NOVEL NUCLEOSIDE TRANSPORT SYSTEMS IN THE INNER MEMBRANE OF *ESCHERICHIA COLI*

MARIANNA RALLI

Thesis presented for the degree of Doctor of Philosophy

University of Edinburgh

2006

DECLARATION

I declare that the composition of this thesis and the research within is my own work except where otherwise stated.

ACKNOWLEDMENTS

First of all I would like to thank my supervisor Dr Maurice Gallagher for giving me the chance to work on this project and for guiding me through the difficulties of it. Also I would like to thank my second supervisor Dr Millie Masters for her help, views and support. Many thanks also go to my friends and colleagues in our lab and other laboratories in Darwin and Swann buildings for their willingness to help, enlighten and joke when needed.

A special thank you goes to Prof Ken Murray and the Darwin Trust for funding this project. Without them this would not have been possible.

A special thank you also goes to Dr Steve Baldwin and all the people in his group in the University of Leeds for their hospitality and for having me over and helping me through various issues of this work.

I cannot thank Dr Christophe Merlin enough for his help and for enlightening me in various subjects of science. His enthusiasm, positive attitude and charisma in teaching me made me love the field of molecular biology.

I would also like to give a big thank to my mother and father for supporting me through the years I lived in Edinburgh, for their love and invaluable advice in everything I did. Also I would like to thank my sister Myrto for giving me inspiration and power to continue with our discussions over everything. This thesis is dedicated to them.

Finally a big part of getting through with this work is thanks to my friends who are my life and joy. Adam, Irini, James, Katerina, Michali, Shirin, Wes, Yana, Zoë, here is to you.

CONTENTS

FIGURES and TABLES

CHAPTER 1- INTRODUCTION

CHAPTER 3- Construction of mutants lacking nucleoside transport systems

ABBREVIATIONS

Å	Angstroms
aa	amino acid
ABC	ATP- binding cassette
Ap^R	Ampicillin resistant
ATP	adenosine triphosphate
AZT	Zidovudine (3'-azido-2', 3'-dideoxythymidine)
bp	base pairs
CAT	Chloramphenicol acetyltransferase
CFTR	cystic fibrosis transmembrane regulator
CFU	colony forming unit
CNT	concentrative nucleoside transporter
Ct	C terminal
Da	Daltons
ddC	zalcitabine
ddI	2'3'-dideoxyinosine
DMSO	Dimethyl sulfoxide
ELISA	Enzyme linked immunosorbant assay
ENT	equilibrative nucleoside transporter
FLP	flip
FRT	*flp* Recombination Targets
GFP	green fluorescent protein
HIV	human immunodeficiency virus
HR	homology regions
IPTG	isopropyl-beta-D-thiogalactopyranoside
KAN^R	kanamycin resistant
Kb	kilobases
LB	Luria-Bertani medium
LPS	lipopolysaccharide
LUM	luminescence
MDR	multidrug resistance transport protein
MFS	major facilitator superfamily
NAD	nicotinamide adenine dinucleotide
NBMPR	Nitrobenzylthioinosine
Nt	N-terminus
OD	optical density
ORF	open reading frame
PFU	plaque forming unit
PHX	predicted highly expressed
PS1, PS2	priming site regions
PTS	phosphotransferase system
RLU	relative light units
SEAP	secretable alkaline phosphatase
TC	Transporter Classification
TMS	transmembrane segments
URFs	unidentified reading frames
VIC	voltage-gated ion channel

ABSTRACT

Nucleoside transporters mediate cellular uptake of physiological nucleosides for nucleic acid synthesis in the salvage pathways in many cell types. These transporters also play an important role in *in vivo* disposition and intracellular targeting of many nucleoside analogs used in anticancer and antiviral drug therapy. Nucleosides and their derivatives are very important compounds for various functions of the cell. They are used as core molecules for nucleic acid production, and as components of cofactors such as NAD, thiamine and folic acid. They may also be used as a direct source of carbon mainly ribose and deoxyribose.

Transport proteins mediating nucleoside transport are classified into two groups: one group catalyses facilitated diffusion down a concentration gradient (ENT, equilibrative nucleoside transporters), whereas the members of the other group (CNT, concentrative nucleoside transporters) are able to transport nucleosides against an existing concentration gradient. Thus, the activity of the latter requires metabolic energy. Both sodium and proton coupled nucleoside transport catalysed by CNT type carriers have been identified. (Griffith and Jarvis 1996, Baldwin *et al,* 1999).

Nucleosides are transported across the inner membrane of *Escherichia coli* by at least two separate energy-driven systems known as NupC and NupG. These systems are encoded by the genes *nupC* and *nupG* respectively.

CNT1 (a mammalian transporter isolated from the jejunum) is a homologue of NupC. Interestingly, CNT1 is able to transport the nucleoside analogue AZT, used to combat HIV infection. CNT1 and NupC share substrate specificity and they have amino acid sequence identity of 27%, particularly in their C-terminal region. With the completion of the *E. coli* genome project paralogues of NupC and NupG have been found, designated as YeiM and YeiJ, and XapB and YegT respectively (Maurice P Gallagher-unpublished data). It is worth noting that XapB has been proposed to be involved in xanthosine catabolism (Seeger, *et al,*1995).

To functionally characterise the putative nucleoside transporter- encoding genes *xapB, yegT, yeiM* and *yeiJ* and study whether they are involved in nucleoside transport as the *nupG* and *nupC* products are, mutants were constructed in which the coding sequence of these genes had been deleted from their chromosome. The mutants acquired were used to investigate if genes *xapB, yegT, yeiM* and *yeiJ* were centrally involved in nucleoside transport in *E. coli*.

From this work it was established that all four putative nucleoside transporters XapB, YegT, YeiM and YeiJ were able to transport adenosine in the assays at similar levels to the positive control. The use of complementation strains in experiments involving nucleoside uptake as a sole carbon source, demonstrated that the collection of genes studied in this project were indeed nucleoside permeases. Furthermore in kinetic analysis of the transporters it was possible to determine Km values for the four transporters with adenosine as substrate.

Following the work on gene characterisation on a functional level the mutant strains were used to construct a bioassay, in which their ability to transport nucleoside analogues was assessed. AZT was used as model substrate since it has been studied before and it is a well known compound used to combat HIV infections. Preliminary experiments gave an insight in developing a high throughput bioassay which can be developed further in order to produce a useful tool for nucleoside analogue drugs assessment.

Chapter 1
Introduction

1. INTRODUCTION

An essential feature of each cell is the presence of membranes that define the boundaries of the cell and delineate its various internal compartments. One of the main functions of the membrane is the regulation of movement of solutes in and out of the cell. It is not enough to think of membranes as permeability barriers though, since it is crucial for the proper function of the cell to be able to overcome the permeability barrier for specific molecules and ions so that they can be moved selectively in the appropriate direction (in or out of the cell). In other words membranes serve not only as barriers to the movement of substances into and out of the cell but also, they provide the controlled passage of specific molecules and ions.

In the following introductory part of this project, focus is given on the membrane structure, so that it can then be related to the ways substances are selectively moved across the membrane and the significance of such transport processes to the life of the cell. Furthermore transporters will be discussed according to an extensive system (the TC system) that classifies all of them. Particular attention will be given to active transport and more specifically, to nucleoside transporters that are the main topic of this work.

In addition to a broad review of nucleoside transporters, their use and the use of nucleoside analogues as agents for disease therapy will also be given. Also a review of contemporary use of bioassays and biosensors will be outlined since part of the project deals with an attempt to relate the use of therapeutic nucleoside analogues with an efficient bioassay prototype. Finally a brief review on gene replacements is given since the study of nucleoside transport in this project involved a considerable amount of gene manipulation and gene deletions from the chromosome of *Escherichia coli*.

1.1 Membrane: Structure, function and chemistry

Membranes have a variety of functions and their structure and chemical features make them suitable for these roles. Membranes define the boundaries of the cell and define its compartments, they serve as loci of specific functions, provide for and regulate the movement of substances into and out of the cell and its compartments, contain the receptors required for the detection of external signals and, provide mechanisms for cell to cell communication.

The modern approach to describing the membrane structure is as a semi-fluid "sea" of phospholipids with proteins "floating" in it. This contemporary concept can be understood best if seen in the course of its scientific evolution through time. Figure 1.1 gives a chronological summary of membrane studies that started over a century ago with the understanding that lipid layers are parts of membranes. Today scientists investigate membrane proteins and their roles and structure.

A starting point in the historical overview is the work done by Charles Overton in the 1890s. Overton was aware that cells seemed to be enveloped by some sort of selectively permeable layer that allowed the passage of some substances but not others. He reasoned that the ability of a substance to cross the membrane might be related in some way to its chemical affinity for the membrane. Working with cells of plant root hairs, Overton found that lipid-soluble substances penetrated readily into cells, whereas water-soluble substances did not. From his studies Overton concluded that lipids were present on the cell surface as some kind of "coat" (Benos *et al,* 1999). A second important advance came about a decade later through the work of Irving Langmuir (Langmuir, 1917) who studied the behaviour of lipids by spreading them out as a thin layer on a water surface of a so called Langmuir trough. He dissolved the lipids in benzene and layered samples of the lipid-benzene solution onto the surface of the water. When the benzene evaporated the molecules were left as a lipid monolayer on the water surface with their hydrophilic heads in the water and their hydrophobic tails in the air. The lipid monolayer became the basis for further thought about membrane structure in the early years of the twentieth century.

In 1925, two Dutch physiologists E. Gorter and F.Grendel read Langmuir's papers and thought that his approach might help answer a question regarding the surface of the red blood cell, with which they worked. When Gorter and Grendel used Langmuir's method to determine the surface area of the erythrocytes, they found that the monolayer of lipid on the water was about twice that estimated (Gorter and Grendel, 1925). They therefore concluded that the erythrocyte plasma membrane consisted of two monolayers of lipid, giving rise to the important concept of the lipid bilayer. Assuming a bilayer structure Gorter and Grendel reasoned that it would be thermodynamically favourable for the nonpolar hydrocarbon chains of each layer to face inward, away from the aqueous environment on either side of the membrane. The polar hydrophilic groups would face outward, towards the aqueous environment. The lipid bilayer that the two scientists envisioned became the basic underlying assumption for each successive refinement in our understanding of membrane structure until today.

Though a main feature of the membrane structure, the lipid bilayer could not explain all the properties of membranes particularly those related to solute permeability and electrical resistance. The permeability of membranes quite often depends on very subtle chemical differences and it was quite difficult for early membrane scientists to imagine how a lipid bilayer could be capable of such differential permeability.

However, soon scientists began to wonder whether membrane components with binding sites similar to enzymes (for which specificity had been established by that time), might be present in the membrane to facilitate transport. Similarly biological membranes are much more permeable to ions than artificial lipid bilayers. To explain these additional features, Hugh Davson and James Danielli invoked the presence of proteins in membranes, proposing in 1935 that biological membranes consist of lipid bilayers coated on both sides with a layer or lamellae of proteins (Danielli and Davson, 1935). The original Davson-Danielli model was really a protein-lipid "sandwich".

With the arrival of electron microscopy in the 1950s it became clear that most subcellular organelles of eukaryotic cells are bounded by membranes. J. David Robertson suggested the first general structure for all membranes, the unit membrane (Robertson, 1957). In this model, membranous components of the cell were regarded as a continuous lipid bilayer, with functional proteins adhering to the surface. When first proposed, the unit membrane appeared to agree remarkably well with the Davson-Danielli model: the electron micrographs they were getting showed two dark lines equated with the outer protein layers which appeared electron dense due to their affinity for the stains, while the space seen in between was thought to correspond to the hydrophobic part of the lipid molecules, which presumably did not take up the stain.

An attractive feature of the unit membrane model was its apparent universality; membranes from widely different types of cells showed the same characteristic features when examined by electron microscopy. However the model did not explain the distinctiveness of different kinds of membranes. As more membranes were studied it became more and more difficult to explain the varying findings according to the unit membrane model. As membrane proteins were isolated it became apparent that most of them were globular proteins with sizes and shapes that were inconsistent with the theory of thin layers or sheets of proteins on the two surfaces of the membrane. These findings along with the others on membrane fluidity caused biologists to rethink membrane structure. Eventually in 1972, the fluid mosaic model was first proposed by S. Jonathan Singer and Garth Nicolson (Singer and Nicolson. 1972). The two key features of the model are that the basic lipid bilayer structure is retained but membrane proteins are not thought of as continuous layers on the membrane surface, but are recognised as globular entities that associate with the membrane, to an extent dictated by their affinity for the hydrophobic interior of the lipid bilayer. Also rather than being rigidly locked in place, the lipid components of a membrane are in constant motion, capable of movement parallel to the membrane surface.

The fluid mosaic model recognised two broad categories of membrane proteins that differ in their affinity for the hydrophobic interior of the membrane and therefore, in the extent to which they interact with the lipid bilayer. Integral membrane proteins have one or more hydrophobic regions with an affinity for the hydrophobic interior of the lipid bilayers and because of their close association to the membrane they cannot be solubilised readily without treatment with a detergent. In contrast peripheral membrane proteins lack discrete hydrophobic sequences and therefore do not penetrate into the lipid bilayer. Instead they associate with the membrane surfaces through weak electrostatic forces, binding either to the polar heads of the membrane lipids or to the hydrophilic portions of integral proteins that extend out of the membrane. Peripheral proteins can be removed from the membrane much easier than integral proteins and can be solubilised by aqueous salt solutions.

In the 1970s scientists began to understand an important property of membrane proteins, that each of the hydrophobic sequences of an integral membrane protein is a transmembrane segment that spans the lipid bilayer anchoring the protein to the membrane and holding it in proper alignment within the lipid bilayer.

Few, if any, cells' membranes only comprise of the plasma membrane. Most have some kind of structure that is external to the membrane but is nonetheless an integral part of the cell. These structures differ among organisms, but in most eukaryotes they contain long, rigid fibres embedded in an amorphous matrix of molecules that are usually glycoproteins or polysaccharides. In animal cells this extracellular structure is primarily made of fibres of collagen embedded in a network of glycoproteins called proteoglycans. In plants, algae, and prokaryotes the extracellular structure is called the cell wall. In higher plants the cell wall mainly comprises of cellulose embedded in a complex network of polysaccharides called hemicelluloses or pectins.

The cell walls of plants and other organisms bring rigidity to the cells, they protect cells from physical damage and attack by viruses and infectious organisms. In animal cells the extracellular matrix forms a wide variety of structures and performs a variety of functions depending on the tissue or organ in which it is found.

A bacterial cell is also surrounded by a cell wall, consisting of complex polysaccharides that provide support and protection to the cell. The bacterial cell wall also protects the cell from bursting due to osmotic pressure and from viral invasion (Salton 1987). Also it serves as a permeability barrier and plays a role in cell-cell recognition and adhesion. However the bacterial cell wall differs strikingly from the plant cell wall. In fact the chemical nature of the cell wall differs among bacterial cells too, and this difference provides the basis of classifying bacteria into two wide categories, based on their response to the Gram stain. The critical difference between the two groups is the chemical nature of the cell wall. The cell wall of Gram-positive bacteria (for example, that of the representative Gram-positive species *Staphylococcus aureus*) is a multilayered, cross-linked network of peptidoglycans located external to the plasma membrane of the cell. These peptidoglycans consist of long chains of N-acetylglycosamine (GlcNAc) and N-acetylmuramic acid (MurNAc) in strictly alternating order, cross-linked by short peptides. The peptide cross-links consist of a tetrapeptide group on each MurNAc that is joined to adjacent tetrapeptides, usually by a chain of five glycines. Also the cell wall of Gram-positive bacteria contains teichoic acids, which are polymers of glycerol or a 5-carbon alcohol called ribitol. The teichoic acid polymers protrude from the membrane through the wall and are cross-linked to the peptidoglycan backbone. Typically the cell wall of a Gram-positive bacterium consists of 20-40 peptidoglycan layers and is about 50nm thick (Antony and Hill, 1988).

The cell wall of Gram-negative bacteria also contains peptidoglycans but is much thinner, consisting typically of just one or a few peptidoglycan layers, which are bound to the inner side by the plasma membrane and covered by an outer membrane. The outer leaflet of the outer membrane is composed of a lipopolysaccharide layer providing the cell with an efficient barrier against penetration of lipophilic agents such as antibiotics and other toxic substances (Nikaido, 1994). The tetrapeptide chains attached to the MurNAc units are joined to one another directly by a peptide bond. Because the cell wall of a Gram-negative bacterium is bounded by the plasma membrane on one side and the outer membrane on the other, its location is described as the periplasmic space. The thin cell wall is porous enough to allow molecules as

large as proteins to move within the periplasmic space. Especially important in this regard are the periplasmic proteins that act as shuttles for amino acids and sugars. These solutes cross the outer membrane through protein channels called porins, and then they are shuttled across the wall by the periplasmic proteins and finally are delivered to the transport proteins in the plasma membrane (Salton, 1987, Williams, 1988, Hendler 197).

Transport proteins located in the plasma membrane play a key role in facilitating and controlling the movement of molecules and ions into and out of cells. Further details on transport modes are given below, followed by a review of transport systems through classification, in order to get an idea of their vast diversity, specialization and abundance across all species of living organisms.

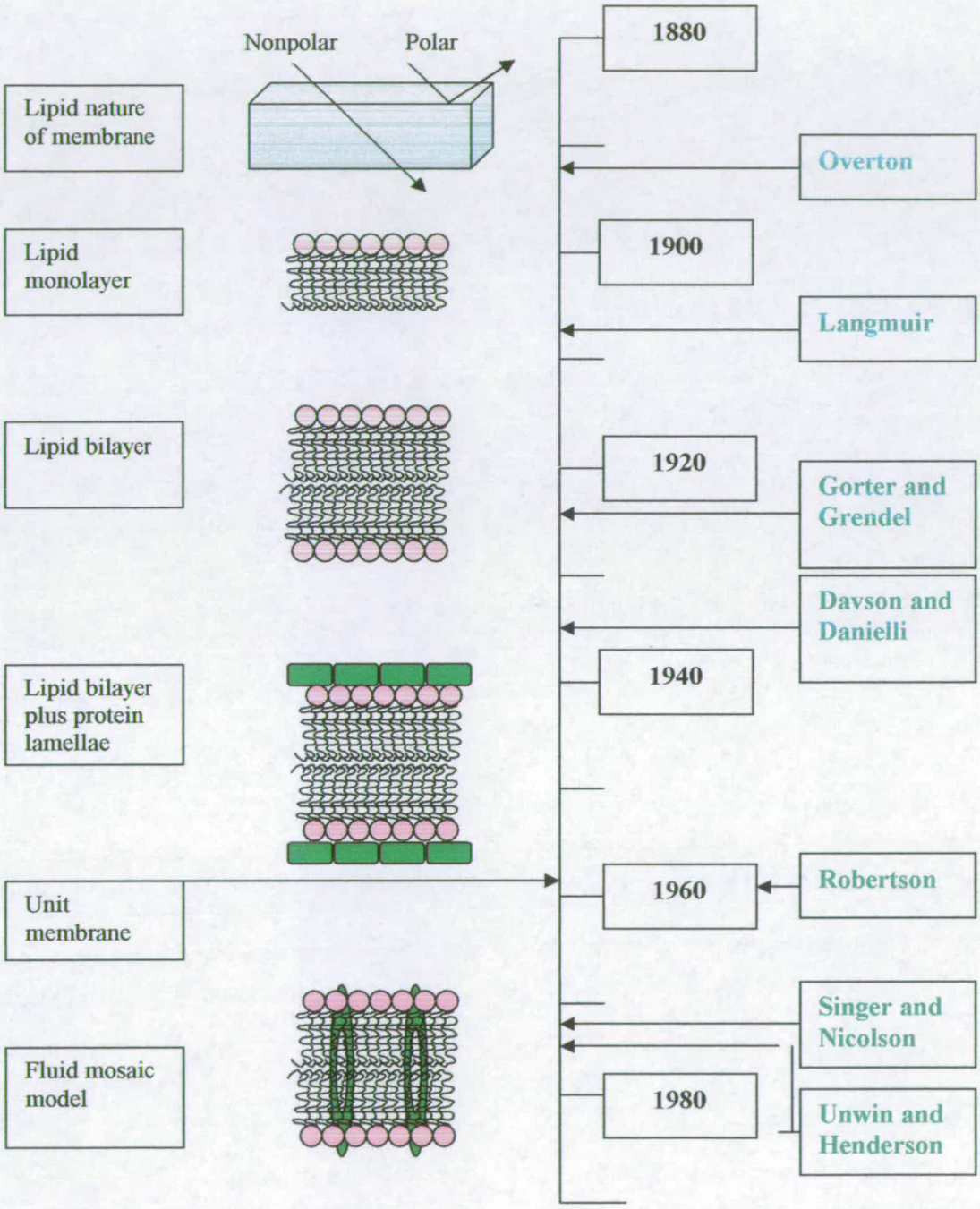


Figure 1.1: (modified from "The world of the cell" by Becker *et al,* 2000) **Timeline for the development of our understanding of membrane chemistry and structure.**

1.1.1 Solute transport across membranes

A central role of every cell and subcellular compartment is the ability to accumulate a variety of substances at concentrations that are significantly different from those in their surrounding environment. The substances that move across membranes are not only macromolecules and fluids but also ions and small organic molecules. Some of the most common ions are sodium, potassium, calcium, chloride and hydrogen ions. Most of the small organic molecules are metabolites, in other words substrates, intermediates and products of metabolic pathways (Ames, 1986).

A central aspect of cell function is transport, the ability to move ions, and organic molecules across membranes selectively. Except for a few small or non-polar molecules the transport of ions and metabolites across membranes is largely possible because of the presence of membrane proteins that can recognise substances with great specificity and mediate their movement across the membrane.

Membranes form major barriers to the free movement of solutes into and out of cellular compartments. Sometimes the barrier is necessary for cell function, such as the maintenance of ion gradients. In other cases the barrier must be overcome so that a solute can move into or out of the cell (Sadee *et al*, 1999).

Transport of solutes across a membrane can be divided into two major categories according to the requirement for energy. Passive transport does not require energy because of the tendency for dissolved molecules to move or diffuse from higher to lower concentrations. Most passive transport depends on specialised membrane proteins to facilitate the passage of ions or polar molecules across the hydrophobic interior of the membrane. However some molecules can diffuse across the membrane unaided because they are very small, non-polar or both. Such molecules include water, ethanol, O_2, CO_2 etc. Passive transport, whether it is facilitated or unaided by a membrane protein, is always exergonic (Becker *et al,* 2000).

Active transport on the other hand requires the input of energy and in the absence of an energy source such transport would not be feasible. Active transport is catalysed by specialised membrane proteins that are often referred to as pumps. Active transport of ions is especially important as the transport process and resulting concentration gradient can produce an electrical voltage or membrane potential across the membrane. The concentration gradient together with the membrane potential is referred to as the electrochemical gradient. The stored energy of such gradients is used to drive a variety of processes in cells, such as ATP synthesis during cellular respiration and photosynthesis in animal and plant cells respectively (Hernandez and Fischbarg, 2005).

1.1.2 Passive transport

Many molecules and ions move exergonically across membranes into and out of cells and organelles. Passive transport can be thought of as a diffusion process, in which a flux of molecules or ions occurs continuously across a membrane in both directions with net movement in the direction dictated by the concentration (or for ions the electrochemical) gradient until equilibrium is reached (Nikaido and Saier, 1992).

Some substances can cross the membrane unaided because they are so small or nonpolar, via a process known as simple diffusion. In simple diffusion a small, (relatively) non-polar molecule that is in aqueous solution on one side of the membrane simply permeates the phospholipid bilayer, diffuses passively across the bilayer and emerges on the other side in aqueous solution. Simple diffusion is only possible for solutes that are readily permeable in the membrane.

Three primary properties of solutes must be taken under consideration when thinking of the bilayer's permeability towards solutes. These properties are the relative size of the molecule, the polarity and the ion permeability. Generally speaking for lipid bilayers molecules up to the size of glucose (180 daltons) are able to diffuse across

membranes at reasonable rates (Becker *et al,* 2000). Also lipid bilayers are relatively permeable to nonpolar molecules. Nonpolar molecules dissolve more readily in the hydrophobic phase of the lipid bilayer which increases their ability to cross the membrane barrier. Finally lipid bilayers are not very permeable to ions, because a great deal of energy is required to move ions from an aqueous environment into a nonpolar environment. Individual molecules diffuse randomly in both directions, but net flux is always down the concentration gradient for the molecule in question.

Most substances in cells are too large or too polar to cross the membrane by simple diffusion, and move into and out of cells with the assistance of specialised membrane proteins. This is called facilitated diffusion because the solute still diffuses from higher to lower concentration but its movement is facilitated by a specific transport protein in the membrane.

Most substances required by cells are polar or charged and move across membranes only if cells have specific means of facilitating that movement. Facilitated diffusion therefore differs from simple diffusion in that specific proteins are required to support the passage of particular molecules. Facilitated diffusion uses the same driving force as simple diffusion that is the movement from a higher to a lower concentration. However transport proteins facilitate the diffusion of polar or charged solutes across an otherwise impermeable barrier.

Transport proteins are integral membrane proteins embedded in the membranes of cells. These proteins have at least two different domains, one to anchor the protein within the appropriate membrane and the other to recognise, bind and translocate the solute. (Nikaido and Saier, 1992). Transport proteins or permeases become saturated as the concentration of the transportable solute rises, because the number of transport proteins is limited and each has a limited maximum velocity at which it can function. As a result permease-facilitated transport follows Michaelis-Menten kinetics, with an upper velocity and a Michaelis constant. Transporters bind one or more solute molecules on one side of the membrane and then undergo a conformational change that transfers the solute to the other side of the membrane.

Channel proteins on the other hand, form hydrophilic channels through the membrane that allow the passage of solutes without any change in the conformation of the protein. Some of these channels are quite large and non-specific, such as the pores found in the outer membranes of bacteria. Pores are formed by transmembrane proteins called porins and allow selected hydrophilic solutes with molecular weight of about 600 Da to diffuse across the membrane Channel proteins may be gated, however (Jap and Walian, 1996).

1.1.3 Active transport

Active transport involves the movement of substances into or out of cells against a concentration or electrochemical gradient. The pumps involved in active transport always require the input of energy. In other words, active transport is thermodynamically unfavourable and occurs only when coupled with an exergonic process. Therefore the membrane proteins involved in active transport must provide not only for translocation of the desired solute molecules across the membrane but also, for coupling of that translocation to an energy-yielding reaction (Baldwin, 1990).

Active transport makes the uptake of essential nutrients from the environment possible even when the concentrations in the surroundings of the cell might be much lower than inside the cell. Also, active transport allows various substances and waste products to be removed from the cell, regardless of the concentration outside the cell. Thirdly it enables the cell to maintain constant, non-equilibrium intracellular concentrations of specific inorganic ions, such as potassium, hydrogen, and calcium. This is a very important aspect of active transport. Whereas passive transport results in an increase in similarity of the outer and internal cellular environment, active transport establishes differences in solute concentration and electrical potential across membranes and lead to a non-equilibrium steady state. Active transport has directionality unlike passive transport, which is inherently non-directional.

An overview of the properties of active and passive transport can be found in table 1.1. Also passive and active transport are further analysed below in view of their classification.

1.1.3.1 Direct (primary) active transport and indirect (secondary) active transport

Active transport mechanisms differ primarily in the source of energy that is being used and whether two solutes are transported concurrently. Based on the energy source active transport is regarded as either direct or indirect (Saier 2000).

In direct active transport (or primary active transport) the accumulation of solute molecules or ions on one side of the membrane is coupled directly to an exergonic chemical reaction, most likely the hydrolysis of ATP. Transport proteins that are driven by ATP hydrolysis are known as transport ATPases.

There are at least four types of transport ATPases known which are responsible for most direct active transport in both prokaryotic and eukaryotic cells. These are P (phosphorylation)-type ATPases, V (vesicle)-type, F (factor) - type and ABC-type cassettes (Saier 2000).

P-type ATPases are reversibly phosphorylated by ATP as part of the transport mechanism. They have 8-10 transmembrane segments in a single polypeptide that winds back and forth across the membrane and all of them are cation transporters. Mostly they are found in eukaryotic cells and they are located in the plasma membrane where in most cells, they maintain the ion gradient.

V-type ATPases pump protons into organelles such as vesicles, lysosomes and the Golgi apparatus. They have two multi-subunit components, an integral component

embedded within the membrane and a peripheral component that protrudes from the membrane surface.

Properties	Passive Transport		Active Transport
	Simple Diffusion	Facilitated Diffusion	
Solutes transported			
Small polar (e.g. oxygen)	Yes	No	No
Large nonpolar (e.g. Fatty acids)	Yes	No	No
Small polar (e.g. Water)	Yes	No	No
Large polar (e.g. Glucose)	No	Yes	Yes
Ions (e.g. N^+, K^+, Ca^{2+})	No	Yes	Yes
Thermodynamic properties			
Direction relative to electrochemical gradient	Down	Down	Up
Effect on entropy	Increased	Increased	Decreased
Metabolic energy required	No	No	Yes
Intrinsic directionality	No	No	Yes
Kinetic properties			
Carrier-mediated	No	Yes	Yes (pump)
Michaelis-Menten kinetics	No	Yes	Yes
Competitive inhibition	No	Yes	Yes

Table 1.1 (Based on a diagram from "The world of the cell" by Becker *et al*, 2000) Properties of Passive and Active transport

F-type ATPases are found in bacteria, mitochondria and chloroplasts where they are an integral part of the mechanism that conserves the energy of solar radiation or of substrate oxidation. They are involved in proton transport and they have two main components, both of which are multi-subunit complexes. These are known as the F_0 and the F_1 components. F-type ATPases can use energy of ATP hydrolysis to pump protons against their electrochemical gradient, and they can also catalyse the reverse process to drive ATP synthesis. These proteins therefore function as ATP synthases too, being involved in cellular energy metabolism.

The fourth major class of ATP-driven pumps is the ABC-type ATP or ABC transporters (ATP-binding cassette), which here annotates the fact that the protein binds ATP as an integral part of the transport process. ABC transporters are a large superfamily mostly belonging to prokaryotic species but also being reported for eukaryotes. The typical ABC transporter has four domains, two of which are embedded in the membrane and are highly hydrophobic, and two peripherally located on the cytoplasmic side of the membrane. Whereas the other three classes of ATPases transport only cations, the ABC transporters handle a remarkable variety of solutes, including ions, sugars, amino acids, peptides and polysaccharides. ABC transporters are of considerable medical interest as some pump antibiotics or other drugs, making the cell drug resistant. In fact the earliest ABC transporter identified in eukaryotes is called the multidrug resistance transport protein (MDR), which uses the energy of ATP hydrolysis to pump hydrophobic drugs out of the cells and therefore reduces the cytoplasmic concentration of the drugs and hence the drug's effectiveness. John Walker received a Nobel Prize in 1997 for his work on enzymes involved in ATPase structure. He found out that ATPase function is based on an enzymatic rotary mechanism which takes energy from the proton motive force (Abrahams *et al,* 1994).

Indirect active transport (or secondary active transport) depends on the co-transport of two solutes with the movement of one solute down its gradient driving the movement of the other solute up or down its gradient. The transport process is either symport or antiport, depending on whether the two solutes move in the same or

opposite direction. In most cases, one of the two solutes is an ion (usually Na^+ or H^+) which moves down its electrochemical gradient. The second solute is driven against a concentration gradient. In animal cells, sodium symport is the usual option, whereas in many bacteria, fungi and plants, the driving force is often an electrochemical proton gradient.

The uptake of such compounds is regarded as indirect active transport because it is driven not by the hydrolysis of ATP but by the pre-established electrochemical gradient of sodium or hydrogen ion.

The *Escherichia coli* lactose-permease can be considered as the prototype of H^+ coupled solute secondary active transport systems in bacteria and because it utilises H^+ produced by primary active transport it is referred to as a secondary active transporter (Linton and Higgins, 1998). The lactose permease symporter uses the proton gradient established across the *E. coli* membrane by fuel oxidation and electron transport (during primary active transport). The lactose permease LacY, is solely responsible for all translocation reactions catalyzed by the galactoside transport system in *E. coli*. Like many members of the MFS LacY couples the free energy released from downhill translocation of protons in response to a proton electrochemical gradient to drive the energetically uphill lactose transport (Kaback and Wu, 1999). The crystal structure of LacY has been unravelled by Abramson *et al,* (2003) and is reviewed by Vazquez-Ibar *et al,* (2003). Therefore a proton motive force established from the electron transport chain drives further secondary active transport of lactose. The principal coupling ion in animal cells appears to be Na^+ rather than H^+ (Ganapathy and Leibach 1991). The electrochemical Na^+ gradient (sodium-motive force), the energy source for the Na^+-coupled transport systems, which exists across the animal cell plasma membrane, is generated by the Na^+-K^+ pump. In recent years, evidence has accumulated which indicates that the difference in the preferred inorganic ion between bacterial (H^+) and animal (Na^+) systems is not absolute. A sizeable number of Na^+-coupled solute transport systems are known in bacteria especially in marine organisms. Likewise, a number of organic-solute

transport systems in animal cells have been shown to be coupled to H^+ rather than Na^+.

Prior to the discovery of the existence of energy-transducing mechanisms in the animal cell plasma membrane to generate the proton-motive force, there appeared to be no reason to question the idea of Na^+ being the sole coupling ion in solute transport in animal cells. With the identification of the Na^+-H^+ exchanger and the H^+ pump and with the impressive number of H^+-coupled solute transport systems, which are now known to be present in the animal cell plasma membrane, there is enough reason to believe that H^+ is also an important coupling ion for solute transport in animal cells (Canapathy and Leibach 1991).

1.2 Classification of transporters

Transport systems serve the cell in a number of ways. Firstly they allow entry of essential minerals into the cytoplasmic region and subsequently into the organelles. They provide a means for the regulation of metabolite concentrations by catalysing the excretion of end or waste products of metabolic pathways from cells and organelles. Also they mediate the active excretion of drugs and other toxic substances from the cytoplasm across the plasma membrane (Sadee *et al*, 1999).

Transport systems also mediate the influx and efflux of ions that must be maintained at concentrations that differ significantly from those in the external environment of the cell. A membrane potential, ion concentration gradients and appropriate cytoplasmic concentrations of all essential minerals that participate as cofactors in metabolic process are central to maintenance of life. These conditions are required for the generation of bioelectricity as well as for the maintenance of enzymatic activities.

