replacing His-322 with the aromatic groups Tyr or Phe, which cannot be protonated, gave no active transport but produced coupled facilitated diffusion of lactose or melobiose (i.e. while proton cotransport occurred, lactose or melobiose could not be transported against their concentration gradients). These experiments suggest that His-322 is important for energy transduction (and also high substrate affinity), but does not need a dissociable proton or electronegative atom, as initially proposed by Kaback's laboratory.

The preparation of a functional lactose permease devoid of eight native Cys residues has allowed analysis of the permease in a number of ways. Each residue in lac permease has been replaced with a single Cys residue, demonstrating that only six amino acids in the permease are irreplaceable with respect to active transport, with all of these being charged. Four of these are thought to be involved in proton translocation and coupling with substrate translocation (Glu-269, Arg-302, His-322 and Glu-325) while Glu-126 and Arg-144 are indispensable for substrate binding, and are also likely to form a charged pair (273). It is proposed that one guanidino  $-NH_2$  group on Arg-144 ion pairs with Glu-126, while the other guanidino group H-bonds with the galactoside moiety of the substrate. The conservative substitution R144K prevents substrate binding and was much more detrimental than the E126D mutant, giving some support to this theory (220). Creating double Cys mutant combinations of charged residues that disrupted transport (without totally inhibiting it) showed that the neutralisation of two Lys-Asp pairs allowed lactose uptake, suggesting they may form salt bridges. (See Chapter 6 introduction, Section 6.1, for more details).



It became obvious that static and dynamic information at a high resolution was needed in order to define the transport mechanism. The Cys-free permease and the single Cys mutants have been important for the use of spectroscopic approaches to establish proximity relationships between TM helices, and to study the molecular dynamics of the protein. Site-directed chemical labeling with thiol reactive agents such as NEM and fluorescence labeling (MIANS, pyrene maleimide) used in conjunction with various hydrophilic or hydrophobic quenching agents of single Cys mutants has been used to show that ligand binding increases the reactivity of V315C, E269C and H322C (112, 113) and also causes a conformational change in the N-terminus of TM 1 (298). These techniques were also used to highlight the presence of two substrate binding sites (one high and one low affinity) which can be occupied at the same time. Proximity relationships have been determined by pyrene maleimide labeling of double Cys mutants, which has led to a proposed tertiary model for the C-terminal half of lactose permease (114). Spatial relationships have also been confirmed by engineering divalent metal-binding sites (bis-His residues) within the permease, and through the use of site-directed spin labeling and thiol cross-linking experiments (92, 116, 278, 301).

Hydropathy analysis, and even the more sophisticated predictive algorithms available today do not take into account charge pairing. Thus the secondary structure model was altered to include charged pair residues Asp-237 and Asp-240 in TM 7. The charge pair Glu-126 and Arg-144 also needed to be moved into the membrane (296, 307).

The helix-loop boundaries of lactose permease have been more tightly defined by the use of single amino acid deletions. Wolin *et al* (296, 297) showed that deletion of single or multiple amino acid residues within extramembranous loops had little effect on protein activity. However deletion of single residues over a narrow range of positions near the boundary of helical domains greatly decreased activity. In typical Kaback laboratory style, all the lactose permease helix-loop boundaries were mapped in this way. This approach caused some  $\alpha$ -helices in the topological model predicted by the HMMTOP 1.1 algorithm to be moved by as much as 10 residues, although most were in general agreement. Unexpectedly, deletion analysis of the whole of TM 6 showed the deletion of only three residues abolished activity. This TM domain is



predicted to be on the periphery of the 12-helix bundle and may make few functionally important contacts (297).

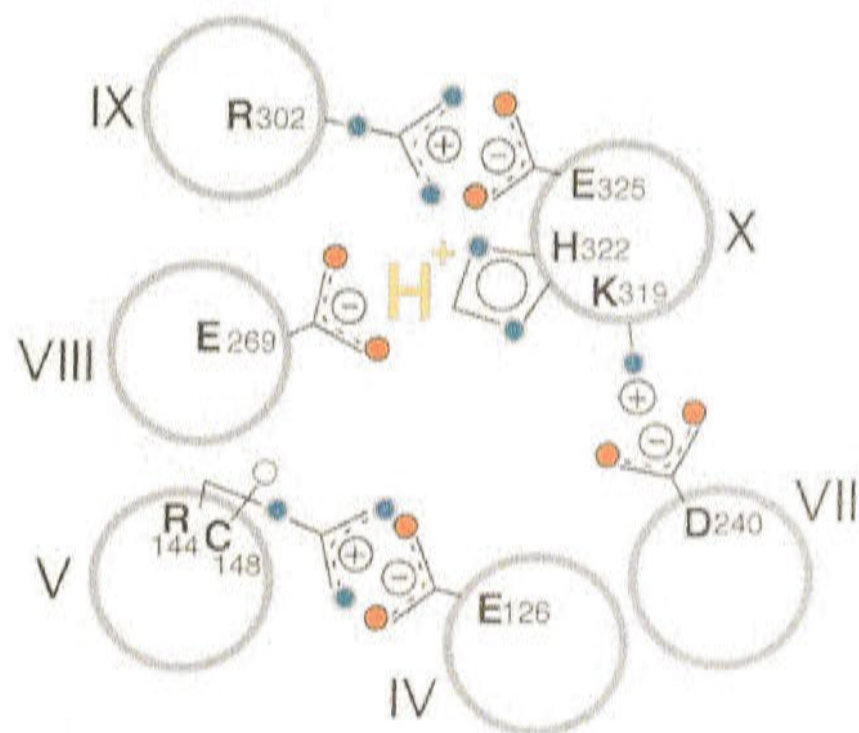
Recently the pH dependence of ligand binding by mutants in the four residues that are considered to be key participants in either proton translocation or the coupling between sugar and proton translocation (Glu-269, Arg-302, His-322 and Glu-325), was studied. A very sensitive binding assay involving the substrate protection of Cys-148, (which is not essential for transport but is located in a ligand binding site (299)), against alkylation was crucial to the success of this project. Wild type permease bound ligand with a pKa of 8.1 indicating that the permease needs to be protonated for binding. Neutral and charged substitutions at Glu-325 and Arg-302 had little effect on substrate binding, while mutations at His-322 and Glu-269 greatly reduced the binding of substrate at any pH. The pKa of 8.1 is greater than that of histidine, suggesting the proton may be shared with Glu-269. This information was combined with the structural and dynamic information previously determined (e.g. the face of TM 7, containing Glu-269, undergoes a ligand induced conformational change (69)), allowing Kaback's group to propose the following mechanism for lactose permease:

In the ground state (see Figure 1.5 from Sahin Toth *et al* (223)) the permease is protonated, with the proton being shared between His-322/Glu-269, while Arg-302 and Glu-325 are charge paired. Substrate binding on the periplasmic surface with relatively high affinity causes a conformational change, leading to the transfer of the proton from His-322/Glu-269 to Glu-325 and the reorientation of the sugar binding site to the cytoplasmic surface with a decrease in substrate affinity, releasing the sugar. These changes lead to deprotonation of Glu-325 as it re-associates with Arg-302. The cycle is returned to the original state by His-322/Glu-269 being reprotonated from the periplasmic surface (223). The role of Arg-302 in this model was subsequently tested by characterising small neutral and polar substitutions. These (R302A, R302S) were defective in active uptake of lactose and downhill influx or efflux, but catalyse equilibrium exchange – similar to neutral replacements for Glu-325. This lends strong support for the argument that Arg-302 interacts with Glu-325 to facilitate deprotonation of the carboxylic acid at the cytoplasmic face of the membrane. It is interesting to note that replacing Arg-302 with the larger amino side chain Leu produced a mutant

## Figure 1.5

### Proposed mechanism for lactose permease by Kaback's research group.

This diagram is from Sahin-Toth *et al* (223). For clarity, only 6 of the 12 helices in the permease are shown. Glu-126 (helix IV) and Arg-144 (helix V) are charge paired and with Cys-148 (helix V) comprise the major components of the substrate binding site. Arg-302 (helix IX) is charge paired with Glu-325 (helix X), while protonated His-322 (helix X) interacts with Glu-269 (helix VIII). Thus the relevant  $H^+$  is shared by His-322 and Glu-269. Also shown are the charge pair between Asp-240 (helix VII) and Lys-319 (helix X), which are not essential for the mechanism. The mechanism is explained in detail in Section 1.9.2.



defective in all aspects of transport, as did the conservative R302K replacement, indicating the potential complexity of interpretation involved in mutagenesis studies (221). This mechanism may also explain the results of King *et al*'s (128) aromatic replacements of His-322. The Phe or Tyr substitutions may be able to hold a proton with Glu-269, allowing low levels of ligand binding, without allowing proton translocation.

However there are mutagenesis studies that argue against the above model. Brooker's laboratory propose that an alternative mechanism of lactose permease cation binding and transport may occur, possibly involving the hydronium ion ( $H_3O^+$ ) (110), an idea also put forward by Kaback (115). This is based on the characterisation of a K319N/E325Q mutant that catalyses a sugar-dependent proton leak. Thus, an acidic residue at position 325 is not required for proton transport, which argues against an essential role for Glu-325 in proton binding (109). Several suppressor mutants at His-332 prevented this proton transport with lactose. However if the substrate was changed to  $\beta$ -d-thiodigalactoside (TDG) several of these mutants continued to allow the proton leak, and the triple mutant K319N/H322Q/E325Q catalysed strong proton transport with TDG. It would be hard to identify potential hydronium ion binding sites via site-directed mutagenesis studies as backbone carbonyl groups may be involved in forming a coordination complex (110).

Brooker's laboratory has also taken a different approach to solving the mechanism of lactose permease. They propose that the 'internal repeat' of 6 TM domains found in the Major Facilitator Superfamily (MFS), of which lactose permease is a member, may represent the transformation of a functional dimer into a single protein, allowing each half of the dimer to evolve independently. Sixty-five members of this family were analysed by a predictive program (MEMSAT (111)) to identify putative TM  $\alpha$ -helices. The least hydrophobic TM region exhibiting high amphipathicity, (which are more likely to be potential channel-lining domains) were then identified, and patterns within the family were noted. The lengths of hydrophilic loops connecting TM  $\alpha$ -helices were then compared throughout the family, with short loops connecting adjacent helices and long loops allowing distant arrangements in the tertiary structure. This data



was combined with information on potential salt bridges and other helical interactions proposed by the mutagenesis and cross-linking studies described earlier to develop a model where lactose permease consists of two halves that fold in a rotationally symmetrical manner. Movement of these two moderately rigid domains of 6 TM helices changes the solute binding site from the outside to the inside of the cell (Figure 1.6)(77).

This lactose permease model suggests that TM 2 slides against TM 11 and TM 8 slides against TM 5. Subsequently a conserved motif in the MFS that is found at the interface of TM 2 and the 2/3 cytoplasmic loop, and is repeated in cytoplasmic loop 8/9 has been analysed. Mutations in the first two conserved amino acids inhibit or substantially decrease the rate of lactose transport without affecting substrate affinity (107, 108). This motif is proposed to play a general role in conformational changes rather than playing a role in substrate recognition. Mutations along the face of TM 2 and TM 8, which align with the first residue in this conserved sequence, were the most defective in lactose transport (80).

Thus a crystal structure of lactose permease would greatly clarify the tertiary structure and mechanism. Attempts to crystallise lactose permease from *E. coli* include the insertion of the small soluble protein Cytochrome b562 into the central hydrophilic loop of lactose permease to increase protein-protein interactions (202). This 'red permease' has been solubilised in detergent and reconstituted into vesicles and well-ordered 2D crystals. Cryo-EM using both single particle analysis and 2D crystal analysis has been carried out, but the resolution is quite low, only revealing the basic shape of this crystallised lac permease protein (311).

### ***1.9.3 The success of indirect techniques in structure and function predictions***

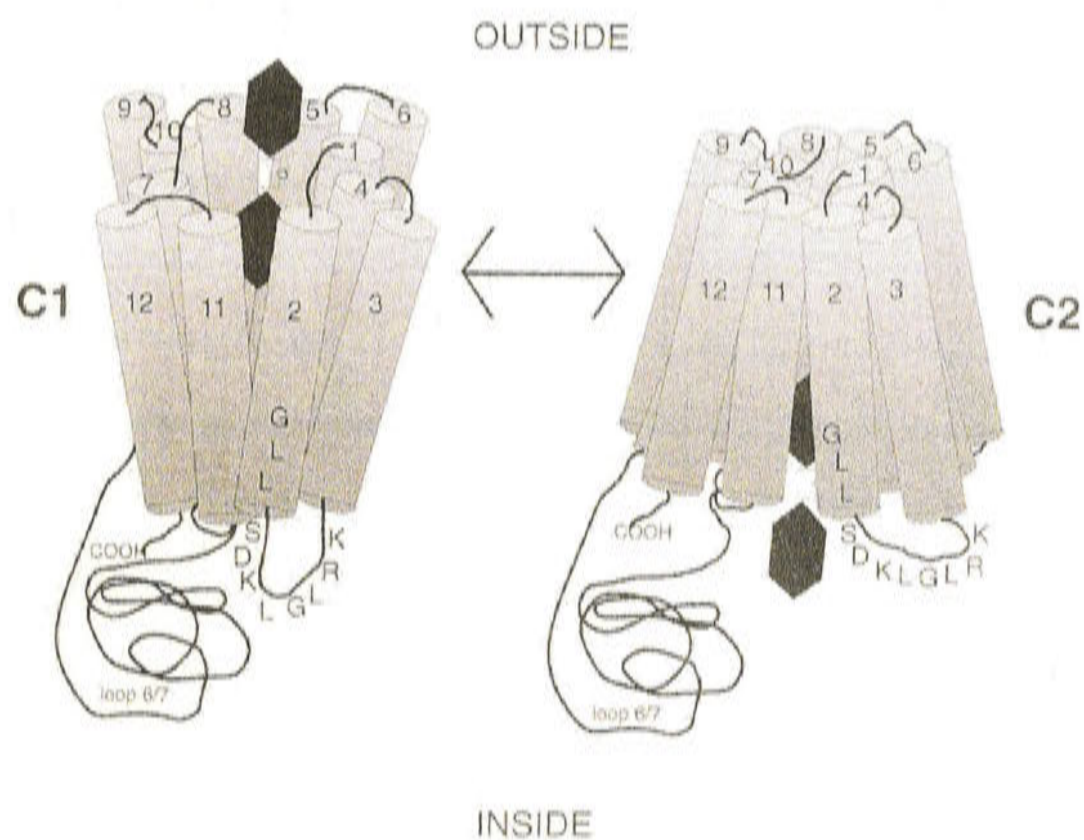
Now that some membrane protein structures have been solved at the atomic level the success of indirect methods at structure/function analysis can be examined for individual proteins.

## Figure 1.6

### Proposed mechanism for lactose permease by Brooker's research group.

This diagram is from Jessen-Marshall *et al* (108). The three dimensional arrangement of transmembrane domains has been proposed by the methods explained in Section 1.9.2. The channel lining segments are TM-1, TM-2, TM-4, TM-5, TM-7, TM-8, TM-10, TM-11, while the scaffolding segments are TM-3, TM-6, TM-9 and TM-12. The two halves of the protein exhibit rotational symmetry around a central axis. The C1 conformation has its solute binding site accessible from the outside while the same site in the C2 conformation is accessible from the inside.

(The highlighted residues in loop 2/3 form part of a conserved motif discussed in Chapter 7.)





Bacteriorhodopsin, a light driven proton pump from Halobacteria, was extensively studied with spectroscopic and mutagenic techniques. While this gave some conflicting data (25), the model and proposed mechanism was in general agreement with the structure produced by cryo-EM of purple membranes (82, 126). Interestingly, atomic resolution X-ray diffraction of crystals produced two quite different structural models, with one being more strongly supported by the independent data (140). These atomic structures revealed features important in defining the mechanism such as bound water molecules and hydrogen bond interactions.

Several mechanisms have been suggested for substrate translocation by the multi-drug resistance transport family, based on labeling and mutagenesis data for a number of different proteins. The ‘flippase’ model proposed that lipid substrates are accepted laterally from the inner leaflet of the cytoplasmic membrane and flipped to the outer leaflet (95). The hydrophobic vacuum cleaner model suggests that drugs may be accepted laterally from the inner leaflet and then released directly into an aqueous chamber exposed to the external environment (79). The substrate binding site may move between the inner lipid and outer lipid/aqueous environments by either a tilting or rotational conformational change (96). The atomic structure of the lipid flippase MsbA from *E. coli* (32) is generally consistent with the biochemical data for the flippase model with a tilting conformational change, while revealing additional insights such as the linking structure between the transmembrane and the nucleotide binding domains. However, MsbA crystals, which were formed in the absence of lipid, consist of a homodimer with a large cytoplasmic cavity that is closed to the external environment. This conflicts with evidence from low resolutions structures for P-glycoprotein and MRP1 that have chambers exposed to the external environment (96), and cross-linking studies which suggest a different arrangement of  $\alpha$ -helices (153). This closed chamber may be an artifact, the result of a different dynamic conformation or may be due to different mechanisms of action within this family. The targeted use of site-directed mutagenesis and other methods can now be used to test this mechanism in MsbA and other related proteins in the multi-drug resistance-ABC transporter family.

The atomic structure of the *Streptomyces lividans* KcsA potassium channel is in excellent agreement with results from functional and mutagenesis studies on eukaryotic



$K^+$  channels. These studies had defined many of the amino acids involved in the pore opening, and the critical  $K^+$  signature sequence which forms the  $K^+$  selectivity filter, as predicted ((50) and references therein). The generally hydrophobic character of the pore lining was anticipated by hydrophobic cation binding studies (5) and the proposed presence of multiple ions queuing inside a long narrow pore in single file, seems true for the selectivity filter. However from the X-ray crystal structure it was apparent that it is the main chain carbonyl oxygen atoms on the residues of the  $K^+$  signature sequence that coordinate with the  $K^+$  ions, not the actual side chains. These side chains seem to be important for constraining the size of the selectivity channel, so that a dehydrated  $K^+$  ion fits with proper coordination but the  $Na^+$  ion is too small to undergo dehydration. The presence of a wide water filled cavity after the narrow selectivity filter was also not anticipated from indirect methods. The 3D structure revealed the basic KcsA channel architecture of four identical subunits forming a pore which narrows to a selectivity filter, then opens into a wide water filled cavity lined with hydrophobic amino acids (50). The structure also helped in the understanding of how the channel could be highly selective for  $K^+$  over  $Na^+$  while maintaining a throughput rate approaching the diffusion limit. The mechanism of action of the selectivity filter has been further defined by crystal structures obtained at 2Å resolution which were carried out with different concentrations of  $K^+$  (310).

Indirect techniques have not been so successful at defining the mechanism and structure of chloride channels. Mutational, electrophysiological and biochemical analyses confirmed that ClC chloride channels consist of a dimer with two separate channels. However, further structural definition remained ambiguous as mutations that changed ion permeation properties were scattered over much of the amino acid sequence and their affects were not very pronounced. The crystal structures of two prokaryotic ClC chloride channels revealed that several  $\alpha$ -helices stop mid-membrane to form part of the  $Cl^-$  ion selectivity filter and that many of the  $\alpha$ -helices are often severely tilted within the membrane (53) (Figure 1.7). Many residues that line the hour-glass shaped pore were identified in earlier mutation studies. However, the antiparallel symmetry and a negatively charge glutamate side chain located just above the  $Cl^-$  binding site (which may act as a gate) were not predicted from these studies. This chloride channel

## Figure 1.7

### Structure of the *Salmonella enterica* serovar *typhimurium* ClC chloride ion channel.

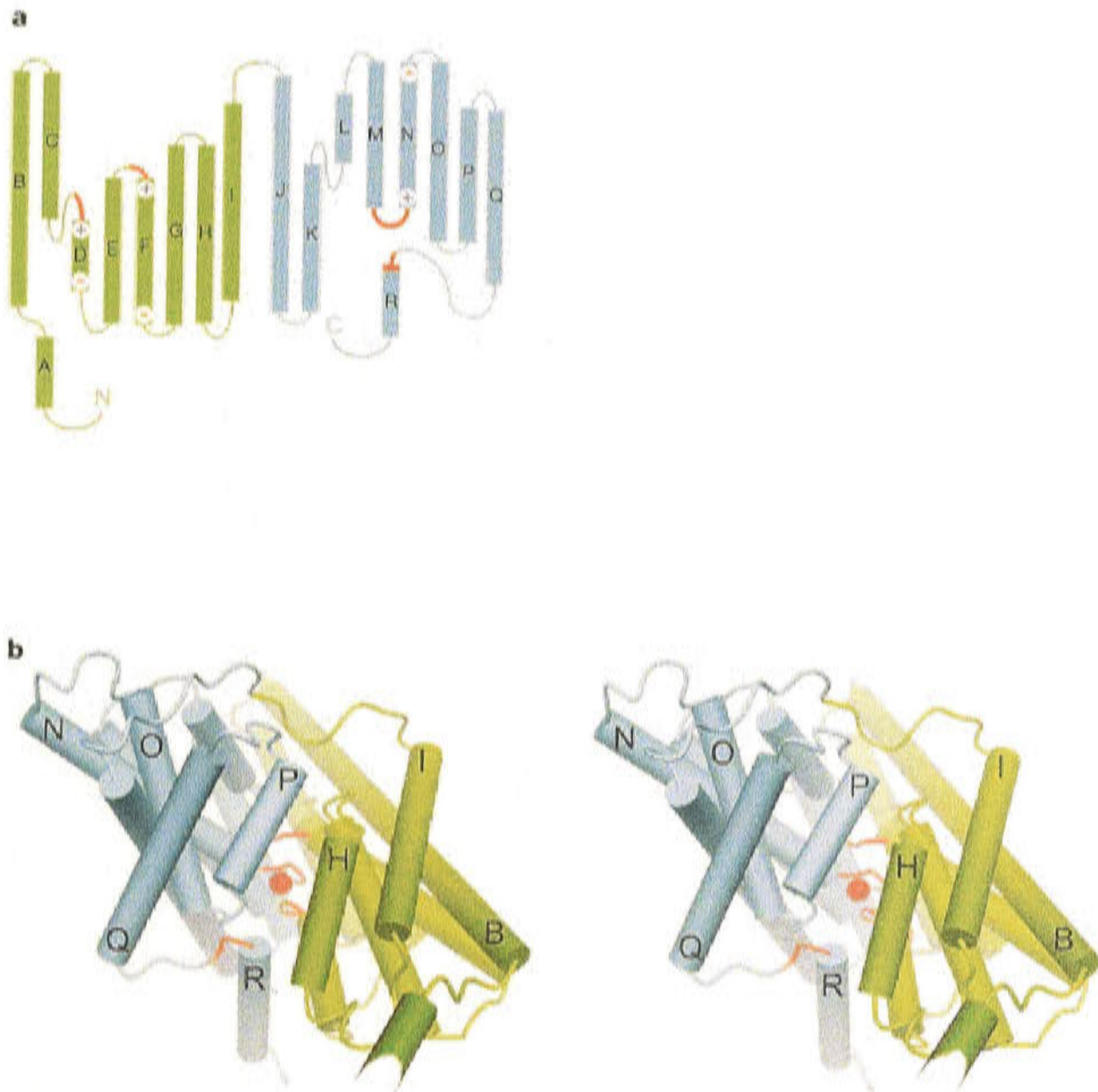
This diagram is from Dutzler *et al* (53).

#### A Helices

The helices are drawn as cylinders with the extracellular region above and the intracellular region below. The two halves of the subunit are green and cyan, and regions forming the chloride ion selectivity filter are red. Partial charges at the end of helices involved in chloride ion binding are indicated by + and - (end charges) to indicate the sense of the helix dipole.

#### B Stereo view of the three dimensional structure.

This view is from within the plane of the membrane from the dimer interface with the extracellular solution above. The helices are drawn as cylinders, loop regions as cords (with the selectivity filter in red), and the chloride ion as a red sphere.





has a very different architecture to the KcsA potassium channel, which has parallel symmetry and a large aqueous cavity at one side of the pore (Figure 1.8).

Non-crystallographic approaches correctly predicted the membrane topology of the small outer membrane protein OmpA (178), and were fairly close with PhoE (266). However, the predicted models of LamB (guided by the OmpF structure) and FhuA (with a predicted loop instead of a plug domain) were not close to the structures determined by crystallography. Predictions failed to identify many functionally important features of various outer membrane proteins, such as the transverse electric field inside the general porins, the LamB substrate translocation pathway – which consists of a row of ‘greasy slide’ aromatic amino acids lined up by an “ionic track” of polar residues (173) and the ‘plug’ domain in the TonB-dependent receptors (130, 150).

These results indicate that when no direct structural evidence is available, the success of indirect techniques at predicting structure is variable, even when some residues important for function have been identified. Atomic structures reveal new and unique architectures that are difficult to conceptualise through indirect techniques, particularly if they are subtle and involve residues from disparate regions of the polypeptide sequence. Atomic structures are also able to confirm and refine protein arrangements that have been proposed through the information provided by indirect techniques. In turn, functional evidence that supports structures revealed through X-ray crystallography, cryo-EM or NMR lends support for these structures having some physiological significance. Indirect techniques such as site-directed mutagenesis can help refine the models for individual proteins when an atomic structure is available for one protein within a family. These techniques also provide a powerful arsenal for examining the dynamics and putative mechanisms of known structures. For example, low resolution cryo-EM of P-glycoprotein isolated at different stages of the transport cycle supported by biophysical and mutagenic experiments suggest that ATP binding causes a large reorganisation of the TM domains of this multi-drug resistance transporter, probably through rotation of the  $\alpha$ -helices (213). Therefore these two sets of techniques are complementary, and are both required if we are to fully understand how membrane proteins function.



## Figure 1.8

### Comparison of the architecture of the CIC Cl<sup>-</sup> channels with the prokaryotic KcsA K<sup>+</sup> channel.

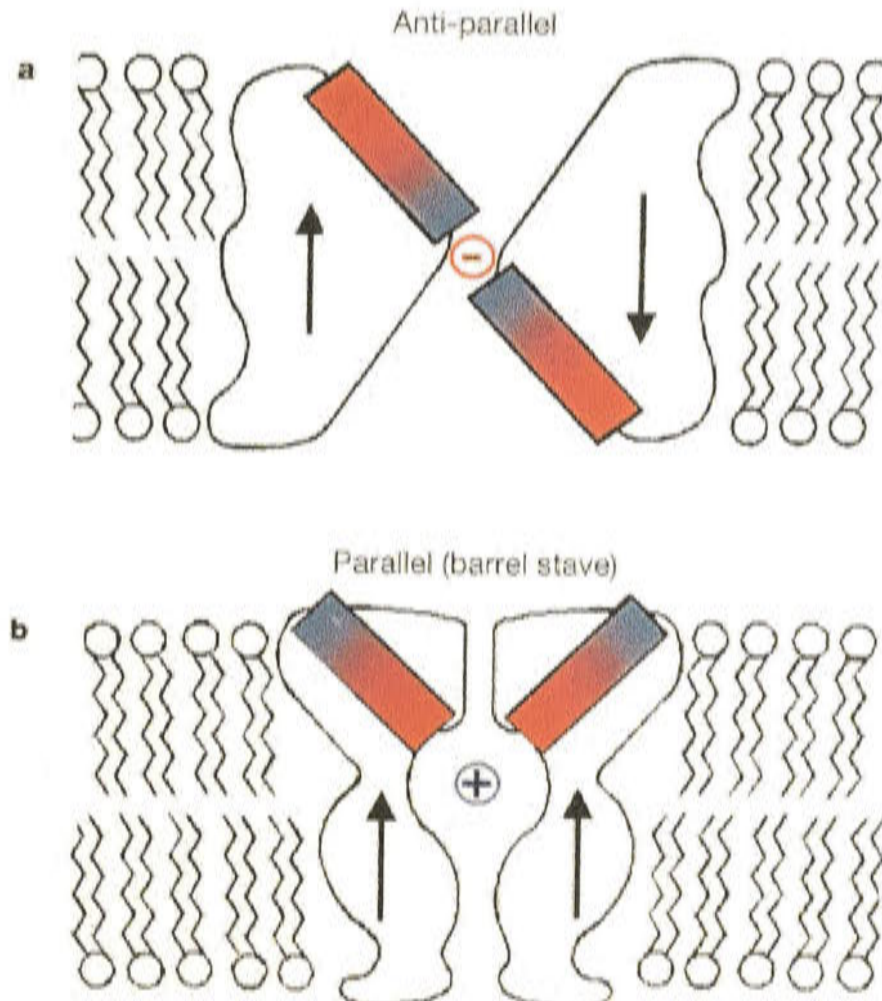
This diagram is from Dutzler *et al* (53).

#### A CIC Cl<sup>-</sup> channel

The antiparallel architecture of the CIC Cl<sup>-</sup> channels contains structurally similar halves with opposite orientations in the membrane (arrows). This architecture permits like ends (same dipole sense) of helices to point at the membrane centre from opposite side of the membrane.

#### B KcsA K<sup>+</sup> channel

The parallel or barrel stave architecture of the KcsA K<sup>+</sup> channel contains structurally similar or identical subunits with the same membrane orientation (arrows). Helices point at the membrane centre from the same side of the membrane. Helices are depicted as dipoles with blue (positive) and red (negative) ends.



## **1.10 Aims of the project**

Pit is the low affinity inorganic phosphate (Pi) transporter of *E. coli*. The seed for this project was sown when a second *E. coli* Pit transporter was isolated and sequenced by the Cox laboratory (JCSMR, ANU). These genes, subsequently named *pitA* and *pitB*, have a putative amino acid identity of 81% and preliminary experiments indicated that the proteins have significantly different  $K_m^{\text{app}}$  values for Pi. Therefore PitA and PitB form a good comparative system where relationships between the kinetic parameters, amino acid sequences and the Pi translocation mechanism of a membrane transporter may be systematically explored.

The aim of this project was to characterise the structure and function of Pit through analysis and comparison of PitA and PitB. The kinetic parameters of the PitA and PitB proteins were to be measured, and purification of each protein was attempted to allow N-terminal sequencing of the mature protein, thus defining the open reading frames of each gene and providing protein for future experimentation. The nature of the relationship between the *pitA* and *pitB* genes were also explored (i.e. are they independent genes or alleles?).

Targeted site-directed mutagenesis of PitA may be carried out to identify key amino acids in the  $\text{MePO}_4$ /proton translocation mechanism that are shared by both proteins. Preparation of PitA/PitB chimeras may be engineered to identify which variable region/s of these sequences confer the difference in substrate affinity.

However, while previous research had found Pit activity to be constitutive (290, 131, 211, 212), results from the above experiments suggested that *pitB* is regulated. Thus, the aims of the project shifted towards defining the expression and mechanism of regulation for the *pitB* gene on plasmids and the genome. This included exploring the possibility that the repression of *pitB* was mediated by the *phoB-phoR* two component regulatory system.

---

# *Chapter two*

## *Experimental Procedures*

---



## **Chapter two**

### **Experimental Procedures**

#### **2.1 Bacterial strains and plasmids**

The bacterial strains used in this study were all derivatives of *Escherichia coli* K-12 or K-10 and their relevant characteristics are listed in Tables 2.1 and 2.2. The plasmids and bacteriophages used and constructed are described in Tables 2.3 and 2.4.

Short-term storage of strains was on solid medium at 4°C. Long term storage of strains was carried out by resuspending a freshly grown lawn of bacteria from solid medium into Luria Bertani media containing 30% glycerol. These suspensions were stored in vials at -70°C.

The cell density of liquid cell cultures was measured as turbidity, against a blank of the liquid medium, using a Klett-Summerson colourimeter (Klett Manufacturing, NY, USA). Cell densities are expressed as Klett units, where  $10^8$  cells/ml is approximately equivalent to a reading of 40 Klett units.

#### **2.2 Oligonucleotide sequences**

Oligonucleotides used in this study for sequencing, site-directed mutagenesis and polymerase chain reactions are listed in Table 2.5.

#### **2.3 Growth media**

Strains were grown in the rich media of Luria Bertani (157), 2xTY, SOB or Z broth. Luria Bertani media consisted of 10g tryptone, 5g yeast extract and 5g NaCl per litre, adjusted to pH7 with aqueous 10M NaOH. This was supplemented with 33mM glucose after sterilisation by autoclaving at 121°C for 45 min. 2xYT medium contained 16g Bacto-tryptone, 10g yeast extract and 5g NaCl, adjusted to pH7 with aqueous 10M NaOH, and sterilised as described above. SOB medium contained 20g tryptone, 5g yeast extract, 0.6g NaCl and 0.5g KCl per litre. This is sterilised and just before use

### **Strain K-12**

*E. coli* K-12 MG1655 was sequenced by the Blattner laboratory because it approximates wild-type *E. coli* and "has been maintained as a laboratory strain with minimal genetic manipulation, having only been cured of the temperate bacteriophage lambda and F plasmid by means of ultraviolet light and acridine orange, respectively." (18). The mutations listed in the genotype are present in most K-12 strains and were probably acquired early in the history of the laboratory strain. MG1655 was derived and named by Mark Guyer from strain W1485, which was derived in Joshua Lederberg's lab from a stab-culture descendant of the original K-12 isolate. This original *E. coli* strain K-12 was obtained from a stool sample of a diphtheria patient in Palo Alto, CA in 1922 (Bachmann, B., pp. 2460-2488 in Neidhardt et al.1996, *Escherichia coli* and *Salmonella*: Cellular and Molecular Biology, ASM Press).

### **Strain K-10**

Information is from the CGSC database (CGSC#: 4234)

<http://cgsc.biology.yale.edu:80/cgi-bin/sybgw/cgsc/Strain/9448>

K-10 is a spontaneous *pit-10* derivative of strain CS101 (CGSC#: 5107)

(Skaar *et al*, see \*)

**TABLE 2.1:** Description of *E. coli* strains used in this study

Strain	Relevant characteristics	Construction/ Reference
AN248	<i>pit</i> <sup>+</sup> <i>pst</i> <sup>+</sup> K-12 strain	(211)
AN259	<i>pit</i> <sup>+</sup> <i>arg</i> <sup>-</sup> <i>entA</i> <sup>-</sup>	Laboratory strain
AN346	<i>pit</i> <sup>+</sup> <i>arg</i> <sup>-</sup> <i>entA</i> <sup>-</sup> <i>ilv</i> <sup>-</sup> <i>pyrE</i> <sup>-</sup> Str <sup>r</sup> Mu <sup>r</sup>	Laboratory strain
AN3066	<i>pitA1</i> $\Delta$ <i>pstC345</i> <i>srl::Tn10</i> <i>recA</i> Tc <sup>r</sup>	(284)
AN3901	<i>pitB::Cat</i> <sup>r</sup> derivative of JC7623	By recombination
AN3902	<i>pitA1</i> <i>pitB::Cat</i> <sup>r</sup> $\Delta$ <i>pstC345</i>	P1 AN3901 X AN3020
AN3926	<i>pitB::Cat</i> <sup>r</sup> $\Delta$ <i>pstC345</i>	P1 AN3901 X AN2537
AN4080	<i>pitA1</i> <i>pitB::Cat</i> <sup>r</sup>	P1 K10 X AN3902
AN4081	<i>pitA1</i> $\Delta$ <i>pstC345</i> $\Delta$ ( <i>phoB-phoR</i> ) Kan <sup>r</sup>	P1 ANCH1 X AN3020
AN4085	<i>pitA1</i> <i>pitB::Cat</i> <sup>r</sup> $\Delta$ <i>pstC345</i> $\Delta$ ( <i>phoB-phoR</i> ) Kan <sup>r</sup>	P1 ANCH1 X AN3902
ANCH1	$\Delta$ ( <i>phoB-phoR</i> ) Kan <sup>r</sup>	(304)
HR152	<i>pstA</i> <i>zhg::Tn10</i> <i>proC</i> <i>purE</i> <i>thyA</i> <i>gyrA</i> <i>metB</i> <i>his</i> <i>pyrF</i> Tc <sup>r</sup>	(59)
AN2537	$\Delta$ <i>pstC345</i>	(253)
AN3020	<i>pitA1</i> $\Delta$ <i>pstC345</i>	(284)
GS5	<i>pitA1</i> <i>pstA</i> <i>proC</i> <i>purE</i> <i>thyA</i> <i>gyrA</i> <i>metB</i> <i>his</i> <i>pyrF</i>	(59)
HR187	<i>pitA1</i> <i>pstA</i> <i>proC</i> <i>purE</i> <i>thyA</i> <i>gyrA</i> <i>metB</i> <i>his</i> <i>pyrF</i> <i>srlA::Tn10</i> <i>recA</i> Tc <sup>r</sup>	(59)
K-10	<i>pitA1</i>	(284)
JC7623	<i>recB21</i> <i>recC22</i> <i>sbsB15</i> <i>sbcC201</i> <i>thr-1</i> <i>ara-14</i> <i>leub6</i> <i>lacY1</i> <i>tsx33</i> <i>supE44</i> <i>galK2</i> <i>hisG4</i> <i>rfbD1</i> <i>rpsL21</i> <i>kdgK51</i> <i>xyl-5</i> <i>mtl-1</i> <i>argE3</i> <i>thi-1</i> $\Delta$ ( <i>gpt-</i> <i>proA</i> )62 not sure if 62/G2 or his64/G4	(137)
M15	<i>Nal</i> <sup>s</sup> <i>Str</i> <sup>s</sup> <i>Rif</i> <sup>s</sup> <i>thi</i> <sup>-</sup> <i>lac</i> <sup>-</sup> <i>Ara</i> <sup>+</sup> <i>gal</i> <sup>+</sup> <i>Mtl</i> <sup>r</sup> <i>F</i> <sup>-</sup> <i>recA</i> <i>Uvr</i> <sup>+</sup> <i>Lon</i> <sup>+</sup>	QIAexpress data sheet
JS5	<i>araD139</i> $\Delta$ ( <i>ara leu</i> )7697 $\Delta$ ( <i>lac</i> )X <sub>74</sub> <i>galU</i> <i>galK</i> <i>hsdR2</i> ( <i>rk</i> <sup>-</sup> <i>mk</i> <sup>-</sup> ) <i>mcrBC</i> <i>rpsL</i> (Str <sup>r</sup> ) <i>thi</i> <i>recA1/F::Tn10</i> (Tet <sup>r</sup> ) <i>proAB</i> <i>lacI</i> <sup>q</sup> <i>lacZ</i> $\Delta$ M15	Life Technologies
HB101	<i>F</i> <sup>-</sup> <i>leu</i> <i>b6</i> <i>pro</i> <i>A2</i> <i>recA13</i> <i>thi-1</i> <i>ara-14</i> <i>lacY1</i> <i>galK2</i> <i>xyl-5</i> <i>mtl-1</i> <i>rpsL20</i> <i>supE44</i> <i>hsdS20</i>	Life Technologies
K-12	<i>F</i> <sup>-</sup> <i>lambda</i> <sup>-</sup> <i>ilvG</i> <sup>-</sup> <i>rfb-50</i> <i>rph-1</i> (MG1655)	18
K-10	<i>Hfr</i> <i>garB10</i> <i>fhuA22</i> <i>ompF627</i> (T2R) <i>fadL701</i> (T2R) <i>relA1</i> <i>pit-10</i> <i>spoT1</i> <i>metB1</i> <i>rrnB-2</i> <i>mcrB1</i> <i>creC510</i>	(284, 289) and see below*

\* Skaar, P.D., A. Garen 1956. The orientation and extent of gene transfer in *Escherichia coli*. Proc.Natl.Acad.Sci.USA 42:619-624

Koch, J.P., H. Hayashi, E.C.C. Lin 1964. The control of dissimilation of glycerol and L-alpha-glycerophosphate in *Escherichia coli*. J.Biol.Chem. 239:3106-3108



**TABLE 2.2:** Description of *E. coli* strains containing plasmids referred to in this study

<b>Strain</b>	<b>Plasmid</b>	<b>Background strain</b>	<b>Construction/Reference</b>
AN3514	pBR322 control	<i>pitA1</i> $\Delta$ <i>pstC345</i>	pBR322/AN3066
AN3135	<i>pitB</i> <sup>+</sup> - long upstream	<i>pitA1</i> $\Delta$ <i>pstC345</i>	pAN656/AN3066
AN3171	<i>pitA</i> <sup>+</sup>	<i>pitA1</i> $\Delta$ <i>pstC345</i>	pAN686/AN3066
AN3531	<i>pitA</i> <sup>+</sup> (opposite orientation)	<i>pitA1</i> $\Delta$ <i>pstC345</i>	pAN920/AN3066
AN3511	<i>pitA</i> H225Q	<i>pitA1</i> $\Delta$ <i>pstC345</i>	pAN906/AN3066
AN3512	<i>pitA</i> D229N	<i>pitA1</i> $\Delta$ <i>pstC345</i>	pAN907/AN3066
AN3513	<i>pitA</i> K232Q	<i>pitA1</i> $\Delta$ <i>pstC345</i>	pAN908/AN3066
AN3776	<i>pitA</i> G213D	<i>pitA1</i> $\Delta$ <i>pstC345</i>	pAN1121/AN3066
AN3778	<i>pitA</i> I413D, G414L	<i>pitA1</i> $\Delta$ <i>pstC345</i>	pAN1123/AN3066
AN3937	<i>pitA1</i> PCR from AN3066	<i>pitA1</i> $\Delta$ <i>pstC345</i>	pAN1243/AN3066
AN3938	<i>pitA</i> G220D	<i>pitA1</i> $\Delta$ <i>pstC345</i>	pAN1244/AN3066
AN3903	pBR322 control	<i>pitA1</i> <i>pitB</i> ::Cat <sup>r</sup> $\Delta$ <i>pstC345</i>	pBR322/AN3902
AN3904	<i>pitA</i> <sup>+</sup>	<i>pitA1</i> <i>pitB</i> ::Cat <sup>r</sup> $\Delta$ <i>pstC345</i>	pAN920/AN3902
AN3905	<i>pitB</i> <sup>+</sup> - short upstream	<i>pitA1</i> <i>pitB</i> ::Cat <sup>r</sup> $\Delta$ <i>pstC345</i>	pAN1116/AN3902

**TABLE 2.3:** Description of *E. coli* plasmids used in this study

<b>Plasmid</b>	<b>Relevant characteristics</b>	<b>Construction/Reference</b>
pBR322	4.36 kb high copy number cloning vector, pMB1 replicon: Ap <sup>r</sup> , Tc <sup>r</sup>	(208)
pAN686	<i>pitA</i> <sup>+</sup>	<i>Cla</i> I/ <i>Sph</i> I fragment from pCE27 ((59)) in pBR322
pAN906	<i>pitA</i> (H225Q)	<i>Sal</i> I/ <i>Bam</i> H1 fragment from <i>pitA</i> (92-147) SDM in pBR322
pAN907	<i>pitA</i> (D229N)	<i>Sal</i> I/ <i>Bam</i> H1 fragment from <i>pitA</i> (92-148) SDM in pBR322
pAN908	<i>pitA</i> (K232Q)	<i>Sal</i> I/ <i>Bam</i> H1 fragment from <i>pitA</i> (92-149) SDM in pBR322
pAN920	<i>pitA</i> <sup>+</sup> (opposite orientation to pAN686)	<i>Sal</i> I/ <i>Bam</i> H1 fragment from pAN686 in pBR322
pAN910	<i>pitA</i> :: <i>hexahis</i> ( <i>pitA</i> <sup>+</sup> with a C terminal hexahistidine tag), in pQE60 with own ribosomal binding site	<i>Eco</i> RI/ <i>Bgl</i> II fragment from <i>pitA</i> (91-62/93-20) SDM in <i>Eco</i> RI/ <i>Bgl</i> II of pQE60
pAN683	pGEX vector with 30 bases of the spinach chloroplast $\epsilon$ -subunit fused to C terminal of GST.	A.J.W. Rodgers – this laboratory
pAN909	<i>pitB</i> <sup>+</sup> piggy backed behind the GST $\epsilon$ -30 fusion gene	<i>Eco</i> RI fragment from <i>pitB</i> (91-70) SDM in <i>Eco</i> RI of pAN683
pAN730	<i>pitB</i> <sup>+</sup> in the opposite orientation to pAN909	<i>Eco</i> RI fragment from <i>pitB</i> (91-70) SDM in <i>Eco</i> RI of pAN683
pAN914	<i>pitA</i> :: <i>hexahis</i> (no promoter) piggy backed behind the GST $\epsilon$ -30 fusion gene	<i>Eco</i> RI/ <i>Pvu</i> II fragment from pAN910 in <i>Eco</i> RI of pAN683
pAN915	<i>pitA</i> :: <i>hexahis</i> in opposite orientation to pAN914 (ie no read through from <i>tac</i> promoter)	<i>Eco</i> RI/ <i>Pvu</i> II fragment from pAN910 in <i>Eco</i> RI of pAN683
pAN918	<i>pitA</i> :: <i>hexahis</i> in pBR322	<i>Mlu</i> I/ <i>Pvu</i> II fragment from pAN910 in <i>Mlu</i> I/ <i>Pvu</i> II of pAN686
pAN919	<i>pitA</i> <sup>+</sup>	Remove <i>Pvu</i> II fragment from pAN686
pAN656	<i>pitB</i> <sup>+</sup> - long upstream region (1403 nucleotides)	<i>Cla</i> I/ <i>Bam</i> H1 fragment from AN2538 in pBR322
pAN1116	<i>pitB</i> <sup>+</sup> - short upstream region (206 nucleotides)	<i>Ssp</i> I/ <i>Cla</i> I fragment from pAN656 in <i>Eco</i> RV/ <i>Cla</i> I sites of pBR322
pAN1121	<i>pitA</i> G213D mutation, <i>Bgl</i> II site	<i>Bgl</i> II site at A213 in pAN686
pAN1123	<i>pitA</i> I413D, G414L mutation, <i>Bgl</i> II site	<i>Bgl</i> II site at I413 in pAN686

pAN1243 *pitA1*

pAN1244 *pitA* G220D mutation

PCR product from AN3066  
in pBR322  
Site-directed mutagenesis of  
G658A in *pitA*<sup>+</sup> ORF on  
pAN686

---



**TABLE 2.4:** Description of bacteriophages used in this study

<b>Bacteriophage</b>	<b>Relevant characteristics</b>	<b>Source/Reference</b>
M13mp18	7.25 kb, <i>lacpoZ'</i>	(188)
M13mp19	7.25 kb, <i>lacpoZ'</i>	(188)
P1 <sub>kc</sub>	Derivative of bacteriophage P1 which forms relatively clear plaques but is able to lysogenise bacterial hosts	(103)

**TABLE 2.5:** Description of oligonucleotides used in this study (sorted by function)

<b>Name</b>	<b>Sequence (5' to3')</b> <sup>a b c</sup>	<b>Function</b>	<b>Source/Ref.</b>
95-87	ctgaggatccgcccgcgttcacgtct	PCR full <i>pitA</i> + strand	Sections 3.3, 4.2
95-88	tgacggATCcgtttggtgcgtacgattacag	PCR full <i>pitA</i> - strand	Sections 3.3, 4.2
95-89	gtcaggatccatgcgtccgttcgtaattc	PCR full <i>pitB</i> + strand	Sections 3.3, 4.2, 5.2
95-90	cgccggatccgggcattttcaggaag	PCR full <i>pitB</i> - strand	Sections 3.3, 4.2, 5.2
97-73	agagGgatCCtgaaccgtaattg	PCR full <i>pitB</i> plus upstream 319 bp	Section 4.2
97-74	cactggaTcCggtggttggtgatg	PCR full <i>pitB</i> plus upstream 319 bp	Section 4.2
91-62	gataatgcgcccGAAttcatgtcc	<i>EcoRI</i> SDM to insert <i>pitA</i> in pQE60	Section 3.4
93-20	cagttcctgAGatcTtacgcacccaaaac	<i>BglII</i> SDM to insert <i>pitA</i> in pQE60	Section 3.4
93-107	cgtaacgtcctGtgctacatttg	SDM <i>pitA</i> start (ATG to GTG)	Section 3.5
93-108	cgtaacgtcctCtgctacatttg	SDM <i>pitA</i> start (ATG to CTG)	Section 3.5
93-135	atctaatatGtgctaaatttatttg	SDM <i>pitB</i> start (ATG to GTG)	Section 3.5
93-136	atctaatatCtgctaaatttatttg	SDM <i>pitB</i> start (ATG to CTG)	Section 3.5
94-55	ccaacgcccgtgTAAaccgttateta	SDM <i>pitA</i> in frame TAA stop at 118bp	Section 3.5
94-56	gctgggtggtTAGagtgttgcc	SDM <i>pitA</i> in frame TAG stop at 208bp, <i>pitA</i> sequencing	Section 3.5 Section 4.2
94-57	atatgggatAgtctcatggcctt	SDM <i>pitA</i> in frame TAG stop at 268bp, <i>pitA</i> sequencing	Section 3.5 Section 4.2
95-99	tcgacaaagtggatgaagatg	<i>pitA</i> sequencing	Section 4.2
94-104	accgtgaccagcattctg	<i>pitA</i> sequencing	Section 4.2
94-102	gttgcctatgccattgtcc	<i>pitB</i> sequencing	Section 4.2
94-105	tcggcctggatcattgcgg	<i>pitB</i> sequencing	Section 4.2
94-106	aacatgaatgcgtccggc	<i>pitB</i> sequencing	Section 4.2
93-130	ctagcgaaactgccagge	<i>pitB</i> sequencing	Section 4.2
95-5	cagtgtctatcggtcttgc	<i>pitB</i> sequencing	Section 4.2
95-101	tccgctatcgAcgtggcgttttc	SDM <i>pitA</i> G220D , recreating G to A at 658 bp in <i>pitA</i> ORF ( <i>pitA1</i> )	Section 4.3.2
92-147	cgtttctgcaGggcgcgaacg	SDM <i>pitA</i> H225Q	Section 6.3.1
92-148	cggcgcgaacAatggtcagaa	SDM <i>pitA</i> D229N	Section 6.3.1
92-149	cgatggtcagCaaggcattgg	SDM <i>pitA</i> K232Q	Section 6.3.1

Name	Sequence (5' to 3') <sup>a b c</sup>	Function	Source/Ref.
96-02	cccgcgcgatgcgAGAtcTgctcgccgtgg ttatg	SDM <i>pitA</i> S50D, Q51L ( <i>Bgl</i> II site in loop 1)	Section 6.3.2
95-25	atatgggatacAGAtcTtggccttgcc	SDM <i>pitA</i> S91D, H92L ( <i>Bgl</i> II site in loop 2)	Section 6.3.2
95-26	ccatctctcaAGATctgattggcgcga	SDM <i>pitA</i> H122Q, T123D ( <i>Bgl</i> II site in loop 3)	Section 6.3.2
96-28	ccgaaagtattaagtatAGATCTttctctgat cgtttccc	SDM <i>pitA</i> F156D, G157L ( <i>Bgl</i> II site in loop 4)	Section 6.3.2
96-01	gctcgcctactggagAgATCTcaagaaacg cgccccg	SDM <i>pitA</i> S182R, G183D, T184L ( <i>Bgl</i> II site in loop 5)	Section 6.3.2
94-51	ctggacgcgtatAgATctgatcctttccgc	SDM <i>pitA</i> A213D ( <i>Bgl</i> II site in loop 5)	Section 6.3.2
94-52	gaagaaactgaagtcAgaTCtgettagcacca tcg	SDM <i>pitA</i> M175L ( <i>Bgl</i> II site in loop 6/7)	Section 6.3.2
94-53	cgactatcggtgagaaaGATCTtaagaaag gcatacctacgc	SDM <i>pitA</i> I413D, G414L ( <i>Bgl</i> II site in loop 7/8)	Section 6.3.2
95-83	gtctatcggcctggcAGAtCTtaccgggatg ccggttcc	SDM <i>pitA</i> S438D, Y439L ( <i>Bgl</i> II site in helix 8/9)	Section 6.3.2
95-28	aaccgtgacAGATCTtctgatggcctgg	SDM <i>pitA</i> S472D, I473L ( <i>Bgl</i> II site in loop 8/9)	Section 6.3.2
95-84	ctggctctccttgcaAGATctgtaatcgtacg cacc	SDM <i>pitA</i> F498D ( <i>Bgl</i> II site at C- terminus)	Section 6.3.2
92-150	gcttcgtggtTaacaatgaatgc	<i>pitA</i> <i>Hpa</i> I at V251, <i>pitA</i> sequencing	Section 6.4.1 Section 4.2
92-151	cgcgctggcCttaggtatcgggt	<i>pitA</i> <i>Sau</i> I at L394, <i>pitA</i> sequencing	Section 6.4.1 Section 4.2
93-18	gcttcgtcgtTaaCatgaatgcgctc	<i>pitB</i> <i>Hpa</i> I at V251	Section 6.4.1
92-153	ggtagcactggcCTtAggcattggcacc	<i>pitB</i> <i>Sau</i> I at L394	Section 6.4.1
93-76	cgcgctgttGaccggctcacc	<i>pitB</i> $\Delta$ <i>Hpa</i> I at L138	Section 6.4.1
93-77	ctggaacctgggCacGtggtactttggtttac	<i>pitA</i> <i>Bbr</i> P1 at G110	Section 6.4.1
93-78	ggaacctgggCacgtggttctctcg	<i>pitB</i> <i>Bbr</i> P1 at G110	Section 6.4.1
93-79	ggtctgatttcttctgctTcgAAgGtactggagc ggcacc	<i>pitA</i> <i>Sfu</i> I at R178	Section 6.4.1
93-80	gatattcctgctTcgaAggtactggagcggg	<i>pitB</i> <i>Sfu</i> I at R178	Section 6.4.1



- <sup>a</sup> Recognition sequences for restriction endonuclease enzymes are underlined.
- <sup>b</sup> Altered bases in site-directed mutagenesis oligonucleotides are capitalised.
- <sup>c</sup> Bold nucleotides indicate the focal sequence for the site-directed mutagenesis.

10mM MgCl<sub>2</sub> and 10mM MgSO<sub>4</sub> was added. Glucose-Z broth consisted of Luria Bertani media plus 33mM glucose and 2.5mM CaCl<sub>2</sub>.

The minimal media used was the half-strength Medium 56 of Monod *et al* (177), consisting of 30mM KHPO<sub>4</sub>, 20mM NaH<sub>2</sub>PO<sub>4</sub>, 0.4mM MgSO<sub>4</sub> and 8mM NH<sub>4</sub>SO<sub>4</sub>. Trace elements, prepared as a 1000 fold stock were added to this medium immediately prior to sterilisation, to give final concentrations of 13.9µM ZnSO<sub>4</sub>, 1.0µM MnSO<sub>4</sub>, 4.69µM H<sub>3</sub>BO<sub>3</sub>, 0.68µM CoSO<sub>4</sub>, 2.45µM CaCl<sub>2</sub> and 1.78µM FeCl<sub>3</sub> (75).

Soft agars were prepared by adding 1mM MgCl<sub>2</sub> and 8g Bacto-agar per litre of minimal media (made as described above). Soft H top agar was prepared with 10g tryptone, 8g NaCl and 15g Bacto-agar per litre of water. H plates were prepared by increasing the Bacto-agar to 15g.

“Pi media” was an inorganic phosphate free media supplemented with 500µM NaH<sub>2</sub>PO<sub>4</sub>. This consisted of 100mM Tris-HCl pH7.4, 10mM KCl, 15mM (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>, 1mM MgCl<sub>2</sub>, 0.1% Casamino acids, 3µM thiamine and 33mM glucose. When appropriate 50µg/ml chloramphenicol, or 70µg/ml ampicillin were added. Auxotrophic requirements which were supplied where necessary included 0.008% arginine, 0.002% uracil, 30µM 2,3-dihydroxybenzoate, 0.17mM methionine, 0.008% glycine, 0.003% valine and 0.003% isoleucine (211). 1mM α-glycerol-3-phosphate (G3P) was used as an organic phosphate source for potential Pi auxotrophs.

Solid media was prepared by the addition of 20g Bacto-agar per litre liquid media prior to sterilisation.

TE buffer consists of 10mM Tris-HCl and 1mM EDTA, pH8.0, sterilised by autoclaving.

Antibiotics were prepared as stock solutions and stored at -20°C. Ampicillin stocks were 100mg/ml in H<sub>2</sub>O, and were normally used at 100µg/ml in culture and solid media. Chloramphenicol was prepared at 60mg/ml in ethanol, and was often used at 100µg/ml

in culture and solid media, although weaker dilutions were also prepared. Kanamycin was stored at 100mg/ml in H<sub>2</sub>O, and the standard dilution was to 50µg/ml. Streptomycin stock was 100mg/ml in H<sub>2</sub>O, and was used at 200µg/ml unless otherwise stated. Tetracycline was stored at 10mg/ml in 70% ethanol, and diluted to 10µg/ml.

## **2.4 General molecular biology techniques**

### **2.4.1 Small scale purification of plasmid DNA or replicative form M13 DNA**

Small scale isolation of plasmid or M13 replicative form (RF) DNA was carried out using the Magic and Wizard Minipreps DNA Purification Systems (Promega Corporation, Madison, WI, USA) following the manufacturers instructions. Strains containing plasmids were grown to stationary phase at 37°C in 5ml of Luria Bertani medium with the appropriate antibiotics/supplements. Alternatively, M13 RF DNA was purified from plaques by inoculating 5mls of TG1 cultured in 2xYT media with a single M13 plaque or 10µl of phage stock, then incubating this culture at 37°C with shaking for 3.5 hours. 1.5ml of either the plasmid or M13 RF culture was transferred into a polypropylene microcentrifuge tube and pelleted by centrifugation for 1 min at 14000xg in an Eppendorf microcentrifuge. The cells were resuspended in 200µl of cell resuspension solution (50mM Tris-HCl (pH7.5), 10mM EDTA, 100mg/ml RNase A). 200µl of cell lysis solution (0.2M NaOH, 1% SDS) was gently mixed through the suspension by inverting the tube several times until the solution became clear. A thick white precipitate formed after addition of 200µl of neutralisation solution (1.32M potassium acetate, pH4.8) and after thorough mixing this was pelleted by centrifugation at 14000xg for 5 mins. The supernatant was transferred to a new tube and 1ml of Magic or Wizard Minipreps DNA purification resin was added. The DNA/resin mixture was loaded onto a minicolumn attached to a vacuum manifold via a disposable syringe barrel, and the resin was washed with 2ml of column wash solution (84mM NaCl, 8.4mM Tris-HCl (pH7.5), 5mM EDTA, 55% v/v ethanol). The minicolumn was then attached to a microcentrifuge tube and briefly centrifuged to remove all wash solution. A fresh microcentrifuge tube was attached to the minicolumn, which was incubated with 50µl of preheated (65°C) water for 1 min, before eluting the DNA by centrifugation at 14000xg.



### **2.4.2 Large scale purification of plasmid DNA by Caesium chloride (CsCl) density gradient**

The method used for the large scale preparation of high purity plasmid or M13 replicative form (RF) DNA was a modification of Selker *et al* (238). A 1 litre culture of the appropriate *E. coli* strain was grown to stationary phase in Luria Bertani media plus the appropriate antibiotics and supplements. Bacteria were harvested by centrifugation at 8000xg (5000rpm in GSA Sorvall rotor) for 5 mins at 4°C, and the bacterial pellet was resuspended in 20ml of buffer containing 50mM Tris-HCl (pH8) and 25% sucrose, and incubated on ice for 5 mins. 8ml of lysozyme solution (10mg/ml in 0.25M Tris-HCl, pH8) was gently swirled into the cell suspension, which was left on ice for 5 mins. An addition of 4ml 0.25M EDTA was made, with swirling, and the spheroplasts were incubated on ice for a further 5 mins. 32ml of Triton lysis buffer (50mM Tris-HCl (pH8), 62.5mM EDTA, 0.3% v/v Triton X-100) was added and stirred for 5 mins to lyse the spheroplasts. Cell debris was pelleted by centrifugation for 20 mins at 48000xg at 4°C. One quarter volume of 5M NaCl was added to the supernatant with stirring, followed by one quarter volume of 50% PEG-6000 in 50mM Tris-HCl (pH8). This was mixed gently in ice, and incubated on ice for at least 2 hrs to precipitate the plasmid DNA. The precipitate was collected by centrifugation at 3000xg at 4°C for 5 mins, and the pellet was resuspended in 7ml TE buffer (10mM Tris-HCl (pH8), 1mM EDTA) and transferred to a Sorvall SE12 centrifuge tube. 7.2g CaCl<sub>2</sub> was dissolved into this solution and centrifuged for 20 mins at 2300xg at room temperature. The layer of protein at the top of the tube was removed with a spatula and the remaining solution was transferred into a 50Ti Beckman centrifuge tube and 300µl of 10mg/ml ethidium bromide was added. The tubes were protected from light, filled with paraffin oil, sealed and centrifuge at 140000xg at 15°C for approximately 42hrs (brake off). The lower band containing supercoiled plasmid DNA was removed with an 18 gauge needle attached to a syringe through the side of the tube. Ethidium bromide was removed by the addition of isopropanol (stored over NaCl-saturated TE) followed by vortexing, and this extraction was repeated 5 times, or until all red colouration had been removed from the aqueous phase. Dialysis against 1 litre of TE buffer at 4°C was used to remove CsCl, and the plasmid solution was stored at 4°C.

### **2.4.3 Purification of chromosomal DNA from *E. coli***

Genomic DNA from *E. coli* strains was isolated from a stationary phase culture by resuspending the cells pelleted from 1.5mls of culture in 567 $\mu$ l TE, then adding 30 $\mu$ l of 10% SDS and 3 $\mu$ l of 20mg/ml proteinase K, giving a final concentration of 100 $\mu$ g/ml proteinase K in 0.5% SDS. This was mixed thoroughly and incubated 1 hr at 37°C. 100 $\mu$ l of 5M NaCl was added to this viscous solution to prevent nucleic acids precipitating with the cell wall debris, denatured protein and polysaccharides upon the addition of hexdecyltrimethylammonium bromide (CTAB). 80 $\mu$ l of 10% CTAB in 0.7M NaCl was added to this mixture and incubated 10 min at 65°C. An approximately equal volume (~0.8ml) of 24:1 chloroform/isoamyl alcohol was added, mixed thoroughly and spun 5 mins in a microcentrifuge. The aqueous supernatant was placed in a fresh tube and an equal volume of 25:24:1 phenol/chloroform/isoamyl alcohol was added, mixed thoroughly and spun in a microcentrifuge for 5 min. The supernatant was transferred to a fresh tube and 0.6 vol isopropanol was added to precipitate the nucleic acids, which become visible upon shaking as a stringy white DNA precipitate. This DNA was transferred to a fresh tube containing 70% ethanol, spun for 5 mins at room temperature to pellet the DNA, and dried in a vacuum desiccator. The chromosomal DNA was then resuspend in 100 $\mu$ l TE buffer.

### **2.4.4 Preparation of single-stranded M13 DNA**

Single stranded DNA used in site-directed mutagenesis or DNA sequencing was prepared from M13 clones by the method of Sanger *et al* (231) as described in the technical notes for the Oligonucleotide-directed *in vitro* mutagenesis system version 2.1 (Amersham). 50 $\mu$ l of a fresh overnight culture of TG1 grown in 2xYT medium was added to 5ml 2xYT media. This was inoculated with either a single M13 plaque grown on a lawn of TG1 cells, or 10 $\mu$ l of phage stock, and cultured for 5 hrs at 37°C. Cells were transferred to 5 microcentrifuge tubes and pelleted by centrifugation for 5 mins in a benchtop centrifuge. The supernatants were transferred to 5 fresh tubes, being careful to keep these free of whole cells, and 200 $\mu$ l of PEG/NaCl (20%w/v polyethylene glycol 6000, 2.5M NaCl) was added, shaken and then left to stand for 15 mins. After centrifugation for 5 mins, the supernatant was discarded and the tubes were centrifuged another 2 mins to facilitate removal of all traces of PEG. Pellets were resuspended in 100 $\mu$ l TE buffer and contaminants were extracted with phenol followed by diethyl



ether. The single-stranded M13 DNA was then precipitated with 10 $\mu$ l 3M sodium acetate, pH6, and 250 $\mu$ l ethanol, and the precipitate was collected by centrifugation. The resulting DNA pellets were washed with ethanol, dried in a vacuum desiccator and resuspended in a total volume of 100 $\mu$ l TE buffer.

#### **2.4.5 Spectrophotometric analysis of DNA**

Appropriately diluted DNA was assayed at 260 and 280nm using a Hitachi U-1100 Spectrophotometer. An absorbance of 1.0 at 260nm was taken as equivalent to a concentration of 50 $\mu$ g/ml for double stranded DNA and 20 $\mu$ g/ml for oligonucleotides. DNA purity was assessed by dividing  $A_{260}$  by  $A_{280}$ , which gives a ratio of between 1.8 and 2.0 for DNA uncontaminated by RNA or protein (230). When a more exact oligonucleotide concentration was required 15.6  $A_{260}$  units/ $\mu$ l per base of oligonucleotide was considered equivalent to 1pmol/ $\mu$ l DNA (from Amersham Sculptor<sup>TM</sup> *in vitro* mutagenesis manual).

#### **2.4.6 Agarose gel electrophoresis**

Agarose gel electrophoresis was used to separate DNA fragments by size. Generally an agarose concentration of 0.8% was used in TAE buffer (40mM Tris-HCl, 20mM sodium acetate, 1mM EDTA, pH7.8), although the agarose concentration could vary between 0.8 and 2%. Loading dye (470mM sucrose, 0.75mM bromophenol blue, 48mM EDTA) was used at a ratio of approximately 1:5, with sample volumes of 10 - 20 $\mu$ l. Gels were run at a constant current of approximately 180mA, until the bromophenol blue dye front had migrated three-quarters of the length of the gel. DNA was stained in 1 $\mu$ g/ml ethidium bromide solution for 20 mins and visualised by exposure to ultra-violet light.

#### **2.4.7 "Rapid cracking" of phage or colonies**

This method was used as a rapid screen for the desired DNA product after a subcloning reaction (checking for the presence of insert in the phage/plasmid) and was used when the expected efficiency of the ligation reaction was low. A small amount of cells from a single colony or 10 $\mu$ l of the phage supernatant from a culture greater than 5 hrs old was suspended in 40 $\mu$ l colony cracking buffer (50mM NaOH, 10% glycerol, 5mM EDTA, 0.5% SDS, 0.025% bromocresol green) and heated to 70°C for 30-60 mins.



This was then run on an 0.8% agarose gel and DNA bands were visualised as described in Section 2.4.6.

### **2.4.8 Synthesis and preparation of oligonucleotides**

Oligonucleotides were synthesised by the Biomolecular Resource Facility at the ANU and were supplied as 40nmol oligonucleotide in an ammonia solution. These were freeze-dried and resuspended in 600 $\mu$ l water for purification by HPLC on a MAC7Q ion exchange column. The HPLC column was equilibrated with 40mM Tris-HCl pH8 before loading 300 $\mu$ l of oligonucleotide solution. A gradient of zero to 1M NaCl was applied, and the absorbance at 280nm was used to monitor the elution of the oligonucleotide, which usually formed a discrete peak. The collected DNA solution (max volume 2.5ml) was desalted and freeze dried overnight. This DNA was resuspended in approximately 400 $\mu$ l of a 10 times dilution of TE, and DNA concentration was determined by spectrophotometry as described in Section 2.4.5. When necessary oligonucleotides were phosphorylated, as described in Section 2.5.4.

### **2.4.9 Polymerase chain reaction (PCR)**

Polymerase chain reactions (180, 228) were performed in a Perkin Elmer Cetus Thermal Cycler. *Taq* polymerase supplied by Promega, was used as the thermostable DNA polymerase in experiments where fidelity in the sequence was not an absolute requirement (such as when screening for the presence or absence of a gene). However the high base incorporation error rate associated with this enzyme ( $2.0 \times 10^{-5}$  (156)) precluded its use when accuracy in the DNA sequence was vital. *Pfu* polymerase, supplied by Stratagene, has been shown to have a 12-fold lower error rate ( $1.6 \times 10^{-6}$  (156)) and was used to produce DNA fragments whenever sequence fidelity was important.

A typical reaction mixture, using 2 units of *Taq* polymerase, contained 20pmol of each oligonucleotide primer, 250ng template DNA, 200 $\mu$ M of each dNTP, buffer (supplied with the enzyme -50mM KCl, 10mM Tris-HCl pH9.0, 0.1% Triton X-100) and 2.5mM MgCl<sub>2</sub> in a total volume of 50  $\mu$ l. *Pfu* polymerase reactions were as above, using 2 units enzyme, buffer (supplied with enzyme - 20mM Tris-HCl pH8.2, 10mM KCl, 6mM NH<sub>4</sub>SO<sub>4</sub>, 2mM MgCl<sub>2</sub>, 0.1% Triton X-100, 10ng/ml nuclease free Bovine Serum

Albumin). Mineral oil was overlaid to prevent evaporation. An estimate of the optimum annealing temperature for a primer was calculated by adding 2°C for each A and T nucleotide present in the primer and 4°C for each G and C nucleotide, and subtracting 5°C from the result. Generally 30 cycles of denaturation, annealing and extension were performed. Denaturation was at 94°C for 1 min, annealing at 55-65°C (depending on the sequence of the oligonucleotide primers) for 1 min and extension at 72°C for 2 mins.

The *pitA* and *pitB* genes were amplified from genomic DNA, which was either purified (Section 2.4.3) or supplied by using approximately 2µl of a small amount of a single *E. coli* colony mixed in 100µl of water.

#### **2.4.10 Purification of PCR products**

All traces of mineral oil were removed from PCR reactions by running the solutions over a clean sheet of Parafilm® laboratory film. The oil adheres to the film, allowing the aqueous solution to be isolated. Purification was carried out using a BresaClean Kit (Bresatech). Three volumes of NaI solution was added to the DNA solution, followed by 5µl of resuspended BresaClean silica resin. The suspension was mixed and incubated for 5 mins, then centrifuged for 5 secs in a bench top centrifuge and the supernatant was discarded. The silica/DNA pellet was resuspended in 500µl of BresaClean wash solution followed by centrifugation. The wash solution was removed as completely as possible and the pellet was dried and then resuspended in sterile distilled water and incubated for 5 mins at 50°C. The silica resin was pelleted by centrifugation for 1 min at top speed, and the supernatant containing the PCR products was transferred into a fresh tube and centrifuged again to ensure complete removal of any residual silica resin.

#### **2.4.11 Site-directed mutagenesis**

Site-directed mutagenesis was carried out using either the Oligonucleotide-directed in vitro mutagenesis system version 2.1 (Amersham), or the more recent Sculptor™ in vitro mutagenesis system (Amersham). These systems are based on the phosphorothioate technique of Fritz Eckstein (234). The methods are described in detail in the handbooks accompanying the kits, and are summarised as follows. A

phosphorylated mutagenic oligonucleotide was annealed to single-stranded DNA (ssDNA) template. Elongation and ligation of the second strand of DNA was carried out using the Klenow fragment (Version 2.1) or T7 DNA polymerase (Sculptor™) and T4 DNA ligase, plus dATP, dGTP, dTTP and dCTP $\alpha$ S. Any remaining ssDNA (non-mutant) was removed either by filtration (Version 2.1) or digestion with T5 exonuclease (Sculptor™). The non-mutant strand of the double-stranded DNA (dsDNA) template was then nicked with *Nci*I, as the incorporated thionucleotide dCTP $\alpha$ S protects *Nci*I sites on the mutant DNA strand. Exonuclease III was used to remove the majority of the nicked non-mutant DNA, leaving primers for repolymerisation by DNA Polymerase I followed by ligation with T4 DNA ligase. The dsDNA was transformed into competent TG1 cells, and ssDNA was isolated and sequenced from single plaques, as described in Sections 2.4.4, 2.4.12 and 2.4.13.

#### **2.4.12 Manual sequencing of DNA**

Nucleotide sequencing of single-stranded and double-stranded DNA was carried out with the T7 Sequencing Kit (Pharmacia) which uses the dideoxynucleotide chain termination method of Sanger *et al* (232). Sequencing reactions were performed according to the methods described in the handbook accompanying the kit, and are summarised below:

For single stranded DNA, the annealing reactions contained 1.5-2 $\mu$ g template, and 2-4pmol oligonucleotide primer in annealing buffer. Double stranded DNA template was denatured with NaOH as described in the handbook, using 1.5-2 $\mu$ g DNA. This DNA was ethanol precipitated and resuspended in water. 5-10pmol of primer was used in the annealing reaction. Labeling reactions were carried out using a mixture of dGTP, dTTP, dCTP and [<sup>35</sup>S]dATP $\alpha$ S with the enzyme T7 DNA polymerase. Chain termination reactions were commenced by transferring aliquots from this tube into 4 labeled tubes, each containing the 4 deoxynucleoside-5'-triphosphates, and one of the dideoxynucleoside-5'-triphosphates. Reactions were halted by the addition of stop solution, which contained formamide, EDTA, xylene cyanol and bromophenol blue.



Electrophoresis gels were prepared using Long Ranger™ gel solution (AT Biochem) which is stated by the manufacturer to contain a modified acrylamide, a co-monomer and a cross-linker (unspecified). To prepare 50ml of gel solution 21g of urea was dissolved in 6ml 10xTBE (Tris-Borate-EDTA buffer), 5ml Long Ranger™ solution, and water to 50ml. 10xTBE buffer consisted of 108g Tris Base, 55g Boric acid, 40ml of 0.5M EDTA pH8 and water to 1000ml. The gel solution, containing a final concentration of TBE buffer at 1.2x, was polymerised by the addition of 25µl TEMED and 300µl of 10% ammonium persulphate.

Electrophoresis gels were run in a Base Runner Nucleic Acid Sequencer (IBI/Kodak) between 21.6cm x 45cm or 21.6cm x 60cm glass plates, using 0.4mm spacers. Gel-Slick solution (IBI/Kodak) was applied to one plate to prevent adherence of the gel. A flat 0.4mm spacer was placed in the top of the poured gel, which was allowed to polymerise for one hour at room temperature. The gel was pre-run with 1xTBE buffer in the buffer reservoir tanks at 30W (for 21.6cm x 45 cm plates) or 45W (for 21.6cm x 60cm plates) for 30 mins to pre-warm the gel. A sharks tooth comb was then placed at the top of the gel to allow loading of up to 32 samples.

Samples were heated at 100°C for 2 mins to denature the DNA and 3µl of each sample was loaded onto the gel, which was electrophoresed, using the conditions described above for up to 5 hrs, depending on the length of sequence to be read. Where appropriate, second applications of the samples were made 2 hrs after the initial loading, to increase the length of DNA sequence able to be read from this gel. Also 50ml of 3M sodium acetate could be added to the bottom buffer tank 90 mins before the expected finish of the electrophoresis to compress the bands at the bottom of the gel, thus increasing the length of readable sequence.

The gel was then adhered to 3MM chromatography paper (Whatman), and dried under vacuum at 80°C for 2 hrs. Autoradiography was carried out using Hyperfilm MP autoradiography film (Amersham) for 20-25 hrs, then developed in an X-omat M20 automatic processor (Eastman Kodak).

### **2.4.13 Automated sequencing of DNA**

Double-stranded DNA (dsDNA) was prepared for sequencing on an Applied Biosystems DNA Sequencing System using the PRISM™ Ready Reaction DyeDeoxy™ Terminator Cycle Sequencing Kit (Applied Biosystems). Primer and template (usually dsDNA from either ‘minipreps’, see Section 2.4.1, or CsCl purified plasmid DNA, see Section 2.4.2 or single-stranded DNA (ssDNA), see Section 2.4.4) are added to supplied reagents to perform fluorescence-based dideoxy sequencing reactions with the thermally stable enzyme AmpliTaq® DNA Polymerase. Reactions were prepared as described in this protocol, except the amount of primer used for dsDNA templates was increased from 3.2pmol to 6pmol. Cycle sequencing was carried out on a Perkin-Elmer Cetus thermal cycler for 25 cycles using the program recommended in this manual. Phenol/choloroform extraction was carried out as described, except that the mineral oil used during the thermocycling was removed by rolling the reaction mixture plus 80µl H<sub>2</sub>O down a piece of clear Parafilm® laboratory film. The oil sticks to the film, so the aqueous solution can easily be removed with a pipette.

Electrophoresis and imaging was carried out by the Biomolecular Resources Facility at the John Curtin School of Medical Research, ANU.

## **2.5 Recombinant DNA techniques**

### **2.5.1 Restriction endonuclease digestion of DNA**

Restriction endonuclease digestion of DNA was used to prepare suitable ‘blunt’ or ‘sticky’ 5’ and 3’ ends for subcloning DNA fragments into suitable vectors, to screen for successful subclonings, and to verify certain DNA sequences or site-directed mutagenesis experiments. Digestion of DNA for analytical purposes was usually carried out using 0.1-1µg DNA in a 20µl volume. Restriction endonucleases were at 0.5-5 units per reaction in the appropriate buffer supplied by the manufacturer, and were normally incubated at 37°C for 60 mins. 5-10µl was then analysed by agarose gel electrophoresis (Section 2.4.6). Preparative scale digestion of DNA was usually carried out in 50-100µl volume with 1-10µg DNA and 1-10 units of restriction endonuclease in the appropriate buffer. These reactions were incubated at 37°C for between 1 and 3 hrs

and when required specific DNA fragments were isolated using agarose gel electrophoresis (Section 2.4.6).

### **2.5.2 Elimination of 5' and 3' protruding ends**

DNA fragments with protruding 5' termini were converted to blunt ends by filling the ends in with the DNA polymerase activity of the Klenow fragments of *E. coli* DNA polymerase 1. Up to 1 $\mu$ g of DNA was incubated with 1-2 units of the Klenow fragment, 25  $\mu$ M dNTPs in a volume of 50 $\mu$ l. The reaction was incubated at room temperature for 10 mins, then stopped by heating to 70°C for 5 mins to inactivate the enzyme. The DNA was then purified with the BresaClean method as described in Section 2.4.10.

Removal of protruding 3' termini was carried out with bacteriophage T4 DNA polymerase, which can also fill in protruding 5' termini with a 5' to 3' polymerase activity. T4 DNA polymerase was added directly to the DNA after restriction endonuclease digestion at a concentration of 2 units per  $\mu$ g DNA, along with the desired dNTPs at 0.1mM. The mixture was incubated at 12 for 15 mins, then heated to 75°C for 10 mins to heat kill all enzymes.

### **2.5.3 Removal of terminal phosphates from linearised DNA**

Terminal phosphates were removed from linearised vector DNA to minimise religation of the vector DNA. Linearised vector DNA was dephosphorylated by adding approximately 1 unit shrimp alkaline phosphatase per microgram DNA to the completed restriction endonuclease digestion (as this enzyme is active in almost all buffer conditions) and this mixture was incubated at 37°C for 60 mins, then heat killed by incubation at 65°C for 15 mins. The DNA was then purified using the BresaClean method as described in Section 2.4.10.

### **2.5.4 Phosphorylation of oligonucleotides**

When necessary oligonucleotides were phosphorylated using 2-4.5 units of T4 Polynucleotide kinase (Pharmacia), kinase buffer (100mM Tris-HCl pH8.0, 10mM MgCl<sub>2</sub>, 7mM dithiothreitol, 1mM ATP) and 48pmol oligonucleotide for 15 mins at



37°C, followed by heat inactivation at 70°C for 10 mins. The phosphorylated oligonucleotide was used immediately or stored at -20°C.

### **2.5.5 Purification of DNA fragments from agarose gels**

DNA fragments were separated by agarose gel electrophoresis and visualised via staining with ethidium bromide, as described in Section 2.4.6. The appropriate DNA band was excised from the gel using a sterile scalpel and transferred into a tube wrapped in tin foil. Purification of the DNA was carried out using the BresaClean Kit (BresaTech). NaI solution was added to the gel slice at 1ml per gram and this was incubated at 55°C with occasional mixing until the gel slice had dissolved. 5µl of resuspended BresaClean resin was added, and the DNA was isolated as described in Section 2.4.10.

### **2.5.6 Ligation of DNA**

#### *'Sticky' ends*

DNA fragments with sticky ends were ligated with a molar ratio of 1 part vector per 2 to 15 parts insert with T4 DNA ligase (0.5 to 1 unit) in a total volume of 20µl. The reaction mix, which also contained 30mM Tris-HCl pH7.8, 10mM MgCl<sub>2</sub>, 10mM DTT and 1mM ATP was incubated at 16°C overnight, or at room temperature for 4-5 hrs.

#### *'Blunt' ends*

Blunt end ligations were carried out with approximately 100ng dephosphorylated vector and a 10 to 20 fold molar excess of insert. The buffer contains 50mM Tris-HCl pH7.6, 10mM MgCl<sub>2</sub>, 5% PEG 8000, 1mM DTT and 1mM ATP in a total volume of 20µl. Approximately 6 units of T4 DNA ligase was added, and this mixture was incubated at 16°C overnight, then the enzyme was heat killed at 65°C for 15 mins.

### **2.5.7 Preparation of competent *E. coli***

Competent cells were prepared by several methods. When small amounts of an unusual strain were needed, fresh competent cells were prepared using the CaCl<sub>2</sub> method described in the Amersham oligonucleotide-directed *in vitro* mutagenesis system version 2.1 booklet. However, for standard strains where it was convenient to prepare and store large volumes of cells either a large scale CaCl<sub>2</sub> method or the RbCl method

(Amersham oligonucleotide-directed *in vitro* mutagenesis system version 2.1 booklet) was used, depending on the bacterial strain being prepared.

#### **CaCl<sub>2</sub> competent cells for use within 24 hours**

40mL of 2xYT medium was inoculated with 400µl of an overnight culture of the appropriate strain, and shaken at 37°C until the A<sub>550</sub> was approximately 0.3. Cells were pelleted by centrifugation at 3000xg for 5 mins, resuspended in 20ml ice cold 50mM CaCl<sub>2</sub> and left on ice for 20 mins. Cells were spun down at 3000xg for 2 mins, resuspended in 4ml cold CaCl<sub>2</sub> and incubated on ice for several hours before use.

#### **CaCl<sub>2</sub> competent cells for long term use**

The appropriate strain was inoculated into 200ml 2xYT media plus any required supplements, and incubated with shaking at 37°C until an A<sub>600</sub> of 1 was obtained. The culture was chilled on ice for 10 min and cells were harvested by centrifugation at 8000xg for 10 min. The supernatant was removed and cells were gently resuspended in 20ml of cold 0.1M CaCl<sub>2</sub> and incubated in an ice/water bath for 20 min. These were then pelleted by centrifugation at 8000xg for 5 min, and gently resuspended in 2ml of cold 0.1M CaCl<sub>2</sub>:glycerol (85:15). Aliquots of 50µl were made into pre-chilled eppendorf tubes, snap frozen in liquid nitrogen and stored at -70°C.

#### **RbCl competent cells for long term use**

The appropriate strain was inoculated into 30mL SOB media and grown overnight. 8ml of this culture was then transferred into 200ml SOB and grown to an A<sub>550</sub> of approximately 0.3. The culture was divided between four tubes, chilled in an ice/salt/water bath and incubated on ice for 15 min. The cells were pelleted by centrifugation at 3000xg for 5 min at 4°C and each pellet was resuspended in 16ml of transformation buffer 1 (100ml contains 1.2g RbCl, 0.99g MnCl<sub>2</sub>.4H<sub>2</sub>O, 3ml of 1M potassium acetate, pH7.5, 0.15g CaCl<sub>2</sub>.2H<sub>2</sub>O and 15g glycerol adjusted to pH5.8 with acetic acid and filter sterilised). This suspension was incubated on ice for 15 min and pelleted by centrifugation at 3000xg for 5 min at 4°C. The cells were then resuspended in a total of 16ml of transformation buffer 2. (100ml contains 2ml of 0.5M MOPS pH6.8, 0.12g RbCl, 1.1g CaCl<sub>2</sub>.2H<sub>2</sub>O and 15g glycerol. Adjust the final pH to pH6.8

with NaOH and then filter sterilise). 300µl aliquots were snap frozen in liquid nitrogen and stored at -70°C.

### **2.5.8 Transformation of competent *E. coli* with plasmid DNA**

#### **CaCl<sub>2</sub> competent cells (for use within 24 hours)**

300µl of fresh competent cells were aliquotted into chilled eppendorf tubes. Plasmid DNA (0.01-0.1µg) or ligation mix (5-20µl) was added to these cells and incubated on ice for 40 min. Cells were heat shocked at 42°C for 45 secs and placed on ice for 5 min. 1mL Luria Bertani media was added and the suspension was allowed to stand at 37°C for 1 hour. Cells were pelleted by centrifugation at 5000rpm for 2 min, gently resuspended in 250µl of Luria Bertani media or minimal media and spread onto the appropriate solid media.

#### **CaCl<sub>2</sub> competent cells (long term use)**

A 50µl aliquot of frozen competent cells was thawed on ice for 5 mins, then 0.9ml of cold 0.1M CaCl<sub>2</sub> was added to each tube. 5-20µl of ligation mix, or the appropriate amount of plasmid DNA (0.1-0.1µg) was gently mixed into a 100µl aliquot of cells, and incubated on ice for 30 min. Cells were heat shocked at 43.5°C for 45 secs and immediately placed on ice for 5 min. 1mL Luria Bertani media was added and the mixture was allowed to stand at 37°C for 1 hour. Cells were pelleted by centrifugation at 5000rpm for 2 min, gently resuspended in 250µl of Luria Bertani media or minimal media and spread onto the appropriate solid media.

#### **RbCl method**

300µl aliquots of competent cells were thawed on ice, and the appropriate amount of plasmid DNA or mutagenised phage (20µl) was added as described above. These were left on ice of 15 mins, heat shocked at 42°C for 2 min, then placed on ice for another 15 mins. Incubation in 5mls Luria Bertani media was carried out for an hour at 37°C to allow the expression of plasmid markers. The cells were pelleted by centrifugation, resuspended in a small volume of Luria Bertani media or minimal media and then plated out on the appropriate solid media. Transfection with M13 derivatives was



carried out by resuspending the cells in minimal media and plating them onto solid H plates with H top agar overlays containing 200 $\mu$ l of mid-exponential phase TG1 cells.

## **2.6 Genomic recombinant DNA techniques**

### **2.6.1 Genetic recombination on the *E. coli* genome**

The genetic recombination method of Oden *et al* (189) was used to replace genes on the *E. coli* genome via double crossover events utilising plasmid DNA. The *recBC sbcBC* strain JC7623 is deficient in the *recBC* nuclease, which degrades linear DNA, and the *sbcBC* genes which make the cell “recombination proficient” by derepressing the *recF* gene (98). A plasmid containing the desired form of a particular gene was transformed into a *recBC sbcBC* strain (JC7623 – see Table 2.1) and this was plated onto solid selection media to screen for recombinants, which normally take 2 days incubation at 37°C to appear. The ORF of the selected gene was amplified by PCR (Section 2.4.9) from a number of these isolates, to identify those which contained only the modified gene. Restriction endonuclease digestion was also used to verify the PCR products (Section 2.5.1). Phage transduction (Section 2.6.2) was then used to transfer the modified gene from JC7623 into the desired background strain.

### **2.6.2 Bacteriophage $P1_{kc}$ transduction**

#### *Determination of phage titer*

Recipient cells were cultured in 10ml glucose-Z broth (Luria Bertani media, 2.5mM CaCl<sub>2</sub>) to Klett 100 (about 2.5x10<sup>8</sup> cells/ml), and the bacteriophage  $P1_{kc}$  lysate was diluted in glucose-Z broth to 10<sup>-3</sup>, 10<sup>-5</sup>, 10<sup>-7</sup> and 10<sup>-9</sup>. 950 $\mu$ l cells were aliquotted into pre-warmed tubes at 37°C, 50 $\mu$ l of diluted P1 lysate was added and the mixture was incubated at 37°C for 30 mins to allow adsorption of the phage to the recipient cells. 200 $\mu$ l of each incubation was mixed with 2ml soft agar at 45°C, and poured onto pre-warmed glucose-Z plates (5g Bactotryptone, 5g NaCl, 2.5g Yeast Extract, 5g Bacto-agar made up to 500ml pH7 and autoclaved, then after cooling add 5ml 0.25M CaCl<sub>2</sub>, 10ml 2M glucose). Plaques were counted after 6 hrs incubation at 37°C. The phage titer, measured in plaque forming units per ml lysate (PFU/ml) was equal to the plaque count divided by the phage dilution (i.e. 10<sup>-7</sup>), multiplied by 100.

### *Preparation of transducing phage lysate*

Donor cells were grown to Klett 100 in 20mls of glucose-Z broth. The cells were pelleted (2000rpm for 10 mins in a benchtop centrifuge) and resuspended in 1ml glucose-Z broth. P1 phage was spun down and the aqueous phase degassed for chloroform by incubation at 37°C with shaking for 1 hr. Between 300µl and 1ml of phage lysate was added to 500µl of cells, using glucose-Z broth in the control tube, and these were incubated at 37°C for 30 mins to allow adsorption of phage onto the cells. 400µl aliquots of the cells were mixed into 6ml pre-warmed soft agars at 45°C and then overlaid on pre-warmed glucose-Z plates. These were incubated at 37°C for 6 hrs and stored overnight at 4°C before preparing the phage lysate. To isolate the phage 4ml glucose-Z broth was added to each plate and the soft agar layer was scraped into sterile SS34 tubes. Chloroform was added at 100µl per 4mls lysate and vortexed thoroughly for several minutes, then left at room temperature for 2 hrs. Centrifugation was carried out at room temperature at 6000xg (7500rpm in SS34 rotor, Sorvall centrifuge) for 20 mins. The supernatant was carefully decanted and the remaining cell debris was removed by filtration through a Millipore 0.2µm low protein binding filter. A few drops of chloroform were added to the phage lysate before storage at 4°C. The phage titer was then determined for this lysate, and 100µl was plated out on the appropriate media to verify the absence of viable donor cells.

### *Generalised transduction*

The recipient strain was grown in 40ml of glucose-Z broth plus the appropriate antibiotics and auxotrophic requirements to Klett 100, and cells were pelleted by centrifugation at 500xg (2000rpm in SS34 rotor, Sorvall centrifuge) for 10 mins. The recipient cells were resuspended in 2ml glucose-Z broth. Phage lysate was spun down and residual chloroform was evaporated from the aqueous phase by incubation at 37°C with shaking for 1 hr. The required amount of phage lysate to give approximately 1-12 PFU/bacterium (up to a maximum of 1.5ml) was added to 500µl of cells, with glucose-Z broth used for the control, and these tubes were incubated at 37°C for 30 mins. 20µl of 1M citrate was then added to each tube to chelate Ca<sup>2+</sup> and stop further infection of P1, followed by 2 volumes of Luria Bertani media. The cultures were incubated with shaking at 37°C for 30 mins. The cells were pelleted by centrifugation as described

above, resuspended in 100mM Tris-HCl pH7.4 and aliquots of 100 $\mu$ l were plated out onto the appropriate selective media and incubated at 37°C for two days.

## **2.7 Protein biochemistry techniques**

### **2.7.1 Gel electrophoresis of proteins**

SDS-polyacrylamide gel electrophoresis was carried out according to the method of (139) using either pre-cast 10% discontinuous or 4-12% gradient Novex polyacrylamide gels with a Novex Xcell-II apparatus, or manually poured discontinuous polyacrylamide gels with a Bio-Rad Mini-PROTEAN<sup>®</sup> II Electrophoresis Cell. Denaturing conditions (containing SDS) were used for all gels. Novex gels buffered with Tris-Glycine were run in 24mM Tris base (pH8.3), 192mM glycine and 0.1% SDS. The NuPAGE<sup>™</sup> MES system uses NuPAGE gels buffered with MES SDS running buffer, pH7.2 (50mM 3-(N-morpholino) propane sulfonic acid (MES), 50mM Tris base, 3.5mM SDS, 1mM EDTA). The 2x loading buffer was composed of 40mM Tris-HCl (pH6.8), 10% (w/v) glycerol, 1.6%SDS and 4%  $\beta$ -mercaptoethanol. Electrophoresis was carried out at a constant voltage of 125V for approximately 100 minutes. Manually poured gels consisted of a 4% stacking gel layer and a 9%-10% resolving gel layer, using a 29:1 acrylamide:bisacrylamide ratio. The stacking gel buffer contained 37.5mM Tris-HCl (pH8.8) and 0.1% SDS, while the resolving gel buffer consisted of 12.5mM Tris-HCl (pH6.8) and 0.1% SDS. APS and TEMED were added to initiate polymerisation of the acrylamide. Electrophoresis was carried out at a constant current of 35mA for approximately 40 mins. Protein samples were mixed with 0.1M fresh DTT and SDS/Z buffer (200mM Tris-HCl (pH6.8), 50% glycerol, 8% SDS and 0.01% bromophenol blue) in a ratio of 1 part buffer to 4 parts sample prior to loading. Low weight makers and prestained molecular weight markers were from Biorad and Novex (SeeBlue<sup>™</sup>).

### **2.7.2 Coomassie staining of polyacrylamide gels**

Protein gels were stained with Coomassie dye using 1.25% w/v Coomassie R-250 (PhastGel<sup>™</sup> BlueR, Pharmacia Biotech), 10% v/v acetic acid 50% v/v methanol and 25% v/v ethanol, and destained overnight in 25% v/v ethanol, 8% v/v acetic acid. Gels



to be dried were desalted in water and equilibrated in gel drying solution (5% w/v glycerol, 30% v/v methanol) then dried between sheets of cellophane pre-soaked in the drying solution using a Novex gel drying frame.

### **2.7.3 Measurement of protein concentration**

Measurements of protein concentration were carried out using the Bio-Rad protein assay system, based on the dye-binding assay method of Bradford (23) using either the full-scale or microtitre process. All assays were performed in duplicate, using bovine serum albumin (BSA) to prepare a standard curve of 0, 20, 40, 60, 80 and 100µg protein. The full scale method mixed 100µl of suitably diluted sample with 5mls diluted dye reagent (1:4 reagent:water). After 2-5 mins the absorbance at 595nm was measured and the protein concentrations were calculated by reference to the BSA standard curve. The microtitre assay involved aliquotting 50µl of diluted protein sample/BSA sample into square bottomed microtitre plates, adding 200µl diluted dye reagent and measuring the absorbance between 405nm and 630nm after 5-30 minutes in a microtitre plate counter spectrophotometer, using the zero protein sample as the blank.

### **2.7.4 Detection of protein by Western blot**

The membrane fraction of various *E. coli* strains was solubilised at 200µg/ml protein (PitA western blots) or 1mg/ml protein (PitB western blots) by 3 cycles of vortexing for 1 min then boiling for 3 mins. Electrophoresis was carried out using a 10% SDS polyacrylamide gel, as described in Section 2.7.1. The proteins were then transferred onto polyvinylidene difluoride (PVDF) membrane by semi-dry electroblotting using a Multiphor<sup>®</sup> II electrophoresis system (Pharmacia) as follows. Sheets of 3MM paper were cut to the size of the gel and presoaked in semi-dry transfer buffer (39mM glycine, 48mM Tris/HCl pH7.5, 0.0375% w/v SDS, 20% v/v methanol). 6 sheets of filter paper were layered in the apparatus, followed by the PVDF membrane, the gel and 6 more sheets of filter paper. Electrophoretic transfer proceeded at 0.8mA/cm<sup>2</sup> for 50-120 mins. Non-specific sites on the PVDF blot were protected by incubation in blocking buffer (10% milk powder, 10mM Tris/HCl pH7.5, 0.9% NaCl) and the PVDF membrane was then washed in Tris-buffered saline (10mM Tris/HCl pH7.5, 0.9% NaCl). The primary antibody was applied by incubation with either polyclonal

antipeptide PitA antibody diluted 1/500 in blocking buffer or polyclonal antipeptide PitB antibody at 1/100 dilution at either 4°C overnight or 2 hrs at room temperature. The blot was washed in Tris-buffered saline prior to incubation with goat anti-rabbit immunoglobulin conjugated to alkaline phosphatase (DAKO), diluted at 1/1500 (PitA) or 1/1000 (PitB) in blocking buffer. The blot was washed with Tris-buffered saline and then immunostained with Western Blue<sup>®</sup> stabilized alkaline phosphatase substrate (Promega). The blot was then washed in water and air dried.

### ***2.7.5 Preparation of cytoplasmic membranes from E. coli by centrifugation***

Cytoplasmic membranes were isolated by centrifugation of cells broken open by a Sorvall Ribi Cell Fractionator. Strains were inoculated into 1 litre of rich media (Luria Bertani media, 34mM glucose, appropriate antibiotics and/or growth requirements) and grown to a cell density of around 250 Klett units (approximately  $6 \times 10^8$  cells/ml). Cells were pelleted by spinning at 5900xg for 15 mins at 4°C and were then washed with potassium phosphate buffer containing the protease inhibitor phenylmethanesulfonylfluoride (PMSF) (0.05M KHPO<sub>4</sub>, 1mM PMSF, pH8). This wash and spin was repeated. On occasion the protease inhibitors aprotinin (2mg/ml), leupeptin (0.5µg/ml) and pepstatin (0.7µg/ml) were also added. The washed cell pellet was then resuspended in 30ml phosphate buffer and cells were broken open by passage through a Sorvall Ribi Cell Fractionator. Unlysed cells and cell debris was separated by centrifugation at 39000xg for 30 mins at 4°C. The supernatant, containing both membrane vesicles and the soluble cytoplasmic fraction, was collected.

Membrane vesicles were separated from the cytoplasm by ultracentrifugation at 4°C for two hours at 360000xg. The cytoplasm was removed and stored at -20°C. The membrane pellet was resuspended in less than 2.5mls of phosphate buffer and stored in 0.5ml aliquots. The protein concentration of the cytoplasmic membranes was assayed as previously described (Section 2.7.3).

### **2.7.6 Preparation of cytoplasmic membranes from *E. coli* by ammonium sulfate precipitation**

Cytoplasmic membranes were isolated through selective ammonium sulfate precipitation as described by Cox *et al.* (45). Strains were inoculated into 2 litres of rich media and the cells were grown, pelleted and fractionated as described above (Section 2.7.5) except that the phosphate buffer used was 0.1M potassium phosphate, pH7 and cells were resuspended in 25mls of buffer before fractionation. The supernatant of the centrifugation was saved and either used immediately or stored at -20°C.

The cytoplasmic membranes were selectively precipitated by adding solid  $(\text{NH}_4)_2\text{SO}_4$  slowly to five 30% saturation at 4°C. The mixture was stirred for a further 30 mins, followed by centrifugation at 14500rpm for 15 mins in an SS34 rotor of a Sorvall centrifuge (25000xg). The pellet was carefully isolated and resuspended in 1ml of phosphate buffer per original 1g wet weight of cells. The protein concentration was determined, and the membrane samples were aliquotted and stored at -20°C.

### **2.7.7 Solubilisation of cytoplasmic membranes**

Cytoplasmic membrane fractions were diluted to approximately 5mg/ml protein in phosphate buffer (50mM  $\text{KHPO}_4$ , 300mM NaCl, pH8.0). Solubilisation was carried out by adding detergent (such as 0.5% Triton X-100, or 1.35% octylglucoside) and incubating 1 to 3 hrs at 4°C with gentle agitation, until there was no further clearing of the solution. This solution was diluted approximately two-fold and the solubilised cytoplasmic fraction was separated by a 360000xg centrifugation for one hour at 4°C. The supernatant was carefully isolated, the pellet was resuspended in approximately 100 $\mu$ l of phosphate buffer, and both the supernatant and pellet were stored at -20°C.

## **2.8 Over-expression and purification of protein using the 6xHistidine affinity tag**

An affinity tag of six consecutive histidine residues was attached to the C-terminus of the putative *pitA* ORF using the QIAGEN vector pQE60. The resulting plasmid was transformed into the host strain MC15, which contains pREP4, a plasmid which



constitutively expresses the *lac* repressor at high levels. This gives tighter control of protein expression, which is often needed for hydrophobic proteins (QIAexpress® product guide).

Protein expression was carried out by growing cultures in rich media (Luria Bertani media, 34mM glucose, 25µg/ml kanamycin and 100µg/ml ampicillin) to Klett 100-200, and then inducing protein production from the plasmid by the addition of 0.1mM Isopropyl-β-D-thiogalactopyranoside (IPTG), (one tenth of the IPTG concentration recommended by QIAGEN.) Cells were grown for a further 1-2 hours before harvesting by centrifugation at 5900xg for 15 mins at 4°C. The preparation of cytoplasmic membranes, and isolation of the various cell fractions was carried out as described in Section 2.7.5. The protein concentration of cytoplasmic membrane fractions were determined (Section 2.7.3), and membrane fractions were then solubilised (Section 2.7.7).

Purification of 6xHis tagged protein from solubilised membrane fractions was carried out by chromatography using Ni-NTA metal-chelate resin, as described in the QIAGEN Expressionist manual. (Triton X-100 up to 2% is compatible with the Ni-NTA-6xHis interaction.) A 0.45ml Ni-NTA column was prepared (0.15ml bed volume per ml of solubilised protein solution - (151)) and the column was pre-incubated with pH8 phosphate buffer (0.05M KHPO<sub>4</sub>, 300mM NaCl) containing 0.25% Triton X-100. (All subsequent buffers contain 0.25% Triton X-100.) After loading the membrane fraction, the column was washed with more than 20 volumes of pH8 phosphate buffer and 20 volumes of pH6 phosphate buffer. Elution was carried out with 5 volumes each of 0.1M imidazole, pH6, and 0.5M imidazole, pH6 in phosphate buffer. Alternatively, 10% glycerol was added to the pH6 wash and all subsequent buffers and/or the Triton X-100 concentration was increased to 0.5%. The elution was also carried out in more graduated steps using 5mls (approximately 10 column volumes) each of 0.1M, 0.25M and 0.5M imidazole, or 0.1M, 0.2M, 0.3M and 0.5M imidazole in pH6 phosphate buffer. Eluted samples were concentrated in Centricon 10 concentrators at 4°C until volumes were reduced about 10-fold.

## **2.9 *In vitro* transcription/translation of plasmid DNA**

The *E. coli* S30 Extract System for Circular DNA (Promega) was used to confirm the expression of protein from genes cloned into plasmid vectors. Only plasmids purified using caesium chloride equilibrium density gradient centrifugation (Section 2.4.2) were used as templates for this procedure. The S30 extract was prepared from *E. coli* by modifications of the method described by Zubay (312) and contain all the components needed for protein synthesis, except the DNA template. The amino acid premix used also lacks methionine, so [<sup>35</sup>S]-methionine can be used to label the protein. The Promega handbook describes the method in detail, which is summarised as follows:

The synthesis of labeled protein was carried out by combining 2µg of DNA template with the S30 extract, amino acid premix (minus methionine) and [<sup>35</sup>S]-methionine, and incubating this mixture at 37°C for 60 minutes. The whole volume (50µl) was then prepared for SDS Gel analysis of the translation products by acetone precipitation. The dried pellet was resuspended in 100µl SDS loading buffer (40mM Tris-HCl (pH6.8), 10% (w/v) glycerol, 1.6%SDS and 4% β-mercaptoethanol), heated at 100°C for 2 mins and 10µl was loaded onto a 9% or 10% SDS acrylamide separating gel with a 5% stacking gel (Section 2.7.1). Electrophoresis was carried out at 35mA constant current for around 40 mins. Visualisation of protein bands was sometimes carried out by staining with Coomassie blue dye (Section 2.7.2) followed by drying the gel for autoradiography. Otherwise the gel was dried directly after electrophoresis (2 hrs at 60°C under vacuum in a gel drier) and autoradiography was carried out by exposure to Hyperfilm MP autoradiography film (Amersham) for between 8-20hrs and then developed in an X-omat M20 automatic processor (Eastman Kodak).

## **2.10 Identification of *E. coli* cells able to transport inorganic phosphate (Pi)**

Pi uptake from genes located on the chromosome was carried out using the particular *E. coli* strain. Plasmid-borne genes were tested for Pi uptake by transforming the plasmid of interest into a Pi auxotrophic background strain - initially AN3066 (*pitA1 ΔpstC345*) and then AN3902 (*pitA1 pitB::Cat<sup>r</sup> ΔpstC345*). Transformed cells were always

cultured in the presence of 1mM  $\alpha$ -glycerol-3-phosphate (G3P) to prevent selection pressure for Pi transport, until the Pi uptake phenotype of the strain was known. Pi uptake was selected for by plating individual colonies onto solid minimal media containing 500 $\mu$ M Pi plus any required antibiotics or supplements (Pi media - described in Section 2.3). All isolates were also plated onto Pi media containing 1mM G3P, to confirm that these strains could grow on minimal media when there was no requirement for Pi uptake. Control strains with/without Pi uptake were present on all media plates, to allow accurate comparisons of cell growth.

### ***2.11 Inorganic phosphate (Pi) transport assay***

Cells were grown overnight in a 50% dilution of minimal media (described in Section 2.3) supplemented with 20mM glucose, 1.5 $\mu$ M thiamine, 5% Luria Bertani media plus the addition of 1mM  $\alpha$ -glycerol-3-phosphate and/or 50 $\mu$ g/ml ampicillin where appropriate. Phosphate free buffered “uptake media” was then used in all cell washes, dilutions and assays. This consisted of 50mM triethanolamine, 15mM KCl, 10mM  $(\text{NH}_4)_2\text{SO}_4$ , and either 1.8mM  $\text{MgSO}_4$  or 10mM  $\text{MgSO}_4$ . Citric acid was used to alter the pH to either pH6.6 or pH7.0. Uptake media, at 37°C, was used to wash cells three times and to dilute them to an absorbance (660nm) of 0.35. These washed suspensions were then incubated at 37°C with shaking for 2 hours to deplete any internal stores of Pi, and were again diluted to an absorbance (660nm) of 0.35. Cell suspensions were sometimes stored at 4°C for less than 30 mins, but were usually assayed immediately. Cells were diluted into the uptake media (usually to a cell density (660nm) of approximately 0.035) and preincubated at 37°C with shaking for 5 mins.  $^{33}\text{Pi}$  (NEN) at approximately 0.035 mCi/ml in a 2.5mM  $\text{K}_2\text{HPO}_4$  (PitB) solution which may then be diluted to 0.5mM  $\text{K}_2\text{HPO}_4$  (PitA) was filtered through a 0.22 $\mu$ M membrane, stored at 4°C and refiltered before each experiment. Uptake assays were started by the addition of the appropriate  $^{33}\text{Pi}$  solution to the preincubated cells. Samples taken at 15 and 25 secs were immediately filtered onto 0.45 $\mu$ M membrane filters and washed twice with a wash solution consisting of 50mM triethanolamine-HCl, 10mM KCl, 1mM  $\text{MgSO}_4$  and 10mM  $(\text{NH}_4)_2\text{SO}_4$  at either pH6.6 or pH7.0. Damp filters were placed in 5ml of Emulsifier Safe (Packard) scintillation cocktail for aqueous samples and counted for 2 min. Uptake by cells was carried out under conditions in which Pi uptake was linear



over time. Pi concentrations for  $K_m^{\text{app}}$  and  $V_{\text{max}}^{\text{app}}$  determinations were 1-15 $\mu\text{M}$  Pi (PitA), 10-100 $\mu\text{M}$  Pi (PitB long) or 2-50 $\mu\text{M}$  Pi (PitB short). Data from these kinetic experiments was analysed by non-linear regression using the Michaelis-Menten equation and the Graphpad Prism program. Most individual experiments had at least two measurements of around 5 Pi concentrations.

## **2.12 *phoA* alkaline phosphatase assay**

Retention of *pst* mutations by recipient strains was tested by screening colonies for constitutive synthesis of *phoA*-encoded alkaline phosphatase on agar plates as described by Bracha *et al.* (22). This rapid spray assay relies on the fact that expression of *phoA* is repressed by Pi but negatively controlled by genes comprising the *pst* operon (162). Therefore strains carrying a *pst* mutation will synthesise alkaline phosphatase constitutively on high phosphate media.

Single cells grown overnight on rich media (Luria Bertani media, 34mM glucose, and 1mM G3P where appropriate) were transferred to filter paper and sprayed with a mixture of fast blue B salt (50mg fast blue B salt in 5ml H<sub>2</sub>O) and a naphthol salt of phosphate (10mg  $\alpha$ -naphthol phosphate in 5ml 1M Tris-HCl pH8) which were freshly made and mixed immediately before use. A colour change occurs immediately, with purple indicating the presence of *phoA* alkaline phosphatase activity, and yellow indicating the absence of *phoA* activity.

## **2.13 Production of antipeptide polyclonal antibodies**

### **2.13.1 Peptides used for raising antibodies**

The PitA peptide ARIHLTPAEREKKDC (from A188 to D201) and the equivalent sequence in PitB, DRIHRIPEDRKKKKC (from D188 to K201) are from an extramembraneous loop in the putative folded structure (Figure 3.15). This region is highly charged (theoretically more antigenic) and exhibits significant sequence variability between PitA and PitB. C-terminal cysteines were added to each sequence to allow chemical conjugation to the core matrix or carrier protein (see below).

### **2.13.2 Immunisation of rabbits with Multi-Antigenic Peptides**

#### *PitA.*

Rabbits were immunised with the PitA peptide coupled to a poly-lysine core matrix via a C-terminal cysteine. The multiple antigen peptide system (MAP) uses a high molar ratio of peptide antigen attached to the immunogenically inert core, removing the need for a carrier protein to elicit an antibody response (154, 256). The PitA peptide-MAP conjugate was prepared by the Biomolecular Resources Facility, John Curtin School of Medical Research. It was dissolved in phosphate buffered saline (PBS), emulsified with an equal volume of Freund's adjuvant (68) and 200µg was injected subcutaneously. Standard sampling and injecting protocols were followed (88), including the preparation of pre-immune sera by bleeding the rabbit prior to the first injection of MAP conjugate. Sera were allowed to clot at 37°C for an hour. The clot was shrunk by separating the clot from the sides of the tubes followed by storage at 4°C overnight. Sera were then isolated by centrifugation at 10000xg for 15 min, and the supernatant was stored at -20°C.

#### *PitB*

The PitB peptide was attached to Imject® maleimide activated keyhole limpet hemocyanin (peptide-KLH, Pierce) following the manufacturer's instructions, after attempts using a PitB peptide-MAP conjugate which had been prepared by the Biomolecular Resources Facility produced no PitB antigenic response in several rabbits. 200µg of conjugate was partially dissolved in dimethylsulfoxide with sonication, then diluted with one volume of phosphate buffered saline (PBS). Injections and screenings were carried out as previously described for the PitA peptide.

### **2.13.3 Detection of antipeptide polyclonal antibodies in rabbit serum**

The sera from rabbits injected with either PitA or PitB peptide-MAP conjugates were initially screened for positive responses to the relevant peptide by enzyme-linked immunosorbent assay (ELISA) (64). The wells of 96 well micro-titre plates were coated with the appropriate peptide by incubating 50µl of 50mg/ml peptide in PBS overnight at 4°C or for 2 hrs at room temperature, with control wells being exposed to

PBS only. All wells were washed twice with PBS, then 50µl serial dilutions of the pre-immune and post-immune sera were added to the control and peptide coated wells. Dilutions were prepared over a range of 1:10 to 1:1000 in PBS. These were incubated at room temperature for 1 hr, then washed twice with PBS, then once with a freshly prepared pH9.5 buffer (10mM diethanolamine, 0.5mM MgCl<sub>2</sub>). 50µl of alkaline phosphatase substrate (1mg/ml p-nitrophenylphosphate in pH9.5 buffer) was added to each well. After incubation at room temperature the optical density between 405nm and 630nm was recorded in a micro-titre plate reader.

Western blotting against the membrane fractions of AN3531 (PitA) AN3135 (PitB) and AN3514 (control) was then used to assess the antigenicity of the sera to either the PitA or PitB membrane protein. The sera from the PitB peptide-KLH conjugate was only assessed by Western blot.

#### ***2.13.4 Purification of polyclonal antipeptide antibodies from serum by immunoaffinity chromatography***

Sera containing PitA antipeptide antibody were not purified. The PitB equivalent was isolated from sera by immunoaffinity purification (SulfoLink® Kit, Pierce) following a two-stage ammonium sulfate precipitation (88) and dialysis against PBS. The synthetic peptide column was prepared following the manufacturer's instructions, and the dialyzed PitB antipeptide antibody solution was passed through the column 3 times.

#### ***2.14 Analysis of DNA sequences using MacTargsearch 2.0***

MacTargsearch is used to identify DNA sequences that are putative matches to a 'target' sequence. It provides several target sequences including the *E. coli* promoter consensus sequence. Target sequences can also be customised. This is a compiled Microsoft BASIC program created by James A. Goodrich, William R. McClure and Michael L. Schwartz in 1989 at Carnegie Mellon University, and modified by Dr Peter Markiewicz, UCLA.



### ***E. coli* promoter consensus sequence**

This target sequence was used to search for the putative promoter sequence/s in PitA and PitB. This sequence is listed at the top of Table 5.3 and consists of two regions of base pair scores separated by a spacer region of variable length. These regions contain the -35 and -10 hexamers respectively, but are larger than the hexamers, so the printed spacer length is 6bp shorter than the true spacer length between the -10 and -35 regions. Similarity scores for each base and spacer length are calculated as described in (179) and (76).

### ***Pho* box consensus sequence**

MacTargsearch was used to identify putative *pho* boxes in and around the putative promoter regions of PitA and PitB. This sequence contains 18 nucleotides and is listed at the top of Table 5.4. The weighting for every nucleotide in each position was calculated from the information content provided for 10 *pho* box consensus sequences ((280) and references therein).

---

# ***Chapter three***

***Escherichia coli has two pit genes  
with distinct characteristics***

---

## Chapter three

# *Escherichia coli* has two *pit* genes with distinct characteristics

### 3.1 Introduction

The Pit transport system was first reported by Willsky *et al* (289) when mutations in the *Pst* system of several *E. coli* K-12 strains revealed the presence of a second Pi transporter, named Pit. Introduction of *pst* mutations into strain K-10 produced no measurable Pi transport, indicating that K-10 has a mutation that effects Pit Pi transport. This lesion mapped at minute 77.1-77.3 on the *E. coli* genome (249, 59). All known *pit* mutations mapped in this locus of the *E. coli* chromosome and, as described below, Elvin *et al* (59) showed that these mutations were complemented by transformation with a plasmid carrying a small fragment of DNA which directs the synthesis of the PitA protein.

*pitA* was originally isolated from *E. coli* K-12 by transforming cosmid DNA from HR152 (*pstA*) into the Pi auxotrophic strain GS5 (*pit pstA*) (59). The cosmid containing wild type *pitA* was selected by its ability to allow strain GS5 to grow on minimal media supplemented with 500  $\mu$ M Pi. In an attempt to re-isolate the *pitA* gene from *E. coli* K-12 Dianne Webb transformed cosmid DNA from AN2537 ( $\Delta$ *pstC345*) (253) into AN3020 (*pit*  $\Delta$ *pstC345*) (284), cloning a gene with a similar but distinct DNA sequence to *pitA* (89). Thus, the *pitB* gene was isolated and shown to transport Pi.

The presence of a second Pit transporter had not been suspected until the discovery of the *pitB* gene, and led to two possible alternatives. *pitB* may be an allele of *pitA*, or may represent a second previously undetected *pit* gene on the *E. coli* genome.

This chapter examines whether *pitA* and *pitB* are alleles or individual genes. It defines and compares individual characteristics of *pitA* and *pitB*, including nucleotide and amino acid sequences, and the kinetic parameters  $K_m^{app}$  and  $V_{max}^{app}$ . Over-expression and isolation of the PitA and PitB proteins is attempted.



### ***3.2 Comparison of the nucleotide sequences of pitA and pitB***

The DNA fragments containing *pitA* and *pitB* Pi uptake activities were previously reduced to 2.27kb and 3.06kb fragments, respectively, and sequenced (59, 89). These two nucleotide sequences were then compared using the DNA Pustell Matrix algorithm (Figure 3.1), which revealed a high degree of similarity in sequence coinciding with an open reading frame (ORF) of 1500 nucleotides found in each DNA fragment. These ORFs are located at 573-2073bp in the *pitA* DNA fragment and 1404 – 2903bp in the *pitB* DNA fragment. Some patches of sequence similarity also occur upstream of these ORFs in the putative promoter regions of each gene. While other ORFs are present these are the two longest ORFs, and the 75% sequence identity between them makes these the most likely candidates for *pitA* and *pitB* (Figure 3.2).

### ***3.3 Are pitA and pitB genes or alleles?***

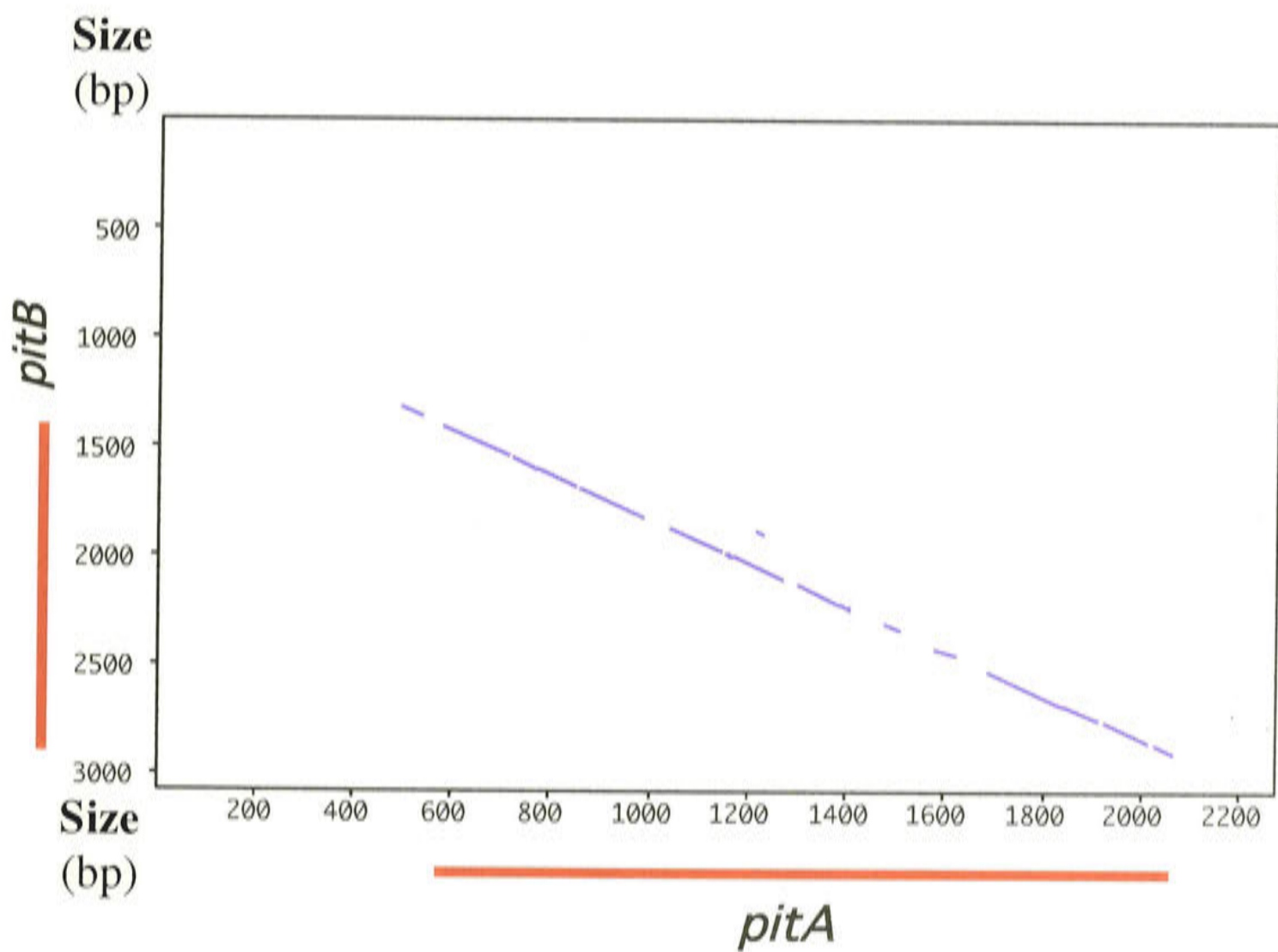
The method used to isolate *pitA* and *pitB* could not distinguish whether these sequences were alleles or distinct genes, although it does indicate that they are unlikely to form an operon, as each gene was expressed individually. There is no previous evidence for the presence of two distinct *pit* loci. However, the 75% sequence identity between the putative *pitA* and *pitB* ORFs, is much less than that found in most alleles and the lack of sequence identity in the surrounding DNA indicates that *pitA* and *pitB* are likely to be found at separate loci, therefore are distinct genes. To clarify this, PCR experiments were carried out to detect the presence of *pitA* and/or *pitB* on the genomes of various *E. coli* strains. AN236 and AN259 were selected because they were used in the preparation of strain HR152, which supplied the wild type *pitA* gene subsequently sequenced and analysed by Elvin *et al* (59). AN248 is a K-12 strain with wild type *pit* and *pst* genes. K-10 is the strain where the *pit* mutation was first detected, while HR187 is a *recA* derivative of GS5 (*pit pstA*), the Pi auxotrophic recipient strain used for isolating wild type *pitA*. AN3066 is a *recA* derivative of AN3020 (*pit ΔpstC345*), the background strain for expression of the *pitB* gene. Thus there are three strains that may have wild type *pit* transport (AN236, AN259, AN248) and three strains with known defects in *Pit* transport (HR187, AN3066, K-10). Genomic DNA was purified

### Figure 3.1

#### Pustell DNA Matrix alignment of the sequences containing *pitA* and *pitB*.

The sequenced DNA fragments containing the *pitA* and *pitB* genes were compared by the Pustell DNA Matrix algorithm (MacVector 6.0), using a window size of 30, a minimum score of 65%, a hash value of 6 and a jump value of 1. Both strands were analysed.

— Highlights the large open reading frame found in each DNA fragment.



## Figure 3.2

### DNA alignment of the putative open reading frames of *pitA* and *pitB*.

The DNA alignment was carried out using the program ClustalW (MacVector 6.0) with an open gap penalty of 10, extending gap penalty of 5 and weighted transitions. Each putative open reading frame (ORF) contains 1500 bases, with *pitB* containing a single addition relative to *pitA* at nucleotide 978 in the ORF, and a deletion at nucleotide 1021 (which forms part of the large variable region between PitA and PitB in the deduced amino acid sequences).

- A Identical bases between *pitA* and *pitB*.
- A Nucleotides unique to the putative *pitA* ORF.
- A Nucleotides unique to the putative *pitB* ORF.

The DNA alignment of the upstream sequences of *pitA* and *pitB* is illustrated in Figure 5.9, with the putative RNA polymerase promoter sequence highlighted.



```

pitA 1 ATGCTACATTTGTTTGCTGGCCTGGATTTGCATACCGGGCTGTTATTATT 50
pitB 1 ATGCTAAATTTATTTGTTGGCCTTGATATATACACAGGGCTTTTGTATT 50
***** ** * ** * ** * ** * ** * ** * ** * ** * ** * ** * ** *

pitA 51 GCTTGCACTGGCTTTTGTGCTGTTCTACGAAGCCATCAATGGTTTCATG 100
pitB 51 GCTTGCTCTGGCATTGTTGTTGTTCTACGAAGCAATCAATGGTTTCATG 100
***** ** * ** * ** * ** * ** * ** * ** * ** * ** * ** * ** *

pitA 101 ACACAGCCAACGCGGTGGCAACCGTTATCTATAACCGCGCGATGCGTTCT 150
pitB 101 ACACGGCGAATGCGGTGGCAGCCGTTATTTATACTCGTGCCATGCTACCA 150
**** * ** * ** * ** * ** * ** * ** * ** * ** * ** * ** *

pitA 151 CAGCTCGCCGTGGTTATGGCGGCGGTATTCAACTTTTTGGGTGTTTGGCT 200
pitB 151 CAACTTGCTGTGGTGATGGCGGCATTTTAACTTTTTGGCGTGTTATT 200
** * ** * ** * ** * ** * ** * ** * ** * ** * ** * ** *

pitA 201 GGGTGGTCTGAGTGTTGCCTATGCCATTGTGCATATGCTGCCGACGGATC 250
pitB 201 GGGCGGACTTAGCGTTGCCTATGCCATTGTCATATGTTGCCAACCGATT 250
*** * ** * ** * ** * ** * ** * ** * ** * ** * ** * ** *

pitA 251 TGCTGCTTAATATGGGATCGTCTCATGGCCTTGCCATGGTGTCTCTATG 300
pitB 251 TGTTGCTGAATATGGGGTCAACCACGGCCTGGCGATGGTCTTTCCATG 300
** * ** * ** * ** * ** * ** * ** * ** * ** * ** * ** *

pitA 301 TTGCTGGCGGCGATTATCTGGAACCTGGGTACCTGGTACTTTGGTTTACC 350
pitB 301 CTGCTGGCGGCGATTATCTGGAACCTGGGAACGTGGTTCTTCGGTTTACC 350
***** ** * ** * ** * ** * ** * ** * ** * ** * ** * ** * ** *

pitA 351 TGCATCCAGCTCTCATACGCTGATTGGCGCGATCATCGGGATTGGTTTAA 400
pitB 351 GGCTCCAGTTCGCACACCTTGATTGGTGCGATTATCGGCATCGGTTTAA 400
** * ** * ** * ** * ** * ** * ** * ** * ** * ** * ** *

pitA 401 CCAATGCGTTGATGACCGGGACGTCAGTGGTGGATGCACTCAATATCCCG 450
pitB 401 CCAACGCGCTGTTAACCGGCTCATCGGTGATGGATGCGTTAAACCTGCGT 450
**** * ** * ** * ** * ** * ** * ** * ** * ** * ** * ** *

pitA 451 AAAGTATTAAGTATTTTCGGTCTCTGATCGTTTCCCCTATTGTCGGCCT 500
pitB 451 GAAGTGACCAAATTTCTCCTCGCTGATTGTTTCCCCTATCGTCGGCCT 500
**** * ** * ** * ** * ** * ** * ** * ** * ** * ** * ** *

pitA 501 GGTGTTTGCTGGCGGTCTGATTTTCTTGCTGCGTCGCTACTGGAGCGGCA 550
pitB 501 GGTCAATTGCGGGAGGCCTGATAATTCCTGCTGCGACGCTACTGGAGCGGGA 550
*** * ** * ** * ** * ** * ** * ** * ** * ** * ** * ** *

pitA 551 CCAAGAAACGCGCCCGTATCCACCTGACCCAGCGGAGCGTGAAAAGAAA 600
pitB 551 CGAAAAGCGTGACCGTATTCACCGCATTCCGGAAGATCGCAAAAAGAAA 600
* ** * ** * ** * ** * ** * ** * ** * ** * ** * ** * ** *

pitA 601 GACGGCAAGAAAAGCCGCGTTCTGGACGCGTATTGCGCTGATCCTTTC 650
pitB 601 AAAGGCAAACGTAAACCGCCATTCTGGACGCGTATTGCGCTGATTGTTTC 650
* ***** ** * ** * ** * ** * ** * ** * ** * ** * ** * ** * ** *

pitA 651 CGCTATCGGCGTGGCGTTTTTCGCACGGCGCGAACGATGGTCAGAAAGGCA 700
pitB 651 CGCTCGGGCGTGGCGTTTTTCGCACGGCGCGAACGACGGACAAAAGGGA 700
**** * ** * ** * ** * ** * ** * ** * ** * ** * ** * ** *

pitA 701 TTGGTCTGGTTATGTTGGTATTGATTGGCGTCGCGCCAGCAGGCTTCGTG 750
pitB 701 TCGGCCTGGTAATGCTGGTACTGGTGGGGATTGCCCTGCTGGCTTCGTG 750
* ** * ** * ** * ** * ** * ** * ** * ** * ** * ** * ** *

```



```

pitA 751 GTGAACATGAATGCCACTGGCTACGAAATCACCCGTACCCGTGATGCCAT 800
pitB 751 GTCAATATGAATGCGTCCGGCTATGAAATTACCCGTACCCGC GATGCCGT 800
      ** ** ***** * ***** ***** ***** ***** *

pitA 801 CAACAACGTCGAAGCTTACTTTGAGCAGCATCCTGCGCTGCTCAAACAGG 850
pitB 801 TACCAACTTCGAACACTACTGCAACAGCATCCTGAACTGCCGCAGAAGT 850
      * ***** ***** *** * * ***** ***** **** * **

pitA 851 CTACCGGTGCTGATCAGTTAGTACCGGCTCCGGAAGCTGGCGCAACGCAA 900
pitB 851 TGATTGCGATGGAACCTCCATTGCCTGCAGCATCGACTGATGGCACGCAA 900
      * * ** * * * * * * * * * * * * * * * * * * * *

pitA 901 CCTGCGGAGTTCCTACTGCCATCCGTGCAATACCATTAACGCGCTCAACCG 950
pitB 901 GTAACAGAGTTTCACTGTCATCCGGCAATACCTTTGATGCTATTGCGCG 950
      * ***** ***** ***** * ***** ** * * * * **

pitA 951 CCTGAAAGGTATGTTGACCACCGATGT-GGAAAGCTACGACAAGCTGTCG 999
pitB 951 CGTTAAAACGATGCTGCCAGGCAATATAGGAAAGTTACGAGCCGTTAAGC 1000
      * * *** ** * * * * * * * * * * * * * * * * * *

pitA 1000 CTTGATCAACGTAGCCAGATGCGCCGCATTATGCTGTGGTT TCTGACAC 1049
pitB 1001 GTGAGTCAGCGCAGCCAGCTG-GCCGCATTATGCTGTGCATC TCTGATAC 1049
      * ** * * * * * * * * * * * * * * * * * * * * * *

pitA 1050 TATCGACAAAGTGGTGAAGATGCCTGGCGTGAGTGCTGACGATCAGCGCC 1099
pitB 1050 CTCCGCGAAGCTAGCGAAACTGCCAGGCGTCAGTAAAGAAGACCAGAACC 1099
      ** ** * * * * * * * * * * * * * * * * * * * * * *

pitA 1100 TGTTGAAGAACTGAAGTCCGACATGCTTAGCACCATC GAGTATGCACCG 1149
pitB 1100 TGCTGAAAAA ACTTCGCAGCGATATGTTAAGCACCATTGAGTACGCTCCG 1149
      ** ***** ***** *** * * * * * * * * * * * * * * *

pitA 1150 GTGTGGATCATCATGGCGGTC GCGCTGGCGTTAGGTATCGGTACGATGAT 1199
pitB 1150 GTGTGGATCATCATGGCGGTA GCACTGGCGCTCGGCATTGGCACCATGAT 1199
      ***** ***** * * * * * * * * * * * * * * * *

pitA 1200 TGGCTGGCGCCGTGTGGCAACGACTATCGGTGAGAAAATCGGTAAGAAAG 1249
pitB 1200 TGGCTGGCGTCGTGTAGCGATGACCATCGGTGAGAAGATGGTAAGCGCG 1249
      ***** ***** * * * * * * * * * * * * * * * * *

pitA 1250 GCATGACCTACGCTCAGGGGATGTCTGCCAGATGACGGCGGCAGTGTCT 1299
pitB 1250 GCATGACGTATGCGCAAGGCATGGCGGCACAAATGACGGCGGCAGTGTCT 1299
      ***** ** * * * * * * * * * * * * * * * * * * * *

pitA 1300 ATCGGCCTGGCGAGTTATAACGGGGATGCCGGTTTCCACTACTCACGTA 1349
pitB 1300 ATCGGTCTTGCCAGTTATAATTGGGATGCCCGTCTCCACAACACAGTCCT 1349
      ***** ** * * * * * * * * * * * * * * * * * * * *

pitA 1350 CTCCTCTTCTGTCGCGGGGACGATGGTGGTA GATGGTGGCGGCTTACAGC 1399
pitB 1350 CTCGTCTGCAAGTGCAGGGACGATGGTGGTG GACGGCGGTGGGTTACAGC 1399
      *** ** * * * * * * * * * * * * * * * * * * * *

pitA 1400 GTAAAACCGTGACCAGCATTCTGATGGCC TGGGTGTTTACCCTTCCGGCT 1449
pitB 1400 GTAAAACGGTAACCAGCATCCTGATGGCG TGGGTATTTACTTTACCGGCG 1449
      ***** ** * * * * * * * * * * * * * * * * * * * *

pitA 1450 GCGGTA CTGCTTTCCGGCGGGCTGTACTGGCTCTCCTTGCAGTTCCTGTAA 1500
pitB 1450 GCAATTTTCTTTCTGGTGGGCTGTACTGGATAGCATTGCAGTTGATTAA 1500
      ** * * * * * * * * * * * * * * * * * * * * * *

```



as described in Section 2.4.3. PCR amplification of the putative open reading frames for each gene was carried out using *Taq* polymerase (primers 95-87/95-88 for *pitA* and 95-89/95-90 for *pitB*, see Table 2.5 for primer details, Figure 4.1 for amplification details), and the PCR products were separated and visualised on a 1% agarose gel (Figure 3.3, Table 3.1).

While the water blanks contained no DNA bands, all strains tested produced DNA of approximately 1.5kb when oligonucleotides for either *pitA* or *pitB* genes were used. Thus *pitA* and *pitB* DNA approximating the putative ORF size of 1500 nucleotides was present in all strains, including those with known defects in *pit* transport. All these PCR products were purified and restriction endonuclease digestion was used to identify the putative *pitA* and *pitB* gene products, to ensure no cross contamination occurred. Thus it was shown that all the strains tested contain approximately full sized putative *pitA* and *pitB* ORFs (Figure 3.4), indicating that *pitA* and *pitB* are separate genes, and that the *pit* mutation/s in K-10, AN3066 and GS5 do not involve large deletions.

### **3.4 Over-expression and purification of PitA**

Attempts were made to over-express and purify PitA and PitB to allow identification of the translation start of the genes by N-terminal sequencing of the proteins and to provide protein for further experimentation.

An affinity tag of six consecutive histidine residues was added to the C-terminus of the putative *pitA* ORF using QIAGEN vector pQE60, to allow the over-expression and purification of a PitA6xHis protein which retains its native N-terminus. This plasmid, pAN910, (Figure 3.5) was screened for potential *pitA* protein products by *in vitro* transcription/translation using the *E. coli* S30 Extract System for Circular DNA (Figure 3.6). The insertion of putative *pitA* DNA into plasmid pQE60 produced a strong protein product at approximately 46kDa, with some possible degradation products. The expected molecular weight of the putative *pitA::6xHis* protein is 54kDa, but many membrane proteins undergo anomalously fast migration during electrophoresis, so this difference is not unusual. The 46kDa protein band was isolated and N-terminal sequencing was attempted by the Biomolecular Resource Facility at the ANU, but the amount of protein present was too low to produce a clear result (results not shown).



## Figure 3.3

### Screening for the presence of *pitA* and *pitB* DNA in various strains using PCR.

**A.** 1% agarose gel analysis of the PCR products from reactions using primers for the proposed *pitA* open reading frame (95-87, 95-88 – see Table 2.5 for details) and the genomic DNA of the strains listed below in Table 3.1. Each lane represents a separate PCR reaction, including a water blank control. 2µl of reaction mix was loaded per well. The 1.5kb product representing an approximately full-length putative *pitA* open reading frame is indicated with an arrow.

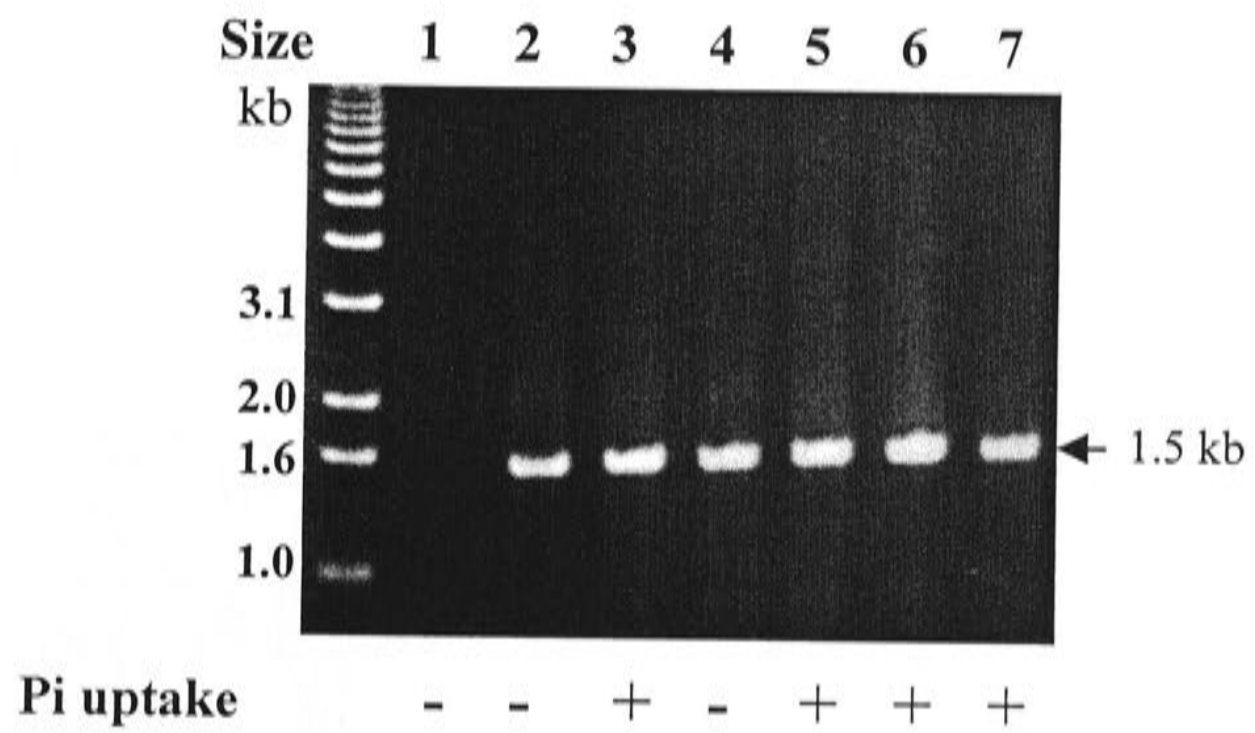
**B.** As above, except primers for the proposed *pitB* open reading frame were used (95-89, 95-90 – see Table 2.5 for details). The 1.5kb product representing an approximately full-length putative *pitB* open reading frame is indicated with an arrow.