Transporters also play roles in movement of larger molecules such as proteins, complex carbohydrates, and lipids into and beyond the cytoplasmic membrane and

those macromolecules serve a variety of biologically important roles in environmental protection, in communication with members of the same and other species, in degrading foodstuffs, in adhesion, in pathogenesis and, in many other roles.

Indeed transport of nucleic acid polymers across cell membranes, allowing genetic exchange between organisms has been pivotal in promoting diversification and the format of biological cells and organisms, as we know them today.

In multicellular organisms clusters of cells have often evolved to form specialised organs that communicate with other regions of the organisms via the release, uptake or detection of transported molecules such as pheromones, hormones, neurotransmitters and a variety of other signalling molecules.

Transport of organic molecules across cell membranes proceeds via specialised active and passive transporters. The completion of genome sequences and the information obtained from the vast array of biochemical and molecular genetics investigations, and the powerful *in silico* analysis resulting from such information, have revealed many aspects of commonality amongst transport systems. This is discussed more extensively below (Busch and Saier, 2002).

1.2.1 Transport nomenclature

Prior to 1969 when Ron Kaback (Kaback, 1969) provided the first clear evidence for protein-mediated solute transport, the lactose permease, the nature of how solutes entered into the cells was undefined. Over 30 years later it is now clear that all transmembrane transport processes, other than diffusion, are mediated by integral membrane proteins, sometimes functioning in conjunction with extracytoplasmic receptors or receptor domains and/ or with cytoplasmic energy-coupling and

regulatory proteins or domains. Each such complex of proteins or protein domains is referred to as a transport system, transporter, permease or permease system.

Indeed enzymatic processes were considered a likely strategy and hence the term "permease". A permease is a protein or protein complex that catalyses a chemical or electron transfer reaction that drives the vectorial process. Thus many transport systems can be thought of as catalytic proteins or protein complexes, analogous to enzymes or enzyme complexes.

Passive transport can occur by equilibrative protein-mediated diffusion. This is not coupled to metabolic energy and cannot give rise to concentration gradients of the transported substrate across the membrane. In such instances the substrate flows from high to low concentration across the membrane until equilibrium is reached. Where elements of specificity are provided by the transporter or carrier, this is referred to as facilitated diffusion.

Facilitated transport can be subdivided into two categories of biological systems, channel type and carrier type. In channel type facilitated diffusion the solute passes in a diffusion limiting process from one side of the membrane to the other via a channel or pore. The structures of several such channel proteins have now been examined and the proteins crystallised (Saier 2000, Doyle, 2004, Kuo *et al,* 2005) On the other hand in carrier-type facilitated diffusion some part of the transporter is presumed to pass through the membrane together with the substrate.

Carriers usually exhibit rates of transport that are several orders of magnitude lower than those of channels. Moreover they also exhibit stereo-specific substrate specificities as opposed to most channels. Although both channels and carriers exhibit the phenomenon of saturation kinetics, this is a more common characteristic of carriers. Very few carriers have been shown to be capable of functioning by a channel-type mechanism and the few that do, generally do so only after the protein has been modified. Observations such as these lead to the conclusion that channels and carriers are fundamentally different.

Recapping on energy expenditure if energy is coupled to transmembrane solute translocation, then a system can become an active transporter. Such a system is considered to be a primary active transporter if a primary source of energy is being utilised (i.e. a chemical reaction, light absorption or electron flow). The system is considered a secondary active transporter if a secondary source of energy is coupled to the solute at the expense of the primary energy source (for example an ion electrochemical gradient, such as proton motive force or sodium motive force which is established in the first instance by the action of the electron transport chain).

Secondly active transporters couple solute transport to the input of energy and they are divided into two classes: ion-coupled and ATP-dependent transporters. Ion-coupled transporters are coupled to the cotransport of H^+, Na^+, Cl^- and the countertransport of K^+. ATP-dependent transporters are directly energised by the hydrolysis of ATP and translocate a heterogeneous set of substrates (Hediger 1994).

Active transporters can function by uniport, symport or antiport mechanisms. Uniporters catalyse the transport of a single molecular species and transport occurs independently of the movement of other molecular species. Symporters, also classically known as cotransporters, mediate the transport of two or more molecular species in the same direction. A single point mutation (Busch and Saier, 2002) in a symporter can convert a carrier into a uniporter, emphasising the fine distinction between those two types of carriers. Antiporters catalyse the transmembrane exchange of one or more molecular species for another.

Antiport processes can be subdivided into antiport of like molecules and antiport of unlike molecules. Many uniporters and symporters also catalyse solute-solute antiport, sometimes at rates that are significantly greater than those for uniport or symport. Some carriers catalyse solute-solute antiport at rates that exceed those of uniport or symport by 10^3 to 10^5- fold and uniport via these carriers is of little or no physiological consequences. Such systems are said to be obligatory antiporters or exchangers (Busch and Saier, 2002).

Accelerative solute-solute antiport or counter transport has been considered to be a diagnostic characteristic of carriers. Early transport kineticists concluded that its demonstration eliminated the possibility that a transporter functions by channel-type mechanism and suggested clear boundaries exist between carriers and channels. A few carriers modify their substrates during transport. The best characterised such system is the bacterial phosphotransferase system (PTS) which is an example of a group translocation system (Busch and Saier, 2002).

1.2.2 The TC System

Early studies revealed that transport proteins could be grouped into families based solely on the degrees of similarity observed in their amino acid sequence. Later however it was revealed that some of the families had arisen independently of each other at different times in evolutionary history, and following different routes.

The availability of complete catalogues of genomic sequences and extensive databases of gene sequences have provided an immense catalogue of transport system information. This has been the subject of extensive bioinformatics analysis. Perhaps the most influential of such analytical reviews has been the recent work of Milton Saier's group (Saier, 2000). According to the classification of transporters as described by Saier, (2000) the Transporter Classification (TC) system, transporters are grouped on the basis of five criteria and each of these criteria corresponds to one of the five entries within the TC number for a particular permease. These factors include the transporter mode of transport, the energy coupling mechanisms, the permease family or superfamily, the phylogenetic cluster within the family (or a family within the superfamily) and finally, the substrate or range of substrates transported, as well as the polarity of transport. Any two transport proteins in the same subfamily of a permease family that transport the same substrate, using the same mechanism, are given the same TC number regardless of whether they are

paralogues (homologue genes belonging in the same species) or orthologues (homologue genes belonging in different species).

1.2.2.1 Categories of the TC system

The TC system was approved by the transport nomenclature panel of the International Union of Biochemistry and Molecular Biology in Geneva in 1999. A summary can be seen table 1.2 below. There are categories that have not been assigned with any entry so that they can be filled in when novel types of transporters are discovered. The TC system is constantly updated and scientists can add their findings on line. An updated form of the TC system is available on the Web at http://www.biology.ucsd.edu/~msaiertransport/titlepage.html.
The four major classes of transporters that have been recognised are channels, porters, primary active transporters and group translocators. The primary level of classification in the TC system is based on mode of transport and energy coupling source. All categories can be seen in table 1.2 below.

Category 1 comprises channels and pores, includes α-Type channels, β-Barrel porins, pore forming toxins and non-ribosomally synthesized channels. Lately a fifth subclass was added which includes holins. This category includes transmembrane pores and channels found in all types of organisms from bacteria to higher eukaryotes. Transport in this category is an energy independent process.

Category 2 includes the electrochemical potential-driven porters. Subcategories of category 2 include porters (these can be uniporters, symporters and antiporters), non-ribosomally synthesised porters and ion gradient-driven energisers. Category 2 includes the primary active transporters that use a primary source of energy to drive active transport of a solute against a concentration gradient. Primary energy sources known to be coupled to transport are chemical, electrical, and solar.

Category 3 includes diphosphate bond hydrolysis-driven transporters, decarboxylation-driven transporters, methyl transfer-driven transporters, oxidoreduction-driven transporters and light absorption-driven transporters.

Category 4 includes group translocator systems (GTS) and in particular the phosphotransfer-driven group translocators or PTS families. These families are only identified in bacterial species and no porters of the GTS have been identified in the archaeal or eukaryotic domain. Substrates for these transporters are sugars such as glucose, fructose, lactose, glucitol, galactitol, mannose and sorbose.

Category 5 is a newly added group that includes transmembrane electron transfer carriers, such as oxidoreductases.

Notably no assignment has been made for categories 6 and 7. These are reserved for novel types of transporters, yet to be discovered which do not fall within categories 1 to 5.

Category 8 includes all accessory factors involved in transport; usually proteins that in some way facilitate transport across one or more biological membranes but do not participate directly in transport. These proteins always function directly in conjunction with one or more established transport systems. Recently this category became more specific in that it is the one that contains auxiliary transport proteins.

Finally category 9 includes incompletely characterised transport proteins, which fall into transporters of unknown biochemical mechanisms, putative but uncharacterised transport proteins and finally functionally characterised transport protein with unidentified sequences.

Thus the comprehensive system of permease classification proposed by Saier has the potential to encompass all types of transporters, both those that are recognised and those that are yet to be discovered.

1.2.2.2 Analysis of TC system categories

Category 1 of the TC system includes all channels and pores. In particular these include α-Type channels, β-Barrel porins, pore-forming toxins and non-ribosomally synthesised channels.

α-Type channels are transmembrane proteins that are found in membranes of all types of organisms from bacteria to higher eukaryotes. These transporters usually catalyse the movement of solutes by an energy-independent process via a passage through a transmembrane aqueous pore, without evidence of carrier-mediated mechanisms. These channel proteins consist largely of α-helical "spanners", although β-strands may be present and may even contribute to the channel. Such channels may be gated as a means of control however.

In Gram-negative bacteria a class of proteins called porins forms non-specific diffusion channels across the outer membrane to permit entry of nutrients from the environment in to the periplasm. Although composed of β-barrel structures these are relatively hydrophilic proteins approximately 36-38 kDa which lack the long stretches of hydrophobic segment found typically in membrane transport proteins, such as mammalian Na^+ /glucose co transporters. Crystal structures of porins such as OmpF in bacteria reveal they may form trimers of identical subunits. Each subunit contains relatively short membrane –spanning segments (approximately 16 amino acids long) that traverse the lipid bilayer.

β-Barrel porins are proteins forming transmembrane pores that usually allow the energy independent passage of solutes across a membrane. The transmembrane portions of these proteins consist exclusively of β-strands that usually form β-barrels. Porin-type proteins are found in the outer membranes of Gram-negative bacteria, mitochondria, plastids and acid-fast Gram-positive bacteria.

Pore-forming toxins are proteins and peptides synthesised by one cell and secreted for insertion into the membrane of another cell where they form transmembrane

pores. They exert their toxic effect by allowing free flow of electrolytes and other small molecules across the membrane. Alternatively they may allow entry into the target cell cytoplasm of a toxin protein that ultimately kills or controls the cell. Protein toxins and ribosomally synthesised peptides are included in this category.

Non-ribosomally synthesised channels also lie within this category. These molecules are chains of L- and D-amino acids, as well as other small molecular building blocks such as hydroxy acids, and they form oligomeric transmembrane ion channels.

Category 2 includes electrochemical potential-driven porters. In this category lie porters, non-ribosomally synthesised porters and ion-gradient driven energizers. Porters include transport systems that utilise a carrier-mediated process to facilitate uniport (a single species is transported either by facilitated diffusion or in a membrane potential-dependent process if the solute is charged), antiport (where two or more species are transported in opposite directions in a tightly coupled process, not coupled to a direct form of energy other than chemiosmotic energy) or symport (two or more species are transported together in the same direction in a tightly coupled process, not coupled to a direct form of energy other than chemiosmotic energy).

Non-ribosomally synthesised porters may be depsipeptides or non-peptide-like substances. Such a porter complexes a solute, such as a cation, in its hydrophilic interior and facilitates translocation of the complex across the membrane by exposing its hydrophobic exterior and moving the solute from one side of the bilayer to the other. If the free porter can cross the membrane in the uncomplexed form, the transport process can be electrophoretic i.e. the charged molecule moves down its electrochemical gradient. If only the complex can cross the membrane, transport may be electroneutral because one charged substrate is exchanged for another.

Finally in category 2 there are ion gradient-driven energisers. Normally, outer membrane porins of Gram-negative bacteria catalyse passive transport of solutes across the membrane, but coupled to energisers they may accumulate their substrates

in the periplasm against large concentration gradients. These energisers use the proton motive force across the cytoplasmic membrane, probably by allowing the electrophoretic transport of protons, conveying conformational changes to the outer membrane receptors or porins.

In category 3 transporters use a primary source of energy to drive active transport. In this category there are diphosphate bond hydrolysis-driven transporters, decarboxylation-driven transporters, methyl transfer-driven transporters, oxidoreduction-driven transporters and light absorption-driven transporters.

Transport systems that hydrolyse the diphosphate bond of inorganic pyrophospate, ATP or another nucleoside triphosphate to drive the active uptake and extrusion of solutes fall into the subclass of diphosphate bond hydrolysis-driven transporters. The transport protein may or may not be transiently phosphorylated but the substrate is not phosphorylated. These kinds of transporters are found universally in all kinds of organisms.

Transport systems that drive solute (ions for example) uptake or extrusion by decarboxylation of a cytoplasmic substrate, form the second subclass of this category. Currently these transporters are thought to be restricted to prokaryotes.

Coenzyme M methyltransferase is the sole protein that falls into the subclass of methyl transfer-driven transporters. The energy used is Na^+ ions and this type of transporter is thought to be restricted to members of the *Archaea*.

Transport systems that drive transport of a solute (e.g. an ion) energised by the exothermic flow of electrons from a reduced substrate to an oxidised substrate are included in the subclass of oxidoreduction-driven transporter. These transporters are universal, although some families are restricted to one domain.

The last subclass in category 3 comprises the light absorption transporters, which utilise light energy to drive transport of a solute. One family is found in archaea and eukaryotes but the other is found only in bacteria and chloroplasts of eukaryotes.

As mentioned above category 4 of the TC systems includes just one subclass of phosphotransfer-driven group translocators. This basically includes the phosphotransferase systems (PTS) for sugar uptake in bacteria.

Category 5 includes all transmembrane electron transfer carriers, such as the phagocyte NADPH oxidase family and the eukaryotic cytochrome b561 (Cytb561) family. This category also includes proteins such as tumour suppressors (Tsp10 pending protein).

Category 8 includes all accessory factors involved in transport. These auxiliary transport proteins may provide a function connected with energy coupling to transport, play a structural role in complex formation, or serve a biogenic or stability function in regulation.

Finally category 9 includes all incompletely characterised transport proteins. Transporters of unknown biochemical mechanism are included in the first subclass of this category. These families will be reclassified to appropriate groups once the mode of transport and energy-coupling mechanism of their members have been characterised. These families include at least one member for which a transport function has been established but either the mode of transport or the energy-coupling mechanism is not yet known.

Putative transport protein families are grouped in the second subclass of putative but uncharacterised transport proteins. These families include members for which a transport function has been suggested but evidence for such a function is not yet compelling.

Finally transporters of particular physiological significance are included in this last subclass of functionally characterised transport proteins with unidentified sequences. When sequences are identified transporters are assigned to established families. This protein subclass includes individual proteins rather that protein families.

A more detailed analysis of some of the most well known transporter families is presented below. Emphasis is given to the families that include nucleoside transporters, which are the focus of this work. As mentioned above in table 1.2 there is also a comprehensive version of the classes and subclasses of transporters with some examples for each type.

1.2.2.3 Families of transporters and distinctive examples

There are more than 250 entries in Saier's classification system of transport system families. Some of these families are large superfamilies with more than a thousand currently sequenced members such as the major facilitator superfamily (MFS), the voltage-gated ion channel (VIC) and the ATP- binding cassette (ABC) transporters. Most families are of intermediate size with between 5 and 500 sequenced members. However others are small families with only one or a few sequenced members.

The criteria used to set the limits of a family have been defined in the past (Saier, 2000). In brief, for two proteins to belong to the same family they must exhibit a region of 60 residues or more in comparable portions of the two proteins that have a comparison score in a excess of 9 standard deviations. At this value the probability that the degree of sequence similarity that occurs for those two proteins is by chance is less than 10^{-9}. The number of 60 residues was arbitrarily selected because many protein domains in water-soluble proteins are of about this size.

In Category 1 that includes channels and pores some typical examples include the holin functional superfamily proteins that are found in phage and bacteria. The lysis

protein S of phage lambda is a typical example of these transporters. In this category there are also various outer membrane porins and export proteins such as Tsx, a nucleoside-specific, channel-forming outer membrane porin. Also, porins OmpC and OmpF which allow the quick passage of small hydrophilic molecules in the outer membrane form a family belonging to this group of porins. Finally various toxins ranging from haemolysins of *E. coli*, such as HlyA, apoptosis regulators of *Homo sapiens* and colicins E1 and V of *E. coli*, belong in this category too (Baty *et al,*. 1990)

In particular Tsx is a substrate-specific channel. The 272-residue Tsx protein is a minor component of the *E. coli* outer membrane but it is essential for the efficient uptake of deoxyribonucleosides and nucleosides at submicromolar concentrations. The Tsx protein enables the bacterial cell to efficiently scavenge these substrates from the environment. Characterisation of Tsx has revealed that it has 14 membrane spanning β-strands which are highly conserved (Nieweg and Bremer, 1997).

The second category of transporters known as electrochemical potential- driven transporters includes some well-known families of transporters. In particular the major facilitator superfamily (MFS) that includes the lactose permease of *E. coli* LacY and the drug efflux permease of *E. coli* EmrD, belongs here.

The MFS also called the uniporter -symporter-antiporter family includes transporters that are single-polypeptide secondary carriers capable of transporting small solutes in response to chemiosmotic ion gradients. Together with the ABC superfamily the MFS account for nearly half of the solute transporters encoded within the genomes of microorganisms but members are also prevalent in higher organisms (Marger and Saier, 1993).

The MFS was originally believed to function primarily in the uptake of sugars. Follow up studies showed that drug efflux systems and transporters of Krebs cycle metabolites belong to this family too. Then phosphate exchangers and oligosaccharide H^+ symport permeases were added. So far 29 families have been

assigned to the MFS superfamily. The MFS transporters transport a wide array of substrates with either inwardly or outwardly directed polarity or without polarity (Pao *et al,* 1998). These substrates include sugars, drugs, an array or metabolites, amino acids, nucleosides, vitamins, and both organic and inorganic cations.

Such compounds are also transported by the ABC type permeases but ABC transporters can also transport macromolecules such as proteins, carbohydrates and phospholipids. Although some MFS transporters do transport polymers their sizes are much smaller, limited to about 4 amino acyl or sugar residues. Several MFS families are restricted to certain types of organisms, namely Gram-negative bacteria and yeasts. Pao *et al,* (1998) proposed that most of the individual MFS families diverged from each other more than 2 billion years ago before the bacteria diverged from eukaryotes and archaea.

Within the same category there are also two families of transporters specific for transporting nucleosides. These belong to the second category and are called the concentrative nucleoside transporters (CNT) and the equilibrative nucleoside transporters (ENT). These two are analysed further in a later part of this review as part of the nucleoside transporters section (Hyde *et al*, 2001).

The ATP binding cassette (ABC) superfamily is a large and diverse group of proteins, whose members mediate a wide range of transport functions. ABC proteins use energy derived from ATP hydrolysis to drive transport and modulate the activity of heterologous channels. The ABC family permeases are generally multicomponent in bacteria. However important eukaryotic examples such as P-Glycoproteins transporters and CFTR are monocomponent. Nevertheless, they are identifiable by their similarity in overall structure and by the presence of the ATP-binding cassette. Thus they equate with multicomponent systems in which the individual components have been fused at the genetic level to produce large multi-domain proteins. (Theodoulou, 2000)

Over 100 ABC proteins have been identified to date, in taxa ranging from bacteria to humans. The completion of the yeast and *Escherichia coli* genomes has revealed 29 and 79 ABC transporters respectively (Decottignies and Goffeau, 1997, Linton and Higgins, 1998).

Substrates assigned to members of this large family include compounds which are as varied as peptides, sugars, lipids, heavy metal chelates, polysaccharides, alkaloids, steroids, inorganic acids and glutathione conjugates (Theodoulou 2000). Such diversity is reflected in the size of this group. While many ABC transporters are specific, others are able to transport chemically unrelated compounds. ABC transporters have been extensively characterised because of their roles as prokaryotic periplasmic permeases, involved in bacterial nutrient uptake.

However recently much interest has been given to members of this group because of their clinical significance such as the cystic fibrosis transmembrane regulator (CFTR) and transporters from humans (P-glycoprotein), parasites, yeast and bacteria that mediate multiple drug resistance (MDR). Cellular resistance to anti-microbial agents or anti- neoplastic drugs via drug export, is one of the reasons why treatments of infectious diseases or malignant tumours fail (Lage, 2003). Multidrug transporters mediate a unidirectional transport of a broad range of compounds across biological membranes. They are members of a limited number of families, and they have been shown to exist in all organisms that have been studied. They can be divided into two major classes. The first class is composed of primary active transporters and they are energised by ATP hydrolysis. These primary transporters are represented by the superfamily of ATP-binding cassette (ABC) (Dean 2005). The second class consists of secondary active transporters that mediate drug efflux across membranes in a coupled exchange with protons or sodium ions down a concentration gradient as symport or antiport. Members of this second class include the MFS groups, as well as others (Lage, 2003). ABC- transporters play the most important role in drug efflux in eukaryotic cells including human cancer cells, whereas in prokaryotic organisms the secondary active multidrug transporters are predominantly responsible for drug extrusion (Lage, 2003).

Classes and subclasses of transporters in the TC system	Examples
1. Channels and pores	
1.A α-type channels	VIC family
1.B β-Barrel porins	*Tsx*
1.C Pore-forming toxins (proteins and peptides)	Colicin E1 (*E. coli*)
1.D Non-ribosomally synthesised channels	
*1.E Hollins	
2. Electrochemical potential-driven transporters	
2.A Porters (uniporters, symporters, antiporters)	MFS, *CNT, ENT, NupG, NupC*
2.B Non ribosomally synthesised porters	
2.C Ion gradient-driven energisers	Outer membrane energisers of *E. coli* (TonB)
3. Primary active transporters	
3.A Diphosphate bond hydrolysis-driven transporters	ABC
3.B Decarboxylation-driven transporters	Oxaloacetate decarboxylase (*Salmonella*)
3.C Methyl transfer-driven transporters	Co-enzyme M
3.D Oxidoreduction-driven transporters	NDH
3.E Light absorption-driven transporters	
4. Group Translocators	
4.A Phosphotransfer-driven group translocators	PTS
5. Transport electron carrier	Tumour suppressors, NADPH oxidase family
*5.A Transmembrane 2-Electron transfer carriers	
*5.B Transmembrane 1- Electron transfer carriers	
8. Accessory factors involved in transport	Energy coupling transporters, regulation related transporters
*8.A Auxiliary transport proteins	
9. Incompletely characterised transport systems	
9.A Recognised transporters of unknown biochemical mechanisms	Lipopolysaccharide exporters (*E. coli*)
9.B Putative but uncharacterised transport proteins	
9.C Functionally characterised transporters lacking identified sequences	

Table 1.2: the TC classification system according to Saier 2000.
Some deviation from the review paper indicated by asterisks is based on current updates published on the web. No assignment has been made for categories 6 and 7, which are reserved for novel transporters yet to be discovered. Nucleoside transporters are highlighted by italics on this table.

1.3 Nucleoside transport systems

Transport across human, animal, fungal and bacterial cell membranes has been intensively characterised in the past at the physiological and molecular level. Nucleosides are molecules consisting of a purine or a pyrimidine base linked to a five-carbon sugar (a ribose or a deoxyribose) through a glycosylic bond. Nucleosides and their derivatives are very important compounds for various functions of the cell. They are used as core molecules for nucleic acid production, and as components of cofactors such as NAD, thiamine and folic acid. They may also be used as a direct source of carbon mainly ribose and deoxyribose (Neuhard and Nygaard, 1987).

Transport proteins mediating nucleoside transport are classified into two groups: one group catalyses facilitated diffusion down a concentration gradient (ENT, equilibrative nucleoside transporters), whereas the members of the other group (CNT, concentrative nucleoside transporters) are able to transport nucleosides against an existing concentration gradient. Thus, the activity of the latter requires metabolic energy and both sodium and proton coupled nucleoside transport catalysed by CNT type carriers have been identified. (Griffith and Jarvis 1996, Baldwin *et al,* 1999).

ENTs are also known as facilitated diffusion systems because they mediate a faster flux of nucleosides across the lipid bilayer than would occur by simple diffusion. In mammals equilibrative transporters (ENT) are found in a wide variety of cells and they can be subdivided into two groups according to their sensitivity to inhibition by a ribonucleoside called nitrobenzylthioinosine (NBMPR). Transporters that are inhibited by the potent action of NBMPR are called equilibrative sensitive (*es*) whereas those that are not affected by NBMPR, are called equilibrative insensitive (*ei*). Both transporter types have broad substrate specificity for pyrimidine and purine nucleosides (Crawford *et al,* 1998).

CNTs are membrane proteins that catalyse the co-transport (symport) of a nucleoside and an ion, either sodium or hydrogen across the cell membrane. Because ion

gradients can drive the transport of nucleosides against the concentration gradient, CNTs are also known as active transporters. Concentrative nucleoside transporters (CNT) have a limited distribution among cells types and are insensitive to NBMPR. They can be divided into three categories based on their substrate specificity. The concentrative transporters, insensitive to NBMPR, which accept thymidine as a permeant are designated as *cit* and they accept pyrimidine nucleosides as well as adenosine, although adenosine is a poor substrate. The *cif* transporters (concentrative, insensitive to NBMPR, accepting formycin) accept purine nucleosides and uridine. Finally there are the *cib* transporters (concentrative, insensitive to NBMPR, with broad range of permeants) that accept both purine and pyrimidine nucleosides (Baldwin *et al,* 1999).

In animals concentrative nucleoside transporters have a more limited tissue distribution than equilibrative transporters and have been mainly found in specialised cells such as intestinal and renal epithelia and liver cells. Generally members of the family with *cit* –type activity are designated CNT1, those with *cif*-type activity CNT2 and those with *cib*-type activity CNT3 (Smith *et al*, 2004, Ritzel *et al,* 2001a). Genome sequencing projects have revealed that CNT family members are present in eukaryotes including *Drosophila melanogaster* and *Caenorhabditis elegans* (Xiao *et al,* 2001) and also in prokaryotes. The best characterised and paradigm member of a bacterial CNT, is the nucleoside transporter NupC which was cloned from *E. coli* (Craig *et al,* 1994).

The first examples of CNTs were identified in rat tissues (Huang *et al,* 1994). A cDNA from rat jejunal epithelium was found to encode a Na^+ dependent nucleoside transporter, which was named CNT1 (Huang *et al,* 1994). Also two more Na^+ dependent subtypes implicated in adenosine transport were identified in rat jejunal and kidney tissues after expression in *Xenopus* oocytes (Yao *et al,* 1996). Hamilton *et al,* (2001) later proposed a topological model of rCNT1, which possesses 13 transmembrane helices with the hydrophilic N-terminal and C-terminal domains on the cytoplasmic and extracellular sides of the membrane, respectively. Also work by Yao and collegues revealed the first transport protein to be characterised in detail in

the Pasific hagfish (*Eptatretus stouti*). The transporter, designated hfCNT1 is a Na^+ nucleoside symporter and revealed the same range of kinetic properties as recombinant mammalian CNT proteins (Yao *et al,* 2002). Also a PCR- based strategy was used to isolate a mouse cDNA, for a sodium-dependent, purine nucleoside concentrative transporter (mCNT2). The genomic organisation and chromosomal localisation of mCNT2 was also elucidated in that work (Patel *et al,* 2000). The human CNT1 contains 650 amino acids and is 83% identical to the rat CNT1. The human CNT2 is 72% identical to rat CNT2 and the human CNT3 78% identical to rat CNT3. (Yao *et al,* 2002, Ritzel *et al,* 2001b). Based on functional studies in cell and tissue preparations, five Na^+-dependent nucleoside transport subtypes (NT1-5) were recognized (reviewed by Kong *et al,* 2004). The CNT1 (or *cif*) subtype is mainly purine-selective but also transports uridine. The CNT2 (or *cit*) subtype is pyrimidine-selective but also transports adenosine. The CNT3 subtype (or *cib*) is broadly selective, transporting both purine and pyrimidine nucleosides. The CNT4 subtype transports pyrimidine nucleosides and the purine nucleosides adenosine and guanosine. CNT4 was characterized in brush border membrane vesicles isolated from human kidney. The CNT5 subtype, which is NBMPR sensitive and transports formycin B and cladribine, was described only in human leukemia cell lines. To date, genes encoding CNT1-CNT3 transporters have been cloned and belong to a CNT gene family (Loewen *et al,* 1999, Loewen *et al*, 2004, Hamilton *et al,* 2001, Toan *et al,* 2003). Genes encoding CNT4 and CNT5 activities have not been identified (Kong *et al,* 2004).

Nucleosides have important roles in human physiology. An example of their roles is adenosine's potent effect in neurotransmission or regulation of cardiovascular activity (Mubagwa *et al,* 1996; Shryock and Belardinelli, 1997). In bone marrow cells, enterocytes, and cells lacking biosynthetic pathways for purine nucleosides, nucleosides are readily scavenged from the extracellular fluids in order to make nucleic acids (Josephsen *et al*, 1983). This is energy efficient for the cell, as it is more energy consuming to make nucleosides *de novo*.

The function of nucleoside import into animal cells is mainly to allow nucleotide synthesis via salvage pathways in cells lacking the capacity for *de novo* synthesis. This is also important in pathogenic protozoans that lack the capacity for purine synthesis (Carter *et al,* 2001). Moreover, adenosine functions as an extra-cellular molecular messenger and neurotransmitter, and is known to cause cell toxicity and immunosuppression, to influence cell morphology, to act as a vasodilator, and to stimulate hormone secretion (Mubagwa *et al,* 1996; Shryock and Belardinelli, 1997). Therefore, transport of these compounds plays an important role in regulation of cellular signalling, regulation of metabolism and development (Centelles *et al,* 1992).

In contrast, knowledge about the function of nucleosides in plant cells and tissues is much more limited. In general, plant cells harbour enzymes allowing nucleoside salvage and some of the genes coding for these enzymes have been identified at the molecular level. In plants, nucleoside uptake into the phloem tissue has been detected (Boldt and Zrenner, 2003). The phloem represents a tubular system, consisting of living cells, which allows long distance transport of organic products between plant organs. These products are either generated by photosynthesis, or liberated by degradation of various storage products. Cotyledons prepared from developing castor bean seedlings import nucleosides which have previously been liberated from mobilising endosperm tissue. In addition, intact petunia pollen cells are able to import exogenous nucleosides via both active import and facilitated diffusion (Kamboj and Jackson, 1985). Members of the CNT family have not been characterised in plants, but recently Wormit *et al,* (2004) described the uptake of nucleosides in *Arabidopsis thaliana* reporting the biochemical properties of three more uncharacterised ENT *Arabidopsis* transporters.

This project focuses on the function and specific characteristics of nucleoside transporters in *E. coli*. The following section is focused on nucleoside transporters and in particular what is known about them in *E coli*.

1.3.1 Nucleoside transport in bacteria

The cell envelope of enteric Gram-negative bacteria contains two distinct membrane structures, the outer membrane and the inner membrane. These two membranes are separated by a rigid porous structure, the peptidoglycan layer. The outer membrane is partially anchored to the peptidoglycan layer through a lipoprotein and contain porin proteins. Also the Tsx receptor in the outer membrane of *E. coli* has been shown to facilitate nucleoside diffusion and the enzymes responsible for nucleoside catabolism are located near the cytoplasmic membrane (Mygind and Munch-Petersen, 1975). The space between the inner and outer membranes is called the periplasmic space or murein layer, and the binding proteins of ABC porters involved in transport of nutrients are located there (Munch-Petersen and Mygind, 1983).

The inner or cytoplasmic membrane is the main selective permeability barrier for entry of nutrients and other substances required by the cell and for waste secretion. Also in this membrane there are energy systems e.g. the electron transport chain necessary for active transport (Munch-Petersen and Mygind, 1983). There is evidence that both membranes of *E. coli* are involved in nucleoside transport.

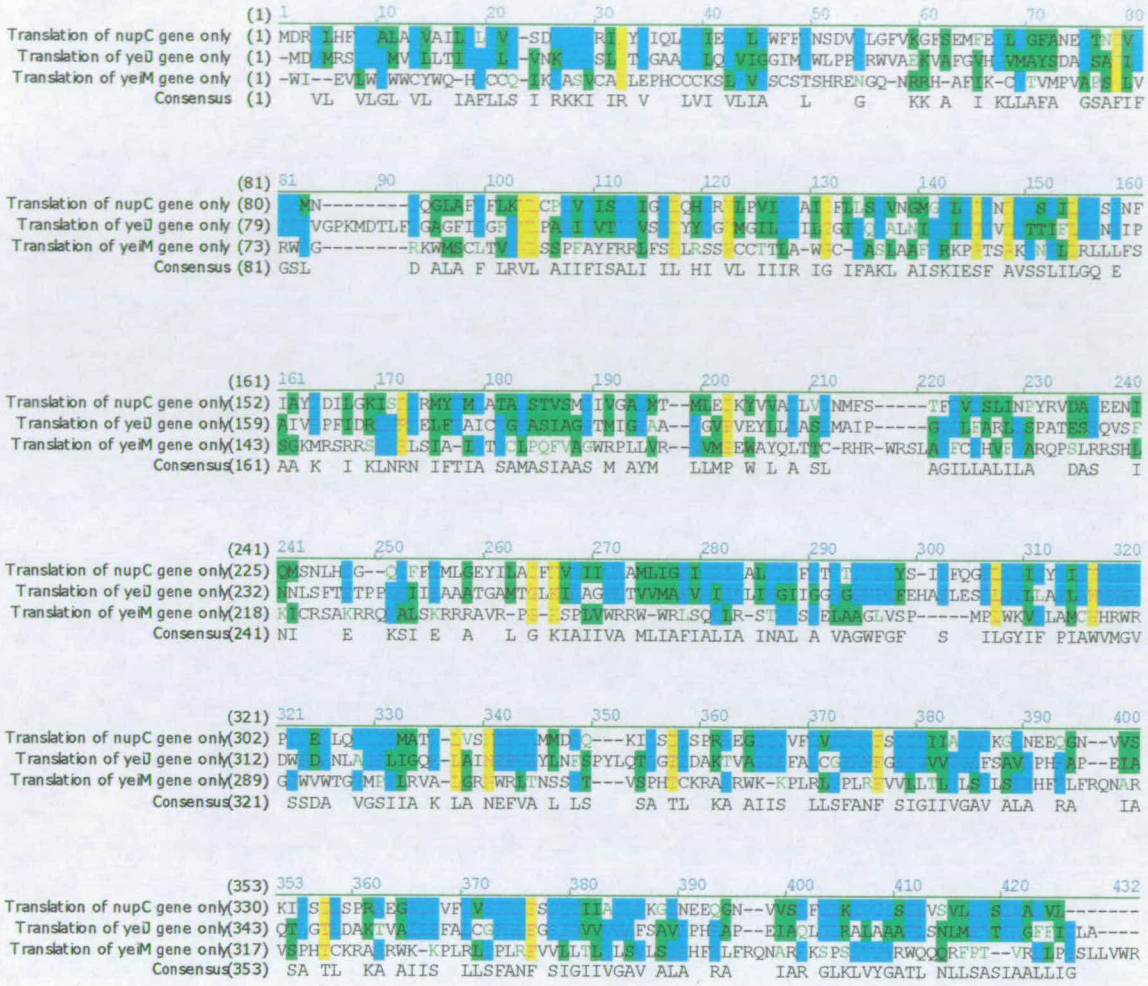
Nucleosides are transported across the inner membrane of *E. coli* by at least two separate energy-driven systems known as NupC and NupG (Patching *et al*, 2005). These systems are encoded by the genes *nupC* and *nupG* respectively. Studies have reported that NupC is unable to transport guanosine efficiently, but is able to transport other common nucleosides and their deoxyribonucleotide derivatives. NupG on the other hand is thought to be able to transport all common nucleosides and their deoxyribonucleoside derivatives and it is able to take up the antibiotic showdomycin (Craig *et al,* 1994).