### Table 3.1

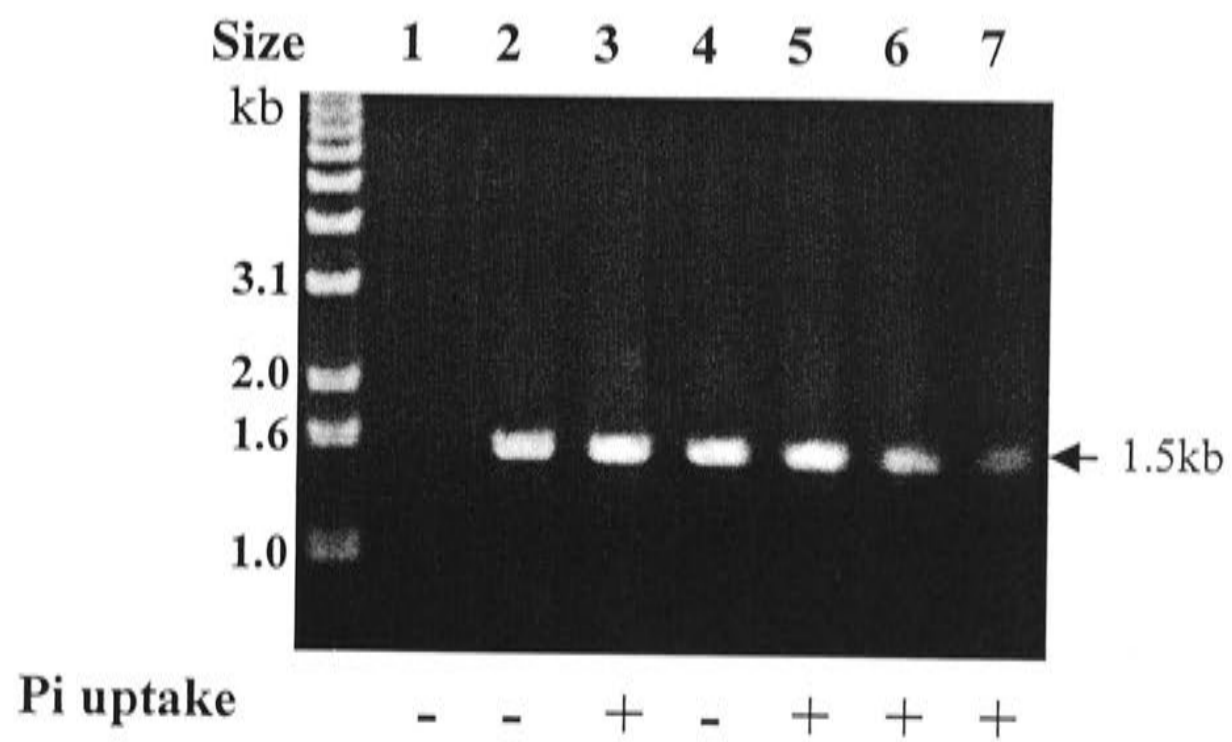
#### Strains tested for the presence of *pitA* and *pitB* DNA by PCR

Lane	Strain	Inorganic phosphate transport		PCR products	
		Phenotype	Known genotype	<i>pitA</i> (kb)	<i>pitB</i> (kb)
1	water	-	-	1.5	1.5
2	HR187	-	<i>pit pstA</i>	1.5	1.5
3	K-10	+	<i>pit</i>	1.5	1.5
4	AN3066	-	<i>pit _pstC345</i>	1.5	1.5
5	AN248	+	wild type	1.5	1.5
6	AN259	+	wild type	1.5	1.5
7	AN236	+	wild type	1.5	1.5

**A** Presence of the *pitA* open reading frame.



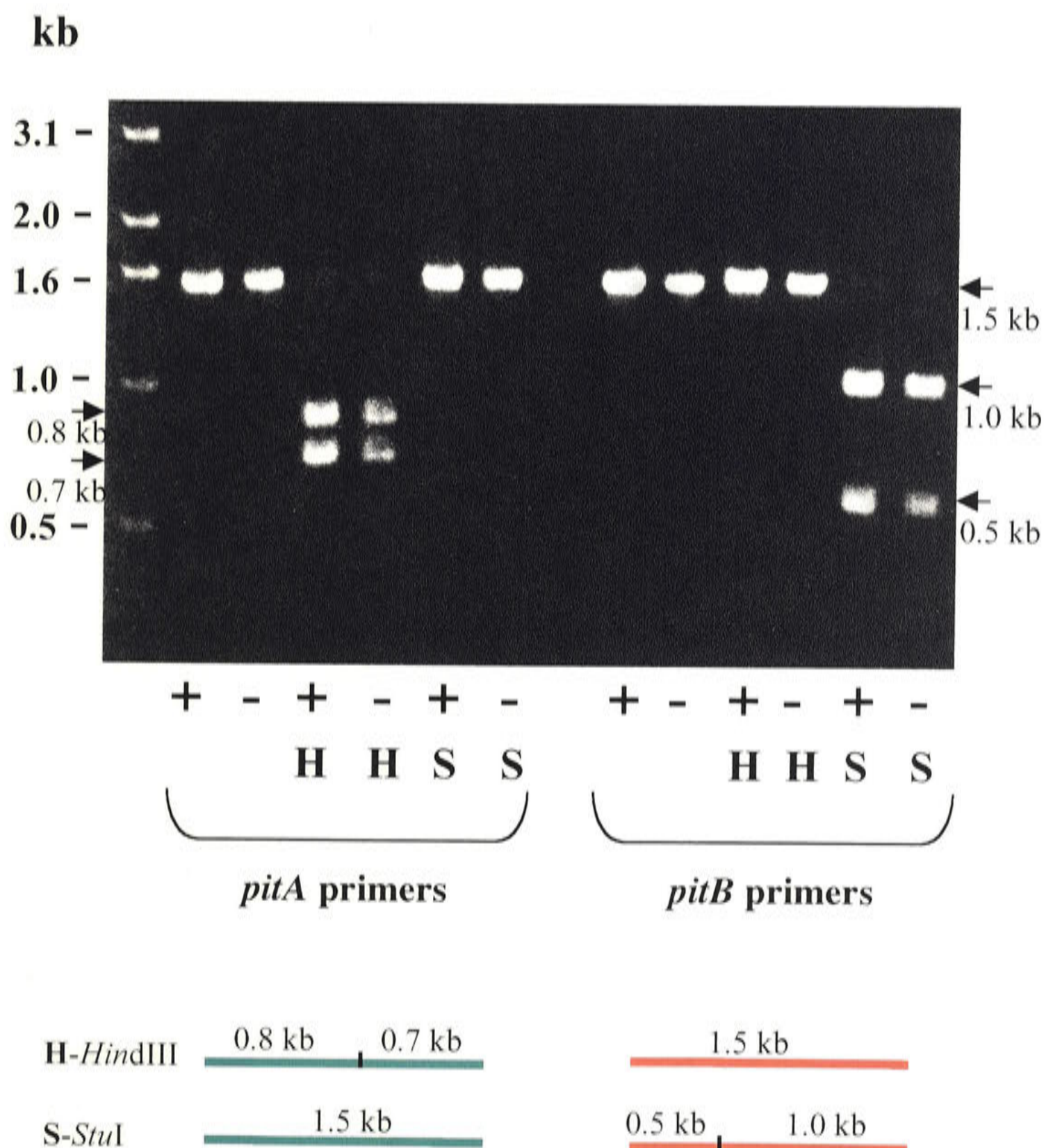
**B** Presence of the *pitB* open reading frame.



### Figure 3.4

#### Restriction endonuclease analysis of genomic *pitA* and *pitB* PCR products amplified from wild type and *pit* phenotype strains.

The PCR products from wild type (AN248) and *pit* phenotype (AN3066) genomic DNA using primers for the open reading frames of either *pitA* (95-87/95-88, see Table 2.5 for details) or *pitB* (95-89/95-90) were digested with either *Hind*III or *Stu*I then separated on a 0.8% agarose gel.



- + wild type genomic DNA (AN248)
- *pit* genomic DNA (AN3066)



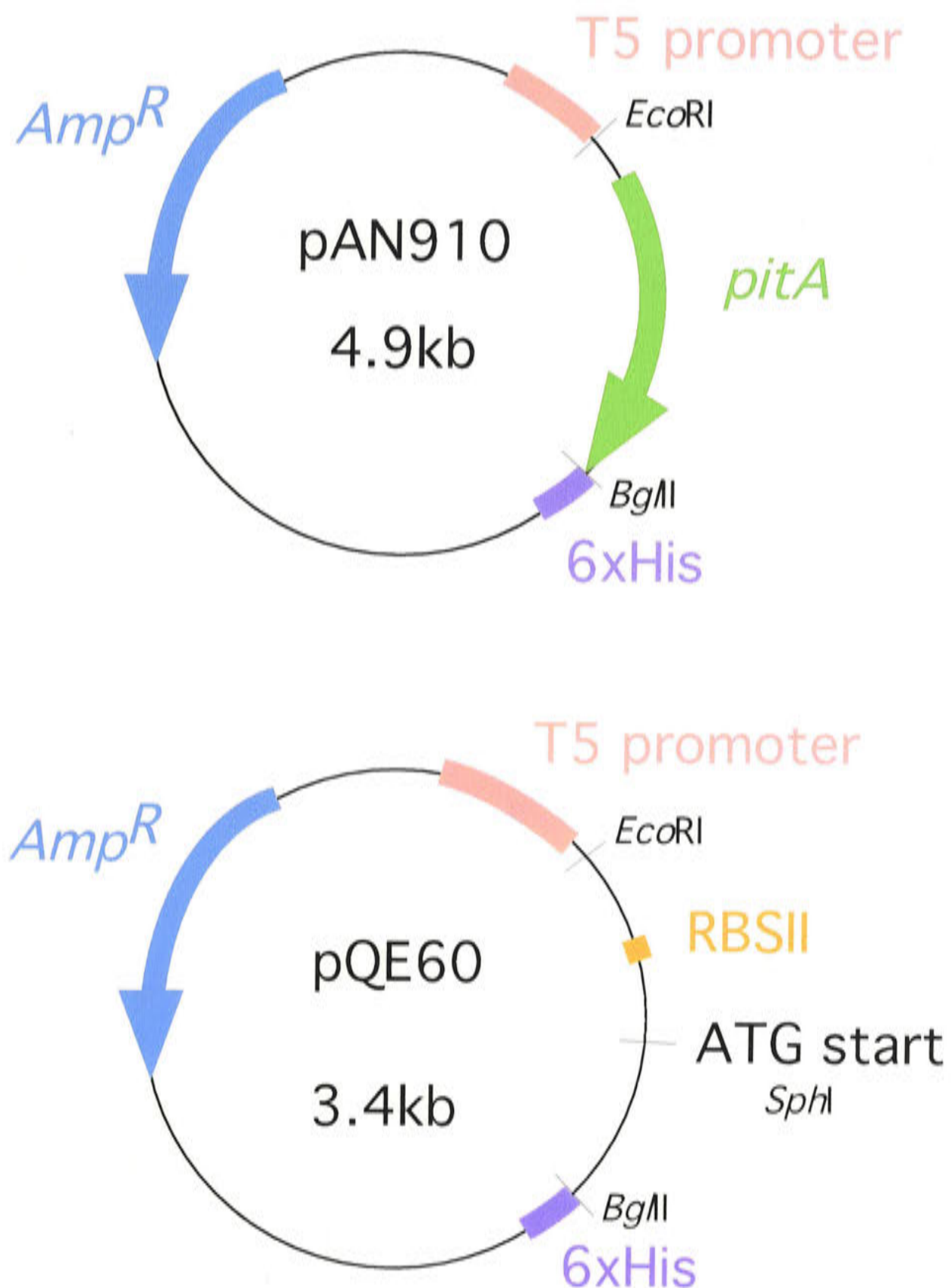
## Figure 3.5

### Plasmid diagrams for pAN910 and vector pQE60

*pitA* without its putative promoter region was ligated into vector pQE60 (QIAGEN), to give *pitA* with its native open reading frame start and a C-terminal 6xhistidine tag.

An *EcoRI* site was introduced 31 nucleotides upstream of the putative ATG start codon and a *BglII* site was introduced at the TAA stop of the putative open reading frame by site-directed mutagenesis (oligomers 91-62 and 93-20, see Table 2.5 for details). The *pitA EcoRI/BglII* fragment was ligated into *EcoRI/BglII* cut pQE60, which removes the engineered ribosomal binding site of this vector, but leaves the T5 promoter region. Thus *pitA* would be expected to use the T5 promoter and its own ribosomal binding site.

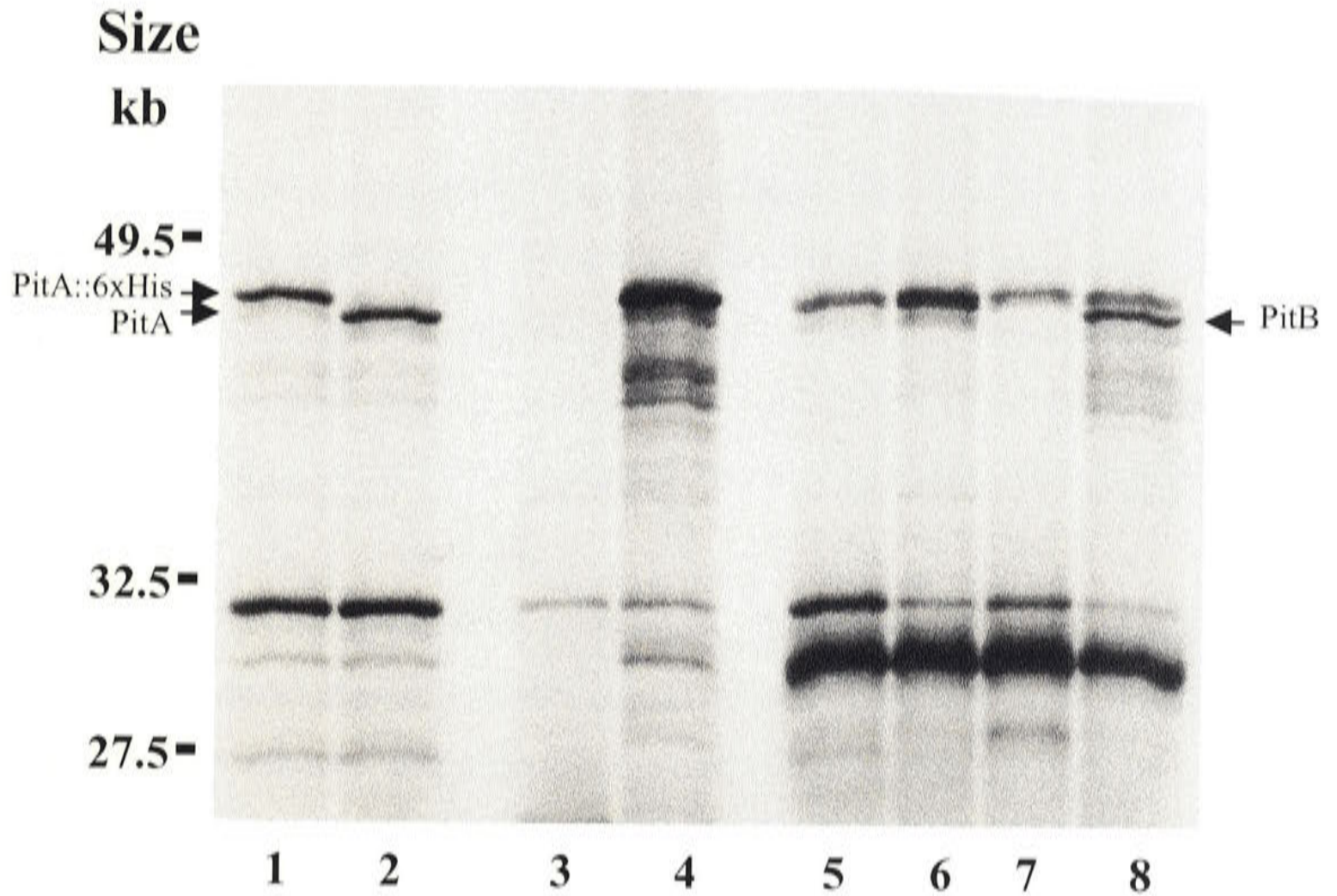
Putative *pitA* promoters are listed in Table 5.3 and the putative *pitA* promoter region is shown in Figure 5.9.



## Figure 3.6

### Gel analysis of *in vitro* transcription/translation products from *pitA*<sup>+</sup>, *PitA::6xhis*<sup>+</sup> and *pitB*<sup>+</sup> plasmids.

*In vitro* transcription/translation products from caesium chloride prepared plasmids labeled with [<sup>35</sup>S]-methionine were separated by PAGE on a 9% gel, and visualised by autoradiography with an 8 hour exposure.



Lane	Plasmid	Pit Genotype
1	pAN918	PitA::6xHis
2	pAN919	PitA
3	pQE60	-
4	pAN910	PitA::6xHis
5	pAN683	-
6	pAN914	PitA::6xHis
7	pAN683	-
8	pAN909	PitB



Purification of *pitA::6xHis* was attempted. A variety of conditions were used for both the induction of protein expression by Isopropyl- $\beta$ -D-thiogalactopyranoside (IPTG) and the subsequent purification of 6xHis tagged protein with nickel-NTA chromatography (See Section 2.8). Coomassie stained SDS-PAGE did not reveal any obvious over-expression or purification of protein between 40kDa-55kDa (results not shown). At this point in time there was no simple mechanism to specifically identify PitA protein, as commercial antibodies for the 6xHis tag were not yet available and the PitA polyclonal antibody had not been prepared.

Plasmid pAN910 was also transformed into the Pi auxotroph AN3066, to determine if the PitA6xHis protein was able to produce a functional Pi transporter. Several independent experiments showed approximately half of the transformants were able to grow on 500 $\mu$ M Pi minimal media supplemented with an IPTG gradient. Thus it is possible that this plasmid allows Pi transport but may also be detrimental to cell health.

A number of other PitA6xHis plasmids were constructed. The *pitA::6xHis* gene lacking the putative *pitA* promoter region was subcloned into pGex plasmid pAN683 (constructed by Andrew Rodgers) behind a glutathione-S-transferase fusion gene (GST $\epsilon$ 30) and in the same orientation as this gene (pAN914 - Figure 3.7). (This approach had been successfully used for other membrane proteins in the laboratory.) The pAN915 control was made by inserting the *pitA::6xHis* gene in the opposite orientation to the GST $\epsilon$ 30 gene. *pitA::6xHis* with a wild type *pitA* promoter region (Figure 5.9) was also subcloned into vector pBR322, creating pAN918 (Figure 3.8). A wild type *pitA* control plasmid was also prepared (pAN919).

*In vitro* transcription/translation of pAN918 and pAN919 showed that pAN919 (*pitA*) has a protein at 45kDa and pAN918 (*pitA::6xHis*) produces a protein of 46kDa, clearly indicating that PitA protein and the slightly larger PitA6xHis are expressed by these plasmids. This result also suggests that even though these proteins appear smaller than their molecular weights (PitA is 53kDa while PitA6xHis is 54kDa), they are full-length products. *In vitro* transcription/translation was also carried out on pAN914 and vector

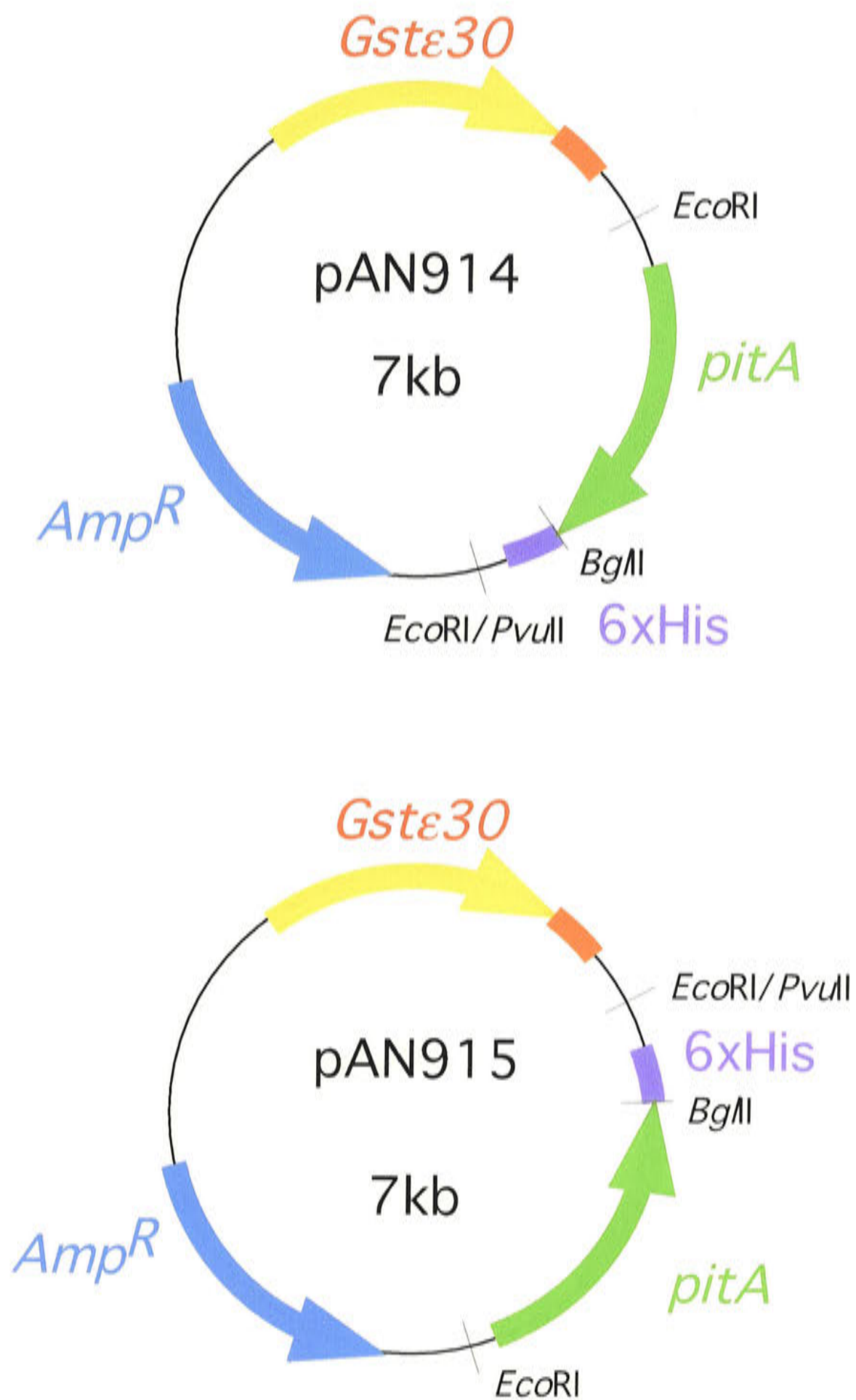


## Figure 3.7

### Plasmid diagrams for pAN914 and pAN915.

*pitA::6xhis* without its putative promoter region was ligated into pAN683 behind the glutathione-S-transferase fusion gene (*Gstε30*) in the same orientation (pAN914) and opposite orientation (pAN915) as this fusion protein.

The *pitA PvuII/EcoRI* fragment from pAN910 was ligated into the *EcoRI* site of pAN683, and restriction endonuclease digestion was used to select plasmids with *pitA::6xhis* in either orientation.



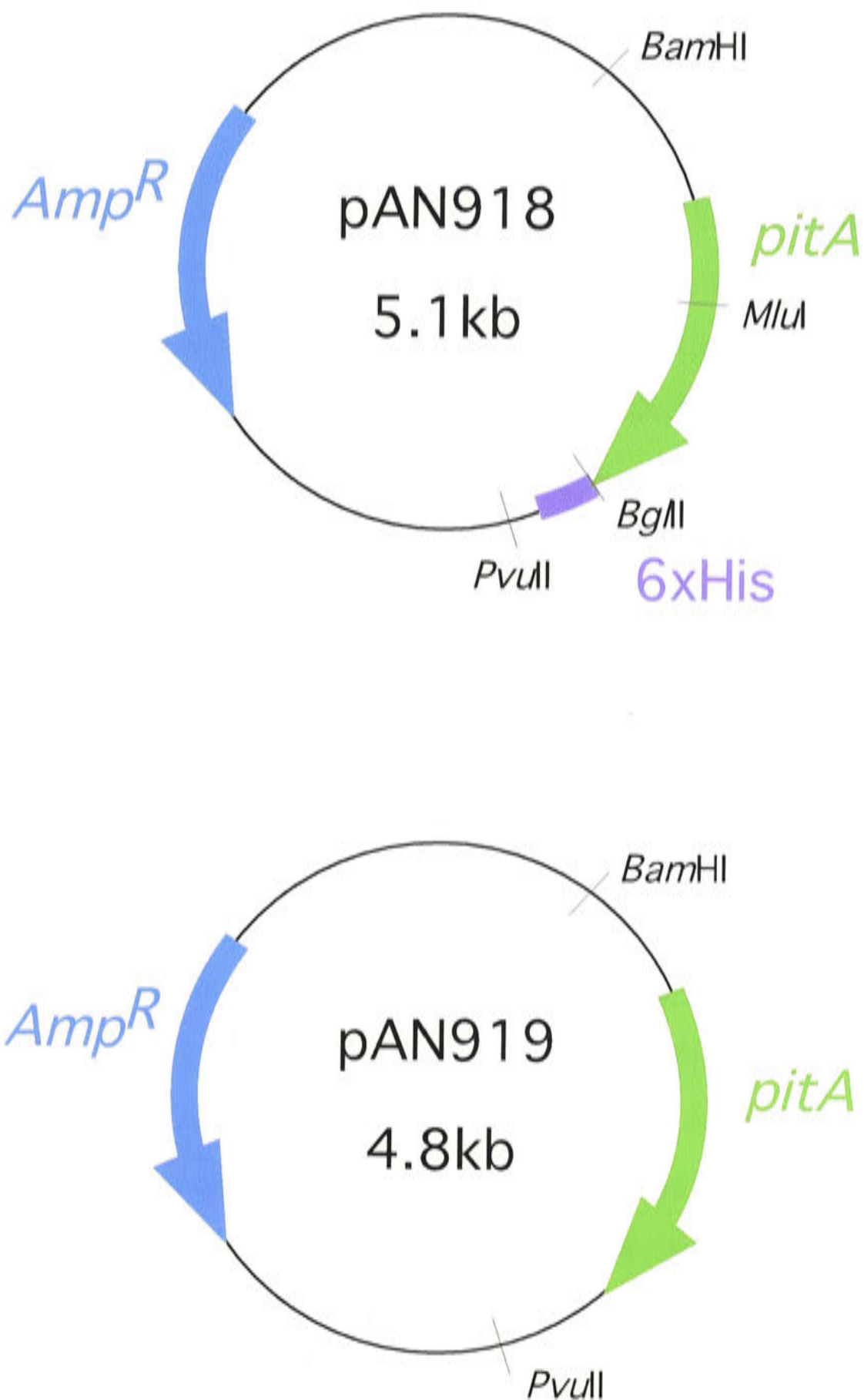
## Figure 3.8

### Plasmid diagrams for pAN918 and pAN919.

*pitA::6xhis* was transferred into pBR322 so it was expressed using its wild type promoter (pAN918), and an equivalent *pitA* control strain was also constructed (pAN919).

The C-terminal part of *pitA::6xhis* was isolated from pAN910 by digestion with *MluI/PvuII*. This was ligated into *pitA* plasmid pAN686 digested with *MluI/PvuII* to remove the N-terminal half of *pitA*, creating pAN918.

A similarly sized control plasmid was made by digesting pAN686 with *PvuII* and religating the large fragment, to create pAN919.



pAN683 (Figure 3.6). Unfortunately the lac repressor protein appears to run in an identical location to the PitA6xHis protein, so there is no definitive proof of protein expression for this construct.

*pitA::6xHis* plasmid pAN918 transformed into the Pi auxotrophic strain AN3066 was unable to grow on Pi media. However AN3066 transformed with the wild type *pitA* control (pAN919) was able to grow on Pi media, suggesting that the C-terminal 6xHis tag has a detrimental effect on the activity of PitA. This inhibition of growth by the 6xHis tag is not consistent across all plasmids. *pitA::6xHis* in pGex plasmid pAN914 allowed growth of AN3066 on Pi media that was induced by IPTG, indicating a link between gene expression and Pi uptake. The pAN915 control showed negligible growth on Pi media (results not shown). Thus different *pitA::6xHis* plasmid constructs induced different Pi uptake phenotypes.

No attempts were made to purify PitA6xHis from pAN918. Induction and purification of PitA6xHis from pAN914 produced no obvious protein product between 45kDa and 55kDa using a variety of growth and purification conditions. Thus it was decided to leave the purification at this stage of development until the PitA6xHis protein could be specifically identified.

Attempts to over-express *pitB* by removing its putative promoter region and placing it behind the GSTε30 fusion gene in pAN683, to create pAN909 (Figure 3.9), produced a protein by *in vitro* transcription/translation experiments at approximately 45kDa which seemed negligible in pAN683 (Figure 3.6). Therefore PitB is a similar size to the wild type PitA protein by SDS-PAGE analysis. While pAN909 seems to produce a functional Pi transporter, allowing the Pi auxotrophic strain AN3066 to grow on solid Pi media, the growth of these cells was noticeably impaired when compared to AN3066 containing wild type *pitB* plasmids pAN656 or pAN1116 (plasmid diagrams - Figure 5.5). When the pAN909/AN3066 transformants were grown on rich media, which eliminates any requirement for Pi transport, these also grew slowly and exhibited spotty uneven growth, even in the absence of IPTG (results not shown). IPTG induction of pAN909 transformed into strain JS5 did not produce an obvious protein band when analysed by Coomassie stained SDS-PAGE, and it was decided to leave the attempted



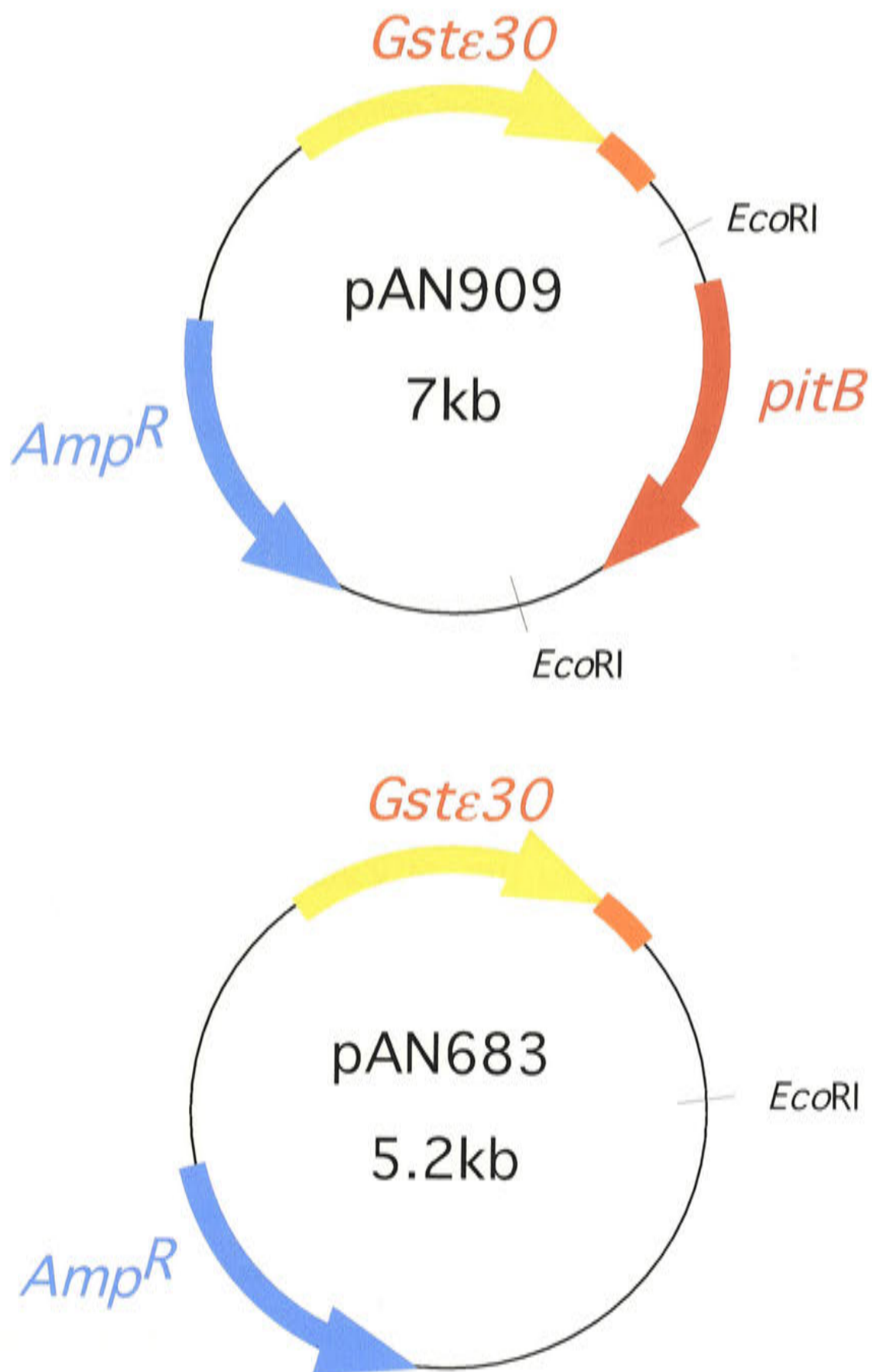
## Figure 3.9

### Plasmid diagrams for pAN683 and pAN909.

*pitB* without its putative promoter region was placed behind a glutathione-S-transferase fusion gene (*GSTε30*) to create pAN909.

An *EcoRI* site was inserted 23 nucleotides upstream from the putative open reading frame of *pitB* (this mutagenesis was carried out by Di Webb), allowing *pitB* without its putative promoter region to be isolated on an *EcoRI* fragment. This was ligated into the *EcoRI* site of pAN683 behind the glutathione-S-transferase fusion gene (*GSTε30*). Restriction endonuclease analysis was used to isolate a plasmid with *pitB* in the same orientation as the *GSTε30* fusion gene (pAN909).

pAN683 was created by Andrew Rodgers.



over-expression of PitB until a suitable method was developed for the tagged PitA protein.

### ***3.5 Identification of the translation starts of the *pitA* and *pitB* open reading frames using site-directed mutagenesis***

As over-expression of PitA and PitB protein was proving difficult, an alternative approach was used to identify the translation starts of these genes. Functional analysis of alleles carrying mutated start codons has previously been used to identify ATG translation start codons (233). This approach is only reasonable if the targeted ATG codon has a high probability of being the translation start of the gene. Both *pitA* and *pitB* have an ATG codon at the beginning of a 1500 nucleotide ORF, with the nucleotide sequence similarity between the two cosmid DNA fragments being strongest over this region. Thus, it is highly probable that these ATG codons may initiate translation of the *pit* genes.

These putative ATG start codons were changed to GTG and CTG by site-directed mutagenesis, using oligonucleotides 93-107, 93-108 for *pitA* and 93-135, 93-136 (see Table 2.5 for sequence details) and single stranded *pitA* or *pitB* M13 DNA. Manual DNA sequencing of mutated alleles on single stranded M13 DNA was carried out over the region of oligonucleotide binding to check that the correct mutation had been inserted, and that no unintended mutations had occurred. Double stranded RF M13 DNA was prepared for each mutated allele, which was then subcloned into the vector pBR322 using *SalI/BamHI* digestion for *pitA* and by ligating the *SspI/HindIII* fragment from *pitBM13* RF into the *EcoRV/HindIII* sites of pBR322 (Figure 3.10). These plasmids were transformed into the Pi auxotroph AN3066 and these cells were cultured in the presence of  $\alpha$ -glycerol-3-phosphate (G3P), providing an alternative source of phosphate to remove any selection pressure for Pi transport. All strains were able to grow when plated onto minimal media containing 500 $\mu$ M Pi, indicating that all mutated alleles resulted in the synthesis of some protein. The initial rates of Pi uptake (where uptake was linear) were then measured for wild type *pitA* and *pitB*, and the associated alleles. The initial rates of Pi uptake were measured for several isolates for each allele,



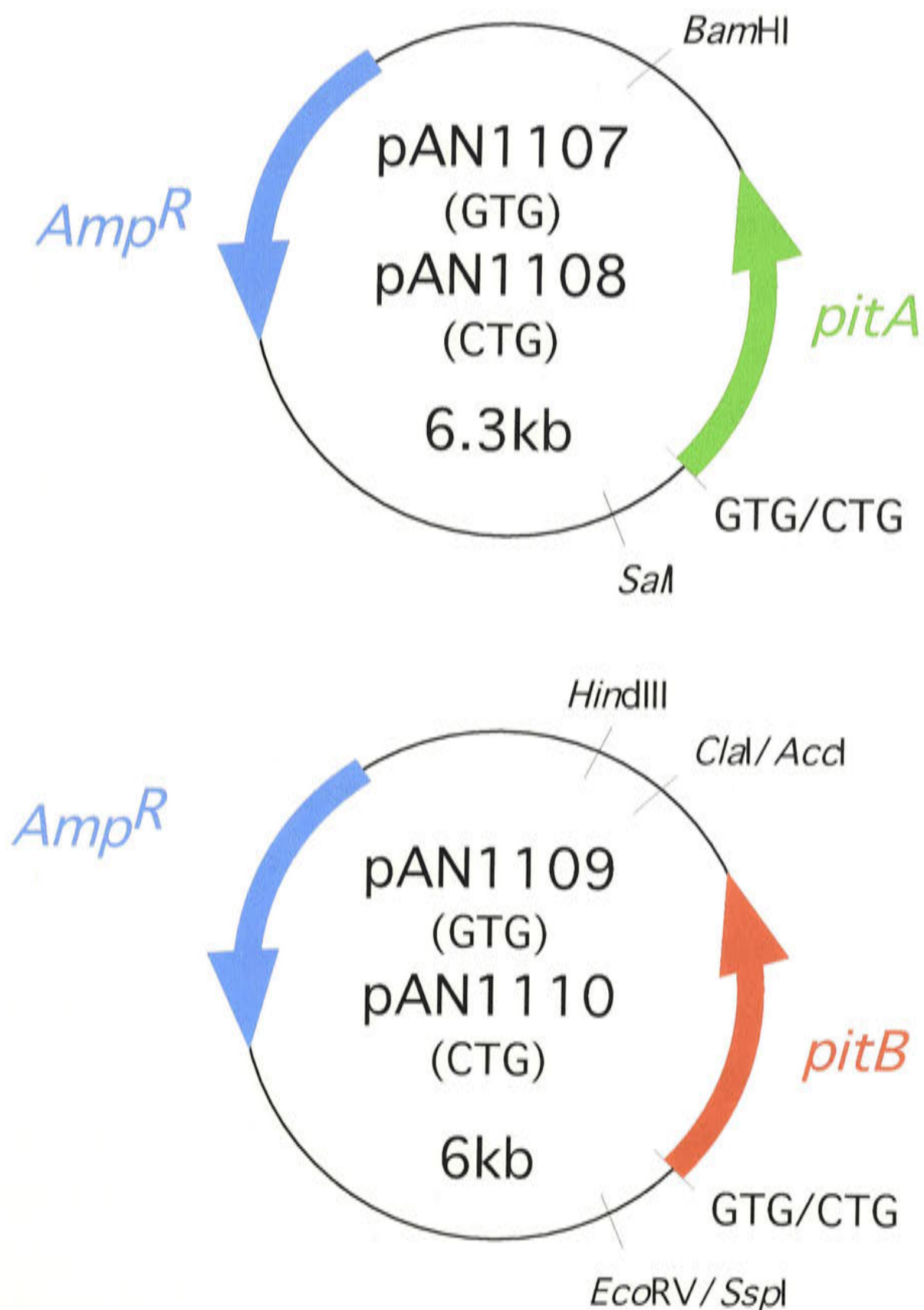
### Figure 3.10

#### Plasmid diagrams for pAN1107, pAN1108, pAN1109 and pAN1110.

The putative ATG start codons of *pitA* and *pitB* were altered to GTG or CTG by site-directed mutagenesis.

The putative start codon of wild type *pitA* in M13mp18 was altered using oligonucleotides 93-107 (GTG) or 93-108 (CTG - see Table 2.5 for details). The *pitA* *SalI/BamHI* fragments were ligated into the *SalI/BamHI* sites of vector pBR322 to create pAN1107 (GTG) and pAN1108 (CTG).

The putative start codon of wild type *pitB* in M13mp18 was altered using oligonucleotides 93-135 (GTG) or 93-136 (CTG - see Table 2.5 for details). The *pitB* *SspI/HindIII* fragments were ligated into the *EcoRV/HindIII* sites of vector pBR322 to create pAN1109 (GTG) and pAN1110 (CTG).





as the full DNA sequence of each mutant had not been verified. For each mutated start codon the isolates gave similar rates of uptake (data not shown) and repeat assays were subsequently performed on only one isolate for each allele. Examples of the initial rates of Pi uptake are given in Figure 3.11, with a summary of the data (a compilation of three experiments for *pitB* and 6 experiments for *pitA*) in Table 3.2. These results show that for both *pitA* and *pitB* changing the putative ATG start codon to GTG and then CTG caused a progressive drop in the Pi uptake activity of these mutants. This is likely to occur when the mutated ATG codon forms the start of the translated protein, as the GTG and CTG changes progressively reduce affinity for the ribosomal initiation complex, effectively decreasing translation of the encoded proteins.

*pitA* has several in frame ATG codons within this large open reading frame at 166, 235, 286 and 298 nucleotides. To check if these are possible starts to the *pitA* gene site-directed mutagenesis was used to place in frame termination codons before (at 118 nucleotides) and between (at 208 and 268 nucleotides) these methionine codons (Figure 3.12) using oligonucleotides 94-55, 94-56 and 94-57 (Table 2.5). The mutations were confirmed, subcloned into pBR322 and transformed into AN3066 by the methods previously described in this section. All mutated *pitA* strains failed to grow on Pi media, indicating all *pitA* Pi uptake activity has been abolished. In addition, two termination codons inserted in frame only 13 nucleotides after *pitB*'s putative ATG start completely abolished all Pi uptake activity by this gene (89).

The combination of the above results indicates that translation of the *pit* genes commences at the beginning of the identified ORFs.

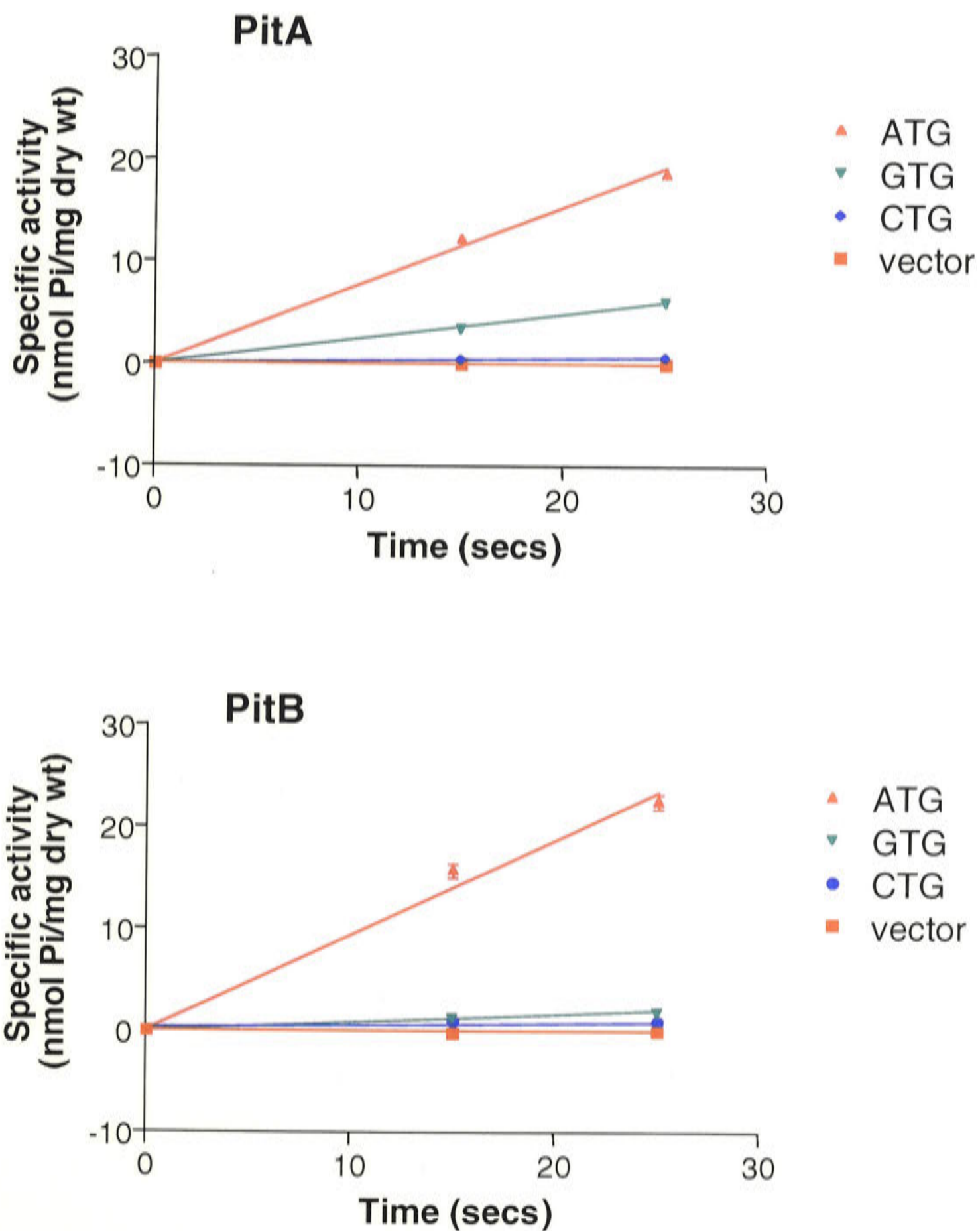
### ***3.6 Comparison of the deduced amino acid sequences of PitA and PitB***

The open reading frames of *pitA* and *pitB* were translated into amino acid sequences and aligned, using the ClustalW program in MacVector 6.0 (Figure 3.13). Both consist of 499 amino acids with 81% identity in the deduced amino acid sequences. Sequence variability appears to occur in patches, with a large variable region between alanine 266 and serine 373.

### Figure 3.11

#### Initial rates of phosphate uptake from *pitA* and *pitB* with altered putative start codons.

The putative ATG start of each large open reading frame on the DNA fragments containing either PitA or PitB activity were mutated to GTG or CTG by site-directed mutagenesis, then assayed for Pi transport. The initial rates shown below were used to formulate Table 3.2. These rates were calculated using six (*pitA*) or three (*pitB*) individual determinations using linear regression in the Graphpad Prism program. Error bars represent the standard error of the mean (SEM).



**TABLE 3.2:** Phosphate uptake activity of PitA and PitB with altered putative start codons

Putative start codon	Comparison with wild type Pi uptake activity (%)	
	PitA <sup>a</sup>	PitB <sup>b</sup>
vector	<0.5	<0.5
ATG	100 ± 1.6	100 ± 2.7
GTG	32.4 ± 0.6	8.4 ± 1.0
CTG	3.9 ± 0.1	4.4 ± 0.5

a assayed at 6µM Pi, SEM, n = 6  
100% = 45nmol Pi min<sup>-1</sup> mg dry weight<sup>-1</sup>

b assayed at 50µM Pi, SEM, n = 3  
100% = 55nmol Pi min<sup>-1</sup> mg dry weight<sup>-1</sup>



## Figure 3.12

### Analysis of putative *pitA* ATG starts within the putative open reading frame by the insertion of stop codons.

The largest open reading frame (ORF) on the DNA fragment containing *pitA* activity has an in frame stop codon just before the putative ATG start, and several in frame ATG codons at 166, 235, 286 and 298 nucleotides within this ORF. In frame stop codons were inserted at 118, 208 and 268 nucleotides using the oligonucleotides 94-55, 94-56 and 94-57 (see Table 2.5 for details).

```

                                1 (ORF start)
TAA CTG CCT ATG CTA CAT TTG TTT GCT GGC CTG
25
GAT TTG CAT ACC GGG CTG TTA TTA TTG CTT GCA
58
CTG GCT TTT GTG CTG TTC TAC GAA GCC ATC AAT
91
                                118
                                TAA
GGT TTC CAT GAC ACA GCC AAC GCC GTG GCA ACC
124
GTT ATC TAT ACC CGC GCG ATG CGT TCT CAG CTC
157
                                166
GCC GTG GTT ATG GCG GCG GTA TTC AAC TTT TTG
190
                                208
                                TAG
GGT GTT TTG CTG GGT GGT CTG AGT GTT GCC TAT
223
                                235
GCC ATT GTG CAT ATG CTG CCG ACG GAT CTG CTG
256
                                268
                                TAG
                                286
CTT AAT ATG GGA TCG TCT CAT GGC CTT GCC ATG
289
                                298
GTG TTC TCT ATG TTG CTG GCG GCG ATT ATC TGG
```

## Figure 3.13

### Alignment of PitA and PitB deduced amino acid sequences.

- A Represents amino acids identical between PitA and PitB.
- a Represents amino acids unique to PitA.
- a Represents amino acids unique to PitB.

This alignment was carried out with ClustalW (MacVector 6.0) using a Blosum 30 scoring matrix with an opening gap penalty of 10 and an extending gap penalty of 0.1.

	1					
PitA	MLhLFaGLDl	htGLLLLLLAL	AFVLFYEAIN	GFHDTANAVA	tVIYTRAMrs	
PitB	MLnLFvGLDi	yTGLLLLLLLAL	AFVLFYEAIN	GFHDTANAVA	aVIYTRAMqp	
	50					
PitA	QLAVVMAAvF	NFLGVLLGGL	SVAYAI VHML	PTDLLLLNMGS	sHGLAMVFSM	
PitB	QLAVVMAAfF	NFFGVLLGGL	SVAYAI VHML	PTDLLLLNMGS	tHGLAMVFSM	
	100					
PitA	LLAIIWNLG	TWYFGLPASS	SHTLIGAIIG	IGLTNALMTG	tSVvDALNip	
PitB	LLAIIWNLG	TWFFGLPASS	SHTLIGAIIG	IGLTNALLTG	sSVmDALNlr	
	150					
PitA	kVlsIFgSLI	VSPIVGLVfa	GGLIFLLRRY	WSGTKKRARI	HltPaEREKK	
PitB	eVtkIFsSLI	VSPIVGLVIA	GGLIFLLRRY	WSGTKKRdRI	HriPeDRkkK	
	200					
PitA	dGkkKPPFWT	RIALIISAiG	VAFSHGANDG	QKGIGLVMLV	LiGvAPAGFV	
PitB	kGKrKPPFWT	RIALIVSAaG	VAFSHGANDG	QKGIGLVMLV	LvGiAPAGFV	
	250					
PitA	VNMNAtGYEI	TRTRDAInNv	EaYFeQHPaL	lkqatgaDql	vPApeagaTQ	
PitB	VNMNASGYEI	TRTRDAVtNf	EhYLqQHPeL	pgkliamEpp	lPAastdgTQ	
	300					
PitA	paEFHCHPsN	TinAlnRlKg	MLttDvesyd	klsldQrsqm	rRIMLCVSDT	
PitB	vtEFHCHPaN	TfdAiaRvKt	MLpgnIgklr	avkrEsaqpa	gRIMLCISDT	
	350					
PitA	idKvvKMPGV	SaDDQrLLKK	LkSDMLSTIE	YAPVWIIMAV	ALALGIGTMI	
PitB	saKlaKLPGV	SkEDQnLLKK	LrSDMLSTIE	YAPVWIIMAV	ALALGIGTMI	
	400					
PitA	GWRRVAtTIG	EKIGKkGMTY	AQGMSAQMTA	AVSIGLASyT	GMPVSTTHVL	
PitB	GWRRVAmTIG	EKIGKrGMTY	AQGMaAQMTA	AVSIGLASyI	GMPVSTTHVL	
	450					
PitA	SSsVAGTMVV	DGGGLQRKTV	TSILMAWVFT	LPAAVLLSGG	LYWlsLQFl	
PitB	SSaVAGTMVV	DGGGLQRKTV	TSILMAWVFT	LPAAIFLSGG	LYWiaLQLi	



Membrane proteins are predicted to contain hydrophobic regions capable of spanning the membrane as  $\alpha$ -helices or  $\beta$ -sheets. The deduced amino acid compositions of PitA and PitB indicate that these proteins are moderately hydrophobic, with nonpolar residues making up 53% and 54% of each protein, respectively. Hydropathy profiles were determined by the methods of Kyte and Doolittle (138), and Goldman, Engleman and Steitz (GES) (63). The average hydrophobicity of a moving window of 19 amino acids was calculated against every position in each sequence. These profiles produced similar results and suggest that both PitA and PitB contain 9 or 10 hydrophobic regions large enough to span the cytoplasmic membrane as an  $\alpha$ -helix (Figure 3.14).

von Heijne's TopPred IV program combines the GES analysis with the 'positive charge inside' rule (276) to determine the orientation of a protein in the cytoplasmic membrane and the positioning of the putative  $\alpha$ -helices. von Heijne found that positively charged residues were four times more prevalent in the cytoplasmic loops of bacterial membrane proteins of known topology than in the periplasmic loops, while there appears to be no preference for negatively charged side chains on either side of the membrane (276). The MEMSAT program, which optimally 'threads' a polypeptide chain through a set of topology models, was also used (111). The putative models produced by these programs for PitA and PitB are very similar and a topological model based on the results for PitA is shown in Figure 3.15, containing 10 transmembrane domains. The amino acid residues which differ between PitA and PitB are highlighted in this figure, showing that most variability occurs in the putative hydrophilic loops. The topology of PitA is explored in more detail in Chapter 6.

### **3.7 Kinetic parameters of PitA and PitB**

To characterise the kinetic properties of PitA and PitB with respect to Pi uptake, the relevant DNA fragments were subcloned into plasmid vector pBR322 to produce pAN656 and pAN686 (Figure 3.16). These plasmids were transformed into the Pi auxotrophic strain AN3066 (*pitA1*  $\Delta$ *pstC345*), which does not grow in minimal media supplemented with 500 $\mu$ M Pi (Pi media), but will grow with the addition of G3P. Both pAN686 (*pitA*) and pAN656 (*pitB*) support the growth of this strain on Pi media (data not shown). These cells expressing *pitA* or *pitB* were then cultured, deprived of

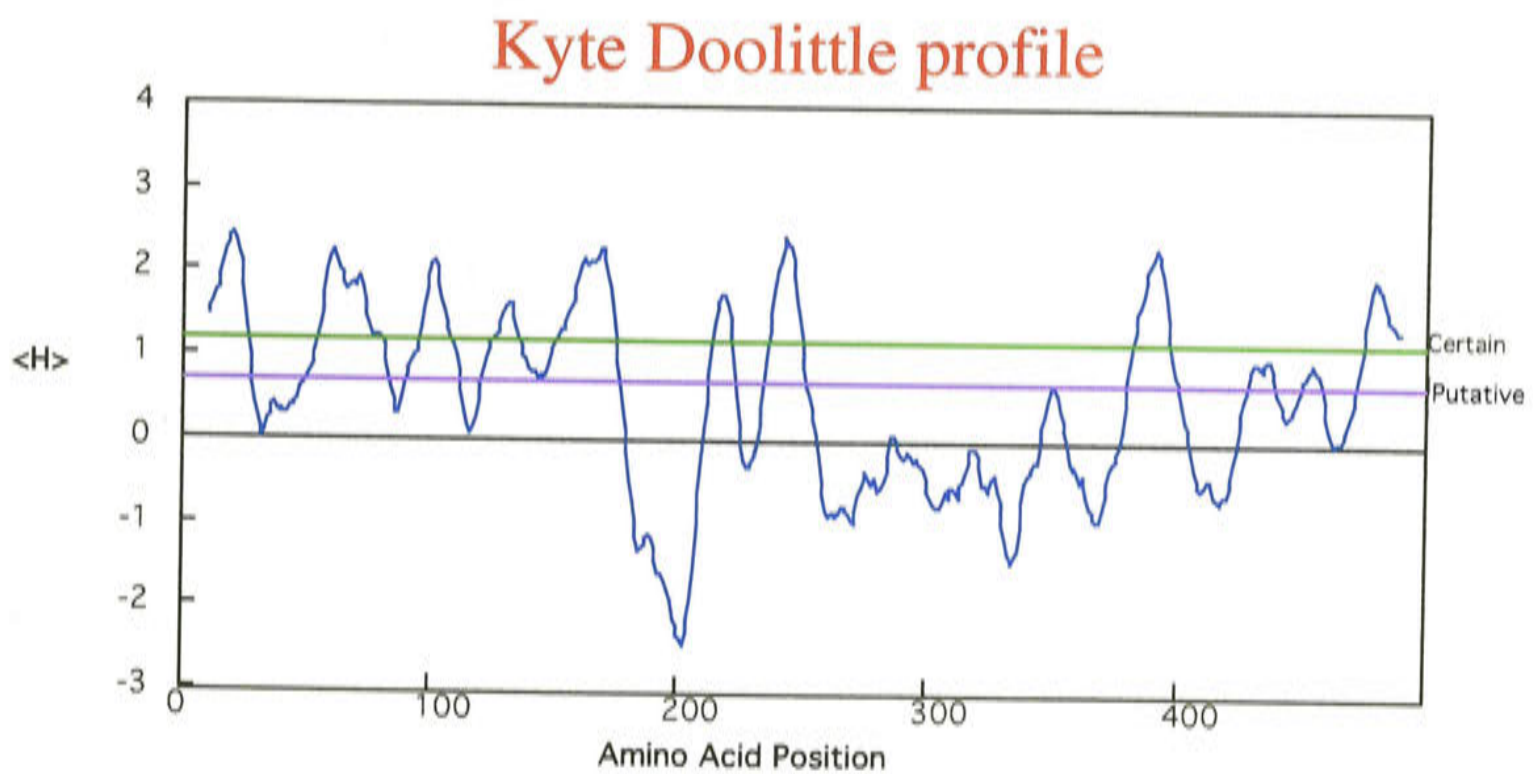
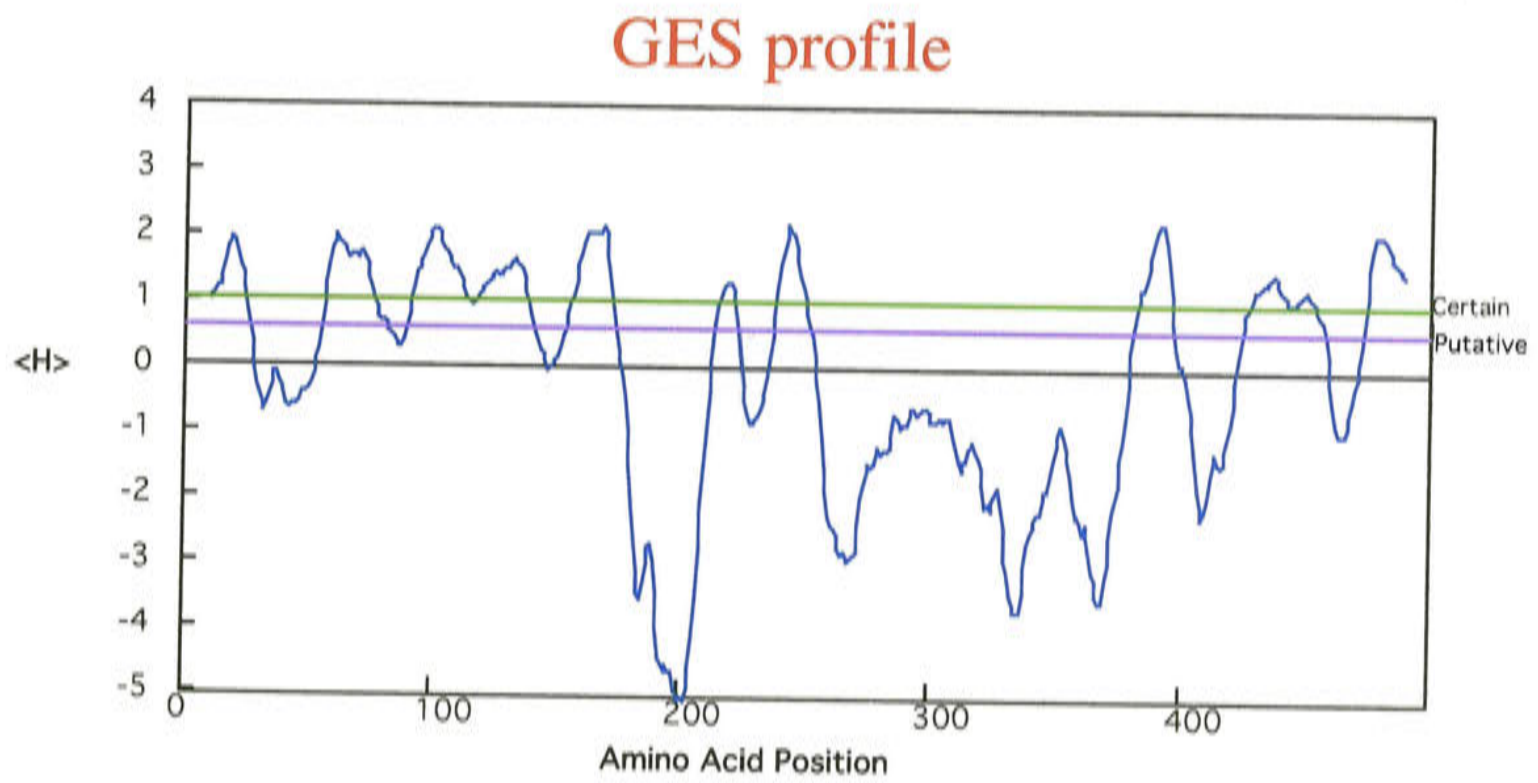


## Figure 3.14

### Hydropathy profiles for the deduced amino acid sequences of *pitA* and *pitB*.

Each amino acid sequence was analysed by the methods of Goldman, Engleman and Steitz (GES) (63) and Kyte and Doolittle (KD) (138), using a moving window of 19 amino acids. This was done through the TopPred IV program.

#### A PitA hydropathy profiles

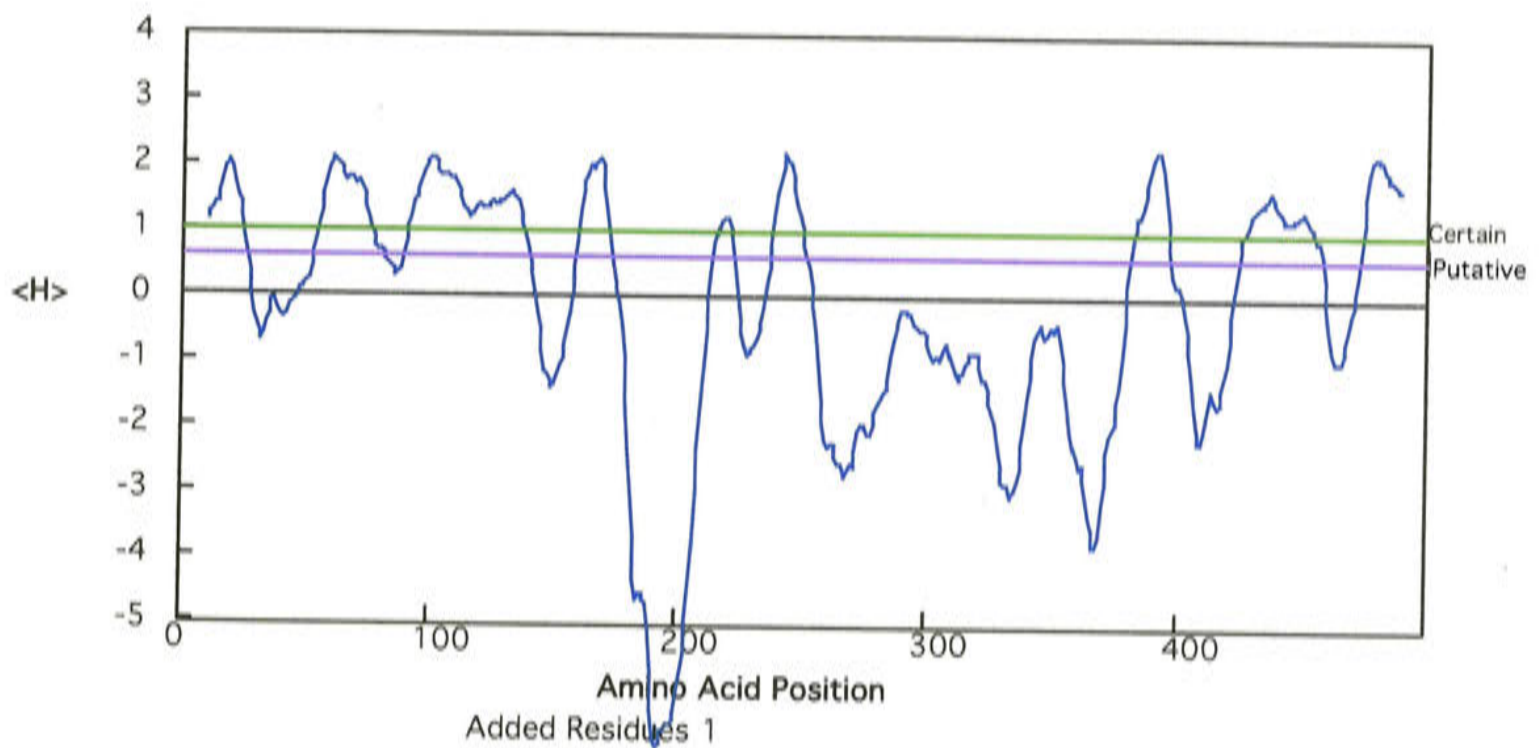


## Figure 3.14

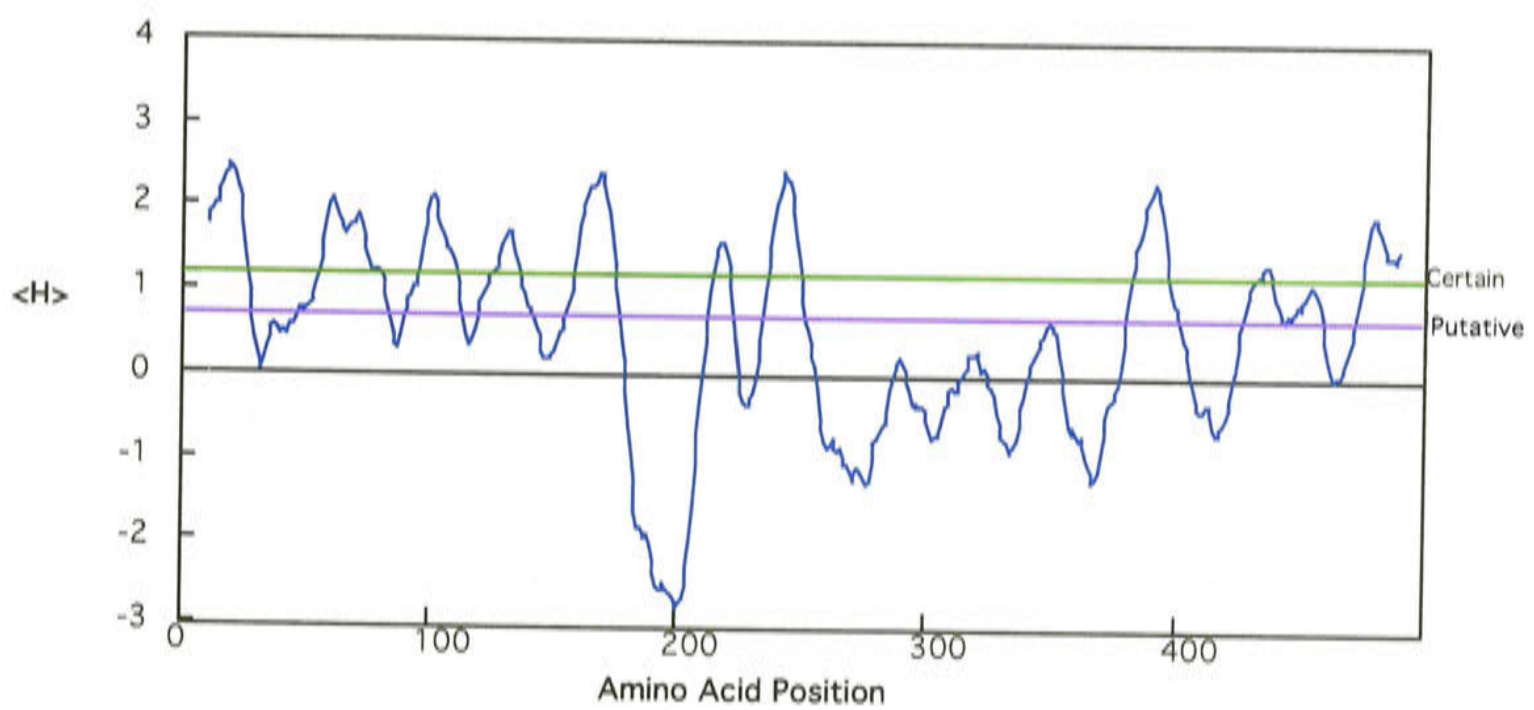
Hydropathy profiles for the deduced amino acid sequences of *pitA* and *pitB*.

### B PitB hydropathy profiles

#### GES profile



#### Kyte Doolittle profile



## Figure 3.15

### Topological model of PitA.

This model was formulated by combining information from GES (63) and Kyte-Doolittle (138) hydrophathy profiles (using a moving window of 19 amino acids) with the results of the prediction programs TopPred IV (276) and MEMSAT (111).

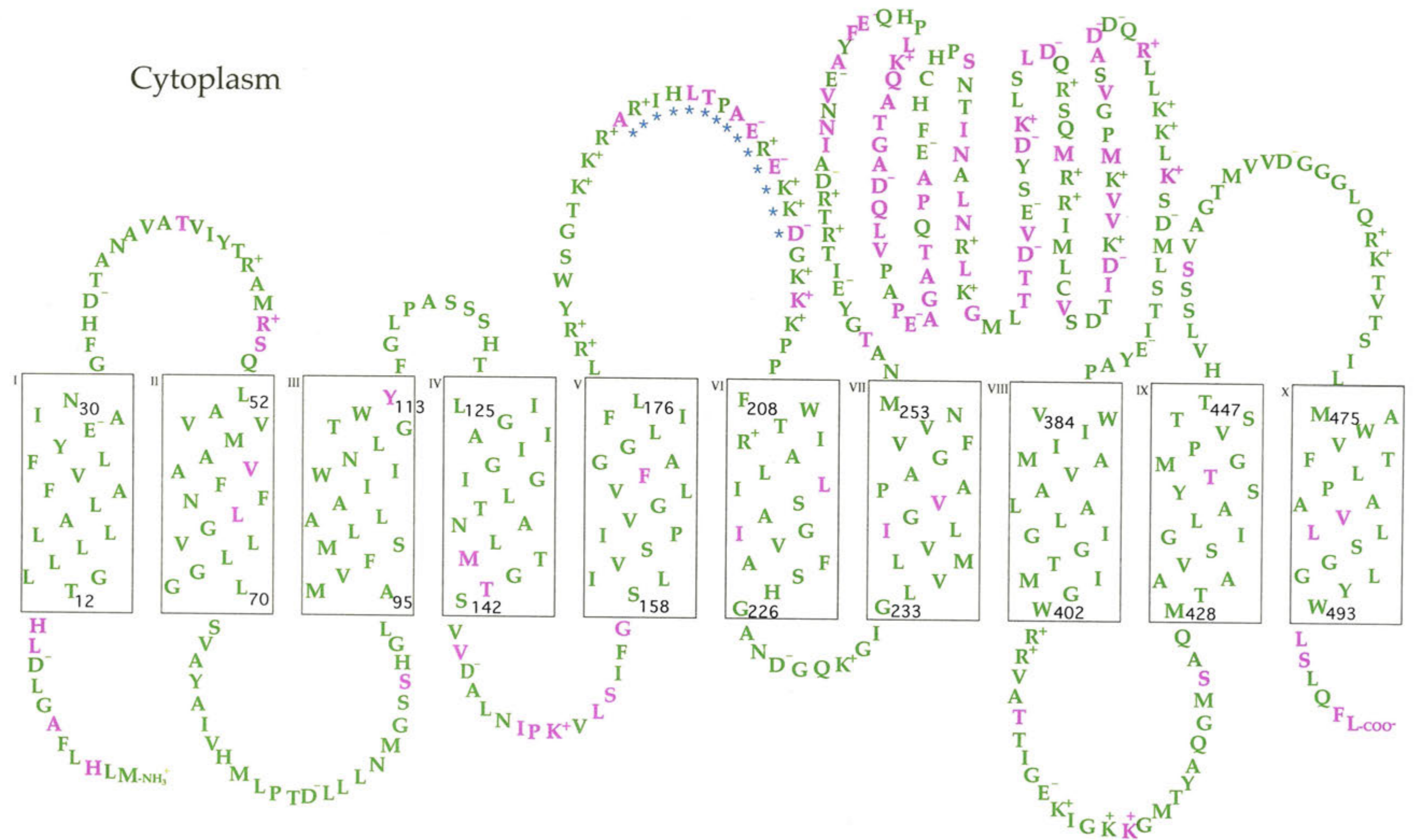
The models proposed for PitA and PitB were identical in topology, so the PitA sequence is shown, highlighting the amino acids which show variability when compared with the PitB sequence.

- A** Amino acids identical between PitA and PitB.
- A** PitA amino acids which are different in the PitB sequence.
- \*** Highlights the peptide sequence used to create the polyclonal PitA antipeptide antibody. The equivalent region of PitB was used to create the polyclonal PitB antipeptide antibody.

The boxes represent the putative transmembrane  $\alpha$ -helices, which have been drawn to contain 19 amino acids, the predicted minimum number needed to cross the cytoplasmic membrane. The single letter amino acid code has been used.