Orthologues of NupC have been identified in a number of other bacteria including *Bacillus subtilis*. However nucleoside transport activity in B. *subtilis* is lower than that of *E. coli* (Schuch *et al,* 1999). Apparently *B. subtilis* actively transports nucleosides by a mechanism which has the characteristics of both NupC and NupG.

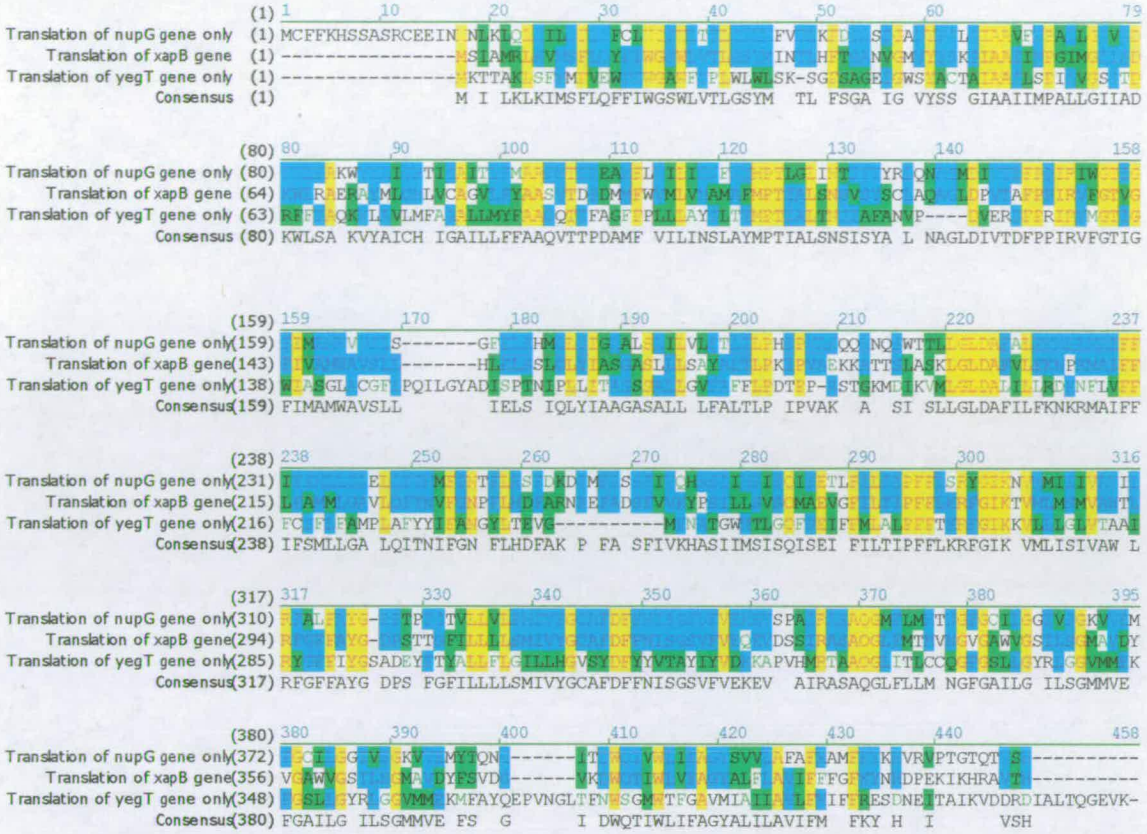
Also orthologues of NupC have been found in *Salmonella typhi* and *Shigella* (J. Craig, PhD thesis, 1993). *Escherichia*, *Salmonella* and *Shigella*, all inhabitants of the intestinal tract of mammals, which probably produces an environment rich in nucleic acid degradation products and this might be why these bacteria have evolved efficient and well regulated mechanisms for nucleoside transport and nucleoside metabolism (Joint Genome Institute, www.jgi.doe.gov)

In *E. coli* NupC and NupG are proton gradient-driven systems allowing the accumulation of nucleosides in the cell, even when the concentration in the external environment is very low. In these systems the flow of protons into the cell provide the necessary energy for the nucleosides to be taken up by the cell. This is a distinction from the concentrative mammalian nucleoside transport systems that are driven by sodium ion gradients (Baldwin *et al,* 1999). CNT1 (a mammalian transporter isolated from the jejunum) is a homologue of NupC. Interestingly, CNT1 is able to transport the nucleoside analogue AZT, used to combat HIV infection. CNT1 and NupC share substrate specificity and they have amino acid sequence identity of 27%, particularly in their C-terminal region, indicating that this probably contains the substrate binding region (Huang *et al,* 1994).

With the completion of the *E. coli* genome project paralogues of NupC and NupG have been found, designated as YeiM and YeiJ, and XapB and YegT respectively (Maurice P Gallagher-unpublished data). The alignment of the six proteins encoded *by nupC, nupG, xapB. yegT, yeiM* and *yeiJ* showing homology of the proteins is shown in figure 1.2. It is worth noting that XapB has been proposed to be involved in xanthosine catabolism (Seeger *et al,*1995). The genetic control of these transport systems is described in order to get a better understanding of how expression of these systems is regulated.


Translation of nupC gene only
Translation of yeiJ gene only
Translation of yeiM gene only
Consensus

A.



B.

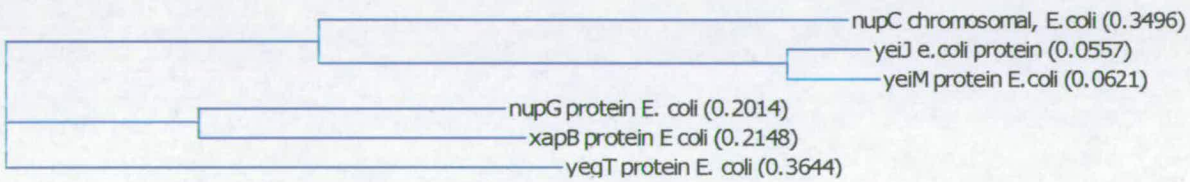


Figure 1.2 Alignment of the six proteins encoded *by nupC, nupG, xapB. yegT, yeiM* and *yeiJ* showing homology of the proteins. As seen from the phylogenetic tree, *nupC* is related in its aa sequence with *yeiJ* and *yeiM* whereas *xapB* and *yegT* are paralogues of *nupG.*

(KEY: yellow= complete identity, blue=conserved aa properties).

1.3.2 Regulation of nucleoside transport systems in *E. coli*

Nucleoside transport and catabolism are coupled in *E. coli* and they are controlled by two repressors, *cytR* and *deoR*. *nupG* is regulated by *cytR* as well as *deoR*, but *nupC* is only regulated by *cytR* (Munch-Petersen and Mygind, 1976). The transcription of genes needed for utilization of nucleosides in *E. coli* is regulated by the repressor protein, CytR, in concert with the cyclic AMP (cAMP) activated form of cAMP receptor protein (CRP).

The *deoR* gene may cause a rise in transport rates of cytidine and uridine (Munch-Petersen and Mygind, 1976). *cytR* has been shown to regulate nucleoside catabolising enzymes.

Experiments showed that *cytR deoR* double mutants show a five fold increase in the nucleoside transport rate, compared to that found in the wild type strains. Strains of *E. coli* which cannot metabolise cytidine and uridine retain their capacity for transporting these nucleosides, showing that nucleosides are transported by active transport systems and that catabolism is not essential for transport (Mygind and Munch-Petersen, 1975; Wohlhueter and Plagemann, 1980).

Experiments involving membrane vesicles have shown that the presence of an electron donor system was required for function of NupC and NupG. From these results it can be concluded that these systems are energised by the proton motive forces (Munch-Petersen and Mygind, 1976).

In 1980 a purine nucleoside phosphorylase (XapA), which could cleave xanthosine, was found. This xanthosine phosphorylase activity is only present when *E. coli* is grown in the presence of xanthosine. Gene *xapA*, which is proposed to be the gene controlling the xanthosine phosphorylase, and *xapB*, which encodes a membrane protein similar to nucleoside transporter NupG, are located adjacent to but clockwise of a gene identified as *xapR* in relation to the origin of replication (Jørgensen and

Dandanell, 1999). *xapR* encodes a positive regulator belonging to the LysR family and is needed for expression of genes *xapA* and *xapB* (Seeger *et al,* 1995, Dandanell *et al,* 2005). LysR is a large family of proteins that regulate genes with very diverse functions in prokaryotes. LysR family members have a number of common characteristics, such as a size of about 300 amino acids, a helix-turn-helix DNA-binding domain in the N terminal part, and the requirement for a small molecule to act as an inducer (Schell, 1993). The homology between LysR family proteins is generally high, with the highest similarity in the N-terminal part. The amino acid sequence of XapR is highly homologous (70%) to that of AlsR from *E. coli*. Of the 60 N-terminal amino acids in XapR and AlsR, 68% are identical. The work done on LysR family proteins has been hampered by the fact that most members of this family are very insoluble and have a tendency to form inclusion bodies. This is also the case for XapR, and it has not yet been possible to purify the protein for characterization and *in vitro* work such as footprinting (Schell, 1993, Jensen and Nygaard, 1975).

A number of *xapR* mutants containing single amino acid substitutions of XapR have now been sequenced and analysed. This work revealed two regions that are important for correct interactions between the inducer and XapR (Jørgensen and Dandanell, 1999). These regions were modelled into the three dimensional structure of the protein and they are most probably directly involved in binding the inducer xanthosine (Dandanell *et al,* 2005).

1.4 Nucleosides as agents for disease therapies

Nucleoside analogues are a pharmacologically diverse family which includes cytotoxic compounds, antiviral agents and immunosuppressive molecules. Modified nucleoside analogues have been used as therapeutic agents in the last few decades due to their ability to incorporate themselves in cells infected by viruses or parasites or cancer cells and to act as inhibitors of enzymes involved in viral replication. The variety of ways via which nucleoside analogues act include incorporation into

nucleic acids resulting in transcription and replication disruption or error, DNA-strand breakage and perturbation of intracellular nucleotide pools (Baldwin *et al,* 1999). This trait has made them invaluable in applications in the medical field.

Knowledge gained from the identification and molecular characterisation of human nucleoside transporters is an important step in future developments in the use of nucleoside drugs in selective treatment of cancer as well as viral and parasitic infections. Routinely used nucleoside analogues include the purine drugs, cladridine and fludarabine, and the pyrimidine drug, cytarabine, which are used in the treatment of haematological malignancies but lack activity against solid tumours. However anticancer drugs such as the pyrimidine drug gemcitabine, a deoxycytidine analogue, have a broad range of activities against solid tumours including lung, breast, bladder and ovarian cancers.

An analysis of the use of nucleoside analogues as antiviral, antiparasitic and chemotherapeutic agents and of their means of transport is given below. An overview of some principal drugs, which are part of purine or pyrimidine groups of nucleoside analogues, is also provided.

1.4.1 Mechanisms of action of nucleoside analogues

The process with which nucleoside analogues are used as therapeutic agents is a three step one, including uptake, metabolic activation (if a pro-drug is involved) and pharmacological action. The molecular basis of these steps is not always fully realised because the transport of a drug across the plasma membrane limits its bioavailability and action. Furthermore the carriers involved in the uptake of nucleoside analogues in mammalian cells have not been purified and have been cloned only recently. It should also be brought to attention that due to our lack of knowledge of the targets for nucleoside analogues, different levels of sensitivity or

resistance to a drug could be the result of varied bioavailability which stems from the different cell types used and their differentiation states.
Zidovudine (3'-azido-2', 3'-dideoxythymidine), also known as AZT, was the first nucleoside analogue approved for AIDS treatment. A variety of purine and pyrimidine analogues have been derived from AZT, including stavudine (d4T), zalcitabine (ddC) and lamivudine (3TC). These nucleosides are converted to their triphosphate form within the cell so as to be able to be incorporated into the growing viral DNA. As such they function as a toxic chain terminator through inhibition of retroviral reverse transcriptase.

Bioavailability of activated metabolites could be limited by inefficient nucleoside phosphorylation. For example AZT monophosphate is readily synthesised by the cell but it is then a poor substrate for the next enzyme in the chain, thus compromising the intracellular concentrations of AZT diphosphate and AZT triphosphate. From this example it can be assumed that in certain cases it is the metabolic activation rather than transport processes that are key steps in mediating the antiviral action of nucleoside analogues (Liu *et al*, 2005).

Figure 1.3 shows a general pathway for nucleoside analogue action. It shows characteristics including transport mediated by membrane transporters, activation by intracellular metabolic steps that retain the nucleoside residues in the cell and the formation of active phosphate derivatives. Nucleoside analogues are generally hydrophilic molecules and require the specialised nucleoside transporters mentioned above to enter the cell.

Each compound also has unique drug-target interactions that help explain their differences in activity against various diseases. For example the cytotoxic effects of the purine analogues, fludarabine and cladribine, on non-dividing cells may be explained by interaction with targets involving DNA repair rather than replication and direct or indirect effects on mitochondria.

The potential of nucleoside analogues to be incorporated into nucleic acids by the DNA repair machinery makes them interesting candidates for combination with DNA damaging agents. Once incorporated, nucleoside analogues are fairly resistant to repair excision and cause irreversible damage by interfering with nucleic acid synthesis, thus promoting anti-proliferative effects or resistance to virus replication in infected cells (Pastor-Anglada *et al,* 1998). Gemcitabine for example causes "masked chain termination" where an additional nucleotide is incorporated into the DNA chain before replication is interrupted. Inhibition of DNA repair by nucleoside analogues may also increase accumulation of DNA lesions induced by DNA-damaging agents and slow their removal, and therefore enhance the cytotoxic effects (Galmarini *et al,* 2002).

Therapy with combinations of different analogues has been proven to be successful in enhancing the cytotoxic effects of nucleoside drugs. Nucleoside analogue combinations such as fludarabine and cladribine have achieved responses in patients with refractory and hairy-cell leukaemias. Other examples include the combinational use of cytarabine with an antracyclin (Galmarini *et al,* 2002) or AZT, ddC and ddI, which are used either in monotherapy or in combination therapy for AIDS (Yatvin *et al,* 1999).

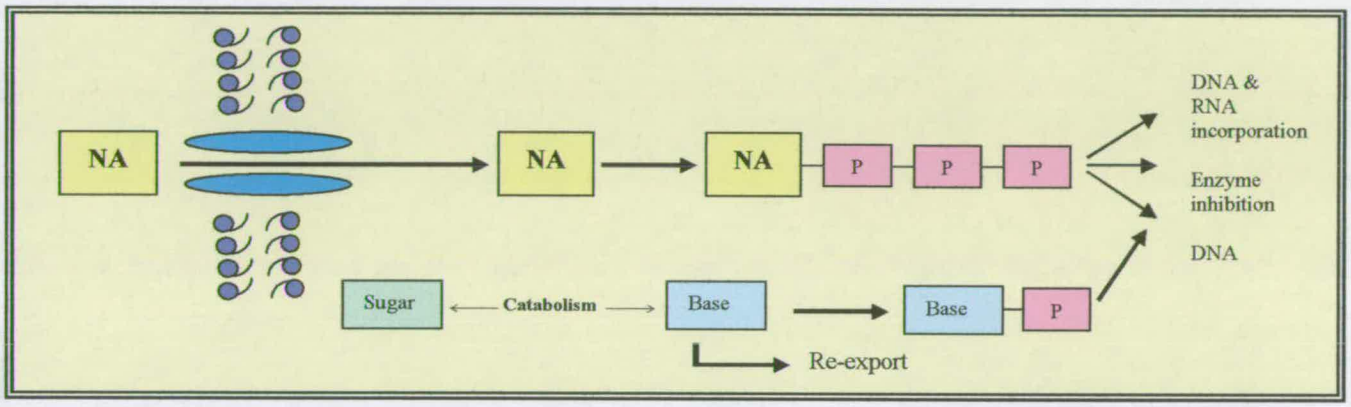


Figure 1.3: Entry and utilisation of nucleoside analogues in the cell (adapted from Galmarini *et al*, 2002).

Most nucleoside analogues (NA) are hydrophilic molecules and therefore need specialised transporter proteins to enter the cell. Once inside they are activated by intracellular metabolic steps to triphosphate derivatives (the phosphates are shown here as pink boxes "P"). Active derivatives can exert cytotoxic activity by being incorporated into and altering the DNA and RNA macromolecules or by interfering with various enzymes involved in synthesis of nucleic acids, such as DNA or RNA polymerases and ribonucleotide reductases. These actions result in inhibition of DNA synthesis and apoptotic cell death.

1.4.2 Nucleoside analogues and their use as chemotherapeutic, antiviral and antiparasitic agents

The growing importance of nucleoside analogues as cytotoxic agents was derived both from the development of new compounds with broad applicability to common cancers and from understanding their mechanisms of action, enabling pharmacological intervention to exploit the antitumour effects of these compounds (Galmarini *et al,* 2002). In terms of antiviral drug developments a wide range of drugs against HIV, hepatitis B and other diseases have been developed and are now used in combination (HIV tools.com). Most of them inhibit the action of the enzyme reverse transcriptase. However a number act on thymidine kinase. Also topical antiviral nucleoside analogues include acyclovir, a synthetic purine nucleoside analogue, which is used for treatment of cold sores against Herpes simplex virus-1 (HSV-1). An important goal in the development of antiparasitic drugs is the identification of unique biochemical targets in parasites that might improve the selectivity of therapy (Carter *et al,* 2001).

Most nucleoside analogue drugs used for cancer therapy are hydrophilic molecules and in the absence of transporter-mediated uptake their access to intracellular targets would be slow. For drugs such as cytarabine, cladribine, fludarabine and gemcitabine active transport is a key factor that determines their effectiveness and clinical efficacy. Knowledge of the abundance of different transporters in different tissues and their abilities to transport nucleoside analogues provides vital information for the selectivity and efficacy of individual drugs.

For example cytarabine, cladribine, and fludarabine are taken up into leukaemic cells by *es* transporters and the *in vitro* sensitivity of cells to the drugs has been correlated with the cell surface abundance of these transporters. To make such correlations *es* transporters have been quantified using tight-binding *es*-specific transport inhibitors as reporter molecules. Scientists now want to specifically detect other transporters through assays using antibiotics. Such assays would be valuable for identifying patients for whom therapy with a specific nucleoside analogue would not be

effective. Failure of response might reflect the low abundance of a nucleoside transporter species capable of drug uptake but also a decrease in drug abundance due to low levels of drug metabolism in specific tissues (Baldwin *et al,* 1999).

To increase cytotoxicity of anticancer nucleoside analogues there are several common approaches. Currently five routes have been identified and are discussed in a review by Galmarini *et al,* (2002). These routes include a high dose of the administered drug, the degradation of inhibitors, metabolism modulation, DNA repair inhibition, and the construction of degradation- insensitive compounds.

Nucleoside analogues are widely used in viral infections. One of the best types investigated are analogues used against the HIV virus. Anti-HIV drugs can be classified into three categories. There are nucleoside analogue- based reverse transcriptase inhibitors, non-nucleoside reverse transcriptase inhibitors and protease inhibitors.

The nucleoside analogue drugs used as anti-HIV agents are dideoxynucleoside analogues of endogenous nucleosides that have to undergo *in vivo* phosphorylation to form active triphosphates, in order to function as inhibitors of reverse transcriptase (Chan *et al,*, 1999). Agents in this category are therefore prodrugs which are metabolised in the following sequence. First the nucleoside analogue is phosphorylated by a nucleoside thymidine kinase to form a nucleoside monophosphate. Then follows a second phosphorylation, by nucleoside monophosphate kinase, to form a nucleoside diphosphate and finally, a third phosphorylation occurs, by a diphosphate kinase, to form a nucleoside triphosphate analogue (Chan *et al,* 1999).

In general the initial phosphorylation reaction is quite inefficient except for the case of AZT. However due to high turnover number, poor substrates may still be metabolised adequately so that the final active triphosphate product may reach therapeutic levels. Most data gathered for nucleoside analogues are focused on AZT, due to its high efficiency in the first phosphorylation phase. The active form of AZT

is the AZT-TP (triphosphate), which inhibits HIV reverse transcriptase via a competitive mechanism. It also inactivates target enzymes via azido agents forming covalent modifications (Chan *et al,*, 1999).

In AIDS therapy the clinical efficacy of AZT is limited by its toxic effects, which are mainly directed to the bone marrow and skeletal cardiac muscle. As a result of long-term treatment of HIV patients with AZT, anaemia, leukopenia, myalgia, muscle weakness and elevated creatine kinase levels in the blood serum, occur (Barile *et al,* 1998).

To prevent AZT side effects, clarification of the processes leading to AZT cytotoxicity is needed and because of that, different hypotheses have been suggested to account for the biochemical mechanisms responsible for the cytotoxic effects induced by AZT in human host cells. These hypotheses include AZT incorporation into newly synthesised host cell DNA (Sommadossi *et al,* 1989) as well as the possibility that mitochondria are cellular targets of AZT (Barile *et al,* 1998).

The fundamental biochemical divergence between parasites and their mammalian host cells offers a promising strategy for therapeutic utilisation of parasitic disease. One basic difference between mammalian cells and protozoan parasites is the inability of pathogenic parasites to synthesise purines *de novo*. Instead each genus of parasites has a unique distinctive network of purine transporters and salvage enzymes that enable them to scavenge preformed purines from the host. The salvage of host purines is initiated by their translocation, either across the parasite plasma membrane or in the case of intracellular parasites, across the host plasma membrane (Carter *et al,* 2001).

The salvaged purines are in the form of either nucleobases or nucleosides because phosphorylated nucleosides are unlikely to cross the lipophilic membranes. The transport of nucleosides is mediated in the case of parasites by nucleoside transport proteins. Biochemical studies have revealed heterogenicity in the nucleoside

transporters with respect to number of paralogues within a genus, substrate specificities, ligand affinities and sensitivity to inhibitors (Carter *et al,* 2001).

For example *Leishmania donovani* produces two high affinity nucleoside transporters of non-overlapping substrate specificity. *Trypanosoma brucei* uses two types of adenosine transporter, one that transports purine nucleosides and one that recognises adenine and several antitrypanosomal drugs. *Plasmodium falciparum* has an adenosine transporter that recognises the L-isomer of adenosine. These variations among parasite nucleoside transporters reflect the unique host environments which these parasites inhabit.

However sophisticated biochemical studies on nucleoside transport in intact intracellular parasites are often compromised due to the existence of multiple permeation mechanisms that operate simultaneously on the same ligand, by the presence of mammalian membrane nucleoside transport systems, or by nucleoside metabolising enzymes of either parasitic or mammalian origin. Overcoming these problems has been slowed down due to the lack of identification of nucleoside transporter genes and suitable heterologous expression systems to characterise such nucleoside transporters (Carter *et al,* 2001).

Several currently used antiparasitic drugs are being used to target the salvage pathway of nucleoside synthesis. For example, studies of the kinetics of transport in the bloodstream of *Trypanosoma brucei brucei* have revealed the presence of two high-affinity transporters. One of these, P1, transports adenosine and inosine while the other one, P2, transports not only adenosine and adenine but also melarsoprol, the front line arsenical drug used in treatment of late-stage sleeping sickness. Also P2 transports pentamidine, which is one of the most common drugs used in the early stages of the disease (Carter and Fairlamb, 1993).

Similarly nucleoside transporters are responsible for the uptake of drugs which are toxic to *Leishmania* and *Trypanosoma* but not to mammalian cells. Resistance of *T. b. brucei* to such drugs has been linked to the loss or alteration of the P2 transporter.

Generally speaking use of nucleoside analogues can become more effective by expanding knowledge on their way of entry into the cell so as to prevent uninfected cells from taking up toxic compounds. Nucleoside analogues can be categorised on whether they are purine or pyrimidine based. This characteristic usually accounts for their means of transport into the cell and the target nucleoside transporter that they utilise in their cellular uptake strategy.

1.4.3 Purine nucleoside analogues

There are two main groups of purine nucleoside analogues used in cancer therapy. The first group includes nucleoside analogues called thiopurines and the second deoxyadenosine derivatives.

Thiopurines play an important role in the management of acute leukaemias. In childhood acute leukaemia they are used as part of standard consolidation and maintenance schedules. In acute myeloid leukaemia, thioguanine is used as part of different schemes for remission therapy.

The cytotoxicity of thiopurines is thought to depend mainly on incorporation of their phosphorylated derivatives into DNA, thereby interfering with the function of DNA polymerases, ligases and endonucleases. Thiopurines may also cause toxic effects by inhibiting other enzymes which are involved in *de novo* purine synthesis (Pettitt, 2003).

Two deoxyadenosine derivatives are currently used alone or in combination, for the treatment of specific malignant disorders of the blood. These derivatives are fludarabine, used for chronic lymphocytic leukaemia, and cladribine, for hairy-cell leukaemias. These drugs share activity against other lymphoid malignant disorders, including non-Hodgkins lymphomas, but lack activity against multiple myelomas and solid tumours (Galmarini *et al,* 2002).

Fludarabine, unlike cladribine, is administered as a soluble form. Both nucleosides are taken up by the target cells through nucleoside membrane transporters and subsequently, are activated into their triphosphate forms. Once fludarabine or cladribine are incorporated into the DNA, chain elongation mediated by DNA polymerase is terminated, inducing apoptosis of the target cells. Both compounds also indirectly impair DNA replication by inhibiting ribonucleotide reductase which normally cleaves nucleotides to form nucleosides, therefore reducing the deoxynucleotide triphosphates available for viral DNA synthesis.

Fludarabine and cladribine also alter gene transcription, resulting in depletion of proteins required for cell survival. In the case of fludarabine, incorporation of its monophosphate form into RNA results in premature termination of the RNA transcript impairing its function as a template for protein synthesis. On the other hand fludarabine triphosphate inhibits RNA synthesis by suppressing the activity of RNA polymerase II. Cladribine also has a role in direct alterations in mitochondrial function that may trigger apoptosis.

The toxicity profiles of both drugs are similar and include moderate myelosuppression and prolonged immunosuppression. One result of this immunosuppression is an increase in opportunistic infections and potentially, increased risk of secondary cancers. Severe neurotoxicity occurs in higher doses.

1.4.4 Pyrimidine nucleoside analogues

A diverse group of drugs constitute the pyrimidine nucleoside analogues. In cancer treatment, these are considered in two groups, the deoxycytidine derivatives and fluoropyrimidine nucleosides.

Cytarabine belongs to the group of deoxycytidine derivatives and it is an analogue commonly used in the treatment of malignant haematological diseases but is without activity in solid tumours. This drug is one of the most active single agents in the treatment of acute myeloid leukaemia. When cytarabine is administered in combination with an anthracycline, the remission rate can be as high as 65-70% in previously untreated patients and 30-50% in patients with relapsed acute myeloid leukaemia (Rosenblum *et al,* 2001).

Intracellular penetration of cytarabine depends on the plasma concentrations. Once inside the cell, cytarabine is broken down to the non-toxic metabolite uracil arabinoside. Cytarabine cytotoxicity is caused by direct inhibition of DNA polymerases and incorporation of its metabolites into DNA which leads to chain termination and DNA synthesis arrest.

The main side effects of conventional doses of cytarabine are leukopenia, nausea, vomiting, diarrhoea, mucositis and hair loss. High dose of cytarabine is associated with other side effects as well, such as pericarditis and neurotoxicity.

The cytotoxic activity of cytarabine is limited by characteristics such as metabolic deamination, low affinity for deoxycitidine kinase and rapid elimination of the triphosphate derivative. In order to deal with this problem further deoxycytidine analogues were designed. Unlike cytarabine, gemcitabine has shown consistent activity in previously treated patients. However gemcitabine is inactive on solid tumours.

Gemcitabine is more lipophilic than cytarabine and is a better substrate for membrane nucleoside transporters. This results in more efficient accumulation and

longer retention of gemcitabine triphosphate in tumour cells and enhances its cytotoxicity (Plunkett *et al,* 1995).

After gemcitabine is deaminated by deoxycytidine deaminase, it is incorporated into nucleic acids, causing chain termination. Also gemcitabine inhibits DNA synthesis indirectly through inhibition of ribonucleotide reductase, blocking the *de novo* DNA synthesis pathway (Plunkett *et al,* 1995).

In the group of fluoropyrimidines, the nucleobase fluorouracil and the nucleoside floxuridine have activity in patients with colorectal, pancreatic, breast, head and neck cancers.

Briefly fluorouracil is incorporated into nucleic acids, mainly into RNA, and affects several processes including transcription, intracellular distribution and translation. Fluorouracil is the main treatment for many common malignant diseases and when given for long durations, it has greater antitumour activity. New orally administrated fluoropyrimidine nucleosides achieve higher concentrations of fluorouracil in tumours than in normal tissues so they are more effective and less toxic (Parker and Cheng, 1990).

The main problems with nucleoside analogues used as drugs are their bioavailability and their serious side effects. One of the approaches to increase the concentrations of active derivatives of nucleoside analogues is by modulating the enzymes that facilitate their activation and inactivation. Also a better understanding of the processes of nucleoside analogue transport may uncover more therapeutic possibilities and eliminate some of the side effects (Galmarini *et al,* 2002).

The different kinetic properties of nucleoside transporters might well provide a key way of designing chemotherapeutic schedules "tailored" to the specific transporters present in the neoplastic cells. Protection of normal cells from the toxic effects of nucleoside analogues might also be enhanced by the co-administration of inhibitors that block ENT but not CNT mediated transport.

In future it should be possible to detect transporters or toxicity levels of the nucleoside analogues to be used, though assays. Such assays should prove useful in identifying patients for whom therapy with a particular nucleoside analogue will fail. There is an increased need to develop not only new molecules which are candidates as therapeutic agents but also, detection systems that determine which drug is active and hopefully which transporter controls activity.

1.5 Bioassays and Biosensors

The need for new and accurate methods in order to monitor toxic components and measure toxicity caused by chemicals has led scientists to develop new screening systems. In order to identify toxic components, methods combining bioassays with chemical analysis of target compounds are used routinely.

The biological responses induced in living organisms by chemical substances are diverse and depend on the biological system's sensitivity to the particular toxin. Bioassays provide a measure of the toxic impact through a variety of readouts such as pH indicators, solubility, light emission, and bioavailability (Namiesnik *et al,* 2005). Some commonly used bioassays are supplemented by microbial bioassays and biosensors.

The processes of classical bioassays are too laborious to be applicable in a wider way or for high throughput analysis. Commonly used microbial testing methods are plating techniques such as agar well diffusion assay and agar spot test. During these methods the indicator strain is inoculated in appropriate agar and the culture supernatant to be studied is placed into agar-cut wells (agar well diffusion assay) or a small inoculum of bacteria is spotted on the agar (agar spot test) (Vesterlund *et al,* 2004). The inhibitory activity of tested substances on the bacterial strains that are used is then read after a certain incubation period. The inhibitory activity is shown by an inhibitory zone around the sample. Although these methods are simple and inexpensive, several factors can influence the results indirectly, making the assays less sensitive or reliable. Such factors are the pH of the medium, the concentration of substances examined and the strain used (Vesterlund *et al,* 2004).

To improve sensitivity in testing, more sensitive and rapid methods have been developed which are based on flow cytometry, measurement of ATP and bioluminescence. These new tests can assess toxicity rapidly by their effects on living cells or organelles directly (Namiesnik *et al,* 2005).

In terms of bacterial-based bioassays a range of methods have been developed for toxicity screening. Measurement of the effects of toxic action on microbial metabolic activity, represent a direct, rapid, sensitive and cost-effective approach. Furthermore this approach is useful for screening large numbers of samples for toxicity. Several bioassays using microorganisms have been standardised and are classified according to the parameter that is measured. For example there are tests based on growth inhibition, such as the *Pseudomonas* cell multiplication inhibition test or the activated sludge test (Leonard *et al,* 2003).

Biosensors are devices that detect, record and transmit information regarding a physiological or biochemical change. Technically they are probes that utilise a biological component, such as a whole cell, an enzyme or an antibody, and transmit information on changes in the physiology of the cell or system, or the presence of various chemical or biological materials in the environment. Typically biosensors are used to monitor changes in environmental conditions. For example they can detect and measure accurately concentrations of bacteria, hazardous chemicals, or acidity levels (D'Souza, 2001a).

Genetically engineered bacteria are commonly used as biosensors because of their ability to respond in specific ways to changes in their environment. Such bacteria commonly give off or can be modified to produce an easily detected signal, such as light. For example there are bacteria that glow when they consume specific pollutants giving a direct indication of the pollutant's presence and its concentration. Biosensors are now used in a variety of sectors and businesses in environmental, medical, military and other processes (Alocilja and Radke, 2003).

1.5.1 Reporter genes in biological sciences

The term reporter gene is used to define a gene which has a readily measurable phenotype that can be distinguished easily. There are a variety of reporter genes available to scientists and they are selected on the basis of their sensitivity, convenience, and reliability for the assay that they are going to be used in. The reporter gene's activity is responsive directly or indirectly to alterations in the normal gene pattern of expression in host cells. For example a number of hormones and growth factors have been shown to stimulate target cells by activating second messenger pathways that in turn regulate the phosphorylation of specific nuclear factors thereby altering gene transcription. This overall event may be measured by placing a reporter gene under the regulation of the nuclear factor.

Table 1.3 summarises some of the advantages and disadvantages of the most commonly used reporter genes. In general, reporter genes have the advantage of low background activity in cells but amplify the signal from the cell to produce a highly sensitive, easily detectable, response. The choice of reporter used, however, depends on the cell line being studied, the nature of the experiment (for example gene expression versus transfection efficiency), and the adaptability of the assay to the question being addressed.

Chloramphenicol acetyltransferase (CAT) was the first reporter gene used to monitor transcriptional activity in cells. CAT is a bacterial enzyme that can detoxify chloramphenicol, an inhibitor of prokaryotic protein synthesis, by catalysing the transfer of acetyl groups from acetyl CoA to the 3-hydroxyl position of chloramphenicol. The enzyme is stable and there is no endogenous expression in mammalian cells. However, the assay relies on radiochemicals, and although an automated ELISA became available later, making the assay more manageable, the linear range and sensitivity of the assay are not as broad as for other reporters (Naylor 1999).

β-Galactosidase is a well-characterised bacterial enzyme and has been one of the most widely used reporter genes in molecular biology. β-galactosidase is a glycosidase that catalyzes the hydrolysis of β-galactoside sugars such as lactose. β-galactosidases have been found in numerous microorganisms, animals and plants. The importance of lactose fermentation tests as a diagnostic tool for the Enterobactericeae sp. led to extensive investigation of β-galactosidases in other Gram-negative bacteria (Sahin-Tóth *et al*, 2000).

Amongst Gram-negative bacteria, β-galactosidases have been reported in strains of the Enterobacteriaceae, Pseudomonodaceae, Parvobacteriaceae, and Neisseriaceae (Corbett & Catlin, 1968). β-galactosidase encoded by *lac Z* of *E. coli* has been the most studied. This enzyme is tetrameric, with a subunit size of 116,000 daltons and a total molecular weight of 465,000 daltons. The substrate 5-bromo-4-chloro-3-indoyl β-galactoside (X-Gal) is often used for *in vivo* analyses. Enzymatic hydrolysis of X-Gal followed by oxidation produces a blue precipitate, causing colonies expressing the enzyme to turn blue and thus to be easily identified.

Over the last decade simple colorimetric assays of poor sensitivity and narrow dynamic range have largely been replaced with more sensitive and adaptable bio- or chemi-luminescent assays which can be automated and lend themselves to high throughput screening (Purohit and Rocke, 2003). Aequorin and green fluorescent protein (GFP) are proteins that have been isolated from the bioluminescent jellyfish *Aequorea victoria*. In the presence of $Ca^{2+}$and the cofactor coelenterazine, aequorin emits blue light, with a peak at 469 nm, which can be detected using conventional luminometry (Leveau and Lindow, 2002). GFP, however, is unique among light-emitting proteins in that it is autofluorescent and, therefore, does not require the presence of any cofactors or substrates for the generation of its green light. It is detectable at 510 nm. Therefore GFP provides a convenient way by which intracellular events can be studied in intact living cells. The gene for GFP was cloned in 1999, and several GFP mutants are now available, which exhibit improved fluorescence properties, altered stability features, or different emissions spectra compared to wild-type GFP (Leveau and Lindow, 2002). The greatest advantage of

GFP is that in the absence of cell lysis, non-invasive monitoring of gene expression in living tissues is possible.

Many of the commonly used reporter genes have been assayed in combination. For example, secretable alkaline phosphatase (SEAP) has been used, in conjunction with luciferase and β-galactosidase to normalise transfection efficiency, and β-galactosidase, and Renilla luciferase have been combined with firefly luciferase for multiple readouts from a single well of cells (Naylor 1999).

The reporter gene chosen for bioassay development will largely depend upon the suitability of the test system for a particular investigation. Factors such as sensitivity, reliability, ease of use, and reporter dynamics have a central role in the selection. The stability of the reporter gene will determine whether it is appropriate for monitoring transcriptional kinetic of gene expression, whether high throughput screening is appropriate or, if all-or-none type expression in gene transfer experiments are preferable. For example, firefly and Renilla luciferases have relatively short half-lives (3 and 5–6 h, respectively) compared with CAT (50 h) in mammalian cells, as a result of their sensitivity to proteolysis, making these proteins more suitable for kinetic or transfection efficiency experiments, respectively.

Likewise, wild-type GFP, which has a much narrower linear range of fluorescence and greater stability than luciferase, is less suitable for the quantitative measurement of gene transcription. The lack of an enzymatic amplification process with GFP tends to make it less sensitive than luciferase and can limit its use only for highly expressed genes. Moreover, the stable accumulation of GFP in cells makes it less suitable for high throughput screening assays.

With reporter proteins and detection systems being constantly improved, luminescent and fluorescent assays are becoming more prevalent because of the ability to visualise reporter gene activity inside cells. Reporters such as GFP and luciferase, which provide a highly sensitive but non destructive way of monitoring gene transfer and expression, are becoming increasingly popular.

Several forms of firefly luciferase, in which the protease- sensitive regions of the protein have been mutated, yield larger and more stable signals compared with the wild-type protein. Similarly, modifications in GFP Reporter Gene Technology have produced mutant proteins with increased stability, altered spectral properties and signal intensity that are useful for monitoring the expression and localisation of multiple intracellular proteins.

Reporter Gene	Advantages	Disadvantages
Chloramphenicol acetyltransferase (CAT) (bacterial)	No endogenous activity, Automated ELISA available.	Narrow linear range, use of radioisotopes, stable.
β- Galactosidase (bacterial)	Well characterised; stable; simple colometric readouts; sensitive bio- chemiluminescent assays available.	Endogenous activity (mammalian cells).
Luciferase (firefly)	High specific activity; no endogenous activity; broad dynamic range; convenient assays.	Requires substrate (luciferin) and presence of O_2 and ATP.
Luciferase (bacterial)	Good for measuring/ analysing prokaryotic gene transcription.	Less sensitive than firefly; not suitable for mammalian cells.
Alkaline phosphatase (human placental)	Secreted protein, inexpensive colorimetric and highly sensitive luminescent assays available	Endogenous activity in some cells; interference with compounds being screened
Green fluorescent protein (GFP) (jellyfish)	Autofluorescent (no substrate needed); no endogenous activity; mutants with altered spectral qualities available.	Requires post-translational modification; low sensitivity (no signal amplification).

Table 1.3: (adapted from Naylor 1999). Summary of some commonly used reporter genes, stating their main positive and negative points. For the purpose of the experiments described here the *lux* genes encoding luciferase were used.

1.5.2 Use of the *lux* gene encoding luciferase as a reporter system

Luciferase refers to a family of enzymes that catalyse the oxidation of various substrates (e.g. luciferin), resulting in light emission. The most commonly used luciferases for reporter gene assays are the bacterial luciferases, the firefly (*Photinus pyralis*) luciferase, and more recently, the Renilla luciferase (from the bioluminescent sea pansy, *Renilla reniformis*).

Luminescence tests have the advantage of being rapid, sensitive and reproducible. Methods combining toxicity assays and chemical analysis are now commonly used to detect pollutants in water. Commonly toxicity tests are based on the inhibition of bioluminescence of luminescent bacteria *Vibrio fischeri* (former *Photobacterium phosphoreum*) and *Photobacterium phosphoreum*. These marine bacteria naturally emit light because of the enzyme luciferase which catalyses the following reaction.

$$FMNH_2 + O_2 + RCOH \rightarrow FMN + R\text{-}COOH + H_2O + \underline{LIGHT}$$

Light emission is directly proportional to the metabolic state of the cell, thus any inhibition of cellular activity is reflected in a decrease of bioluminescence. This light emitting reaction involves an intracellular luciferase-catalysed oxidation of the reduced form of flavin mononucleotide ($FMNH_2$) and reduction of a long chain aliphatic aldehyde (RCHO), by molecular oxygen.

The genes of the *lux* operon from three bacterial genera *Photobacterium, Vibrio*, and *Xenorrhabdus*, have been cloned and sequenced (Stewart and Williams, 1992). In these genera the *lux* operon consists of the *luxA* and *luxB* genes, which encode the luciferase α and β subunits, as well as genes that code for a reductase (*luxC*), a transferase (*luxD*), and a synthetase (*luxE*). The latter three genes are required for the re-conversion of fatty acids into the long chain aldehyde, required for bioluminescence (Wall *et al*, 1986). The order of the genes that code for luciferase (*luxAB*) and the fatty acid reductase (*luxCDE*) is the same in the presently

characterised operons, with *luxC* and *luxD* upstream of the luciferase genes and *luxE* downstream of the luciferase genes. Although the luciferase and fatty acid reductase genes are the only *lux* genes common to all luminescent bacteria, other *lux* genes (*luxF, luxG, luxH, luxI, luxR*), as well as unlinked *lux* gene loci have been identified in luminescent strains.

Regulatory genes controlling the expression of bacterial luminescence from *V.fischeri* and *V. harveyi* have been cloned and sequenced, and induction of luminescence has been characterized for the *V.fischeri lux* system (Meighen, 1993). Two regulatory genes, designated as *luxI* and *luxR*, are located upstream of the *luxCDABEG* genes.

The bacterial luciferases are generally heat- labile, dimeric proteins. A sensitive reporter system capable of monitoring transcriptional activity within living bacterial cells has been developed using *luxCDABE*, encoding bacterial luciferase and the genes for aldehyde recycling. Using this self controlled sysem the aldehyde substrate required for the assay is produced constitutively inside the cell when the operon is expressed.

The firefly luciferase has been one of the most popular reporter genes used in mammalian cells because of its high sensitivity. Deficiencies in the original assay arose from the requirement for cell lysis because of lack of permeability of the substrate luciferin and also production of a sudden "flash" response. However, the development of more membrane-permeable and photolysable firefly luciferin esters has removed the need for cell disruption (Yang and Thomason, 1993). Secondly, the development of "glow" reagents from industry has increased the duration and stability of the flash response such that it can be detected in a scintillation counter, also making it suitable for high throughput screening. Renilla luciferase may be a particularly appropriate reporter for intact living cell systems because this luciferase catalyses the oxidation of coelenterazine, which is membrane permeable. Like other luciferases, there is no endogenous activity in mammalian cells.

1.6 Gene replacements: an overview of the theory behind the methodology used

In recent years a vast amount of information on DNA sequences has led scientists to develop molecular techniques for dissecting the functions encoded by particular sequences. Small genomes can now be sequenced within days, although subsequent analysis in terms of gene functions of these genomes can be a long process. High throughput techniques such as microarrays for transcriptional analysis, and methods developed for proteomics allow the collection of enormous amounts of information. Although such techniques are now available, nevertheless large amounts of sequenced genomes belong to genes of putative or unknown function.

This is a fact that also applies to *E. coli* which is considered one of the most used and analysed bacteria in science. Since the *E. coli* genome annotation in 1997, only 4% of the genes have had definite functions assigned on the basis of experiment or very close homology to genes of verified function. Furthermore 1, 250 open reading frames (ORFs) have been assigned possible functions which still need confirmation and 850 ORFs still have entirely unknown functions (Serres *et al,* 2001).

One approach to functional analysis of these genes involves simple and efficient gene manipulation via targeted alterations of sequences of interest in their chromosomal location. A systematic mutational analysis of genes on the chromosome allows investigators to delete or replace the coding sequence of these genes and then look for subsequent phenotypic alterations. A number of approaches have been currently used in order to pinpoint the most efficient and precise method for this.

The more frequently experiments are performed using these tools, the more information is gathered on genes of unknown or hypothetical function. Furthermore the use and continuous testing of appropriate molecular techniques leads to their modification and improvements. Therefore the field of inactivation of chromosomal genes is a very dynamic and challenging area in which knowledge and novel approaches are added constantly.

The approaches and techniques that have been used in the past and present are recapitulated in the following sections. In addition, an analytical description of the method used in the work described in this thesis is given, highlighting the main points of the technique. More comments on the results acquired from this work as well as the efficiency of this method are given in chapter 3.

1.6.1 Gene replacements: the past and present

Gene replacements have generally relied on the use of specific bacterial strains with particular genetic background, such as *polA, recD sup*$^{+}$, or F'. An *in vitro* altered sequence would be used to replace the wild-type sequence in the mutant background. The altered chromosomal region was then transduced into a wild type genetic background. Unfortunately these methods often require the transduction of a marker along with the mutant allele. This marker can obscure the phenotype of the mutant allele because it may itself cause a phenotypic alteration (Link *et al,* 1997, Reyrat *et al,* 1998).

Insertional, frameshift, nonsense or antisense disruptions of ORFs within an operon can affect upstream and downstream gene expression as well as the gene that is targeted for inactivation. These polar effects could confuse detection of the mutant phenotype that is expected from a specific gene disruption.

Hamilton *et al,* (1989) described a method for gene replacements in wild type *E. coli* that uses homologous recombination between the bacterial chromosome and a plasmid, carrying cloned chromosomal sequences, whose replication ability is temperature sensitive. Cells maintain drug resistance only if the plasmid integrates into the chromosome by homologous recombination between the cloned fragment and the bacterial chromosome. Excision of the integrated plasmid is then allowed at the permissive temperature. Depending on the position of the recombination event that excises the plasmid, the chromosome retains either the wild type sequence or

inherits the altered sequence from the plasmid. This method can be applied to wild type strains but does not have a selection process for the loss of the excised plasmid. Lack of a selection for plasmid excision was approached by utilisation of the *Bacillus subtilis* gene *sacB* which encodes levansucrase, an enzyme that catalyses the hydrolysis of sucrose and levan elongation. When expressed in *E. coli* growing in media supplemented with sucrose, the gene *sacB* is lethal. In particular Blomfeld *et al,* (1991) developed a system for using a temperature-sensitive plasmid and the counter selectable *sacB* marker in the chromosome to facilitate allelic exchange. Later, Link *et al,* (1997) reduced this system to one component by incorporation of the *sacB* gene into a gene replacement plasmid. The limitation of this method is the strain-medium requirements and the temperature dependence.

Generally gene replacements can be achieved in a variety of ways. One of the easiest procedures for gene manipulation in *E. coli* is to place a drug resistance marker near, within or in place of a cloned gene of interest. Linear DNA can be used instead of plasmids, which contain a mutated gene flanked by homologous regions of the chromosome. This linear DNA is then transformed into recombination-proficient strains which are usually genotypically *recBC*, or *recD* negative. Recombination between the bacterial chromosome and both ends of the linear DNA fragment results in gene replacement. Then the recombinants can be selected by the presence of the drug resistance marker. However this type of replacement is restricted to RecBCD nuclease-deficient and recombination-proficient genetic backgrounds and it is not applicable in all cases. Additionally this technique has the disadvantage of limiting the number of mutations to the availability of useable antibiotic resistance markers.

A different method generally used, involves integration of a plasmid containing a selectable marker and the mutant gene of interest into the bacterial chromosome by homologous recombination, followed by resolution of the cointegrate. Since plasmid sequences are removed during resolution of the cointegrate, gene replacement can be made free from antibiotic resistance markers. The key to this procedure involves the use of vectors that cannot replicate under conditions used for selection of the cointegrate. This method has some disadvantages since the replacement is actually a

composite gene generated at the cointegrate resolution event. Furthermore cloning of the gene is required before replacement, and resolution of the cointegrate plasmid occurs at low frequencies not giving a desired replacement (Murphy, 1998).

Murphy (1998) identified a new method by which gene replacements can be generated in nearly any *E. coli* strain at high frequencies. This method uses linear DNA, does not rely on a cointegrate, and does nor require cloning of the gene. More specifically, the bacteriophage λ recombination Red functions, *exo, bet* and *gam*, are expressed from a multicopy plasmid and they are used to promote gene replacements in wild type *E. coli* cells after transformation with linear DNA substrates. Homologous recombination is elevated by the addition of the λ functions rather than induced by alterations to the host functions. Murphy also reported the integration of the P_{lac}-red operon, in place of *recBCD*, from a plasmid into the chromosome via P1 transduction. Chromosome encoded λ Red promotes even higher rates of recombination as compared to plasmid encoded λ Red (Murphy, 1998).

Replacement via linear DNA transformations requires a double cross-over event and variations of the method depend on the host cells, require long PCR primers and introduce a marker along with the mutant gene into the genome. This marker may have polar effects or limit multiple manipulations of the genome. To combat this problem a study by Pósfai *et al,* (1999) presented a method in which the mutant gene is carried from a circular plasmid into a homologous locus in the chromosome resulting in a direct duplication.

An I-*Sce*I site is included within the plasmid and resolution of the cointegrate occurs via intramolecular recombination. This is controlled by introducing a unique double-stranded break into the chromosome by meganuclease I-*Sce*I. This enzyme recognises an 18-bp sequence and generates a double stranded break with a 4 base 3' hydroxyl overhang. However the source of the meganuclease is provided through plasmids which give rise to different doses of the I-*Sce*I gene which leads to differences in the number of surviving cells and therefore in the rate with which these cells resolve the co integrate.

In 2000, two separate groups described precise replacement of genes in *E. coli* (Yu *et al,* 2000, Datsenko and Wanner, 2000). The use of the λ Red system was reported again. However this time short homology extensions were exploited. Furthermore with the use of an additional tool of site specific recombination the approach provided the option to make multiple replacements using the same antibiotic marker as well as completely delete the ORF from the chromosome. This method uses a recombination proficient strain and small sequences of homology. The principles of the method as described by the two groups are outlined below as it was used in this project.

1.6.2 Gene replacement and bacteriophage λ recombination systems

As many as 15 to 20 proteins are thought to be directly involved in homologous recombination of linear sequences in *E. coli*. In particular, RecA is strictly required for recombination to take place. Certain linear recombination events can take place in *recA* mutants but they generally require bacteriophage recombination functions for efficiency (Poteete *et al,* 1999). Several of the recombination proteins in *E. coli* are helicases and recombinant formation depends on the activity of helicase and the RecBCD nuclease complex. In the absence of RecBCD recombination can be efficient if the genes *sbcB*, *sdcC* or *sbcD*, encoding other nucleases are mutated (Poteete *et al,* 1999).

Historically gene replacement was carried out using linear DNA containing the mutated or replaced gene flanked by homologous regions to the chromosome. This was transformed into recombination-proficient strains such as *recBC*, *sbcBC* or *recD* mutants. However this replacement procedure restricts experiments to those involving specific RecBCD nuclease-deficient, recombination-proficient backgrounds and therefore was not applicable in all cases.

Murphy (1998) first reported high frequency gene replacements in nearly any *E. coli* strain without prior cloning of the gene and without dependency on a cointegrate. More recently the construction of a strain that is 10- to 100- times more recombinogenic after transformation with linear DNA, as compared to *recBCD*, *sdcBC* and *recD* strains was reported (Murphy *et al,* 2000). This strain was constructed by replacing the *recBCD* genes of *E. coli* with the λ *red* recombination functions encoded by *exo* and *bet*.

The basic principle of the method described by Yu *et al,* (2000) and Datsenko and Wanner (2000) involves the expression of the λ recombination functions. This system involves the λ Gam protein which shuts down all the enzymatic activities of RecBCD (Poteete *et al,* 1999). The λ recombination system begins with double stranded breaks in DNA, which are the initiation sites for recombination. λ Exonuclease (*exo*) degrades progressively from the 5' ends of these break sites and *bet* binds to the remaining 3' single strand tail, protecting and preparing the recessed DNA for homologous strand invasion (Carter and Radding, 1971).

The λ recombination system has been shown to be very efficient at gene replacements using linear substrates with lengths of flanking homology of more than 1,000 bp (Murphy, 1998). To test the efficiency of short regions of flanking homology (less than 100 bp long) as substrates for λ mediated recombination, Yu *et al,* (2000) modified the λ prophage to express high levels of phage recombination functions for a controlled amount of time. In this prophage the p_L operon is expressed under the control of a temperature-sensitive λ cI-repressor. A deletion removed genes *cro* to *bioA*, removing the lytic genes of the prophage, since the functions encoded by these lytic genes are toxic to the cell and cause cell death(Yu *et al,* 2000).. Another property of the modified phage is that the Cro-repressor is absent. This allows p_L operon expression to be derepressed when the cI-repressor is inactivated at 42°C. Functions of the p_L operon are also toxic to the cell however only after 60 minutes of continuous induction at 42°C. The resulting strain, named DY329 contained the defective prophage harbouring the recombination genes under the temperature sensitive control.

Datsenko and Wanner (2000) constructed plasmids which carry the λ Red recombination system under a controllable promoter induced by arabinose. The low copy number plasmids (pKD20 and pKD46) contain ara-P_{ara} and the genes *gam*, *exo* and *bet*. In their work transformants carrying the Red helper plasmids were induced with arabinose in culture and then were made electrocompetent. Also they constructed template plasmids (pKD3, pKD4) that contain Flp reaction targets (FRT)-flanked antibiotic resistance genes (kanamycin or chloramphenicol respectively) which were used to construct linear DNA cassettes for gene replacements.

The λ red-gam system and some of the genetic tools described by the two research groups in the section above underpins the gene replacement strategy of the current project and will be discussed in more specific detail later.

1.6.3 Gene knockouts using PCR-mediated gene replacements and linear DNA homologous recombination

The aim of the project was to delete the genes *nupC, nupG, yeiM, yeiJ, xapB* and *yegT* from the *E. coli* chromosome so as to look for effects on the *E. coli* mutants when these putative or partly characterised nucleoside transporters were missing from the chromosome. To achieve these deletions, linear pieces of DNA containing a cassette with a kanamycin resistance gene and homology to upstream and downstream regions of the targeted genes were used.

The method chosen was that of Datsenko and Wanner (2000) and Yu *et al,* (2000), which uses PCR primers which have homology to the targeted gene. The genes encoding phage λ Red recombinase, which have been placed under the control of an inducible promoter in a low-copy number plasmid, are present in the host strain during transformation. This is because *E. coli* does not normally tolerate transformation with linear DNA. One reason for that is because of the presence of intracellular exonucleases such as the exonuclease RecBCD, that degrades linear DNA (Datsenko and Wanner, 2000). Contrary to this, genes in yeasts such as *Saccharomyces cerevisiae* are directly disrupted by PCR fragments encoding a selectable marker and short flanking homology regions to the chromosome (Vickers *et al*, 2001).

There are strains of *E. coli* mutants which lack exonuclease V of the RecBCD recombination complex. However recombinants from transformation with linear DNA are rare in such mutants, and such recombination requires large regions of base pair homology (Yu *et al,* 2000).

Many bacteriophages have their own homologous recombination systems and the λ Red (*β, γ, exo*) function has been shown to promote an enhanced rate of recombination. A method has been developed which exploits this function. It involves the use of the strain of *E. coli* DY329 that contains a λ prophage harbouring the recombination genes *exo*, *bet*, and *gam* under the control of a temperature-

sensitive λ cI-repressor (Yu *et al,* 2000). The *exo*, *bet*, *gam* genes are switched on at 42°C and off at 32°C. When the genes are turned on for 5 minutes, cells become more recombinogenic and take up linear DNA without its destruction (Yu *et al,* 2000).

Plasmid pKD4 (Datsenko and Wanner, 2000) was used as the source of the kanamycin resistance cassette for selecting for gene replacement. This is comprised of a kanamycin resistance gene, directly flanked by two FRT sites (Flp Recombinase Targets). Two priming sites (PS1 and PS2) are placed at either end of the cassette (figure 1.4). Firstly primers are designed that have regions of homology (called HR1 and HR2 and are 40bp) to the upstream and downstream regions of the target gene. They also contain regions at the 3' end that correspond to priming sites 1 and 2 of the plasmid pKD4. Conventionally the primers are called Ct and Nt and are in the present study used for construction of the cassettes. An outline of the method is shown in figure 1.4. PCR with the Ct and Nt primers and using pKD4 DNA as a template, will result in production of a linear piece of DNA which will consist of the cassette with tails which are homologous to the target gene.

This PCR product can then be transformed into DY329 cells (Yu *et al,* 2000) and hopefully will replace the target gene by recombination with the linear DNA product. Successful integrants will convey kanamycin resistance and can be selected for, after plating onto appropriate media containing the antibiotic (figure 1.5).Because the regions with homology to the target gene are short there is the possibility that the cassette will integrate elsewhere in the chromosome, such that the strain acquires the antibiotic resistance but the target gene remains intact. To verify that the cassette is correctly inserted at the right position, PCR with two sets of primers are used, as shown in figure 1.6. The first set is designed to regions outside the target gene (Co and No) and the second set of internal primers is homologous to regions of the cassette's kanamycin resistance gene (primers K1 and K2). By performing three PCR reactions the replacement is checked for the fragments NoK1 and CoK2 to confirm that the cassette is integrated as a whole, and also fragment NoCo to confirm that the overall size of the region is correct (figure 1.6).

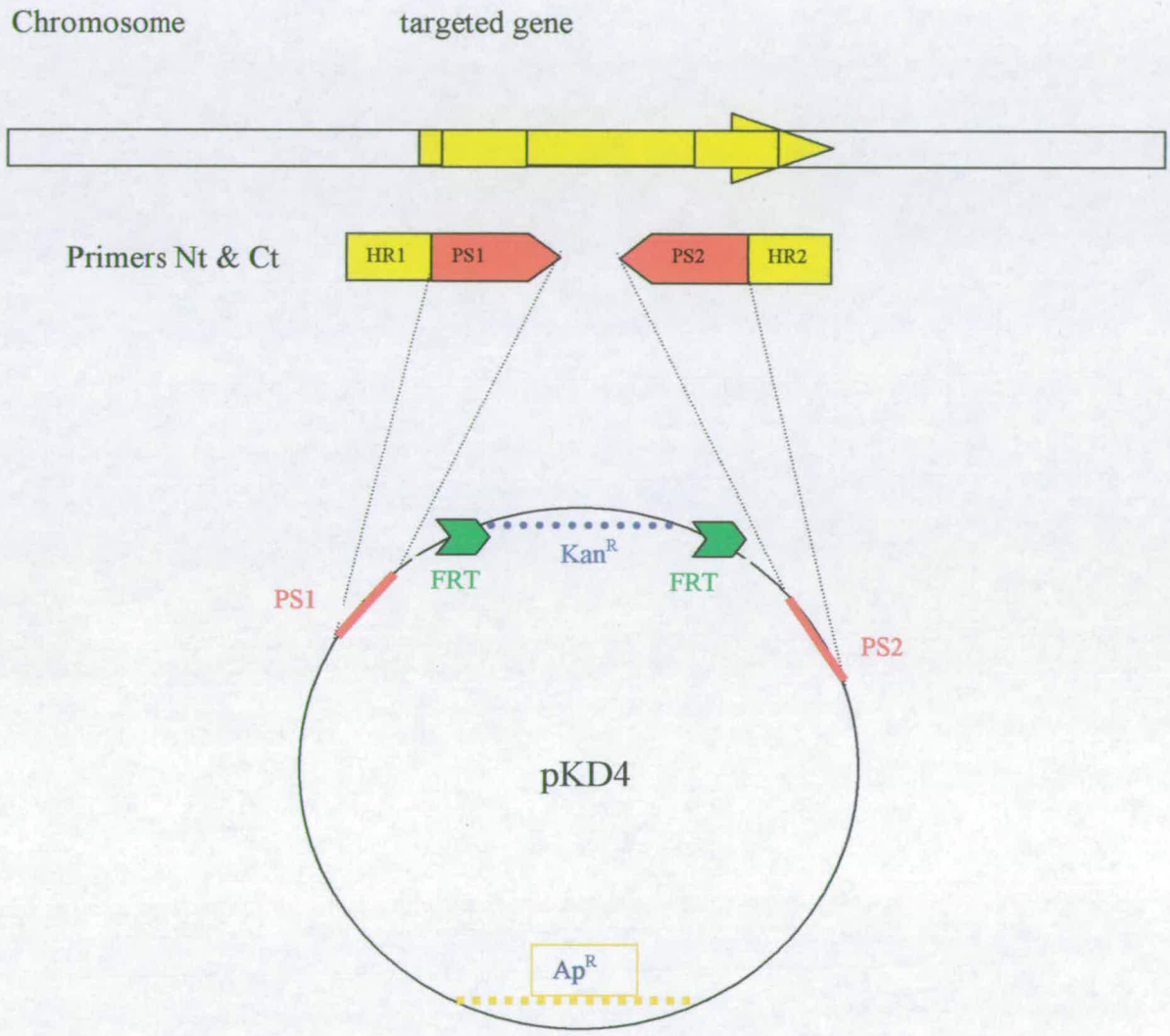


Figure 1.4: Plasmid pKD4 showing homology to primers Nt and Ct.
The primers Nt and Ct are composed of 40bp regions (HR1 and HR2) which are homologous to the target gene, and also regions PS1 and PS2 which are homologous to the regions on the plasmid that flank the cassette (as shown in the picture). The cassette is comprised of the kanamycin resistance gene flanked by two FRT sites (on which the Flp recombinase can promote intramolecular recombination). The cassette is about 1.5Kb in size.

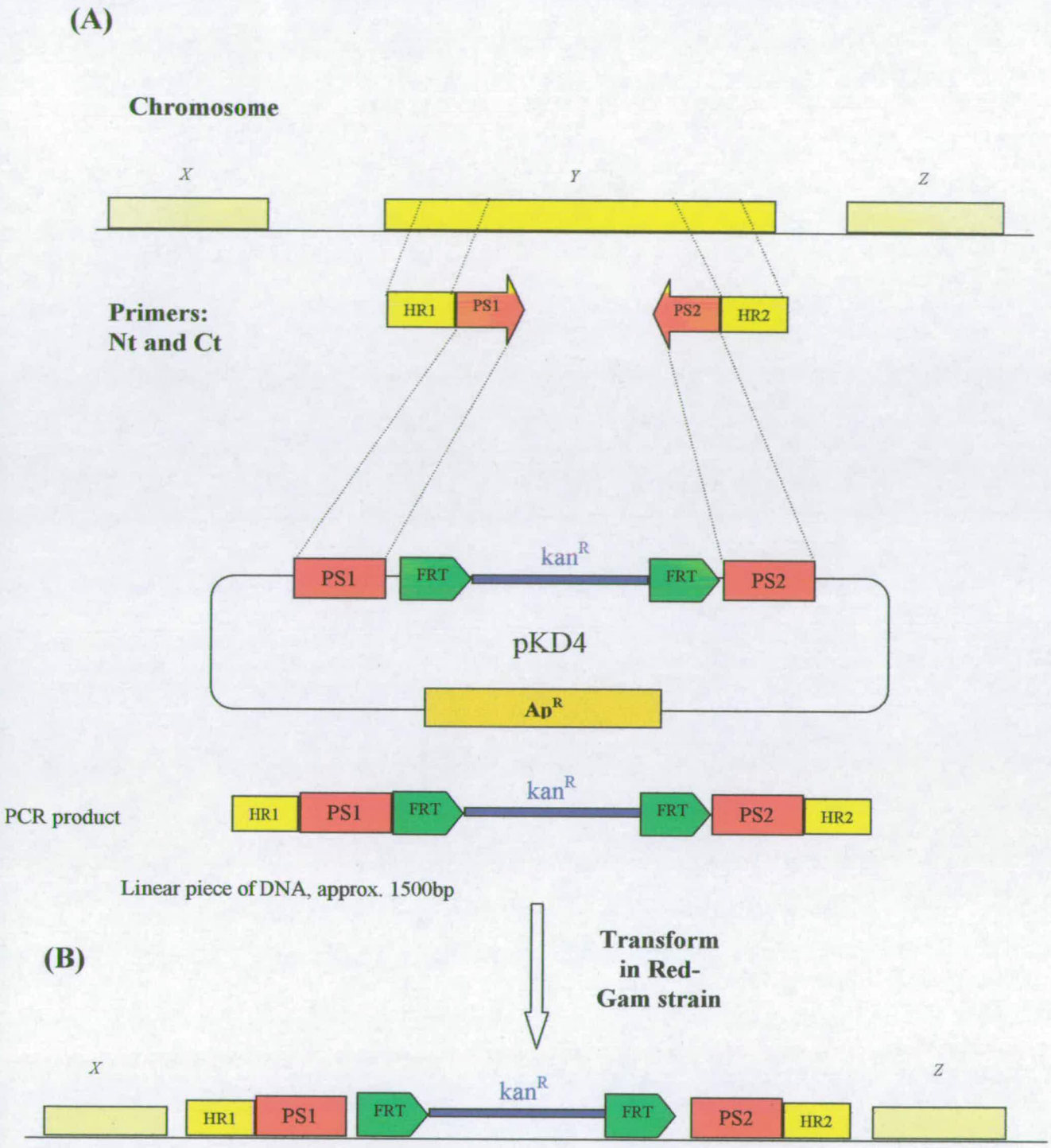


Figure 1.5: A schematic representation of the method used for the replacement of genes.
(A) Primers Nt and Ct are prepared with homologous regions to both the upstream and downstream regions of the gene respectively. These primers also contain the priming sites present in the plasmid pKD4. After PCR with pKD4 as template and the two primers Nt and Ct, a linear piece of DNA is obtained that contains tails homologous to the target gene's start and end.
(B) The linear piece of DNA containing the cassette and the homology tails is used for transformation into suitable host cells that express the Red-Gam system and after successful electroporation, due to double cross-over events, the target gene *Y* is replaced by the cassette.
KEY: The yellow shading signifies the target gene/ chromosomal homology areas (HR1 and HR2), red areas are the priming sites (PS1 and PS2) which belong to the pKD4- derived cassette, while the kanamycin resistance gene is the blue shading, flanked by the FRT sites in green

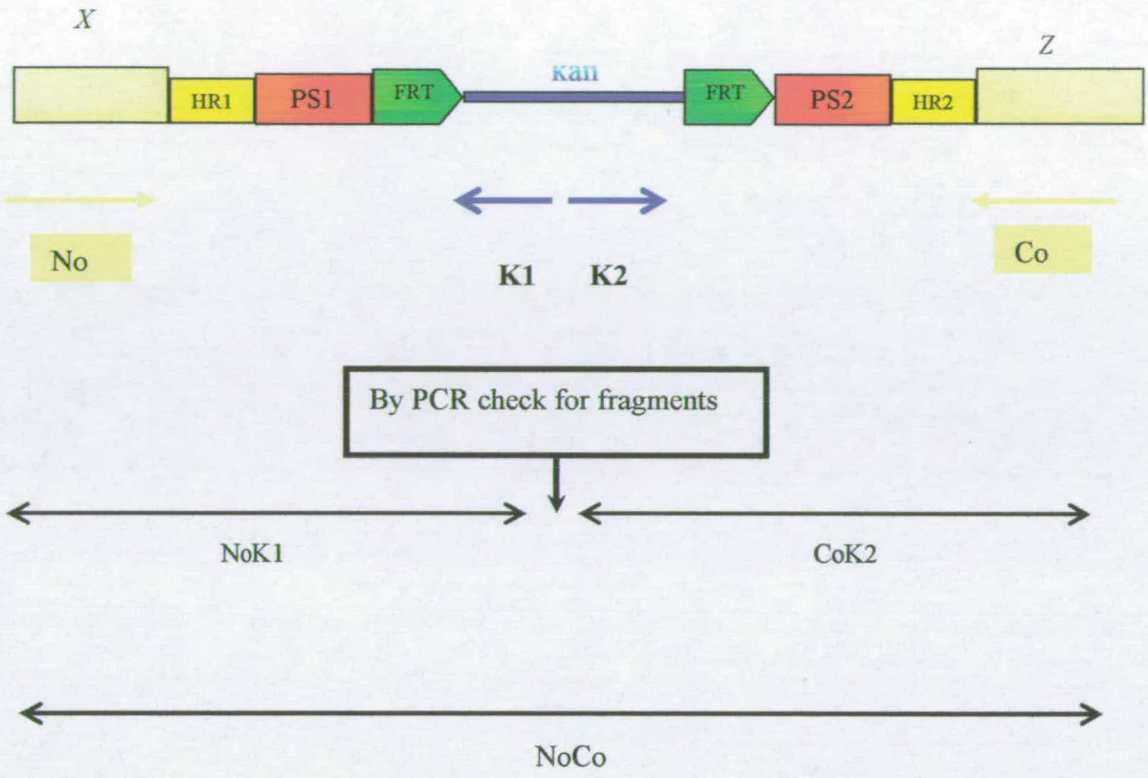


Figure 1.6: Schematic representation of verification via PCR after gene replacements
To certify that the cassette is correctly inserted into the targeted region of the chromosome, which in this case is gene *Y*, three PCR reactions are carried out per replacement. One involves the two external primers shown as No, Co, and the other two reactions involve the two smaller fragments, which in this sketch are represented by the couples NoK1 and CoK2
KEY: The yellow shading signifies the target gene/ chromosomal homology areas (HR1 and HR2), red areas are the priming sites (PS1 and PS2) which belong to the pKD4- derived cassette, while the kanamycin resistance gene is shown in blue shading, flanked by the FRT sites in green. The arrows roughly indicate the DNA bands that are the results of the verification PCR.