Cytoplasm



Periplasm

## Figure 3.16

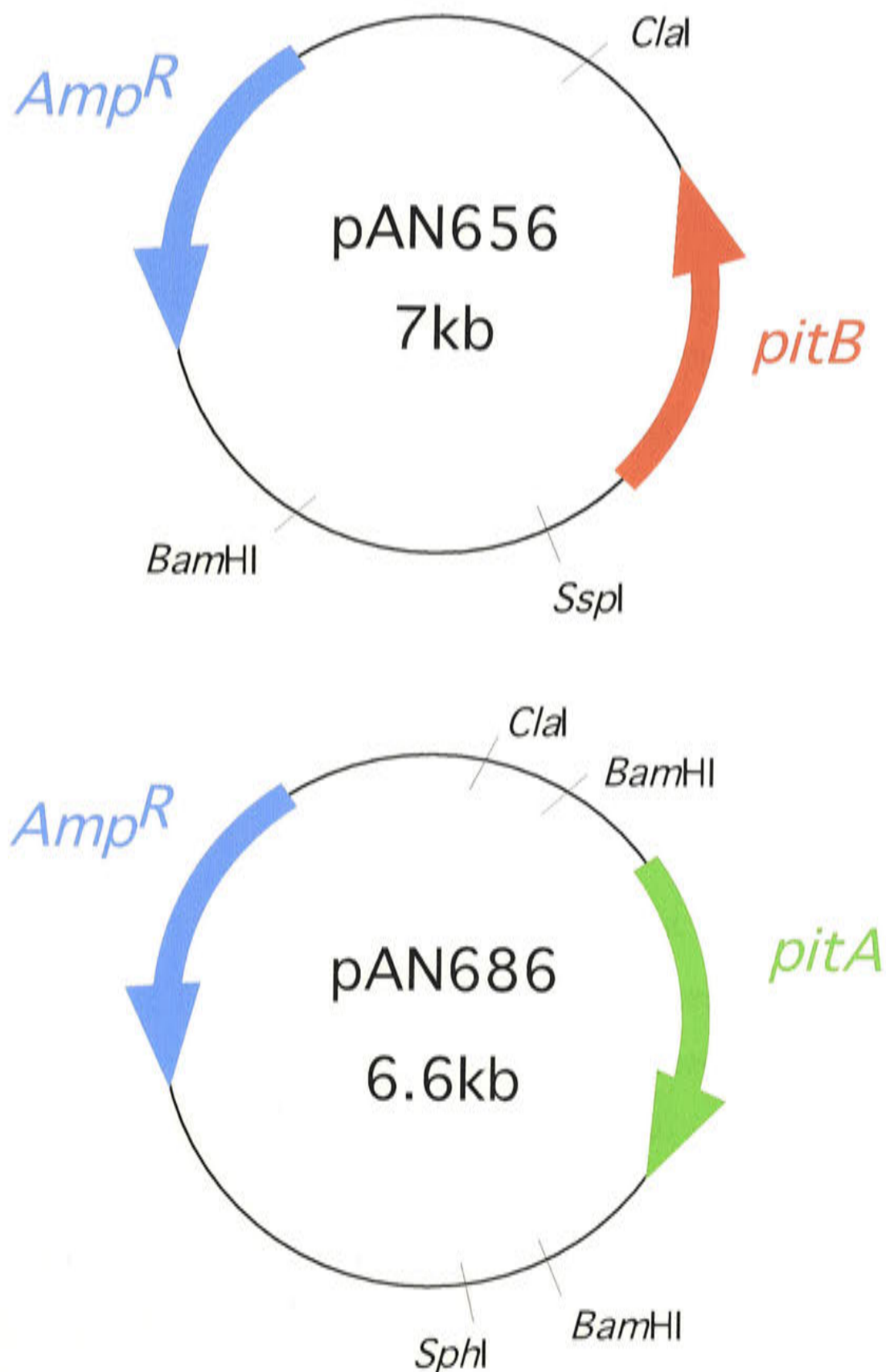
### Plasmid diagrams for pAN656 and pAN686.

Wild type *pitA* or *pitB* on vector pBR322.

*pitA* was isolated from pCE27 (Elvin *et al* (59)) on a *ClaI/SphI* fragment, which was then ligated into the *ClaI/SphI* sites of pBR322 to form pAN686. (This plasmid effectively represents genomic *pitA* DNA on a *BamHI* fragment cloned into the pBR322 *BamHI* site, as pCE27 consists of a *pitA* *BamHI* fragment within cosmid vector pHc79, which is a pBR322 derivative.)

*pitB* was isolated on a cosmid *ClaI/BamHI* fragment (Harris *et al* (89)) and ligated into the *ClaI/BamHI* of pBR322, creating pAN656.

Both the above plasmids were constructed by Dianne C. Webb.



phosphate, and assayed for the initial rates of Pi uptake as described in Section 2.11. Measurement of  $^{33}\text{Pi}$  uptake by cells was carried out under conditions in which Pi uptake was linear over time. Examples of the experiments used to determine  $K_m^{\text{app}}$  and  $V_{\text{max}}^{\text{app}}$  are given in Figure 3.17. Statistical analysis was used to combine the  $K_m^{\text{app}}$  and  $V_{\text{max}}^{\text{app}}$  parameters from individual experiments to produce the values listed in Table 3.3.

The PitA protein has a  $K_m^{\text{app}}$  about 14 fold lower than PitB on pAN656 in whole cells. While the  $V_{\text{max}}^{\text{app}}$  for PitA was relatively stable, PitB  $V_{\text{max}}^{\text{app}}$  values were variable, as indicated by the large error values. Magnesium concentration and pH were altered to allow comparison with kinetic parameters measured by other researchers. While the  $K_m^{\text{app}}$  values remained similar under these conditions, the  $V_{\text{max}}^{\text{app}}$  was lower for both PitA and PitB when the magnesium concentration was increased from 1.8mM to 10mM and the pH was increased from pH6.6 to pH7.0 (Table 3.3).

### **3.8 Discussion**

*pitA* and *pitB* are individual genes which code for Pi transporters that function in the Pi auxotrophic strain AN3066 of *Escherichia coli*. They have differing kinetic parameters, with the plasmid-borne PitA  $K_m^{\text{app}}$  of 2 $\mu\text{M}$  being approximately 14 fold lower than the PitB  $K_m^{\text{app}}$  of 28 $\mu\text{M}$ . Plasmid-borne PitB also has a lower  $V_{\text{max}}^{\text{app}}$  in these whole cell assays, indicating that this protein maybe less active in Pi transport than PitA. Translation of both *pit* genes occurs from the beginning of each 1500 base ORF. Several topology prediction programs use the deduced amino acid sequences of PitA and PitB to propose that they both have a secondary structure of 10 transmembrane  $\alpha$ -helices connected by hydrophilic loops, with most amino acid variability occurring in a few putative hydrophilic loops. The differing  $K_m^{\text{app}}$  values and the 81% identity in the deduced amino acid sequences means that PitA and PitB form a good comparative system where differences in mechanism, kinetic parameters and amino acid sequence may be systematically explored.

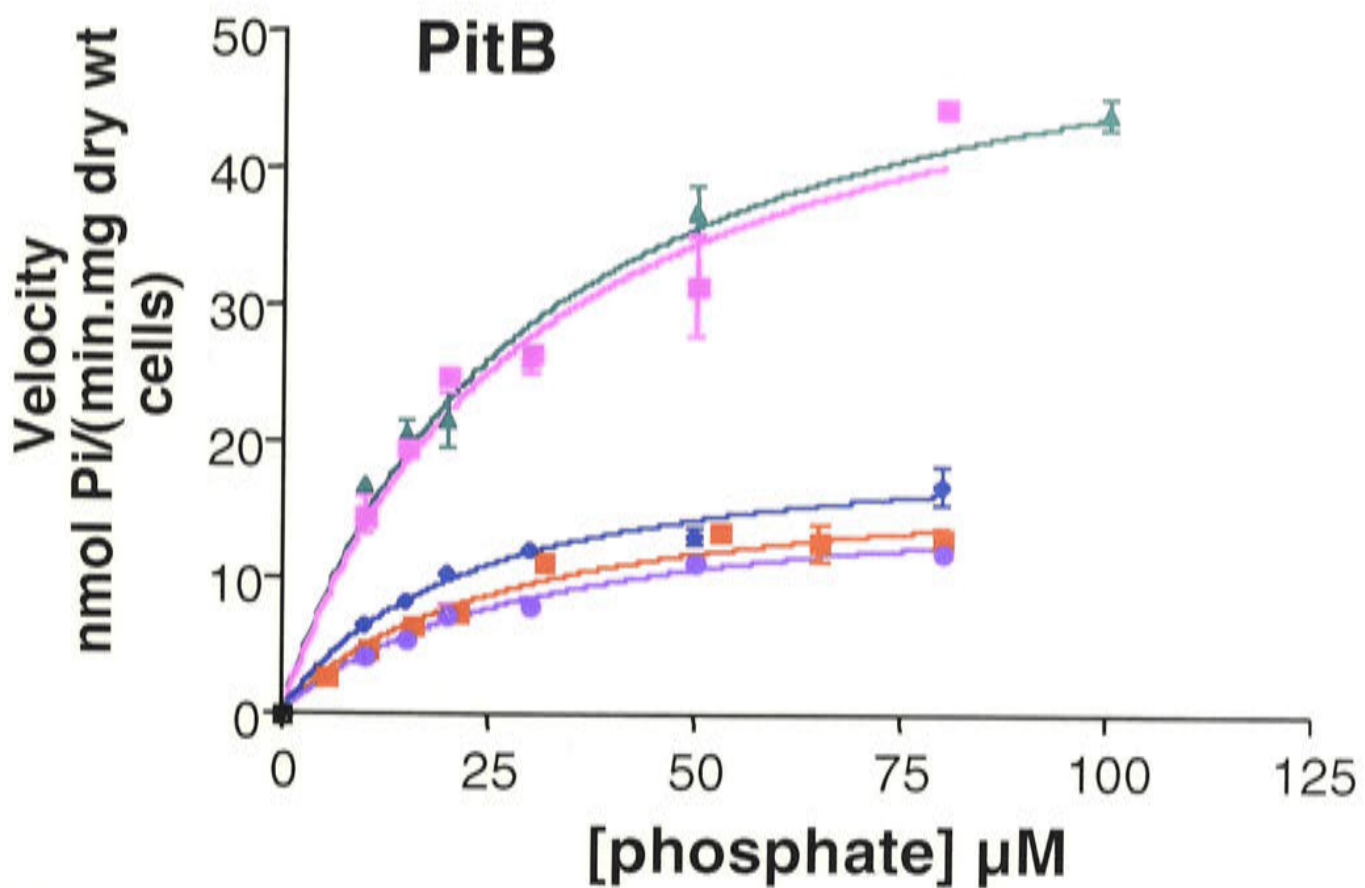
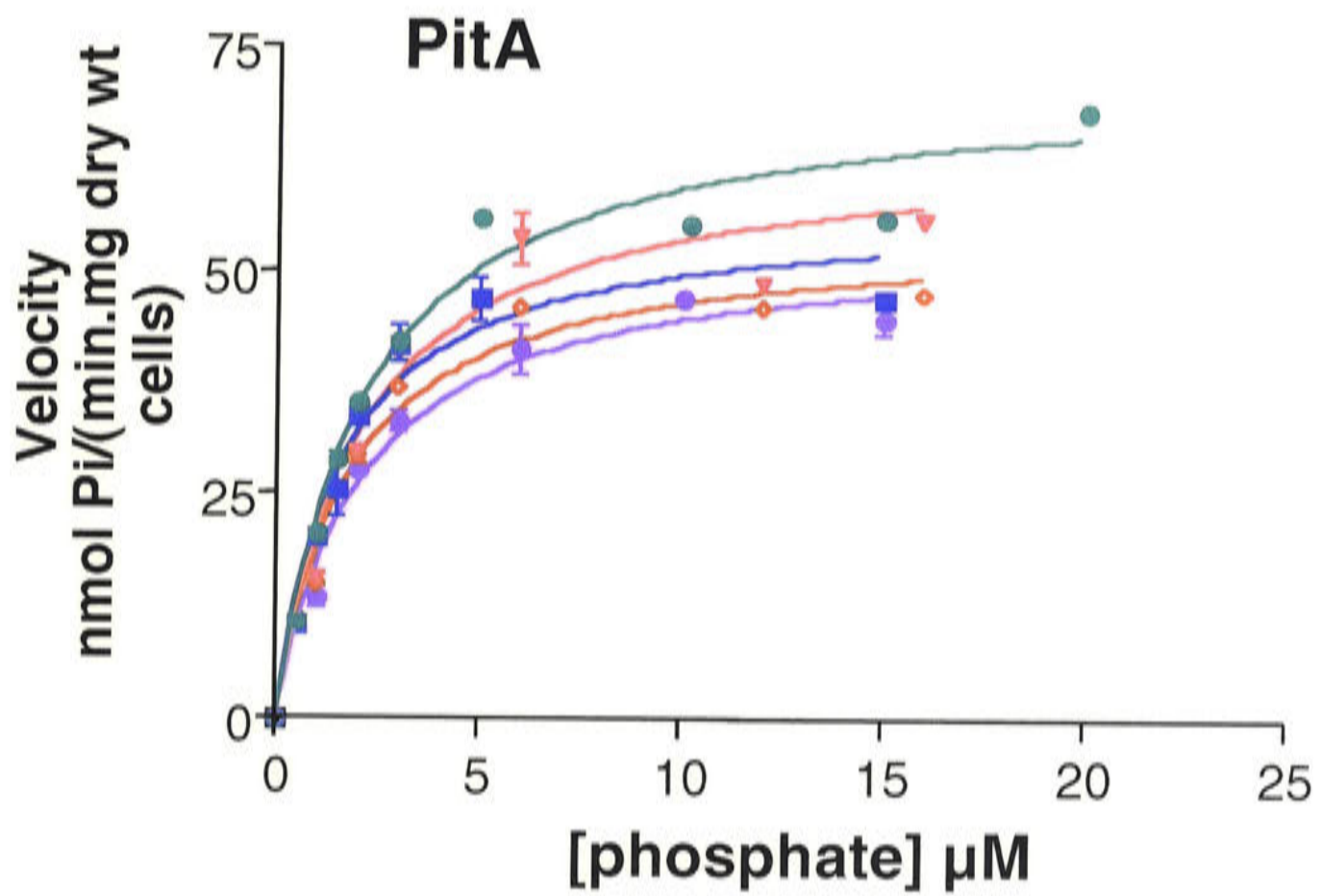
The presence of a second Pit transporter reveals that Pi uptake activity previously attributed to a single Pit transporter may be a combination of distinct PitA and PitB



### Figure 3.17

#### Experiments used in the determination of $K_m^{app}$ and $V_m^{app}$ for PitA and PitB.

Shown below are the individual experiments used to determine the kinetic parameters listed in Table 3.3, using the conditions of pH6.6 and 1.8mM magnesium. Error bars represent the standard error of the mean (SEM). The initial rates were calculated using one to three (usually three) individual determinations, taken at 15 and 25 sec intervals. The kinetic parameters were calculated by nonlinear regression using the Graphpad Prism program.



**TABLE 3.3:** Kinetic parameters for PitA and PitB

Plasmid <sup>a</sup>	Phenotype	Assay conditions		$K_m^{\text{app}}$ <sup>b</sup> $\mu\text{M}$	$V_{\text{max}}^{\text{app}}$ <sup>b</sup> $\text{nmol Pi min}^{-1} \text{mg dry weight}^{-1}$
		pH	[Mg] mM		
pAN686	<i>pitA</i> <sup>+</sup>	6.6	1.8	1.9±0.1 (n=5)	58±4 (n=5)
pAN686	<i>pitA</i> <sup>+</sup>	7.0	10.0	1.9	39
pAN656	<i>pitB</i> <sup>+</sup>	6.6	1.8	28.6±1.4 (n=5)	33±9 (n=5)
pAN656	<i>pitB</i> <sup>+</sup>	7.0	10.0	28.1±1.0 (n=3)	17±3 (n=3)

<sup>a</sup> host strain is AN3066 for all experiments

<sup>b</sup> SEM

activities, or may even be the result of a heteromeric transporter and this will be examined in the next two chapters. A screening of 34 completely sequenced genomes revealed that several other bacteria, including *Pseudomonas aeruginosa*, contain more than one *pitA* homologue (99), and multiple Pi transporters with overlapping activities have been reported in *Saccharomyces cerevisiae* (169) and *Neurospora crassa* (274).

The nucleotide sequences and chromosomal locations of *pitA* and *pitB* were published for *E. coli* K-12 by the *E. coli* genome sequencing project during the course of this work. These nucleotide sequences proved to be identical to those determined in this laboratory (89), with the genomic locations being 78.51 min for *pitA* (245) and 67.44 min for *pitB* (18).

While the 6xHis affinity tag has been successfully used to isolate other membrane proteins (151) and is reported to rarely interfere with protein structure or function, the Pi auxotrophic phenotype of pAN918 in AN3066 was clearly attributable to the 6xHis tag. However other *pitA::6xHis* plasmid constructs procured Pi uptake. The reasons for this variability are unclear. The poor growth of some strains may be due to toxicity effects caused by the over-expression of a membrane protein. Harry Rosenberg's group, which produced a strain over-expressing wild type PitA (59), found these *E. coli* cells had highly invaginated cytoplasmic membranes, as shown by electron micrographs (personal communication, G. B. Cox). Attempted over-expression and purification of PitA6xHis protein from plasmid constructs which showed some Pi uptake activity produced no obvious band of protein by Coomassie stained SDS-PAGE analysis. This protein purification work, which was placed on hold until PitA6xHis could be specifically identified, was never resumed as the regulation of *pitB* became more central to this thesis (Chapter five). The production of the polyclonal PitA and PitB antipeptide antibodies (Sections 4.3.3, 5.4.2) and the commercial availability of Penta.His or Tetra.His antibodies (QIAGEN) now allows for identification of the PitA, PitA6xHis and PitB proteins. This, combined with subsequent refinements in the 6xHis technique (QIAGEN), means that growth conditions and protein isolation can now be optimised should protein purification become a goal in the future.



Chapter four

Identification of the

*E. coli* K-10

---

# ***Chapter four***

***Identification of the *E. coli* K-10***

***Pit mutation***

---

## Chapter four

# Identification of the *E. coli* K-10 Pit mutation

### 4.1 Introduction

*E. coli* relies on the Pst and Pit systems for the uptake of inorganic phosphate (Pi). Another two transporters accept Pi as a low affinity analogue for either glycerol-3-phosphate (*glpT*) (91) or glucose-6-phosphate (*uhpT*) (199, 295) but in the absence of Pit and Pst activity these latter two systems cannot support cell growth when cells are supplied with Pi alone (249). When the Pst system of *E. coli* strain K-10 is inactivated, cells cannot utilise Pi for cell growth (289), suggesting that K-10 has a mutation that affects Pit activity. Conjugational mapping placed this defect at about minute 77 on the *E. coli* genome (249). The *E. coli* genome sequencing project has located *pitA* at minute 78.51 (245), indicating this K-10 lesion is likely to be in the *pitA* gene. The mutation has yet to be defined.

The discovery of *pitB*, a second Pit Pi transporter located at 67.44 minutes on the *E. coli* genome (18) means that this gene should also be examined for mutations if the true nature of the K-10 Pit mutant is to be understood.

It is important that the K-10 Pit mutant is fully defined as all assays to characterise *pitA* and *pitB* have been carried out in strain AN3066, a Pi auxotroph derived from K-10. While AN3066 has negligible Pi transport, this does not guarantee that there is no expression of wild type *pitA* or *pitB*, or that the addition of plasmid-borne *pit* will not alter genomic *pit* expression. There are a variety of possible scenarios. Mutations within a post-transcriptional regulatory gene or protein may inhibit PitA and/or PitB protein. If only hetero-dimers are functional, expression of either *pitA* or *pitB* would not provide function. Mutations within regulators or regulatory regions of either genomic *pit* gene may be overcome by multicopy *pit* plasmids. Thus, *pitA* or *pitB* introduced on a plasmid has the potential to interact with the genomic *pit* system in a number of unintended ways.

This chapter identifies the mutation/s in the K-10 *pit* genes and analyses them with respect to Pi uptake activity and protein assembly.

## **4.2 Identification of the *E. coli* K - 1 0 *pit* mutation/s**

Strain AN3066 was the standard background strain used for measuring the kinetic parameters of plasmid-borne *pitA* or *pitB* as described in Chapter three. It is unable to grow on minimal media supplemented with Pi as the sole source of phosphate, therefore has no Pi transport. AN3066 is a derivative of strain K-10, which has been modified by a deletion in the *pstC* gene ( $\Delta pstC345$ ), producing a nonfunctional Pst system, and the introduction of a *recA* mutation (*srl::Tn10 recA*) to hinder genetic recombination events (284).

The *pitA* and *pitB* nucleotide sequences were amplified from the genomic DNA of strain AN3066 by PCR, using *Pfu* DNA polymerase in order to minimise the introduction of errors (156). Oligonucleotide primers complementary to the 5'- and 3'-regions flanking the open reading frames (ORFs) of *pitA* and *pitB*, including approximately 40 upstream nucleotides and ending just after the translation termination codons, were used for this amplification. *Bam*HI sites were included at the distal ends of each primer (Figure 4.1, Table 2.5). The 1.5kb *pitA* and *pitB* PCR products were confirmed by restriction analysis, purified, digested with *Bam*HI and subcloned into the *Bam*HI site of dephosphorylated M13mp18. RF DNA was screened for the correct size and orientation of the insert by restriction analysis. ssDNA was purified from several phage isolates, and one isolate was fully sequenced. Manual sequencing was carried out over regions where the cycle sequencing results were ambiguous.

DNA sequencing revealed that the *pitB* ORF is wild type. To investigate whether regulatory sequences were mutated a longer *pitB*<sup>+</sup> PCR product including 319 nucleotides upstream of the ORF was also prepared by PCR (Figure 4.1B, primers listed in Table 2.5). A DNA fragment of 1.95kb was amplified and verified as *pitB* DNA by restriction analysis. This DNA was purified and subcloned into M13mp18,



## Figure 4.1

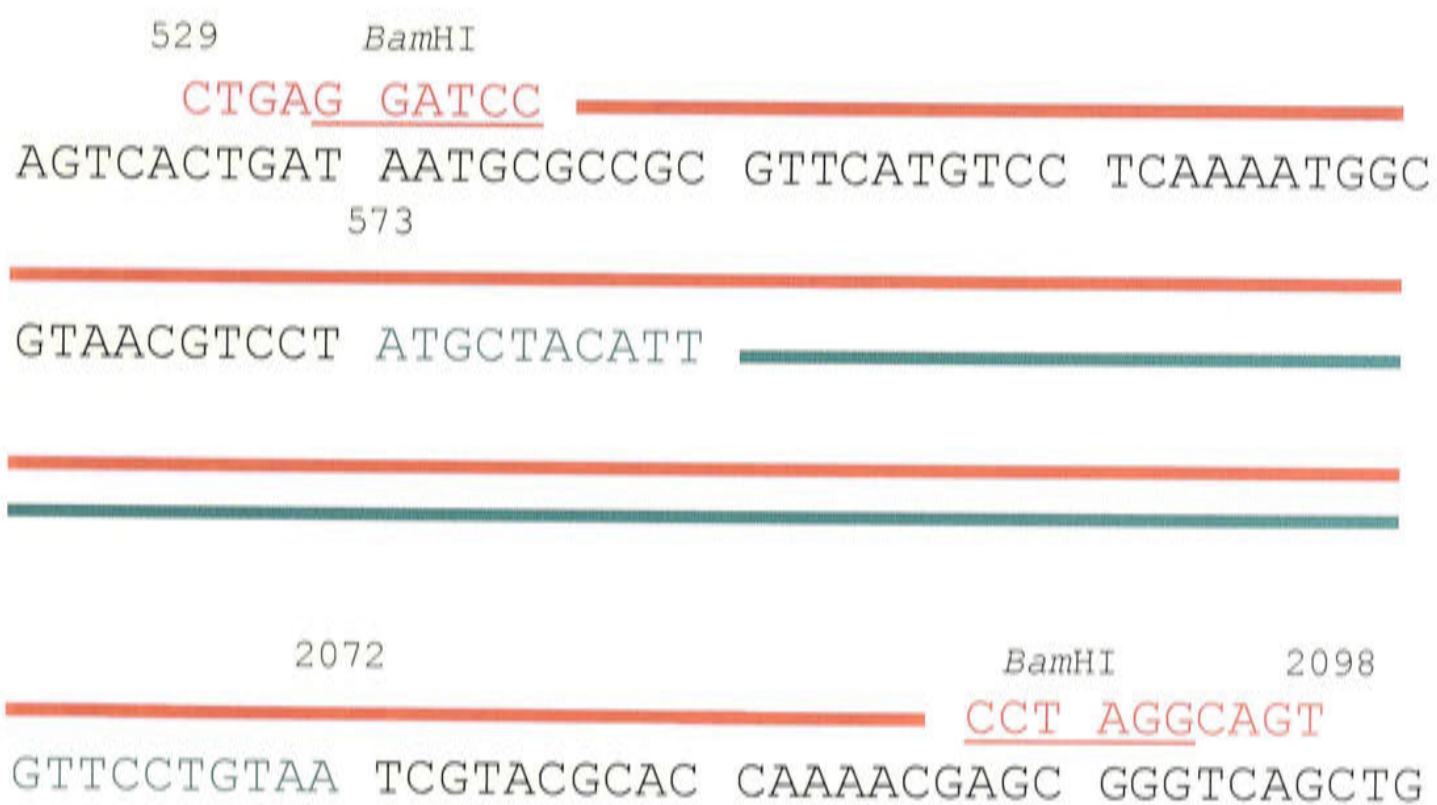
PCR amplification of the putative *pitA* and *pitB* open reading frames.

### A *pitA* PCR amplification.

The largest open reading frame (ORF) on the DNA fragment containing *pitA* activity was amplified by PCR using the oligonucleotides 95-87/95-88 (see Table 2.5 for details). The nucleotides amplified at the beginning and end of this open reading frame are shown below.

CTG — AGT PCR amplification product.

ATG — TAA Open reading frame.



## B *pitB* PCR amplification.

The largest open reading frame (ORF) on the DNA fragment containing *pitB* activity was amplified by PCR using the oligonucleotides 95-89/95-90 (see Table 2.5 for details). The nucleotides amplified at the beginning and end of this open reading frame are shown below. The larger *pitB* PCR product containing 319 nucleotides upstream of the ORF start, using oligonucleotides 97-73/97-74 is also shown. The putative stem loop sequence is underlined.

CTG — AGT PCR amplification product for the *pitB* ORF.

ATG — TAA Open reading frame.

AGA — CAC PCR amplification product for the longer *pitB* fragment (ORF plus 319 upstream nucleotides).

1075 BamHI  
AGAGGGATCC —————  
AAGAGAGTGATATTGAACCGTTAATTGTGGTGAAAAAGTAATAC  
—————  
ATGTTGTAGCAACCCCGTACTGATTATGAAGAATAATCAGTACG  
—————  
GGGATTAATCAAATATTCAATATTTATATATTCCAATATTTAT  
—————  
TCATTTCATATGTGAATATATTAACCAGTGGAATACCTGTGTTT  
—————  
TGATTATTTCTAAAGGTTTTGAATGAATGCTTATTGTCTGATAC  
—————  
ACAAGAAATAACACTCTTTTTATCGTTAAAAAATGATATTTAC  
1354 BamHI  
GTCAGGATCC —————  
TTTGCCCATGCCGTTAAAGTCACTGATAATGCGTCCGTTTCGTAA  
1404  
—————  
ATTCAAATGGCGTAATCTAATATATGCTAAAT —————  
2903 BamHI 2930  
CCTAGGCCGC  
TTGATTTAACCTTCCTGAAAATGCCCGGTCCTGGCGATCGGGCA  
—————  
TTTCCATTTTTGACTAGTGATAACCACGCGCGGTCATAAAATCC  
BamHI 3022  
CCTAGGTCAC  
GTAATCGCTTTTTTCTGCATCAACCAACACCTGTTCCAGTGGCTG



screened for the correct orientation and insert size and the M13-40 primer was used to cycle sequence three isolates over the upstream region of *pitB* and into the *pitB* ORF. These 319 upstream nucleotides were wild type in all isolates. Therefore *pitB*, including both the promoter region and ORF, appears to be wild type in AN3066 even though this strain exhibits no Pi transport activity.

The *pitA* sequence contained a single point mutation of G to A at nucleotide 658 in the ORF of the published sequence. This was confirmed by sequencing three independent PCR isolates over this region, all of which contained the G to A mutation (Figure 4.2). The point mutation creates an amino acid change of glycine to aspartic acid (G220D) at residue 220 in the coding sequence of PitA (*pitA1*), and this amino acid substitution is located within a putative transmembrane  $\alpha$ -helix of the PitA membrane protein, using the proposed topological model of PitA (Figure 3.15).

### **4.3 Functional analysis of *pitA1* from K-10 strain AN3066**

#### **4.3.1 *pitA1* Pi transport**

The *pitA1* gene from AN3066, which contains the G to A point mutation, was then tested for its ability to produce a functional Pi transporter. The PCR isolate that had been fully sequenced was subcloned from the M13 RF into the *Bam*HI site of pBR322 (pAN1243) and then transformed into AN3066. The desired orientation was selected using restriction endonuclease analysis (Figure 4.3). This new strain, AN3937, was unable to grow on Pi media, indicating that the amplified *pitA1* gene from AN3066 was nonfunctional.

#### **4.3.2 Preparation of *pitA(G220D)* by site-directed mutagenesis**

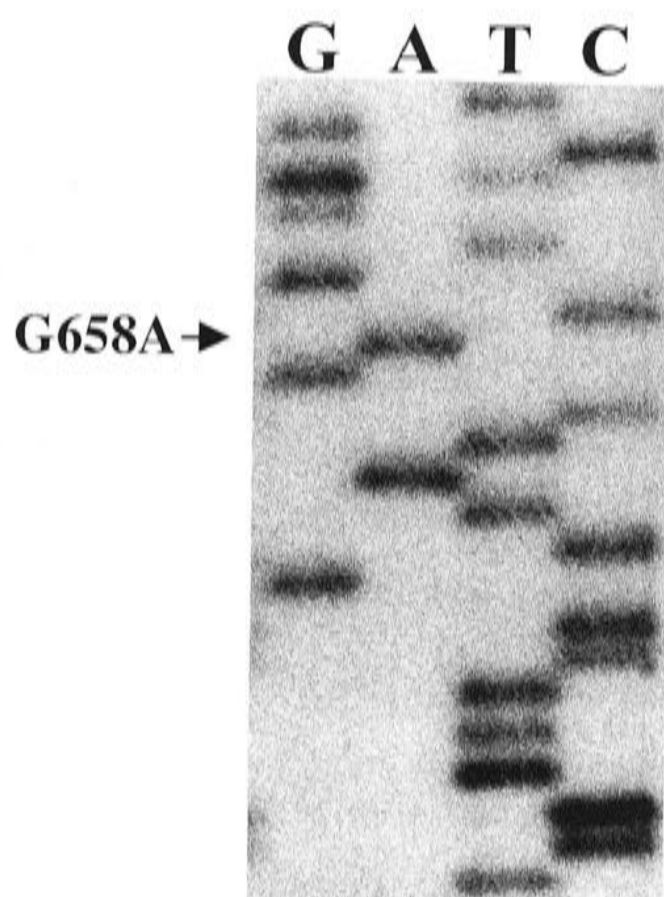
The lack of Pit function in strain K-10 is quite stable. Point mutations are often unstable due to their ability to revert to wild type at a moderately high frequency when placed under selection pressures. Therefore it seemed reasonable to check that this single base change could fully account for the stable loss of Pit function found in K-10.



## Figure 4.2

### Sequencing of K-10 plasmid-borne *pitA* mutation.

Manual sequencing gel autoradiograph of PCR amplified *pitA* from Pi auxotrophic strain AN3066, which contains a G658A point mutation. This mutation was present in the three independent PCR isolates that were sequenced.



### Sequence

(reading from bottom to top)

<i>pitA</i> mutant	TCCTTTCCGCTATCG <sup>+</sup> ACGTGGCGT
wild type <i>pitA</i>	TCCTTTCCGCTATCG <sup>+</sup> GCGTGGCGT

Plasmid pAN686 has *E. coli* K-12 wild type *pitA* sequence over its full length, from 573 nucleotides upstream of the ORF to 199 nucleotides downstream of the ORF (245). Double stranded site-directed mutagenesis was used to recreate the G to A point mutation in this wild type *pitA* gene, creating plasmid pAN1244 (Figure 4.3, oligonucleotide 95-101 in Table 2.5). The mutation was confirmed by sequencing and pAN1244 was then transformed into strain AN3066 so that Pi uptake could be assayed. This new strain, AN3938, did not grow on Pi media, indicating that this single point mutation can disrupt the function of the wild type *pitA* gene. As this gene is wild type for at least 573 nucleotides upstream of the *pitA* ORF, it is unlikely that another mutation in the regulatory region is needed to reproduce the K-10 *pitA1* phenotype. However, *pitA1* and *pitA*(G220D) will be further characterised by comparing their levels of protein expression by Western blot using PitA polyclonal antipeptide antibody.

### **4.3.3 Analysis of PitA expression using polyclonal PitA antipeptide antisera**

The nonfunctional PitA protein has glycine 220, a small hydrophobic amino acid, replaced by aspartic acid which has a medium sized negatively charged side chain. Amino acid substitutions may affect the assembly of the protein, and/or its enzyme activity. To determine if this *pitA1*(G220D) mutation affected the assembly of PitA, the levels of PitA protein in the membrane fraction of whole cells containing either wild type or mutant PitA were compared by Western blotting.

Polyclonal antipeptide PitA antibody was used to visualise the PitA protein. This was prepared by immunising rabbits with a multiple antigen peptide system (MAP) conjugate consisting of the peptide ARIHLTPAEREKKD-C attached to a polylysine core via a C-terminal cysteine (Section 2.13.2). This PitA peptide forms part of an extramembranous loop in the putative folded structure (Figure 3.15) and is in a region of variability between the PitA and PitB sequences. After initial testing of the sera by an ELISA assay against the PitA peptide, the specificity of the antipeptide PitA sera was investigated by Western blot. Sera was applied to the membrane fractions of the AN3066 background strain, which has no Pi uptake, containing either a vector control or wild type *pitA* on a plasmid. Polyclonal antibodies in this sera bound strongly to a

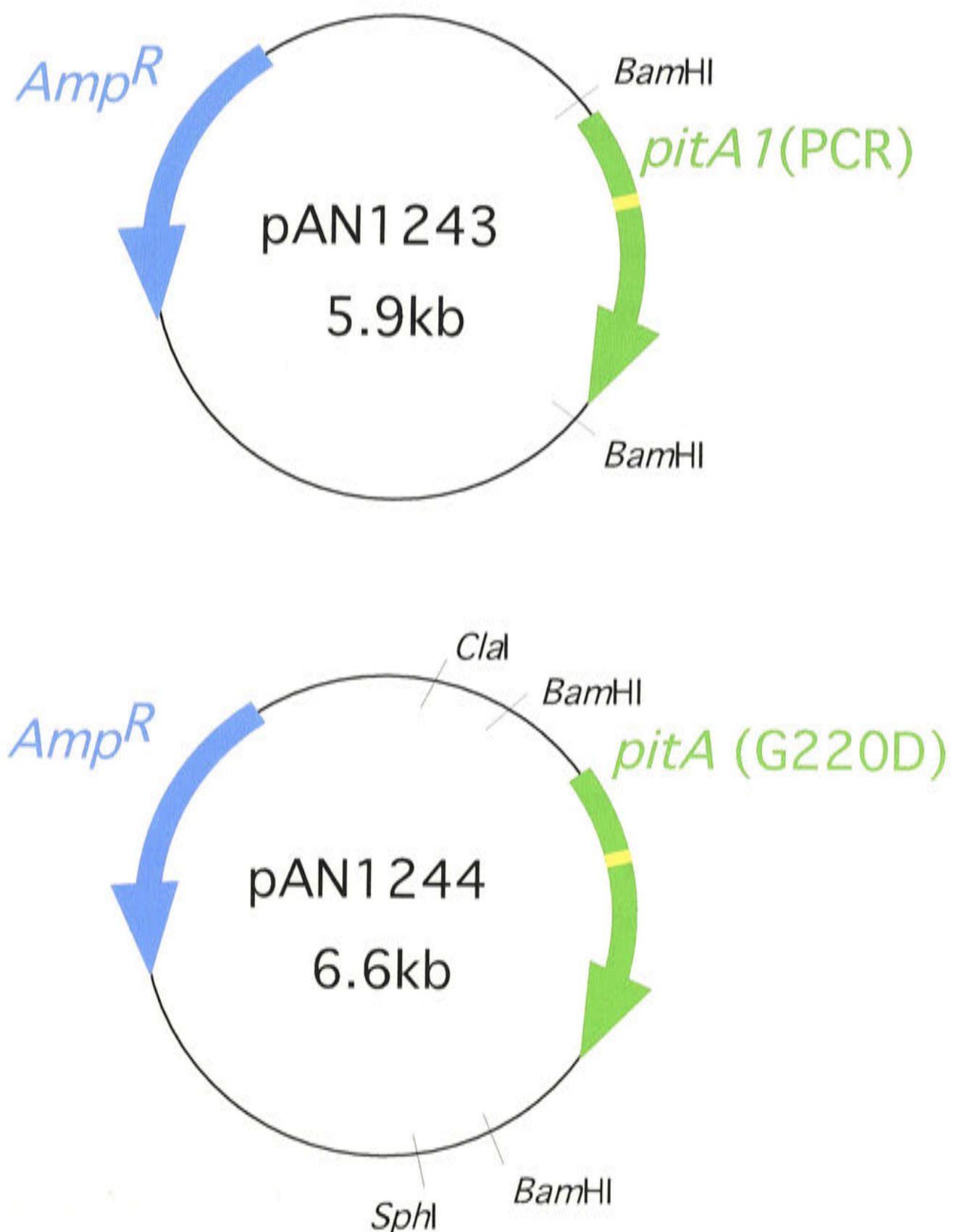
## Figure 4.3

### Plasmid diagrams for pAN1243 and pAN1244.

*pitA* (PCR) or *pitA*(G220D) on vector pBR322.

The *pitA* PCR amplification product from AN3066 (using primers 95-87/95-88, See Table 2.5 for details) was digested with *Bam*HI and ligated into the *Bam*HI site of pBR322. The desired orientation was selected by restriction endonuclease analysis, forming pAN1243.

The *pitA*(G220D) mutation was created by using site-directed mutagenesis to make a G658A nucleotide change in the open reading frame of wild type *pitA* on plasmid pAN686, creating pAN1244.





protein at approximately 48kDa (from a 10% manually poured polyacrylamide gel with a Tris buffer system) (Figure 4.4). There were no equivalent bands in Western blots exposed to pre-immune sera. The deduced molecular weight of the PitA protein is 53kDa. The smaller than expected apparent molecular weight is not unusual for membrane proteins, which often undergo anomalously fast migration during electrophoresis. This may be due to the reformation of some of the  $\alpha$ -helical structure during electrophoresis, or by the binding of excess molecules of SDS by the highly hydrophobic regions of the protein. No further purification of the sera was carried out as this Western blot showed minimal binding of other membrane proteins by the polyclonal antibody.

Possible cross-reactivity with the PitB protein was investigated by a Western blot of the sera against the membrane fractions of AN3066 containing either vector (pBR322), wild type *pitA* (pAN920) or the long *pitB* plasmid (pAN656) (Figure 4.4). The membrane fraction of the pAN920/AN3066 strain had a strongly labeled band at approximately 48kDa representing the PitA protein. The vector and long *pitB* membrane fractions had weakly labeled bands in the same position. While this does not completely rule out any cross reactivity of the PitA antipeptide antibody with PitB protein, it does indicate that under the experimental conditions used here cross reactivity is negligible.

#### ***4.3.4 Comparison of PitA assembly in strains containing either wild type pitA or pitA1***

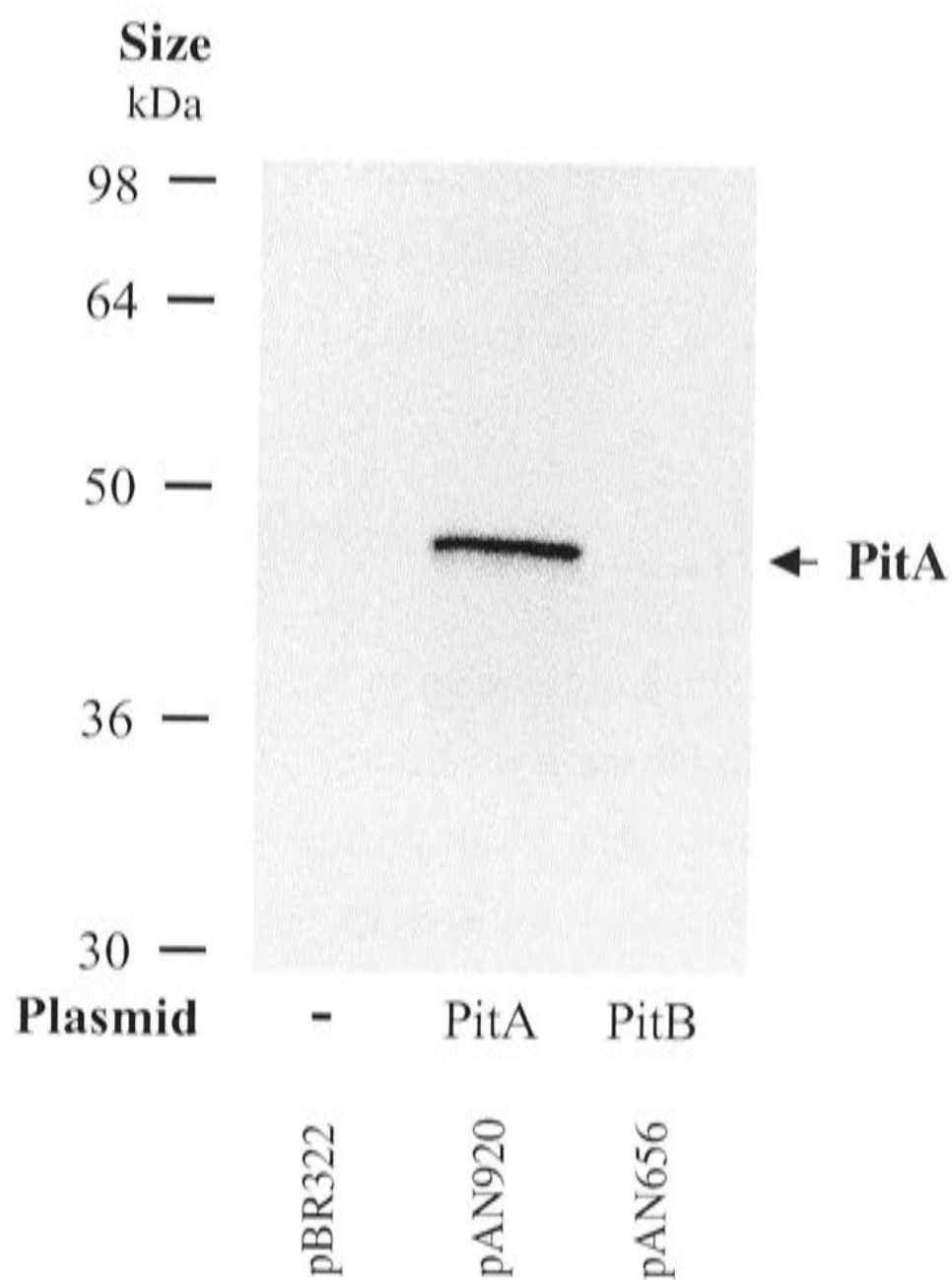
Plasmids containing wild type *pitA*, *pitA1* or *pitA*(G220D) were transformed into the Pi transport negative background strain AN3066 (*pitA1 pst*) so the levels of PitA protein assembly conferred by each plasmid could be compared. As the cytoplasmic membrane of AN3066 may contain a nonfunctional PitA protein, an *E. coli* strain with wild type *pitA* on the genome plus the AN3066 background strain were also included in this comparison.

The membrane fractions of whole cells were isolated from strains grown in rich media via an ammonium sulfate precipitation procedure (Section 2.7.6). The protein concentrations were determined and membrane fractions were diluted into 2mg/ml

## Figure 4.4

### Western blot analysis of polyclonal PitA antipeptide antisera against PitA and PitB protein.

The membrane fractions of the Pi auxotrophic strain AN3066 containing plasmids pBR322 (vector), pAN920 (PitA) or pAN656 (PitB) were solubilised and separated by PAGE on a 10% polyacrylamide gel (Tris buffer system) and transferred onto PVDF membrane. Protein was visualised by incubation with polyclonal PitA antipeptide antisera, followed by incubation with alkaline phosphatase conjugated goat anti-rabbit secondary antibody and the application of Western blue alkaline phosphatase substrate.



aliquots and stored at -70°C. Western blots were carried out by denaturing samples diluted to 200µg/ml protein, separating these proteins by polyacrylamide gel electrophoresis (NUPAGE 4-12% gradient gel with MES buffer system), and carrying out semi-dry electroblotting and visualisation using an alkaline phosphatase coupled secondary antibody (Section 2.13.3).

The membrane fraction of the control strain AN3066 plus vector pBR322 (AN3514) contained very little PitA protein (Figure 4.5), while the wild type *pitA* plasmid produces significant levels of PitA (AN3531). AN248, which has wild type PitA on the genome, produced an intermediate level of PitA protein. This level of protein expression was much greater than that produced by genomic *pitA1* from AN3066, but was less than that produced from the wild type *pitA* plasmid. Cells with the *pitA1* plasmid or the site-directed *pitA*(G220D) plasmid had low levels of protein similar to the AN3066 background strain (AN3937, AN3938). Therefore the G220D mutation seems to greatly reduce the insertion of PitA protein into the membrane, and protein levels between the site-directed mutant and the AN3066 *pitA1* mutation appear similar. Although the membrane fractions of these mutant strains did contain a small amount of PitA protein, they have negligible Pi transport and it is unlikely that this mutated membrane protein is functional.

#### **4.4 Discussion**

Strain K-10 is defective in Pit transport, as outlined in the introduction to this chapter. Both the *pitA* and *pitB* genes were examined for potential mutations. While *pitB* had wild type sequence for the ORF and at least 319 upstream nucleotides, DNA sequencing of the *pitA* ORF from the K-10 strain, AN3066, identified a single point mutation of G to A at nucleotide 658 in the published sequence. This causes an amino acid change from glycine to aspartic acid at residue 220 in the PitA protein.

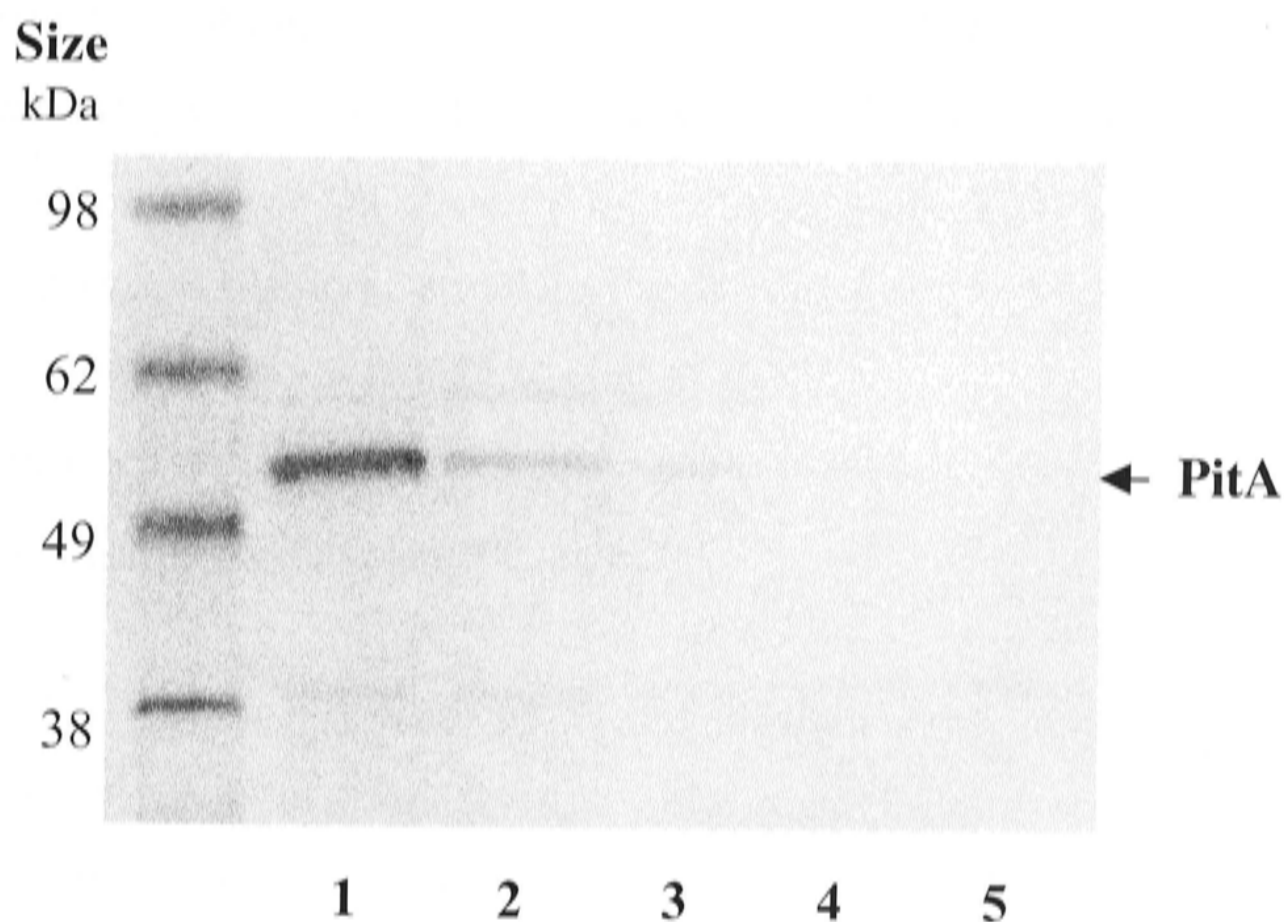
Western blots using polyclonal PitA antipeptide antisera showed that the mutated PitA protein is found in cell membranes at a much lower level than in an equivalent strain containing wild type *pitA*. This suggests that inhibiting the insertion of the PitA protein into the plasma membrane is a major effect of the *pitA1* mutation. However, a small amount of PitA protein is visualised in the membrane fraction of AN3937, which



## Figure 4.5

### Western blot analysis of PitA protein expression by the plasmid-borne K-10 *pit* mutant and *pitA*(G220D) site-directed mutant.

Membrane protein fractions of the wild type strain AN248 and the Pi auxotrophic strain AN3066 containing the plasmids listed below were solubilised and separated by PAGE on a 4-12% gradient polyacrylamide gel (MES buffer system) and transferred onto PVDF membrane. Protein was visualised by incubation with polyclonal PitA antipeptide antibody followed by incubation with alkaline phosphatase conjugated goat anti-rabbit antibody and application of Western blue stabilised alkaline phosphatase substrate.



Lane	Strain	Genome		Plasmid
		genotype	phenotype	
1	AN3531	<i>pitA1</i>	-	<i>pitA</i> <sup>+</sup>
2	AN248	<i>pitA</i> <sup>+</sup>	+	-
3	AN3514	<i>pitA1</i>	-	pBR322
4	AN3937	<i>pitA1</i>	-	<i>pitA1</i> (PCR)
5	AN3938	<i>pitA1</i>	-	<i>pitA</i> (G220D)

contains the *pitA1* mutant plasmid in the *pitA1* AN3066 background strain. This level of protein is only a little less than that found in the membrane fractions of AN248, a strain with wild type *pitA* on the genome, and normal Pit function. Therefore it is likely that the substitution of aspartic acid for glycine 220 affects both the assembly and function of the PitA protein.

Putative topological models of PitA places glycine 220 within transmembrane (TM)  $\alpha$ -helix 6 (Figures 3.15, 6.2). Aspartic acid is a medium sized negatively charged amino acid, while glycine has a small and uncharged side chain. The insertion of charged residues into the hydrophobic environment of the membrane has a high energy cost (62). Thus, the substitution of a negatively charged aspartic acid for glycine 220 may prevent this section of  $\alpha$ -helix from inserting into the membrane, or greatly reduce its stability or efficiency of insertion. The bulk and charge of the aspartic acid side chain may interfere with the  $\alpha$ -helical packing of the TM  $\alpha$ -helix and/or interhelix interactions within the mutant PitA protein. Thus, some PitA may insert in the membrane, but have a disrupted structure that creates a nonfunctional protein. This region of PitA may be very sensitive to change as the G220D substitution is located near several charged residues within putative TM  $\alpha$ -helix 6 that have been found to be important in PitA structure/function (Figure 6.2).

These postulations assume that the G220D point mutation is the only mutation in *pitA1*. The lack of Pit function in strain K-10 is quite stable. Point mutations are often unstable due to their ability to revert to wild type at a moderately high frequency when placed under selection pressures. As the ORF of only one isolate of *pitA1* was fully sequenced, there remained the possibility that other mutations may have been missed, either within the ORF or in the promoter region of the gene. Therefore it seemed prudent to determine if the G220D substitution alone could create the changes in function and protein expression observed in the *pitA1* mutant. A site-directed mutation changing nucleotide G658 to A658 in wild type *pitA*, showed that this *pitA*(G220D) gene behaved identically to the *pitA1* mutant, being unable to support cell growth on Pi media, and having similarly low levels of protein inserted in the cell membranes. This gene had wild type sequence for at least 573 nucleotides upstream of the *pitA*(G220D)

ORF, so the need for additional mutations in the *pitA* regulatory region seems unlikely. Subsequently, another group of researchers have identified the same K-10 *pitA* point mutation (99).

The *pitB* gene from AN3066 was sequenced and is identical to the functional *pitB* gene isolated by cosmid cloning, and to the *pitB* gene sequenced as part of the *E. coli* K-12 genome sequencing project (18). The open reading frame and 319 upstream nucleotides were sequenced, which should include the promoter and most regions, which may potentially interact with regulatory proteins. However, this *pitB* gene does not function as a Pi transporter on AN3066. *pitB* was cloned due to the ability of the 3kb fragment to complement for Pi uptake when located on vector pHc79 (pCE27) in strain AN3066 (89) and Pi uptake assays confirmed that a functional Pi transporter was encoded on this fragment. The only known differences between AN3066 and pCE27/AN3066 are the location of the *pitB* gene on either a plasmid or the genome and the number of *pitB* copies. Therefore genomic *pitB* may be under regulation that is disrupted when the gene is placed on a plasmid. The gene copy number may be most important, as Hoffer *et al* (99) identified a *pitA* *pst* mutant able to transport Pi which had a DNA rearrangement creating multiple copies of the wild type *pitB* gene. Experiments to investigate the regulation of *pitB* are described in the next chapter.

Defects in translation may also be responsible for this lack of PitB activity. These defects may occur as mutations in the either the mRNA or 16S rRNA sequence motifs to decrease complementarity, which reduce the efficiency of initiation of translation, or could occur as defects in peptide transfer, translocation of the protein through the ribosome, the stability of the mRNA or the mRNA/ribosome complex, or termination and removal of the ribosome. It is unlikely that the decrease in PitB Pi uptake activity is caused by the mRNA Shine Dalgarno rRNA binding site. Both *pitA* and *pitB* appear to have poor Shine Dalgarno (SD) sequences, consisting of GCGG (promoter sequences are in Figure 5.9) rather than the more favourable AGGAGG sequence that is found 4 to 7 nucleotides 5' of the initiator AUG of many mRNAs. However, the identity between these *pitA* and *pitB* SD sequences shows that this aspect of translation is not the source of differential activity between the two *pit* genes. Possible post-transcriptional modification defects are discussed in "Section 7.2 Regulation of the Pit system". Comparison of the levels of mRNA transcription (via northern blots) with translation (via western blots, the use of reporter genes or by testing Pi uptake activity) for various *pitA* and *pitB* constructs could be used to determine if the reduced PitB activity is caused by defects in translation.



---

# ***Chapter five***

***Expression and regulation of pitA and  
pitB***

---

## Chapter five

### *Expression and regulation of pitA and pitB*

#### **5.1 Introduction**

The Pit system was previously characterised as the product of a single locus which is constitutively expressed (59, 211, 212, 249, 289). However, evidence presented in the previous two chapters indicates that *E. coli*, has two *pit* genes, and at least one of these may be regulated. Therefore both *pitB* and *pitA* should be analysed for any potential interactions or regulation. Previously presented kinetic data which could represent a combination of PitA and PitB activities may, upon re-examination, provide clues as to what environmental or genetic conditions influence Pit regulation.

This analysis should be open to the possible involvement of the Pst system and the genes that regulate it. The Pst operon is induced by the *phoBR* two component regulatory system when external Pi concentrations become limiting (0.1 to 1mM) (291) (2, 254, 259). The *phoB-phoR* operon controls the *pho* regulon, activating over 100 genes involved in phosphate metabolism and assimilation (271, 280). There is also evidence to suggest that several members of the *pho* regulon are repressed under these conditions (271, 291). The process of Pi-dependent activation of the *pho* regulon is described in Section 1.4. Inhibition of the *pho* regulon occurs when external Pi conditions rise above 0.1-1mM. The Pst system, which is thought to play a role in the detection of the external Pi concentration, interacts with PhoU and PhoR to cause the dephosphorylation of PhoB and subsequent inhibition of the *pho* regulon (Figure 1.2). The majority of mutations in PhoU or any of the four Pst proteins result in high constitutive *pho* regulon gene expression (181, 282). This may be relevant to the regulation of PitA or PitB as the Pst system needs to be inactivated to measure the kinetics of the *E. coli* Pit system. Pst activity can be inhibited by growing the cells in media containing high concentrations of Pi, using membrane vesicles which removes a binding protein that is essential for Pst activity, or by the mutation of Pst genes to create a nonfunctional Pst system. As any change to the Pst system has the potential to affect *pho* regulon activity, previous experiments isolating Pit activity may have unintentionally created environments with very different Pi regulatory conditions.

It is possible that *pitB* may be repressed by processes which activate most members of the *pho* regulon. *Rhizobium meliloti* contains a Pit-like Pi transporter that is repressed under conditions of Pi limitation. This repression is relieved in *phoB* mutants of *R. meliloti* (8). All Pi uptake assays to measure Pit activity undertaken in Chapter 3 have been carried out in the background strain AN3066, which contains the  $\Delta pstC345$  Pst system deletion. This deletion prevents repression of the *pho* regulon, so it is constitutively expressed at all Pi concentrations (44) rather than being expressed only when Pi concentrations are very low. If *pitB* is repressed when the *E. coli pho* regulon is activated, a constitutively expressed *pho* regulon could repress PitB activity under all the conditions we have used to assay Pit transport. This could explain the lack of activity from genomic *pitB* that is described in Chapter four.

With this frame of reference the work of past researchers was examined for any evidence of *pitB* activity. In 1974 Willsky *et al* (292) described a *pitA1* strain grown and assayed at 1mM Pi which had a  $K_m^{app}$  of 0.4 $\mu$ M Pi, indicative of Pst system activity. When a *pstS72* mutation was introduced the rate of uptake fell and the  $K_m^{app}$  increased to 18 $\mu$ M, which is similar to the  $K_m^{app}$  values we have obtained for *pitB*. The *pstS72* mutation in the above strain reduces the activity of the Pst system but does not eliminate all Pi uptake (211). If *pstS72* has no effect on the *pho* regulon (which is certainly possible, as the regulatory and transport roles of the Pst system could be uncoupled by specific amino acid substitutions in *pstC* or *pstA* (43, 44)) *pitB* could be expressed in 1mM Pi, which is within the 0.1-1mM Pi concentration range where the *pho* regulon starts being repressed. In K-10, which has a wild type Pst system and the *pitA1* mutation, a component of the Pi uptake attributed to the Pst system at high Pi concentrations could also represent PitB activity (211). *phoBR* controls the *pho* regulon, so its influence on *pitB* activity will be investigated.

The presence of a wild type *pitB* gene in AN3066, which has been the background strain used in all Pi assays reveals several potential complications. Pi transport from *pitA* plasmids has been measured in the presence of a genomic wild type *pitB* gene. Thus Pi uptake activity previously attributed to *pitA* could conceivably result from an interaction between the plasmid's PitA protein and PitB protein expressed from the



genome. It is not known if genomic wild type *pitA* can function as a Pi transporter in the absence of the wild type *pitB* gene. AN3066 is therefore not an appropriate background strain and a new background strain which lacks *pst*, *pitA* and *pitB* genes is required to replace AN3066.

This chapter examines the ability of PitA and PitB to function as independent Pi transporters. The expression and regulation of *pitB* is analysed, and the possible involvement of the *pho* regulon is investigated.

## **5.2 Preparation of a Pi transport triple mutant**

Previous experiments have shown that the *pitB* gene of AN3066 has wild type sequence (Section 4.2). Although no conditions have yet been found where the genomic wild type *pitB* gene on AN3066 is functional, insertional inactivation of the *pitB* gene removes all chances of ambiguity in future experiments.

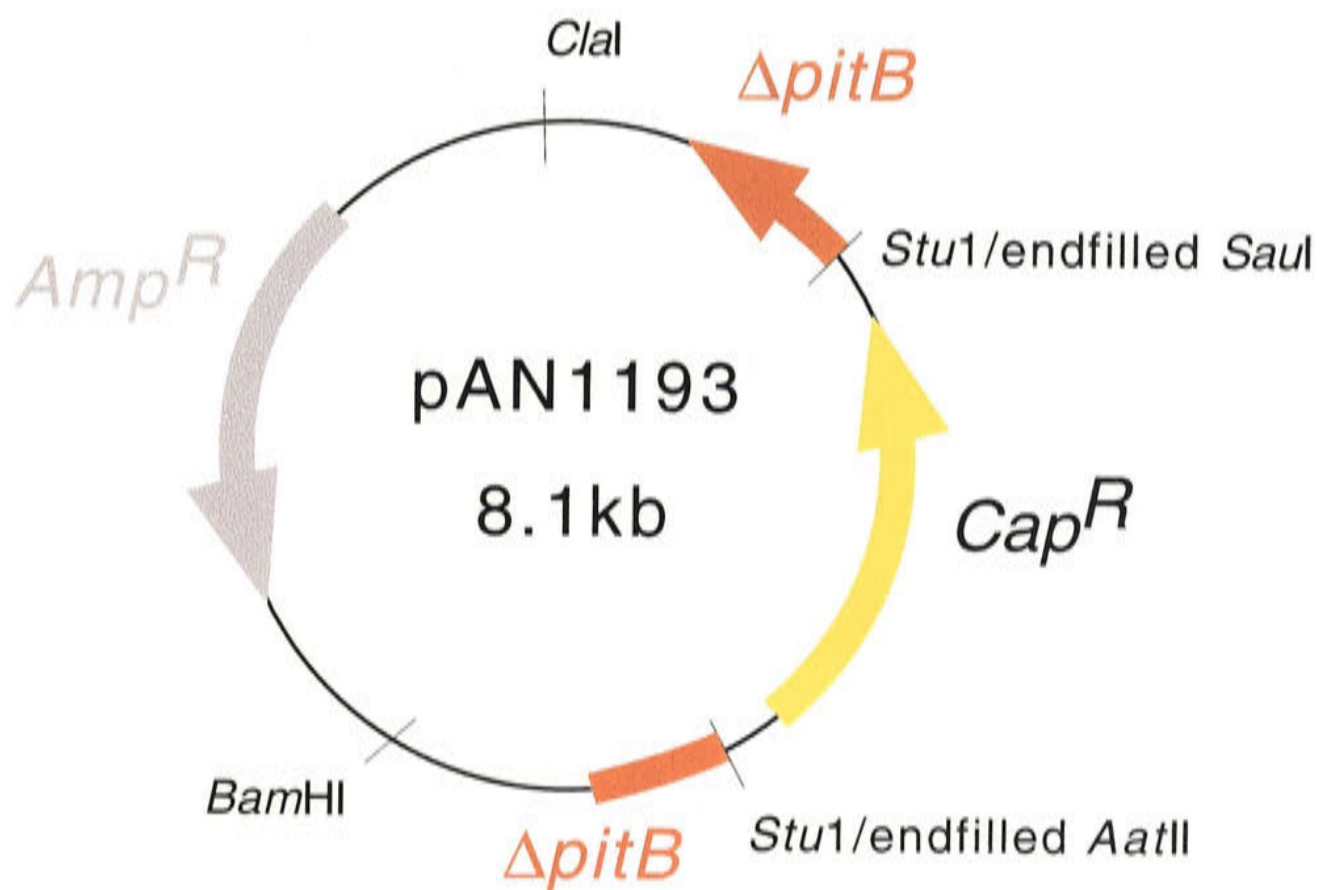
The *pitB* gene was insertional inactivated by ligating the chloramphenicol resistance gene from pBR328 (cut with restriction endonucleases *AatII* and *SauI*, then endfilled with T4 DNA polymerase) into the *StuI* site of pAN656, creating pAN1193 (Figure 5.1). This placed the chloramphenicol resistance gene 514 nucleotides into the open reading frame of the *pitB* gene (*pitB*::Cat<sup>r</sup>). Recombination with genomic *E. coli* DNA was carried out by transforming pAN1193 into competent JC7623 cells. This strain contains *recBC sbcBC* mutations which allow double crossover events to occur between covalently closed circular plasmid DNA and the genome (Section 2.6.1) (189). Recombinant JC7623 cells were selected by plating the transformation mix onto rich media containing the antibiotics chloramphenicol (selecting *pitB*::Cat<sup>r</sup>) and streptomycin (selecting the JC7623 background). The absence of plasmid pAN1193 was verified by screening transformants for sensitivity to ampicillin, as pAN1193 contains an ampicillin resistance gene. Those isolates which were Cat<sup>r</sup> Str<sup>r</sup> Ap<sup>s</sup> were streaked to single colonies and the presence of the nonfunctional *pitB*::Cat<sup>r</sup> gene was confirmed by PCR analysis, using the primers for the *pitB* ORF (Table 2.5). The PCR reaction results for four isolates are shown in Figure 5.2. All isolates amplified an approximately 2.7kb DNA fragment, indicating that the *pitB* gene contained the Cat<sup>r</sup> insert. No 1.5kb DNA band was amplified, which showed that the wild type *pitB* gene

## Figure 5.1

### Plasmid diagram for pAN1193.

Insertional inactivation of wild type *pitB* with a chloramphenicol resistance gene.

The chloramphenicol resistance gene from pBR328 was removed on an *AatII/SauI* fragment and then endfilled with T4 DNA polymerase. This was ligated into the *StuI* site located within *pitB* on pAN656. Restriction endonuclease analysis was used to choose the desired orientation, forming pAN1193.

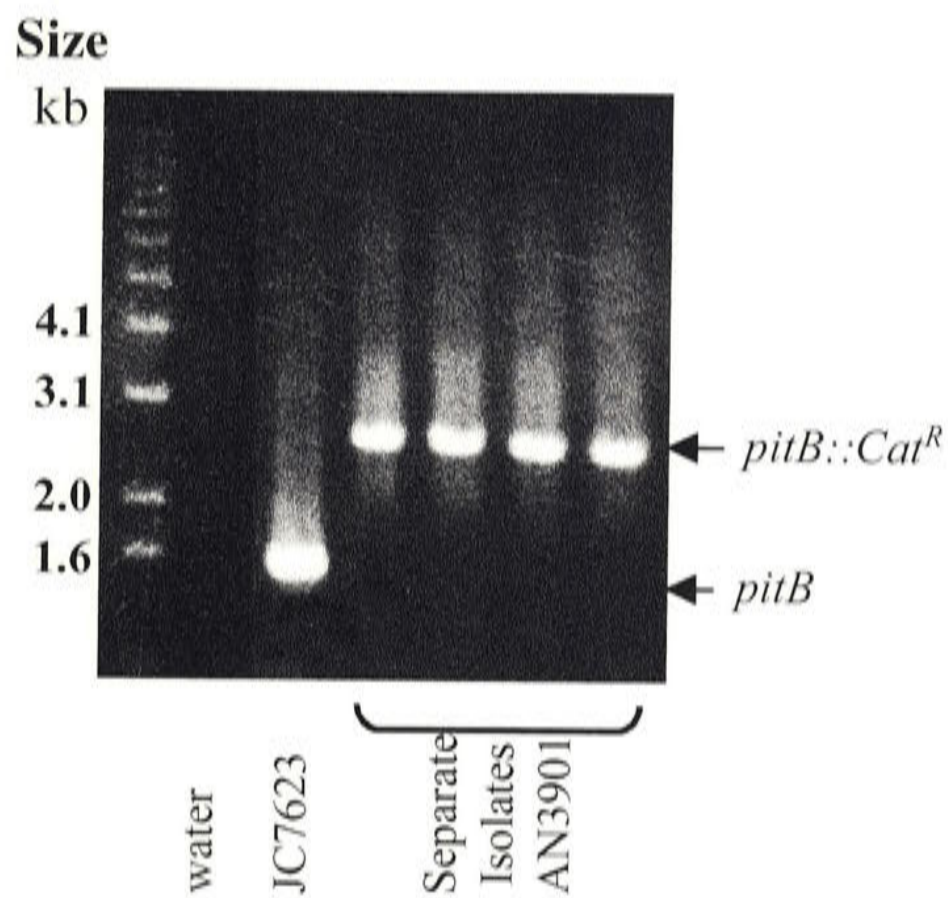


## Figure 5.2

PCR and restriction endonuclease analysis of recombinant JC7623 (AN3901) to confirm the replacement of *pitB* by *pitB::Cat<sup>R</sup>*.

### A PCR amplification of *pitB* and/or *pitB::Cat<sup>R</sup>*

0.8% agarose gel analysis of the PCR products from reactions using the primers for the *pitB* open reading frame (95-89/95-90, See Table 2.5 for details) with genomic DNA from JC7623 and isolates of AN3901, to confirm that the *pitB* (1.5kb) of JC7623 has been replaced by *pitB::Cat<sup>R</sup>* (2.7kb).







was completely replaced. These PCR products were verified by restriction endonuclease digestion.

P1<sub>kc</sub> transduction was used to transfer the *pitB::Cat<sup>r</sup>* gene into the *pitA pstC* strain AN3020, creating *pitA pitB::Cat<sup>r</sup> pstC* (AN3902). Transformants were selected by resistance to 50µg/ml chloramphenicol on rich media containing 1mM G3P. The absence of donor cells was verified by screening for the  $\Delta$ *pstC345* gene of AN3020 using the rapid spray alkaline phosphatase assay. *phoA* alkaline phosphatase is normally inhibited when phosphate concentrations are greater than 0.1-1mM (291), but many mutations in the Pst system, including  $\Delta$ *pstC345*, cause constitutive synthesis of *phoA*. This alkaline phosphatase assay showed that *phoA* was active at high phosphate conditions, indicating the presence of  $\Delta$ *pstC345* (Figure 5.3). Complete removal of wild type *pitB* was confirmed by PCR, as described above (Figure 5.4).

This triple mutant was unable to grow on Pi media and subsequently replaced AN3066 (*pitA pstC recA*) as the background strain for assaying plasmid-borne *pit* genes for Pi function.