1.6.4 Use of Flp for excision of antibiotic resistance cassettes

Recent studies have resulted in the use of site-specific recombination systems as a means of deleting antibiotic resistance cassettes following gene replacement. The yeast enzyme FLP is encoded by the 2-μm plasmid of *S. cerevisiae* and it is a member of the Integrase family of "conservative" site-specific recombinases (Tribble *et al,* 2000). The family has a wide variety of functions but they all utilise a common biochemical mechanism, mediating breakage and joining between two double helical DNA partners harbouring their respective target sequences (Tribble *et al,* 2000).

The FLP enzyme is a site-specific recombinase that promotes recombination at a specific site and has been shown to promote specific recombination between FRT sites *in vivo* in *E. coli* and in several eukaryotic species (Cherepanov and Wachernagel, 1995). The enzyme operates on FRT sites arranged as short direct repeats (Cherepanov and Wachernagel, 1995). The minimal functional FRT site contains only 34 bp and FLP can promote intramolecular recombination between two of these (Huang *et al,*1997).

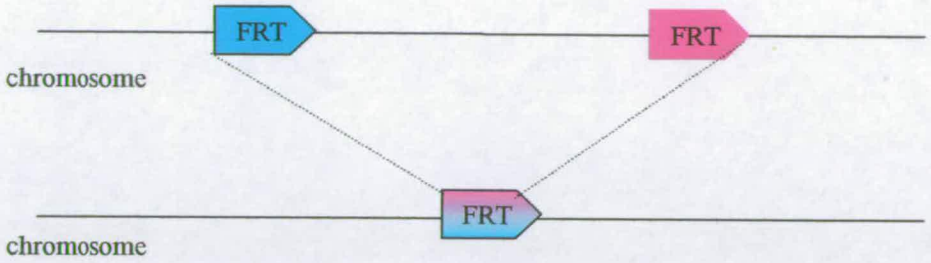
Biochemically, four recombinase monomers participate in the reaction to complete two sequential pairs of strand cleavage and joining steps. A Holliday junction is therefore a basic intermediate in the recombination process (Tribble *et al,* 2000). When two FRT sites are used to flank a selectable marker such as the kanamycin cassette this recombination system can be used to delete the intervening cassette. This circumvents problems where there is a limited availability of antibiotic resistance markers and when multiple deletions are required within the same strain.

Thus with the technology of FRT-FLP it is possible to use the same cassette sequentially as a mechanism for deleting multiple genes by constructing replacements sequentially, one by one, and eliminating the cassette between each replacement and subsequent deletion. However one disadvantage of the FRT system for repeated deletions is that each deletion leaves a conserved FRT site at the point of recombination. Where multiple sites exist (following several deletions) subsequent

recombination events may occur between FRT sites at different points on the same chromosome with or without FLP. This could lead to unwanted deletions or inversions of a large chromosomal segment depending on the orientation and positions of the FRT sites (Wild *et al,* 1998, Hoang *et al,* 1998). Figure 1.7 shows a diagram describing this principle.

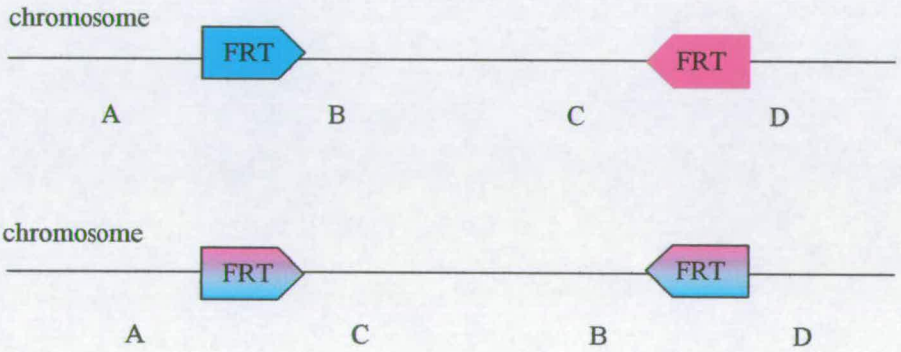
It is therefore important when integrating the same cassette into a second gene in the same host, to place the second cassette in the same orientation as the first one. This would allow deletions, but not inversions, to occur via recombination events at the FRT scars. To facilitate this, particular attention has to be paid to the orientation of the genes that are to be deleted, as well as to the design of the primers, in order to obtain appropriate cassettes specific for each gene.

(A)



If FRT sites have the same orientation then the integrase (flp) may cause deletion of the DNA fragment between them

(B)



If FRT sites have opposite orientations then the integrase may cause inversion of the sequence between them

Figure 1.7: Effect of FLP on FRT sites of same (A) and opposite (B) orientation.
(A) As seen in the simple cartoon above, FLP may promote recombination when two FRT sites are in the same orientation resulting in deletion of the DNA sequence between them.
(B) When FRT sites have opposite orientations FLP will cause inversion of the DNA sequence between them via recombination.

1.7 Aim of the project

The principal aim of this project was to further characterise the known nucleoside transporters and study the putative nucleoside transporters of *E. coli*. Several approaches were used. Firstly strains were constructed with all or all-but-one nucleoside transporter genes knocked out.This was a powerful tool that could be used to dissect the role of transporters in nucleoside uptake.

After construction of the nucleoside transporter-negative stain (null strain) as well as the construction of monogene strains (in which one transporter was left intact) the project aimed to investigate what effect the lack of some or all of the six genes of interest would have on the function of the cell in relation to its ability to transport nucleosides.

Such experiments might facilitate an understanding of the function of the putative transporters XapB, YegT, YeiM and YeiJ as well as provide further information on the known transporters NupC and NupG. This information might also shed light on the mechanism(s) and specificity of toxic analogue uptake.

Another aim of this work was to explore the possibility of developing a bioassay with the utilisation of the mutant strains. The rationale behind this part of the experiments was that ultimately, a bioassay which used *E. coli* monogene strains could be easily manipulated and used as a rapid first screening method, to test a range of nucleoside analogues, as well as nucleoside analogue drugs for effectiveness and toxicity levels.

Several approaches were used to characterise the known and putative transporters. Firstly radioactive nucleosides were used as substrates in order to obtain a precise measurement of specificity and uptake of the nucleosides into the cell. Secondly, the monogene strains and the nucleoside transporter-negative strain were tested for their ability to use physiological nucleosides as a sole carbon source for growth.

Utilisation of nucleosides as the carbon source reflects the cells' ability to uptake and catabolise these molecules.

Thirdly two forms of bioassay were examined for use in a high-throughput manner, using a 96 well plate. Firstly it was investigated whether bioluminescence can be used as a direct indication of whether the cell was able to take up toxic nucleosides such as AZT. Alternatively a bioassay was investigated to see if the toxic effect of AZT can be measured through optical density and if OD is a more reliable measure than bioluminescence, after administration of the toxic molecule.

Chapter 2
Materials and Methods

2.1 MATERIALS

2.1.1 Oligonucleotides

Primers for sequencing and PCR reactions were obtained from MWG (www.mwg-biotech.com) and Sigma-Genosys Biotechnologies Ltd, stored at -20°C and are listed in Tables 2.1-2.3.

Gene targeted Accession no.	Primers for cassette Nt sense	Ct antisense
xapB **EG13159**	5'-CCTCGGCTCTTACATGATTAATACTCTTCATTTCACCGGC-TGTAGGCTGGAGCTGCTTCG-3' (71, 109)	5'-CCAGCCAGATAGTTTGCCAGTCTTTTACGCCATCCACCGA-CATATGAATATCCTCCTTAG-3' (1155, 1116)
yeiM **EG12032**	5' GGTGGTGTTACTGGCAATAGCATTTCTGTTGTCAGTGAAT-TGTAGGCTGGAGCTGCTTCG-3' (3, 43)	5' TTACGCCAGACCAATAAAGAACCCGGCAATAGTCGCACTC-CATATGAATATCCTCCTTAG 3' (1251, 1112)
yeiJ **EG12029**	5' GGGAATGGTGGTATTGCTGACGATTGCGTTTTTACTGTCA-TGTAGGCTGGAGCTGCTTCG 3' (30, 69)	5' GCAAGTAAATGTAGGCCTGATAAGCTTGCGCATCAGGCAA-CATATGAATATCCTCCTTAG 3' (1251, 1112)
yegT **EG14063**	5' TCGCGTACCTTACATTACCTGTCATGAAGGAATTAAAAGA-CATATGAATATCCTCCTTAG 3' (-39, 1)	5' GCGATCATCGACCTTGATAGCCGTAATTTCGTTGTCGGAT-TGTAGGCTGGAGCTGCTTCG 3' (1245, 1205)
nupC **EG11971**	5' TGTGCACGGAAATTTAACCTGCCTCATATTTGGAGCAAAT-CATATGAATATCCTCCTTAG 3' (-40, -1)	5' TTACAGCACCAGTGCTGCGATTGACGCAGACAGCACACTC-TGTAGGCTGGAGCTGCTTCG 3' (1203, 1164)
nupG **EG10664**	5' TCCGCGAAGTTGATAGAATCCCATCATCTCGCACGGTCAA-CATATGAATATCCTCCTTAG 3' (-40, -1)	5' TTAGTGGCTAACCGTCTGTGTGCCTGTCGGGACACGAACG-TGTAGGCTGGAGCTGCTTCG 3' (1305, 1265)

Table 2.1 Primers used to produce cassettes. Shown is a list of primers used to synthesise cassettes from genes *nupC, nupG, xapB, yegT, yeiM, yeiJ*. The numbers within the parentheses refer to the position at which the primer lies with respect to the published gene. Underlined sequences state the gene regions whereas the rest is part of the template vector sequence. (This layout has been used for all primer tables).

Gene name (accession number)	External primers for verification of replacement	
	No sense	Co antisense
xapB **EG13159**	5' GGGTGAGGCCAAATGATGCG 3' (287-297)	5' CTGAAGAGCAGAGCCGCGAA 3' (160-180)
yeiM **EG12032**	5' GCGGTGACAGCCTGAAATCC 3' (10276--10296)	5' CGCGCAGTTCATTCACGACGG 3' (1486-1507)
yeiJ **EG12029**	5' AAACCCAGGAGATGTACGTCG 3' (2985-3006)	5' ATGCTGACCGTAATACCGCAG 3' (4927-4948)
yegT **EG14063**	5' CGGTGCTGTAAAAGGTTGTC 3' (1178-1198)	5' AGGAAGAAAACGGTCAATCC 3' (2876-2896)
nupC **EG11971**	5' GCAGAGAGGATCTGAATCAGC 3' (4174-4195)	5' GCCGAAGTATCCAGTTCTTCG 3' (6042-6063)
nupG **EG10664**	5' CGCAGGCAGCGTGCTGCGAG 3' (977-997)	5' AGAAGGGGCCTTGCCTTATCC 3' (2842-2863)

Table 2.2: Primers used for gene replacement/deletion verification: Shown is a list of primers used to verify gene replacement by the cassettes and gene deletion in the chromosome for genes *nupC, nupG, xapB, yegT, yeiM, yeiJ.*

Gene name	Cloning primers	
	Sense	Antisense
xapB	5' CTGGAATTCGCAGAACTTCATCAAC 3" *EcoRI*	5' CCCAAGCTTTTAATGAGTCACCGCT 3' *HindIII*
yeiM	5' GGGGTACCGTACTCATCAAGCATA 3' *KpnI*	5' ACGGGTCGACATTACGCCAGACCAAT 3' *SalI*
yeiJ	5' GGGGTACCGACATGAAGAGATAAT 3' *KpnI*	5' ACGGGTCGACTCAAGCTAAACCAATA 3' *SalI*
yegT	5' CTCGAATTCCTTACATTACCTGTCA 3' *EcoRI*	5' CCCAAGCTTTCATTTAACTTCCCCT 3' *HindIII*

Table 2.3: Primers used for cloning : Shown is a list of primers used to clone the ORFs of genes *xapB, yegT, yeiM, yeiJ.* into vector pBAD18. Restriction digest sites shown as underlined areas are the sites used to clone each one of the genes.

2.1.2 Bacterial Strains and Media

Bacterial strain	Genotype	Reference
***E. coli* DH5α**	*supE*44 Δ*lacU*169 (Φ80 *lacZ*ΔM15) *endA*1 *hsdR*17 *recA*1 *gyrA*96 *thi*-1 *relA*1	Hanahan (1985)
***E. coli* BLR (DE3)**	*RecA- lon- ompT-* BL21 derivative	Studier *et al.* (1990)
***E. coli* K-12 MG1655**	F- lambda- *ilvG- rfb-50 rph-1*	Blattner *et al* (1977)
***E. coli* DY329**	ΔlacU169 nad::Tn10 gal490 λcI857 Δ(cro-bioA)	Yu *et al,* (2000)

Table 2.4. Bacterial strains used within this study. The table above lists the strains of *E. coli* used within this study. *E. coli* strains K-12 and DY329 were used for the purpose of constructing gene deletion mutants, while DH5α was used in the cloning processes. Finally BLR cells were used for transport assays and for the strain harbouring pucΔ320 obtaining the light emission strain.

All bacterial strains, shown in Table 2.4, were grown in Luria-Bertani (LB) medium (pH7) consisting of 10g bacto-tryptone, 10g NaCl and 5g yeast extract, dissolved in 1litre of distilled water unless stated otherwise for experimental purposes. Agar plates were made by adding $15gl^{-1}$ bacto-agar and pouring 20ml into petri dishes All media were autoclaved to maintain sterility prior to use.

Minimal media used for growth studies were prepared according to the following recipe: 4x M9 Salts (100ml), 20% Glucose (4ml), $MgSO_4$ 1M (1ml), $CaCl_2$ 1M (50µl). Glucose was replaced by appropriate nucleoside when needed as sole carbon source in the media to a final concentration of 1mg/ml.

Cultures were grown routinely by inoculating 5ml LB broth with a single bacterial colony. These were then incubated overnight at 37°C with shaking. Cultures were stored at -70°C by placing 1ml culture and 70µl dimethyl sulfoxide (DMSO) in a cryovial (NUNC). Cells were recovered by streaking a sample on LB agar and incubating overnight at 37°C. Cells were streaked onto LB plates and kept at 4°C for short-term storage. Antibiotics were added where appropriate at a final concentration of $50\mu g\ ml^{-1}$.

2.1.2 Plasmids

All plasmids were stored in either TE (10mM Tris.Cl, 1mM EDTA pH7.5) or dH_20 at -20°C, unless in use when stored on ice. Plasmids used within this study are listed in Table 2.5 below.

Plasmid	Relevant features	Purpose	Source
pGJL25	*nupG* PCR from pGJL24: *Eco*RI, *Xbal* site 3' of His tag and stop codon, product inserted into *Eco*RI/*Xbal* digested pTTQ18	Expression vector of *nupG*	Gallagher collection
pGJL16	*nupC* PCR from pYZ15: *Eco*RI site engineered at start codon so that gene is inserted as in-frame fusion with *lacZ* in pTTQ18	Expression vector of *nupC*	Gallagher collection
pMRB	pBAD18 derivative *xapB* ORF engineered in *Hind*III and *Eco*RI cloning sites	Expression vector of *xapB*	This work
pMRT	pBAD18 derivative *yegT* ORF engineered in *Hind*III and *Eco*RI cloning sites	Expression vector of *yegT*	This work
pMRM	pBAD18 derivative *yeiM* ORF engineered in *Sal*I and *Kpn*I cloning sites	Expression vector of *yeiM*	This work
pMRJ	pBAD18 derivative *yeiJ* ORF engineered in *Sal*I and *Kpn*I cloning sites	Expression vector of *yeiJ*	This work
pKD4	Ap^R Km^R *oriR* plasmid containing FRT-aph-FRT cassette	Cassette construction for all genes	Datsenko and Wanner 2000
pCP20	Ap^R Cm^R *repA* (Ts) pSC101-based plasmid expressing the Flp recombinase	Expressed to produce Flp and transformed into replacement mutant to excise cassette	Cherepanov and Wachernagel 1995
pucΔ320	Expressing lux operon	Bioluminescence –based bioassay	Shaw and Kado, 1986

Table 2.5 Recombinant vectors used in this project: Vectors used within the study for the purpose of radiolabelled nucleoside uptake studies as well as vectors used for gene replacements and luminescence work

2.1.3 Strains constructed for the purposes of this project

All knockout strains were constructed in this study as described in the relevant chapters and in the methods section. All strains were stored at -20°C or on agar plates that were streaked regularly to remain fresh. Strains that are lacking five out of the six genes of interest are referred to as monogene strains in this study. The name of each monogene strain reflects the chronological order in which each mutation was added to the mutant strain.

Single replacement mutants	Relevant features	Purpose	Source
MPG585	*E. coli* DY329 background, ORF of *xapB* replaced by Kanamycin cassette derived from pKD4	Single replacement mutants used as source of ORF replacement for further transduction	This project
MPG590	*E. coli* DY329 background, ORF of *yegT* replaced by Kanamycin cassette derived from pKD4	Single replacement mutants used as source of ORF replacement for further transduction	This project
MPG586	*E. coli* DY329 background, ORF of *nupG* replaced by Kanamycin cassette derived from pKD4	Single replacement mutants used as source of ORF replacement for further transduction	This project
MPG588	*E. coli* DY329 background, ORF of *nupC* replaced by Kanamycin cassette derived from pKD4	Single replacement mutants used as source of ORF replacement for further transduction	This project
MPG587	*E. coli* DY329 background, ORF of *yeiM* replaced by Kanamycin cassette derived from pKD4	Single replacement mutants used as source of ORF replacement for further transduction	This project
MPG589	*E. coli* DY329 background, ORF of *yeiJ* replaced by Kanamycin cassette derived from pKD4	Single replacement mutants used as source of ORF replacement for further transduction	This project

Multiple deletion strains	Genes deleted	Remaining functional NT gene	Source
MPG584 *E. coli* MG1655/ΔBTGMCJ	*xapB, yegT, nupG, yeiM, nupC, yeiJ*	none	This project
MPG578 *E. coli* MG1655/ΔTCGMJ	*yegT, nupC, nupG, yeiM,, yeiJ*	*xapB*	This project
MPG579 *E. coli* MG1655/ΔBCGMJ	*xapB,, nupC nupG, yeiM,, yeiJ*	*yegT*	This project
MPG580 *E. coli* MG1655/ΔTCBJM	*yegT, nupC , xapB, yeiJ yeiM*	*nupG*	This project
MPG582 *E. coli* MG1655/ΔBTGMJ	*xapB, yegT, nupG, yeiM,, yeiJ*	*nupC*	This project
MPG581 *E. coli* MG1655/ΔTCBJG	*yegT, nupC, xapB, yeiJ, nupG*	*yeiM*	This project
MPG583 *E. coli* MG1655/ΔBTGMC	*xapB, yegT, nupG, yeiM, nupC*	*yeiJ*	This project

Table 2.6 Knockout strains in *E. coli* DY329 and MG1655: knockout strains were constructed for the purposes of this study.

Complementation Mutants			
MPG591	*E. coli* BLR background and pMRB	Overexpression of XapB for uptake and kinetics studies	This project
MPG592	*E. coli* BLR background and pMRT	Overexpression of YegT for uptake and kinetics studies	This project
MPG593	*E. coli* BLR background and pMRM	Overexpression of YeiM for uptake and kinetics studies	This project
MPG594	*E. coli* BLR background and pMRJ	Overexpression of YeiJ for uptake and kinetics studies	This project
Lux mutants			
MPG560	*E. coli* MG16565 background with pΔ320	Wild type background which continually expresses the x operon when growing under 33°C for bioassay	This project
MPG567	MPG584 (null) background with pΔ320	Wild type background which continually expresses the lux operon when growing under 33°C for bioassay	This project
MPG566	Monogene MPG578 (*xapB*+) background with pΔ320	Wild type background which continually expresses the lux operon when growing under 33°C for bioassay	This project
MPG564	Monogene MPG579 (*yegT*+) background with pΔ320	Wild type background which continually expresses the lux operon when growing under 33°C for bioassay	This project
MPG 565	Monogene MPG580 (*nupG*+) background with pΔ320	Wild type background which continually expresses the lux operon when growing under 33°C for bioassay	This project
MPG 561	Monogene MPG582 (*nupC*+) background with pΔ320	Wild type background which continually expresses the lux operon when growing under 33°C for bioassay	This project
MPG 562	Monogene MPG581 (*yeiM*+) background with pΔ320	Wild type background which continually expresses the lux operon when growing under 33°C for bioassay	This project
MPG 563	Monogene MPG583 (*yeiJ*+) background with pΔ320	Wild type background which continually expresses the lux operon when growing under 33°C for bioassay	This project

Table 2.7: *E. coli* strains complementation studies in kinetics and uptake experiments and pucΔ320-containing mutants constructed for the bioassay work.

2.1.4 Solutions

- All solutions were made up using distilled H_20 and sterilised by autoclaving prior to use. Heat labile components such as antibiotics were filter sterilised. All solution components are detailed within the text describing the protocol in which they were used. Ampicillin and Kanamycin were both diluted in dH_20 to a concentration of 100μg ml^{-1}.

- Nucleoside solutions were prepared as shown in the table below.

Nucleoside	Stock solution
Adenosine	10mg/ml in hot water
Thymidine	50mg/ml in water at RT
Cytidine	50mg/ml in water at RT
Uridine	50mg/ml in water at RT
2-Deoxy guanosine	50mg/ml in water
Inosine	50mg/ml in hot water
Xathosine	50mg/ml in water, pH 6.5

Table 2.8 Summary of natural nucleoside stock solutions: All solutions were prepared fresh before inoculation of media, filter sterilised and added to a final concentration of 1mg/ml in the media when used as sole carbon source.

- Radiolabelled nucleoside stock solutions

Typically uptake experiments were performed using a single low concentration of permeant usually 50μM. Kinetic parameters were obtained using a concentration range of permeant starting well below the typical Km and raising to at least 5-fold greater than the Km. These experiments were performed using a fixed concentration of radioisotope and varying concentrations of unlabelled permeant. Adenosine unlabelled stocks were prepared in KCl/MES buffer, divided into aliquots and stored in deep freezing at -80 C.

Radioactive nucleoside stocks containing the appropriate concentration of unlabelled permeant plus radiolabelled permeant were prepared. The concentration of the

radiolabelled is 10μCi/ml for ^{3}H -labelled permeant are was Adenosine in this set of experiments (2, 8-^{3}H Adenosine, NET064 from NEN, received as 1mCi in 1ml of aqueous solution with specific radioactivity of 25-50Ci/mmol).

An appropriate volume of 10mM unlabelled nucleoside stock is diluted n KCl/MES to give a final concentration of 2.05mM (this will yield a final concentration of 50μM in the assay). For experiments involving estimation of kinetic parameters of transport use a range of concentrations 41-fold greater than the final concentration desired in the assay.

Just before carrying out the experiment dry down a volume of ^{3}H labelled permeant containing 2.4μCi (typically 2.5μl) in a glass tube under a gently stream of nitrogen in the fume cupboard so as to remove any tritiated water, Dissolve in 250μl of diluted stock to give a 10μCi/ml solution. This solution is then sufficient for 50 uptake points as described in the methods section.

2.2 METHODS

2.2.1 DNA manipulation, cloning and gene deletions

i) Nucleic Acid Amplification: PCR

PCR for verification of correct gene replacements, deletions and gene cloning were carried out in 50µl volumes consisting of 1µg template DNA, 5µl 10x buffer, 4µl dNTP's (0.2mM of each), 1.5µl 5' primer (1mM), 1.5µl 3' primer (1mM), 0.5µl *Taq* polymerase (Boeringer Mannheim) and 36.5µl dH_20. Reactions were carried out in a PCR machine (Techne) as follows (concentrations in brackets represent stock concentrations): Initial denaturation was for 10 minutes at 94°C, followed by 25 cycles of (i) *Denaturation*: 1 minute at 4°C, (ii) *Annealing*: 1 minute at between 50°C - 68°C, depending upon the primers, (iii) *Elongation*: 2 minutes at 72°C. The reaction was completed with a final elongation step of 10 minutes at 72°C. PCR products were observed by agarose gel electrophoresis.

To prepare the linear cassettes the primers shown in table 2.1 were used for each of the genes targeted. Each of the Nt and Ct primers are joined with the suitable priming sites (PS1 and PS2) as described earlier in the report in figures 6 and 7. The reagents used in the PCR were as follows:

PCR for cassette construction were performed with approximately 1µl template DNA pKD4 (0.1µg) per reaction together with 1µl of each of the primers. 1 µl of dNTPs per reaction of final concentration of 5mM. *Pfu* DNA polymerase (Promega) was used for amplification. Initial denaturation was for 1minute at 95°C, followed by 30 cycles of (i) *Denaturation*: 1 minute at 94°C, (ii) *Annealing*: 1 minute at 50°C,(iii) *Elongation*: 2 minutes at 72°C. The reaction was completed with a final elongation step of 10 minutes at 72°C. PCR products were observed by agarose gel electrophoresis.

The linear DNA fragment was purified using a QIAquick™ PCR Purification Kit (QIAGEN) before it was used in transformation. The amplification program used is described in table 2.1 below.

ii) Agarose Gel Electrophoresis

DNA was resolved using a 1% (w/v) agarose gel (FMC BioProducts) in TBE (0.1M Tris-borate, 2mM EDTA). Ethidium bromide (Sigma) was included to a final concentration of 0.5μg ml^{-1} to allow visualisation. DNA was loaded into wells formed by a comb in the gel tray in 5μl loading solution (25% sucrose, 0.5% bromophenol in dH_20). Fragment sizes were determined by running pre-digested *Hin*dIII phage lambda DNA fragments and 100bp ladder (New England Biolabs) concurrently. The gel was run in an electrophoresis chamber (Pharmacia), covered in 1X TBE buffer at 50V for 2 hours. DNA was visualised using a UV light box (300-360nm) and photographed using a digital camera (UVP Laboratory products).

iii) Restriction Endonuclease Digestion of DNA

Digestions were carried out in 10μl volumes containing 1μl 10X buffer, 1 unit of restriction enzyme (Roche and New England Biolabs), 1μg DNA and 7μl dH_20 and incubated in a water bath at 37°C. In the case where two restriction enzymes were incompatible in the same buffer, the product of the first digestion was purified using the PCR purification kit from Qiagen and a second digestion set up. Digestions were heat inactivated at 65°C for 20 minutes where appropriate.

iv) Recovery of DNA from Agarose Gel

DNA was recovered from agarose gel containing ethidium bromide (0.5μg/ml) by viewing under a UV light box, isolating the band using a scalpel and purifying the band using a Qiagen Gel Purification Kit, following the manufacturer's recommendations.

v) Dephosphorylation of Plasmid DNA

After restriction enzyme digestion plasmids were dephosphorylated using shrimp alkaline phosphatase (SAP; Roche) to prevent reannealing during ligation. Reactions were performed in 20µl volumes containing 0.2-10µg DNA, 1x alkaline phosphatase buffer and 1.0 unit of SAP at 37°C for 1 hour. The SAP was subsequently inactivated at 65°C for 20 minutes.

vi) Ligation of Insert into Vector

Ligations were performed in 10µl volumes containing vector and insert DNA (at a ratio of 5:1), 1x T4 DNA ligase buffer and 1 unit of T4 DNA ligase (Roche) and incubated overnight at 16°C.

vii) Blue/white Screen for Recombinant Plasmids

The vectors pGEM-T and pMAL are specifically designed to allow selection of plasmids containing a gene of interest. The cloning sites lie within the coding region of the β-galactosidase α-peptide so that if a fragment is inserted which disrupts the open reading frame, the *lacZα* gene is inactivated. Thus in *E. coli* strains lacking *lacZ* but carrying constitutively expressed lacZβ, recombinant clones can be selected for by screening for β-galactosidase activity. The screen involves selection of transformant *E. coli* DH5α cells on LB medium containing 50µg ml^{-1} carbenicillin. Screening is provided by the addition of 40µl of a 2% X-gal solution (w/v, in dimethylformamide) and 60µl of a 2% IPTG solution (w/v) to the LB medium. After overnight incubation of the plates at 37°C, colonies harbouring pGEM-T or pMAL appear as blue colonies, whilst cells bearing these plasmids with inserts appear white.

viii) Induction of λ recombination and preparation of electrocompetent cells

Overnight cultures of DY329 were grown with shaking at 30°C in Luria-Bertani medium (LB) with the addition of tetracycline at final concentration of 10µg/ml. 50ml of fresh LB containing tetracycline was then inoculated with 1ml of the overnight culture and were incubated at 30°C with shaking until an OD_{600} of 0.5-0.6.

The culture was transferred into smaller flasks and incubated in a 42°C waterbath for 15 minutes to induce the functions of λ. The culture was transferred into smaller flasks to increase the surface area of the culture that comes in touch with the waterbath's water. Immediately the flasks were transferred into an ice water slurry for 10 minutes to cool down and the samples were swirled.

30ml of the sample were centrifuged at 5,000rpm, at 3°C for 10 minutes. Cells were suspended in 2ml ice cold water, then they were divided into 1ml aliquots in Eppendorf tubes. The samples were then centrifuged for 20 seconds, at 4°C at maximum speed in a micro centrifuge and suspended in 1 ml ice-cold water. The previous wash step was repeated twice. To finally collect electrocompetent cells, samples were spun at maximum speed, at 4°C and resuspended in 200µl of ice-cold water. Therefore beginning with 30ml of culture 200µl of competent cells were obtained. Fresh competent cells were prepared for transformation and used the same day.

ix) Transformation of E. coli by Electroporation

Transformation of linear DNA for cassette integration into the bacterial chromosome was performed by mixing purified linear DNA (100ng) with 50µl of competent cells on ice and then pipetting into a pre cooled electroporation cuvette (0.1cm). Electroporation was performed using a Bio-Rad Gene Pulser set at 1.8kV, 25µF, with the pulse controller at 200 ohms. After the pulse 1ml of LB+0.4% glucose was added to the cell suspension. The cells were allowed to recover at 30°C for 2 hours. Then they were serially diluted and plated out onto LB/kanamycin plates and incubated for two days at 30°C.

Any remaining cured cells, were incubated at room temperature, without shaking overnight and they were plated out on one LB/kanamycin plate. Many recombinants were isolated from this overnight incubation. Any transformed colonies present on the LB/kanamycin plates were purified twice on LB/kanamycin plates by streaking for single colonies.

Electroporation of competent cells after cloning of ORFs into multicopy plasmids for radioactive uptake experiments was done by diluting a 5 ml overnight culture in 500ml LB medium and growing the bacterial strain until an optical density of 0.4 at 600nm was obtained. Cells were chilled for 1 hour on ice before being pelleted at 4000 x *g* for 10 minutes at 4°C. Cells were then washed by resuspending in 150ml of chilled sterile water and repelleted. The cells were washed a further 3 times, before being resuspended in 0.5ml chilled distilled water and divided into 100µl volumes. DNA (10ng) was added to each 100µl aliquot, before gently mixing and leaving for 5 minutes on ice. Cells were then placed in an electroporation cuvette and electroporated at 2.5 kV, 200Ω for 5 seconds using a BIORAD Gene-Pulser. Subsequently, cells were incubated in 800µl SOC medium (20g Difco bacto-tryptone, 5g Difco yeast extract, 10mM NaCl, 2.5mM KCl, 10mM $MgS0_4$ and 20mM glucose in 1L dH_20), with shaking at 37°C for 1 hour, and subsequently plated on LB agar plates containing 50µg ml^{-1} carbenicillin.

x) Transformation of E. coli with heat-shock for cassette excision

After the replacements were transduced into strain MG1655 then the plasmid in strain BT340 (DH5-α/pCP20) was used to transform and excise the cassette. pCP20 encodes for the enzyme Flp described above. Strain BT340 was grown in LB with ampicillin (100 µg/ml), chloramphenicol (25 µg/lml), or both, at 30 °C. Plasmid extraction was performed using "Wizard® Plus SV Minipreps DNA Purification System" by PROMEGA. A detailed protocol of the heat shock transformation for cassette excision is shown in Table 2.9

DAY 1

Transforming with pCP20

- make up culture 1/50 from overnight of MG1655/cassette and grow up to OD_{540} =0.5 –0.6
- put culture in ice for 5'
- take 2ml culture, spin at full speed, at 4C, for 20" (2 x 1ml)
- add 1ml precooled $MgCl_2$, 0.1M and resuspend pellet (1 x 1ml-pool)
- spin at 4C, for 20", full speed
- add 100µl precooled $CaCl_2$ 0.1M and resuspend the pellet

DO NOT VORTEX-DO NOT PIPETTE-FLICK WITH FINGERTIP

- leave on ice for at least 2 hours (can leave o/n but not more than 12 hours)
- add 10µl pCP20, mix with finger
- 30' on ice
- 2' in 42 C waterbath (for heat-shock)
- 2' in ice
- add 900µl LB+0.4% glucose
- let cells recover for 1-2 hours @30°C
- plate out onto LB/Amp at 30°C

DAY 2

Purifying the colonies

- pick colony from plate (1-2 colonies)
- purify it once on LB/Amp (should be positive) RESISTANT
- check on LB/kanamycin for sensitivity

DAY 3

Eliminating the plasmid

- pick colony from LB/Amp and plate on LB incubate at 42 C
- pick colony and streak again on LB @ 37C, and also check LB/Kanamycin LB/Amp at 37C to check sensitivity

Table 2.9 Protocol followed for expression of Flp recombinase from pCP20 and heat shock transformation of *E. coli* MG1655 for cassette elimination.

Site-specific recombination systems were used to delete antibiotic resistance cassettes following gene replacement. The yeast enzyme encoded by the FLP protein is an example of such a system and it operates on FRT sites arranged as short direct repeats. The minimal functional FRT site contains only 34 bp and the Flp can promote intramolecular recombination.

The Flp enzyme is a site-specific recombinase which promotes recombination at a specific site and it is shown to promote specific recombination between FRT sites *in vivo* in *E. coli* and in several eukaryotic species.

The Flp protein is encoded by the 2-μm plasmid of *Saccharomyces cerevisiae* and it is a member of the Integrase family of "conservative" site-specific recombinases . When two FRT sites are used to flank a selectable marker such as the kanamycin cassette this recombination system can be used to delete the intervening cassette. This circumvents problems of the limited availability of antibiotic resistance markers.

xi) DNA transduction in E. coli using phage P1

Preparation of P1 lysates

The replacements from DY329 were moved to a wild type strain *E. coli* MG1655 via*E. coli* phage P1. P1 is a temperate phage which will pack DY329 DNA and then after transduction will hopefully recombine it with the wild type strain. Thus each one replacement can be transduced one after the other and by making use of the FRT sites the cassette can be excised and be used as a marker for the others

DAY 1
• grow *E.coli* DY329/ Δ gene strain overnight in LB and kanamycin • next day inoculate 5ml LB + Ca^{2+}/kanamycin antibiotic and grow up to OD_{600}=0.9 • mix 1 ml of the cells with 1.6µl of P1 lysate (provided lysate was $6x10^{10}$) • also mix 1ml of cells with 16µl of the P1 lysate • incubate them for 15' at 30°C • add 3ml LB+Ca^{+2}, and 3ml LC top agar +Ca^{+2} • pour onto LB plate and let them set • incubate at 30/37°C overnight without inversion
DAY 2
• transfer top agar layer into a sterile beaker and wash plate with 5 ml Ca^{2+}, transfer wash into the same beaker • add 0.5ml $CHCl_3$ and shake at 30°C for 20 minutes gently • spin for 15minutes at 4,000 rpm • transfer supernatant into sterile bijou, add 0.1ml $CHCl_3$ and store at 4°C

Table 2.10 Protocol followed to prepare P1 lysates

Transduction

The recipient strain MG1655 was prepared in a LB medium enriched with Ca^{2+} which is necessary for the survival of the P1 virions. The recipient cells were mixed with the lysates and selected for on LB/kanamycin plates. Survival of the strain on kanamycin shows that MG1655 has acquired the replacement cassette. The protocol in table 2.11 summarises the processes followed. Note that this step was followed by the elimination of the cassette via pCP20 transformation. Part of this protocol states that the strain should be incubated at 42°C in order to eliminate the pCP20 helper plasmid. Furthermore this step ensures that if the transductants isolated on the LB/kanamycin plates were in fact DY329 and not MG1655, they would then die, since DY329 will not grow at this temperature.

- Grow MG1655 strain up to OD_{600} =1.000 spin down (4,000rpm for 15') and resuspend in 1/10 volume of LB+Ca^{2+}
- To 0.1ml of recipient strain add 0.1ml of the lysate, mix and incubate at 37°C for 15'
- Add 1ml of phage buffer spin down and resuspend cells in 1ml LB + glucose (10ml LB+gluc=9.82ml LB + 0.18ml of 20% glucose)
- Incubate at 37°C for 1hour, spin, resuspend in 100µl of LB+Ca^{2+} and plate out on LB+Kanamycin plate

Table 2.11 Protocol followed for transduction of P1 into *E. coli* MG1655

xii) Small-Scale Plasmid Preparation

A single colony was removed from a plate and grown overnight at 37°C in 5ml LB broth, containing the relevant antibiotic (50µg ml^{-1}). Of this culture, 1.5 ml was placed in an eppendorf tube before centrifuging at 20000 x *g* for 10 minutes. Cells were resuspended in 100µl chilled GTE (25mM Tris.Cl, 50mM glucose, 10mM EDTA, pH 8.0), before mixing with 200µl of chilled, fresh lysis solution (0.2M NaOH, in 1% SDS). After 5 minutes on ice, 300µl of neutralising buffer (3M potassium phosphate, 5M sodium acetate, pH 4.8) was added and incubated for 5

minutes at 4°C to neutralise the suspension. The lysate was centrifuged for 5 minutes at 20000 x *g* and the supernatant removed and mixed with an equal volume of phenol:chloroform (1:1). After centrifugation at 20000 x *g* for 10 minutes, the supernatant was removed and mixed with 2 volumes of absolute ethanol (room temperature) before a further centrifugation. The pellet was washed in 1ml 70% ethanol (room temperature), dried and resuspended in 100µl nuclease free water.

2.2.2. Determination of DNA Concentration

Plasmid DNA was diluted 1:200 in dH_2O and the absorbance was measured at 260nm. The concentration was calculated by multiplying the absorbance by 50 (1 unit = 50µg/ml) and by the dilution factor. Purity of the DNA was calculated by measuring absorbance at 280nm.

2.2.3. Sequencing of Double-Stranded DNA

Sequencing was carried out on vectors to determine that the correct sequence was encoded, using BIGDYE v3.1 sequencing kit (Applied Biosystems). Sequencing primers were designed to allow complete sequencing of encoded genes. Primers were complementary to flanking regions of plasmid ~50bp either side of the gene of interest. Plasmid DNA (~500ng) was added to 4µl of BIGDYE sequencing reaction mix, with 1mM primer and made up to 10µl with dH_20. Reactions were carried out in a PCR machine (Techne) under the following conditions: 25 cycles of (i) *Denaturation*: 1 minute at 94°C, (ii) *Annealing*: 30 seconds at 50°C (iii) *Elongation*: 5 minutes at 72°C. After the reaction products were purified by gel filtration, dried using a vacuum dessicator and then stored in the dark at 4°C. Samples were analysed by the sequencing service within the ICMB or ICAPB, University of Edinburgh, and the sequencing information received was read using the program Vector NTI

2.2.4 Radioactive nucleoside transport assays

The method described below explains how bacterial cultures were prepared for nucleoside transport assay using radioactively marked nucleosides as the substrate. More details are found in the paper describing sugar uptake in *E. coli* by Henderson *et al*, 1977.

i) Preparation of Bacterial culture

E. coli BLR harbouring the nucleoside transporters were grown overnight at 37°C in 3ml LB containing 100µg/ml ampicillin shaking at 200rpm in 15ml Falcon tubes. Flasks containing 20ml M9 medium with glycerol ($CaCl_2$ to final concentration of 0.2mM, glycerol to final concentration of 0.2%, $MgSO_4$ 2mM, and casamino acids to final concentration of 0.2% made up to 1l with sterile milliQ water) were inoculated the next day with 1ml of the overnight culture. One culture was used as the uninduced control, the other induced with IPTG or arabinose. Also included were vector-only cultures (no insert) so as to control for background levels of transport. The cells were grown for 3 hours at 37°C with shaking as before whereupon induction occured with IPTG (final concentration 200mg/ml) or arabinose (final concentration 10mM) for about 1hour. Cells were harvested by centrifuging in 50ml Falcon tubes for 10 minutes at 3000rpm at 4°C. Pellets are resuspended in 15ml 150mM KCl/5mM MES buffer pH 6.6 by pipetting gently to avoid damage to the cells. Centrifugation and washes were repeated 3 times in total and then resuspension was in 5ml KCl/MES. A 1:10 dilution for the cell preparations was checked for OD_{680} and adjusted (as needed by diluting cells) to 1.8-2.2.

ii) Transport assay

196µl of bacteria were aliquoted into a 5mlbijou bottle containing 4µl 1M glycerol (final concentration 20mM). Cells were mixed and left to equilebrate at room temperature before carrying out the assay. For each uptake point the bijou bottle was placed in the water-jacketed glass holder at 25°C. The clock was started and the bacteria renenrgised by bubbling with air for 3minutes using a fresh yellow tip. The suspension should be bubbling slowly. A membrane filter (Gelman GN-6, pre soaked

in KCl/MES) is placed on the vacuum manifold. In 5µl ^{3}H-nucleoside (in this case adenosine) and at 3 minutes exactly the radiolabelled nucleoside was added to the cell suspension. This was time 0. Cell were allowed to uptake nucleoside for as long as desired (typically 15sec, 30sec, 1min, 3min for uptake and 15sec for kinetics and competition experiments). In particular at desired timepoints 80µl of the cells were put onto the filter and the bijou bottle was returned to the jacketed holder replacing the air supply. Immediately the filter was rinsed with 3 times 2ml of KCl/ MES and when the filter was dry it was placed in a scintillation vial. 10ml of scintillation fluid (Emulsifier Safe, Perkin Elmer) was added, and the vial was caped and vortexed briefly. Vials were placed in the dark for 5hours and vortexed again before counting in the scintillation counter. Controls were included of standards in duplicate (5µl of radiolabelled nucleoside stock pipetted directly onto the filter) as well as filter blank controls (min 200µl KCl/MES and 5µl of radiolabelled nucleoside, pipetted 80µl onto a filter and washed the same as for the samples).

2.2.5 Growth studies

1ml of overnight cultures were used as inocula to start cultures in 250ml glass flasks which contained 50ml of pre warmed LB at 30°C. The strains used for these purposes were the ones which had been transformed with the plasmid pucΔ320 which will continuously emit light at 405nm. Cultures were incubated at 30°C in a shaking water bath and 1ml aliquots were taken every 20 to 25 minutes to monitor optical density (OD) at wavelength of 600nm. Immediately before measuring OD, 50μl of the aliquot was placed in a 96 well plate and light emission was measured in the luminometer (Luminoscan). Growth profiles were drawn from these measurements OD_{600} and some data were derived in terms of the light emission profiles that correspond to the growth of each strain.

For OD_{600}-based growth profiles in minimal media with a nucleoside as the sole carbon source measured in the 96 well plate, overnight cultures were prepared from a single colony in a bijou bottle containing minimal medium and incubated at 37°C. These overnight cultures were then used to inoculate other minimal media containing nucleosides as the sole carbon source which were left to grow at 37°C for 4 hours. 50μl were then transferred into an ELISA plate and were placed into the ELISA reader (ELISA microplate reader, Alpha Diagnostics) in Edinburgh University's COIL facilities choosing the settings so that the machines would take measurements of OD_{600} every 20 minutes for an appropriate number of hours (usually 14 to 24 hours). The readings were then obtained using KC5 software.

Chapter 3

Construction of mutants lacking nucleoside transport systems

3.1 Introduction

A common and popular approach to the study of uncharacterised genes is to delete their coding sequence from the genome and look for resulting phenotypic changes. The methods used for deletions in the *E. coli* genome vary and over the years they have been improved so that gene deletions are as precise as possible. The main idea behind most such techniques is the exchange of altered genetic material constructed *in vitro*, that usually contains a selectable marker, with the targeted chromosomal area.

So far two nucleoside permeases, NupC and NupG have been characterised in some detail (Munch-Petersen and Mygind, 1983) while Nørholm and Dandanell (2001) have been studying the specificity of XapB and compared it to NupC and NupG. In the latter work they expressed the ORF of each one of these genes in low copy number plasmids and subsequently transformed them into transport-defective strains which were mutated by point mutations in the three ORFs for genes *nupC*, *nupG* and *xapB* (Nørholm and Dandanell, 2001).

The work of this thesis has focused on the previously studied genes *nupC*, *nupG* and *xapB*, as well as genes *yeiM*, *yeiJ* and *yegT* that are putative nucleoside transporters as revealed after a BLAST search (http://genolist.pasteur.fr/Colibri). Within the contents of this work the six genes were studied by precisely engineering deletions of their coding sequence so that mutants can be clear of any obscuring marker sequences or remnants of ORF of the actual genes that might interfere with mutant phenotypes.

To functionally characterise the putative nucleoside transporter-encoding genes *yeiM*, *yeiJ*, *yegT*, as well as genes *xapB*, *nupC*, *nupG* (for which there is some degree of understanding of function), the construction of *E. coli* mutants from which these genes were deleted from the bacterial chromosome was required (Baldwin *et al,*

1999, Seeger *et al,* 1995). Ultimately the mutant needed for the purposes of the series of characterisation experiments was a "null" *E. coli* strain from which all six genes were deleted. Also six monogene mutants from whose chromosomes five of the genes were deleted and where only one functional gene remained, were constructed.

To delete the chromosomal regions encoding genes *nupC*, *nupG*, *xapB*, *yeiM*, *yeiJ* and *yegT*, the method of linear recombination published by Murphy (1998) was followed. This technique outlines the method by which open reading frames can be deleted from the chromosome of *E. coli* using specially constructed cassettes. The cassettes are linear pieces of DNA constructed *in vitro*, which are homologous with the regions flanking the targeted chromosomal ORFs. These cassettes replace the ORF via recombination events, resulting in the replacement of the ORF by the cassette which harbours an antibiotic resistance gene and which can be easily selected for.

Because the antibiotic resistance gene on the cassette is directly flanked by recombination sites called FRT (standing for *flp* Recombination Targets), on which *flp* recombinase acts specifically, the step of replacement can be followed by a step of deletion of the cassette, including the antibiotic resistance gene, via recombination between the two FRT sites (Cherepanov and Wachernagel, 1995, Wild *et al,* 1998). Only one antibiotic resistance gene is necessary for multiple deletions since the approach is cyclical- each deletion event is separate and follows the excision of the cassette (Yu *et al,* 2000). This technique allows accurate alterations in chromosomal areas of interest where the genes coding for known and putative nucleoside transporters lie.

Previous studies have attempted to analyse phenotypic changes in bacterial nucleoside transporter mutants. However such studies were based on mutants which contained only point mutations. Also no data whatsoever exist on the function of

newly found genes encoding putative transporter paralogues *yegT, yeiM* and *yeiJ*. An analysis of how the method was applied in this particular project and the results acquired, is given in the following sections.

3.2 Replacements of genes

The sequence of events for the replacement of the six genes was to firstly construct six cassettes, with plasmid pKD4 as the template which provided the following key characteristics to the cassettes: a) the antibiotic resistance gene *kan* encoding for kanamycin resistance and b) two recombination sites directly flanking the kanamycin gene, called FRT sites on which *flp*, an enzyme which acts on such sites promoting recombination events at FRT sites. The *flp*-mediated recombination event was put in use later in the process, when replacement of the genes had occurred. Importantly the cassettes possessed long homology regions (HR) to DNA sequences outwith the ORFs of each of the genes to be deleted (hereafter called HR1 and HR2). Finally the cassettes also comprised two priming sites directly flanking the FRT sites that belonged to the pKD4 vector and were used as initiation points for the amplification of the PCR reaction which resulted in the production of the linear DNA, genes-specific cassettes (for a diagram of pKD4 see figure 1.3).

The linear cassettes were transformed into a highly recombinogenic strain of *E. coli* called DY329. DY329 was an ideal candidate strain for the purpose of this work as it is deficient for genes encoding for exonucleases (Yu *et al,* 2000). In DY329 transformation with linear DNA fragments was feasible, as the linear cassettes did not get degraded by exonucleases that normally act on foreign pieces of linear DNA and destroy it.

The result of these processes was the construction of six strains of *E. coli* DY329 in which the cassette has replaced each one of the six genes of interest. These strains were now lacking the ORF of the particular gene and had acquired a kanamycin resistance gene. A schematic representation of these processes has been included in Chapter 1 (1.6.3).

3.2.1 Construction of the cassettes and primer design

The cassettes were each constructed based on plasmid pKD4. This plasmid contains the antibiotic resistance gene *kan* encoding for kanamycin resistance that is directly adjacent to, and between two FRT sites. On either side of the FRT sites there are two small priming site regions, PS1 and PS2, which were used as a base for the construction of the cassette (Datsenko and Wanner, 2000). These areas also constitute the amplifying basis for the PCR-mediated cassette construction used for the replacement of the genes.

In order to replace the genes with the cassette via homologous recombination it was necessary for the cassette to contain regions of the sequence homologous with flanking regions of the ORF to be replaced. Therefore PCR reactions were performed with the pKD4 plasmid as template DNA and primers were designed which had homology to PS1 and PS2 but also had long tails incorporated to their ends which were corresponding to the sequence adjacent to the ORFs start and end.

Thus when transformation of the amplified linear cassette was performed with the recombinogenic *E. coli* strain DY329, the homology regions of the cassette acted as a good template for homologous recombination between the linear DNA cassette and the homologous target gene region. Through double cross over events in these areas, the ORF was replaced by the cassette (Yu *et al,* 2000).

It is of great importance to stress that when carried out sequentially, the end product of this process would be a strain in which all six genes were missing. For this reason it was important that in the final strain, the FRT scars face in the same orientation so as to avoid unwanted chromosomal inversions between FRT scars in the multiple knockout mutant as a result of natural recombination. To do so the cassettes would have to be positioned on the chromosomal area in such a way so that the FRT scars resulting from different replacement events have the same orientation regardless of the gene orientation. For that reason the orientation of each gene on the chromosome had to be taken under consideration in the primer design. As seen from figure 3.1, it is therefore desired to have FRT scars in the same orientation in the multiple mutants so that the only possibility for recombination between the FRT scars would result in chromosomal areas to be deleted and not flipped over. Deletions of such large areas would be detectable by PCR fragments reducing in size dramatically or, by cell death if the chromosomal area that was deleted was a vital one in the cell. Flipping over of the chromosomal areas between two FRT scars would be likely to go unnoticed since there would not be any means to check for it. If such an inversion occurred any subsequent phenotypic change could be attributed incorrectly to one of the mutations of the genes of interest whereas in reality it might be a direct result of the flipping over of the DNA sequence, making any phenotypic change inconclusive.

As mentioned above, the primers designed for the genes' replacements were constructed by using DNA sequences from two separate sources; the DNA homologous to the gene that acted as recombination sites, and the priming sites belonging to the plasmid which amplified the cassette plasmid cassette area during the PCR (see Figure 3.1).

Figure 3.1 below shows a general situation for two genes of opposite orientation and indicates how the primer design determines the orientation of the cassettes on the

chromosome after gene replacement and thereby the orientation of the FRT scars on the chromosome after deletion of the genes.

The regions of target gene homology were about 40 bp long, and they were homologous to regions adjacent to the start and end of the ORF. Such long DNA sequences were necessary, as these areas would subsequently act as the templates for homologous recombination after transformation of the cassettes into *E. coli* DY329. Yu *et al* (2000) recommend that these homology regions are at least 36bp long. It should be noted at this point that according to Datsenko and Wanner (2000) similar number of recombinants were obtained when using primers with 36- to 50- bp of homology.

Within the oligonucleotides used for PCR, the target gene homology regions were followed by shorter priming areas, PS1 and PS2, which annealed to the plasmid during PCR so as to facilitate construction of the cassette. These priming sites were 20bp long. Details about the exact DNA sequence of the primers are given in the Materials and Methods chapter, table 2.1.

The orientation of the genes varied, in that *nupC*, *nupG* and *yegT* were of positive orientation and so, transcribed clockwise in relation to the origin whereas *yeiJ*, *yeiM* and *xapB* were of negative orientation, transcribed anticlockwise in relation to the origin of replication (genolist.Pasteur.fr/Colibri). This meant that particular attention had to be paid in to the way the priming sites were designed in order to avoid subsequent unwanted chromosomal deletion via recombination between FRT sites. As illustrated in figure 3.1 priming sites were designed such that the cassette sat on the chromosome in an orientation so that the residual FRT scars resulting from cassette deletions always faced the same direction. The orientation of the priming sites determine if the cassette will be incorporated in the chromosomal locus in positive or negative orientation.

Thus, if two genes are of opposite orientation on the chromosome, then "outer" priming sites 1 and 2 will have to be designed using the opposite ends of the two genes so as to ensure proper orientation of the FRT sites. Genes *nupG* and *xapB* for example, have opposite orientations on the chromosome; *nupG* being positive and *xapB* being negative. When primers for *xapB* were designed PS1 was connected with homology regions corresponding to the upstream region of the gene and PS2 to the downstream region of the gene. This led to the cassette's FRT sites having a negative orientation on the chromosome. Ultimately this would lead to the FRT scar having a negative orientation on the chromosome too.

As a result when designing the primers for the second gene, *nupG*, the same negative orientation had to be maintained for cassette incorporation. For that reason PS1 was connected to the homology region corresponding to the downstream region of *nupG*, and PS2 to the upstream region of the gene. This led to the cassette's FRT sites, and hence the FRT scar, acquiring a negative orientation. Figure 3.2 summarises the gene orientation in this project and indicates how priming sites were joined to suitable areas of the genes.

After constructing the six replacement mutants then the mutations are transduced into a recipient strain one by one and this highlights the value of keeping the orientation of FRT sites constant. After the first replacement is transduced, the cassette is excised through *flp*-mediated recombination, leaving one remaining FRT scar. This will determine the orientation of the initial FRT sites. Although constructed in a separate strain initially, the second replacement is then transduced into the strain carrying the FRT scar, and again through *flp*-mediated recombination the cassette can be excised leaving a second FRT scar in the same orientation.

By maintaining the same orientation for all FRT sites, the FRT scar orientations are maintained as well. This ensures that if unwanted recombination events occur subsequently between FRT scars and not (or as well as) FRT sites of a newly

transduced cassette, the DNA region between them can only be deleted and not inverted.

Construction of cassettes for the six genes was undertaken with the use of plasmid pKD4 as described above. The primers used for each gene and PCR amplification programs are summarised in Materials and Methods, 2.2.1(i) and table 2.1.

In figure 3.3A and B the cassettes constructed for the six genes are shown. Lanes 1, 4, 5 and 6 (Fig.3.3A) represent the amplified linear DNA cassette constructed by PCR with the appropriate primers as described in the Materials and Methods chapter, for genes *xapB, nupG, nupC* and *yegT* respectively. The two bands corresponding to cassettes constructed for genes *yeiM* and *yeiJ* are shown in figure 3B (lanes 1 and 2 respectively). *Pfu* polymerase enzyme was used in this step to ensure the accurate synthesis of the DNA sequence in the cassette. All cassettes were 1.525 kbp in size.

(I)

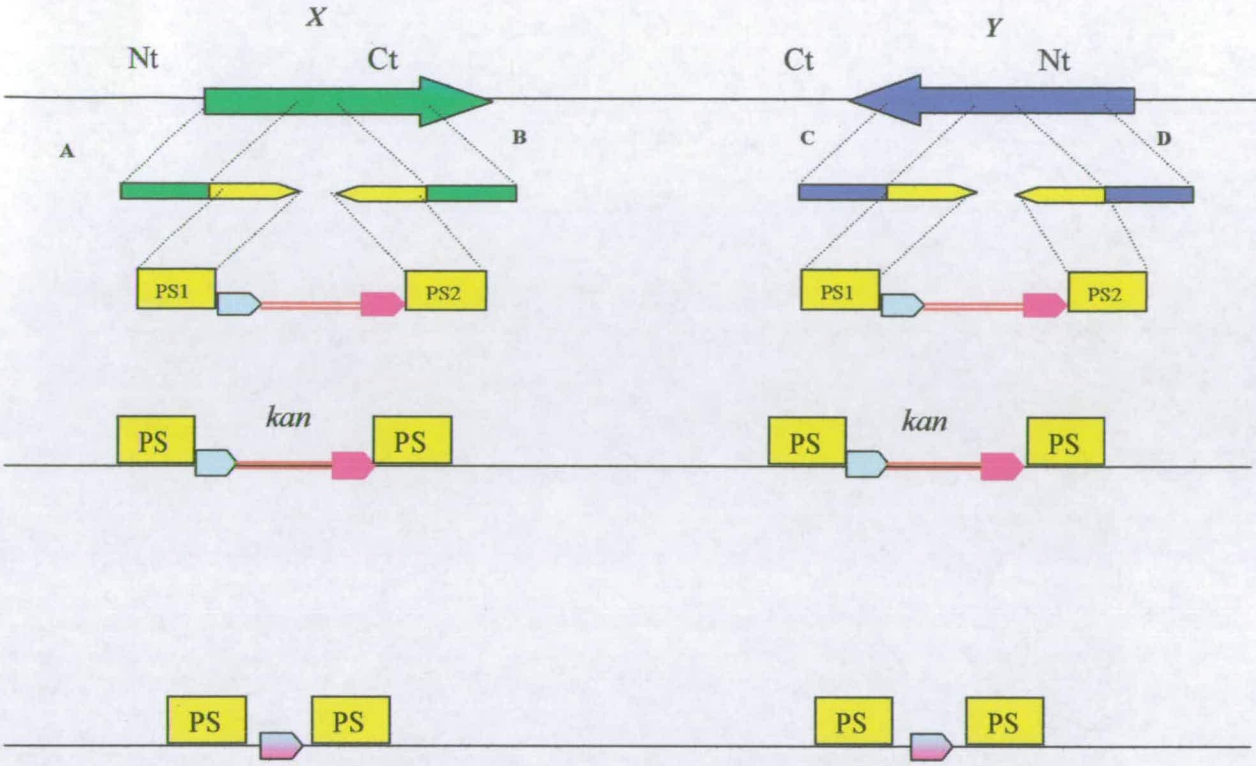


Figure 3.1: Sketch showing primer and cassette design for gene replacements and deletion in a double mutant situation. (I) Genes *X* and *Y* are of opposite orientations on the chromosome. The two cassettes always have the same orientation therefore when PS1 is joined with the N terminus of gene *X* and PS2 with the C terminus of gene *X*, then PS1 will be joined with the C terminus of gene *Y* and PS2 with the N terminus of gene *Y*. (II) In the replacement step of the process, the cassettes have replaced the two genes and have the same orientation on the chromosome. (III) After excision of the cassettes the FRT scars have the same orientation on the chromosome

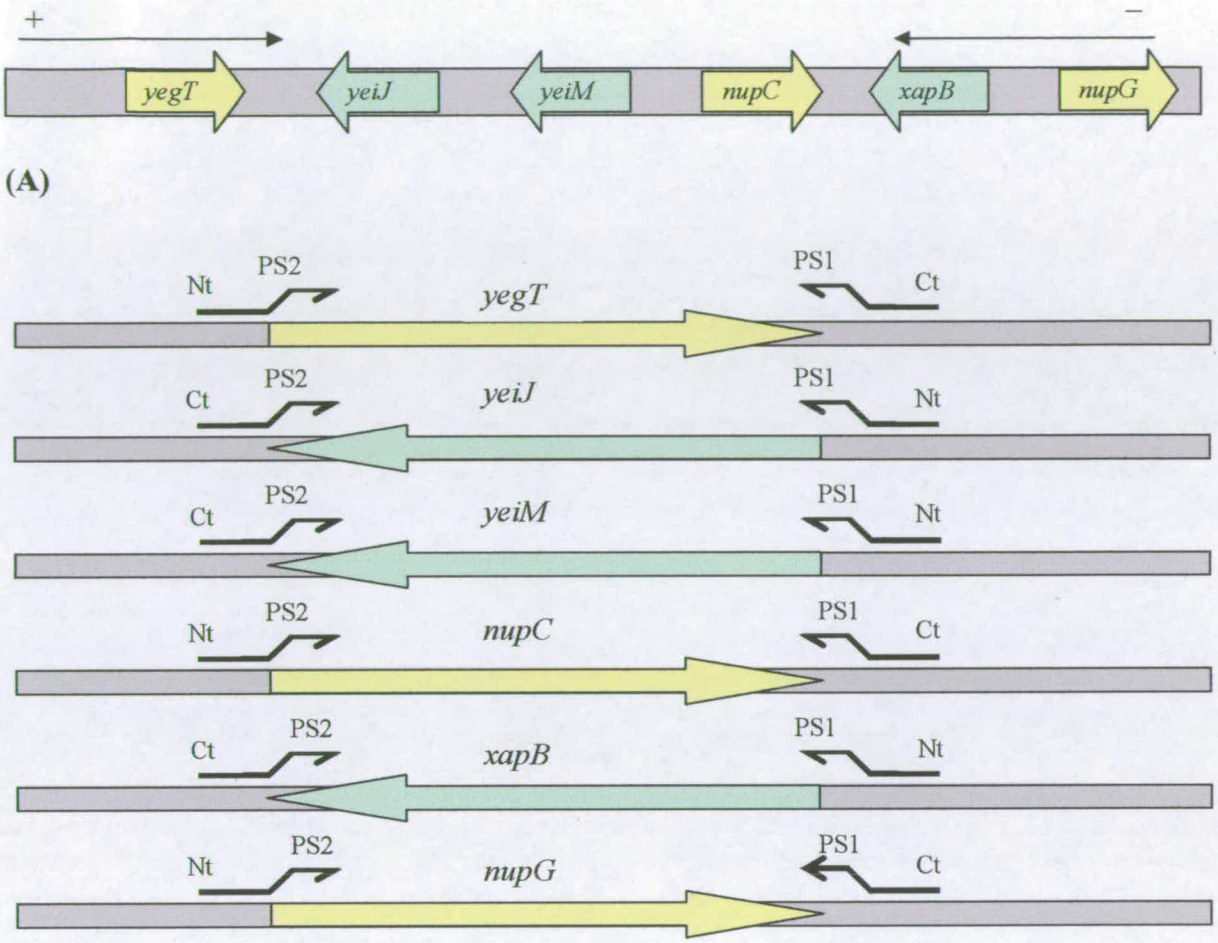
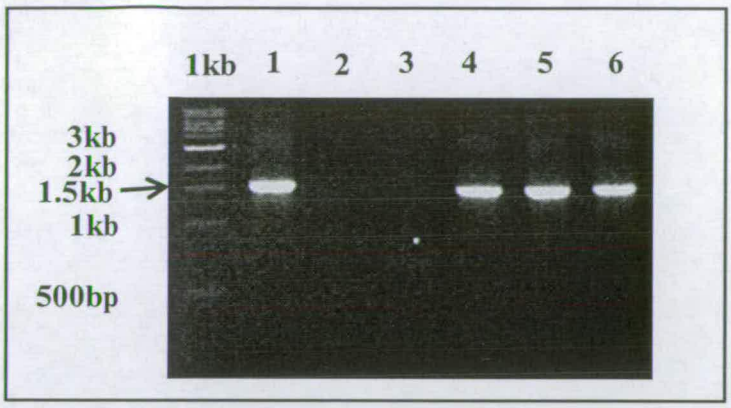


Figure 3.2: Gene orientation and primer design (A) Cartoon showing the orientation of the genes on the chromosome transcribed clockwise/ anticlockwise from the origin of replication. **(B)** Genes and their orientations aligned to show orientation of PS1 and PS2 sequences. PS1 and PS2 were designed and joined to the start and end of the genes according to orientation. The diagram shows how primers were connected to the appropriate side of the gene so as to always align FRT sites in the same direction. C terminal (Ct) signifies the end of the ORF whereas Nt the N-terminus. The direction of each gene is shown according to the arrowhead.



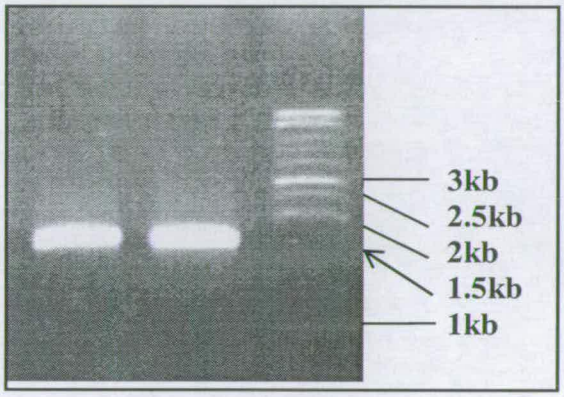


Figure 3.3 Cassette construction: All cassettes were constructed after PCR with pKD4 DNA as template and primers corresponding to regions of the plasmid and each one of the genes of interest. **A)** Cassettes constructed for genes *xapB, nupG, nupC* and *yegT* in lanes 1, 4, 5, 6 respectively. Tracks 2 and 3 are non loaded gel wells **B)** Cassettes constructed for genes *yeiM* and *yeiJ*.