### **5.3 Expression of wild type *pitA***

#### **5.3.1 Expression of wild type *pitA* from the K-12 genome**



The K10 strain AN3066, which contains a mutant *pitA1* and a wild type *pitB* on the genome, is unable to grow with Pi as the sole source of phosphate. Yet, *pitB* expresses a functional Pi transporter when located on a plasmid. Could *pitA* also prove nonfunctional when it is the sole *pit* gene on the genome? To explore this further, a *pitA<sup>+</sup> pitB pstC* mutant was created by insertional inactivation of the *pitB* gene in *pstC* strain AN2537.

Bacteriophage P1<sub>kc</sub> transduction was used to transfer the *pitB::Cat<sup>r</sup>* gene (previously prepared in Section 5.2) into the *pstC* strain AN2537. Recipients were selected for the presence of *pitB::Cat<sup>r</sup>* by their resistance to 50µg/ml chloramphenicol on rich media containing 1mM G3P. They were also screened for the presence of  $\Delta$ *pstC345* which

### Figure 5.3

Confirmation that P1 transduced strain AN3902 is not the donor strain AN3901 by analysis of alkaline phosphatase activity.

Cells were grown on rich media and single colonies were patched onto filter paper for the rapid spray alkaline phosphatase assay (22).

Colour	Alkaline phosphatase activity
	-
	+

Recipient strain  
AN3020

Donor strain  
AN3901



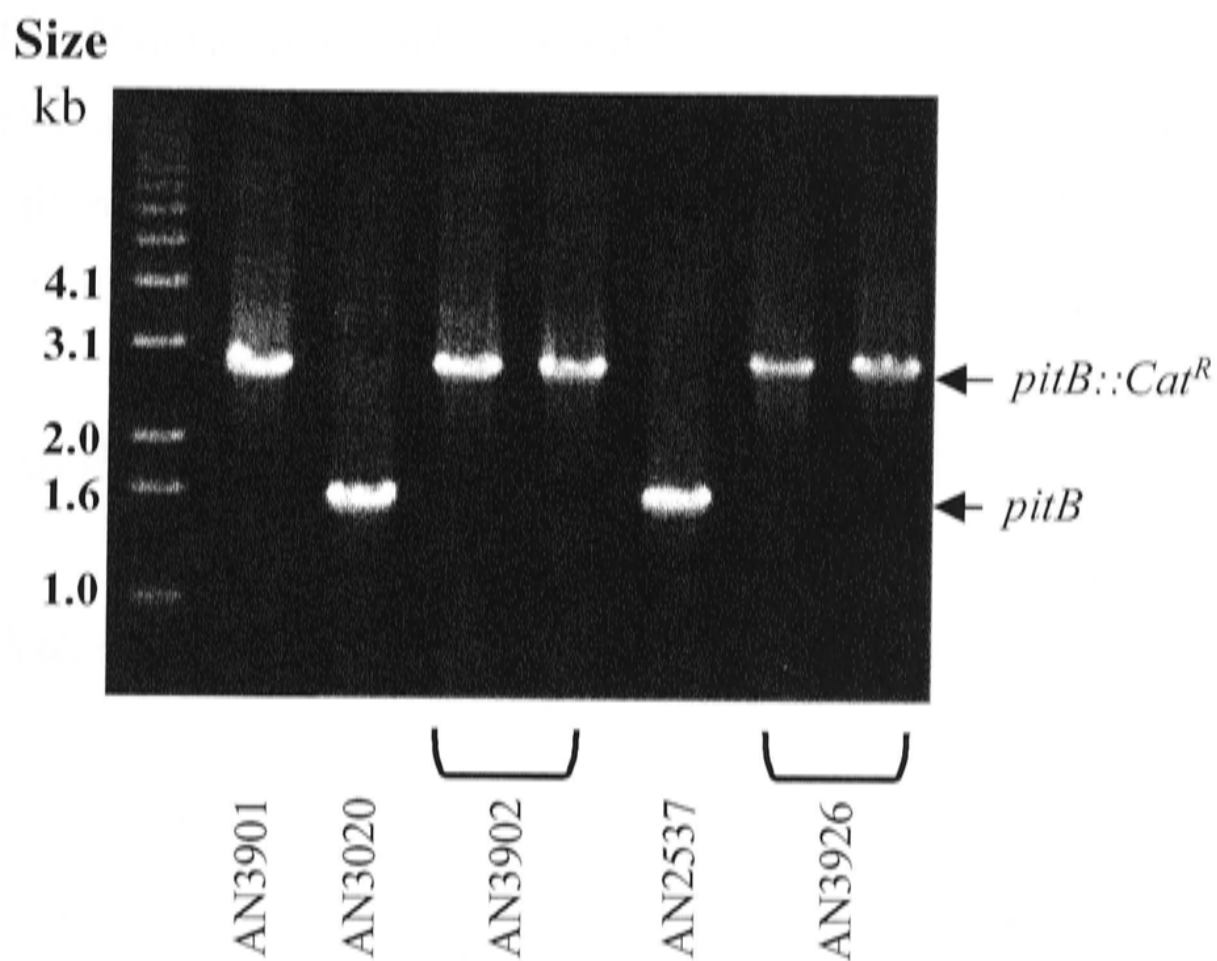
Transduced isolates  
AN3902



## Figure 5.4

### PCR amplification analysis of the P1 transduction of *pitB::Cat<sup>R</sup>* into strains AN3020 and AN2537.

0.8% agarose gel analysis of PCR products from reactions using primers for the *pitB* open reading frame (95-89/95-90, See Table 2.5 for details) with genomic DNA from *pitB::Cat<sup>R</sup>* donor strain AN3901, the recipient strains AN3020 (*pitA1*  $\Delta$ *pstC345*) and AN2537 ( $\Delta$ *pstC345*) and two isolates of each of the P1 transduction products AN3902 (*pitB::Cat<sup>R</sup>* *pitA1*  $\Delta$ *pstC345*) and AN3926 (*pitB::Cat<sup>R</sup>*  $\Delta$ *pstC345*). *pitB* is 1.5kb in size while *pitB::Cat<sup>R</sup>* is 2.7kb.



causes a positive response in the rapid spray alkaline phosphatase assay when it is performed on cells grown on rich media. Inactivation of wild type *pitB* was confirmed by PCR, as described in Section 5.2 (Figure 5.4).

This new strain, AN3926, which contains wild type *pitA* and nonfunctional *pitB* and *pst* was able to grow on Pi media. This Pi uptake activity can be attributed to PitA, as the *pitA pitB::Cat<sup>r</sup> pstC* triple mutant, AN3902, cannot grow on Pi media. The ability of *pitA*<sup>+</sup> strain AN3926 to transport Pi shows that genomic *pitA* is expressed in the presence of a mutated *pitB* gene and that the PitA protein can transport Pi in the absence of PitB.

### **5.3.2 Expression of plasmid-borne *pitA* or *pitB* in a Pi transport triple mutant strain**

Wild type *pitA* or *pitB* plasmids transformed into the *pitA pitB::Cat<sup>r</sup> pstC* triple mutant AN3902 could restore growth on Pi media, as they did in AN3066, confirming that both PitA and PitB can transport Pi independent of each other.

## **5.4 Expression and activity of wild type *pitB***

### **5.4.1 Variability of *pitB* kinetic parameters $K_m^{app}$ and $V_{max}^{app}$**

The lack of Pi transport from the genomic wild type *pitB* located on AN3066 indicates that *pitB* may be repressed and/or inhibited under certain conditions. In addition, the maximum apparent velocity ( $V_{max}^{app}$ ) and the apparent  $K_m$  ( $K_m^{app}$ ) of the PitB protein were unexpectedly altered by reducing the DNA upstream of the *pitB* ORF from 1403 to 207 nucleotides.

*pitB* was isolated on a *ClaI/BamHI* cosmid fragment and subcloned into vector pBR322 with 1403 nucleotides of genomic DNA upstream of the open reading frame to form pAN656. pAN1116 was created by subcloning *pitB* on a *SspI/ClaI* fragment into the *EcoRV/ClaI* site of pBR322 (Figure 5.5). This reduced *pitB*'s upstream genomic DNA from 1403 to 207 nucleotides, removing restriction endonuclease sites, which would

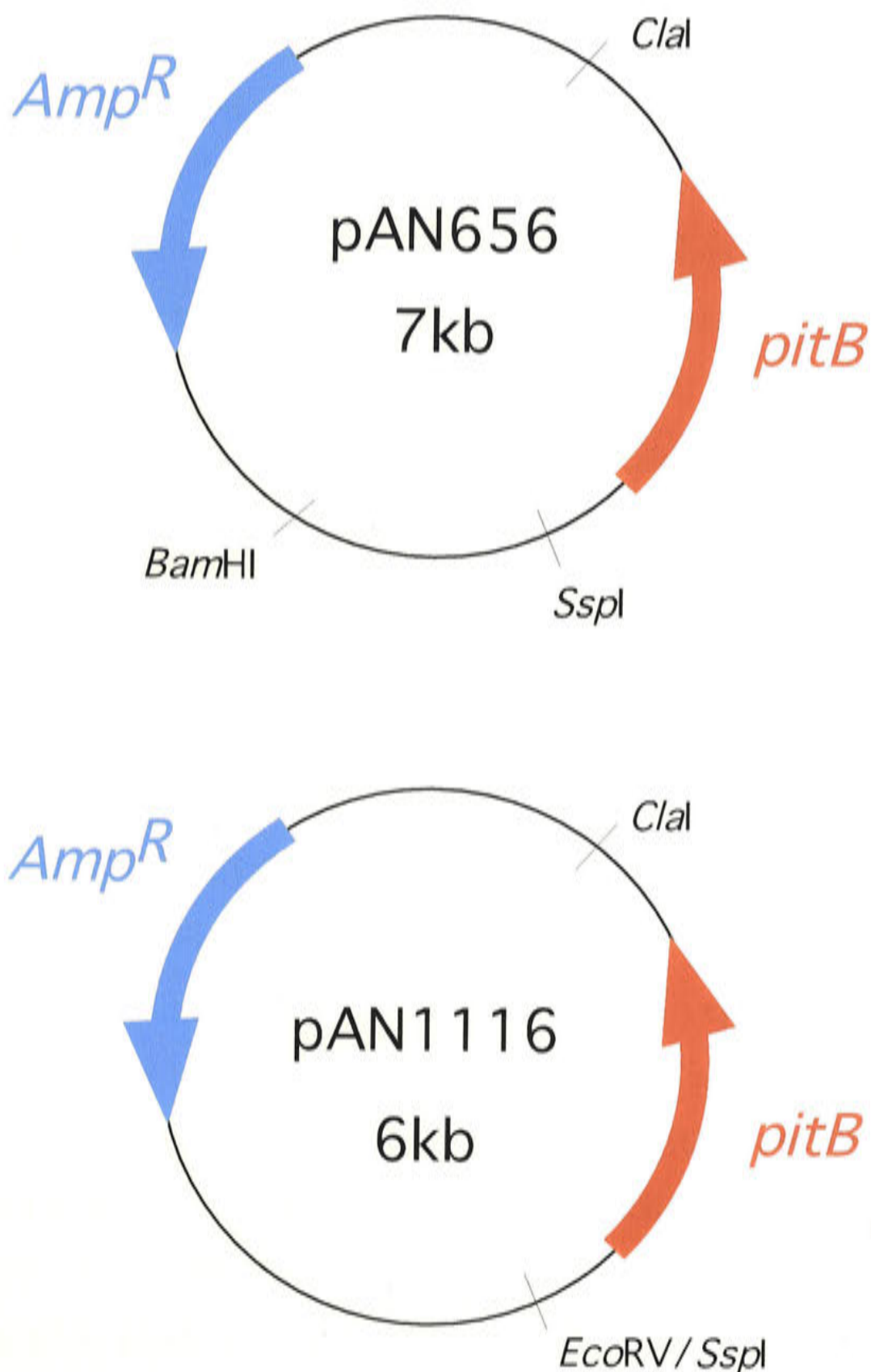
## Figure 5.5

### Plasmid diagrams for pAN656 and pAN1116.

Wild type *pitB* with either 1403 (pAN656) or 207 (pAN1116) genomic nucleotides upstream of the putative *pitB* open reading frame.

*pitB* was isolated on a cosmid *Cla*I/*Bam*HI fragment (Harris *et al* (89)) and ligated into the *Cla*I/*Bam*HI of pBR322, creating pAN656. This plasmid was constructed by Dianne C. Webb.

The *Ssp*I/*Bam*HI region of genomic DNA upstream of the *pitB* open reading frame was deleted by removing *pitB* from pAN656 on an *Ssp*I/*Cla*I fragment, and ligating it into pBR322 digested with *Eco*RV/*Cla*I, producing pAN1116.





interfere with the production of *pitA/pitB* chimeras (Section 6.4). Pi uptake assays on these plasmids showed that removal of these 1196 nucleotides increased the  $V_{\max}^{\text{app}}$  by four-fold (Table 5.1). The variability in  $V_{\max}^{\text{app}}$  recorded from individual experiments was also reduced as shown by the decrease in error values between pAN1116 (*pitB* short) and pAN656 (*pitB* long). This reduction in  $V_{\max}^{\text{app}}$  variability was greater than these error values suggest, as five experiments with pAN656 produced Pi transport activity too low to calculate a  $K_m^{\text{app}}$ , thus were not included in Table 5.1. pAN1116 never produced very low Pi transport activity.

This reduction in *pitB*'s upstream DNA also decreased the  $K_m^{\text{app}}$  of PitB 5-fold (Table 5.1), a significant change that cannot be ascribed to experimental variation.  $K_m^{\text{app}}$  is normally expected to remain constant even if the  $V_{\max}^{\text{app}}$  is greatly altered as  $K_m^{\text{app}}$  effectively represents the characteristic mechanism of the transporter. This alteration in  $K_m^{\text{app}}$  indicates that the transport mechanism of PitB has changed.

Different plasmid constructs may change the copy number of a gene or the amount of transcription, thus altering the levels of protein expression and the maximum apparent velocity ( $V_{\max}^{\text{app}}$ ) of a transporter in whole cell assays. Alternatively, modifications may occur at the protein level, such as oligomerisation or the interaction with activators, which alters both the PitB  $V_{\max}^{\text{app}}$  and  $K_m^{\text{app}}$ , or a combination of the above may take place. To determine if the increase in  $V_{\max}^{\text{app}}$  and the reduction in  $K_m^{\text{app}}$  correlated with an increase in protein expression, the levels of PitB protein inserted in the cell membranes of strains containing either pAN656 or pAN1116 were compared.

#### **5.4.2 *PitB* antipeptide antibody production**

Polyclonal antipeptide PitB antibody was prepared to visualise the PitB protein. The peptide used is in the equivalent position to the PitA peptide, and forms part of an extramembranous loop in the putative folded structure (Figure 3.15). Initially, antibody preparation was attempted by immunising rabbits with a multiple antigen peptide system (MAP) conjugate consisting of the peptide DRIHRIPEDRKKKKK<sub>C</sub> attached to a polylysine core via a C-terminal cysteine, as prepared for PitA. However, this MAP conjugate failed to produce an antigenic response in several rabbits, whose sera was

**TABLE 5.1:** Kinetic parameters for PitB with either the long or short upstream nucleotide sequence.

Plasmid <sup>a</sup>	Phenotype	Assay conditions		$K_m^{\text{app}}$ <sup>b</sup> $\mu\text{M}$	$V_{\text{max}}^{\text{app}}$ <sup>b</sup> $\text{nmol Pi min}^{-1}$ $\text{mg dry weight}^{-1}$
		pH	[Mg] mM		
pAN656	<i>pitB</i> <sup>+</sup> long <sup>c</sup>	6.6	1.8	28.6±1.4 (n=5)	33±9 (n=5)
pAN656	<i>pitB</i> <sup>+</sup> long <sup>c</sup>	7.0	10.0	28.1±1.0 (n=3)	17±3 (n=3)
pAN1116	<i>pitB</i> <sup>+</sup> short <sup>d</sup>	7.0	10.0	6.0±0.5 (n=3)	67±2 (n=3)

a host strain is AN3066 for all experiments

b SEM

c 1403 upstream nucleotides

d 207 upstream nucleotides

screened by ELISA and by Western blot against cell membrane fractions. The PitB peptide was then attached to maleimide activated keyhole limpet hemocyanin and this conjugate was used for immunisation as described in Section 2.13.2. Detection of sera antibodies to PitB was carried out by a Western blot against the membrane fractions of AN3066 containing either vector (AN3514), wild type *pitA* (AN3531) or the long *pitB* plasmid (AN3135) (Figure 5.6). A protein of approximately 45kDa was detected in AN3135 (long *pitB* plasmid) exposed to post-immune sera which was not present in either the *pitA* plasmid strain or the background strain. No protein at an equivalent position was visible in the AN3135 membrane fractions exposed to pre-immune sera. While PitB protein has a deduced molecular weight of 53.5kDa, *in vitro* transcription/translation carried out on wild type *pitB* plasmid pAN909 also produced a protein at approximately 45kDa which was negligible in its control (Figure 3.6) and many membrane proteins undergo anomalously fast migration during electrophoresis. The western blots of most sera samples revealed several other strong bands of protein. These were approximately uniform in intensity for each strain, and were also present in the pre-immune sera, indicating that nonspecific binding by other polyclonal antibodies was occurring. Therefore PitB antipeptide polyclonal antibody was purified from pooled sera by immuno-affinity purification following selective ammonium sulfate precipitation, as described in Section 2.13.4, greatly decreasing the presence of nonspecific bands. There was no cross-reactivity of this purified antibody with the PitA protein, as shown in lane 5 of Figure 5.7.

### **5.4.3 Investigation of PitB protein expression**

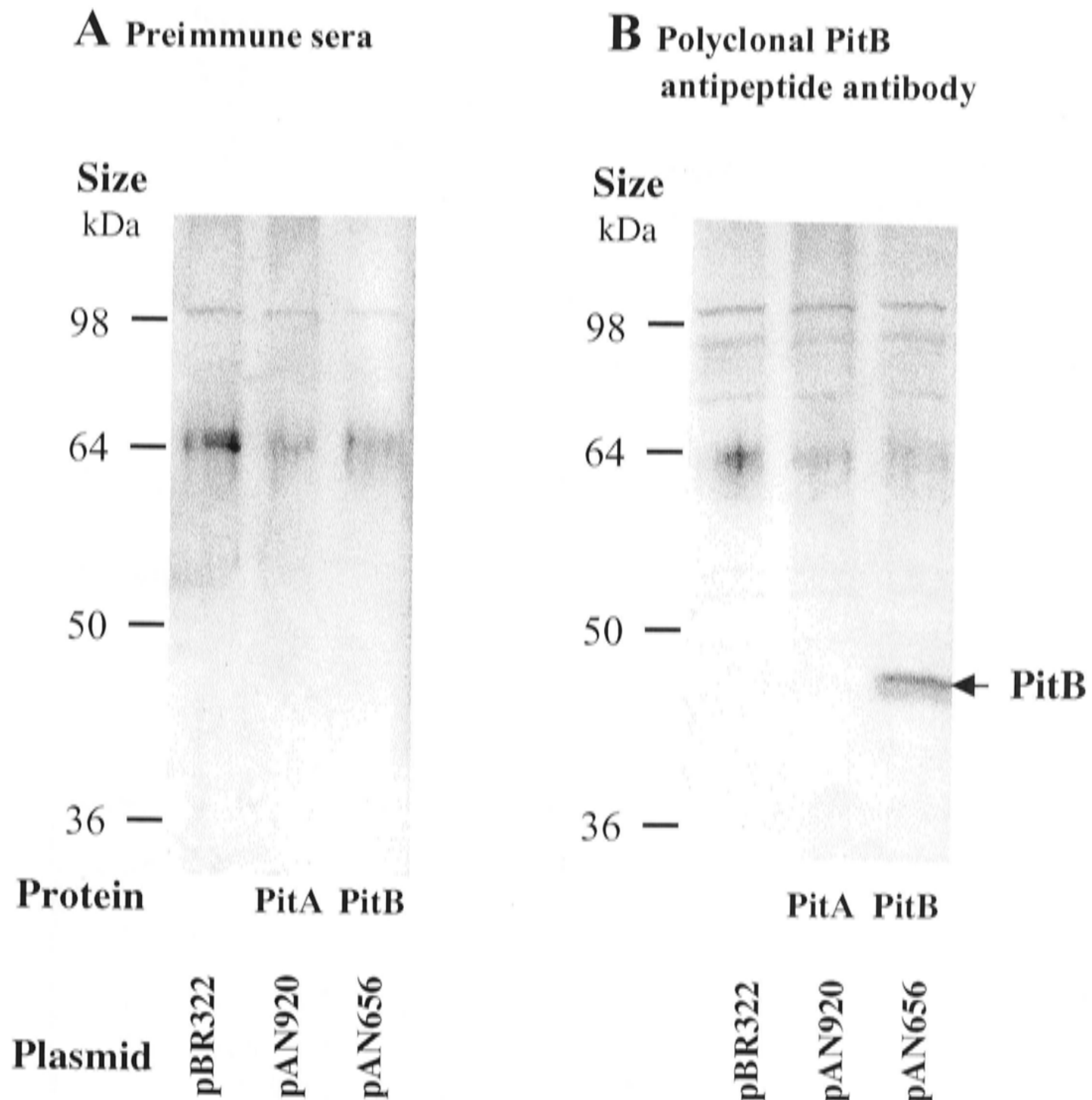
Strains with no Pi uptake containing vector pBR322 and either genomic *pitB::Cat<sup>r</sup>* (AN3903) or genomic wild type *pitB* (AN3514) in the *pitA1 ΔpstC345* background were analysed by western blot for the presence of PitB protein (lanes 1 and 2 - Figure 5.7). Only negligible levels of PitB protein could be detected by polyclonal antipeptide PitB antibody in the membrane fractions of these strains. Therefore genomic wild type *pitB* does not express PitB protein in a *pitA1 ΔpstC345* background, and the lack of Pi uptake is due to the absence of protein rather than the presence of an inactive protein. In these circumstances the presence of genomic wild type *pitB* gene is indistinguishable from *pitB::Cat<sup>r</sup>* for both PitB function and PitB protein expression. A strain with wild type *pitA* on a plasmid was also included in this western blot, to show that the PitB



## Figure 5.6

### Western blot analysis of polyclonal PitB antipeptide antisera against PitA and PitB protein.

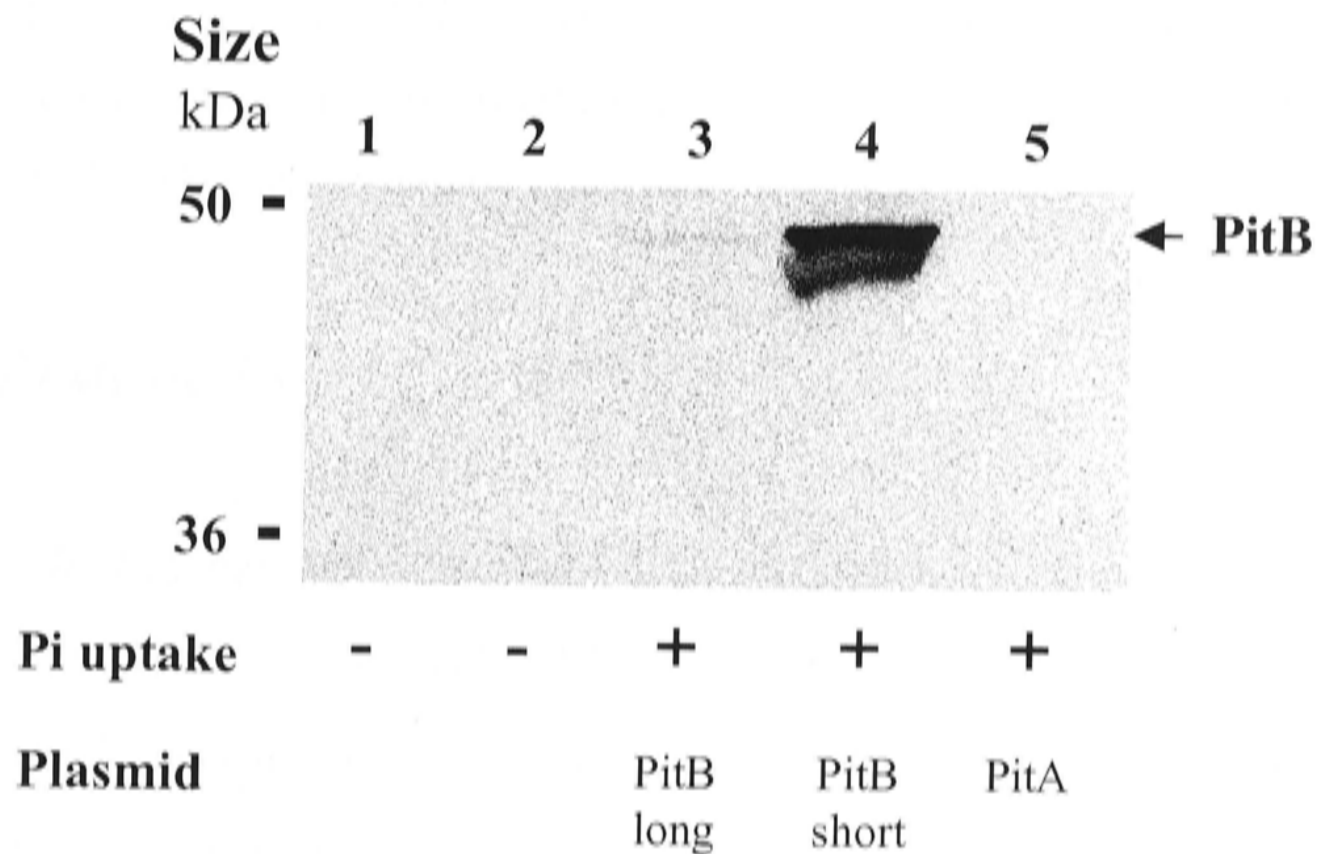
The membrane fractions of the Pi auxotrophic strain AN3066 containing the plasmids pBR322 (vector), pAN920 (PitA) or pAN656 (PitB) were solubilised and separated by PAGE on a 10% polyacrylamide gel and transferred onto PVDF membrane. Protein was visualised by incubation with either preimmune sera or sera taken after immunisation with conjugated PitB peptide/keyhole limpet haemocyanin, followed by incubation with alkaline phosphatase conjugated goat anti-rabbit secondary antibody and the application of Western blue alkaline phosphatase substrate.



## Figure 5.7

### Western blot analysis of PitB protein expression from various strains using purified polyclonal PitB antipeptide antibody.

The membrane fractions of the strains listed below were solubilised and separated by PAGE on a 10% polyacrylamide gel and transferred onto PVDF membrane. Protein was visualised by incubation with polyclonal PitA antipeptide antibody, followed by incubation with alkaline phosphatase conjugated goat anti-rabbit secondary antibody and the application of Western blue alkaline phosphatase substrate. Plasmid containing *pitB* (long) has genomic DNA for 1403 nucleotides upstream of the open reading frame while *pitB* (short) has 207 nucleotides.



Lane	Strain	Pi transport genotype	
		Genome	Plasmid
1	AN3903	<i>pitA1 pitB::Cat<sup>R</sup> ΔpstC345</i>	pBR322
2	AN3514	<i>pitA1</i>	pBR322
3	AN3135	<i>pitA1</i>	<i>pitB<sup>+</sup></i> (long)
4	AN3905	<i>pitA1 pitB::Cat<sup>R</sup> ΔpstC345</i>	<i>pitB<sup>+</sup></i> (short)
5	AN3904	<i>pitA1 pitB::Cat<sup>R</sup> ΔpstC345</i>	<i>pitA<sup>+</sup></i>

antipeptide polyclonal antibody had no cross reactivity with PitA protein. The absence of a band at approximately 48kDa indicates that genomic *pitB* does not express PitB protein in a  $\Delta pstC345$  background, as well as the *pitA1*  $\Delta pstC345$  background.

Western blots were also used to visualise the PitB protein in the membrane fractions of the above background strains containing either pAN656 (1403 upstream nucleotides) or pAN1116 (207 upstream nucleotides) (lanes 3 and 4 - Figure 5.7). The presence of plasmid pAN656 (wild type *pitB* plus 1403 upstream nucleotides) resulted in low levels of PitB protein in the cell membranes, while large amounts of PitB were produced when the upstream *pitB* DNA was decreased to 207 nucleotides in pAN1116. When the length of upstream DNA of *pitB* was decreased by 1196 nucleotides the  $V_{\max}^{\text{app}}$  increased 4-fold (Table 5.1). Therefore this western blot showed that the higher Pi uptake activity from pAN1116 correlated with increased PitB protein expression. However, the increase in PitB protein expression was much greater than the 4-fold elevation in  $V_{\max}^{\text{app}}$  noted in uptake experiments (Section 5.4.1).

## **5.5 Regulation of *pitB***

### **5.5.1 Deletion of the *phoB-phoR* operon from *pitA1* $\Delta pstC345$ strains with/without genomic *pitB***

The *pho* regulon consists of many genes involved in phosphate assimilation that are induced via activation of the *phoB-phoR* two component regulatory system when external Pi concentrations are limiting. This regulon may also include genes that are repressed under these conditions (271, 291) such as the Pit-like Pi transporter of *Rhizobium meliloti* (8). The potential regulation of *pitB* by the *phoB-phoR* operon was investigated by mutating the *phoB-phoR* operon of AN3020, which contains a wild type *pitB* and the *pitA1*  $\Delta pstC345$  Pi transport mutations. AN3020 is unable to grow on Pi media. The *phoB-phoR* operon was also deleted from AN3902, the *pitB::Cat<sup>r</sup> pitA1*  $\Delta pstC345$  triple mutant, which was used as a *pitB* control. This strain is also unable to grow on Pi media.



Strain ANCH1 contains a *phoB-phoR* operon deletion which replaces these genes with a kanamycin resistance gene (304). This *phoB-phoR* deletion was transduced into AN3020 and AN3902 with bacteriophage P1<sub>kc</sub>. Transductants were selected by their resistance to kanamycin on rich media containing 1mM G3P (which prevents selection pressure for Pi transport). Inactivation of the *pho* regulon was determined by testing single colonies grown on rich media for *phoA* activity using the rapid spray alkaline phosphatase assay. A negative result indicated that the *phoBR* operon had been deleted, as both AN3020 and AN3902 contain the  $\Delta$ *pstC345* mutation that causes constitutive induction of *phoA*. These new strains were AN4081 (*pitA1*  $\Delta$ *pstC345*  $\Delta$ (*phoB-phoR*) Kan<sup>r</sup>) and AN4085 (*pitB::Cat<sup>r</sup>* *pitA1*  $\Delta$ *pstC345*  $\Delta$ (*phoB-phoR*) Kan<sup>r</sup>).

### 5.5.2 Effect of the *pho* regulon on *pitB* activity

AN3020 and AN3902 have negligible Pi transport and do not grow on Pi media. The introduction of the *phoB-phoR* deletion into AN3020, which contains wild type *pitB*, enabled the new strain to grow on Pi media. Therefore the presence of *phoB-phoR* does repress Pi transport in AN3020 (*pitA1*  $\Delta$ *pstC345*). However deleting *phoB-phoR* from AN3902, which has a mutated *pitB* gene, also enabled this new strain to grow on Pi media. Thus the inactivation of the *pho* regulon allowed one or more systems to transport Pi in both these strains. No conclusions can be made about the possible regulation of *pitB* through the *pho* regulon as AN4081(*phoB-phoR* *pitA* *pstC*) and AN4085 (*pitB* *phoB-phoR* *pitA* *pstC*) cannot be differentiated by a Pi complementation test.

Pi uptake assays revealed that AN4085, which contains no Pit or Pst transporters, had Pi uptake that was not significantly above background levels (Table 5.2, Figure 5.8). However the *pitB*<sup>+</sup> strain, AN4081, had a significant rate of Pi transport. The fact that the control strain was able to grow on Pi media but had no measurable Pi uptake may be attributed to the difference in Pi concentrations used in these experiments. Growth was assessed at 500 $\mu$ M Pi, while Pi uptake was measured at 20 $\mu$ M Pi. Thus, the growth of the control strain may be due to the presence of a transport system which has a lower affinity for Pi than either PitA or PitB. As the *pho* regulon contains a large number of genes involved in Pi assimilation more than one system may be involved.

**TABLE 5.2:** Effect of the *phoBR* operon on PitB inorganic phosphate uptake activity

Strain	Relevant genotype	Alkaline phosphatase activity <sup>a</sup>	Growth on minimal media with the following source of phosphate		Phosphate uptake <sup>d</sup> nmol Pi min <sup>-1</sup> mg dry weight <sup>-1</sup>
			Pi and G3P <sup>b</sup>	Pi <sup>c</sup>	
AN3066	<i>pitB</i> <sup>+</sup> <i>phoB</i> <sup>+</sup> <i>phoR</i> <sup>+</sup>	+	+	-	0.6 ± 0.1
AN4081	<i>pitB</i> <sup>+</sup>	-	+	+	4.4 ± 0.3
AN4085	-	-	+	+	0.9 ± 0.2

a rapid spray assay on cells grown on Luria Bertani media (22).

b 500μM Pi and 1mM glycerol-3-phosphate.

c 500μM Pi.

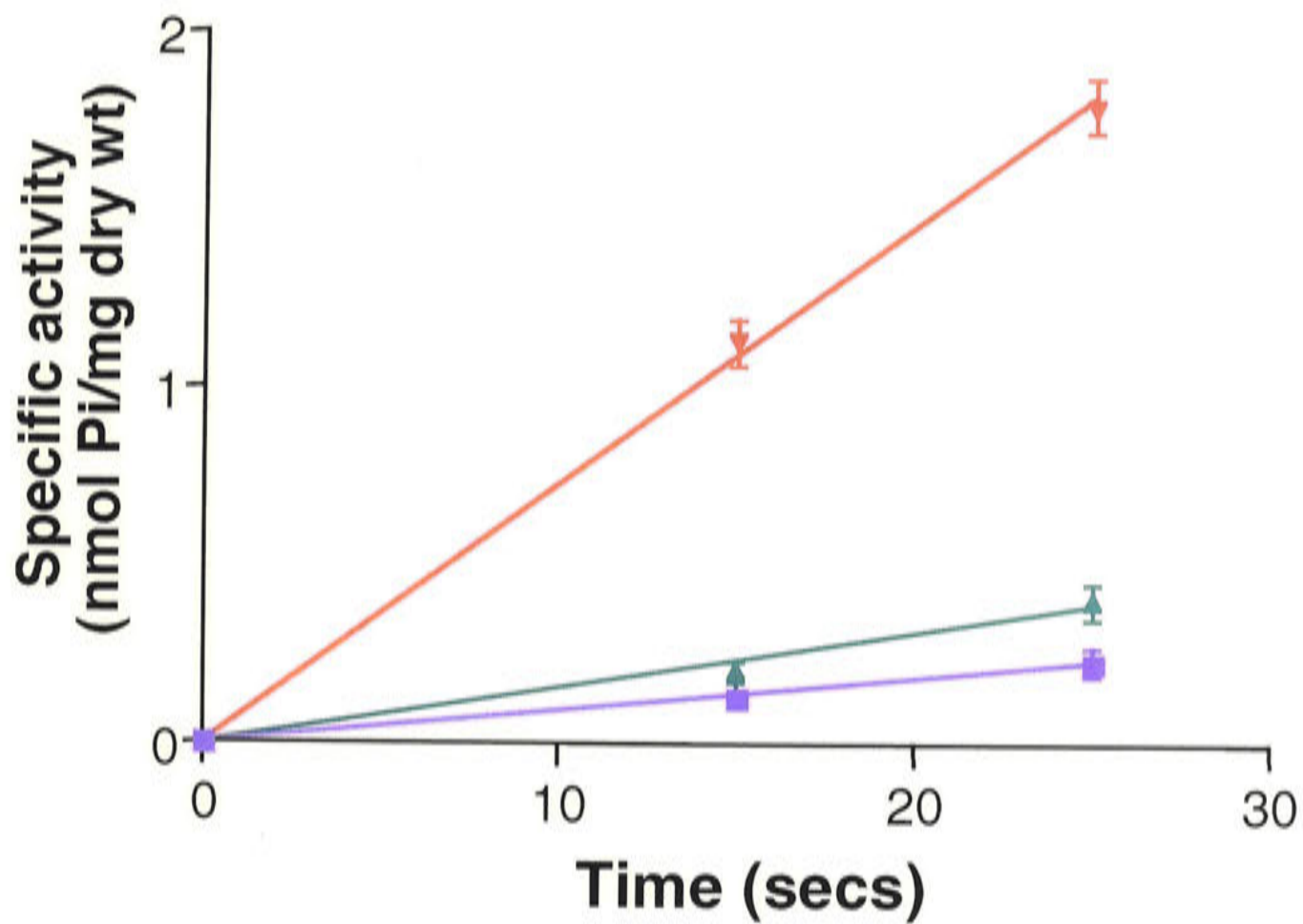
d assayed at 20μM Pi, SEM, n=4.

## Figure 5.8

### Effect of the *phoB-phoR* operon on PitB Pi uptake activity.

Strains containing mutations in *pitA* and *pstC* were analysed for the effect of the *phoB-phoR* operon on *pitB* activity, by deleting the *phoB-phoR* operon and/or *pitB* and measuring the initial rates of Pi uptake, using 20 $\mu$ M Pi.

Shown below are the assays used to determine the initial rates listed in Table 5.2. These were calculated using four experiments (each having two to three individual determinations per time point) by nonlinear regression in the Graphpad Prism program. Error bars represent the standard error of the mean (SEM).



- ▼ AN4081 *-pitB*<sup>+</sup>
- ▲ AN4085
- AN3066 *-pitB*<sup>+</sup> *phoB*<sup>+</sup> *phoR*<sup>+</sup>



*pitB*<sup>+</sup> strain AN4081 exhibited Pi transport of 4.4nmol Pi min<sup>-1</sup> mg dry weight<sup>-1</sup> at 20μM Pi. By comparison, Pi uptake by PitB expressed from pAN656 was 5-25nmol Pi min<sup>-1</sup> mg dry weight<sup>-1</sup> at 20μM Pi, and PitB on pAN1116, which produced greater protein expression and a lower  $K_m^{app}$ , had Pi uptake rates of 46-59nmol Pi min<sup>-1</sup> mg dry weight<sup>-1</sup> at 20μM Pi (assay results isolated from  $K_m^{app}$  experiments, data not shown). Thus, these experiments show that chromosome-encoded *pitB* is active in the absence of the *phoBR* operon.

## 5.6 Discussion

Results described here indicate that genomic *pitB* is regulated. The lack of Pi uptake in strains with *pitB*<sup>+</sup> *pitA1*  $\Delta$ *pstC345* correlates with negligible levels of PitB protein, suggesting regulation is at the level of transcription or translation. This regulation is likely to be mediated through the *pho* regulon, since a deletion which inactivates the *phoB-phoR* operon allows Pi uptake activity that is attributable to the presence of a *pitB* gene. *pitB* may be directly repressed by PhoB, or repressed/inhibited by a *pho* regulon intermediate. Alternatively, *pitB* mRNA may be stabilised.

Construction of a new background strain, AN3902, which has mutations in all three *E. coli* Pi transporters (*pitA1 pitB::Cat<sup>r</sup> ΔpstC345*) allows independent expression of PitA and PitB on plasmids, confirming that both proteins can function independently as Pi transporters. In addition, genomic *pitA* was shown to be free of the regulatory control that represses/inhibits *pitB* in the  $\Delta$ *pstC345* background, as strain AN3926 (*pitA*<sup>+</sup> *pitB::Cat<sup>r</sup> ΔpstC345*) is able to transport Pi.

PitB activity may be modulated by more than one form of regulation, as deletion of DNA 207 nucleotides upstream of *pitB*'s ORF on plasmid pAN656 unexpectedly increased the substrate affinity of PitB for Pi, as well as increasing the maximum velocity of the transporter and the level of PitB protein expression. This change in  $K_m^{app}$  indicates that the Pi transporter's mechanism has been altered.

While most studies on the *pho* regulon in *E. coli* have focussed on the activation of genes involved in phosphate assimilation, there is mounting evidence that some genes within the *pho* regulon are repressed. 2D protein experiments have shown that conditions of Pi limitation induce about 118 proteins of the *pho* regulon and repress around 19 proteins (all with pI's less than seven) (271). This does not include all potentially regulated proteins as many proteins have pI's greater than seven, such as PitA and PitB, which have estimated pIs of 9.5 and 9.9 respectively. More specifically, Willsky and Mallamy (291) have shown that two proteins which are repressed under Pi limiting conditions are not repressed in *phoB* or *phoR* strains under the same conditions. Smith and Payne (244) propose that these are the periplasmic peptide binding proteins OppA and DppA. Putative *pho* box sequences have been identified within the promoter regions of the *E.coli opp* and *dpp* operons and the *R. meliloti orfA-pit* operon (9, 244). These *pho* boxes were found in atypical locations when compared with genes activated by PhoB. Genes activated within the *pho* regulon often lack an efficient -35 promoter region and have a *pho* box located 10 bases upstream from the -10 promoter region. A dimer of phosphorylated PhoB interacts with the *pho* box and with the  $\sigma^{70}$  subunit of the RNA polymerase holoenzyme, and seems to create a bending of the DNA which may assist in initiation of transcription (159, 191). The presence of multiple upstream *pho* boxes can increase the induction by Pi limitation (236). These *pho* boxes may be placed at regular intervals so that PhoB dimers always bind on the same face of the DNA. This has been shown to be the case for the formation of efficient catabolite repressor cAMP-CRP binding sites, which occur at different locations in a number of promoters (e.g. -41 *gal* locus, -65 *lac* operon (145)).

Those operons which may be repressed within the *pho* regulon have putative *pho* boxes upstream of the promoter, or overlapping the proposed -35 region, -10 region or within the transcribed region. It has been suggested that this atypical positioning may reflect the negative regulation by PhoB, but there is no experimental evidence in support of this proposition so far (9, 244). Most *E. coli* transcription repressors are homo-dimers which bind to an inverted repeat located between +30 to -50bp relative to the transcriptional start of +1 (152). The position of a transcription factor binding site relative to the transcriptional start site of a gene can be critical for the precise effect created by this binding. The CreBC two component regulatory system has been shown

to activate and repress different genes. Its consensus “cre-tag” binding sequence, which is similar in organisation to the *pho* box, having a highly conserved TTCAC repeat separated by six nucleotides, occurs between -40 and -80 in *Aeromonas* genes which are activated by this regulon when expressed in *E. coli*. However the putative cre-tag binding site of *malE*, which is repressed by the *cre* regulon, is downstream (i.e. within the transcript) of the transcription start site (7). Atypical binding sites can also cause repression by activating a divergent promoter (87, 134, 148, 285). While most members of the TyrR regulon have two adjacent TyrR boxes overlapping the putative RNA polymerase binding site of the principal promoter P1, the amino acid transporter *aroP* has two adjacent TyrR boxes that are downstream from and outside this binding site. This allows TyrR-mediated activation of a divergent promoter, P3, located on the opposite DNA strand, which directs the RNA polymerase away from promoter P1 (279). Dual activation/repression functions are also carried out by the cI repressor protein of lambda phage. Dimers bind to the P<sub>RM</sub>, P<sub>R</sub>/O<sub>R</sub> regulatory region, repressing the P<sub>R</sub> promoter which controls transcription along one strand, while activating the P<sub>RM</sub> promoter to transcribe the cI mRNA in the other direction (145). Thus, atypical positioning of *pho* boxes in *pho* regulon repressed genes is a feasible mechanism of repression.

It is useful to examine the promoter regions of *pitA* as well as *pitB* for putative *pho* boxes, as *pitA* is reported to be constitutive (211, 212, 249, 289) and seems unaffected by changes in the *pho* regulon. Therefore potential inhibitory *pho* boxes should be unique to *pitB*'s regulatory region. The identification of the RNA polymerase promoters for *pitA* and *pitB* is necessary to determine if these potential *pho* boxes have an atypical location. The *E. coli*  $\sigma^{70}$  RNA polymerase promoter consists of two consensus hexamers separated by a spacer of 17 nucleotides centred at -10 and -35 nucleotides from the site of transcription initiation. Other nucleotides around these hexamers are also conserved, giving a consensus *E. coli* promoter of AAATAATTCTTGACAT(11bp spacer)TTTGGTATAATACA, where the -10 and -35 consensus hexamers are underlined. A search for potential RNA polymerase promoters in *pitA* and *pitB* was carried out with the computer program MacTargSearch. Several of the highest matching possible promoters for *pitA* and *pitB* are listed in Table 5.3. The underlined sequences are identical between *pitA* and *pitB*, or contain only a G/A



**TABLE 5.3:** Comparison of putative RNA polymerase promoter sequences

Sequence	Location <sup>a</sup> (bp)	<i>E. coli</i> promoter sequence <sup>d</sup>			Similarity <sup>e</sup> (%)
		-35 Region <sup>b</sup>	Spacer	-10 Region <sup>b</sup>	
consensus		AAATAATTC <u>TTGACAT</u>	11 bp	TTGGTATAATACA	100.0
<i>pitA</i>					
	<u>34</u> <sup>c</sup>	<u>ATATTTCACTTTGCCC</u>	12 bp	<u>TCACTGATAATGCG</u>	48.5
	<u>34</u> <sup>c</sup>	<u>TATTTCACTTTGCCCCG</u>	11 bp	<u>TCACTGATAATGCG</u>	50.3
	84	AATTATTTTTTGGAGTG	11 bp	GGGGGCAAATCAA	50.3
	85	AATTATTTTTTGGAGTG	10 bp	AGGGGGCAAATCAA	47.3
<i>pitB</i>					
	<u>35</u> <sup>c</sup>	<u>ATATTTCACTTTGCCC</u>	12 bp	<u>TCACTGATAATGCG</u>	48.5
	<u>35</u> <sup>c</sup>	<u>TATTTCACTTTGCCCA</u>	11 bp	<u>TCACTGATAATGCG</u>	51.5
	41	AAATGATATTTCACTT	11 bp	TTAAAGTCACTGAT	43.8
	73	ATAACACTCTTTTTAT	11 bp	GATATTTCACTTTG	46.2
	85	ATTGTCTGATACACAA	11 bp	CTTTTTATCGTTAA	47.9

<sup>a</sup> number of nucleotides from the 3' end of the -10 region to the ORF start

<sup>b</sup> italicised nucleotides represent the -35 and -10 consensus hexamers

<sup>c</sup> underlined sequences are the putative *pitA* and *pitB* promoters that are identical or very similar

<sup>d</sup> from Mulligan *et al* (179) and Goodrich *et al* (76)

<sup>e</sup> calculated by MacTargsearch from consensus sequence information

difference in the last nucleotide of the -35 region. These two matches form virtually the same promoter, as the only difference is the -35 region being displaced sideways by one nucleotide. When the regulatory regions of *pitA* and *pitB* are aligned, the underlined putative promoters have high sequence identity while being surrounded by sequence with lower homology (Figure 5.9). This homology decreases drastically upstream of these putative promoters. Whilst these promoters are speculative, they will be considered in the search for a unique *pitB pho* box.

The 18 nucleotide *pho* box consensus sequence consists of 5'-**CTGTCATA(A/T)A(T/A)CTGTCA(C/T)**-3', where the nucleotides in bold form the direct binding sites for phosphorylated PhoB (159). Further details on the *pho* regulon can be found in Section 1.4. The computer program MacTargsearch was used to look for putative *pho* box sequences in the regulatory regions of *pitA* and *pitB*. Several possible *pho* box sequences with atypical positioning in relation to the putative RNA polymerase promoters on either DNA strand were revealed (Table 5.4). However, there is no obvious *pho* box unique to *pitB*. Either *pitA* and *pitB* have similar *pho* box sequences in the equivalent locations, or there are several *pho* box sequences with similar degrees of identity within the immediate vicinity. Any potential *pho* box involved in negative regulation should be unique to *pitB*, as *pitA* is not repressed by the *pho* regulon. Therefore a simple analysis of the regulatory regions of *pitA* and *pitB* does not reveal a putative *pho* box sequence unique to *pitB* that could be involved in its repression. It should be emphasised that the binding sites for PhoB repression have not been characterized for any gene, so other DNA motifs may be involved. Analysis of the *cre* regulon has revealed that some activated genes contain only half of a "cre-tag" (7). While the active form of PhoB is a phosphorylated dimer, the presence of inhibitory "half sites" cannot be ruled out. Thus, DNA binding studies will be needed to locate any PhoB interactions with the *pitA* and *pitB* genes.

Increasing the copy number of *pitB* by placing it on a plasmid may titrate out *pho* regulon mediated repression, allowing the Pi uptake we have observed with plasmids pAN656 and pAN1116. This titration effect has been recently observed by Hoffer *et al* (99), who analysed a pseudorevertant of a Pi auxotrophic *pitA pstS* strain that conferred the ability to grow with Pi as the sole source of phosphate. They demonstrated that Pi

## Figure 5.9

### Alignment of *pitA* and *pitB* putative RNA polymerase promoter sequences.

Highlighted in red is the putative RNA polymerase promoter sequence that is identical between *pitA* and *pitB*. The open reading frame of each gene is highlighted in green. A second putative RNA polymerase promoter alters the -35 region by one base, so there is a G/A difference between *pitA* and *pitB*, highlighted in blue, and the spacer region is one base shorter. (The -10 region is identical - see Table 5.3 for details.)

```

                                                    -35 region
PitA  473  -----ACAGGGGGCAAATCAAAAAAAGTCTATATTTCACTTTGCCCGCGC
PitB 1297  AAATAACACTCTTTTATCGTTAAAAAATGATATTTCACTTTGCCCATGC
          ***          ***      ****      ***** **
                                                    -10 region
PitA  518  CGCGAAAGTCACTGATAATGCCCGCGTTCATGTCCTAAAATGGCGTAAC
PitB 1347  CGTTAAAGTCACTGATAATGGTCCGTTTCGTAAATTCAAATGGCGTAAT
          ** ***** * * *****
          open reading frame
PitA  568  GT--CCTATGCTACATTTGTTTGGCTGGCCTGGATTTGCATACCGGGCTGT
PitB 1397  CTAATATATGCTAAATTTATTTGTTGGCCTTGATATATACACAGGGCTTT
          * ***** ** ** ** ** ** ** ** ** ** ** ** ** ** ** ** ** ** **
```



**TABLE 5.4:** Comparison of putative *pitA* and *pitB* *pho* box sequences

Sequence	Location <sup>a</sup> (bp)	<i>pho</i> box sequence <sup>b</sup>	Similarity <sup>c</sup> (%)	Identity	
				(binding site)	all
consensus <sup>d</sup>		<u>CTGTCATAAATCTGTCAC</u> T  A          T	100.0	(10)	18
<i>pitA-1</i>	F <sup>e</sup> 2	<u>CTATGCTACATTGTTTG</u>	25.4	(5)	10
<i>pitB-1</i>	F 9	<u>CTCTAATATATGCTAAAT</u>	22.4	(4)	10
<i>pitA-2</i>	F 80	<u>TCTATATTTTCACTTTGCC</u>	<10	(3)	8
<i>pitB-2</i>	F 77	<u>ATGATATTTTCACTTTGCC</u>	20.9	(5)	10
<i>pitA-3</i>	R 97	<u>CTTTTTTTTGATTGCCCC</u>	25.4	(5)	10
<i>pitB-3</i>	R 105	<u>CGATAAAAAGAGTGTTAT</u>	22.4	(6)	11
<i>pitA-4</i>	F 122	<u>TTATTTTTTTGAGTGAAAT</u>	19.4	(5)	9
<i>pitB-4</i>	F 121	<u>GTCTGATACACAAGAAAT</u>	20.9	(5)	9
<i>pitA-5</i>	R 140	<u>TTATCTTATATAATTCAG</u>	31.3	(6)	10
<i>pitA-6</i>	R 144	<u>CTTATATAATTCAGGCAA</u>	29.8	(5)	11
<i>pitB-5</i>	R 134	<u>CAGACAATAAGCATTCAT</u>	26.9	(6)	11

<sup>a</sup> number of nucleotides from the upstream side of the putative *pho* box to the ORF start (i.e. 5' end of + strand sequences, 3' end of - strand sequences)

<sup>b</sup> italicised nucleotides represent the PhoB direct binding sites

<sup>c</sup> calculated by MacTargsearch from consensus sequence information

<sup>d</sup> from Wanner, B.L. (280)

<sup>e</sup> located on the Forward (+ strand) or Reverse (- strand) of the DNA

transport was linked to the presence of multiple genomic copies of *pitB*, caused by an IS5 mediated DNA rearrangement. (The number of *pitB* repeats was not identified.)

Transcriptional repression by PhoB has already been observed for a *pitB*-like *pit* gene in *Rhizobium meliloti*, using a *pit::lacZ* fusion. The *R. meliloti pit* gene was repressed under conditions of Pi limitation but this repression was relieved in a *phoB* mutant (8). As in *E. coli*, mutating a Pst transport system equivalent (*phoCDET*) caused constitutive expression of the *pho* regulon, repressing *pit* at all Pi concentrations. This mechanism of repression may also explain the Pi uptake observed from *pitB* located on various plasmids and cosmids. Transforming a *phoC R. meliloti* strain with plasmid-borne *pit* increases cell growth on high Pi media to normal levels, suggesting that the repression can be overcome by supplying multiple copies of the gene. Repression was also overcome by a single thymidine deletion 54 nucleotides upstream of the *orfA-pit* operon transcription start site which increased the level of *orfA-pit* transcription (9). Only a two- to three-fold increase in *pit* expression was needed to suppress the *phoC* phenotype, overcoming the effects of constitutive expression of the *pho* regulon. Further investigation of *pitB* transcription, using Northern Blots or by replacing genomic *pitB* with a reporter gene such as *lacZ* is required.

*pho* boxes in activated proteins can overlap the binding sites for other transcription factors such as cAMP-CRP (120). The putative *pho* box in *R. meliloti*, which may repress *pit*, overlaps a hepta-thymidine repeat. Deletion of a thymidine from this repeat was shown to increase the transcription of *pit*, allowing Pi uptake to occur in a *phoC* strain (9). A search for an equivalent regulatory region to this hepta-thymidine repeat revealed that *pitA* has a 6 thymidine repeat 117 nucleotides upstream of the ORF, while *pitB* has a 5 thymidine repeat 94 nucleotides upstream of the ORF. As *pitB* undergoes some form of repression and the presence of a 5 thymidine repeat was associated with an increase in the transcription of *R. meliloti pit*, there is no obvious connection between these sequences and the expression of *pitA* or *pitB*.

Post-transcriptional regulation enhancing mRNA stability is a less common, but viable alternative to increasing protein expression through transcriptional activation. Increased stability of mRNA has been shown to occur for the Pi transporter PiT-2 in

human cell lines, as a response to Pi depletion (35). Thus, this mechanism cannot be ruled out for *pitB*.

PitB may also undergo further regulation. Decreasing the length of *pitB*'s upstream DNA from 1403 to 207 nucleotides unexpectedly decreased the  $K_m^{app}$  five-fold, from 28 $\mu$ M to 6 $\mu$ M, making PitB Pi transport more efficient at lower concentrations of Pi. This DNA deletion also raised the  $V_{max}^{app}$  four-fold, showing the overall level of PitB activity had also increased.  $K_m^{app}$  effectively represents the characteristic mechanism of a transporter, and is expected to remain constant even when the protein's level of expression changes. PitB's decrease in  $K_m^{app}$  indicates a different mode of enzyme activity is occurring in the protein expressed from the shortened plasmid pAN1116. There are a variety of ways that a protein's mechanism can be altered, and these usually involve post-translational modification. The increase in PitB protein expression may dilute out an inhibitor, which interacts with the PitB protein, or may allow the transport protein to form a homo-oligomeric complex with a more effective mechanism. Although most secondary transporters are thought to function as monomers, this is not always the case. The sodium proton antiporter from *E. coli* has been crystallised as a dimer (288), and exists in the cytoplasmic membrane as a homo-oligomer (73). Glutamate transporters from the human brain have been shown to form dimers and trimers (90). While lactose permease has been shown to function as a monomer (40), mutants with large deletions are able to complement each other, allowing the possibility that the wild type permease may form oligomers (16). There are also several examples of transporters that have dual-affinity for their substrate and/or two mechanisms (71, 123, 147, 241, 309). Further investigation is needed to determine if this applies to PitB.

There are already clear examples of multiple levels of regulation in other members of the PiT family. The expression of mammalian Pi transporter PiT-2 (which is also a receptor for amphotropic murine leukemia virus) is increased by phosphate starvation, but this is not its only form of regulation. Only a fraction of PiT-2 molecules at the cell surface are capable of processing virus entry, suggesting that the receptor may have active and inactive forms (10). Rodrigues *et al* (207) propose the PiT-2 monomer may be inactive. As Pi concentrations decrease PiT-2 molecules associate with the actin



cytoskeleton network and form high molecular weight complexes. Salaun *et al* (229) suggest that these may be post-translational changes important for receptor activity.

Modulation of Pit activity by a second protein is a distinct possibility. The *pit* gene of *Rhizobium meliloti* forms part of the *orfA-pit* operon. OrfA, also known as the Pit accessory protein, is transcribed at around three times the levels of *pit*, by *lacZ* reporter gene studies (8). Currently its function is unknown, but a Blastp search (NCBI) shows that similar proteins are found in a number of prokaryotes and archaea, including *Pseudomonas aeruginosa*, several *Listeria* and *Pyrococcus* species, *Agrobacterium tumefaciens*, *Haemophilus influenzae*, *Clostridium acetobutyli*, *Vibrio cholerae* and *Archaeoglobus fulgidus*, as well as one human example. The prokaryotic genes are all closely associated with putative *pit* genes whose proteins have some identity to PitA and PitB. Most *pit* genes start within 20 bases of the associated protein's ORF, suggesting they are part of the same operon. Slight variations occur in *Mesorhizobium loti*, which has two "Pit accessory protein" genes preceding a *pit* gene and *Bacillus subtilis* where the *pit* gene occurs before the accessory protein gene. While the "Pit accessory proteins" are linked to a *pit* gene, not all *pit* genes are associated with "accessory protein" genes. Blastp and tBlastn searches (NCBI) produced no matches for "Pit accessory protein" in *E. coli* and neither *pitA* nor *pitB* are closely linked to any other gene. This "accessory protein" could potentially regulate transcription or cause post-translational changes to Pit, or may be involved in another aspect of Pi metabolism.

Western blots using PitB polyclonal antipeptide antibody showed that the four-fold increase in  $V_{\max}^{\text{app}}$ , from 17-67nmol Pi min<sup>-1</sup> mg dry weight<sup>-1</sup>, correlated with increased PitB protein expression in the cell membranes. This increase in expression is not unusual. Different plasmid constructs may change the copy number of a gene or the amount of transcription, altering the levels of protein expression and the  $V_{\max}^{\text{app}}$  of a transporter in whole cell assays. For example, moving *pitA* from a single copy to a multicopy plasmid increased Pi uptake activity at 50μM Pi by 10-fold (59), and genes can be inserted behind strong promoters such as the *tac* promoter, to increase transcription. However, the increase in PitB protein expression was much greater than the four-fold elevation in  $V_{\max}^{\text{app}}$  noted in Pi uptake experiments. This variation may

be due to the different growth conditions used. Pi uptake assays were carried out on cells grown in minimal media while the cells used in western blotting were cultured in Luria Bertani media. The difference in Pi concentrations and other nutrients may have had unintended effects on the regulation of Pi transport. For example, many members of the *pho* regulon, which is constitutively activated in these cells, are also regulated by the cell's carbon and energy sources and the state of the cells metabolism. Analysis of *phoR* mutants also indicates that regulation of the *pho* regulon may sometimes be influenced by these factors (120, 125, 255, 280).

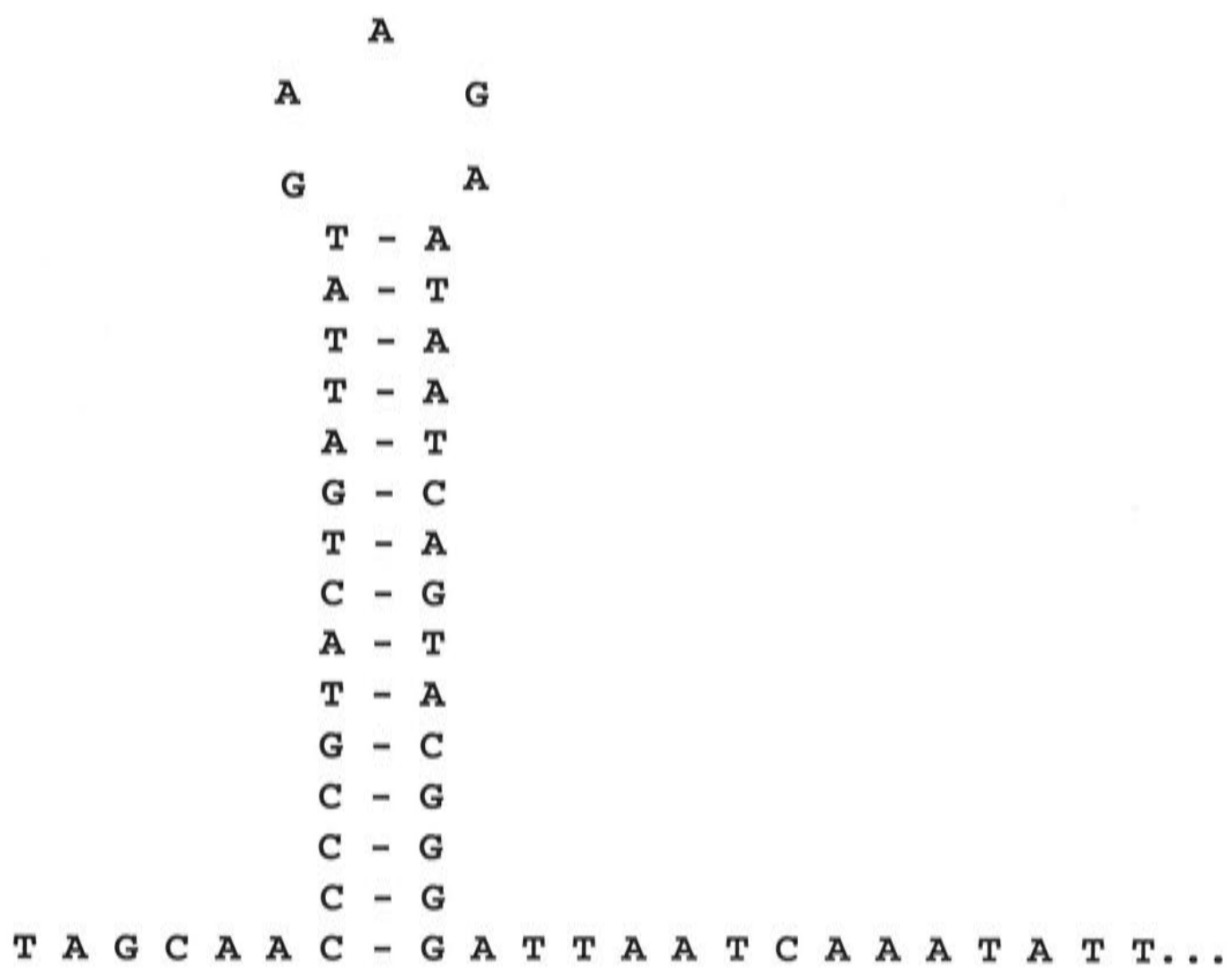
The increase in PitB  $V_{\max}^{\text{app}}$  was also accompanied by a noticeable reduction in the variability of  $V_{\max}^{\text{app}}$  between individual experiments. This stabilisation in  $V_{\max}^{\text{app}}$  was greater than suggested by the standard errors of the mean (SEM) listed in Table 5.1, as pAN656 (*pitB* long) often produced Pi uptake activities too low to allow the calculation of  $K_m^{\text{app}}$  or  $V_{\max}^{\text{app}}$ . These assays on pAN656 may have been carried out under conditions where the regulation of *pitB* was finely balanced, and small changes in growth conditions may have lead to big changes in protein expression. The altered DNA sequence on pAN1116 may have allowed *pitB* to overcome the finely balanced repression which often left PitB activities from pAN656 too low to measure accurately. Research on the *pit* equivalent in *Rhizobium meliloti* noted that only a two- to three-fold increase in *pit* expression was needed to overcome suppression (9). Unfortunately, the levels of PitB protein expression in whole cells was not measured for each  $V_{\max}^{\text{app}}/K_m^{\text{app}}$  determination.

Thus, the deletion of DNA 207 nucleotides upstream of the *pitB* ORF altered both the expression levels and the mechanism of PitB activity, possibly by increasing the levels of *pitB* transcription. While it is unlikely that this deletion affected the *pitB* promoter and any regulatory sequences immediately surrounding it, the excised 1196 nucleotides does contain a putative mRNA transcription terminator (294). This consists of a large stem loop followed by a sequence rich in T and A nucleotides that is appropriately positioned after a large ORF which has the same orientation as the *pitB* ORF (Figure 5.10). Removing this stem loop sequence from pAN656 could increase *pitB* transcription by allowing continuation of any mRNA initiated at the plasmid's ampicillin resistance gene. There is previous evidence for plasmid-borne gene

## Figure 5.10

Structure of the putative stem loop sequence that is present in the upstream DNA of *pitB* on plasmid pAN656, but which is removed on plasmid pAN1116.

The AT rich region continues after the stem loop structure shown below. (See Figure 4.1B for more sequence details.)





expression from non-specified promoters within a vector. The  $F_0F_1$ -ATPase *uncF* gene lacks its own promoter, as it forms part of the  $F_0F_1$ -ATPase operon. This *uncF* gene could restore the ability of an *uncF* mutant strain to grow on succinate as the sole carbon source when it was placed in plasmid pUC18 in the opposite orientation to the *lac* promoter on this vector (206). While the increase in PitB activity and substrate affinity is an interesting result it is also the response to an artificially induced situation that may have little relevance to PitB regulation in a wild type cell. As such, the mechanism causing increased PitB protein expression from pAN1116 was not investigated further.

While *pitA* is not modulated by the *pho* regulon, it is possible that PitA undergoes some form of regulation. PitA's substrate affinity may change when protein expression is increased, as is the case for PitB. Western blots indicate PitA protein expression was greatly elevated by placing it on plasmid pBR322 (AN248 and AN3531, Figure 4.5), and Elvin *et al* (59) showed that initial rates of Pi uptake increased ten-fold when *pitA* was transferred from a single copy to a multicopy plasmid. The  $K_m^{app}$  obtained from plasmid-borne PitA was around 2  $\mu$ M, significantly lower than the 11.9-38 $\mu$ M range recorded by researchers using genomic *pit* genes (211, 290, 268). Further experimentation is needed before any conclusions can be made, as these  $K_m^{app}$  values for genomic *pit* have been measured under a variety of conditions that make it impossible to attribute activity to *pitA* and/or *pitB*.

While our results indicate that PitA is likely to be active under a greater variety of conditions than PitB, and that both transporters can function independently, these do not rule out the possibility that a population of PitA and PitB proteins may interact to form a Pi transporter complex in wild type cells.

---

# ***Chapter six***

***Investigations into the Structure  
and Function of PitA***

---

## ***Chapter six***

# ***Investigations into the Structure and Function of PitA***

## ***6.1 Introduction***

When studying the mechanism of a transporter it is helpful to know the protein structure, so that side chains likely to be involved in the mechanism can be identified and examined. However only a few membrane protein structures have been resolved to atomic resolution and currently none is an ion-coupled transporter. Therefore other methods of analysis must be used. Topology prediction programs utilise the unique features of membrane proteins to propose a basic secondary structure from the amino acid sequence. Further analysis of the protein can then be explored through experimentation.

At the time that PitA and PitB were being analysed, site-directed mutagenesis was being used to elucidate the important amino acids involved in lactose/H<sup>+</sup> symport by lactose permease. The possibility that PitA, which imports a neutral MeHPO<sub>4</sub> complex with a proton (268), may have a similar proton translocation/coupling mechanism to lactose permease was worth exploring. Below is a summary of the work carried out on lactose permease that led to the proposed topological and functional models for PitA and PitB. Current models for the mechanism of lactose permease are described in Section 1.9.2 of the general introduction.

Chemical modification studies showed that histidine residues may be important in coupling proton and lactose translocation. Replacement of His-322 with neutral or positively charged amino acids caused severe loss of activity. H322R carried out facilitated diffusion without proton translocation (117). Glu-325 is located one turn away on the same face of the putative  $\alpha$ -helix and was mutated to neutral side chains which gave mutants exhibiting no active transport or efflux but that were able to undergo wild type levels of exchange and counter flow (29, 30). Therefore these mutants have no proton translocation and are probably unable to release protons.



Kaback proposed that His-322 probably undergoes protonation/deprotonation during symport, and that His-322 and Glu-325 may be components of a charge relay system coupling lactose and proton transport (117).

Mutagenesis of Arg-302, which is located on putative  $\alpha$ -helix 9, showed that mutants with various amino acid substitutions for Arg-302 behaved similarly to the equivalent His-322 mutants. Thus Arg-302 may also be involved in the putative charge relay pathway (171). Lys-319 was analysed as it is located one turn away from His-322 on  $\alpha$ -helix 10, in the opposite direction to Glu-325. Neutral substitutions for Lys-319 produced mutants with no active transport or efflux that exhibited wild type exchange and counterflow. The similarity of these mutants to those produced by neutral substitutions for Glu-325 indicates that Lys-319 may also be involved with proton translocation, either directly or by modulating proton transfer capabilities of nearby residues (209).

Second site suppressor experiments have identified complementary mutations for detrimental mutants, such as the neutral substitution K319L. This approach identified potential interactions between Lys-319 and Asp-240 or Glu-269 (142, 143). Mutagenesis studies showed that the neutral combination of K319C/D240C had activity and normal protein assembly. Therefore this pair may interact as a salt bridge, and these results imply that K-319 is not directly involved in proton translocation. However, the polarity of the interaction could not be interchanged, so the local charges on Lys-319 and Asp-240 are important for normal functioning, possibly by affecting the proton transfer characteristics of nearby side chains (222).

Second site suppressor experiments also discovered a putative Asp-237/Lys-358 salt bridge when neutral suppressor mutations for the detrimental mutant K358T formed at Asp-237 (127). Neutral substitutions at both residues produced an active transporter with low levels of protein insertion in the membrane, indicating that Lys-358 and Asp-237 may form a salt bridge. The Asp and Lys residues could be interchanged, suggesting that the charge of neither residue of the salt bridge is important (52, 127).

These experiments suggest that amino acids involved in the mechanism of lactose permease may include Glu-325, His-322 and Lys-319 on  $\alpha$ -helix 10, plus associated groups Glu-269 on  $\alpha$ -helix 8, Arg-302 on  $\alpha$ -helix 9 and Asp-240 on  $\alpha$ -helix 7. Interactions have been proposed between His-322/Glu-325, His-322/Arg-302, Lys-319/Asp-240, Lys-319/Glu-269 and possibly Lys-319/His-322. These amino acids are highlighted in Figure 6.1. No sulfhydryl/disulfide interconversions are needed for the symport mechanism, and no proline residues were found to be essential for activity (209).

This chapter aims to formulate putative topological models for PitA and PitB using the available tools for secondary structure prediction from amino acid sequence. Residues similar to amino acids involved in the mechanism of the model transporter lactose permease will also be taken into consideration. The functional importance of various amino acids which these models identify as significant will then be examined in PitA. PitA/PitB chimeric proteins will be used to analyse the effects of amino acid differences between these two Pi transporters. Conserved residues and domains within the recently compiled PiT transporter family will then be considered in the light of these findings.

## ***6.2 Putative topological models for PitA and PitB***

A topological model has been formed for PitA based on the GES hydrophobicity profile and predictions from the programs TopPred IV and MEMSAT (Figure 3.15). These programs try to reduce the number of charged or polar residues inserted into the hydrophobic environment of the membrane. However, charged or polar residues often play an important mechanistic role in membrane protein transporters. These side chains may insert into the membrane as uncharged, polar groups or interact with each other to reduce the degree of charge within the hydrophobic environment, by forming salt bridges ((305) and references therein). Therefore polar and charged residues may be under-represented in the membrane regions of the suggested models. Studies on lactose permease, a model protein for symporters energised by the proton-motive force, show that it has three residues on  $\alpha$ -helix 10 which may play a role in the active uptake





of lactose/ H<sup>+</sup> (116). These residues, Lys- 319, His-322 and Glu-325, may be next to each other on the same face of this  $\alpha$ -helix (KxxHxxE). PitA has three similar residues, His-225, Asp-229 and Lys-232, which are located in the periplasm of the TopPred IV models. When placed within an  $\alpha$ -helix, these residues would also line up on the same face of the  $\alpha$ -helix (HxxxDxxK). Equivalent residues are found in PitB and for simplicity, only PitA will be considered here.

Thus, two alternative topological models for PitA have been formulated by modifying the model produced by von Heijne's TopPred IV program to insert these three charged and polar residues within the membrane as part of  $\alpha$ -helix 6 (Figure 6.2). These models are very similar for the first half of PitA and follow the topology predicted by TopPred IV for the first 5 transmembrane (TM) segments and the first four hydrophilic loops (Table 6.1). TopPred IV predicts that His-225, Asp-229, Lys-232 are in loop 6 between TM 6 and TM 7, and moving these residues into the membrane makes the new TM 6 a composite of this putative loop plus the adjacent sections of putative TM 6 and TM 7. Two alternative arrangements have been proposed for the last half of the protein. Model A leaves the remainder of putative TM 7 in the periplasm, so there are 9 transmembrane regions, one large periplasmic loop and a cytoplasmic C-terminus. Model B draws the remainder of putative TM 7 into the membrane so there are 10 transmembrane  $\alpha$ -helices with a medium sized and a large cytoplasmic loop and a periplasmic C-terminus (Figure 6.2). The last three putative  $\alpha$ -helices for each model are similar, but are reversed in orientation within the membrane. The charge ratios (i.e. the number of lysines and arginines within cytoplasmic loops of 60 residues or less subtracted from the number of positively charged residues in similarly sized periplasmic loops) for PitA and PitB in Model A are -14 and -18 respectively, and are -11 and -12 for PitA and PitB in Model B. The TopPred IV charge ratios for PitA and PitB are -10 and -11, respectively. Therefore the new models have more favourable charge ratios.

## Figure 6.2

### Topological models of PitA.

These two models were formulated by combining information from site-directed mutagenesis studies on lactose permease with the information used to determine the topological model illustrated in Figure 3.15 - GES (63) and Kyte-Dolittle (138) hydropathy profiles plus the results of the prediction programs TopPred IV (276) and MEMSAT (111).

The models proposed for PitA and PitB were identical in topology, so the PitA sequence is shown, highlighting the amino acids which show variability when compared with the PitB sequence.

- A Amino acids identical between PitA and PitB.
- A PitA amino acids which are different in the PitB sequence.
- A The amino acids in PitA which may be equivalent to the lactose permease residues R-319, H-322, and E-325, which have been shown to be important for lactose permease function.

The boxes represent the putative transmembrane (TM)  $\alpha$ -helices, which have been drawn to contain 19 amino acids, the predicted minimum number needed to cross the cytoplasmic membrane. The single letter amino acid code has been used.

### Model A

Contains 9 TM domains. This leaves the large variable loop in the periplasm.

### Model B

Contains 10 TM domains. An extra TM domain is inserted after the sixth TM domain, so most of the large variable loop is now in the cytoplasm. The remaining three TM  $\alpha$ -helices are identical to model A, but in the opposite orientation







**Table 6.1:** Predicted transmembrane regions of PitA (all starting with a periplasmic N-terminus)

Transmembrane region	TopPred IV <sup>a</sup>	MEMSAT <sup>b</sup>	Model A <sup>c</sup>	Model B <sup>c</sup>
1	D9 - I29	T12 - F32	T12 – N30	T12 – N30
2	L52 - V72	L52 - I76	L52 – L70	L52 – L70
3	S91 - T111	G93 - L116	A95 – Y113	A95 – Y113
4	H123 - V143	T124 - V144	L125 – S142	L125 – S142
5	G157 - L177	L153 - L177	S158 – L176	S158 – L176
6	F208 - N228	F208 - G226	I219 – V237	I219 – V237
7	G233 - M253	G233 - V251	V384 – W402	G243 – T261
8	A382 - W402	V384 - W402	M428 – T447	V384 – W402
9	Q427 - T447	M428 - L450	M475 – W493	M428 – T447
10	T469 - G489	V470 - W493	-	M475 – W493

a Prediction program combining GES profile with ‘positive inside’ rule, (276).

b Prediction program using ‘dynamic programming’ to optimally thread a polypeptide chain through a set of topology models, (111).

c Models formulated by combining the predictions of TopPred IV and MEMSAT with lactose permease mutagenesis data.

## **6.3 Identification of important amino acids in PitA**

### **6.3.1 Investigation of charged or polar amino acids located within the putative transmembrane helices of PitA**

The putative topological Model A for PitA places three charged amino acids - lysine, aspartic acid and glutamic acid - and two polar histidines within the putative membrane spanning sections of the protein. Three of these amino acids have been specifically placed within putative  $\alpha$ -helix 6 (HxxxDxxK) as they form a similar arrangement to key residues in lactose permease  $\alpha$ -helix 10 (KxxHxxE). Therefore these three PitA residues, His-225, Asp-229 and Lys-232, were targeted for mutagenesis to see if they were important for Pi uptake.

The charged PitA residues His-225, Asp-229 and Lys-232 were individually altered by site-directed mutagenesis to polar amino acid side chains of similar dimensions. His was mutated to Gln (H225Q), Asp was mutated to Asn (D229N) and Lys was changed to Gln (K232Q). This was carried out by single stranded site-directed mutagenesis on wild type *pitA* inserted into M13 DNA (oligonucleotides are listed in Table 2.5). The individual mutations were sequenced and subcloned into the *SalI/BamHI* sites of pBR322 and transformed into AN3066.

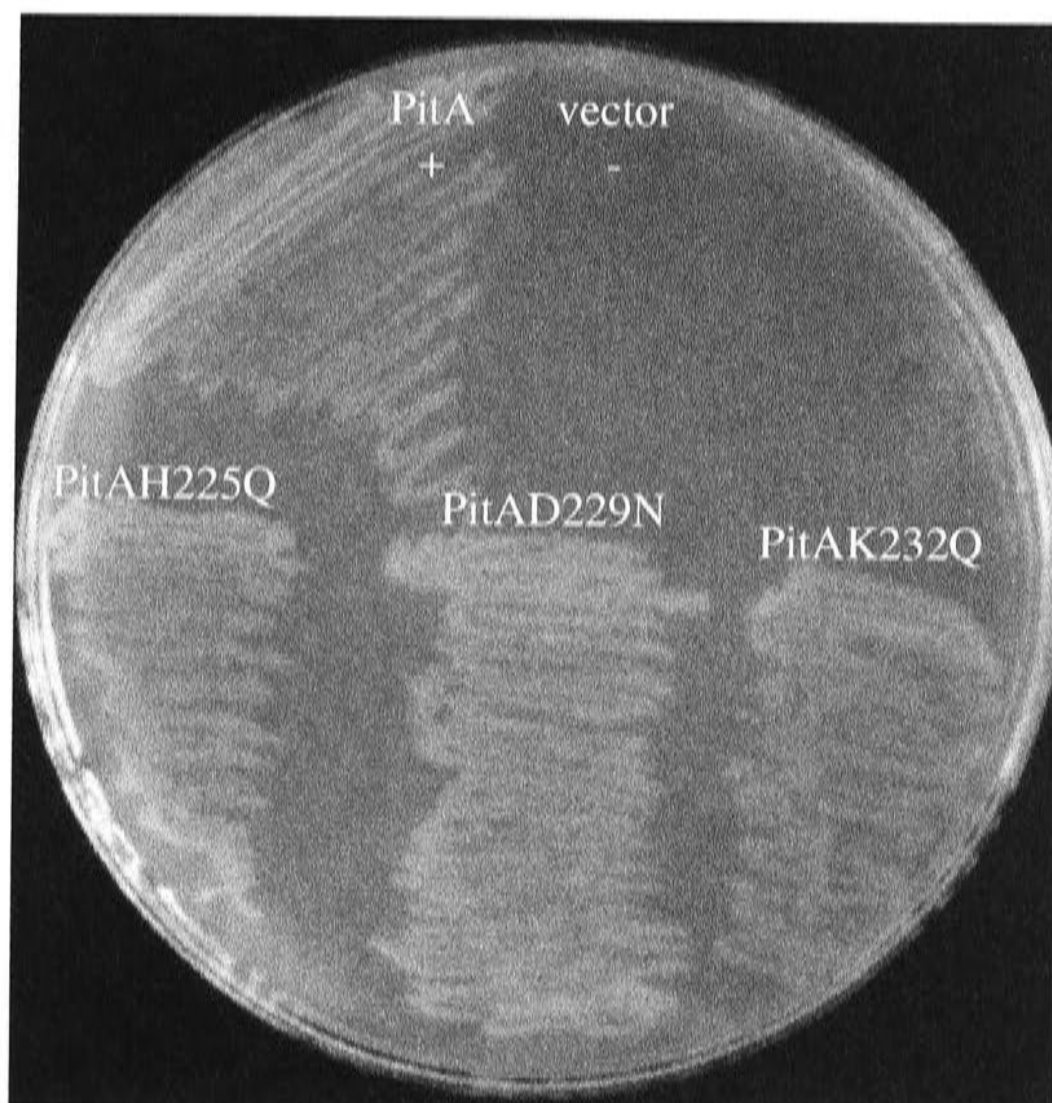
*pitAH225Q*, *pitAD229N* and *pitAK232Q* produced cell growth equivalent to wild type *pitA* on Pi media (minimal media plus 500 $\mu$ M Pi), showing these strains can transport enough Pi for cell growth under these conditions (Figure 6.3). (Intracellular concentrations of Pi in *E. coli* are often around 3-9mM. Cells started to excrete Pi when levels rise to approximately 13mM, with a maximum intracellular Pi concentration of about 30mM, as measured by  $^{31}\text{P}$ nuclear magnetic resonance spectroscopy (262, 302)). Pi uptake assays for each mutated PitA show that initial rates of activity are reduced to less than 3% of plasmid-borne wild type PitA when cells are supplied with 6 $\mu$ M Pi (Table 6.2). However, western blots using polyclonal antipeptide



## Figure 6.3

### Growth of the PitA single mutations H225Q, D229N and K232Q on Pi media.

Plasmids containing PitAH225Q, PitAD229N and PitAK232Q were transformed into the Pi auxotrophic strain AN3066. Single colonies were then streaked onto minimal media containing 500 $\mu$ M Pi, and grown overnight at 37 $^{\circ}$ C. The positive control is AN3531, which has wild type PitA on pBR322 within AN3066, and the negative control is AN3514, which is vector pBR322 within AN3066.



**TABLE 6.2:** Analysis of charged residues within/near putative  $\alpha$ -helix 6 of PitA using site-directed mutagenesis

Strain	Plasmid genotype ( <i>pitA</i> )	Pi uptake			Presence of PitA in cytoplasmic membrane <sup>c</sup>
		Growth on Pi media <sup>a</sup>	Initial rate of Pi uptake nmol Pi. min <sup>-1</sup> mg dry wt <sup>-1</sup>	% wild type (AN3531)	
AN3531	wild type <i>pitA</i>	+	35.0 ± 1.5	100.0	++
AN3511	<i>pitA</i> H225Q	+	0.82 ± 0.07	2.3	++
AN3512	<i>pitA</i> D229N	+	0.63 ± 0.08	1.8	++
AN3513	<i>pitA</i> K232Q	+	0.81 ± 0.05	2.3	+
AN3514	vector only	-	0.15 ± 0.07	0.4	-
AN3776	<i>pitA</i> A213D	-	nd	nd	-
AN3938	<i>pitA</i> G220D	-	nd	nd	-

a Growth on Pi media – minimal media containing 500 $\mu$ M Pi.

b Assayed at 6 $\mu$ M Pi, SEM, n = 5.

c Determined by western blot of cytoplasmic membrane fractions using PitA polyclonal antipeptide antibody. ++ indicates the presence of more protein than + (see Figure 6.4).

PitA antibody against the membrane fractions of these strains show that the levels of protein expression are similar to that of the wild type *pitA* gene on plasmid pAN686 for H225Q and D229N, with K232Q protein expression being less (Figure 6.4). Therefore the drop in Pi uptake activity for PitA H225Q and PitA D229N cannot be attributed to lack of protein insertion in the membrane. While the amount of K232Q PitA membrane insertion is lower than that of the plasmid-borne PitA protein, it cannot account for the greater than 97% drop in initial velocity for Pi uptake. The PitA protein expression from plasmid-borne K232Q PitA is also higher than that found in AN248, which has wild type *pitA* on the genome. The ability of these strains to grow on Pi media while having very low Pi uptake activity may be due to a number of factors. These plasmid-borne proteins are all expressed at high levels when compared to the amount of PitA protein produced by wild type strain AN248 (Figure 6.4). Thus low levels of Pi transport may be compensated for by high levels of protein expression. Transporters impaired in active Pi uptake, but able to undergo facilitated diffusion, may also be able transport enough phosphate from the 500 $\mu$ M Pi media to allow cell growth while exhibiting low initial rates of uptake in assays using 6 $\mu$ M Pi.

### **6.3.2 Investigation of the role of other PitA amino acids**

In the course of placing *Bgl*III restriction endonuclease sites within the *pitA* gene, various residues towards the end of the putative hydrophilic loops were replaced with aspartic acid, leucine and a number of other amino acid substitutions (Figure 6.5). (This was done in preparation for making *phoA* sandwich fusions to determine the PitA topology - an approach that was left when the investigation of PitB regulation became paramount.) These mutations effectively serve as a limited form of random mutagenesis. While most of these altered proteins (including eight with aspartic acid substitutions) could grow on minimal media containing 500 $\mu$ M Pi, two PitA constructs had negligible growth on this Pi media (Table 6.3).

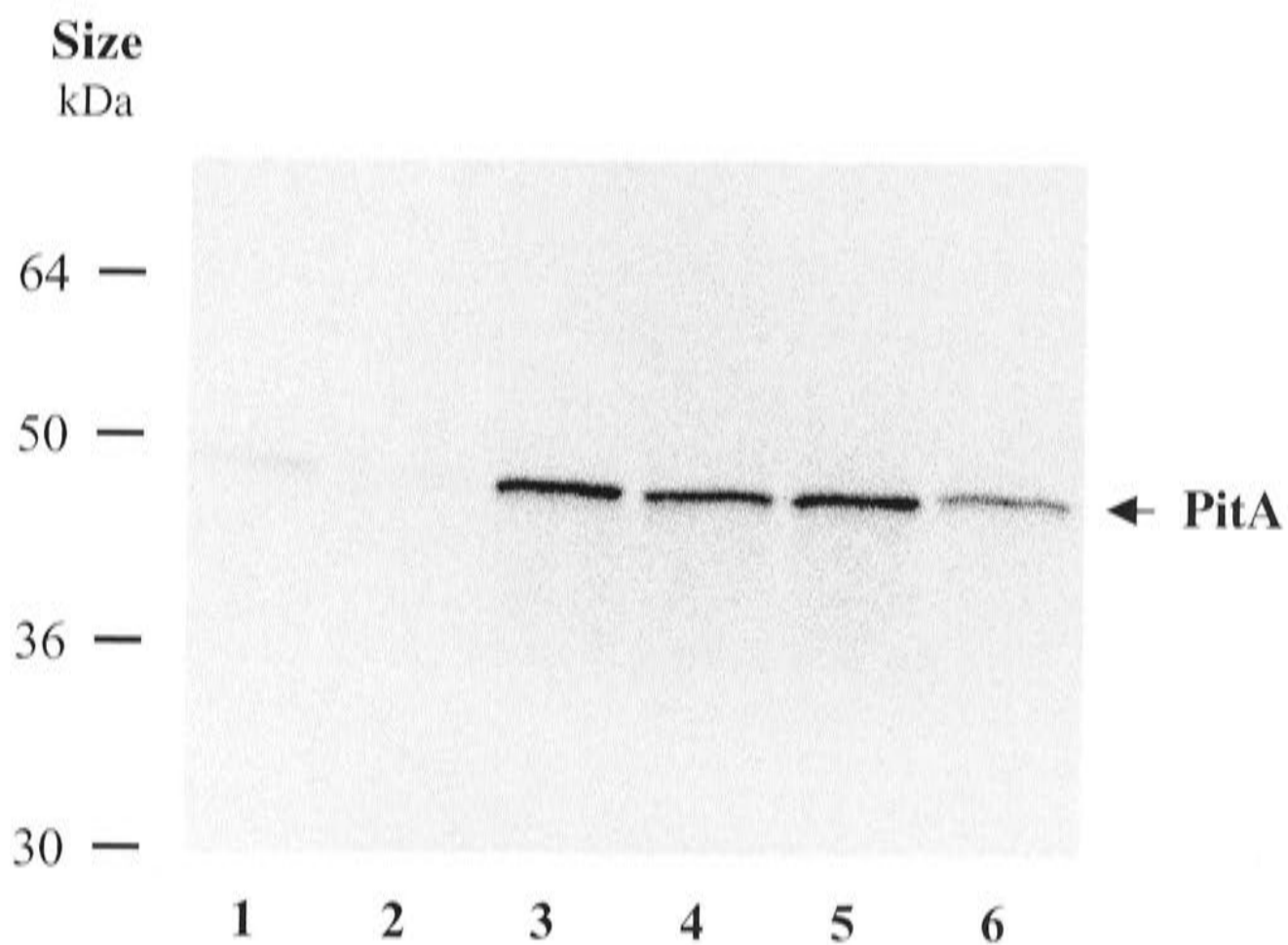
Changing Ala-213 to Asp (A213D) completely disrupts Pi transport by PitA. This mutation is near the beginning of putative  $\alpha$ -helix 6 in Models A and B (but within the transmembrane region of  $\alpha$ -helix 6 in the TopPred IV model). Very little PitA(A213D) protein is inserted in the cytoplasmic membrane, as is shown by Western blot using polyclonal antipeptide PitA antibody against the cytoplasmic membrane fraction of



## Figure 6.4

### Western blot analysis of PitA expression from mutants PitA(H225Q), PitA(D229N) and PitA(K232Q).

Membrane fractions of wild type strain AN248 and Pi auxotrophic strain AN3066 containing plasmids with the PitA mutations listed below were solubilised and separated by PAGE on a 10% polyacrylamide gel (Tris buffer system) and transferred onto PVDF membrane. Protein was visualised by incubation with polyclonal PitA antipeptide antibody, followed by incubation with alkaline phosphatase conjugated goat anti-rabbit secondary antibody and the application of Western blue alkaline phosphatase substrate.



Lane	Strain	Genome PitA phenotype	Plasmid PitA genotype
1	AN248	+	-
2	AN3066	-	-
3	AN3531	-	<i>pitA</i> <sup>+</sup>
4	AN3511	-	<i>pitA</i> (H225Q)
5	AN3512	-	<i>pitA</i> (D229N)
6	AN3513	-	<i>pitA</i> (K232Q)

## Figure 6.5

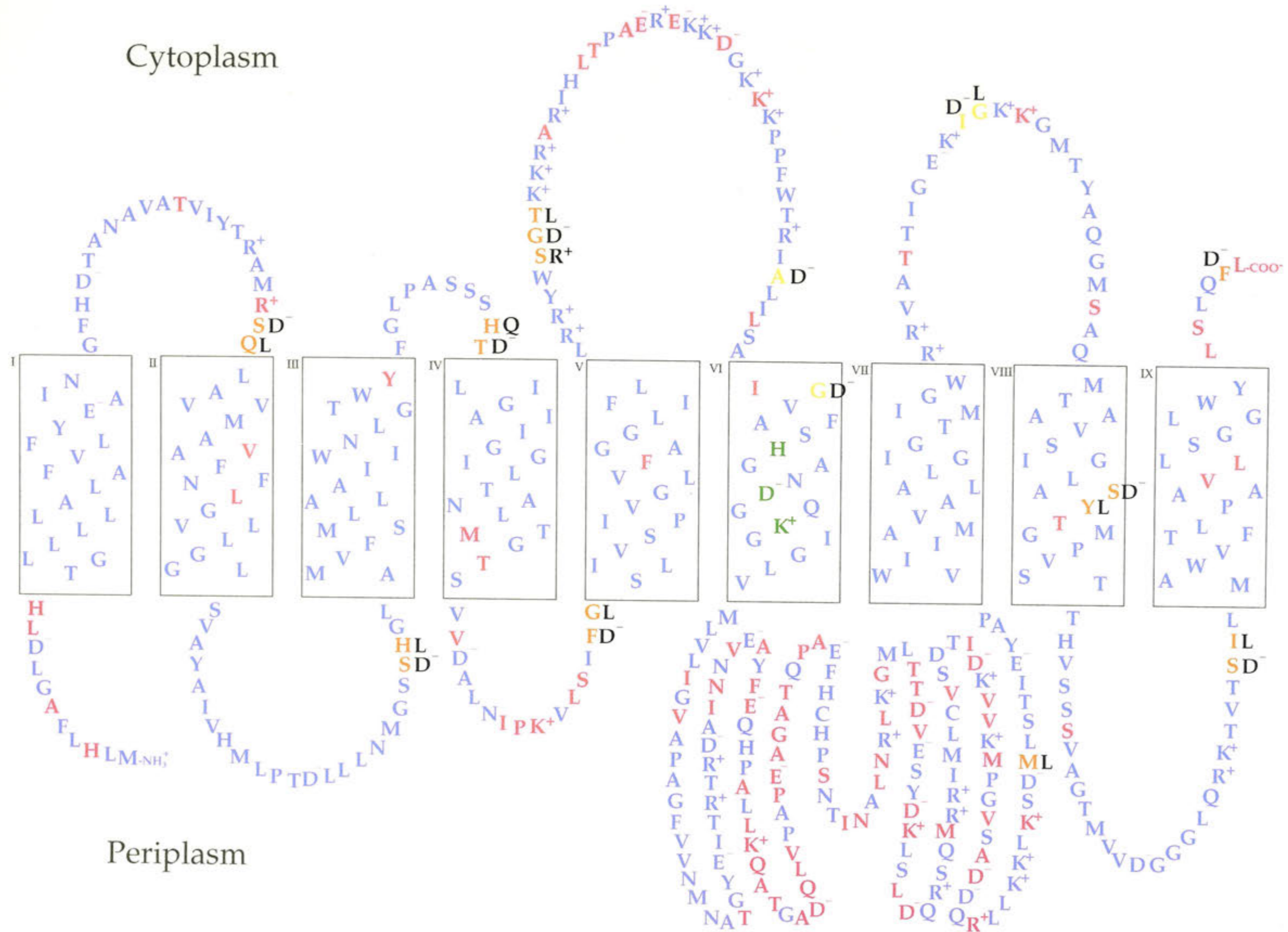
### Topological model of PitA showing site-directed mutagenesis changes.

These site-directed mutagenesis changes, shown in black, were the consequence of inserting *Bgl*III sites into PitA, with the exception of G220D (within TM 6), which is the *pitA1* mutation. Therefore each group of mutations occurs in a separate construct (e.g. PitA(S50D, Q51L), PitA(S91D, H92L). Different colours have been used to show the effect of these mutations on Pi uptake, and this is covered in more detail in Table 6.3.

- A Amino acid which is identical between PitA and PitB.
- A PitA amino acid which is different in the PitB sequence.
- A The amino acids in PitA which may be equivalent to the lactose permease residues R-319, H-322, and E-325, which have been shown to be important for lactose permease function.
- A The amino acid changed by site-directed mutagenesis. These changes allowed growth on 500 $\mu$ M Pi media.
- A The amino acid changed by site-directed mutagenesis. These changes did not allow growth on 500 $\mu$ M Pi media, therefore were detrimental to PitA.
- A The amino acid substitution created by site-directed mutagenesis.

The boxes represent the putative transmembrane  $\alpha$ -helices, which have been drawn to contain 19 amino acids, the predicted minimum number needed to cross the cytoplasmic membrane. The single letter amino acid code has been used.

Cytoplasm



Periplasm



**TABLE 6.3:** Presence of Pi uptake activity for various PitA mutations

<b>Putative location<sup>b</sup></b>	<b>PitA amino acid substitutions</b>	<b>Pi uptake<sup>a</sup></b>
Loop 1	S50D Q51L	+
Loop 2	S91D H92L	+
Loop 3	H122Q T123D	+
Loop 4	F156D G157L	+
Loop 5 early	S182R G183D T184L	+
Loop 5 late	A213D	-
Loop 6/ <u>7</u>	M375L	+
Loop 7/ <u>8</u>	I413D G414L	-
$\alpha$ -Helix 8/ <u>9</u>	S438D Y439L	+
Loop 8/ <u>9</u>	S472D I473L	+
C-terminus	F498D	+

a Growth on minimal media containing 500 $\mu$ M Pi.

b From the putative topological models: Model A and Model B (underlined where different).

strain AN3776 (Figure 6.6). While some membrane protein is assembled, it is at a similar level to that of the *pitA1* G220D mutation identified in K-10, and it is unlikely that this protein is functional.

A second mutated PitA protein, which contains Ile-413 changed to Asp and Gly-414 altered to Leu (I413D G414L), also displays no growth on Pi media. These residues are located within cytoplasmic loop 7 in Model A and periplasmic loop 8 of Model B (Figures 6.5, 6.2B). Membrane protein assembly of this PitA protein approaches that of wild type PitA expressed from a plasmid (Figure 6.6), so these changes cause a nonfunctional protein to be produced, without significantly altering assembly of the membrane protein.

## ***6.4 Investigation of substrate binding by PitA and PitB through the construction and analysis of chimeras***

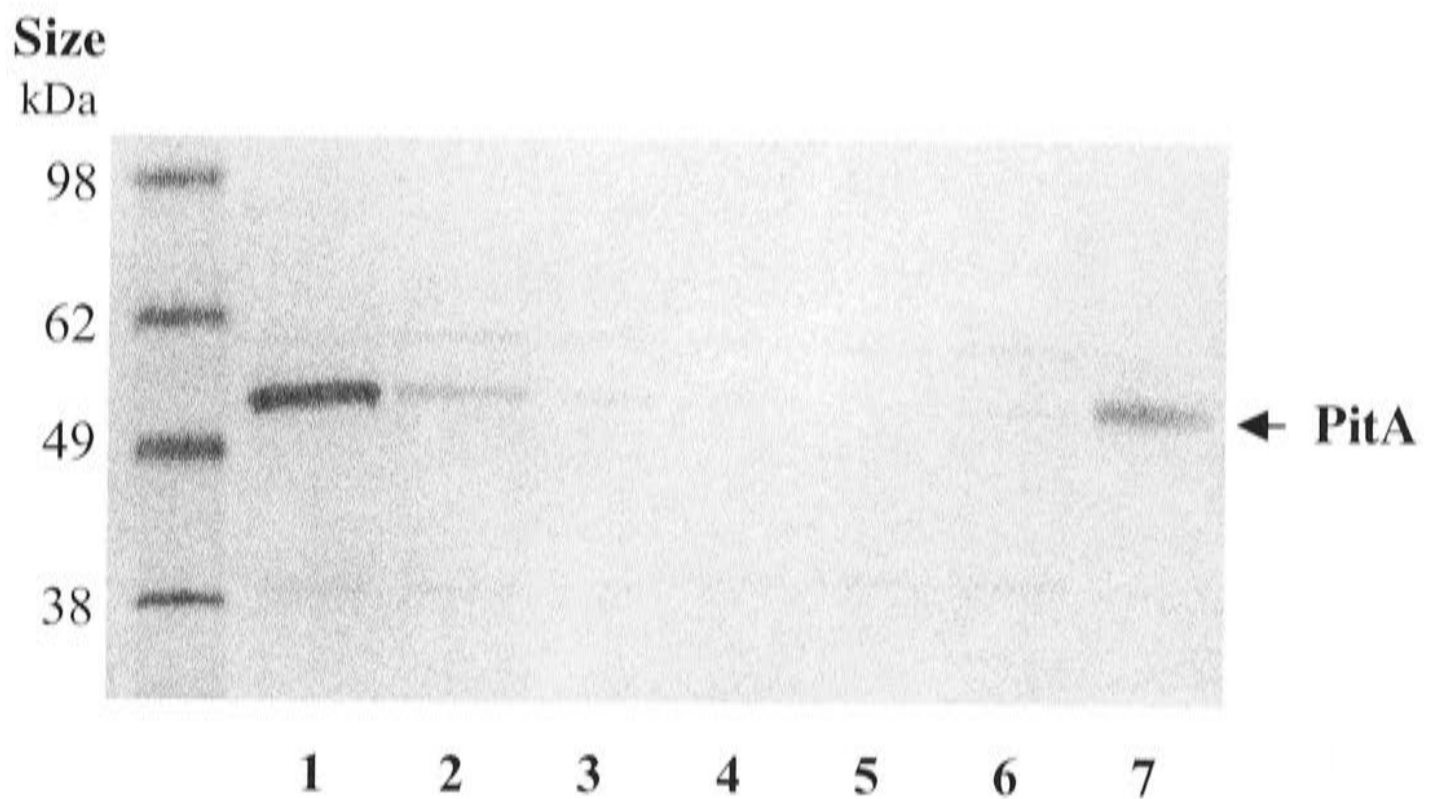
### ***6.4.1 Preparation of *pitA* and *pitB* with silent restriction endonuclease sites***

Chimeric PitA and PitB proteins were prepared for analysis of their kinetic parameters to see if specific regions of variable amino acid sequence between the two proteins could be linked with changes in Pi transport. PitA and PitB share 81% identity in deduced amino acid sequence, with most of this variability occurring within the putative hydrophilic loops. At the time that this experiment was designed the  $K_m^{app}$  of PitA appeared to be 14 fold lower than that of PitB (2 $\mu$ M versus 28 $\mu$ M, Table 3.3). Four silent restriction endonuclease sites were inserted into the *pitA* and *pitB* genes at Gly-110 (*BbrP1*), Arg-178 (*Sfu1*), Val-251 (*Hpa1*) and Leu-394 (*Sau1*) by site-directed mutagenesis, to allow the variable regions of PitA and PitB to be exchanged between the proteins (Figure 6.7, plasmid constructs - Figure 6.8). A *Hpa1* site and a *Sfu1* site were also removed from the PitB coding region by site-directed mutagenesis. (See Table 2.5 for the oligonucleotide sequences.)

## Figure 6.6

### Western blot analysis of the expression levels of various PitA mutants, using polyclonal PitA antipeptide antibody.

The membrane fractions of the wild type strain AN248 and the Pi auxotrophic strain AN3066 containing the plasmids listed below were solubilised and separated by PAGE on a 4-12% gradient polyacrylamide gel (MES buffer system) and transferred onto PVDF membrane. Protein was visualised by incubation with polyclonal PitA antipeptide antibody followed by incubation with alkaline phosphatase conjugated goat anti-rabbit antibody and application of Western blue stabilised alkaline phosphatase substrate.



Lane	Strain	Genome PitA phenotype	Plasmid PitA genotype
1	AN3531	-	<i>pitA</i> <sup>+</sup>
2	AN248	+	-
3	AN3514	-	-
4	AN3937	-	<i>pitA</i> (PCR- <i>pitA1</i> )
5	AN3938	-	<i>pitA</i> (G220D)
6	AN3776	-	<i>pitA</i> (A213D)
7	AN3778	-	<i>pitA</i> (I413D, G414L)





## Figure 6.7

### Topological models of PitA showing the silent restriction endonuclease sites used for PitA/PitB chimera construction.

These restriction endonuclease sites were inserted by site-directed mutagenesis, but cause no changes to the amino acid sequence of either PitA or PitB. The colours highlight the sequence between each unique restriction endonuclease site that could potentially be exchanged between PitA and PitB.

The chimeras tested in this study were:

 loop 5/helix 6 (Model A) exchanged between PitA and PitB.

 loop 6 (Model A) exchanged between PitA and PitB.

 loops 3/4 of PitA placed into PitB.

(See Table 6.4 for details.)

The boxes represent the putative transmembrane  $\alpha$ -helices, which have been drawn to contain 19 amino acids, the predicted minimum number needed to cross the cytoplasmic membrane. The single letter amino acid code is used.







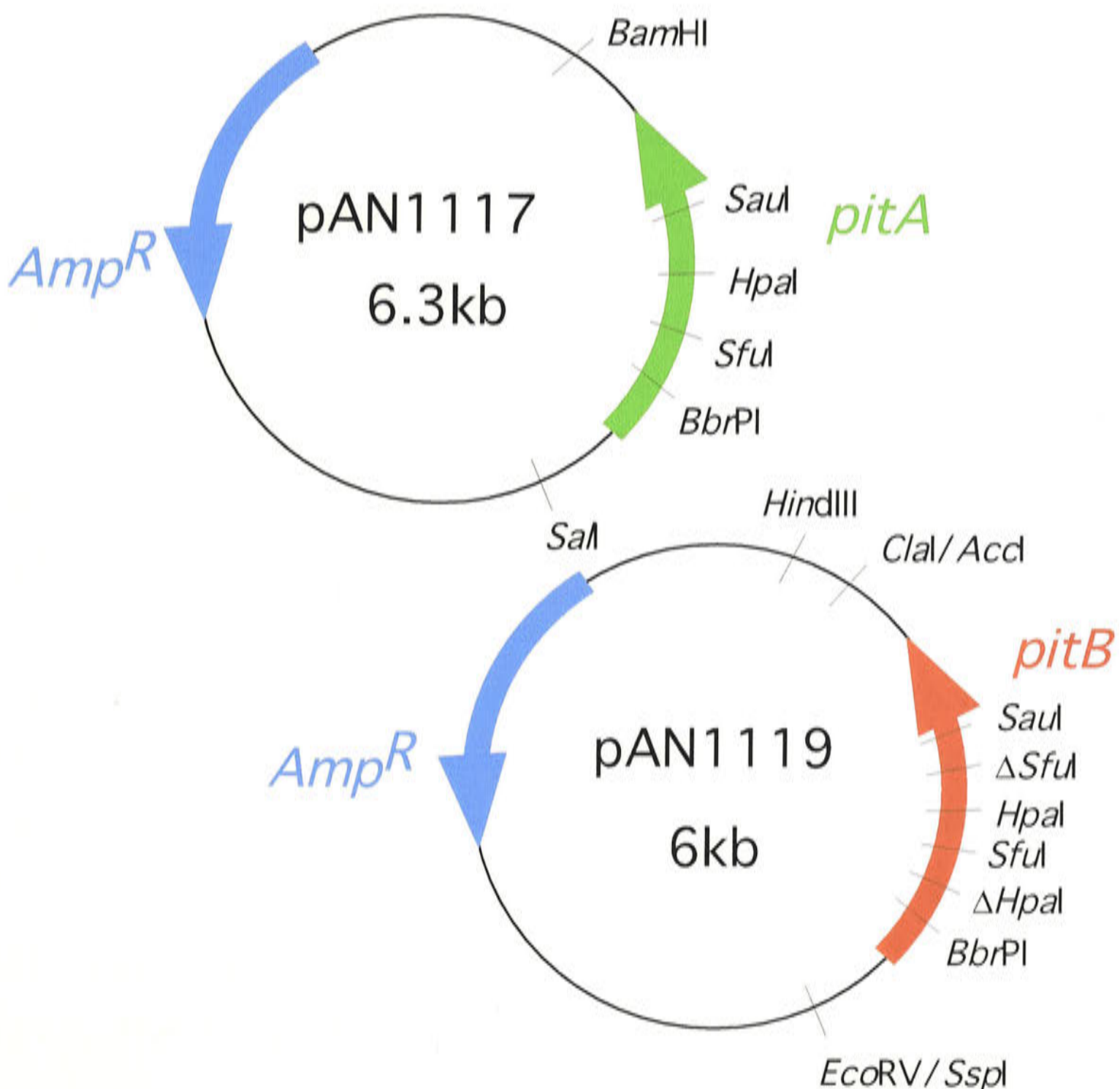
## Figure 6.8

### Plasmid diagrams for *pitA* and *pitB* containing silent restriction endonuclease sites.

Four unique silent restriction endonuclease sites were placed in equivalent positions in the wild type *pitA* and *pitB* genes, to allow the construction of *pitA/pitB* chimeras.

Site-directed mutagenesis was carried out on *pitA* in M13mp18 to insert the unique restriction endonuclease sites *BbrPI*, *SfuI*, *HpaI* and *SauI* (using oligonucleotides 93-77, 93-79, 92-150 and 92-151 respectively, See Table 2.5 for details). The *SalI/BamHI* fragment from the replicative form of this construct was then ligated into the *SalI/BamHI* sites of pBR322, creating pAN1117.

Similar mutagenesis experiments were carried out on *pitB* in M13mp18 to insert the unique restriction endonuclease sites *BbrPI*, *SfuI*, *HpaI* and *SauI* (using oligonucleotides 93-78, 93-80, 93-18 and 92-153 respectively, See Table 2.5 for details) and to remove an *SfuI* site (94-29) and a *HpaI* site (93-76). The *SspI/HindIII* fragment from the replicative form of this construct was then ligated into the *EcoRV/HindIII* sites of pBR322, removing a *HpaI* site from the genomic DNA upstream of *pitB*, to create pAN1119.




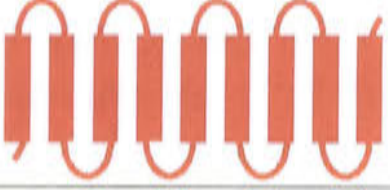


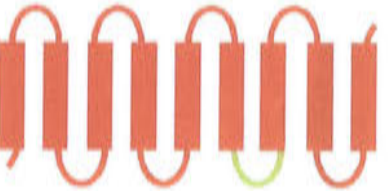


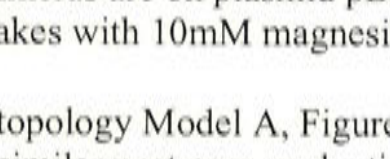



### 6.4.2 Production and analysis of chimeras

Most variability between PitA and PitB proteins occurs in putative hydrophilic loops 5 and 6 (Model A), which lie either side of  $\alpha$ -helix 6, containing the polar and charged residues which have been shown to play a role in Pit function. Putative loop 4 also contains a moderate number of variable residues. Therefore chimera production concentrated on exchanging these regions of the protein (Figure 6.7). Chimeras were prepared by exchanging equivalent sequences between the PitA plasmid pAN1117 and the PitB plasmid pAN1119. Thus *pitA* containing a *pitB* sequence will have a *pitA* promoter, and *pitB* containing an internal *pitA* sequence will be surrounded by *pitB* DNA control regions. These chimeras were then transformed into the Pi auxotrophic strain AN3066, checked for growth on Pi media and then analysed as previously described for  $K_m^{app}$  and  $V_{max}^{app}$  (Table 6.4). All chimeric proteins were functional and had  $V_{max}^{app}$  which were at least 50% of the PitA and PitB values.

This experiment was designed when the  $K_m^{app}$  values for PitA and PitB appeared significantly different (2 $\mu$ M and 28 $\mu$ M respectively - Table 3.3). However, all chimeras produced  $K_m^{app}$  values that were very low, with all except one being well under 10 $\mu$ M. This is an unexpected result. As the two plasmid constructs used for the preparation of *pitA* and *pitB* chimeras were slightly different from those used to determine the kinetic parameters for wild type PitA and PitB, new control plasmids were prepared and analysed for each gene. The orientation of *pitA* in vector pBR322 was reversed to form plasmid pAN920 by subcloning a *SalI/BamHI* wild type *pitA* fragment from pAN686 into the *SalI/BamHI* sites of pBR322 (Figure 6.9A). This had no effect on  $K_m^{app}$  and  $V_{max}^{app}$  (results not shown). The upstream DNA of the *pitB* gene was reduced from 1403 to 207 nucleotides (pAN1116 - Figure 6.9B), as this DNA had been deleted in the preparation of the *pitB* gene containing unique silent restriction endonuclease sites to remove a *HpaI* site. Unexpectedly this plasmid had a  $K_m^{app}$  change for PitB from 28 $\mu$ M to 6 $\mu$ M, explaining why all  $K_m^{app}$  determined for the chimeras were under 12 $\mu$ M. This  $K_m^{app}$  change is discussed in detail in Sections 5.4 and 5.6. Therefore the difference between the PitA and PitB  $K_m^{app}$  is only 3-fold,



**TABLE 6.4:** Kinetic parameters for PitA/PitB chimeras.

Chimera <sup>a d</sup>	$K_m^{app}$ <sup>b c</sup> μM	$V_{max}^{app}$ <sup>b c</sup> nmol Pi min <sup>-1</sup> mg dry weight <sup>-1</sup>
	PitA	1.9
	PitB (short upstream) <sup>e</sup>	6.0±0.5 (n=3)
	PitA loop 5/helix 6 into PitB	3.0±0.06 (n=3)
	PitB loop 5/helix 6 into PitA	46±11 (n=3)
	PitA loop 6 into PitB	5.7±1.6 (n=5)
	PitB loop 6 into PitA	40±5 (n=5)
	PitA loops 3+4 into PitB	12±3.6 (n=3)
	PitB loops 3+4 into PitA	2.3±0.4 (n=4)
	PitA loops 3+4 into PitB	54.5±0.3 (n=4)
	PitA loops 3+4 into PitB	4.7
	PitB loops 3+4 into PitA	58

a All chimeras are on plasmid pBR322 in host strain AN3066

b Pi uptakes with 10mM magnesium, pH7.0

c SEM

d From topology Model A, Figure 6.2

e Has a similar upstream nucleotide composition to the *pitB* chimeras



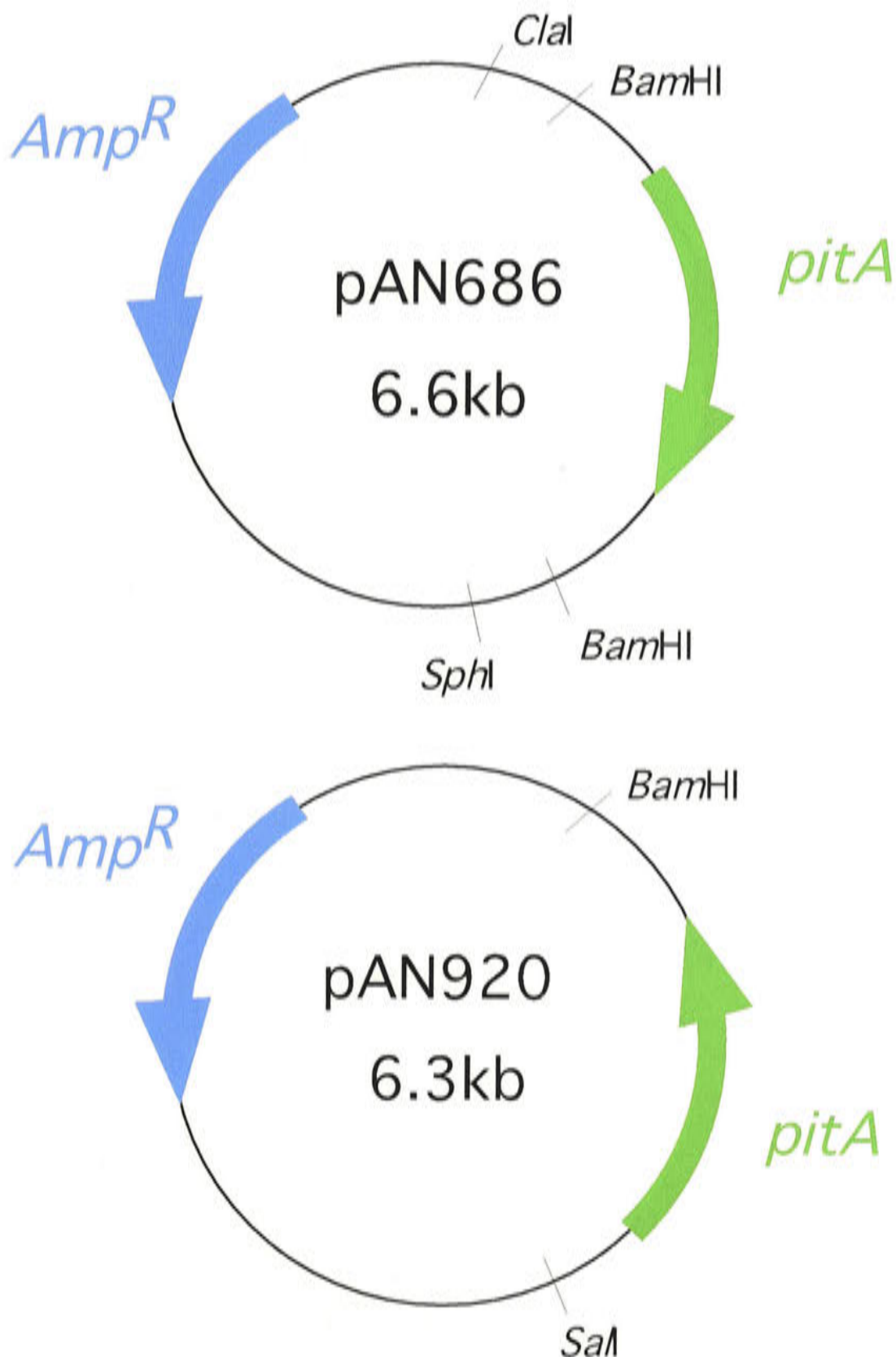
## Figure 6.9

### A Plasmid diagrams for *pitA* plasmids pAN920 and pAN686.

Creation of a control plasmid for pAN1117 (*pitA* containing silent restriction endonuclease sites) which was used for the construction of *pitA* chimeras containing sections of *pitB*. Wild type *pitA* from pAN686 was reversed in orientation (pAN920), effectively creating the same construct as pAN1117.

*pitA* was isolated from pCE27 (59) on a *ClaI/SphI* fragment, which was then ligated into the *ClaI/SphI* sites of pBR322 to form pAN686. (This plasmid effectively represents genomic *pitA* DNA on a *BamHI* fragment cloned into the pBR322 *BamHI* site, as pCE27 consists of a *pitA* *BamHI* fragment within cosmid vector pH79, which is a pBR322 derivative.) This plasmid was constructed by Dianne C. Webb.

*pitA* was removed from pAN686 on a *SalI/BamHI* fragment and ligated into the *SalI/BamHI* sites of pBR322, reversing the direction of wild type *pitA*



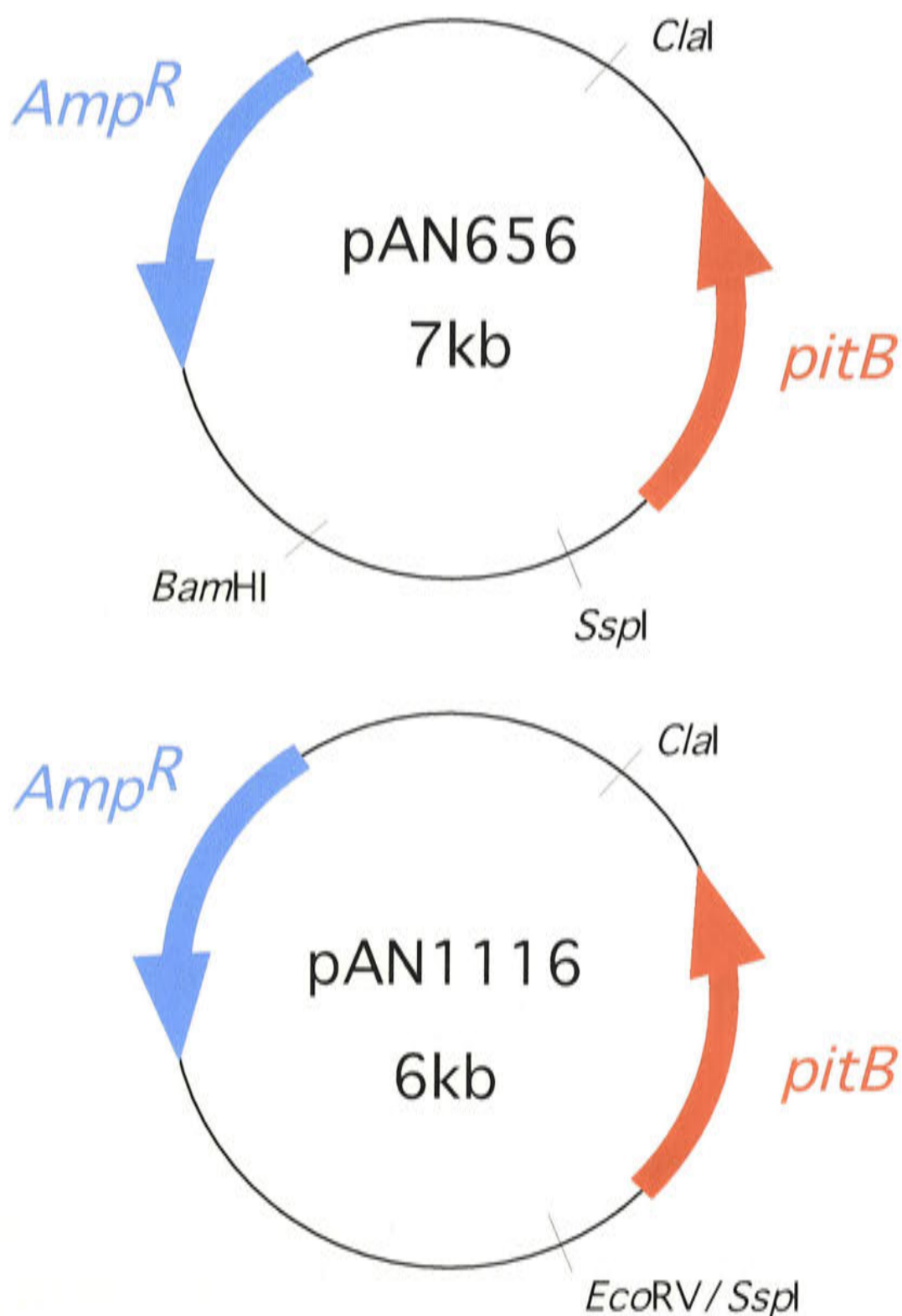
## Figure 6.9

### B Plasmid diagrams for pAN656 and pAN1116.

Creation of a control plasmid for pAN1118 (*pitB* containing the silent restriction endonuclease sites) which was used for the construction of *pitB* chimeras containing sections of *pitA*. 1196 nucleotides of genomic DNA upstream of wild type *pitB* (between *Bam*HI and *Ssp*I) was removed, effectively creating an identical construct to pAN1118, which has 207 rather than 1403 genomic nucleotides upstream of the putative *pitB* open reading frame.

*pitB* was isolated on a cosmid *Cl*aI/*Bam*HI fragment (89) and ligated into the *Cl*aI/*Bam*HI of pBR322, creating pAN656. This plasmid was constructed by Dianne C. Webb.

The *Ssp*I/*Bam*HI region of genomic DNA upstream of the *pitB* open reading frame was deleted by removing *pitB* from pAN656 on an *Ssp*I/*Cl*aI fragment, and ligating it into pBR322 digested with *Eco*RV/*Cl*aI, producing pAN1116.





making it much more difficult to extract information from the chimeric protein  $K_m^{app}$  values, as changes are much smaller and meaningful differences are more likely to be obscured by the variability inherent in these whole cell experiments.

The variable residues in 'loop 5' may influence the different  $K_m^{app}$  of PitA and PitB. Replacing PitB's 'loop 5/helix 6' (Figure 6.7) with the equivalent region from PitA reduces the  $K_m^{app}$  from 6 $\mu$ M to 3 $\mu$ M (approaching the PitA value of 2 $\mu$ M) (Table 6.4). Creating the complementary chimera by placing 'loop 5/helix 6' from PitB into PitA increases the  $K_m^{app}$  from 2 $\mu$ M to 5.7 $\mu$ M (similar to the PitB value of 6 $\mu$ M). Thus the characteristic  $K_m^{app}$  of each protein was transferred with the 'loop 5/helix 6' region, which contains 12 amino acid changes (9 within the putative loop 5, one within helix 6 and 2 conservative changes within the first part of loop 6 – see Figure 6.10). Placing PitB 'loop 6' into PitA has little effect on the  $K_m^{app}$ , while placing 'loop 6' from PitA into PitB increased the  $K_m^{app}$  and decreased the  $V_{max}^{app}$  in comparison to wild type PitB. This decrease in transporter efficiency and activity is characteristic of some chimeric constructs and can represent a general incompatibility between different regions of the chimeric protein, rather than indicating the transfer of a specific region of structure/function.

While further chimeras were prepared, they were not analysed in detail once the lowered  $K_m^{app}$  for PitB became apparent. An initial experiment for the insertion of 'loops 3 and 4' from PitA into PitB showed this change has little effect or may decrease the  $K_m^{app}$  slightly by comparison to wild type PitB (Table 6.4), but further experiments will be needed to confirm this.

## **6.5 Sequence comparison and domain analysis of PitA and PitB with other members of the PiT family**

While the Pho-4 phosphate transporter from *Neurospora crassa* was the only known protein which exhibited any sequence similarity to PitA and PitB when this project was started (21), the PiT family has grown to include more than 71 members (ProDom



## Figure 6.10

### Topological models of the 'loop 5/helix 6' region of PitA, showing the PitB amino acid side chains for the variable residues.

The 12 amino acid differences found between PitA and PitB in the 'loop 5/helix 6' region which may be involved in determining the substrate affinity of each Pi transporter. The topology is shown for Model A and Model B.

**A** 'loop 5/helix 6' region exchanged between PitA and PitB.

**A** PitA amino acids which differ from the PitB sequence.

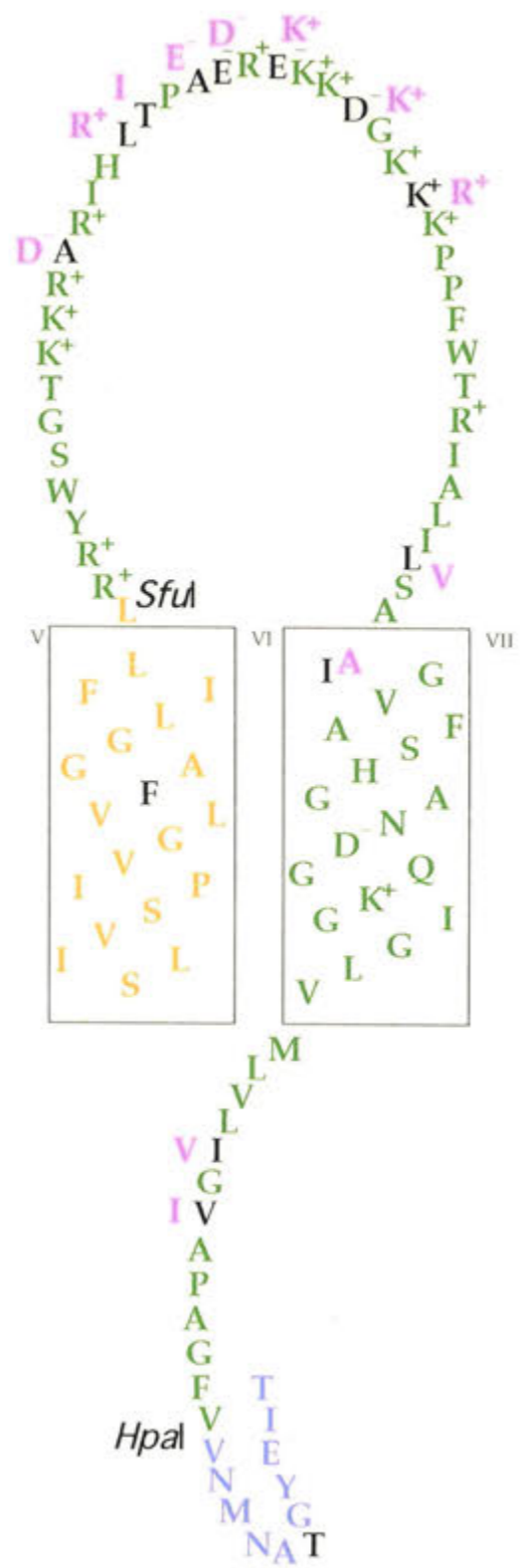
**A** PitB amino acids which differ from the PitA sequence.

```
PitA  RRY WSGTKKRaRI HltPaeReKK dGKkKPPFWT RIALIlSAiG
PitB  RRY WSGTKKRdRI HriPedRkKK kGKrKPPFWT RIALIvSAaG
```

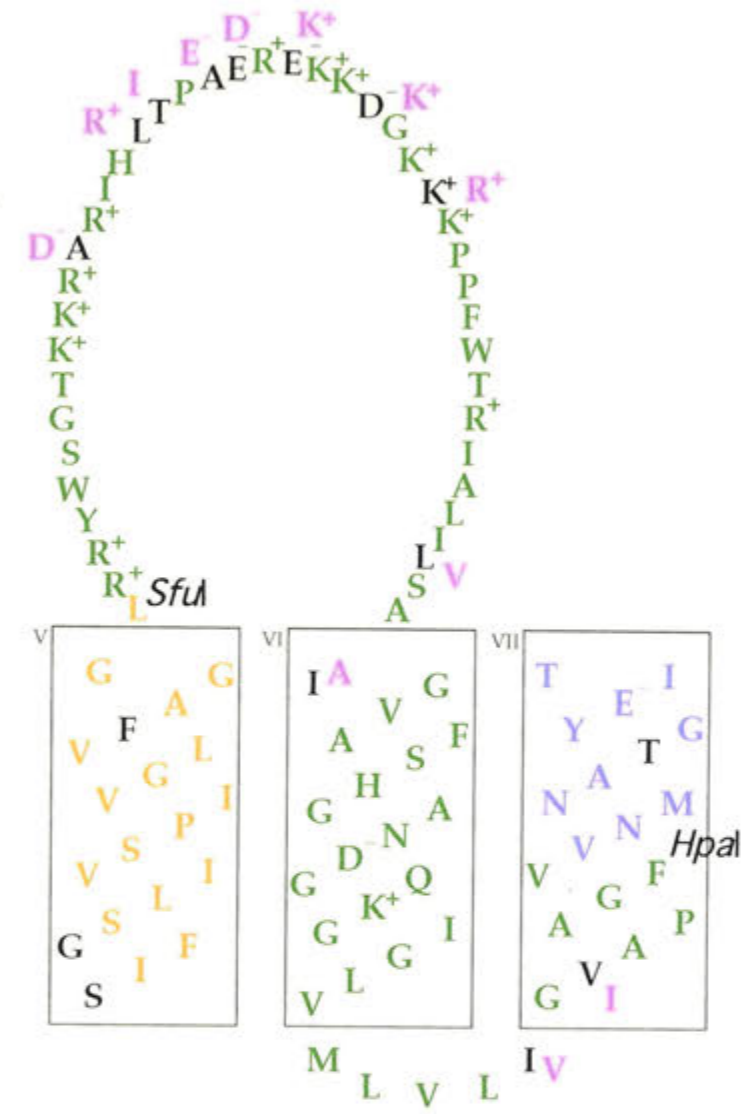
```
PitA  VAFSHGANDG QKGIGLVMLV LiGvAPAGFV
PitB  VAFSHGANDG QKGIGLVMLV LvGiAPAGFV
```

The boxes represent the putative transmembrane  $\alpha$ -helices, which have been drawn to contain 19 amino acids, the predicted minimum number needed to cross the cytoplasmic membrane. The single letter amino acid code is used.

Model A



Model B





database). It is defined by homology to the PD1131 domain, which is duplicated in most of these proteins (229), Transport Commission classification (226). Figure 6.11 shows two multiple sequence alignments using representative sequences from evolutionarily distant organisms over the N- and C-terminal regions, which demonstrate a high similarity across all kingdoms. The intervening central regions between these repeated domains have poor similarity and large insertions/deletions. Members of the PiT family originating from animals, worms, yeast and fungi contain an additional domain, PD7717, in this central region (ProDom database – (229)) and many plant proteins also have an extended N-terminus (47).

It is interesting to note that the C-terminal PD1131 domain of PitA and PitB has been split by a sequence (Figure 6.11B) which forms the large variable hydrophilic loop in the PitA and PitB topological models (Figure 6.12). The multiple sequence alignment in Figure 6.13 shows that this region is present in the putative Pit sequences of closely related Gram-negative bacteria, forming domain PD040191 (ProDom database), but is lacking in the Pit proteins of Gram-positive bacteria and other more distantly related species. Many members of the PiT family exhibit diversity in this location. The mammalian PiT-1 and PiT-2 transporters have a variable 13 amino acid recognition sequence for retroviral glycoprotein (Region A) in this position (51). Sections of the variable loop 5 of PitA and PitB are also lacking in the Gram-positive bacteria (Figure 6.13). *Rhizobium loti*, which contains at least two Pit genes, has proteins with and without these loops. (NCBI Blast analysis performed at SIB - (1)).

Salaun *et al* (229) have compiled 68 of these PD1131 domains (combining the N- and C-terminal domains) to identify four blocks of highly conserved amino acids. These have been numbered and highlighted on the multiple sequence alignments of Figure 6.11. Figure 6.12 relates the N- and C-terminal PD1131 domains to the proposed topology of PitA, and also locates the four highly conserved blocks of amino acids identified by Salaun *et al*, and the ‘signature sequence’ highlighted by Saier *et al* (226). The PitA residues His-225, Asp-229 and Lys-232, which greatly reduced PitA Pi uptake when mutated, are within the first highly conserved sequence of the C-terminal domain. These three residues are conserved in Pit proteins from closely related species (see mauve boxes in Figure 6.13), although Lys-232 becomes an Asn in more distantly







## Figure 6.11

Multiple sequence alignment of the N-PD1131 and C-PD1131 PiT domains from distant species.





protein	taxonomy	N-terminal region	C-terminal region
PitA   <i>E. coli</i>	bacteria	20-180	221-499
PitB   <i>E. coli</i>	bacteria	20-180	221-499
Pit   <i>R. meliloti</i>	bacteria	14-160	187-334
Pht2;1   <i>A. thaliana</i>	plant	162-316	432-587
PiT-1   <i>H. sapiens</i>	human/mammals	14-168	499-652
Pho4   <i>N. crassa</i>	fungus	14-173	432-590
YBR29C   <i>S. cerevisiae</i>	yeast	14-169	415-574

### A N-terminal PD1131 domain alignment

-  Highlights the PD1131 domain.
-  Highlights the conserved sequences identified by Salaun *et al* (229).
-  Highlights the strongly conserved residues within these sequences.
-  Highlights mutated amino acids referred to in Chapter 7.

The multiple sequence alignment was carried out using ClustalW with the following parameters. The Blosum scoring matrix was used, with Open and end gap penalties of 10, and extending and separating gap penalties of 0.05.

### B C-terminal PD1131 domain alignment

-  Highlights the PD1131 domain (which has been split in PitA and PitB).
-  Highlights the conserved sequences identified by Salaun *et al* (229).
-  Highlights the strongly conserved residues within these sequences.
-  Highlights mutated amino acids referred to in Chapter 6.

The ClustalW alignment (MacVector 6.0) was carried out as described above, except the extending and separating gap penalties were dropped to 0.03.

## A – N terminal PD1131 domain alignment

```

PitA|E. coli      LAFVLFYEAINGFHDANAVATVIYTRAMRSQLAVVMAAVFNFLGVLLGGLSVAYAIVHM
PitB|E. coli      LAFVLFYEAINGFHDANAVA AVIYTRAMQPQLAVVMAAFFNFFGVLLGGLSVAYAIVHM
Pit|R. meliloti   IAVALFFDFLNGLDHAANSIATIVSTRVLRPQYAVFWAAFFNFIAFLFFGLHVAETLGTG
Pht2;1|A. thaliana LLFGFYMAWNIGANDVANAMGTSVSGGALTIRQAVMTAAVLEFSGALLMGTHVTSMQKG
PiT-1|H. sapiens  FIIAFILAFSVGANDVANSFGTAVGSGVVTLRQACILASIFETTGSVLLGAKVGETIRKG
Pho4|N. crassa    TIFAALDAWNIGANDVANSWATSVAARSVTYLQAMILGSIMEFAGSVGVGARVADTIRTK
YBR29C|S. cerevisiae MLFAFLDAFNI GANDVANSFASSISSRSLKYWQAMVLAGLCEFLGAVLAGARVSGTIKNN
                  .      * :*.**: .: : : :      * . . . : . : * * ::

PitA|E. coli      LPTDLLLNMGSSHGLAMVFSMLLAAI IWNLGTWYFGLPASSSHTLIGAIIGIGLTNALMT
PitB|E. coli      LPTDLLLNMGSTHGLAMVFSMLLAAI IWNLGTWFFGLPASSSHTLIGAIIGIGLTNALLT
Pit|R. meliloti   IIDPGIVTP-----QVIFAALMGAITWNI VTWVFGIPSSSSHALIGGLVGAGLAKTGFS
Pht2;1|A. thaliana ILMANVFQGDMLLFAGLLS LAAAGTWLQVASYYGWVSTTHCIVGSMVGFGLVYGGAG
PiT-1|H. sapiens  IIDVNLNETVETLMAGEVVSAMVGS AVWQLIASFLRLPISGTHCIVGSTIGFSLVAIGTK
Pho4|N. crassa    VVDTTLFADDPALLMLGMVCAVVASSI YLTMATRFGLPVSTH SIMGGVIGMGIAAVGAD
YBR29C|S. cerevisiae IIDSSIFTNDPAVLMLTMTSALIGSSCW LTFATAIGMPVSTH SIVGGTIGAGIAAGGAN
                  :      :      . : .: : :      * * :* :*: .* .:

PitA|E. coli      GTSVVDALNIPKVL SIFGSLIVS PIVGLVFAGGLIFLLRRY
PitB|E. coli      GSSVMDALNLRVTKIFSSLIVS PIVGLVIAGGLIFLLRRY
Pit|R. meliloti   SIVWQG-----LLKTAGAI VMS PGIGFVLALLLVLIVS--
Pht2;1|A. thaliana AVFWSS-----LAKVASSWVIS PILGALVSFLVYKCIRRF
PiT-1|H. sapiens  GVQWME-----LVKIVASWFIS PLLSGFMSGLLFVLIRIF
Pho4|N. crassa    GVQWVGSSINDGVVSVFLAWVIAPGLAGAFAS IIFLVTKYG
YBR29C|S. cerevisiae GVVWGW S----GVSQIIASWFIAPILAGAI AIVFSISRF-
                  .      : . : .:.* :. .: :

```



## B – C terminal PD1131 domain alignment

```

PitA|E. coli          AFSHGANDGQKGIGLVMLVLIGVAPAGFVVNMNATGYEITRTRDAINNVEAYFEQHPALL
PitB|E. coli          AFSHGANDGQKGIGLVMLVLVGIAPAGFVVNMNASGYEITRTRDAVTNFEHYLQQHPPELP
Pit|R. meliloti       SLGHGGNDAQKTMGIIAVLL--FSQGYLG-----
Pht2;1|A. thaliana   SFAHGGNDVSNNAIGPLAAALSILQNGAAAG-----
Pit-1|H. sapiens     SFAHGGNDVSNNAIGPLVALWLIYKQGGVT-----
Pho4|N. crassa       SFTHGANDIANAIGPYATVFQLWKD GALPE-----
YBR29C|S. cerevisiae SFAHGANDVANATGPLSAVYVIWKTNTIG-----
:: **: ** : *

```

```

PitA|E. coli          KQATGADQLVPAPEAGATQPAEFHCHPSNTINALNRLKGM LTTDVESYDKLSLDQRSQMR
PitB|E. coli          QKLIAMEPPLPAASTDGTQVTEFHCHPANTFDAIARVKTMLPGNMESYEPLSVSQRSQLR
Pit|R. meliloti       -----
Pht2;1|A. thaliana   -----
Pit-1|H. sapiens     -----
Pho4|N. crassa       -----
YBR29C|S. cerevisiae -----

```

```

PitA|E. coli          RIMLCVSDTIDKVVKMPGVSADDQRLKLLKLS DMLSTIEYAPVWI IMAVALALGIGTMIG
PitB|E. coli          RIMLCISDTS AKLAKLPGVSKEDQNL LKLRSDMLSTIEYAPVWI IMAVALALGIGTMIG
Pit|R. meliloti       -----SEFYVPFVWVITCQAAIALGTLFG
Pht2;1|A. thaliana   -----GAEIVIPMDVLAWGGFGIVAGLTMW
Pit-1|H. sapiens     -----QEAATPVWLLFYGGVGICTGLVWV
Pho4|N. crassa       -----K GKADV PVWILVFGASCLVIGLWTY
YBR29C|S. cerevisiae -----AKSEVPVWV LAYGGVALVIGCWTY
:: * . :: : *

```

```

PitA|E. coli          WRRVATTIG EKIGKKGMTYAQGMSAQMTAAV SIGLAS YTGMPVSTTHVLSSSVAGTMVVD
PitB|E. coli          WRRVAMTIG EKIGKRGMTYAQGMAAQMTAAV SIGLAS YIGMPVSTTHVLSSAVAGTMVVD
Pit|R. meliloti       GWRIVHTMGSKITKL--NPMQGFCAETGGAITLFAATWLGIPVSTTHTITGAIIGVGAAR
Pht2;1|A. thaliana   GYRVIATIGKKITEL--TPTRGFAAEFAAASV VLFASKLGLPISATHTLVGAVMGVGFAR
Pit-1|H. sapiens     GRRVIQTMGKDLTPI--TPSSGFTIELASAFTVVIASNIGL PVSTTHCKVGSVAVGWIR
Pho4|N. crassa       GYNIMRNLGNRITLQ--SPSRGFSMELGSAVTVILATRLKLPVSTTQCITGATVGVGLCS
YBR29C|S. cerevisiae GYNI IKNLGNKMILQ--SPSRGFSIELAVAITVMATQLGIP TSTTQIAVGGIVAVGLCN
:: : :* . : . * : : * * : :* * : : . .