3.2.2 Linear transformation and recombination events

Linear cassette construction was followed by transformation of the linear cassettes into the highly recombinogenic strain of *E. coli*, DY329. The chromosome of this strain contains a defective λ prophage harbouring the recombination genes *exo, bet, gam* under the control of a temperature-sensitive λ cI-repressor (Yu *et al,* 2000). The *exo*, *bet* and *gam* genes are switched on at 42°C and off at 32°C. When the genes are turned on for 5 minutes, cells become more recombinogenic and take up linear DNA without its destruction by intracellular exonucleases.

When the linear DNA cassette is taken up by the cell, homologous recombination occurs between the ends of the cassette that are homologous to the target gene. This leads to a double cross over event, one at either side of the cassette. Consequently, the targeted gene is replaced by the cassette and the chromosomal ORF ceases to exist.

Six mutants were produced after the transformation with the cassettes, each one missing one of the six genes of interest *nupC*, *nupG*, *xapB*, *yeiM*, *yeiJ* and *yegT*. These mutants were named *E. coli* DY329/Δ*xapB*, DY329/Δ*yegT,* DY329/Δ*nupG,* DY329/Δ*yeiM*, DY329/Δ*nupC,* and DY329/Δ*yeiJ* and will be known as MPG585, MPG590, MPG586, MPG587, MPG588 and MPG589 respectively.

Linear transformation was not particularly efficient and only a few transformants were acquired after each electroporation. In particular, only one or two transformants were acquired for all gene replacements except *yeiM* and *nupG*. Transformants were incubated at 37°C for at least 36 hours. In some cases plates were left for up to 48 hours to incubate. It should be noted here that many successful replacement transformants were acquired by leaving the remnants of the recovered cells, after plating out, on the bench over night in the bijou bottles and then plating the whole of the remaining cell suspension onto one selection plate. This gave a lot of successful

recombinants. The reason for that could be the long recovery time that cells were given in the bijou before pressure was put upon them to grow in selection media. Equally a small number of replacement mutants may have arisen, which subsequently replicated in liquid culture.

3.2.3 Verification of correct gene replacement of *nupC*, *nupG*, *xapB*, *yeiM*, *yeiJ* and *yegT* in *E.coli* DY329

Because the recombination events on the chromosome are based on 40bp long homology regions, it is important to check that the cassettes have recombined with the target gene and not elsewhere in the chromosome.

To verify the results PCR reactions were run on each mutant with primers designed to amplify at areas outside the gene to be replaced and consequently deleted. The sequence of the primers used for verification of the replacements are shown in the Materials and Methods chapter in table 2.2. Internal primers specific for the cassette were also used to ensure that the cassette had replaced the gene and that the cassette was intact after recombination with the targeted chromosomal area. Therefore three PCR reactions per replacement were performed; two reactions with locus-specific primers were coupled with the respective cassette specific primer to test for both new junction fragments, and the third reaction was carried out with only the flanking locus-specific primers to verify simultaneous loss of the parental, non mutated fragment and gain of the new mutant–specific fragment. In each case controls were run side by side, using the wild type fragment of non- mutated genes so as to directly compare the size difference of the unmodified and modified DNA.

DNA fragments increased in size after replacement and this aided detection of replacements. The sizes of all fragments for the six genes after PCR are summarised in Table 3.1

Gene name	Wild type NoCo (bp)	Replacement mutants		
		NoK1/K2 (bp)	CoK2/K1 (bp)	NoCo-DY329 (bp)
xapB	1785	784	1169	2043
yeiJ	1964	840	1294	2224
yeiM	1496	784	1108	1982
nupG	1787	1215	799	2113
nupC	1940	1455	727	2272
yegT	1715	1220	694	2004

Table 3.1: Table summarising all DNA fragments expected after verification PCR for gene replacements. Next to the gene name is given the size of PCR fragment from the wild type strain with the external primers (wild type NoCo). The next columns show the sizes for the replacement using the kanamycin resistance gene-specific primers and one external primer as well as the external primers amplified on the replacement mutants. N and C designate the 5' and 3' orientation of the genes respectively.
KEY: K designates a primer for Kanamycin resistance gene used for checking the cassette internally.

Figures 3.4, 3.5, 3.6, 3.7 and 3.8 give an outline of the replacements of the six genes in *E. coli* DY329. All replacements were checked by PCR as described above, and PCR reactions were run on a non-mutated strain of DY329 as a control.

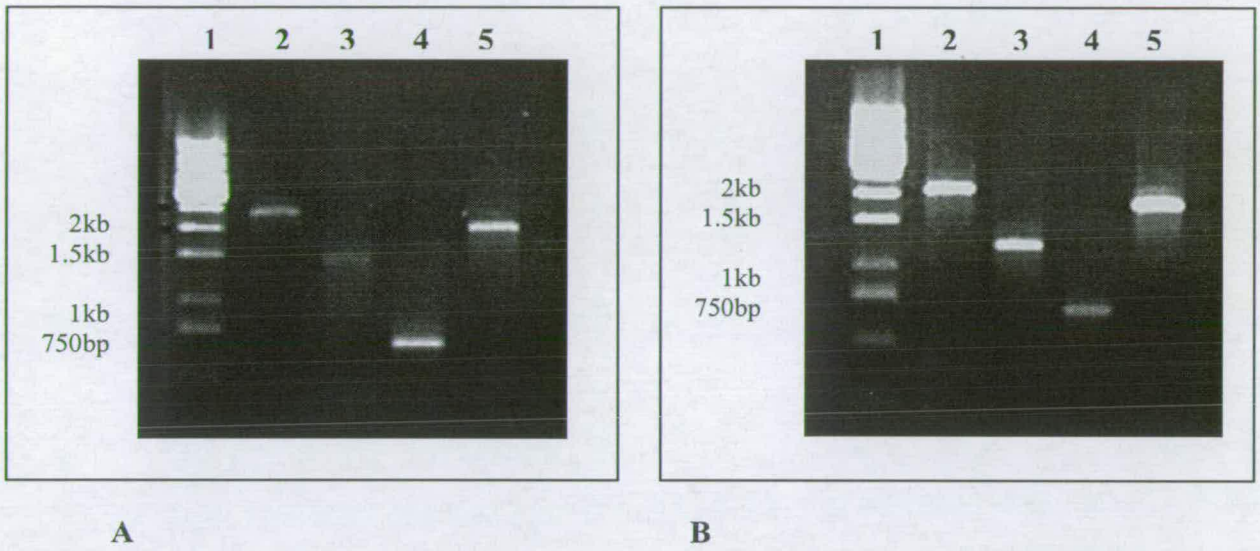


Figure 3.4 A, B
Panel A. Replacement mutant DY329/Δ*nupC* Panel B. Replacement mutant DY329/Δ*yegT*

.

Lane 1 contains a 1kb DNA ladder (Promega). Lanes 2, 3 and 4 represent the mutant DY329 in which *yegT* or *nupC* respectively has been replaced by the kanamycin cassette. Band 2 shows the PCR product resulting from external primers No, Co, whereas bands 3 and 4 are combinations of external and cassette specific primers (CoK2/K1 and NoK1/K2). Lane 5 is a PCR product from the non-mutated (wild type) DY329 gene using external primers.

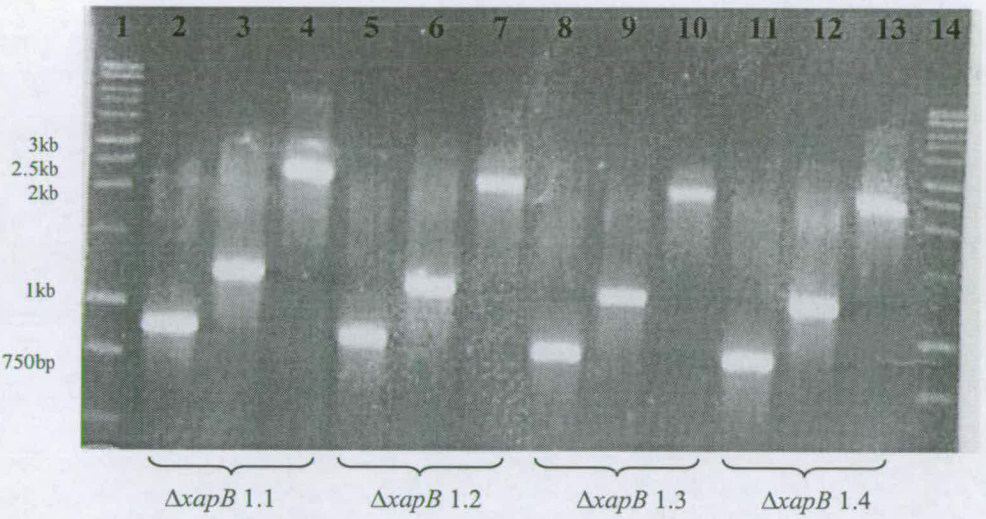


Figure 3.5: Replacements of the *xapB* gene. Four *xapB* mutants are shown above which were checked for replacements at the same time. Lanes 2, 5, 8 and 11 show the replacement fragment for NoK1, lanes 3, 6, 9, 12 the CoK2 and lanes 4, 7, 10, 13 the fragment NoCo. Exact fragment sizes can be found in table 3.1. The non mutated DY329 control is not shown in this gel and lane 1 contains a 1kb ladder.

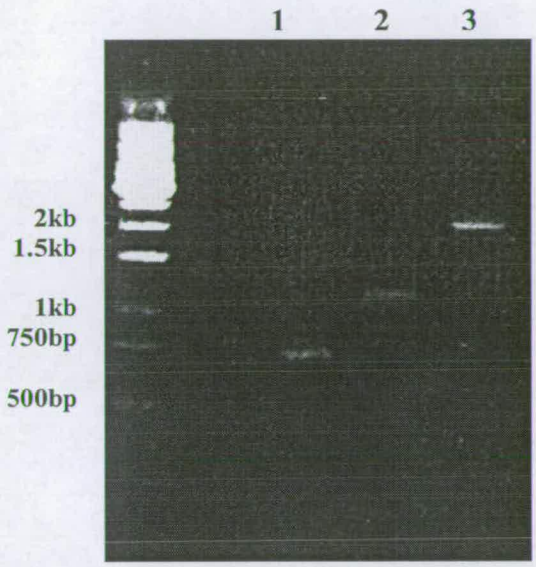


Figure 3.6: Replacement of *yeiM*. Lanes 1 and 2 show the PCR reactions with internal cassette primers and their external pairs. Lane 1 is the PCR fragment with primers NoK1, lane 2 shows CoK2 and lane 3 is from the NoCo, flanking primers outside the *yeiM* locus.

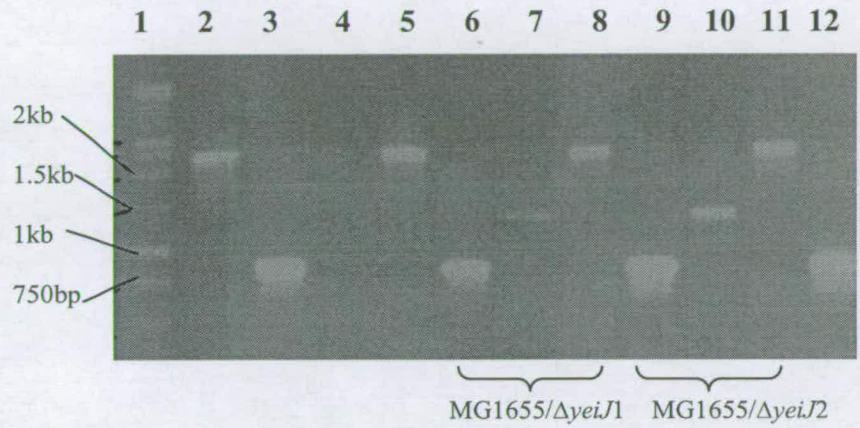


Figure 3.7:
Replacement of gene *yeiJ*. Several mutants were tested for the correct PCR sizes. Of these, two correct mutants for gene *yeiJ* can be seen in bands 6-11. Mutants DY329/Δ*yeiJ*1 and 2 were checked by PCR and lanes 6 and 9 show the PCR product after primers NoK1 were amplified lanes 7 and 10 CoK2 and band 8 and 11 NoCo, the external primers. Exact band sizes can be found in table 3.1. The non mutated DY329 control is not shown in this gel and lane 1 contains a 1kb ladder.

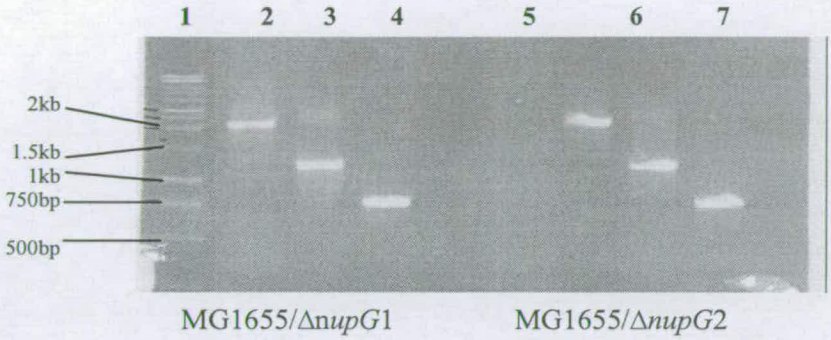


Figure 3.8:
Replacement of gene *nupG*. Mutants DY329/Δ*nupG*1 and DY329/Δ*nupG*2 were checked by PCR for gene replacement. Lanes 2 and 5 show the PCR products after amplification with external primers NoCo, lanes 3 and 6 are after PCR with primers NoK1 and lanes 4 and 7 are PCR products after amplification with CoK2. The non mutated DY329 control is not shown in this gel and the first lane contains a 1kb ladder.

3.3 Gene deletions in the chromosome

Following the step of replacement of genes *nupC*, *nupG*, *xapB*, *yeiM*, *yeiJ* and *yegT* in *E.coli* DY329, gene replacements had to be moved to a single recipient strain of *E. coli* in a cumulative manner.

The sequence of events leading to a complete knockout strain for all six genes was that lysates were prepared with the aid of *E. coli* phage P1 and through transduction, the first replacement was moved to *E. coli* MG1655. MG1655 is a commonly used wild type laboratory strain, lacking the F plasmids. Before the second replacement was added to the recipient strain it was necessary to excise the cassette from the initial transduced replacement, as all the cassettes had the same selectable marker, the kanamycin resistance gene.

The elimination of the antibiotic resistance gene occurred with the aid of a helper plasmid called pCP20 that encodes for the recombinase FLP. FLP acts on the FRT sites, which flank the antibiotic resistance gene, and results in deletion of the antibiotic resistance gene via FRT recombination, thereby leaving a single FRT site in its place (Cherepavov and Wachernagel, 1995).

After the elimination of the resistance gene and deletion of the gene the second replacement can be transduced into MG1655:: FRT recipient strain. Following this same sequence of events six monogene mutant strains were constructed, each one lacking five transporter genes and retaining only one of them intact. Furthermore a complete knock out "null" strain was also constructed which was deficient for all six genes and in whose places there now were FRT scars. It should be mentioned at this point that mutant strains that required both *yeiM* and *yeiJ* to be deleted from the chromosome (these strains include the "null" strain and all monogenes strains except the ones that had either *yeiM* or *yeiJ* intact) still had the kanamycin cassette during the replacement step, and they were therefore not completely deleted from the

chromosome. The reason for this is explained further in section 3.3.3 and in the Discussion section of the chapter.

3.3.1 P1 transduction and antibiotic resistance gene elimination

In order to construct multimutants P1 lysates were prepared from each one of the DY329/Δ strains, using P1 phage, prior to removal of the kanamycin resistance genes as described in the Materials and Methods chapter. The lysates were then used to infect the recipient strain *E. coli* MG1655. Through recombination events between the transduced replacement DNA fragments and the chromosome of MG1655 or derivatives carrying deletions, the targeted gene was replaced by the transduced element. Effectively the replacement was moved from strain DY329 to strain MG1655 or a derivative of it.

Transductants occurring after P1 transduction were checked by selecting for the acquisition of kanamycin antibiotic resistance. Transduction efficiency varies with P1 over a 25-fold range due to fluctuations in the recombination system (Abrahams *et al,* 1994). Fuller details of transduction processes are described in the Materials and Methods chapter.

Before transducing the next cassette for gene replacement into MG1655 or a derivative of it, the strain was transformed with the helper plasmid pCP20, which caused the excision of the antibiotic resistance gene. The helper plasmid was then cured by growing the strains at 42°C. This allowed for the cyclical repetition of the process with the next transduction using DNA fragments that carried a single antibiotic resistance cassette. Removal of the antibiotic resistance gene allows the multiple use of the same cassette, which solves the problem of utilisation of different selection markers if multiple genes are to be deleted. Furthermore this strategy clears

the chromosome of unwanted and unnecessary DNA fragments, which may interfere with the expression of downstream genes, which may belong to the same operon as the targeted genes and are under the control of the same promoter.

FRT sites are direct repeats of DNA flanking the kanamycin resistance gene and they are the only sequence requirement for Flp recombinase to act. Flp is encoded by the helper plasmid pCP20 and requires no accessory proteins to function (Chen, *et al,* 2000). When FRT sites are facing the same way then a general recombination between the two will cause deletion of the intervening DNA sequence. If they have opposite orientations then inversion of the intervening DNA fragment will occur.

pCP20 contains both an ampicillin and a chloramphenicol resistance gene and shows temperature-sensitive replication (inhibition at high temperatures). Induction of FLP synthesis by pCP20 occurred after growth of the strain at 30°C. Transformation of the recipient strain MG1655 which contained the replacement fragment was facilitated by heat shock and then growth over night at 30°C in LB media containing the appropriate antibiotic, in this case ampicillin for plasmid maintenance, thus resulted in the excision of the cassette.

To subsequently eliminate the plasmid following cassette removal, transformants were colony purified at 42°C overnight, a temperature that made the loss of the helper pCP20 plasmid possible. After that, transformants were tested on ampicillin and kanamycin. If the cassette's antibiotic resistance gene was eliminated and the helper plasmid cured, mutants should be sensitive to both of these.

Incubation of transformants at 42°C also validated that the mutants were indeed derivatives of strain MG1655 and not DY329. Growth of DY329 at 42°C would result in the death of the bacterial strain through induction of lysis by phage λ.

The efficiency of pCP20-mediated transformation was very high. However there were cases in which the transformants were not what was expected. In some cases the transformants were shown to still retain kanamycin resistance, which indicated that pCP20 did not produce functional FLP. In other cases the transformants were ampicillin resistant, even after incubation of the transformants at 42°C. This could mean that a mutation had occurred to the plasmid which made it lose its temperature sensitive replication characteristic.

3.3.2 Knockout strains and how they were designed

The aim of the experiment was to construct a strain of *E. coli* MG1655 that had the genes *nupC*, *nupG*, *xapB*, *yeiM*, *yeiJ* and *yegT* deleted from its chromosome. As described above by recursive replacements and deletion, each gene was deleted one by one. In order to accumulate all deletions in the chromosome it was necessary to plan the way this would take place so as to maximise progress and avoid unnecessary deletions that would be time consuming.

All genes were first replaced by the cassettes in stain DY329 and subsequently replacements were transduced to MG1655. Therefore six single mutants were obtained which were named MG1655/Δ*nupC*, MG1655/Δ*nupG*, MG1655/Δ*xapB*, MG1655/Δ*yeiM*, MG1655/Δ*yeiJ* and MG1655/Δ*yegT*. Three of them, MG1655/Δ*xapB,* and MG1655/Δ*yegT* and MG1655/Δ*nupC*, were used as the basis of a sequence of subsequent transductions and deletions, which led to the final mutant lacking all six genes.

Another experimental aim was to construct mutants with five of the genes missing, and only one functional gene present, the monogene mutants. The final aim was to construct an absolute deletion mutant with all six genes knocked out.

For reasons of convenience and due to the large number of deletions accumulating in these strains, the nomenclature follows the rule that the letters that follow 'Δ' signify the order of deletions in that strain. Strain MG1655/ΔBTG for example is a strain in which genes *xapB, yegT* and *nupG* have been deleted in that order. To signify that the kanamycin resistance has not yet been removed from a particular locus, the symbol ":: Kan" is used after the letter. For example strain MG1655/ΔBTGMCJ::Kan is a strain in which genes *xapB*, *yegT*, *nupG*, *yeiM* and *nupC* have been deleted in that order and gene *yeiJ* has been replaced by the kanamycin cassette.

The sequence of transductions and deletions, which led to the construction of the strains in this project, is outlined in figure 3.9. Each letter represents the respective gene (B for *xapB*, T for *yegT*, C for *nupC*, G for *nupG*, J for *yeiJ* and M for *yeiM*). Each letter that is added after the other signifies an added replacement which is transduced through P1 transduction into the MG1655 chromosome. After elimination of the kanamycin resistance gene, the following replacement is added through P1 transduction and so on.

Every replacement that was added through P1 transduction to the recipient strain was checked before and after cassette excision, by PCR. Routinely in multi mutant strains all previous deletions were checked after cassette excision. This also ensured no recombination events occurred between FRT scars that could lead to chromosomal areas being deleted (data not shown). More information on cases where such events have occurred can be found in the discussion of this chapter.

When the monogene strains were constructed it was possible to obtain the complete knockout strain from any of them. However the complete knockout strain finally was constructed after *yeiJ* was added in the strain MG1655/ΔBTGMC. In this null strain, gene *yeiJ* still harbours the kanamycin cassette.

The monogene strains obtained were MG1655/ΔBTGMC, MG1655/ΔBTGMJ, MG1655/ΔBCGMJ, MG1655/ΔTCGMJ, MG1655/ΔTCBJM, and MG1655/ΔTCBJG. These strains were renamed MPG583, MPG582, MPG579, MPG578, MPG580, MPG581 respectively for convenience. The null strain MG1655/ΔBTGMCJ::kan is known as MPG584. (See Appendix 1 for analysis of individual monogene strains)

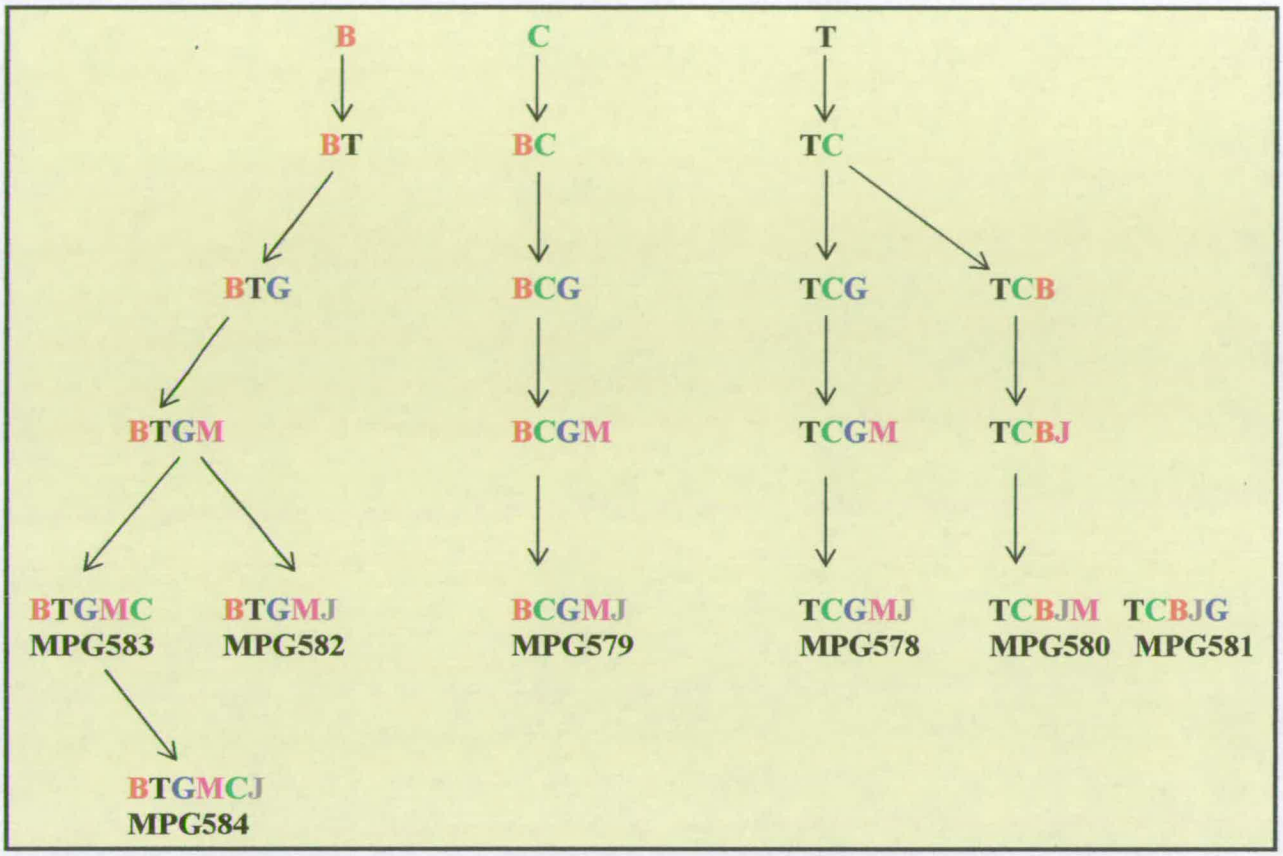


Figure 3.9: Sequence of mutations that led to the monogene strains and null mutant. This rough diagram shows how the monogene strains were acquired after sequential deletions from single mutants leading to the null. Each letter represents a gene as explained above. Every letter added is a deletion added to the *E. coli* mutant. The result is six monogene strains and the null strain. The colour code is used to emphasise the sequential order by which deletions were accumulated the multi mutants. (See Appendix 1, for a representation of all monogene strains). Note that strains BTGMJ, BCGMJ, TCGMJ, TCBJM bear the kanamycin resistance cassette in the last gene transduced, i.e. *yeiJ* for strains BTGMJ, BCGMJ, TCGMJ and *yeiM* for strain TCBJM.
KEY: B (*xapB*), T (*yegT*), G (*nupG*), C (*nupC*), M (*yeiM*), J (*yeiJ*)

3.3.3 Results of gene deletions

The following figures show some of the gene deletions that were checked by PCR as examples in the mutant strain MG1655/ΔBTGMCJ::Kan. Table 3.2 summarizes the sizes of all PCR fragments that were expected before and after the kanamycin gene was excised leaving an FRT scar in the gene's place.

Gene name	Cassette insertion NoCo (bp)	Scar (after gene deletion) (bp)
xapB	2043	684
yeiJ	2224	733
yeiM	1982	570
nupG	2113	568
nupC	2272	600
yegT	2004	544

Table 3.2: Summary of PCR fragments of genes after deletion of the ORF and following the removal of the cassette. This leaves "scar" areas, the sizes of which are shown in column 2. Primer positions are summarised in Chapter 2, Materials and Methods table 2.2

The resulting null mutant was an MG1655 *E. coli* strain in which all six genes were missing (MPG584). Genes *xapB, yegT, nupG, yeiM, nupC* were deleted with a FRT scar in their place, while the last gene to be replaced was *yeiJ* and this was left with the kanamycin cassette intact. This was because attempts to excise the cassette from gene *yeiJ* were not fruitful due to the very low efficiency by which pCP20 would produce functional Flp and the events that would normally follow pCP20 transformation gave no positive result on cassette excision. The mutant's verification results are shown in figure 3.10.

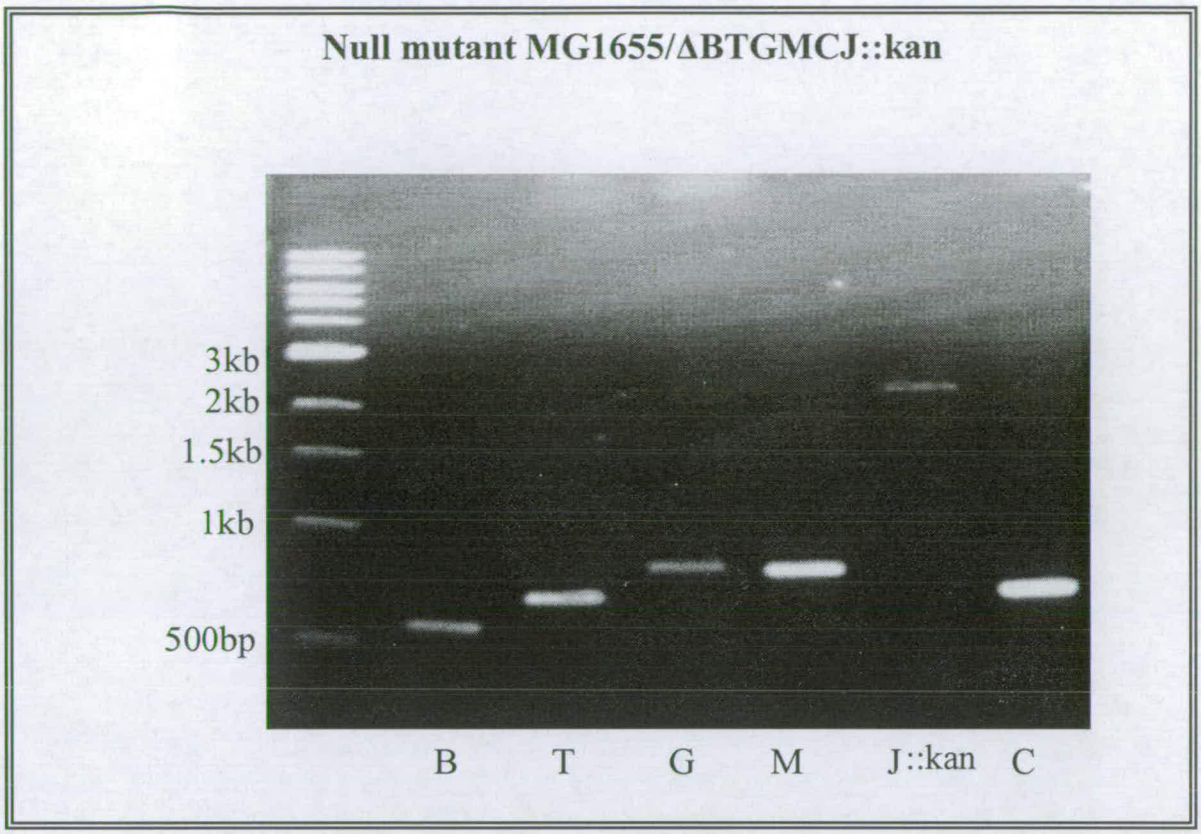


Figure 3.10: The null mutant. In strain MG1655/Δ BTGMCJ::kan genes *xapB, yegT, nupG, yeiM, nupC* have been deleted in that order while the last gene, *yeiJ* was replaced by the kanamycin cassette. Evidence from PCR for each one of the deletions is shown above in the gel using flanking primers (Table 2.2 Chapter 2). The PCR product for *yeiJ* in the fifth lane is large as it harbours the antibiotic resistance gene. All PCR fragment sizes can be found in tables 3.1 and 3.2. The positions of the sizes of markes (1kb marker by Promega) are shown on the left of the diagram.

3.4 Discussion

The *E. coli* strain DY329 was used for the purposes of this work as it has the characteristic of being able to be transformed with linear pieces of DNA. Strain DY329 was constructed by Yu *et al* (2000) and it contains a λ prophage harbouring the recombination genes *exo, bet* and *gam* under the control of a temperature sensitive λ cI-repressor. The genes *exo, bet* and *gam* can be switched on at 42˚C and off at temperatures below 32˚C. When the λ functions are turned on for as short a time as 5 minutes, cells become more recombinogenic and any linear DNA which is taken up at this time persists without its destruction (Yu *et al,* 2000). In particular Gam inhibits the RecBCD nuclease from attacking linear DNA, and Exo and Beta generate recombination activity for the linear DNA (Yu *et al,* 2000). Recombination is effective with DNA homologies as short as 30-50 bp on the ends of linear DNA substrates.

Yu *et al* (2000) also determined optimal recombination methods by examining factors such as induction time needed for the recombination system to be effective, concentration of linear DNA needed for generation of recombinants and recombinant number generation as affected by the homology length.

What they found was that a maximum efficiency of recombinants is achieved when the induction of the recombination systems occurs at a period of between 7.5 to 17.5 minutes, while some reduction occurs if induction continues for longer than that. Expression of the P_L operon for longer than 60 minutes causes cell death. The optimal linear DNA concentration was found to be in the range of 1ng to 100 ng per electroporation after which point a saturation level is reached, in terms of recombinants recovered occurs. Finally it was determined that no recombinants were recovered with 10 bp of DNA homology, and less than ten recombinants with 20 bp of homology. Homologous recombination and thus appearance of recombinants,

increased by four orders of magnitude when homology was between 20 to 40 bp. For 40 to 1000 bp of homology, recombination increased 10-fold.

The optimal conditions that were suggested by Yu *et al* (2000) were used as guideline in the design of this set of experiments. Specific details on quantities and primer design can be found in the Materials and Methods chapter 2.

Also use of plasmids for expression of the λ functions means that it is not necessary to waste one type of drug resistance on maintaining the plasmid in the cell. Indeed in preliminary side-by-side experiments performed in this work with linear DNA transformed into cells which were expressing the red and gam functions from a plasmid (pKD46, from Datsenko and Wanner, 2000) as opposed to cells that were expressing the red and gam functions from a prophage in the chromosome of DY329, from Yu *et al,* 2000) it was confirmed that linear transformations were successful when the latter was being used (data not shown). However it is reported that based on their plasmid expressing red gam functions Datsenko and Wanner (2000) were able to make more than 40 different gene disruptions.

Linear transformation is not a very efficient method. As compared to transformation with circular plasmids, the efficiency is significantly lower. After transformation only a few clones are produced, making this an ineffective method for many purposes. However it is very precise and when it works, it virtually guarantees the desired resulting mutation. Since the first publications describing linear transformation and PCR-mediated gene replacements, the method has been improved in that researchers are now using plasmids, which bear the cassette that will replace the gene. Therefore transformation efficiencies are much higher as well as precise.

In particular, in this study, overall linear recombination was not a particularly efficient method. After transformation with the linear DNA cassette, recombination

at the pre- designed recombination sites was taking place, however in most transformation events the efficiency was as low as acquiring one or two correct recombinant transformants.