```

```

PitA|E. coli          G--GGLQRKTVTSILMAWVFTLPAAVLLSGGLYWLSLQFL-----
PitB|E. coli          G--GGLQRKTVTSILMAWVFTLPAAIFLSGGLYWIALQLI-----
Pit|R. meliloti       R-VSAVRWGLAGNIVVAWVITMPAAALISALCYFAADLVA-----
Pht2;1|A. thaliana   G-LNSVRAETVREIVASWLV TIVPG-ATLAVIYTWIFTKILSFVL----
Pit-1|H. sapiens     S-RKAVDWRLFRNIFVAWFVTVPVAGLFSAAVMALLMYGILPYV----
Pho4|N. crassa       GTWRTINWRLVAWIYMGWFITL PVAGIISGCLMGIIINAPRWGYSG---
YBR29C|S. cerevisiae KDLKSVNWRMVAVWCYSGWELTLP IAGLIAGIINGIILNAPRFGVEYQMT
: . * . * : * .

```



## Figure 6.12

### Topological model A of PitA, highlighting the N- and C-terminal PD1131 domains, and the highly conserved PiT family sequences.

The C-terminal PD1131 domain is split by a large variable region. Salaun *et al* (229) combined 68 N- and C-terminal domains of various PiT family transporters, to propose the presence of four highly conserved domains. Saier *et al* (226) proposed a signature sequence that is located in the first PD1131 domain. Russ *et al* (216) have identified a glycine sequence which seems to be important in helix packing. These sequences are highlighted on the topology model, with the proposed conserved sequences listed below. The boxes represent the putative transmembrane (TM)  $\alpha$ -helices, which have been drawn to contain 19 amino acids, the predicted minimum number needed to cross the cytoplasmic membrane. The single letter amino acid code has been used.

The models proposed for PitA and PitB were identical in topology, so only the PitA sequence is shown.

- A** PD1131 domains.
- A** Highly conserved amino acids within the conserved regions (see below).
- A** Variable amino acids within the conserved regions (see below).
- A** PitA sequence between domains.
- \*** “Signature sequence” for PiT family proposed by Saier *et al* (226) (see below).
- G** Helix packing motif.

### Conserved regions

	<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>
Consensus	G $\phi$ ND $\phi$	GxxxxGxxVxxT	P $\phi$ S	IxxxW $\phi$
PitA N-PD1131	GFHDT	GxxxxGxxVxxA	PAS	IxxxLI
PitA C-PD1131	GANDG	GxxxxWxxVxxT	PVS	IxxxWF

PitB is identical to PitA in the highly conserved amino acids for each region.

### Signature sequence

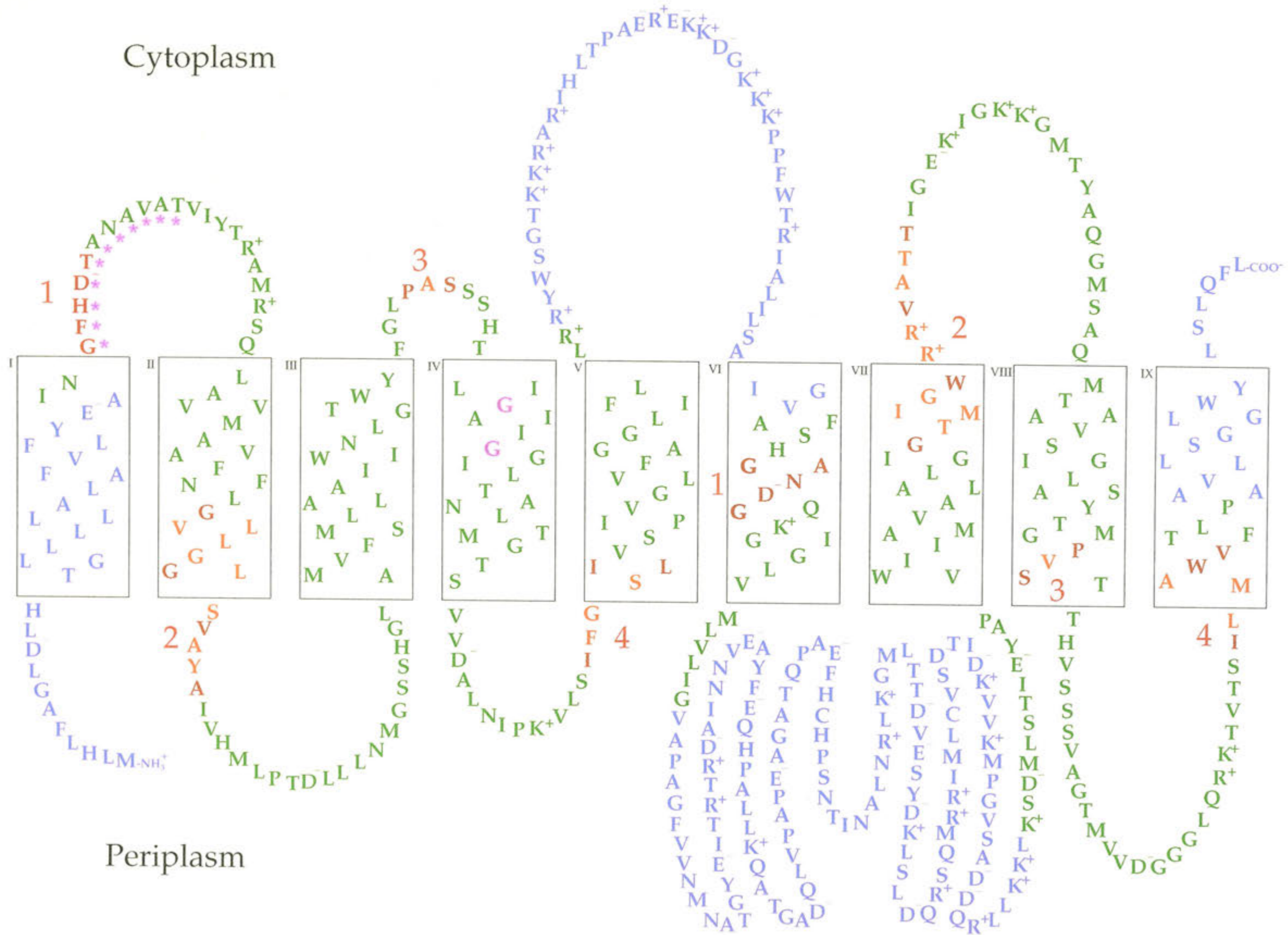
Consensus	<b>GFNDLGNSLAA</b>
	AHNIADAIGT
	G VS M S
	T F
	A W

PitA	<b>GFHDTANAVAT</b>
PitB	<b>GFHDTANAVAA</b>

### Helix packing motif

Consensus	GXXXG
PitA TM 4	GXXXG
PitA TM 6	GXXXG

Cytoplasm



Periplasm



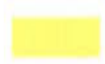


## Figure 6.13

Multiple sequence alignment of Pit proteins from species closely related to *E. coli*.

### A Protein details

This table includes information on the proteins and their organisms, such as accession details and basic taxonomic information .

### B C-terminal PD1131 domain alignment

-  Highlights the repeated PD1131 domains. (The C-terminal PD1131 domain has been split into two parts.)
-  This identifies the PD040191 domain, found in the first 8 sequences.
-  Highlights the conserved His (H), Asp (D) and Lys (K) residues that were changed by site-directed mutagenesis in PitA.

The multiple sequence alignment was carried out using ClustalW with the following parameters. The Blosum scoring matrix was used, with Open and end gap penalties of 10, and extending and separating gap penalties of 0.05. The single letter amino acid code has been used.



## A

Database /accession no.	Name/description	Organism	Taxonomy
PitA - Low-affinity inorganic phosphate transporter 1	sp P37308 PITA_ECOLI	<i>Escherichia coli</i>	Gram -ve bacteria γ-subdivision Enterobacteriaceae
PitA - Low-affinity phosphate transporter	tn AAL22449	<i>Salmonella typhimurium</i> - LT2	Gram -ve bacteria γ-subdivision Enterobacteriaceae
PitB - Low-affinity inorganic phosphate transporter 2	sp P43676 PITB_ECOLI	<i>Escherichia coli</i>	Gram -ve bacteria γ-subdivision Enterobacteriaceae
PIT? - Phosphate transport protein	tn CAC93429	<i>Yersinia pestis</i> (plague)	Gram -ve bacteria γ-subdivision Enterobacteriaceae
PIT? - Low affinity inorganic phosphate transporter	sp P57647 PIT_BUCAI	<i>Buchnera aphidicola</i> (subsp. <i>Acyrtosiphon pisum</i> )	Gram -ve bacteria γ-subdivision Buchnera
PA4292 - Probable phosphate transporter	tr Q9HWA8	<i>Pseudomonas aeruginosa</i>	Gram -ve bacteria γ-subdivision Pseudomonadaceae
PA0450 - Probable phosphate transporter	tr Q9I668	<i>Pseudomonas aeruginosa</i>	Gram -ve bacteria γ-subdivision Pseudomonadaceae
MLL3637 - Phosphate transporter	tr Q98FS6	<i>Rhizobium loti</i> ( <i>Mesorhizobium loti</i> )	Gram -ve bacteria α-subdivision Rhizobiaceae
SA0619 protein	tr Q99VV7	<i>Staphylococcus aureus</i> (strain N315)	Gram +ve bacteria Firmicutes Bacillus/Clostridium
Lmo0405 protein	tn CAC98484	<i>Listeria monocytogenes</i>	Gram +ve bacteria Firmicutes Bacillus/Clostridium
PITH - phosphate transport protein	tr Q9KZW3	<i>Streptomyces coelicolor</i>	Gram +ve bacteria Firmicutes Actinobacteria
PIT? - Probable phosphate transporter	tn CAD00327	<i>Listeria innocua</i>	Gram +ve bacteria Firmicutes Bacillus/Clostridium
PIT? - Probable low-affinity inorganic phosphate transporter	sp O34436 PIT_BACSU	<i>Bacillus subtilis</i>	Gram +ve bacteria Firmicutes Bacillus/Clostridium
PITH protein	PI tr Q9EUS1	<i>Streptomyces griseus</i> subsp. <i>griseus</i>	Gram +ve bacteria Firmicutes Actinobacteria
Phosphate permease - CAC3093	tr Q97EL7	<i>Clostridium acetobutylicum</i>	Gram +ve bacteria Firmicutes Bacillus/Clostridium
PIT? - Probable low-affinity phosphate transporter	tn CAD15015	<i>Burkholderia solanacearum</i> ( <i>Pseudomonas solanacearum</i> )	Gram -ve bacteria β-subdivision Ralstonia
PIT? - Probable low-affinity inorganic phosphate transporter	sp O30499 PIT_RHIME	<i>Rhizobium meliloti</i> ( <i>Sinorhizobium meliloti</i> )	Gram -ve bacteria α-subdivision Rhizobiaceae
PIT? Phosphate transporter	tr Q983D8	<i>Rhizobium loti</i> ( <i>Mesorhizobium loti</i> )	Gram -ve bacteria α-subdivision Rhizobiaceae



B

```

PITA - E. coli -----MLHLFAGLDLHT
PITA - S. typhimurium -----MLHLFAGLDLHT
PITB - E. coli -----MLNLFVGLDIYT
PIT? - Y. pestis -----MLHLFAGLDFHT
PIT? - B. aphidicola -----MLHLFYSYDLNH
PA4292 - P. aeruginosa -----MFDLFSGLDAWV
PA0450 - P. aeruginosa MTS DVL PAD AVASAAGNARKPHLEQRPG--RLTALVFFAVLATGLLFSAYS LMQD VDDVG
MLL367 - R. loti -----MVDVVSSEAGIPRPDHP LHS SSGSAKWFLPAFGVLV L VGVVYGYALSQDLAEAK
SA0619 - S. aureus -----MSYI-
Lmo0405 - L. monocytogenes -----MDTV-
PITH - S. coelicolor -----MDTF-
PIT? - L. innocua -----MEGM-
PIT? - B. subtilus -----MDIL-
PITH - S. griseus -----MDTF-
CAC3093 - C. acetobutylicum -----MLNS-
PIT? - B. solanacearum -----MHTLQ
sp|O30499|PIT_RHIME -----MDAT-
tr|Q983D8 -----MEAT-

```

```

PITA - E. coli G-----LLLLLALAFVLFYEAINGFHDTANAVATVIYTRAMRSQ LAVVMAAVFNFLG
PITA - S. typhimurium G-----LLLLLALAFVLFYEAINGFHDTANAVATVIYTRAMRSQ LAVVMAAVFNFFG
PITB - E. coli G-----LLLLLALAFVLFYEAINGFHDTANAVAAVIYTRAMOPOLAVVMAAFFNFFG
PIT? - Y. pestis G-----LMLVLALLFVLFYEAINGFHDTANAVATVIYTRAMRSQ IAVVMAGVFNFLG
PIT? - B. aphidicola S-----LLVFLALFFVLFYEAINGFHDTANAVSTLIYTRAMSAHVAVIMSGIFNFLG
PA4292 - P. aeruginosa G-----ISLVLALAFVLTFFEFINGFHDTANAVATVIYTKAMSPYRAVMLSGVFNFLG
PA0450 - P. aeruginosa STITWT P FLL LGVALLIALGFEFVNGFHDTANAVATVIYSHSLPPHFVWWSGCFNFLG
MLL367 - R. loti T-----VPWILLGIAL LIALGFEFVNGFHDTANAVATVIYTRSM P AEFVWWSGAFNFLG
SA0619 - S. aureus -----IIVTIAVVI FSLIFDFINGFHDTANAVATAVSTRALTPKTA I LMAAVMNFIG
Lmo0405 - L. monocytogenes -----ILITV IIVIVGLAFDFINGFHDIANAVATS ISTRALKPRVAIGIAAVMNFILG
PITH - S. coelicolor -----ALVVTIGVALFFTYTNGFHDSANAIATSVSTRALTPRAALAMA AVMNLAG
PIT? - L. innocua -----FLITLVIVLAALAFDLINGFHDTANAIATSVSTRALKPRHAI I LAAVMNFVG
PIT? - B. subtilus -----LILTILIVICALAFDFINGFHDTANAIATSVSTRALKPRHAI I LAAVMNFVG
PITH - S. griseus -----ALIVTIGVALGFTYTNGFHDSAQRIATSVSTRALTPRAALAMA AVMNLAG
CAC3093 - C. acetobutylicum -----AVLITIIIVILAI S FDFLNGFHDSANAIATSVSTRVLSMKSAVIMSAILNFLG
PIT? - B. solanacearum I-----SLWVVVLLVALAILFDFMNGFHDAANS IATVVSTGV LKPOQAVAMAAACNVIA
PIT? - R. meliloti L-----AFPLLVLGLIAVALFFDFLNGLHDAANS IATIVSTRVLRPOYAVFWAAFFNFIA
PIT? - R. loti I-----AFPVLIALVAVALFFDFLNGLHDAANS IATIVSTRVLRPOYAVLWAAFFNFIA

```

```

: : : *:* * : : : : : * : . * .

```

```

PITA - E. coli VLLGGLSVAYAI VHMLPTDLLLNMGSSHGLAMVFSMLLAAI IWNLGTWYFGLPASSSHTL
PITA - S. typhimurium VLLGGLSVAYAI VHMLPTDLLLNMGSAHGLAMVFSMLLAAI IWNLGTWYFGLPASSSHTL
PITB - E. coli VLLGGLSVAYAI VHMLPTDLLLNMGSTHGLAMVFSMLLAAI IWNLGTWFFGLPASSSHTL
PIT? - Y. pestis VMLGGLSVAYAI VHLLPTDLLLNVSSTHGLAMVFSMLLAAI IWNLGTWYFGIPASSSHTL
PIT? - B. aphidicola VLLGG LTVAYTIVHLLPNDLLLNATSKNALAMVFSMLLAAI IWNLSTWYFCLPASSSHSL
PA4292 - P. aeruginosa VLLGGVGVAYAI VHLLPVELLINVNTGHGLAMVFSLLAAAI AWNLGTWYFGIPASSSHTL
PA0450 - P. aeruginosa VLLSSGAVAFGI IALLPVELILQVGS SAGFAMVFALLIAAI LWNLGTWVWGLPASSSHTL
MLL367 - R. loti VLTSSGAVAFGILSLLPVELILQVGS SSGFAMVFALLVAAI LWNLGTWVWGLPASSSHTM
SA0619 - S. aureus ALTF-TGVAGTIT----KDIVDPFKLENGLVVLAAILAAI IWNLATWYFGIPSSSSHAL
Lmo0405 - L. monocytogenes AISF-TGVAESLT----KSIVDPFSLNNGEFVVL CGLIAAVIWNLMTWLVGMPSSSSHAL
PITH - S. coelicolor AFMG-SGVAKTVS----EGVIETPEGSKGMGILFAALVGAI VWNLITWYFGLPASSSHTL
PIT? - L. innocua AISF-TGVAKTIT----KDIVNPFDL DHGELVILAALLSAI AWNLITWYFGIPSSSSHAL
PIT? - B. subtilus AMTF-TGVAKTIT----KDIVDPYTLENG SVVILAALLAAI AWNLITWYFGIPSSSSHAI
PITH - S. griseus AFLG-QGVAKTVS----EGLIATPVGQKGMGILFAALVGAI IWNLITWYFGIPSSSSHTL
CAC3093 - C. acetobutylicum AFMS-DKVAKTVG----DGIIDP--SKIVPSVI IVALIAAI IWNLITWYFGIPSSSSHTL
PIT? - B. solanacearum IFIFHLKVATTVG----RDTIDP--SIVDHYVIFGALVGAI AWNLITWYFGIPSSSSHTL
PIT? - R. meliloti FLFFGLHVAETLG----TGIIDP--GIVTPQVIFAALMGAI TWNIWTVFGIPSSSSHTL
PIT? - R. loti FMFFGLHVAETVG----KGIVDV--SIVTPAVIFSA LVGAI VWNITWYFGIPSSSSHTL

```

```

: ** : : : : : * : * : * : * : * : * :

```







PITA - *E. coli* EFHCHPSNTINALNRLKGLMLTDDVESYDKLSLDQRSQMRRLMCLVSDTIDKVVK--MPGV  
PITA - *S. typhimurium* EFHCHPANTINALNRAKGLMLAN-VESYDKLSVEQRSQRRIMLCISDTTDKVVK--LPGV  
PITB - *E. coli* EFHCHPANTFDAIARVKTMLPGNMESYEPLSVSORSQRRIMLCISDTSAKLAK--LPGV  
PIT? - *Y. pestis* QFHCDTSRAMVAINHQAALLTD-LKSYDDLTVDQRSQMRLLMCVAETAGAVAK--LPET  
PIT? - *B. aphidicola* VVRSQSYNSIKNIKNTKLLLN-INSYNDLSIKKRFQLRHYLLCISDSIDKKNV--SSDI  
PA4292 - *P. aeruginosa* FYRCDPKQTEPTINALLRDLRG-VPSYNDLDADERVQVRRYLLCLDDTAKKVGK--LSDL  
PA0450 - *P. aeruginosa* ASPELIPALAALTGSIGQEVKG-YGAFSRVPAEAMGNVRNDMYLTSEAIRLMDKDGVGRF  
MLL367 - *R. loti* WNDQTTAALQTYIHNTTAGLQP-YATVDNVPTDLVSNARNDIYLIGEALKLIDKKKLLPM  
SA0619 - *S. aureus* -----N---  
Lmo0405 - *L. monocytogenes*-----K---  
PITH - *S. coelicolor* -----D---  
PIT? - *L. innocua* -----T---  
PIT? - *B. subtilus* -----T---  
PITH - *S. griseus* -----G---  
CAC3093 - *C. acetobutylicum*-----S---  
PIT? - *B. solanacearum* -----S---  
PIT? - *R. meliloti* -----G---  
PIT? - *R. loti* -----G---

PITA - *E. coli* SADDQRLKLLKSDMLSTIEYAP-VWIIAVALALGIGTMIGWRRVATTIGEKGKGGMT  
PITA - *S. typhimurium* SSDDQRLKLLKTDMLSTIEYAP-VWIIAVALALGIGTMIGWRRVATTIGEKGKGGMT  
PITB - *E. coli* SKEDQNLKLLKRSMLSTIEYAP-VWIIAVALALGIGTMIGWRRVAMTIGEKGKGGMT  
PIT? - *Y. pestis* SAENRRFLNLRDLDLDTVEYAP-TWIIAVALALSLGTMIGWRRVAVTIGEKGKGGMT  
PIT? - *B. aphidicola* CSKDKRFLIHSKKVILKTIEYAP-MWIIILVALSLSIGTMIGWKRIVVTIGEKGKGGMT  
PA4292 - *P. aeruginosa* PAREKADLEKLRKDLTATTEYAP-FWVIIAVALALGIGTMIGWKRIVVTVGEKIGKGGMT  
PA0450 - *P. aeruginosa* DADTRGKLOAFKQKIDATKFIPLWVKIAVAIALGLGTMIGWKRIVVTVGEKIGKGGMT  
MLL367 - *R. loti* EAADLKAVTDYHKAVDNATKFIPLWVKIAVAIALGLGTMIGWKRIVVTVGEKIGKGGMT  
SA0619 - *S. aureus* -----DGSVEPOLWVKFACATAMGLGTAIGGWKIIKTVGGNIMK--IR  
Lmo0405 - *L. monocytogenes*-----ESAGIPFWVQVSCAASMAIGSSVGGYRIKTVGTRKIMK--IT  
PITH - *S. coelicolor* -----YGDPIPLWVKLACALMLSLGTIYAGGWIRMTLGRKIIIE--LD  
PIT? - *L. innocua* -----TDDVQLWVQVSCAIAAIGTSIGGWKIIKTVGGKIMK--IK  
PIT? - *B. subtilus* -----S-ANDIPTWVQFACATAMGLGTSIGGWKIIKTVGGKIMK--IR  
PITH - *S. griseus* -----PNDEIPLWVKIACALMLSLGTIYAGGWIRMTLGRKIIIE--LD  
CAC3093 - *C. acetobutylicum*-----S-FTVPLWVKAACAIAMAFGTSFGGYRIKTMGMNMAK--LA  
PIT? - *B. solanacearum* -----ATASTPPIWVIVCCYVAIGMGTLLGGWRIVRTMGQKITK--LK  
PIT? - *R. meliloti* -----SEFYVPLWVITCQAAIALGTLFGGWIRVHTMGSKITK--LN  
PIT? - *R. loti* -----EKFYVPLWVLTCSALALGTLFGGWIRVHTMGSKITR--LN  
\* : : : \* \* : : \* \* : : . :

PITA - *E. coli* YAQGMSAQMATAAVSIG-LASYTGMPVSTTHVLSSVAGTMVVD-GGGLQRKTVTSILMAW  
PITA - *S. typhimurium* YAQGMSAQMATAAVSIG-LASYTGMPVSTTHVLSSVAGTMVVD-GGGLQRKTVTSILMAW  
PITB - *E. coli* YAQGMATAAVSIG-LASYIGMPVSTTHVLSSAVAGTMVVD-GGGLQRKTVTSILMAW  
PIT? - *Y. pestis* YAQGVSAQMATAAVSIG-IASYTGMPVSTTQVLSSAVAGTMLVD-GGGVQSKTVKSIMLAW  
PIT? - *B. aphidicola* YAQMSAQITASFISIG-IASYTGIPVSTTHILSSVAGTMLID-GDGIQTKTIKNIALAW  
PA4292 - *P. aeruginosa* YAQMSAQITAAAAIG-MANIYSLPVSTTHVLSSVAGTMVAN-KSGLHGGTVRNILMAW  
PA0450 - *P. aeruginosa* YAQGASAETVAMLTIG-AADLYGLPVSTTHVLSSVAGTMAAN-GSGLQWKTIRNLLMAW  
MLL367 - *R. loti* YGQAAAELVAMVTIG-MADRLGLPVSTTHVLSSVAGTMAAN-GSGLQWSTVRNLLLAW  
SA0619 - *S. aureus* PANGAAADLSSALTIPLVASSLHFLSTTHVSSSILGVGASNRAGVKWSTAQRMIITW  
Lmo0405 - *L. monocytogenes* PVTGVASDLSLSSVIM-TATLIHLPVSTTQVIDSSIMGVGTANHKKEVNRWRTGKNMVVTW  
PITH - *S. coelicolor* PPOGFAAETTGASIMFGSAFIFHAPISTTHVITSAIMGVGATRRVNAVWGVAKNIIMGW  
PIT? - *L. innocua* PVNGVAADLSSVIIIF-GATFIHLPVSTTHVISSSILGVGTAHRVKGVKWDTAQRMIITW  
PIT? - *B. subtilus* PVNGVSADLTGAIIIF-GATFIHLPVSTTHVISSSILGVGASHRVKGVNWGTAKRISLHG  
PITH - *S. griseus* PPOAFAAETTGASIMFGSAFLFHAPISTTHVITSAIMGVGATKRVNAVWGVAKNIILGW  
CAC3093 - *C. acetobutylicum* PVNGFAAETGAAVIF-SATMFHAPVSTTQIISTSIMGVAASKRISSVRWVAKDILIAW  
PIT? - *B. solanacearum* PVGGFCAETGGALTLF-FASALGVPVSTTHTITGAIIVGVGAOAKASAVRWGVAGNIVVAW  
PIT? - *R. meliloti* PMQGFCAETGGAITLF-AATWLGIPVSTTHTITGAIIGVGAARRVSAVRWGLAGNIVVAW  
PIT? - *R. loti* PMQGFCAETGGAITLF-AATWLGVPVSTTHTITGAIIGVGAARRVSAVRWGIAGNIVIAW  
. : : . : \* \* : : \* \* : : . :

PITA - *E. coli* VFTLPAAVLLSGGLYWLSLQFL-  
PITA - *S. typhimurium* VFTLPAAILLSGVLYWLSLKI-  
PITB - *E. coli* VFTLPAAIIFLSGGLYWIALQLI-  
PIT? - *Y. pestis* VFTLPIIFLSGALYWIALKFI-  
PIT? - *B. aphidicola* IFTLPSMMLSSFLYWIALFLI-  
PA4292 - *P. aeruginosa* VFTLPTSMALSAGLFWLASQFI-  
PA0450 - *P. aeruginosa* VFTLPAAILLSASLYWLLTRLF-  
MLL367 - *R. loti* VFTLPCSITLAFVLFIVFRQVF-  
SA0619 - *S. aureus* VITLPI SALLAGLLFYILNLF-  
Lmo0405 - *L. monocytogenes* FITLPLAGLLAAVVYWISAAIFL  
PITH - *S. coelicolor* FITMPAAAVVAAVSFWIVNLAVL  
PIT? - *L. innocua* VITLPI SATTIAALIFYVLRFIL  
PIT? - *B. subtilus* SSRFRFQR---HLVPSPTLF---  
PITH - *S. griseus* FITMPAAALVAALSYGAVLLFLG  
CAC3093 - *C. acetobutylicum* FITMPVCAGIAALLALFIR----  
PIT? - *B. solanacearum* VLTIPASAFMAAIAWWIGRQLL-  
PIT? - *R. meliloti* VITMPAAALISALCYFAADLVA-  
PIT? - *R. loti* IVTLPAAAAISALTYFVTRQ---  
:



related species (Figure 6.11B – conserved region 1). It is interesting to note that these amino acids are in the small section of the C-terminal PD1131 domain which is immediately followed by a region of high variability between many PiT family members (Figures 6.12, 6.11B).

## **6.6 Discussion**

Single polar amino acid substitutions of residues His-225, Asp-229 or Lys-232 in PitA produced assembled proteins with greatly decreased rates of Pi uptake, indicating the importance of these charged residues in Pit structure/function. PitA/PitB chimeras show that the adjacent putative loop 5 region plays a role in determining the affinity of these transporters for Pi. Therefore this putative ‘loop 5/helix 6’ section of the protein is important for function and is probably involved in the mechanism of Pi transport. However, where these regions are in relation to the membrane is not known because the topology of PitA and PitB is uncertain.

The individual PitA mutations of H225Q, D229N and K232Q reduce initial rates of Pi transport to less than 3% of wild type activity. This triad has a resemblance to the arrangement of functionally important residues on lactose permease  $\alpha$ -helix 10 (Figure 6.1) consistent with placement of these residues within TM 6 of the putative topological model (Figure 6.2). While secondary structure prediction programs did not place this triad of amino acids within the membrane, this modification was made to our models because the unfavourable insertion of charged residues into a hydrophobic environment can be mediated by interactions with other charged residues, and such interactions cannot be predicted by these programs.

There is some evidence suggesting that Lys-232 may play a different role in Pit structure or function to His-225 or Asp-229. The PitA K232Q mutation has decreased levels of membrane protein assembly when compared with PitA mutants H225Q and D229N, or PitA expressed on plasmids (Figure 6.4). It is interesting to note that Kaback’s most recent model for lactose permease predicts that Lys-319 forms a salt bridge with Asp-240, while His-322 and Glu-325 play more direct roles in lactose/proton transport (Section 1.9.2). Evidence indicating the possibility of a Lys-319/Asp-240 salt bridge came from the double neutral mutation of K319C/D240C

producing activity and a normal amount of protein assembly (222). Lys-319 may still be important for modulating the proton transfer capabilities of nearby residues (e.g. influencing pKa) (209). Thus it is possible that Lys-232 in PitA may form part of a salt bridge that is disrupted by the K232Q mutation. While the Pit HxxxDxxK triad is conserved in the Pit proteins of species closely related to *E. coli* (mauve boxes -Figure 6.13) more distantly related organisms, such as eukaryotes, have the sequence HxxxDxxN (conserved region 1 - Figure 6.11B), where Lys has been replaced by the polar amino acid Asn. This does not necessarily mean that the role of this residue has changed. Recent experiments on models of inter-helical interactions show that Asn residues can form strong hydrogen bonds within membranes, and that these interactions are strong enough to form dimers that can be isolated by SDS gel analysis (20, 36, 308). Perhaps organisms containing an Asn in this strongly conserved sequence have replaced a salt bridge with a hydrogen bond. Further experimentation is required before any conclusions can be made.

It is also possible that His-225, Asp-229 and Lys-232 may be functionally important, but located within the periplasm. Topology prediction programs and hydropathy profiles place these amino acids at the periplasmic surface of PitA/PitB (Figure 3.15), and predictive methods also place the equivalent residues of other members of the PiT family, such as mammalian transporter PiT-2, in a similar location (229). Pit models A and B (Figure 6.2) can easily be adjusted place these residues in the periplasm by pulling the last third of putative loop 5 into the membrane. If so, the last part of putative loop 5 would be a transmembrane helix. This is predicted by the hydropathy profile, which shows that the first two thirds of putative loop 5 is highly charged, (with a lot of amino acid variability between PitA and PitB), while the last third is hydrophobic, (and has a more highly conserved sequence). Located within this putative transmembrane region is Ala-213. Replacement of this residue by Asp prevents assembly of the PitA protein, indicating it may be located in an area sensitive to charge. The *pitA1* mutation G220D, predicted to be within TM helix 6 in all models, also prevents assembly and functioning of the PitA protein. Eight other Asp substitutions in the putative hydrophilic loops of PitA allow the growth of these strains on Pi Media (Table 6.3), showing that Asp amino acid substitutions are rarely so detrimental in the loop regions of this protein.



PitA/PitB chimeras which have exchanged a region containing the putative 'loop 5/helix 6' show that this sequence influences the  $K_m^{app}$  of each protein. The PitA 'loop 5/helix 6' region placed within PitB gives this chimera a more efficient PitA-like  $K_m^{app}$  while the complementary chimera (PitB 'loop 5/helix 6' in PitA) has a higher  $K_m^{app}$  similar to that of PitB (Table 6.4). Thus this region seems to be important for substrate affinity. The exchanged region referred to as 'loop 5/helix 6' contains all the amino acid changes within the putative loop 5 (including L216V in the hydrophobic section), I219A at the 'loop 5/helix 6' interface plus two other conservative changes (I242V, V244I) that reside at either the beginning of the large loop 6 (Model A) or the interface between the small loop 6/helix 7 (Model B) (See Figure 6.10). Most amino acid variation occurs within putative loop 5, which is highly positively charged and placed in the cytoplasm of all topological models. PitA contains 13 positive and 3 negatively charged amino acids and PitB contains 16 positive and 4 negatively charged amino acids in the 42 residue loop. However, this variable region of 'loop 5' is only present in PiT family members which are closely related to *E. coli* (Figure 6.13), so it is unlikely to contain particular residues that are essential to a conserved Pi transport mechanism. Possibly amino acid residues in this loop interact with  $\alpha$ -helix 6 and participate indirectly in the transport mechanism. Mutagenesis of individual amino acids that differ between PitA and PitB is required to determine the role of this region.

These results do imply that putative TM helix 6 and its surrounding sequences are involved in the transport pathway of  $MePO_4/H^+$ . Determining the topology of this region will be an important step in defining the mechanism of Pit.

The small 3-fold difference in  $K_m^{app}$  between PitA and PitB makes it difficult to interpret results for the remaining chimeric proteins, as variations between whole cell assays could obscure small to moderate alterations in kinetic parameters. PitB chimera containing PitA 'loop 6' may slightly decrease the affinity for substrate and lower the activity of the transporter, but this may be due to nonspecific unfavourable interactions between the variable regions of the chimera. The complementary chimera (PitB 'loop

6' in PitA) has kinetic parameters almost identical to wild type PitA protein, suggesting that this highly variable region has little effect on the mechanism of Pi transport in Pit.

Multiple sequence alignments of 68 PiT family PD1131 domains identify eight highly conserved sequences that will also be useful targets for mutagenesis as these are conserved amongst distant species, suggesting that these regions may be important for transporter function (229). The His-225, Asp-229, Lys-232 triad forms part of one of these sequences, and there is evidence that other conserved sequences may be essential. For example the mutated PitA(I413D G414L) protein produces an assembled membrane protein that is nonfunctional, thus these mutations are more disruptive than the H225Q, D229N, or K232Q substitutions. Ile-413 and Gly-414 are predicted to be in either a cytoplasmic (Model A) or periplasmic (Model B) loop containing several charged side chains, and this region is also predicted to be within a cytoplasmic loop in PiT-2 (229). This degree of disruption is unexpected in a hydrophilic loop, as shown by the eight other Asp substitutions in the putative hydrophilic loops of PitA allow the growth of these strains on Pi Media (Table 6.3). However, Ile-413 and Gly-414 are at the end of one of the highly conserved PiT family sequences, and Gly-414 is strongly conserved within many of these proteins. Thus this region may play an important structural or functional role in Pi transport.

The predicted topologies of other PiT family transporters such as PiT-2, Pht2;1 and Pho89 bear some similarities to PitA and PitB models (47, 196, 197, 229). The mammalian type III NaPi transporter PiT-2 is predicted to have a similar topology to PitA for the first five transmembrane helices, three of which form the conserved N-terminal PD1131 domain. This topology has been confirmed experimentally by antibody tagging, identification of glycosylation sites, and C-terminal truncation tagging, showing that the N-terminus plus loops 2 and 4 are extracellular. However experimental evidence indicates that the predicted topology for the remainder of the PiT-2 protein is incorrect. While there is less experimental support for the topology of the latter half of the protein, tagging experiments indicated that the C-terminus and loop 8 are extracellular, while functional evidence places loop 7 intracellular and loop 8 extracellular (229). This model then implies that the C-terminal PD1131 domain of 3 transmembrane helices has an inverted topology. This inverted topology has also been

predicted for the yeast Pho89 protein and the *Arabidopsis* Pht2;1 protein (47, 196, 197). The presence of internal repeats with opposite membrane topologies has also been shown for the *E. coli* YrbG protein. PhoA fusions were used to confirm that this putative Na<sup>+</sup>/Ca<sup>2+</sup> exchanger consists of two sets of five transmembrane segments in opposite orientations (219).

The predicted topology for the first 5 transmembrane helices of PitA and PitB is consistent with experimental data from PiT-2. However the topology is less clear for the C-terminal half of the proteins. The PiT-2 cytoplasmic loop 8, which coincides with the retroviral Region A binding site, is located where the large variable loop of PitA/PitB interrupts the C-terminal PD1131 domain (See multiple sequence alignment, Figure 6.11B). This should unambiguously locate this area of PitA/PitB at the periplasm. However this large variable region, which is only found in the Gram-negative bacteria that are closely related to *E. coli*, contains an extra hydrophobic region which has the potential to form a transmembrane helix. If this potential  $\alpha$ -helix (between G233 and T261) is inserted in the membrane then the remainder of the C-terminal PD1131 domain is in the same orientation as the N-terminal domain (Table 6.1 - transmembrane region 7, Figure 6.2 - Model B), while ignoring this potential  $\alpha$ -helix causes the remainder of the C-terminal PD1131 domain to be inverted (Figure 6.2 - Model A).

A number of topology prediction programs have been developed since the initial analyses used to form models for PitA and PitB. As more information on membrane proteins has become available new methods using neural networks (PHD - (215)) or Hidden Markov Models (TMHMM - (135), HMMTOP - (261)) have been formulated. von Heijne's group (185) have determined that if four or five methods (out of five) agree on their topology predictions for a particular protein then experimental results show that these predictions are over 90% correct, (by analysis of 60 proteins of known topology). However this agreement between programs applies to only half the potential inner membrane proteins of *E. coli*, and PitA and PitB do not form part of this group (web site [http://www.sbc.su.se/~johan/Very\\_Reliable\\_Topol\\_Pred.html](http://www.sbc.su.se/~johan/Very_Reliable_Topol_Pred.html), PitA - EG12230, PitB - EG12883, (185)). Application of the TMHMM program to PitA and PitB produced topologies of 10 transmembrane helices quite similar to those of the



TopPred IV and GES profiles, but with  $\alpha$ -helix 9 moved towards the N-terminus by at least 11 residues (G418 to T440).

Locating the environment of amino acids His-225, Asp-229 and Lys-232, either within a transmembrane helix or within the periplasm, is important for determining the strategy of future mutagenesis studies. Therefore defining the topology of PitA, particularly around TM helix 6 and the latter half of the protein, will be very helpful. Sandwich fusions with bacterial alkaline phosphatase, which is active only when located in the periplasm, could be used to refine this topology (55).

---

# ***Chapter seven***

***Overview and future directions***

---

# Chapter seven

## Overview and future directions

### 7.1 Introduction

PitA and PitB have been characterised as individual genes of *E. coli* with open reading frames of 1500 bases starting at the predicted ATG start codons. Each protein was shown to function as a Pi transporter, providing the first evidence of two *E. coli* Pit transport systems. Kinetic parameters were defined for the plasmid-borne Pit genes, showing PitA to have a  $K_m^{\text{app}}$  of about 2 $\mu\text{M}$  and PitB to have  $K_m^{\text{app}}$  of around 28 $\mu\text{M}$ . However, subsequent experiments indicated that the PitB gene undergoes some form of repression and/or activity modulation that affects these parameters, so at this point the kinetic determinations are relevant only to the specific assay conditions. Kinetic parameters that accurately reflect the characteristics of PitA and PitB in wild type cells can only be defined once these issues of regulation and/or modulation have been addressed. Therefore this thesis has evolved into two projects – the regulation of *pitB* and the characterisation of PitA/PitB structure and function.

### 7.2 Regulation of the Pit system

The *pit* lesion of strain K-10 was identified. K-10, which has no pit activity when its Pst system is mutated, contains a point mutation of G to A at nucleotide 658 in *pitA* causing a G220D substitution in the PitA protein, and a wild type *pitB*. (The *pitA* point mutation is discussed later.) PCR amplification of AN3066 genomic DNA showed that the *pitB* protein coding sequence and putative promoter region (for at least 319 upstream nucleotides) were wild type. *pitB* was originally isolated by its ability to transport Pi, and can function as an independent Pi transporter on plasmid pBR322. This suggested that genomic *pitB* may undergo repression that is disrupted by placing it on a plasmid, and is the first evidence for regulation of the *E. coli* Pit system. The use of a PitB polyclonal antipeptide antibody clearly demonstrated that PitB activity was linked to the levels of PitB protein in cell membranes, so this regulation may occur at the level of transcription. The Pi auxotrophic background strains AN3066 and AN3902 contain a *pstC345* deletion that causes constitutive expression of the *pho* regulon.



While most genes of this regulon are activated there is also some evidence for gene repression (9, 244, 271, 291). The possibility that the *phoB-phoR* two component regulatory system may mediate this repression was explored by deleting this operon from Pi auxotrophic strains that contained either genomic *pitB*<sup>+</sup> (*pitA pstC*) or no Pi transporters (*pitB pitA pstC*). Unexpectedly, the *phoB-phoR* deletion allowed both strains to grow on media containing 500µM Pi. However, assays of Pi uptake using 6µM Pi showed that the *pitB*<sup>+</sup> strain, AN4081, had a significant initial rate of Pi transport while the control strain AN4085 had Pi uptake that was not significantly above background levels (Table 5.2). Thus, the presence of an active *phoB-phoR* operon mediates the repression of genomic *pitB*. This finding is significant, as it provides the first definitive evidence of PitB regulation. This regulation of *pitB* may explain why, despite extensive characterisation of Pit, previous research identified only one *pit* lesion (*pitA*) whose gene appeared to be constitutive in expression/activity (59, 211, 212, 249, 289).

Transcriptional regulation of *pitB* by PhoB is the most likely mechanism of regulation. The above strains could be analysed by Northern blot to determine if an increase in *pitB* mRNA correlates with the presence of Pi uptake mediated by deleting *phoB-phoR*. Reporter genes may also be used to measure levels of transcription. A reporter gene such as β-galactosidase (*lacZ*) could replace part of the open reading frame of *pitB*, or be fused near the C-terminus of this gene (activity requiring a cytoplasmic location). It would be important to place this *pitB::lacZ* fusion on the genome, as plasmid-borne *pitB* is not repressed by the *pho* regulon. The β-galactosidase activity of wild type cells and the mutants described above could then be assayed under a variety of Pi concentrations. *Rhizobium meliloti* contains a *pit* equivalent that is repressed under conditions of Pi limitation. A chromosomal *R. meliloti orfA-pit::lacZ* fusion was used to show that expression of this operon becomes constitutive in a *phoB* mutant (8). It is likely that regulation of *pitB* in *E. coli* is similar.

Repression of *pitB* may occur directly through the binding of PhoB, or may be mediated by another product of the *pho* regulon. It has been proposed that PhoB can cause repression by binding in atypical locations on the promoter. Genes repressed by

an active *phoBR* operon contain putative *pho* boxes in atypical locations (9, 244). Similarly, activation of the CreBC two component regulatory system causes repression of *malE* transcription. This gene has an atypically positioned “cre-tag” binding site within the transcript, rather than between -40 and -80 within the promoter (see Chapter five discussion, Section 5.6, for details) (7). However, there is no direct experimental evidence to support this theory.

*pitA* and *pitB* form a good comparative system for analysis of putative PhoB binding or regulation, as *pitA* appears constitutive while *pitB* undergoes regulation. A preliminary analysis of the promoter sequences of these genes did not reveal any obvious *pho* boxes that are unique to *pitB* (Table 5.4). Thus, experimental analysis of these regions is needed. The putative transcriptional start sites should be confirmed for *pitA* and *pitB* by primer extension analysis, to define the promoters for each gene. The DNA of these promoter regions could be tested for PhoB binding *in vitro* by the use of electrophoretic mobility shift assays. Alternatively, DNase I footprinting could be used. This would also identify the region of PhoB binding, if it occurred. As proposed atypical binding sites occur before the -35 promoter region and after the transcriptional start (7, 9, 244) the analysed *pitA/pitB* DNA sequences should extend past these regions. Potential mechanisms of repression include the presence of divergent promoters, so binding to either DNA strand is a possibility (279).

Profiling *pitB* expression and activity with a normal *pho* regulon under a variety of Pi concentrations will be important for gaining an understanding of *pitB* regulation. While the identification of genomic *pitB* activity in the presence of a *phoB-phoR* deletion is an important first step, mutations in the *phoBR* two component regulatory system can have wide ranging effects that may interfere with a detailed analysis of *pitB* regulation. For example, experiments on *phoR* mutants suggest that histidine kinases from other two component regulators ‘cross talk’ with PhoB in the absence of the PhoR protein. This activation may not be physiologically significant in wild type cells (7, 66, 125, 283). Mutations in over 14 genes have also been shown to dramatically alter alkaline phosphatase synthesis (which acts as a reporter of *pho* regulon activity) in *phoR* mutants via unknown mechanisms which may involve other two component regulatory systems (125). Whether ‘cross talk’ can also occur at the transcriptional activator level

in *phoB* mutants has not been investigated. Deletion of the *phoB-phoR* operon alters the expression of over 100 genes within the *pho* regulon (271). This may have many unexpected consequences, such as was shown by the induction of low level Pi transport when *phoB-phoR* was deleted from the Pi auxotrophic *pitA pitB pstC* triple mutant (Section 5.5.2). Thus, a clear picture of normal *pitB* expression may be obtainable only in the presence of wild type *phoBR* and its associated regulators.

All work described in this study used the *pstC345* deletion to prevent Pi uptake by the Pst system. This mutation causes constitutive derepression of the *pho* regulon. Cox *et al* (43, 44) have described several PstA and PstC mutations which prevent Pi transport without causing constitutive derepression of the *pho* regulon (as measured by alkaline phosphatase activity, a protein that is activated by the *pho* regulon). These Pst mutations may provide a better background strain for analysis of *pitB* regulation. However, as we do not fully understand the role of the Pst system in the control of the *pho* regulon, we cannot be sure that this regulon remains unaffected by these Pst mutations. Thus the *pstSCABU* operon may also need to be wild type for unambiguous results.

Analysis of *pitB* transcription through the use of northern blots or reporter genes allows the detection of *pitB* mRNA or transcription in the presence of wild type Pst, PitA and *phoBR* systems. These methods may be used to compare *pitB* expression between wild type cells and those containing the mutations  $\Delta(\textit{phoB-phoR})$  (no *pho* regulon activation) or  $\Delta\textit{pstC345}$  (constitutive *pho* regulon activity) under different Pi concentrations. The levels of *pitB* mRNA or transcription may also be determined for previously constructed strains with varying levels of PitB protein expression. In particular, analysis of strains containing plasmids pAN656 and pAN1116 could determine if an increase in transcription is involved in the higher levels of Pi uptake and protein expression which occur upon deletion of 1196 nucleotides upstream of the *pitB* open reading frame (See Chapter 5, Section 5.4).

It is probable that increased PitB Pi uptake and protein expression involves an increase in *pitB* transcription. Placing wild type *pitB* on a plasmid allowed a formerly Pi auxotrophic *pitA pstC* strain to transport Pi, while Hoffer *et al* (99) demonstrated that



the presence of multiple genomic copies of wild type *pitB* had a similar effect on a Pi auxotrophic *pitA pstS* strain, consistent with titrating out a repressor protein. An increase in *Rhizobium meliloti pit* transcription, caused by either placing this gene on a plasmid or by the presence of a single thymidine deletion 54 nucleotides upstream of the *orfA-pit* operon transcription start site, has been shown to overcome the effects of constitutive expression of the *pho* regulon (9). This required only a two- to three-fold increase in *pit* expression, as measured by an *orfA-pit::lacZ* fusion.

However, a viable alternative to transcriptional activation is the enhancement of mRNA stability. This post-transcriptional mechanism has been shown to occur for the Pi transporter PiT-2 in human cell lines, as a response to Pi depletion (35). Thus, this mechanism cannot be ruled out for *pitB*. Any increase in *pitB* mRNA found by northern blot analysis could then be checked by the reporter gene system to determine if *pitB* promoter activity has also increased. Timed quantitative Northern blot experiments could also be carried out to detect changes in mRNA stability, or real time PCR could be used.

However, increasing the levels of *pitB* mRNA by transcription and/or enhanced mRNA stability, does not explain all of the regulatory effects observed in this transport system. The decrease in  $K_m^{app}$  noted for PitB expressed with only 207 upstream nucleotides (pAN1116) indicates a different mechanism of activity is occurring. This change in substrate affinity may be due to post-translational modifications stimulated by the increased PitB protein expression. Post-translational activation involving the aggregation of mammalian Pit-2 transporter has already been observed (207, 229). Cross-linking or the use of native polyacrylamide gel electrophoresis may be used to determine if PitB forms a homo-oligomer. An alternative explanation for the altered mechanism of PitB could be that an inhibitor which interacts with the PitB protein may be diluted out by the increased PitB expression. The presence of a second conserved gene, called the 'Pit accessory protein', which is found in the same operon as the *pit* gene in many bacterial species could indicate the involvement of another protein in Pit function or regulation. This situation has recently been demonstrated for the high affinity ammonium transporter, AmtB, of *E. coli*. Most prokaryotes and archaea encode at least one Amt protein, which is almost always found in an operon together

with a second gene (*glnK*) that encodes a small signal transduction protein involved in sensing the cellular nitrogen status (3, 257). Coutts *et al* (41) have demonstrated that GlnK acts as a negative regulator of AmtB activity by binding to the membrane (probably through direct association with AmtB) when nitrogen supplies are plentiful. Under low nitrogen conditions GlnK is uridylylated and is not strongly membrane associated. The *glnk-amtB* operon is also repressed by transcriptional regulation through the NtrB/NtrC two component regulatory system when nitrogen is plentiful. It is suggested that GlnK regulation can fine-tune the activity of AmtB in response to minor or transient fluctuations in ammonium availability. A second protein homologous to GlnK (GlnB) is also present in *E. coli*. Thus, this system could provide a viable model for Pit regulation. However, *E. coli* does not encode a 'Pit accessory protein', either in an operon with *pitA* or *pitB*, or elsewhere on the genome (by Blast sequence similarity analyses using several 'Pit accessory protein' amino acid sequences).

It is also possible that *pitB* expression or activity is moderated by environmental conditions such as pH or sodium ion concentration. The PiT family of transporters contains several members that are induced by Pi limitation but also require specific conditions, such as an alkaline pH, for activity. *pho89*, a Pit-like Na<sup>+</sup>-Pi transporter from *Saccharomyces cerevisiae*, is induced by Pi limitation but needs alkaline conditions and Na<sup>+</sup> to function (168). Similarly, *pho-4* from *Neurospora crassa* codes for a Na<sup>+</sup>/H<sup>+</sup>/Li<sup>+</sup>-Pi transporter induced by low Pi concentrations that is dominant at pH8 and above (274). Thus the activity of PitB (in a strain where genomic *pitB* is expressed) under different pH and sodium ion concentrations could be investigated. However, experiments checking the initial rates of plasmid-borne PitA (6μM Pi) and PitB (30μM Pi) over the pH range of 6.8-7.6 found little change in activity (results not shown). Activity over a wider pH range could be investigated.

Genomic *pitB* may produce a much lower level of Pi uptake than either PitA or the Pst system. For example, the Pit-like *pho89* from *Saccharomyces cerevisiae* has low activity levels and it can only be studied when the unrelated H<sup>+</sup>-Pi transporter *pho84*, which is also induced by Pi limitation, is mutated (168).

While *pitA* does not appear to be regulated by the *phoBR* operon, other forms of activity modulation cannot be ruled out. The plasmid-borne PitA  $K_m^{app}$  of around  $2\mu\text{M}$  is significantly lower than the  $11.9 - 38\mu\text{M}$  range recorded by researchers using genomic *pit* genes (211) (290) (268). Western blot analysis shows that placing *pitA* on a plasmid correlates with a significant increase in protein expression (Figure 4.5 – first two lanes). Further analysis of genomic and plasmid-borne PitA kinetic parameters should determine if this transporter also undergoes changes in substrate affinity. Thus, Pit regulation may be quite complex. As shown above, many organisms contain at least two Pit genes, and some of these genes are accompanied by Pit accessory proteins, whose function has yet to be defined ((8) and NCBI blast searches).

### **7.3 Structure and function studies on PitA and PitB**

PitA and PitB have 81% identity in their deduced amino acid sequences. Similar topological models were produced for PitA and PitB using a combination of hydropathy profiles and topology prediction programs. These models have 10 transmembrane  $\alpha$ -helices, with most variation between PitA and PitB occurring in the hydrophilic loops (Figure 3.15). Measurement of kinetic parameters initially showed that PitA had a  $K_m^{app}$  14-fold lower than that of PitB. Therefore PitA and PitB formed a good comparative system where relationships between the kinetic parameters, amino acid sequences and the Pi translocation mechanism of a membrane transporter could be systematically explored.

It was noted that His-225, Asp-229 and Lys-232, a triad of amino acids shared by PitA and PitB, were similar to Lys-319, His-322 and Glu-325 found in helix 10 of lactose permease. These amino acids play a role in the functioning of lactose permease, with the mutant H322R carrying out facilitated diffusion without proton translocation (117), and neutral substitutions of Glu-325 and Lys-319 exhibiting no active transport or efflux while undergoing wild type levels of exchange and counter flow (29, 30, 209). Subsequent mutagenesis studies showed that the neutral combination of K319C/D240C had activity and normal protein assembly, suggesting this pair may form a salt bridge (222). Kaback's latest model incorporates this data, with His-322 and Glu-325 forming



part of the proton translocation pathway and Lys-319 playing a structural role as part of a salt bridge with Asp-240 (223). (This model is discussed in more detail in Section 1.9.2 and Figure 1.5.)

Polar substitutions of the equivalent amino acids in PitA (H225Q, D229N and K232Q) resulted in decreased initial rates of Pi uptake to less than 3% of wild type PitA, while allowing growth on 500 $\mu$ M Pi media (Table 6.2, Figure 6.3). Thus, these amino acids play an important role in Pit structure or function. The subsequent sequencing of putative Pit genes from many different species reveals that this sequence is highly conserved in the PiT family, although more distantly related organisms such as eukaryotes contain HxxxDxxN rather than HxxxDxxK. The PitA K232Q substitution had lower levels of protein assembly than H225Q, D229N or wild type PitA (Figure 6.4), indicating that this mutation may have some effect on protein assembly. Perhaps Lys-232 forms a salt bridge, as Lys-319 is proposed to in lactose permease. In some PiT proteins this lysine residue is replaced by an asparagine residue, which may form a hydrogen bond rather than a salt bridge (20, 36, 308).

The PitA H225Q, D229N and K232Q single mutants could be further characterised through measuring the Pi uptake, facilitated diffusion, efflux and exchange of proteoliposomes from cells carrying the above mutants or wild type PitA (268). These results may provide information on what aspect of the mechanism has been affected, such as substrate binding, substrate and/or proton translocation. PitB could also be characterised in this way.

The initial mutagenesis experiments on His-225, Asp-229 and Lys-232 indicate that these amino acids play important roles in Pit structure or function. In the proposed PitA topology model, these residues were moved into a transmembrane helix because of the analogy with lactose permease (Figure 6.2). This topology is not predicted by any sequence analysis program. These topology prediction programs often under-represent the presence of charged or polar amino acids within the membrane, as they cannot predict the interactions that make these insertions more favourable.

However, the predicted model is still a viable option. It is possible that His-225, Asp-229 and Lys-232 may be structurally or functionally important, but located within the periplasm. The PiT family contains an internally repeated domain, PD1131, and these three amino acids form part of the first highly conserved sequence in the N-terminal PD1131 domain. The proposed topology for several PiT family members places this N-terminal conserved sequence just inside the periplasm, and the equivalent C-terminal conserved sequence just inside the cytoplasm. In all Pit topology models this C-terminal conserved sequence is also located in the cytoplasm (Figures 3.15, 6.2). However, most of these topologies have yet to be supported by experimental data, and analysis of mammalian PiT-2 showed the last half of the protein's topology differed from that suggested by predictive programs (229).

Therefore defining the environment of His-225, Asp-229 and Lys-232 is crucial for progress in elucidating the mechanism of PitA, and this should be carried out before any further mutagenesis studies are undertaken. This area of PitA is obviously important for the mechanism. As well as containing the above triad of amino acids, this putative transmembrane helix plus the putative loop 5 region contains amino acids determining the difference in  $K_m^{\text{app}}$  between PitA and PitB, as shown by chimera analysis (Section 6.4, Table 6.4, Figure 6.10). Nearby single mutations G220D and A213D cause complete loss of Pi uptake and minimal insertion of protein in the membrane (Figures 6.5, 6.6).

Bacterial alkaline phosphatase sandwich fusions could be used to define the topology of PitA. This has been successfully used in other systems, such as the aromatic amino acid permease, AroP (39), and the phenylalanine-specific permease, PheP (198). The alkaline phosphatase reporter gene would be inserted within *pitA* near the end of putative periplasmic and cytoplasmic loops. The alkaline phosphatase protein is only active in the periplasm, and the sandwich fusion approach has the advantage of including the entire *pitA* sequence, leading to less ambiguous results (55). Analysis of the last half of PitA should easily determine if the protein has 9 or 10 transmembrane regions. Carefully targeted sandwich fusions within the putative transmembrane helix 6 and the following loop should then be able to locate His-225, Asp-229 and Lys-232 within either the membrane or the periplasm, as experimental evidence suggests that

approximately half of a transmembrane domain is needed to translocate the alkaline phosphatase protein through the membrane to the periplasm (118). Alternative approaches that could be trialed if the results of this experiment are ambiguous include the use of site-directed mutagenesis with biophysical labels to determine the environment of these amino acids, and the deletion of single amino acids at the putative helix/loop boundaries of helix 6, as described for lactose permease in Section 1.9.2. The proximity of other  $\alpha$ -helices may also be determined through double labeling experiments, as for lactose permease (Sections 1.8.2, 1.9.2).

Defining the PitA topology will allow more appropriate targeting and analysis of amino acids important for function and/or structure. If His-225, Asp-229 and Lys-232 are located in the membrane, different substitutions for each amino acid, such as H225R or H225F, may help determine if a certain charge or the presence of a dissociable proton is required for function. These amino acids may interact with each other and with other charged amino acids, which may also be located within the membrane. Interacting amino acids are often identified by second site suppressor mutagenesis. These experiments start with a detrimental mutant of an important residue (such as a neutral substitution of a charged residue) and identify spontaneous mutations within the protein that confer activity. However, this powerful system may be difficult to use on PitA. Firstly, conditions would need to be established where a PitA mutant such as K232A produced negligible growth (e.g. media with a very low Pi concentration) or could otherwise be distinguished from suppressor mutants. Then an effective screening assay for the identification of *pitA* suppressor mutants would need to be established. This may prove difficult, as mutations elsewhere in the *E. coli* genome (such as deleting the *phoBR* operon) are also able to confer Pi transport on Pi auxotrophic strains. Hoffer *et al* (99, 100) tried to create suppressor mutations in Pi auxotrophic *pstS* strains, and identified *pitA* and *pitB* pseudo-revertants instead. With lactose permease Lee *et al* (143) could only use this method when they assayed for melobiose fermentation as frequent mutations in the lac repressor created too many background cells with lactose transport, preventing the identification of genuine suppressor mutants. An alternative but more labour intensive approach is identify charged amino acids that may have the potential to interact with His-225, Asp-229 or Lys-232 and replace them with neutral substitutions via site-directed mutagenesis. Possible residues to mutate include Glu-27,



which is within TM 1 of all models and is conserved as a negatively charged residue (Glu or Asp) in the putative Pit sequences of many Gram negative bacteria. Model B (Figure 6.2B) also places Glu-259 within TM 7, so this should be tested. Other residues to check include Glu-380 and Asp-374, if the C-terminal end of the large variable loop is moved into the membrane. Any detrimental single mutants could then be analysed further. If the formation of a salt bridge is suspected a double neutral mutant could be constructed, to see if Pi uptake activity and/or protein assembly is increased.

Russ *et al* (216) have identified GxxxG as a strong helix packing motif, which can be surrounded by a large variety of amino acids. This motif is strongly conserved in the N-PD1131 domain of PiT proteins from distantly related organisms and is predicted to be within a membrane helix (putative TM 4 of PitA – Figure 6.12). The equivalent region in the C-PD1131 domain contains alanine and serine substitutions in some proteins and is predicted to be in an aqueous environment in our models (SxxxG in putative loop 8 of PitA – Figure 6.12). Proteins closely related to PitA and PitB also have this GxxxG motif within the His-225, Asp-229 and Lys-232 sequence that has been shown to be important for Pit function (HGxNDGxQK) (near mauve boxes in Figure 6.13), which may support a membrane location for this sequence. However this particular GxxxG motif is not conserved in all PiT proteins.

If His-225, Asp-229 or Lys-232 are found to be located in the periplasm mutagenesis studies may be used to analyse the role of these residues (such as in substrate binding or structure) and to identify other residues that may be involved in the  $\text{MeHPO}_4/\text{H}^+$  translocation pathway. This model (Figure 3.15) places Arg-211 in TM 6, so this residue should be changed to a neutral or polar residue. However, the amino acids involved in the translocation pathway need not be charged.

Another useful strategy is to examine the other highly conserved regions in the PiT family. Conserved region 1, which contains the His, Asp and Lys triad in the C-terminal PD1131 domain of PitA and PitB, is more highly conserved in the N-terminal PD1131 domain of the PiT family (region 1 in the first cytoplasmic loop - Figure 6.12) and targeted mutagenesis may be useful in defining its role. Disruption of these

conserved regions has already been shown to strongly affect Pit activity. For example, the PitA mutations I413D/G414L, which occur at the end of the third highly conserved region in the N-terminal PD1131 domain, completely disrupted Pi uptake while allowing protein assembly (Figures 6.5, 6.6, 6.12). Further analysis of this area, through more conservative substitutions and mutagenesis of other conserved amino acids in this domain, may determine whether this region has a structural or functional role. Information can also be gained from analysis of the multiple sequence alignments, which may be useful for designing future experiments or models. An initial analysis shows that the putative TM 2 and TM 4 contain conserved glycines and/or alanines which is consistent with close packing between one face of each helix.

It is interesting to note that the first cytoplasmic loop of the Major Facilitator Superfamily also contains a highly conserved region that is repeated later in the protein. This is thought to play a structural role as these proteins have many diverse substrates and often use different ions for energy coupling. Detrimental mutations (D68T or D68S) in this region of lactose permease led to many second site repressor mutations that appear to alter the helical topologies in the second half of the protein to facilitate a better interaction with the first half of the protein (Figure 7.1). Therefore this region may be involved in global conformation changes in the protein (106). The conserved motif is GXXX(D/E)(R/K)XG[X](R/K)(R/K) (See Figure 1.6 for its location) (81), with mutations in GXXXD/E significantly inhibiting activity. The PiT family has a GXNDX motif in this location (conserved region 1). Thus bulky substitutions at G and a conservative D to E substitution could be tried.

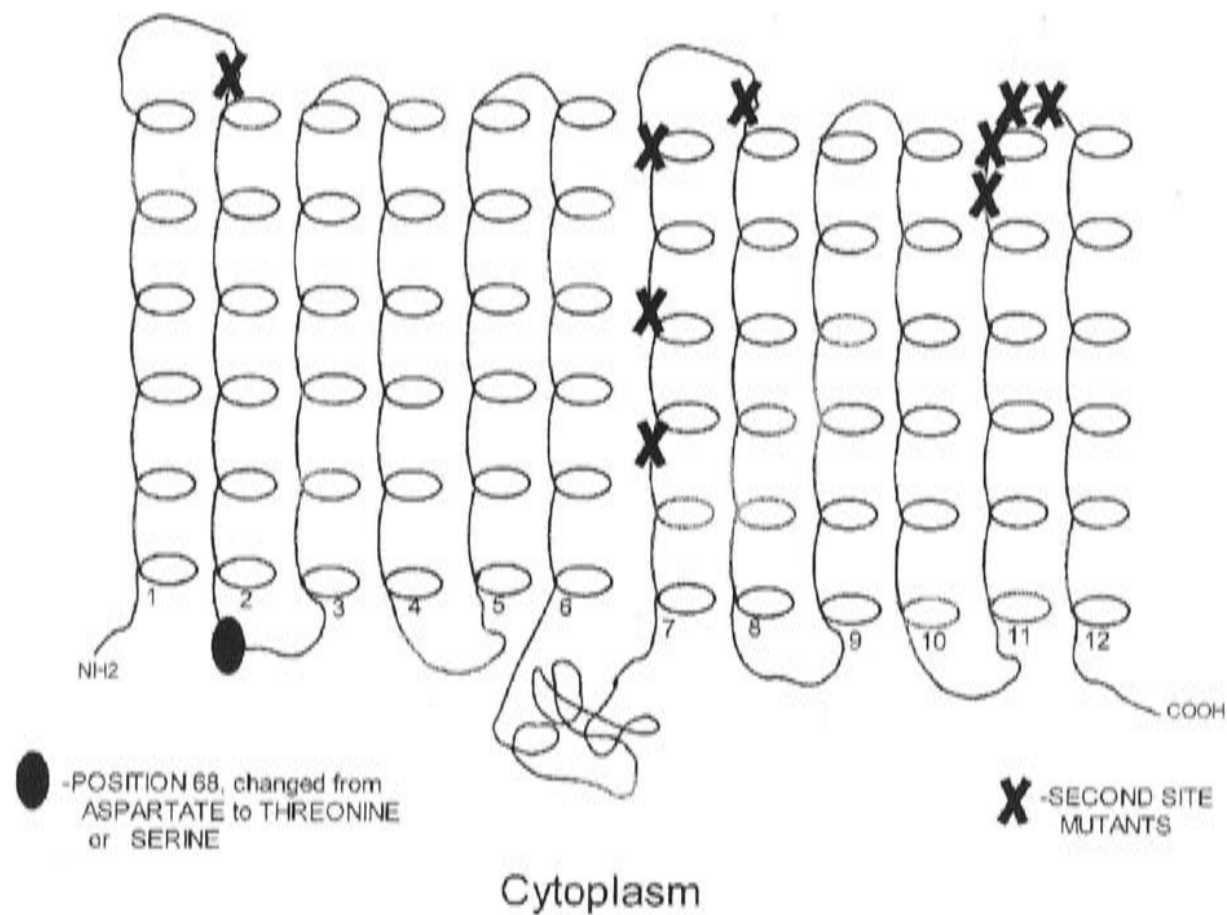
Hoffer *et al* (100) have isolated an interesting pseudo-revertant for the PitA (G220D) mutation. The additional point mutation T41I was able to overcome the original G220D mutation to restore Pi transport, actually causing a greater degree of alkaline phosphatase repression when grown in high Pi media in comparison with the wild type PitA protein (100). This mutation may alleviate the protein assembly problems caused by G220D through either a direct interaction, such as removing an inappropriate charge interaction between the polar threonine and the negatively charged aspartic acid, or through some form of physical hindrance caused by the larger isoleucine side chain. While Thr-41, located several residues after the first highly conserved region within the

## Figure 7.1

### Second site suppressor mutations in lactose permease produced in response to threonine or serine substitutions at Asp-68.

This diagram is from Jessen-Marshall *et al* (106).

Crosses mark the locations of the suppressor mutations which restored the ability of lactose permease with threonine or serine substitutions at Asp-68 to ferment melibiose. The suppressor mutation changes are Thr45Arg, Cys234Trp, Cys234Phe, Gln241Leu, Phe247Val, Gly257Asp, Ser366Phe, Val367Glu and Ala369Pro.





first cytoplasmic loop (Figure 6.12), is conserved in most PiT family transporters, PitB has an alanine in this position (green box in Figure 6.11A). The Pi uptake and protein expression of this pseudo-revertant could be characterised. Perhaps PitB could be mutated to G220D, to see if this point mutation can also prevent Pit protein assembly and function with a nonpolar group at residue 41.

The initial work on protein purification of PitA and PitB, plus the development of polyclonal antibodies for each protein, allows for the future optimisation of these methods. However, at this stage pure Pit protein is not required for the most relevant experiments for elucidating Pit regulation, structure or function. Pure protein could be used to form 3D crystals for X-ray crystallographic analysis, or for the formation of 2D crystalline arrays, tubes or single particles for cryo-EM analysis. However, the initial experiments in Section 3.4 suggest obtaining enough pure Pit protein may be difficult. Providing pure protein or good quality 2D arrays is also just the first step towards solving the crystal structure of a membrane protein, as has been shown for lactose permease and the nicotinic acetylcholine receptor (94). It has taken almost 20 years to obtain an acetylcholine receptor structure with 4.6Å resolution (175) and despite intense effort, the crystal structure of lactose permease has yet to be elucidated.

Initial characterisation of PitA and PitB indicate that these proteins provide a good model system for examining the function and structure of the PiT family of transporters. This study also highlights the potential for *pitA* and *pitB* to form a good comparative system for examining repression mediated by the *E. coli phoBR* two component regulatory system – a form of regulation that has yet to be systematically explored.

---

# ***Bibliography***

---

# Bibliography

1. Altschul, S. F., T. L. Madden, A. A. Schaffer, J. Zhang, Z. Zhang, W. Miller, and D. J. Lipman. 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* **25**(17):3389-402.
2. Amemura, M., K. Makino, H. Shinagawa, A. Kobayashi, and A. Nakata. 1985. Nucleotide sequence of the genes involved in phosphate transport and regulation of the phosphate regulon in *Escherichia coli*. *J Mol Biol.* **184**(2):241-50.
3. Arcondeguy, T., R. Jack, and M. Merrick. 2001. P(II) signal transduction proteins, pivotal players in microbial nitrogen control. *Microbiol Mol Biol Rev.* **65**(1):80-105.
4. Argast, M., and W. Boos. 1979. Purification and properties of the *sn*-glycerol 3-phosphate-binding protein of *Escherichia coli*. *J Biol Chem.* **254**(21):10931-5.
5. Armstrong, C. M. 1971. Interaction of tetraethylammonium ion derivatives with the potassium channels of giant axons. *J Gen Physiol.* **58**(4):413-37.
6. Auer, M., M. J. Kim, M. J. Lemieux, A. Villa, J. Song, X. D. Li, and D. N. Wang. 2001. High-yield expression and functional analysis of *Escherichia coli* glycerol-3-phosphate transporter. *Biochemistry.* **40**(22):6628-35.
7. Avison, M. B., R. E. Horton, T. R. Walsh, and P. M. Bennett. 2001. *Escherichia coli* CreBC is a global regulator of gene expression that responds to growth in minimal media. *J Biol Chem.* **276**(29):26955-61.
8. Bardin, S. D., and T. M. Finan. 1998. Regulation of phosphate assimilation in *Rhizobium (Sinorhizobium) meliloti*. *Genetics.* **148**(4):1689-700.
9. Bardin, S. D., R. T. Voegelé, and T. M. Finan. 1998. Phosphate assimilation in *Rhizobium (Sinorhizobium) meliloti*: identification of a *pit*-like gene. *J Bacteriol.* **180**(16):4219-26.
10. Battini, J. L., P. Rodrigues, R. Muller, O. Danos, and J. M. Heard. 1996. Receptor-binding properties of a purified fragment of the 4070A amphotropic murine leukemia virus envelope glycoprotein. *J Virol.* **70**(7):4387-93.
11. Bauer, K., M. Struyve, D. Bosch, R. Benz, and J. Tommassen. 1989. One single lysine residue is responsible for the special interaction between polyphosphate and the outer membrane porin PhoE of *Escherichia coli*. *J Biol Chem.* **264**(28):16393-8.
12. Beard, S. J., R. Hashim, G. Wu, M. R. Binet, M. N. Hughes, and R. K. Poole. 2000. Evidence for the transport of zinc(II) ions via the *pit* phosphate transport system in *Escherichia coli*. *FEMS Microbiol Lett.* **184**(2):231-5.
13. Benz, R., A. Schmid, and R. E. Hancock. 1985. Ion selectivity of gram-negative bacterial porins. *J Bacteriol.* **162**(2):722-7.
14. Benz, R., A. Schmid, T. Nakae, and G. H. Vos Schepkerkeuter. 1986. Pore formation by LamB of *Escherichia coli* in lipid bilayer membranes. *J Bacteriol.* **165**(3):978-86.
15. Benz, R., A. Schmid, P. Van der Ley, and J. Tommassen. 1989. Molecular basis of porin selectivity: membrane experiments with OmpC-PhoE and OmpF-PhoE hybrid proteins of *Escherichia coli* K-12. *Biochim Biophys Acta.* **981**(1):8-14.
16. Bibi, E., and H. R. Kaback. 1992. Functional complementation of internal deletion mutants in the lactose permease of *Escherichia coli*. *Proc Natl Acad Sci U S A.* **89**(5):1524-8.



17. **Bibi, E., and H. R. Kaback.** 1990. *In vivo* expression of the *lacY* gene in two segments leads to functional lac permease. *Proc Natl Acad Sci U S A.* **87**(11):4325-9.
18. **Blattner, F. R., G. Plunkett, 3rd, C. A. Bloch, N. T. Perna, V. Burland, M. Riley, J. Collado Vides, J. D. Glasner, C. K. Rode, G. F. Mayhew, J. Gregor, N. W. Davis, H. A. Kirkpatrick, M. A. Goeden, D. J. Rose, B. Mau, and Y. Shao.** 1997. The complete genome sequence of *Escherichia coli* K-12. *Science.* **277**(5331):1453-74.
19. **Bowie, J. U.** 1997. Helix packing in membrane proteins. *J Mol Biol.* **272**(5):780-9.
20. **Bowie, J. U.** 2000. Understanding membrane protein structure by design. *Nat Struct Biol.* **7**(2):91-4.
21. **Bowman, B. J., K. E. Allen, and C. W. Slayman.** 1983. Vanadate-resistant mutants of *Neurospora crassa* are deficient in a high-affinity phosphate transport system. *J Bacteriol.* **153**(1):292-6.
22. **Bracha, M., and E. Yagil.** 1973. A new type of alkaline phosphatase-negative mutants in *Escherichia coli*. *Mol Gen Genet.* **122**(1):53-60.
23. **Bradford, M. M.** 1976. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal Biochem.* **72**:248-54.
24. **Brandl, C. J., and C. M. Deber.** 1986. Hypothesis about the function of membrane-buried proline residues in transport proteins. *Proc Natl Acad Sci U S A.* **83**(4):917-21.
25. **Brown, L. S.** 2001. Proton transport mechanism of bacteriorhodopsin as revealed by site-specific mutagenesis and protein sequence variability. *Biochemistry.* **66**(11):1249-55.
26. **Brzoska, P., and W. Boos.** 1988. Characteristics of a *ugp*-encoded and *phoB*-dependent glycerophosphoryl diester phosphodiesterase which is physically dependent on the *ugp* transport system of *Escherichia coli*. *J Bacteriol.* **170**(9):4125-35.
27. **Buck, K. J., and S. G. Amara.** 1994. Chimeric dopamine-norepinephrine transporters delineate structural domains influencing selectivity for catecholamines and 1-methyl-4-phenylpyridinium. *Proc Natl Acad Sci U S A.* **91**(26):12584-8.
28. **Calamia, J., and C. Manoil.** 1990. Lac permease of *Escherichia coli*: topology and sequence elements promoting membrane insertion. *Proc Natl Acad Sci U S A.* **87**(13):4937-41.
29. **Carrasco, N., L. M. Antes, M. S. Poonian, and H. R. Kaback.** 1986. Lac permease of *Escherichia coli*: histidine-322 and glutamic acid-325 may be components of a charge-relay system. *Biochemistry.* **25**(16):4486-8.
30. **Carrasco, N., I. B. Puttner, L. M. Antes, J. A. Lee, J. D. Larigan, J. S. Lolkema, P. D. Roepe, and H. R. Kaback.** 1989. Characterization of site-directed mutants in the lac permease of *Escherichia coli*. 2. Glutamate-325 replacements. *Biochemistry.* **28**(6):2533-9.
31. **Chan, F. Y., and A. Torriani.** 1996. PstB protein of the phosphate-specific transport system of *Escherichia coli* is an ATPase. *J Bacteriol.* **178**(13):3974-7.
32. **Chang, G., and C. B. Roth.** 2001. Structure of MsbA from *E. coli*: a homolog of the multidrug resistance ATP binding cassette (ABC) transporters. *Science.* **293**(5536):1793-800.
33. **Chang, G., R. H. Spencer, A. T. Lee, M. T. Barclay, and D. C. Rees.** 1998. Structure of the MscL homolog from *Mycobacterium tuberculosis*: a gated mechanosensitive ion channel. *Science.* **282**(5397):2220-6.

34. **Chien, M. L., J. L. Foster, J. L. Douglas, and J. V. Garcia.** 1997. The amphotropic murine leukemia virus receptor gene encodes a 71-kilodalton protein that is induced by phosphate depletion. *J Virol.* **71**(6):4564-70.
35. **Chien, M. L., E. O'Neill, and J. V. Garcia.** 1998. Phosphate depletion enhances the stability of the amphotropic murine leukemia virus receptor mRNA. *Virology.* **240**(1):109-17.
36. **Choma, C., H. Gratkowski, J. D. Lear, and W. F. DeGrado.** 2000. Asparagine-mediated self-association of a model transmembrane helix. *Nat Struct Biol.* **7**(2):161-6.
37. **Corpet, F., J. Gouzy, and D. Kahn.** 1998. The ProDom database of protein domain families. *Nucleic Acids Res.* **26**(1):323-6.
38. **Cosgriff, A. J., G. Brasier, J. Pi, C. Dogovski, J. P. Sarsero, and A. J. Pittard.** 2000. A study of AroP-PheP chimeric proteins and identification of a residue involved in tryptophan transport. *J Bacteriol.* **182**(8):2207-17.
39. **Cosgriff, A. J., and A. J. Pittard.** 1997. A topological model for the general aromatic amino acid permease, AroP, of *Escherichia coli*. *J Bacteriol.* **179**(10):3317-23.
40. **Costello, M. J., J. Escaig, K. Matsushita, P. V. Viitanen, D. R. Menick, and H. R. Kaback.** 1987. Purified lac permease and cytochrome o oxidase are functional as monomers. *J Biol Chem.* **262**(35):17072-82.
41. **Coutts, G., G. Thomas, D. Blakey, and M. Merrick.** 2002. Membrane sequestration of the signal transduction protein GlnK by the ammonium transporter AmtB. *Embo J.* **21**(4):536-45.
42. **Cowan, S. W., T. Schirmer, G. Rummel, M. Steiert, R. Ghosh, R. A. Paupit, J. N. Jansonius, and J. P. Rosenbusch.** 1992. Crystal structures explain functional properties of two *E. coli* porins. *Nature.* **358**(6389):727-33.
43. **Cox, G. B., D. Webb, J. Godovac Zimmermann, and H. Rosenberg.** 1988. Arg-220 of the PstA protein is required for phosphate transport through the phosphate-specific transport system in *Escherichia coli* but not for alkaline phosphatase repression. *J Bacteriol.* **170**(5):2283-6.
44. **Cox, G. B., D. Webb, and H. Rosenberg.** 1989. Specific amino acid residues in both the PstB and PstC proteins are required for phosphate transport by the *Escherichia coli* Pst system. *J Bacteriol.* **171**(3):1531-4.
45. **Cox, G. B., I. G. Young, L. M. McCann, and F. Gibson.** 1969. Biosynthesis of ubiquinone in *Escherichia coli* K-12: location of genes affecting the metabolism of 3-octaprenyl-4-hydroxybenzoic acid and 2-octaprenylphenol. *J Bacteriol.* **99**(2):450-8.
46. **Cronan, J.** 1987. , p. 31-55. *In* F. C. Neidhardt, J. L. Ingraham, K. B. Low, B. Magasanik, M. Schaecter, and H. E. Umbarger (ed.), *Escherichia coli* and *Salmonella typhimurium*: cellular and molecular biology. American Society for Microbiology Publications, Washington, D.C.
47. **Daram, P., S. Brunner, C. Rausch, C. Steiner, N. Amrhein, and M. Bucher.** 1999. Pht2;1 encodes a low-affinity phosphate transporter from *Arabidopsis*. *Plant Cell.* **11**(11):2153-66.
48. **Deber, C. M., B. J. Sorrell, and G. Y. Xu.** 1990. Conformation of proline residues in bacteriorhodopsin. *Biochem Biophys Res Commun.* **172**(2):862-9.
49. **Deisenhofer, J., O. Epp, I. Sinning, and H. Michel.** 1995. Crystallographic refinement at 2.3 Å resolution and refined model of the photosynthetic reaction centre from *Rhodospseudomonas viridis*. *J Mol Biol.* **246**(3):429-57.



50. **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<sup>+</sup> conduction and selectivity. *Science*. **280**(5360):69-77.
51. **Dreyer, K., F. S. Pedersen, and L. Pedersen.** 2000. A 13-amino-acid Pit1-specific loop 4 sequence confers feline leukemia virus subgroup B receptor function upon Pit2. *J Virol*. **74**(6):2926-9.
52. **Dunten, R. L., M. Sahin Toth, and H. R. Kaback.** 1993. Role of the charge pair aspartic acid-237-lysine-358 in the lactose permease of *Escherichia coli*. *Biochemistry*. **32**(12):3139-45.
53. **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 Å reveals the molecular basis of anion selectivity. *Nature*. **415**(6869):287-94.
54. **Edelman, A., L. Bowler, J. K. Broome Smith, and B. G. Spratt.** 1987. Use of a beta-lactamase fusion vector to investigate the organization of penicillin-binding protein 1B in the cytoplasmic membrane of *Escherichia coli*. *Mol Microbiol*. **1**(1):101-6.
55. **Ehrmann, M., D. Boyd, and J. Beckwith.** 1990. Genetic analysis of membrane protein topology by a sandwich gene fusion approach. *Proc Natl Acad Sci U S A*. **87**(19):7574-8.
56. **Eiglmeier, K., W. Boos, and S. T. Cole.** 1987. Nucleotide sequence and transcriptional startpoint of the *glpT* gene of *Escherichia coli*: extensive sequence homology of the glycerol-3-phosphate transport protein with components of the hexose-6-phosphate transport system. *Mol Microbiol*. **1**(3):251-8.
57. **Eilers, M., S. C. Shekar, T. Shieh, S. O. Smith, and P. J. Fleming.** 2000. Internal packing of helical membrane proteins. *Proc Natl Acad Sci U S A*. **97**(11):5796-801.
58. **Ellison, D. W., and W. R. McCleary.** 2000. The unphosphorylated receiver domain of PhoB silences the activity of its output domain. *J Bacteriol*. **182**(23):6592-7.
59. **Elvin, C. M., N. E. Dixon, and H. Rosenberg.** 1986. Molecular cloning of the phosphate (inorganic) transport (*pit*) gene of *Escherichia coli* K12. *Mol Gen Genet*. **204**:477-484.
60. **Elvin, C. M., C. M. Hardy, and H. Rosenberg.** 1987. Molecular studies of the phosphate inorganic transport system of *Escherichia coli*, p. 156-158. In A. Torriani-Gorini, F. G. Rothman, S. Silver, A. Wright, and E. Yagil (ed.), *Phosphate metabolism and cellular regulation in microorganisms*. American Society for Microbiology, Washington DC.
61. **Elvin, C. M., C. M. Hardy, and H. Rosenberg.** 1985. Pi exchange mediated by the GlpT-dependent *sn*-glycerol-3-phosphate transport system in *Escherichia coli*. *J Bacteriol*. **161**(3):1054-8.
62. **Engelman, D. M., and T. A. Steitz.** 1981. The spontaneous insertion of proteins into and across membranes: the helical hairpin hypothesis. *Cell*. **23**(2):411-22.
63. **Engelman, D. M., T. A. Steitz, and A. Goldman.** 1986. Identifying nonpolar transbilayer helices in amino acid sequences of membrane proteins. *Annu Rev Biophys Chem*. **15**:321-53.
64. **Engvall, E., and P. Perlman.** 1971. Enzyme-linked immunosorbent assay (ELISA). Quantitative assay of immunoglobulin G. *Immunochemistry*. **8**(9):871-4.



65. **Fiedler, U., and V. Weiss.** 1995. A common switch in activation of the response regulators NtrC and PhoB: phosphorylation induces dimerization of the receiver modules. *Embo J.* **14**(15):3696-705.
66. **Fisher, S. L., S. K. Kim, B. L. Wanner, and C. T. Walsh.** 1996. Kinetic comparison of the specificity of the vancomycin resistance VanS for two response regulators, VanR and PhoB. *Biochemistry.* **35**(15):4732-40.
67. **Foster, D. L., M. Boublik, and H. R. Kaback.** 1983. Structure of the lac carrier protein of *Escherichia coli*. *J Biol Chem.* **258**(1):31-4.
68. **Freund, J., and K. McDermott.** 1942. Sensitization to horse serum by means of adjuvants. *Proc Soc Exp Biol Med.* **49**:548-53.
69. **Frillingos, S., M. L. Ujwal, J. Sun, and H. R. Kaback.** 1997. The role of helix VIII in the lactose permease of *Escherichia coli*: I. Cys-scanning mutagenesis. *Protein Sci.* **6**(2):431-7.
70. **Froshauer, S., G. N. Green, D. Boyd, K. McGovern, and J. Beckwith.** 1988. Genetic analysis of the membrane insertion and topology of MalF, a cytoplasmic membrane protein of *Escherichia coli*. *J Mol Biol.* **200**(3):501-11.
71. **Fu, H. H., and S. Luan.** 1998. AtKuP1: a dual-affinity K<sup>+</sup> transporter from *Arabidopsis*. *Plant Cell.* **10**(1):63-73.
72. **Fu, R., and T. A. Cross.** 1999. Solid-state nuclear magnetic resonance investigation of protein and polypeptide structure. *Annu Rev Biophys Biomol Struct.* **28**:235-68.
73. **Gerchman, Y., A. Rimon, M. Venturi, and E. Padan.** 2001. Oligomerization of NhaA, the Na<sup>+</sup>/H<sup>+</sup> antiporter of *Escherichia coli* in the membrane and its functional and structural consequences. *Biochemistry.* **40**(11):3403-12.
74. **Gerdes, R. G., and H. Rosenberg.** 1974. The relationship between the phosphate-binding protein and a regulator gene product from *Escherichia coli*. *Biochim Biophys Acta.* **351**(1):77-86.
75. **Gibson, F., G. B. Cox, J. A. Downie, and J. Radik.** 1977. A mutation affecting a second component of the F<sub>0</sub> portion of the magnesium ion-stimulated adenosine triphosphatase of *Escherichia coli* K12. The *uncC424* allele. *Biochem J.* **164**(1):193-8.
76. **Goodrich, J. A., M. L. Schwartz, and W. R. McClure.** 1990. Searching for and predicting the activity of sites for DNA binding proteins: compilation and analysis of the binding sites for *Escherichia coli* integration host factor (IHF). *Nucleic Acids Res.* **18**(17):4993-5000.
77. **Goswitz, V. C., and R. J. Brooker.** 1995. Structural features of the uniporter/symporter/antiporter superfamily. *Protein Sci.* **4**(3):534-7.
78. **Gott, P., and W. Boos.** 1988. The transmembrane topology of the *sn*-glycerol-3-phosphate permease of *Escherichia coli* analysed by *phoA* and *lacZ* protein fusions. *Mol Microbiol.* **2**(5):655-63.
79. **Gottesman, M. M., and I. Pastan.** 1993. Biochemistry of multidrug resistance mediated by the multidrug transporter. *Annu Rev Biochem.* **62**:385-427.
80. **Green, A. L., and R. J. Brooker.** 2001. A face on transmembrane segment 8 of the lactose permease is important for transport activity. *Biochemistry.* **40**(40):12220-9.
81. **Griffith, J. K., M. E. Baker, D. A. Rouch, M. G. Page, R. A. Skurray, I. T. Paulsen, K. F. Chater, S. A. Baldwin, P. J. Henderson, and P. J. Henderson.** 1992. Membrane transport proteins: implications of sequence comparisons. *Curr Opin Cell Biol.* **4**(4):684-95.

82. **Grigorieff, N., T. A. Ceska, K. H. Downing, J. M. Baldwin, and R. Henderson.** 1996. Electron-crystallographic refinement of the structure of bacteriorhodopsin. *J Mol Biol.* **259**(3):393-421.
83. **Guan, C. D., B. Wanner, and H. Inouye.** 1983. Analysis of regulation of *phoB* expression using a *phoB-cat* fusion. *J Bacteriol.* **156**(2):710-7.
84. **Haldimann, A., L. L. Daniels, and B. L. Wanner.** 1998. Use of new methods for construction of tightly regulated arabinose and rhamnose promoter fusions in studies of the *Escherichia coli* phosphate regulon. *J Bacteriol.* **180**(5):1277-86.
85. **Hama, H., and T. H. Wilson.** 1994. Replacement of alanine 58 by asparagine enables the melibiose carrier of *Klebsiella pneumoniae* to couple sugar transport to Na<sup>+</sup>. *J Biol Chem.* **269**(2):1063-7.
86. **Han, J. S., J. Y. Park, Y. S. Lee, B. Thony, and D. S. Hwang.** 1999. PhoB-dependent transcriptional activation of the *iciA* gene during starvation for phosphate in *Escherichia coli*. *Mol Gen Genet.* **262**(3):448-52.
87. **Hanamura, A., and H. Aiba.** 1991. Molecular mechanism of negative autoregulation of *Escherichia coli crp* gene. *Nucleic Acids Res.* **19**(16):4413-9.
88. **Harlow, E., and D. Lane.** 1988. *Antibodies. A laboratory manual.* Cold Spring Harbor Laboratory Press, Cold Spring Harbor, USA.
89. **Harris, R. M., D. C. Webb, S. M. Howitt, and G. B. Cox.** 2001. Characterization of PitA and PitB from *Escherichia coli*. *J Bacteriol.* **183**(17):5008-14.
90. **Haugeto, O., K. Ullensvang, L. M. Levy, F. A. Chaudhry, T. Honore, M. Nielsen, K. P. Lehre, and N. C. Danbolt.** 1996. Brain glutamate transporter proteins form homomultimers. *J Biol Chem.* **271**(44):27715-22.
91. **Hayashi, S.-I., J. P. Koch, and E. C. C. Lin.** 1964. Active Transport of L- $\alpha$ -Glycerophosphate in *Escherichia coli*. *J Biol Chem.* **239**:3098-105.
92. **He, M. M., J. Voss, W. L. Hubbell, and H. R. Kaback.** 1995. Use of designed metal-binding sites to study helix proximity in the lactose permease of *Escherichia coli*. 1. Proximity of helix VII (Asp237 and Asp240) with helices X (Lys319) and XI (Lys358). *Biochemistry.* **34**(48):15661-6.
93. **Heller, K. B., E. C. Lin, and T. H. Wilson.** 1980. Substrate specificity and transport properties of the glycerol facilitator of *Escherichia coli*. *J Bacteriol.* **144**(1):274-8.
94. **Heuser, J. E., and S. R. Salpeter.** 1979. Organization of acetylcholine receptors in quick-frozen, deep-etched, and rotary-replicated Torpedo postsynaptic membrane. *J Cell Biol.* **82**(1):150-73.
95. **Higgins, C. F., and M. M. Gottesman.** 1992. Is the multidrug transporter a flippase? *Trends Biochem Sci.* **17**(1):18-21.
96. **Higgins, C. F., and K. J. Linton.** 2001. Structural biology. The xyz of ABC transporters. *Science.* **293**(5536):1782-4.
97. **Hobot, J. A., E. Carlemalm, W. Villiger, and E. Kellenberger.** 1984. Periplasmic gel: new concept resulting from the reinvestigation of bacterial cell envelope ultrastructure by new methods. *J Bacteriol.* **160**(1):143-52.
98. **Hoekstra, W. P., J. E. Bergmans, and E. M. Zuidweg.** 1980. Role of *recBC* nuclease in *Escherichia coli* transformation. *J Bacteriol.* **143**(2):1031-2.
99. **Hoffer, S. M., P. Schoondermark, H. W. van Veen, and J. Tommassen.** 2001. Activation by gene amplification of *pitB*, encoding a third phosphate transporter of *Escherichia coli* K-12. *J Bacteriol.* **183**(15):4659-63.
100. **Hoffer, S. M., and J. Tommassen.** 2001. The phosphate-binding protein of *Escherichia coli* is not essential for P(i)-regulated expression of the *pho* regulon. *J Bacteriol.* **183**(19):5768-71.



101. **Hoffer, S. M., N. Uden, and J. Tommassen.** 2001. Expression of the *pho* regulon interferes with induction of the *uhpT* gene in *Escherichia coli* K-12. *Arch Microbiol.* **176**(5):370-6.
102. **Hustedt, E. J., and A. H. Beth.** 1999. Nitroxide spin-spin interactions: applications to protein structure and dynamics. *Annu Rev Biophys Biomol Struct.* **28**:129-53.
103. **Ikeda, H., and J. I. Tomizawa.** 1965. Transducing fragments in generalized transduction by phage P1. 3. Studies with small phage particles. *J Mol Biol.* **14**(1):120-9.
104. **Island, M. D., B. Y. Wei, and R. J. Kadner.** 1992. Structure and function of the *uhp* genes for the sugar phosphate transport system in *Escherichia coli* and *Salmonella typhimurium*. *J Bacteriol.* **174**(9):2754-62.
105. **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):1609-18.
106. **Jessen Marshall, A. E., and R. J. Brooker.** 1996. Evidence that transmembrane segment 2 of the lactose permease is part of a conformationally sensitive interface between the two halves of the protein. *J Biol Chem.* **271**(3):1400-4.
107. **Jessen Marshall, A. E., N. J. Parker, and R. J. Brooker.** 1997. Suppressor analysis of mutations in the loop 2-3 motif of lactose permease: evidence that glycine-64 is an important residue for conformational changes. *J Bacteriol.* **179**(8):2616-22.
108. **Jessen Marshall, A. E., N. J. Paul, and R. J. Brooker.** 1995. The conserved motif, GXXX(D/E)(R/K)XG[X](R/K)(R/K), in hydrophilic loop 2/3 of the lactose permease. *J Biol Chem.* **270**(27):16251-7.
109. **Johnson, J. L., and R. J. Brooker.** 1999. A K319N/E325Q double mutant of the lactose permease cotransports H<sup>+</sup> with lactose. Implications for a proposed mechanism of H<sup>+</sup>/lactose symport. *J Biol Chem.* **274**(7):4074-81.
110. **Johnson, J. L., M. S. Lockheart, and R. J. Brooker.** 2001. A triple mutant, K319N/H322Q/E325Q, of the lactose permease cotransports H<sup>+</sup> with thiodigalactoside. *J Membr Biol.* **181**(3):215-24.
111. **Jones, D. T., W. R. Taylor, and J. M. Thornton.** 1994. A model recognition approach to the prediction of all-helical membrane protein structure and topology. *Biochemistry.* **33**(10):3038-49.
112. **Jung, H., K. Jung, and H. R. Kaback.** 1994. A conformational change in the lactose permease of *Escherichia coli* is induced by ligand binding or membrane potential. *Protein Sci.* **3**(7):1052-7.
113. **Jung, H., K. Jung, and H. R. Kaback.** 1994. Cysteine 148 in the lactose permease of *Escherichia coli* is a component of a substrate binding site. 1. Site-directed mutagenesis studies. *Biochemistry.* **33**(40):12160-5.
114. **Jung, K., H. Jung, J. H. Wu, G. G. Prive, and H. R. Kaback.** 1993. Use of site-directed fluorescence labeling to study proximity relationships in the lactose permease of *Escherichia coli*. *Biochemistry.* **32**(46):12273-12278.
115. **Kaback, H. R.** 1992. , p. 97-125. In K. W. Jeon and M. Friedlander (ed.), *International Review of Cytology*, vol. 137A. Academic Press, New York.
116. **Kaback, H. R.** 1997. A molecular mechanism for energy coupling in a membrane transport protein, the lactose permease of *Escherichia coli*. *Proc Natl Acad Sci U S A.* **94**(11):5539-43.
117. **Kaback, H. R.** 1987. Use of site-directed mutagenesis to study the mechanism of a membrane transport protein. *Biochemistry.* **26**(8):2071-6.



118. **Kaback, H. R., S. Frillingos, H. Jung, K. Jung, G. G. Prive, M. L. Ujwal, C. Weitzman, J. Wu, and K. Zen.** 1994. The lactose permease meets Frankenstein. *J Exp Biol.* **196**:183-95.
119. **Kadner, R. J., C. A. Webber, and M. D. Island.** 1993. The family of organo-phosphate transport proteins includes a transmembrane regulatory protein. *J Bioenerg Biomembr.* **25**(6):637-45.
120. **Kasahara, M., K. Makino, M. Amemura, A. Nakata, and H. Shinagawa.** 1991. Dual regulation of the *ugp* operon by phosphate and carbon starvation at two interspaced promoters. *J Bacteriol.* **173**(2):549-58.
121. **Kavanaugh, M. P., and D. Kabat.** 1996. Identification and characterization of a widely expressed phosphate transporter/retrovirus receptor family. *Kidney Int.* **49**(4):959-63.
122. **Kavanaugh, M. P., D. G. Miller, W. Zhang, W. Law, S. L. Kozak, D. Kabat, and A. D. Miller.** 1994. Cell-surface receptors for gibbon ape leukemia virus and amphotropic murine retrovirus are inducible sodium-dependent phosphate symporters. *Proc Natl Acad Sci U S A.* **91**(15):7071-5.
123. **Kim, E. J., J. M. Kwak, N. Uozumi, and J. I. Schroeder.** 1998. AtKUP1: an *Arabidopsis* gene encoding high-affinity potassium transport activity. *Plant Cell.* **10**(1):51-62.
124. **Kim, S. K., S. Kimura, H. Shinagawa, A. Nakata, K. S. Lee, B. L. Wanner, and K. Makino.** 2000. Dual transcriptional regulation of the *Escherichia coli* phosphate-starvation-inducible *psiE* gene of the phosphate regulon by PhoB and the cyclic AMP (cAMP)-cAMP receptor protein complex. *J Bacteriol.* **182**(19):5596-9.
125. **Kim, S. K., M. R. Wilmes Riesenber, and B. L. Wanner.** 1996. Involvement of the sensor kinase EnvZ in the in vivo activation of the response-regulator PhoB by acetyl phosphate. *Mol Microbiol.* **22**(1):135-47.
126. **Kimura, Y., D. G. Vassylyev, A. Miyazawa, A. Kidera, M. Matsushima, K. Mitsuoka, K. Murata, T. Hirai, and Y. Fujiyoshi.** 1997. Surface of bacteriorhodopsin revealed by high-resolution electron crystallography. *Nature.* **389**(6647):206-11.
127. **King, S. C., C. L. Hansen, and T. H. Wilson.** 1991. The interaction between aspartic acid 237 and lysine 358 in the lactose carrier of *Escherichia coli*. *Biochim Biophys Acta.* **1062**(2):177-86.
128. **King, S. C., and T. H. Wilson.** 1989. Galactoside-dependent proton transport by mutants of the *Escherichia coli* lactose carrier. Replacement of histidine 322 by tyrosine or phenylalanine. *J Biol Chem.* **264**(13):7390-4.
129. **Koebnik, R.** 1995. Proposal for a peptidoglycan-associating alpha-helical motif in the C-terminal regions of some bacterial cell-surface proteins. *Mol Microbiol.* **16**(6):1269-70.
130. **Koebnik, R., K. P. Locher, and P. Van Gelder.** 2000. Structure and function of bacterial outer membrane proteins: barrels in a nutshell. *Mol Microbiol.* **37**(2):239-53.
131. **Konings, W. N., and H. Rosenberg.** 1978. Phosphate transport in membrane vesicles from *Escherichia coli*. *Biochim Biophys Acta.* **508**(2):370-8.
132. **Koprowski, P., and A. Kubalski.** 2001. Bacterial ion channels and their eukaryotic homologues. *Bioessays.* **23**(12):1148-58.
133. **Kozak, S. L., D. C. Siess, M. P. Kavanaugh, A. D. Miller, and D. Kabat.** 1995. The envelope glycoprotein of an amphotropic murine retrovirus binds specifically

- to the cellular receptor/phosphate transporter of susceptible species. *J Virol.* **69**(6):3433-40.
134. **Krause, H. M., and N. P. Higgins.** 1986. Positive and negative regulation of the Mu operator by Mu repressor and *Escherichia coli* integration host factor. *J Biol Chem.* **261**(8):3744-52.
  135. **Krogh, A., B. Larsson, G. von Heijne, and E. L. Sonnhammer.** 2001. Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. *J Mol Biol.* **305**(3):567-80.
  136. **Kuhlbrandt, W., D. N. Wang, and Y. Fujiyoshi.** 1994. Atomic model of plant light-harvesting complex by electron crystallography. *Nature.* **367**(6464):614-21.
  137. **Kushner, S. R., H. Nagaishi, A. Templin, and A. J. Clark.** 1971. Genetic recombination in *Escherichia coli*: the role of exonuclease I. *Proc Natl Acad Sci U S A.* **68**(4):824-7.
  138. **Kyte, J., and R. F. Doolittle.** 1982. A simple method for displaying the hydropathic character of a protein. *J Mol Biol.* **157**(1):105-32.
  139. **Laemmli, U. K.** 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature.* **227**(259):680-5.
  140. **Lanyi, J. K.** 1998. Understanding structure and function in the light-driven proton pump bacteriorhodopsin. *J Struct Biol.* **124**(2-3):164-78.
  141. **Larson, T. J., G. Schumacher, and W. Boos.** 1982. Identification of the *glpT*-encoded *sn*-glycerol-3-phosphate permease of *Escherichia coli*, an oligomeric integral membrane protein. *J Bacteriol.* **152**(3):1008-21.
  142. **Lee, J. I., P. P. Hwang, C. Hansen, and T. H. Wilson.** 1992. Possible salt bridges between transmembrane alpha-helices of the lactose carrier of *Escherichia coli*. *J Biol Chem.* **267**(29):20758-64.
  143. **Lee, J. I., P. P. Hwang, and T. H. Wilson.** 1993. Lysine-319 interacts with both glutamic acid-269 and aspartic acid-240 in the lactose carrier of *Escherichia coli*. *J Biol Chem.* **268**(27):20007-20015.
  144. **Lemmon, M. A., and D. M. Engelman.** 1994. Specificity and promiscuity in membrane helix interactions. *FEBS Lett.* **346**(1):17-20.
  145. **Lewin, B.** 2000. *Genes VII*, 7 ed. Oxford University Press, New York.
  146. **Li, S. C., N. K. Goto, K. A. Williams, and C. M. Deber.** 1996. Alpha-helical, but not beta-sheet, propensity of proline is determined by peptide environment. *Proc Natl Acad Sci U S A.* **93**(13):6676-81.
  147. **Liu, K. H., C. Y. Huang, and Y. F. Tsay.** 1999. CHL1 is a dual-affinity nitrate transporter of *Arabidopsis* involved in multiple phases of nitrate uptake. *Plant Cell.* **11**(5):865-74.
  148. **Livrelli, V., I. W. Lee, and A. O. Summers.** 1993. *In vivo* DNA-protein interactions at the divergent mercury resistance (*mer*) promoters. I. Metalloregulatory protein MerR mutants. *J Biol Chem.* **268**(4):2623-31.
  149. **Lloyd, A. D., and R. J. Kadner.** 1990. Topology of the *Escherichia coli* *uhpT* sugar-phosphate transporter analyzed by using *TnphoA* fusions. *J Bacteriol.* **172**(4):1688-93.
  150. **Locher, K. P., B. Rees, R. Koebnik, A. Mitschler, L. Moulinier, J. P. Rosenbusch, and D. Moras.** 1998. Transmembrane signaling across the ligand-gated FhuA receptor: crystal structures of free and ferrichrome-bound states reveal allosteric changes. *Cell.* **95**(6):771-8.
  151. **Loddenkotter, B., B. Kammerer, K. Fischer, and U. Flugge.** 1993. Expression of the functional mature chloroplast triose phosphate translocator in yeast internal



- membranes and purification of the histidine-tagged protein by a single metal-affinity chromatography step. *Proc Natl Acad Sci U S A*. **90**(6):2155-9.
152. **Lodish, H., A. Berk, S. L. Zipursky, P. Matsudaira, D. Baltimore, and J. Darnell.** 2000. *Molecular Cell Biology*, fourth ed. W.H. Freeman and company, New York.
  153. **Loo, T. W., and D. M. Clarke.** 1996. Inhibition of oxidative cross-linking between engineered cysteine residues at positions 332 in predicted transmembrane segments (TM) 6 and 975 in predicted TM12 of human P-glycoprotein by drug substrates. *J Biol Chem*. **271**(44):27482-7.
  154. **Lu, Y. A., P. Clavijo, M. Galantino, Z. Y. Shen, W. Liu, and J. P. Tam.** 1991. Chemically unambiguous peptide immunogen: preparation, orientation and antigenicity of purified peptide conjugated to the multiple antigen peptide system. *Mol Immunol*. **28**(6):623-30.
  155. **Luecke, H., and F. A. Quioco.** 1990. High specificity of a phosphate transport protein determined by hydrogen bonds. *Nature*. **347**(6291):402-6.
  156. **Lundberg, K. S., D. D. Shoemaker, M. W. Adams, J. M. Short, J. A. Sorge, and E. J. Mathur.** 1991. High-fidelity amplification using a thermostable DNA polymerase isolated from *Pyrococcus furiosus*. *Gene*. **108**(1):1-6.
  157. **Luria, S. E., and J. S. Burrous.** 1957. Hybridisation between *Escherichia coli* and *Shigella*. *J Bacteriol*. **74**:461-476.
  158. **MacKenzie, K. R., J. H. Prestegard, and D. M. Engelman.** 1997. A transmembrane helix dimer: structure and implications. *Science*. **276**(5309):131-3.
  159. **Makino, K., M. Amemura, T. Kawamoto, S. Kimura, H. Shinagawa, A. Nakata, and M. Suzuki.** 1996. DNA binding of PhoB and its interaction with RNA polymerase. *J Mol Biol*. **259**(1):15-26.
  160. **Makino, K., S. K. Kim, H. Shinagawa, M. Amemura, and A. Nakata.** 1991. Molecular analysis of the cryptic and functional *phn* operons for phosphonate use in *Escherichia coli* K-12. *J Bacteriol*. **173**(8):2665-12.
  161. **Makino, K., H. Shinagawa, M. Amemura, T. Kawamoto, M. Yamada, and A. Nakata.** 1989. Signal transduction in the phosphate regulon of *Escherichia coli* involves phosphotransfer between PhoR and PhoB proteins. *J Mol Biol*. **210**(3):551-9.
  162. **Makino, K., H. Shinagawa, and A. Nakata.** 1985. Regulation of the phosphate regulon of *Escherichia coli* K-12: regulation and role of the regulatory gene *phoR*. *J Mol Biol*. **184**(2):231-40.
  163. **Maloney, P. C.** 1987. Coupling to an energised membrane: role of ion-motive gradients in the transduction of metabolic energy, p. 222-243. *In* F. C. Neidhardt, J. L. Ingraham, K. B. Low, B. Magasanik, M. Schaecter, and H. E. Umbarger (ed.), *Escherichia coli* and *Salmonella typhimurium*: cellular and molecular biology, vol. 2. American Society for Microbiology Publications, Washington DC.
  164. **Maloney, P. C., S. V. Ambudkar, V. Anatharam, L. A. Sonna, and A. Varadhachary.** 1990. Anion-exchange mechanisms in bacteria. *Microbiol Rev*. **54**(1):1-17.
  165. **Manoil, C.** 1990. Analysis of protein localization by use of gene fusions with complementary properties. *J Bacteriol*. **172**(2):1035-42.
  166. **Manoil, C., and J. Beckwith.** 1986. A genetic approach to analyzing membrane protein topology. *Science*. **233**(4771):1403-8.
  167. **Mansilla, M. C., and D. de Mendoza.** 2000. The *Bacillus subtilis cysP* gene encodes a novel sulphate permease related to the inorganic phosphate transporter (Pit) family. *Microbiology*. **146**(Pt 4):815-21.



168. **Martinez, P., and B. L. Persson.** 1998. Identification, cloning and characterization of a derepressible Na<sup>+</sup>-coupled phosphate transporter in *Saccharomyces cerevisiae*. *Mol Gen Genet.* **258**(6):628-38.
169. **Martinez, P., R. Zvyagilskaya, P. Allard, and B. L. Persson.** 1998. Physiological regulation of the derepressible phosphate transporter in *Saccharomyces cerevisiae*. *J Bacteriol.* **180**(8):2253-6.
170. **Medveczky, N., and H. Rosenberg.** 1971. Phosphate transport in *Escherichia coli*. *Biochim Biophys Acta.* **241**(2):494-506.
171. **Menick, D. R., N. Carrasco, L. Antes, L. Patel, and H. R. Kaback.** 1987. Lac permease of *Escherichia coli*: arginine-302 as a component of the postulated proton relay. *Biochemistry.* **26**(21):6638-44.
172. **Merkel, T. J., J. L. Dahl, R. H. Ebright, and R. J. Kadner.** 1995. Transcription activation at the *Escherichia coli uhpT* promoter by the catabolite gene activator protein. *J Bacteriol.* **177**(7):1712-8.
173. **Meyer, J. E., and G. E. Schulz.** 1997. Energy profile of maltooligosaccharide permeation through maltoporin as derived from the structure and from a statistical analysis of saccharide-protein interactions. *Protein Sci.* **6**(5):1084-91.
174. **Miller, D. G., R. H. Edwards, and A. D. Miller.** 1994. Cloning of the cellular receptor for amphotropic murine retroviruses reveals homology to that for gibbon ape leukemia virus. *Proc Natl Acad Sci USA.* **91**(1):78-82.
175. **Miyazawa, A., Y. Fujiyoshi, M. Stowell, and N. Unwin.** 1999. Nicotinic acetylcholine receptor at 4.6 Å resolution: transverse tunnels in the channel wall. *J Mol Biol.* **288**(4):765-86.
176. **Mizuno, T., H. Kasai, and S. Mizushima.** 1987. Construction of a series of *ompC-ompF* chimeric genes by *in vivo* homologous recombination in *Escherichia coli* and characterization of their translational products. *Mol Gen Genet.* **207**(2-3):217-23.
177. **Monod, J., G. Cohen-Bazire, and M. Cohn.** 1951. Sur la biosynthese de la β-galactosidase (lactase) chez *Escherichia coli* la specificite de l'induction. *Biochim Biophys Acta.* **7**:585-599.
178. **Morona, R., M. Klose, and U. Henning.** 1984. *Escherichia coli* K-12 outer membrane protein (OmpA) as a bacteriophage receptor: analysis of mutant genes expressing altered proteins. *J Bacteriol.* **159**(2):570-8.
179. **Mulligan, M. E., D. K. Hawley, R. Entriken, and W. R. McClure.** 1984. *Escherichia coli* promoter sequences predict *in vitro* RNA polymerase selectivity. *Nucleic Acids Res.* **12**(1 Pt 2):789-800.
180. **Mullis, K. B., and F. A. Faloona.** 1987. Specific synthesis of DNA *in vitro* via a polymerase-catalyzed chain reaction. *Methods Enzymol.* **155**:335-50.
181. **Nakata, A., M. Amemura, K. Makina, and H. Shinagawa.** 1987. Genetic and biochemical analysis of the phosphate-specific transport system in *Escherichia coli*, p. 150-155. In A. Torriani-Gorini, F. G. Rothman, S. Silver, A. Wright, and E. Yagil (ed.), *Phosphate metabolism and cellular regulation in microorganisms*. American Society for Microbiology, Washington DC.
182. **Neidhardt, F. C., J. L. Ingraham, and M. Schaechter (ed.).** 1990. *Physiology of the bacterial cell: A molecular approach*. Sinauer Associates.
183. **Nikaido, H.** 1993. Transport across the bacterial outer membrane. *J Bioenerg Biomembr.* **25**(6):581-9.
184. **Nikaido, H., and M. Vaara.** 1985. Molecular basis of bacterial outer membrane permeability. *Microbiol Rev.* **49**(1):1-32.

185. **Nilsson, J., B. Persson, and G. von Heijne.** 2000. Consensus predictions of membrane protein topology. *FEBS Lett.* **486**(3):267-9.
186. **Nogami, T., T. Mizuno, and S. Mizushima.** 1985. Construction of a series of *ompF-ompC* chimeric genes by *in vivo* homologous recombination in *Escherichia coli* and characterization of the translational products. *J Bacteriol.* **164**(2):797-801.
187. **Nollert, P., W. E. Harries, D. Fu, L. J. Miercke, and R. M. Stroud.** 2001. Atomic structure of a glycerol channel and implications for substrate permeation in aqua(glycero)porins. *FEBS Lett.* **504**(3):112-7.
188. **Norrande, J., T. Kempe, and J. Messing.** 1983. Construction of improved M13 vectors using oligodeoxynucleotide-directed mutagenesis. *Gene.* **26**(1):101-6.
189. **Oden, K. L., L. C. DeVeaux, C. R. Vibat, J. E. Cronan, Jr., and R. B. Gennis.** 1990. Genomic replacement in *Escherichia coli* K-12 using covalently closed circular plasmid DNA. *Gene.* **96**(1):29-36.
190. **Ohara, P. J., P. O. Sheppard, H. Thogersen, D. Venezia, B. A. Haldeman, V. Mcgrane, K. M. Houamed, C. Thomsen, T. L. Gilbert, and E. R. Mulvihill.** 1993. The ligand-binding domain in metabotropic glutamate receptors is related to bacterial periplasmic binding proteins. *Neuron.* **11**(1):41-52.
191. **Okamura, H., S. Hanaoka, A. Nagadoi, K. Makino, and Y. Nishimura.** 2000. Structural comparison of the PhoB and OmpR DNA-binding/transactivation domains and the arrangement of PhoB molecules on the phosphate box. *J Mol Biol.* **295**(5):1225-36.
192. **Olah, Z., C. Lehel, W. B. Anderson, M. V. Eiden, and C. A. Wilson.** 1994. The cellular receptor for gibbon ape leukemia virus is a novel high affinity sodium-dependent phosphate transporter. *J Biol Chem.* **269**(41):25426-31.
193. **Ostermeier, C., A. Harrenga, U. Ermler, and H. Michel.** 1997. Structure at 2.7 Å resolution of the *Paracoccus denitrificans* two-subunit cytochrome c oxidase complexed with an antibody FV fragment. *Proc Natl Acad Sci U S A.* **94**(20):10547-53.
194. **Palmer, G., J. P. Bonjour, and J. Caverzasio.** 1997. Expression of a newly identified phosphate transporter/retrovirus receptor in human SaOS-2 osteoblast-like cells and its regulation by insulin-like growth factor I. *Endocrinology.* **138**(12):5202-9.
195. **Pautsch, A., and G. E. Schulz.** 1998. Structure of the outer membrane protein A transmembrane domain. *Nat Struct Biol.* **5**(11):1013-7.
196. **Persson, B. L., A. Berhe, U. Fristedt, P. Martinez, J. Pattison, J. Petersson, and R. Weinander.** 1998. Phosphate permeases of *Saccharomyces cerevisiae*. *Biochim Biophys Acta.* **10**:1-2.
197. **Persson, B. L., J. Petersson, U. Fristedt, R. Weinander, A. Berhe, and J. Pattison.** 1999. Phosphate permeases of *Saccharomyces cerevisiae*: structure, function and regulation. *Biochim Biophys Acta.* **16**(3):255-72.
198. **Pi, J., and A. J. Pittard.** 1996. Topology of the phenylalanine-specific permease of *Escherichia coli*. *J Bacteriol.* **178**(9):2650-5.
199. **Pogell, B. M., B. R. Maity, S. Frumkin, and S. Shapiro.** 1966. Induction of an active transport system for glucose 6-phosphate in *Escherichia coli*. *Arch Biochem Biophys.* **116**(1):406-15.
200. **Popot, J. L.** 1993. Integral membrane protein structure - transmembrane alpha-helices as autonomous folding domains. *Curr Opin Struct Biol.* **3**(4):532-540.
201. **Popot, J. L., and D. M. Engelman.** 2000. Helical membrane protein folding, stability, and evolution. *Annu Rev Biochem.* **69**:881-922.



202. **Prive, G. G., and H. R. Kaback.** 1996. Engineering the lac permease for purification and crystallization. *J Bioenerg Biomembr.* **28**(1):29-34.
203. **Pusch, M., U. Ludewig, A. Rehfeldt, and T. J. Jentsch.** 1995. Gating of the voltage-dependent chloride channel CIC-0 by the permeant anion. *Nature.* **373**(6514):527-31.
204. **Rao, N. N., and A. Torriani.** 1990. Molecular aspects of phosphate transport in *Escherichia coli*. *Mol Microbiol.* **4**(7):1083-90.
205. **Rao, N. N., and A. Torriani.** 1988. Utilization by *Escherichia coli* of a high-molecular-weight, linear polyphosphate: roles of phosphatases and pore proteins. *J Bacteriol.* **170**(11):5216-23.
206. **Rodgers, A. J. W.** 1995. PhD Thesis: Molecular studies on the energy-transducing ATP synthase. The Australian National University, Canberra.
207. **Rodrigues, P., and J. M. Heard.** 1999. Modulation of phosphate uptake and amphotropic murine leukemia virus entry by posttranslational modifications of PIT-2. *J Virol.* **73**(5):3789-99.
208. **Rodriguez, R. L., F. Bolivar, H. M. Goodman, H. W. Boyer, and M. Betlach.** 1976. Construction and characterisation of cloning vehicles, p. 471-477. *In* D. P. Nierlich, W. J. Rutter, and C. Fox (ed.), *Molecular mechanisms in the control of gene expression.* Academic Press, New York.
209. **Roepe, P. D., T. G. Consler, M. E. Menezes, and H. R. Kaback.** 1990. The lac permease of *Escherichia coli*: site-directed mutagenesis studies on the mechanism of beta-galactoside/H<sup>+</sup> symport. *Res Microbiol.* **141**(3):290-308.
210. **Ronson, C. W., B. T. Nixon, and F. M. Ausubel.** 1987. Conserved domains in bacterial regulatory proteins that respond to environmental stimuli. *Cell.* **49**(5):579-81.
211. **Rosenberg, H., R. G. Gerdes, and K. Chegwidan.** 1977. Two systems for the uptake of phosphate in *Escherichia coli*. *J Bacteriol.* **131**(2):505-11.
212. **Rosenberg, H., R. G. Gerdes, and F. M. Harold.** 1979. Energy coupling to the transport of inorganic phosphate in *Escherichia coli* K12. *Biochem J.* **178**(1):133-7.
213. **Rosenberg, M. F., G. Velarde, R. C. Ford, C. Martin, G. Berridge, I. D. Kerr, R. Callaghan, A. Schmidlin, C. Wooding, K. J. Linton, and C. F. Higgins.** 2001. Repacking of the transmembrane domains of P-glycoprotein during the transport ATPase cycle. *Embo J.* **20**(20):5615-25.
214. **Rosenbusch, J. P.** 1988. Secondary and tertiary structure of membrane proteins. *Zbl Bakt.* **17**:259-266.
215. **Rost, B., P. Fariselli, and R. Casadio.** 1996. Topology prediction for helical transmembrane proteins at 86% accuracy. *Protein Sci.* **5**(8):1704-18.
216. **Russ, W. P., and D. M. Engelman.** 2000. The GxxxG motif: a framework for transmembrane helix-helix association. *J Mol Biol.* **296**(3):911-9.
217. **Russell, L. M., and H. Rosenberg.** 1979. Linked transport of phosphate, potassium ions and protons in *Escherichia coli*. *Biochem J.* **184**(1):13-21.
218. **Russell, L. M., and H. Rosenberg.** 1980. The nature of the link between potassium transport and phosphate transport in *Escherichia coli*. *Biochem J.* **188**(3):715-23.
219. **Saaf, A., L. Baars, and G. von Heijne.** 2001. The internal repeats in the Na<sup>+</sup>/Ca<sup>2+</sup> exchanger-related *Escherichia coli* protein YrbG have opposite membrane topologies. *J Biol Chem.* **276**(22):18905-7.
220. **Sahin Toth, M., J. Coutre, D. Kharabi, G. le Maire, J. C. Lee, and H. R. Kaback.** 1999. Characterization of Glu126 and Arg144, two residues that are



- indispensable for substrate binding in the lactose permease of *Escherichia coli*. *Biochemistry*. **38**(2):813-9.
221. **Sahin Toth, M., and H. R. Kaback.** 2001. Arg-302 facilitates deprotonation of Glu-325 in the transport mechanism of the lactose permease from *Escherichia coli*. *Proc Natl Acad Sci U S A*. **98**(11):6068-73.
222. **Sahin Toth, M., and H. R. Kaback.** 1993. Properties of interacting aspartic acid and lysine residues in the lactose permease of *Escherichia coli*. *Biochemistry*. **32**(38):10027-35.
223. **Sahin Toth, M., A. Karlin, and H. R. Kaback.** 2000. Unraveling the mechanism of the lactose permease of *Escherichia coli*. *Proc Natl Acad Sci U S A*. **97**(20):10729-32.
224. **Sahin Toth, M., M. C. Lawrence, and H. R. Kaback.** 1994. Properties of permease dimer, a fusion protein containing two lactose permease molecules from *Escherichia coli*. *Proc Natl Acad Sci U S A*. **91**(12):5421-5.
225. **Saibil, H. R.** 2000. Conformational changes studied by cryo-electron microscopy. *Nat Struct Biol*. **7**(9):711-4.
226. **Saier, M. H., Jr., B. H. Eng, S. Fard, J. Garg, D. A. Haggerty, W. J. Hutchinson, D. L. Jack, E. C. Lai, H. J. Liu, D. P. Nusinew, A. M. Omar, S. S. Pao, I. T. Paulsen, J. A. Quan, M. Sliwinski, T. T. Tseng, S. Wachi, and G. B. Young.** 1999. Phylogenetic characterization of novel transport protein families revealed by genome analyses. *Biochim Biophys Acta*. **25**(1):1-56.
227. **Saier, M. H. J., J. T. Beatty, A. Goffeau, K. T. Harley, W. H. Heijne, S. C. Huang, D. L. Jack, P. S. Jahn, K. Lew, J. Liu, S. S. Pao, I. T. Paulsen, T. T. Tseng, and P. S. Virk.** 1999. The major facilitator superfamily. *J Mol Microbiol Biotechnol*. **1**(2):257-79.
228. **Saiki, R. K., S. Scharf, F. Faloona, K. B. Mullis, G. T. Horn, H. A. Erlich, and N. Arnheim.** 1985. Enzymatic amplification of beta-globin genomic sequences and restriction site analysis for diagnosis of sickle cell anemia. *Science*. **230**(4732):1350-4.
229. **Salaun, C., P. Rodrigues, and J. M. Heard.** 2001. Transmembrane topology of PiT-2, a phosphate transporter-retrovirus receptor. *J Virol*. **75**(12):5584-92.
230. **Sambrook, J., E. F. Fritsch, and T. Maniatis.** 1989. *Molecular Cloning. A Laboratory Manual*, second ed, vol. 1. Cold Spring Harbor Laboratory Press, Cold Spring Harbor.
231. **Sanger, F., A. R. Coulson, B. G. Barrell, A. J. Smith, and B. A. Roe.** 1980. Cloning in single-stranded bacteriophage as an aid to rapid DNA sequencing. *J Mol Biol*. **143**(2):161-78.
232. **Sanger, F., S. Nicklen, and A. R. Coulson.** 1977. DNA sequencing with chain terminating inhibitors. *Proc Natl Acad Sci USA*. **74**:5463-5467.
233. **Sarsero, J. P., P. J. Wookey, and A. J. Pittard.** 1991. Regulation of expression of the *Escherichia coli* K-12 *mtr* gene by TyrR protein and Trp repressor. *J Bacteriol*. **173**(13):4133-43.
234. **Sayers, J. R., W. Schmidt, and F. Eckstein.** 1988. 5'-3' exonucleases in phosphorothioate-based oligonucleotide-directed mutagenesis. *Nucleic Acids Res*. **16**(3):791-802.
235. **Schindler, H., and J. P. Rosenbusch.** 1978. Matrix protein from *Escherichia coli* outer membranes forms voltage-controlled channels in lipid bilayers. *Proc Natl Acad Sci U S A*. **75**(8):3751-5.

236. **Scholten, M., and J. Tommassen.** 1994. Effect of mutations in the -10 region of the *phoE* promoter in *Escherichia coli* on regulation of gene expression. *Mol Gen Genet.* **245**(2):218-23.
237. **Schweizer, H., M. Argast, and W. Boos.** 1982. Characteristics of a binding protein-dependent transport system for *sn*-glycerol-3-phosphate in *Escherichia coli* that is part of the *pho* regulon. *J Bacteriol.* **150**(3):1154-63.
238. **Selker, E., K. Brown, and C. Yanofsky.** 1977. Mitomycin C-induced expression of *trpA* of *Salmonella typhimurium* inserted into the plasmid ColE1. *J Bacteriol.* **129**(1):388-94.
239. **Sen, K., J. Hellman, and H. Nikaido.** 1988. Porin channels in intact cells of *Escherichia coli* are not affected by Donnan potentials across the outer membrane. *J Biol Chem.* **263**(3):1182-7.
240. **Sen, K., and H. Nikaido.** 1991. Trimerization of an in vitro synthesized OmpF porin of *Escherichia coli* outer membrane. *J Biol Chem.* **266**(17):11295-300.
241. **Shelden, M. C., B. Dong, G. L. de Bruxelles, B. Trevaskis, J. Whelan, P. R. Ryan, S. M. Howitt, and M. K. Udvardi.** 2001. *Arabidopsis* ammonium transporters, AtAMT1;1 and AtAMT1;2 have different biochemical properties and functional roles. *Plant and Soil.* **231**:151-160.
242. **Shinagawa, H., K. Makino, and A. Nakata.** 1983. Regulation of the *pho* regulon in *Escherichia coli* K-12. Genetic and physiological regulation of the positive regulatory gene *phoB*. *J Mol Biol.* **168**(3):477-88.
243. **Singer, S. J., and G. L. Nicolson.** 1972. The fluid mosaic model of the structure of cell membranes. *Science.* **175**(23):720-31.
244. **Smith, M. W., and J. W. Payne.** 1992. Expression of periplasmic binding proteins for peptide transport is subject to negative regulation by phosphate limitation in *Escherichia coli*. *FEMS Microbiol Lett.* **79**(1-3):183-90.
245. **Sofia, H. J., V. Burland, D. L. Daniels, G. Plunkett, 3rd, and F. R. Blattner.** 1994. Analysis of the *Escherichia coli* genome. V. DNA sequence of the region from 76.0 to 81.5 minutes. *Nucleic Acids Res.* **22**(13):2576-86.
246. **Sola, M., F. X. Gomis Ruth, L. Serrano, A. Gonzalez, and M. Coll.** 1999. Three-dimensional crystal structure of the transcription factor PhoB receiver domain. *J Mol Biol.* **285**(2):675-87.
247. **Sonna, L. A., S. V. Ambudkar, and P. C. Maloney.** 1988. The mechanism of glucose 6-phosphate transport by *Escherichia coli*. *J Biol Chem.* **263**(14):6625-30.
248. **Sonntag, I., H. Schwarz, Y. Hirota, and U. Henning.** 1978. Cell envelope and shape of *Escherichia coli*: multiple mutants missing the outer membrane lipoprotein and other major outer membrane proteins. *J Bacteriol.* **136**(1):280-5.
249. **Sprague, G. F., Jr., R. M. Bell, and J. E. Cronan, Jr.** 1975. A mutant of *Escherichia coli* auxotrophic for organic phosphates: evidence for two defects in inorganic phosphate transport. *Mol Gen Genet.* **143**(1):71-7.
250. **Steed, P. M., and B. L. Wanner.** 1993. Use of the *rep* technique for allele replacement to construct mutants with deletions of the *pstSCAB-phoU* operon - evidence of a new role for the PhoU protein in the phosphate regulon. *J Bacteriol.* **175**(21):6797-6809.
251. **Stowell, M. H., A. Miyazawa, and N. Unwin.** 1998. Macromolecular structure determination by electron microscopy: new advances and recent results. *Curr Opin Struct Biol.* **8**(5):595-600.
252. **Subramaniam, S., and R. Henderson.** 2000. Molecular mechanism of vectorial proton translocation by bacteriorhodopsin. *Nature.* **406**(6796):653-7.



253. **Surin, B. P., G. B. Cox, and H. Rosenberg.** 1987. Molecular studies on the phosphate-specific transport system of *Escherichia coli*, p. 145-149. In S. Silver, A. Wright, and E. Yagil (ed.), Phosphate metabolism and cellular regulation in microorganisms. American Society for Microbiology, Washington DC.
254. **Surin, B. P., H. Rosenberg, and G. B. Cox.** 1985. Phosphate-specific transport system of *Escherichia coli*: nucleotide sequence and gene-polypeptide relationships. *J Bacteriol.* **161**(1):189-98.
255. **Suziedlien, E., K. Suziedlis, V. Garbencit, and S. Normark.** 1999. The acid-inducible *asr* gene in *Escherichia coli*: transcriptional control by the *phoBR* operon. *J Bacteriol.* **181**(7):2084-93.
256. **Tam, J. P.** 1988. Synthetic peptide vaccine design: synthesis and properties of a high-density multiple antigenic peptide system. *Proc Natl Acad Sci U S A.* **85**(15):5409-13.
257. **Thomas, G., G. Coutts, and M. Merrick.** 2000. The *glnKamtB* operon. A conserved gene pair in prokaryotes. *Trends Genet.* **16**(1):11-4.
258. **Tommassen, J., P. van der Ley, M. van Zeijl, and M. Agterberg.** 1985. Localization of functional domains in *E. coli* K-12 outer membrane porins. *Embo J.* **4**(6):1583-7.
259. **Torriani, A., and D. N. Ludtke.** 1985. *The Molecular Biology of Bacterial Growth.* Bartlett Publishers, Boston.
260. **Traxler, B., D. Boyd, and J. Beckwith.** 1993. The topological analysis of integral cytoplasmic membrane proteins. *J Membr Biol.* **132**(1):1-11.
261. **Tusnady, G. E., and I. Simon.** 1998. Principles governing amino acid composition of integral membrane proteins: application to topology prediction. *J Mol Biol.* **283**(2):489-506.
262. **Ugurbil, K., H. Rottenberg, P. Glynn, and R. G. Shulman.** 1982. Phosphorus-31 nuclear magnetic resonance studies of bioenergetics in wild-type and adenosinetriphosphatase(1-) *Escherichia coli* cells. *Biochemistry.* **21**(5):1068-75.
263. **Ulmschneider, M. B., and M. S. Sansom.** 2001. Amino acid distributions in integral membrane protein structures. *Biochim Biophys Acta.* **2**(1):1-14.
264. **Unwin, N.** 1998. The nicotinic acetylcholine receptor of the Torpedo electric ray. *J Struct Biol.* **121**(2):181-90.
265. **van der Ley, P., P. Burm, M. Agterberg, J. Meersbergen, and J. Tommassen.** 1987. Analysis of structure-function relationships in *Escherichia coli* K12 outer membrane porins with the aid of *ompC-phoE* and *phoE-ompC* hybrid genes. *Mol Gen Genet.* **209**(3):585-91.
266. **van der Ley, P., M. Struyve, and J. Tommassen.** 1986. Topology of outer membrane pore protein PhoE of *Escherichia coli*. Identification of cell surface-exposed amino acids with the aid of monoclonal antibodies. *J Biol Chem.* **261**(26):12222-5.
267. **van Gelder, P., N. Saint, P. Phale, E. F. Eppens, A. Prilipov, R. van Boxel, J. P. Rosenbusch, and J. Tommassen.** 1997. Voltage sensing in the PhoE and OmpF outer membrane porins of *Escherichia coli*: role of charged residues. *J Mol Biol.* **269**(4):468-72.
268. **van Veen, H. W., T. Abee, G. J. Kortstee, W. N. Konings, and A. J. Zehnder.** 1994. Translocation of metal phosphate via the phosphate inorganic transport system of *Escherichia coli*. *Biochemistry.* **33**(7):1766-70.
269. **van Veen, H. W., T. Abee, G. J. Kortstee, H. Pereira, W. N. Konings, and A. J. Zehnder.** 1994. Generation of a proton motive force by the excretion of metal-



- phosphate in the polyphosphate-accumulating *Acinetobacter johnsonii* strain 210A. *J Biol Chem.* **269**(47):29509-14.
270. **van Veen, H. W., T. Abee, G. J. J. Kortstee, W. N. Konings, and A. J. B. Zehnder.** 1993. Mechanism and energetics of the secondary phosphate transport system of *Acinetobacter-johnsonii-210A*. *J Biol Chem.* **268**(26):19377-19383.
271. **VanBogelen, R. A., E. R. Olson, B. L. Wanner, and F. C. Neidhardt.** 1996. Global analysis of proteins synthesized during phosphorus restriction in *Escherichia coli*. *J Bacteriol.* **178**(15):4344-66.
272. **Vanzeijl, M., S. V. Johann, E. Closs, J. Cunningham, R. Eddy, T. B. Shows, and B. Ohara.** 1994. A human amphotropic retrovirus receptor is a 2nd member of the gibbon ape leukemia virus receptor family. *Proc Natl Acad Sci USA.* **91**(3):1168-1172.
273. **Venkatesan, P., and H. R. Kaback.** 1998. The substrate-binding site in the lactose permease of *Escherichia coli*. *Proc Natl Acad Sci U S A.* **95**(17):9802-7.
274. **Versaw, W. K., and R. L. Metzenberg.** 1995. Repressible cation-phosphate symporters in *Neurospora crassa*. *Proc Natl Acad Sci USA.* **92**:3884-87.
275. **Viitanen, P., M. J. Newman, D. L. Foster, T. H. Wilson, and H. R. Kaback.** 1986. Purification, reconstitution, and characterization of the lac permease of *Escherichia coli*. *Methods Enzymol.* **125**:429-52.
276. **von Heijne, G.** 1992. Membrane protein structure prediction. Hydrophobicity analysis and the positive-inside rule. *J Mol Biol.* **225**(2):487-94.
277. **von Heijne, G.** 1994. Membrane proteins: from sequence to structure. *Annu Rev Biophys Biomol Struct.* **23**:167-92.
278. **Voss, J., J. Wu, W. L. Hubbell, V. Jacques, C. F. Meares, and H. R. Kaback.** 2001. Helix packing in the lactose permease of *Escherichia coli*: distances between site-directed nitroxides and a lanthanide. *Biochemistry.* **40**(10):3184-8.
279. **Wang, P., J. Yang, B. Lawley, and A. J. Pittard.** 1997. Repression of the *aroP* gene of *Escherichia coli* involves activation of a divergent promoter. *J Bacteriol.* **179**(13):4213-8.
280. **Wanner, B. L.** 1993. Gene regulation by phosphate in enteric bacteria. *J Cell Biochem.* **51**(1):47-54.
281. **Wanner, B. L.** 1996. Signal transduction in the control of phosphate-regulated genes of *Escherichia coli*. *Kidney Int.* **49**(4):964-7.
282. **Wanner, B. L., and J. A. Boline.** 1990. Mapping and molecular cloning of the *phn* (*psiD*) locus for phosphonate utilization in *Escherichia coli*. *J Bacteriol.* **172**(3):1186-96.
283. **Wanner, B. L., and M. R. Wilmes Riesenber.** 1992. Involvement of phosphotransacetylase, acetate kinase, and acetyl phosphate synthesis in control of the phosphate regulon in *Escherichia coli*. *J Bacteriol.* **174**(7):2124-30.
284. **Webb, D. C., H. Rosenberg, and G. B. Cox.** 1992. Mutational analysis of the *Escherichia coli* phosphate-specific transport system, a member of the traffic ATPase (or ABC) family of membrane transporters. A role for proline residues in transmembrane helices. *J Biol Chem.* **267**(34):24661-8.
285. **Wek, R. C., and G. W. Hatfield.** 1988. Transcriptional activation at adjacent operators in the divergent-overlapping *ilvY* and *ilvC* promoters of *Escherichia coli*. *J Mol Biol.* **203**(3):643-63.
286. **Williams, K. A.** 2000. Three-dimensional structure of the ion-coupled transport protein NhaA. *Nature.* **403**(6765):112-5.
287. **Williams, K. A., and C. M. Deber.** 1991. Proline residues in transmembrane helices: structural or dynamic role? *Biochemistry.* **30**(37):8919-23.

288. **Williams, K. A., U. Geldmacher-Kaufer, E. Padan, S. Schuldiner, and W. Kuhlbrandt.** 1999. Projection structure of NhaA, a secondary transporter from *Escherichia coli*, at 4.0 Å resolution. *Embo J.* **18**(13):3558-63.
289. **Willsky, G. R., R. L. Bennett, and M. H. Malamy.** 1973. Inorganic phosphate transport in *Escherichia coli*: involvement of two genes which play a role in alkaline phosphatase regulation. *J Bacteriol.* **113**(2):529-39.
290. **Willsky, G. R., and M. H. Malamy.** 1980. Characterization of two genetically separable inorganic phosphate transport systems in *Escherichia coli*. *J Bacteriol.* **144**(1):356-65.
291. **Willsky, G. R., and M. H. Malamy.** 1976. Control of the synthesis of alkaline phosphatase and the phosphate-binding. *J Bacteriol.* **127**(1):595-609.
292. **Willsky, G. R., and M. H. Malamy.** 1974. The loss of the PhoS periplasmic protein leads to a change in the specificity of a constitutive inorganic phosphate transport system in *Escherichia coli*. *Biochem Biophys Res Commun.* **60**(1):226-33.
293. **Wilson, C. A., M. V. Eiden, W. B. Anderson, C. Lehel, and Z. Olah.** 1995. The dual-function hamster receptor for amphotropic murine leukemia virus (MuLV), 10A1 MuLV, and gibbon ape leukemia virus is a phosphate symporter. *J Virol.* **69**(1):534-7.
294. **Wilson, K. S., and P. H. von Hippel.** 1995. Transcription termination at intrinsic terminators: the role of the RNA hairpin. *Proc Natl Acad Sci U S A.* **92**(19):8793-7.
295. **Winkler, H. H.** 1966. A hexose-phosphate transport system in *Escherichia coli*. *Biochim Biophys Acta.* **117**(1):231-40.
296. **Wolin, C. D., and H. R. Kaback.** 1999. Estimating loop-helix interfaces in a polytopic membrane protein by deletion analysis. *Biochemistry.* **38**(26):8590-7.
297. **Wolin, C. D., and H. R. Kaback.** 2001. Functional estimation of loop-helix boundaries in the lactose permease of *Escherichia coli* by single amino acid deletion analysis. *Biochemistry.* **40**(7):1996-2003.
298. **Wu, J., S. Frillingos, and H. R. Kaback.** 1995. Dynamics of lactose permease of *Escherichia coli* determined by site-directed chemical labeling and fluorescence spectroscopy. *Biochemistry.* **34**(26):8257-63.
299. **Wu, J., and H. R. Kaback.** 1994. Cysteine 148 in the lactose permease of *Escherichia coli* is a component of a substrate binding site. 2. Site-directed fluorescence studies. *Biochemistry.* **33**(40):12166-71.
300. **Wu, J., D. M. Perrin, D. S. Sigman, and H. R. Kaback.** 1995. Helix packing of lactose permease in *Escherichia coli* studied by site-directed chemical cleavage. *Proc Natl Acad Sci U S A.* **92**(20):9186-90.
301. **Wu, J., J. Voss, W. L. Hubbell, and H. R. Kaback.** 1996. Site-directed spin labeling and chemical crosslinking demonstrate that helix V is close to helices VII and VIII in the lactose permease of *Escherichia coli*. *Proc Natl Acad Sci U S A.* **93**(19):10123-7.
302. **Xavier, K. B., M. Kossmann, H. Santos, and W. Boos.** 1995. Kinetic analysis by *in vivo* <sup>31</sup>P nuclear magnetic resonance of internal Pi during the uptake of *sn*-glycerol-3-phosphate by the *pho* regulon-dependent Ugp system and the *glp* regulon-dependent GlpT system. *J Bacteriol.* **177**(3):699-704.
303. **Yamada, H., N. Oshima, T. Mizuno, H. Matsui, Y. Kai, H. Noguchi, and S. Mizushima.** 1987. Use of a series of OmpF-OmpC chimeric proteins for locating antigenic determinants recognized by monoclonal antibodies against the OmpC



- and OmpF proteins of the *Escherichia coli* outer membrane. *J Biochem.* **102**(3):455-64.
304. **Yamada, M., K. Makino, M. Amemura, H. Shinagawa, and A. Nakata.** 1989. Regulation of the phosphate regulon of *Escherichia coli*: analysis of mutant *phoB* and *phoR* genes causing different phenotypes. *J Bacteriol.* **171**(10):5601-6.
305. **Yerushalmi, H., and S. Schuldiner.** 2000. A model for coupling of H<sup>(+)</sup> and substrate fluxes based on "time-sharing" of a common binding site. *Biochemistry.* **39**(48):14711-9.
306. **Zhang, P., C. Toyoshima, K. Yonekura, N. M. Green, and D. L. Stokes.** 1998. Structure of the calcium pump from sarcoplasmic reticulum at 8-Å resolution. *Nature.* **392**(6678):835-9.
307. **Zhao, M., K. C. Zen, W. L. Hubbell, and H. R. Kaback.** 1999. Proximity between Glu126 and Arg144 in the lactose permease of *Escherichia coli*. *Biochemistry.* **38**(23):7407-12.
308. **Zhou, F. X., M. J. Cocco, W. P. Russ, A. T. Brunger, and D. M. Engelman.** 2000. Interhelical hydrogen bonding drives strong interactions in membrane proteins. *Nat Struct Biol.* **7**(2):154-60.
309. **Zhou, J. J., L. J. Trueman, K. J. Boorer, F. L. Theodoulou, B. G. Forde, and A. J. Miller.** 2000. A high affinity fungal nitrate carrier with two transport mechanisms. *J Biol Chem.* **275**(51):39894-9.
310. **Zhou, Y., J. H. Morais Cabral, A. Kaufman, and R. MacKinnon.** 2001. Chemistry of ion coordination and hydration revealed by a K<sup>+</sup> channel-Fab complex at 2.0 Å resolution. *Nature.* **414**(6859):43-8.
311. **Zhuang, J., G. G. Prive, G. E. Werner, P. Ringler, H. R. Kaback, and A. Engel.** 1999. Two-dimensional crystallization of *Escherichia coli* lactose permease. *J Struct Biol.* **125**(1):63-75.
312. **Zubay, G.** 1980. The isolation and properties of CAP, the catabolite gene activator. *Methods Enzymol.* **65**(1):856-77.