As compared to the transformation efficiencies of the same recipient strain transformed with a positive control plasmid (puc4K) this was truly low (data not shown). Transformation efficiencies of the positive control were as high as 1.46 x 10^7 CFU/ml whereas transformation efficiency of the recipient strain with the linear DNA cassette was at 6 CFU/ml.

Reassuringly enough though, the few transformants acquired after linear recombination events were all found to be correct, that is the cassette had integrated correctly at the targeted gene area for replacement.

Although the method of linear DNA transformations worked for the purpose of this project and gave the desired results, it was a process that was costly in terms of time. It was unreliable in that the frequency of occurrence of recombinants was unpredictable, although the method was followed consistently in each experiment. Since gene replacements are more efficient by plasmid-mediated transformations this would be the preferred approach for future studies (Merlin *et al,* 2002a)

An important aspect for successful transduction is the lysate titre. In order to give rise to transductants the lysate titre has to be at least 10^9PFU/ml (Masters, 1977). Thus for these experiments the titre was always kept to a minimum of 10^{10} PFU/ml. In order to get such a high titre, lysates were prepared according to the protocol included in the Materials and Methods chapter. It should be emphasised at this point that there is a requirement of Ca^{+2} for P1 stability. Lack of calcium ions results in the collapse of the viral microskeleton. It is therefore vital for the success of this method to always include calcium ions in the media and in all solutions until infection is complete (Sambrook *et al,* 1989).

The transductant survives the original phage infection because it does not receive any viral DNA from the viral shell which carries the linear DNA but it will not survive a subsequent or simultaneous P1 infection unless a lysogen forms. Because 99.9% of the P1 virions in any given lysate are functional virulent phages, as soon as the cell is lysed large quantities of P1 are released in the culture. These P1 particles could infect the transductants and lyse them. However if calcium ions are removed from the medium before any cells lyse, but after transducing particles have attached, the transductants are protected against superinfection by P1 virions.

P1 can pack about 100 kb of DNA making it one of the most efficient carriers of lage segments of DNA. This guarantees that flanking areas upstream and downstream of the target gene would be transduced via P1, which were big enough so as to ensure precise recombination. However the negative aspect of this characteristic became apparent in the present study when genes of close proximity were to be transduced separately. In particular, genes *yeiM* and *yeiJ* were only 1940 bp apart. After *yeiM* was successfully transduced into MG1655, the antibiotic resistance cassette was eliminated with the aid of pCP20. Following that, *yeiJ* was transduced into the same strain. This resulted in the successful replacement of *yeiJ* with the transducing element but due to the genes' close proximity and high recombination likelihood, it also led to the inevitable restoration of neighbouring Δ*yeiM* with wild type *yeiM*. This meant that although the *yeiJ* cassette now replaced *yeiJ*, *yeiM*, which was knocked out previously and replaced with a single FRT scar, was recombining with the DNA transported by P1 and being restored to wild type. Figure 3.11 (Masters 1987) shows how cotransduction is highly dependent upon the distance between the genes to be transduced. As seen from the graph the probability of cotransduction between *yeiM* and *yeiJ* is 98%. The same situation was a risk for genes *xapB* and *nupC* which are only 0.21 minutes apart. According to the graph 3.11 this gives them an 80% chance of co transduction (Newman and Masters, 1980).

The gene proximity of genes *yeiM* and *yeiJ* was challenging since the two genes were very close and were co-transduced causing restoration of the wild type genotype in the recipient strain for the first gene when the second of the two genes was transduced into the MG1655 strain. However after a number of repetitions of the experiment some successful transductants that would have one of the genes deleted and the second close by gene replaced were obtained and the experiment could proceed.

Due to this incident coupled with the fact that the cassette excision helper plasmid was losing its efficiency some of the monogene strains that had to have both *yeiM* and *yeiJ* deleted were left with one of the genes deleted (that is the cassette removed) and the other gene simply replaced (without cassette excision).

This is not causing a problem however since it is believed that due to gene orientation the genes do not belong to the same operon and therefore polar effects would not be an issue.

In this set of experiments, repetition of the transduction a few times and screening of a larger number of colonies led to the desirable result. Thus, it was possible to acquire a mutant that had genes of close proximity, such as *yeiM* and *yeiJ* knocked out in one strain.

Problems such as the one described above are very common and not always overcome by simply repeating the experiment a few times. A way to resolve this problem is to initially study the genes that are to be deleted well and if some of them are found to be very close to each other, then in the beginning of the methodology design to be followed, it is advisable to construct cassettes with different antibiotic resistances. This way the first gene can be replaced by the antibiotic resistance but it will not be necessary to excise the resistance gene because the second gene to be deleted will carry a different antibiotic resistance marker. Thus, it becomes possible

to select for both antibiotic resistance genes, ensuring that both genes are replaced by the resistance cassettes. This should not be a problem if the genes belong to different operons and expression of downstream genes is not affected by the presence of the antibiotic resistance cassette. In the current study the antibiotic resistance cassette was left on the chromosome in the final constructs. However since no other genes appear to be downstream from *yeiJ* in this operon no adverse effects were anticipated (http://genolist.Pasteur.fr/Colibri).

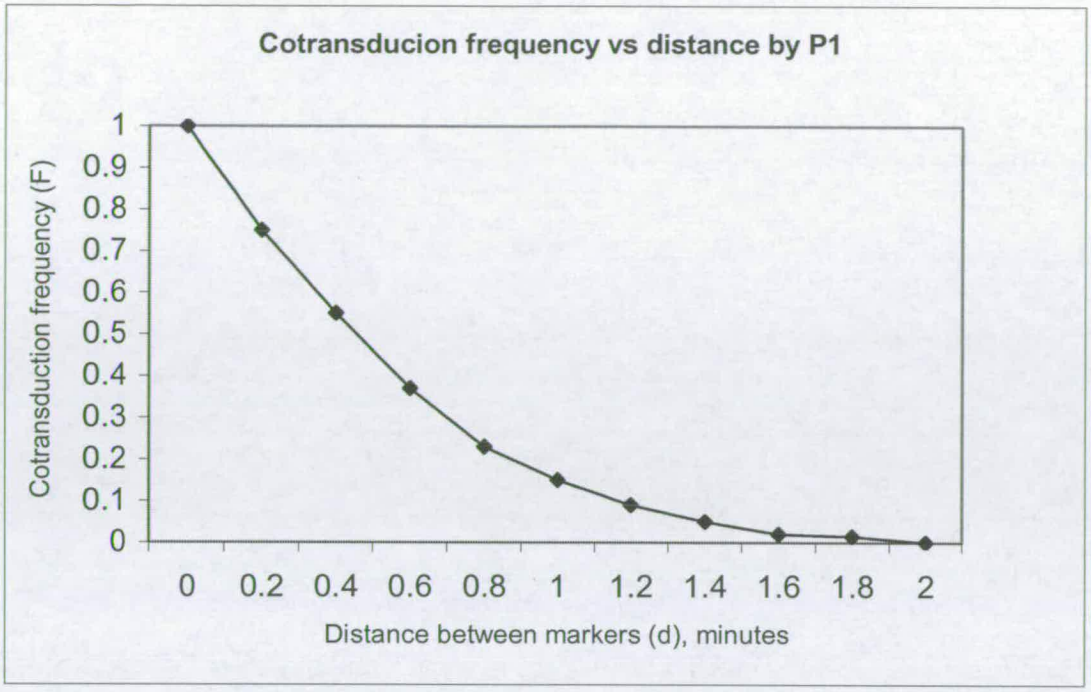


Figure 3.11 Reproduced from Masters (1987): **Cotransduction frequencies in relation to distance between genetic markers:** Cotransduction frequencies (F) are plotted as a function of the distance between two markers according to the formula: $F= (1-d/L)^3$ derived by Wu (1966). The constant *L* represents the length, in map minutes, of the packaged DNA packaged by P1 estimated at 2.1min. From the graph it is shown that the closer two genes are to each other (i.e. the smaller their distance), the higher the cotransduction efficiency. The distance between *yeiM* and *yeiJ* is 0.07 minutes; therefore the cotransduction frequency is almost 98%. That means that there is only a 2% possibility for P1 to pack DNA from these two areas and transduce those two genes separately.

One of the issues raised in the processes leading to gene deletions was that of co transduction and reversion of deleted loci back to wild type genotypes. It is important to emphasise that DNA areas that are considered as high risk of relapsing into wild type genotypes should be checked by PCR, amplified with the flanking primers after every replacement is added to the recipient strain. An example of such screenings of colonies and checks on previous deletions are shown in the figures below.

As seen in figure 3.12, the double mutant ΔBC was checked in the *xapB* locus to ensure that the transduction of the second mutation at the locus of *nupC* via P1 did not cause the "scar" area of the initial *xapB* mutation to revert back to wild type. As can be seen from the gel, seven transductants were checked by PCR, ideally expecting the *xapB* area to be deleted and a PCR fragment the size of the scar to appear. However only two of the mutants tested had *xapB* deleted. In the rest of the mutants, in lanes 1, 3, 5, 6 and 7 the wild type locus has been restored.

Similarly double mutants MG1655/ΔTB and MG1655/ΔBT were tested for deletion of both *xapB* and *yegT*. In mutant MG1655/ΔTB both genes were found to be deleted, however in mutant MG1655/ΔBT while *xapB* was deleted, *yegT* had reverted to wild type (Figure 3.13). This event reflects a weakness of the method and emphasises the importance of validation that no reversion has occurred when constructing a multi mutant by transduction.

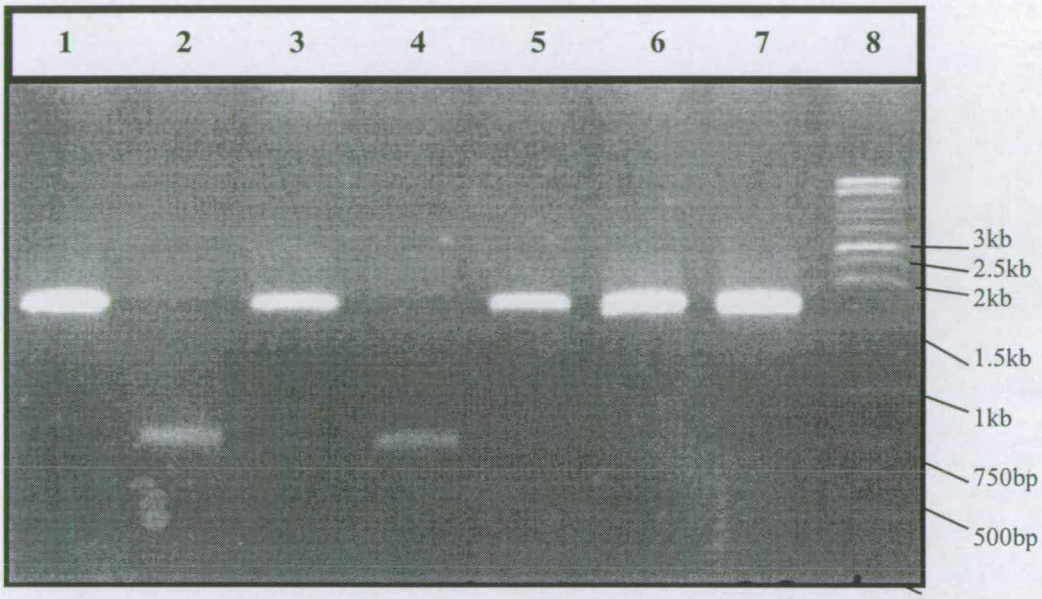


Figure 3.12: **Screening for reversion to wild type of neighbouring genes**. Several colonies were picked and screened through PCR using flanking primers Co and No as described before and summarised in Chapter 2, Table 2.2 for restoration of wild type. From the seven colonies tested by PCR, the PCR products revealed that 5 mutants had the *xapB* operon reverted to wild type (lanes 1, 3, 5, 6, 7 size of 1785) while only two mutants had the *xapB* operon deleted (lanes 2 and 4, size of 684).

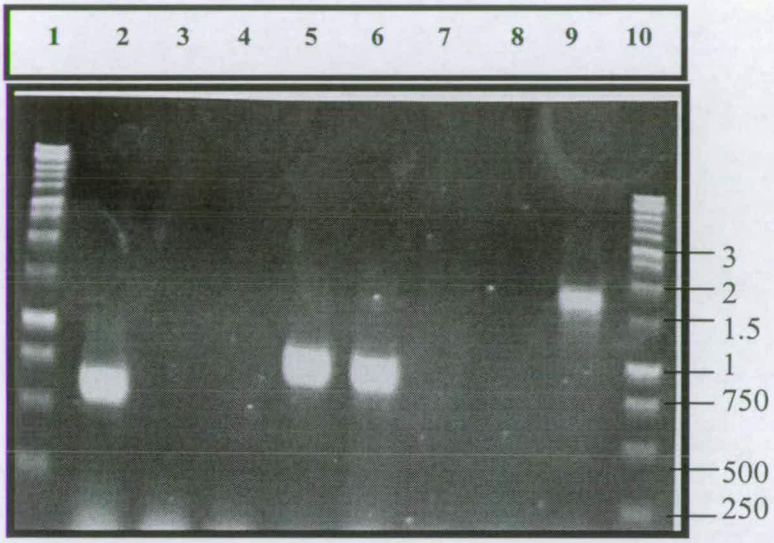


Fig 3.13 Double mutants ΔTB and ΔBT, showing a case of normal gene deletions and a case of reversion to WT. The first and last lanes have a 1kb ladder. The four lanes 2-5(mutant 1) show deletions at *yegT* and *xapB* loci, and the next four lanes (mutant 2) show deletion of the *xapB* locus and reversion of *yegT* locus to WT. Exact sizes of the wild type and deletion DNA bands can be found in tables 3.1 and 3.2.

Verification of gene replacements and subsequent gene deletions was carried out by PCR on the flanking gene regions and looking for DNA fragments of a size that would indicate gene replacement by a cassette of larger size or gene deletion when the PCR fragment was smaller than the gene size meaning that the gene had been deleted and a scar area was left in its place.

As a whole the linear transformation method was successful. However recent literature suggests more efficient methods that can be used (Merlin *et al,* 2002b). As suggested above, plasmid mediated transformation, with a cloning step can be used to construct gene deletions just as precise as the ones performed in this work. However this can be more costly in terms of materials needed and the time required for plasmid construction.

The advantage of the mutants produced in this study as compared to other nucleoside transporter mutants that have appeared in the literature before, is that the genes were deleted completely from the chromosome. Previous work largely depended on poorly characterised point mutations (Link *et al,* 1997, Jørgensen and Dandanell, 1999). Furthermore, deletions of putative genes such as *yeiM, yeiJ* and *yegT* now enable the characterisation of these genes, and also clarify the basis for previously published background problems on nucleoside transport experiments which were attributed to the putative existence of other nucleoside transporter genes on the chromosome of *E. coli.* (Jørgensen and Dandanell, 1999).

The general trend in functional characterisation of unknown genes is to delete their coding sequence and observe how the resulting mutants behave in their environments. The methods used are generally classified in two categories. The first class of methods used DNA circles, usually conditionally replicating plasmids that are not affected by

exonuclease digestion. The plasmids carrying DNA flanking the target sequence are forced to integrate into one of the homologous regions by selection for retention of the antibiotic resistance marker on the plasmid under restrictive conditions. However long regions of homology are required, about 400bp, in order to achieve good rates of recombination using host enzymes. Surviving bacteria allow cointegrate resolution. Selection against the plasmid favours survivors in which allele exchange has occurred and they are identified by screening. A variant of this method can also used which utilises a suicide plasmid carrying the target gene mutation and a unique recognition site for the meganuclease I-*Sce*I. When integrated in the chromosome the plasmid is flanked by the wild type target locus and the constructed replacement that includes DNA homologous to the target. When I-*Sce*I is then provided it results in a double strand cut which stimulates recombination between the alleles and the engineered version can be selected for. Such methods have at least one cloning step and may require a second one if a selective marker is used and they therefore require more time. However a large replacement can be made and since each step can be monitored for success and the number of successful replacements is high, it is easy to judge whether failures have a technical or biological basis.

The second method employs linear DNA which can be exchanged by a double crossover event to replace the chromosomal target with engineered DNA in a single step. The most recent variant of this type of technique is the one used in this study and it essentially prevents the need for cloning altogether by using primers of no more than 50bp of homology flanking the target gene to amplify the selectable marker. As described above the linear product is then introduced into the host strain expressing the plasmid -borne lambda Red-Gam recombination system. This method when successful is very fast since no cloning is required. However the length of sequence to be replaced is limited and good rates of recombination and replacement are not easy to achieve.

In conclusion, the mutants constructed in this work, due to their precise genotypic profile, provided ideal candidates in order to perform further experiments and to attempt clearer characterisation of the unknown or partially characterised genes. Moreover the null strain if shown to be a nucleoside negative strain could have potential to be used as a first screen to check the merits of potential nucleoside analogues that might be used in chemotherapy.

The benefits of this technique are therefore its accuracy rather its high efficiency. This is arguably a factor that could influence other researchers to favour this technical approach when gene deletions are necessary for a study. The ability to modify or subclone large fragments of genomic DNA with precision should facilitate many kinds of genomic experiments that were difficult or impossible to perform previously and aid in studies of gene function in the postgenomic era (Lee *et al,* 2001).

Nevertheless other studies have reported higher transformation efficiencies as well as good results on accuracy of gene replacements by employing conditionally replicating plasmids (which are immune to exonuclease digestion) as delivery systems for recombination events to occur and to promote gene replacements by cassettes. These techniques as described by Link *et al* (1997) require longer regions of homology (about 400bp) in order to achieve good rates of recombination using host enzymes.

Chapter 4

Characterisation of *E. coli* putative nucleoside transporters

4.1 Introduction

Genes *nupC, nupG* have been shown to have a central role in active nucleoside transport across the cytoplasmic membrane of *E. coli* (Munch-Petersen *et al,* 1979, Loewen *et al,* 2003, Xie *et al,* 2004). Also *xapB* has been studied to some extent as a xanthosine transporter (Nørholm and Dandanell, 2001). With the construction of mutant strains of *E. coli* lacking unknown or putative nucleoside transporters *xapB, yeiM, yeiJ* and *yegT*, it seemed possible to characterise these genes through assessment of the effect of mutations in these genes and to establish their involvement in nucleoside transport in *E. coli* via various experimental paths.

Several strategies for analysis were explored. One experimental path included experiments in which the genes were cloned into multi-copy plasmids and transformed into suitable *E. coli* strains, which allows over expression of the cloned gene product. These strains were then used in transport experiments involving radiolabelled nucleosides.

To measure transport activity of proteins encoded by genes *xapB*, *yegT, yeiM* and *yeiJ* and retrieve information on their functions as possible nucleoside transporters, experiments were based on uptake of ^{3}H-radiolabeled adenosine. The activity was measured initially through a time course to check that the over-produced gene product increased uptake of the radiolabelled nucleoside, as compared to background levels of transport.

Concentration-dependent studies were then performed to determine kinetic characteristics of transport for each of the genes. These findings reflect the gene products' ability to transport the radiolabelled substrate.

Furthermore to explore the transporters' specificity, competition experiments were performed in order to investigate whether they have a preference for transporting the

radiolabelled substrate provided or if this uptake is inhibited by other competing nucleosides.

Additionally the growth of monogene mutants, which had only one of the genes intact, and the other five genes deleted from the bacterial chromosome were examined. Cells were grown in minimal media supplemented with physiological nucleosides as the sole carbon source and growth of the bacteria was measured in 96 well plates through OD. This set of experiments gave information on the bacteria's ability to utilise nucleosides as their main carbon-based growth substrate under limiting growth conditions.

The experiments described above give an idea of the function of these genes and suggest that nucleoside transport is the main function of these genes. Below, a description and analysis of the experiments is given, and conclusions attained from the findings.

4.2 Construction of plasmids

Genes *xapB, yegT, yeiM* and *yeiJ* were cloned into the high copy number plasmid pBAD18. Plasmid pBAD18 is a vector commonly used in cloning as it has a multi-cloning site. The vector contains the P_{BAD} promoter of the arabinose operon and its regulatory gene, *araC*. Expression from the P_{BAD} promoter is induced up to 1,000-fold when arabinose is present in the growth medium.

Initially the genes to be cloned were amplified from the chromosomal DNA by PCR, using Taq as the polymerase enzyme so as to get copies of the chromosomal sequence as PCR products. Taq-mediated PCR amplification often results in the addition of deoxyadenosine in a template independent fashion, to the 3' ends of the amplified fragments. Details of the amplification programs, primers and other reagents used can be found in Chapter 2. The amplified sequence was designed in such a way as to include upstream regions of the gene, ensuring that features such as the natural ribosome binding site (Shine-Dalgarno sequence) would be cloned into the host vector as well as the ORF of the gene to optimise the likelihood that the gene would be expressed after plasmid induction. Additionally the PCR products were engineered with restriction sites included which would be used subsequently as sticky ends to join the gene sequences into the expression vector pBAD18. The restriction enzyme pairs engineered were *Hind*III and *Eco*RI for genes *xapB* and *yegT,* and *Sal*I and *Kpn*I for genes *yeiM* and *yeiJ*.

The addition of the extra deoxyadenosine bases on the PCR product facilitated ligation into the helper vector pGEM-T, as an intermediate cloning step. The commercial vector pGEM-T is prepared by cutting with *Eco*RV and adding a 3' terminal thymidine to both ends. These single 3' T- overhangs at the insertion site greatly improve the efficiency of ligation of a PCR product into the plasmids by preventing recircularisation of the vector and providing a compatible overhang for PCR products prepared with polymerases such as Taq, which as mentioned above, add a 3' adenosine.

The cloning site of pGEM-T lies within the coding region of the β-galactosidase α-peptide so that if a fragment is inserted which disrupts the open reading frame, the *lacZα* gene is inactivated. Therefore in *E. coli* strains lacking *lacZ* but carrying constitutively expressed lacZβ, recombinant clones can be identified by screening for β-galactosidase activity. The screen involves selection of plasmid transformant *E. coli* cells on LB medium containing 50μg ml^{-1} carbenicillin. Screening is provided by the addition of 40 μl of 2% X-gal solution (w/v, in dimethylformamide) and 60μl of a 2% IPTG solution (w/v) to the LB agar medium in a Petri dish. After overnight incubation of the plates at 37°C, colonies harbouring pGEM-T appear as blue colonies, whilst cells bearing these plasmids with inserts appear white.

The pGEM-T kit includes a ligation buffer for ligation of PCR products. Reactions using this buffer are performed at 4°C, and will produce the maximum number of transformants if incubated overnight.

Plasmids from white colonies were purified using a Promega DNA extraction Kit. Purified plasmids were then digested with suitable restriction enzymes as described in the Materials and Methods chapter. The restriction enzymes sites were unique in pBAD18 so as to ensure that the fragment could only be cloned in one direction from pGEM-T. It was also ensured that the restriction sites were not part of the actual ORF to avoid unwanted digestion of the PCR product. The product of the digestion reaction gave fragments of DNA that contained the ORF of the genes, including the upstream Shine-Dalgarno sequence, flanked by sticky ends. Cloned fragments were checked by sequencing for correct fragment orientation in pGEM-T and for possible mutations.

In order to subclone from pGEM-T plasmid pBAD18 was digested with the appropriate combination of restriction enzymes as described in the Materials and Methods chapter. In particular the pairs of enzymes used were *Hind*III and *Eco*RI for genes *xapB* and *yegT* and restriction enzymes S*al*I and *Kpn*I for genes *yeiM* and *yeiJ*. The freshly digested vectors were then treated with SEAP in order to dephosphorylate them so as to avoid re-

ligation of the digested vectors to themselves during the ligation reaction. Gene sequences cloned in pGEM-T were also digested with the same combinations of restriction enzymes mentioned above.

The ligation reaction that followed between the digested vector and the gene fragments gave rise to plasmids now containing the desired gene ORFs under the controllable *ara* promoter of the pBAD18. The resulting plasmids were sequenced to determine that the correct sequence was encoded, and samples were analysed by the sequencing service within the ICMB or ICAPB, in the University of Edinburgh.

The plasmids were subsequently transformed into host strain BLR which allows the over production of proteins due to its defect in proteases *lon* and *ompT*. Plasmids were extracted for a final time and digested with the restriction enzyme pairs *Hind*III and *EcoR*I for genes *xapB* and *yegT* and *Sal*I and *Kpn*I for genes *yeiM* and *yeiJ* in order to check that the insertion of the fragments was correct. In figure 4.1 the digested products from all four plasmids when checked for correct insertion of the PCR plasmids are shown. The BLR strains containing plasmids pMRB (overexpressing *xapB*), pMRT(overexpressing *yegT*), pMTM (overexpressing *yeiM*), and pMRJ (overexpressing *yeiJ*), were named MPG591, MPG592, MPG593 and MPG594 respectively. For genes *nupC* and *nupG*, plasmids that were already constructed in the lab, for the purposes of other research that had taken place in the past were used. These plasmids were pGJL16 and pGJL25 with engineered *nupC* and *nupG* respectively and the genes are known to be induced when IPTG is added to exponentially growing cells.

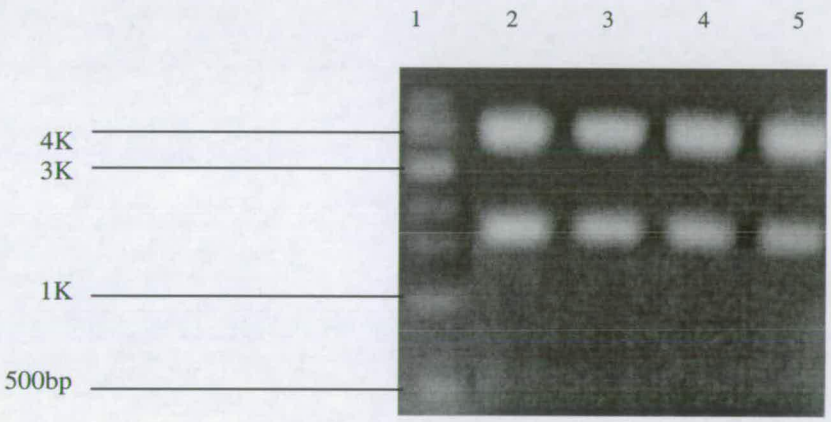


Figure 4.1: Gel showing the DNA bands resulting after digestions of the pBAD plasmids carrying the ORFS of genes *xapB*, *yegT*, *yeiM* and *yeiJ*. The digestion occurred for verification purposes. Plasmids pMRB and pMRT were digested with *Hind*III and *Eco*RI and resulted in two bands. Lanes 2 and 3 show the resulting bands from these digestion reactions. The biggest one is the pBAD vector at around 4.4 kb and the smaller bands are the inserts. Similarly pMRM and pMRJ were digested with restriction enzymes *Sal*I and *Kpn*I and the digestion results are shown in lanes 4 and 5. As mentioned before the top band represents the pBAD vector and the bottom bands are the inserts. The BLR strains containing plasmids pMRB (overexpressing *xapB*), pMRT(overexpressing *yegT*), pMTM (overexpressing *yeiM*), and pMRJ (overexpressing *yeiJ*), were named MPG591, MPG592, MPG593 and MPG594 respectively.

4.3 Transport assays using ^{3}H labelled adenosine as substrate

Uptake studies are a common method used in order to gain knowledge on the function of genes. So far scientists have gathered information for genes *nupC* and *nupG* through this method reaching the conclusion that they are both nucleoside permeases (Loewen *et al,* 2003, Xie *et al,* 2004).

The ORFs of putative nucleoside transporters XapB, YegT, YeiM and YeiJ were cloned into the pBAD18 expression vector as described above, and were then transformed into *E. coli* BLR host cells. Sufficient production of the protein expressed by the vector is necessary in order to be able to measure uptake of the radiolabelled nucleoside in the assay and this is aided by the use of the multicopy plasmid. Thus induction of the pBAD18-based engineered vectors ensured the required production of proteins from genes *xapB, yegT, yeiM* or *yeiJ.* The purpose of the assay is to get an insight into the possible function of the putative nucleoside transporters via their ability to transport radioactive adenosine. This was achieved by the experiment described below.

Overnight cultures of the BLR cells containing the pBAD18 plasmids with the ORF of the nucleoside transporters XapB, YegT, YeiM or YeiJ (strains MPG591, MPG592, MPG593 and MPG594 respectively) were used to set up 50ml starter cultures of M9 medium with glycerol (to final concentration of 0.2%), and ampicillin (100 μg/ml). Cultures were set up in duplicate, one of which was induced with arabinose after 3 hours of growth at 37°C and one remaining uninduced for use as a control. Cells were harvested by centrifugation for 10 minutes at 3000 rpm (2200g) at 4° C. Cells were then resuspended in KCl/MES (pH 6.6). This was repeated a further two times at 4° C. To complete the cell preparation a 1:10 dilution of the cell suspension was checked for OD_{680} and the cell concentration of the undiluted suspension was adjusted so that the OD_{680} would be between 1.8 and 2.2.

The substrate used in these experiments was adenosine prepared to a final concentration of 50μM and labelled to final specific radioactivity of 1μCi/ml with adenosine (2, 8-^{3}H Adenosine, from NEN). The label is on the carbon numbers 2 and 8 of the adenosine molecule. Cells were mixed with glycerol to a final concentration of 20mM at 25 °C and then were reenergized by bubbling with air for 3 minutes. At the appropriate time points the radiolabelled adenosine was added to the cell sample so as to yield a final concentration/specific radio activity of 50μM/ 1μCi/ml in the assay. Uptake was measured over a period of 4 minutes and samples were taken at 15 seconds, 1, 2 and 4 minutes. The initial rate of transport is given by the sample taken at 15 seconds. Radioactivity was measured by liquid scintillation counting. Further details on the methodology and materials used in this experiment are outlined in the Materials and Methods chapter.

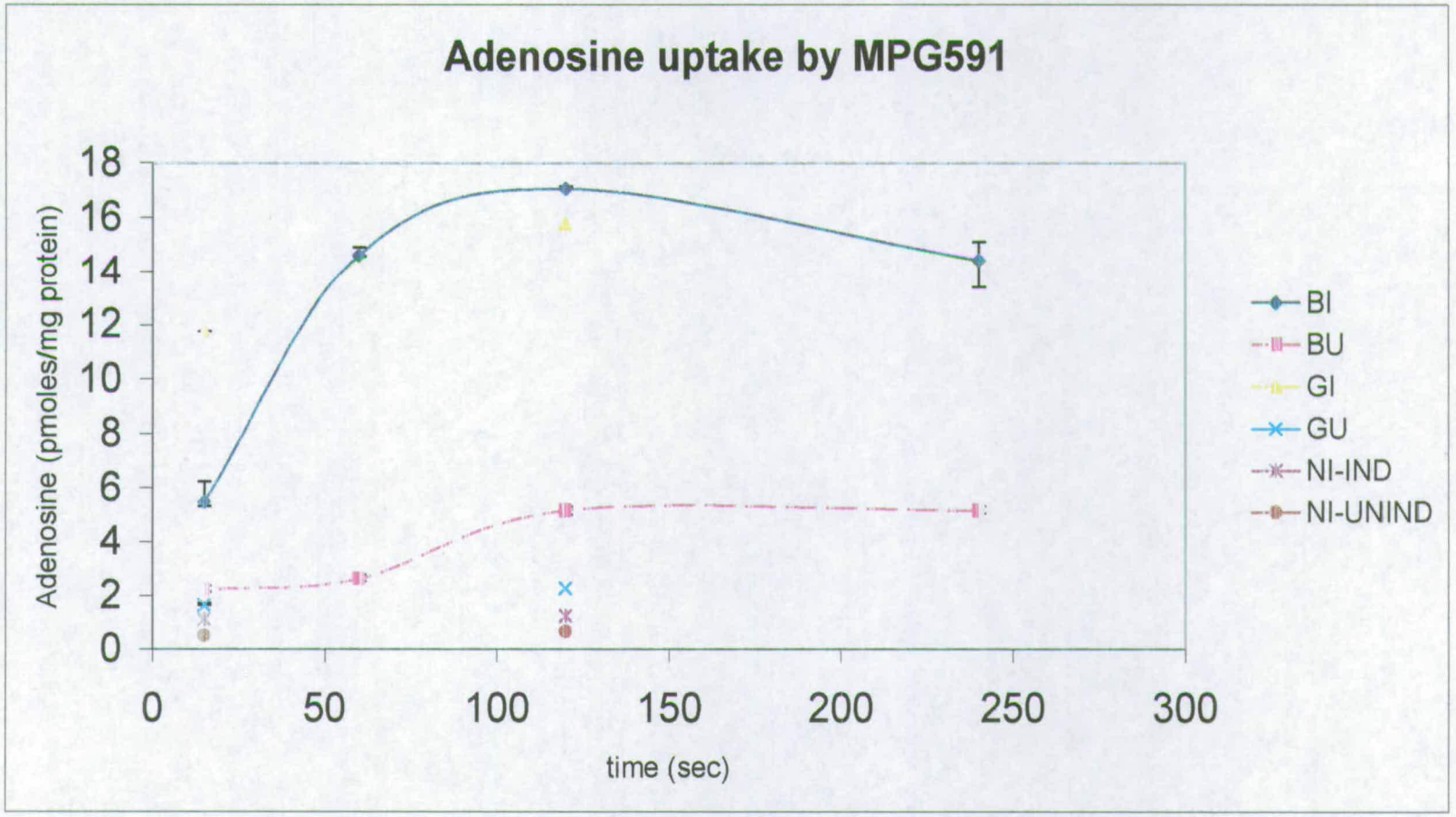
Standards were included in duplicate (radioactive adenosine only) as well as duplicates of filter blanks (Gelman GN-6, pre soaked in KCl/MES). *nupG* cloned in vector pGJL25, a pTTQ18 derivative under a controllable IPTG-induced promoter, was used as a positive control in order to ensure that the assay was working. This positive control was constructed previously in our lab and has been shown to be positive for nucleoside uptake. Also, plasmid pBAD18 without insert was used as a control, with and without arabinose induction, to deduce the background levels of expression and consequent uptake of the substrate.

From the data gathered in figure 4.2 (A-D) it is demonstrated that all four transporters seem to have the ability to transport adenosine, reflecting their likely function as nucleoside transporters. Specifically figure 4.2 panels A-D, show the uptake of ^{3}H adenosine after genes *xapB, yegT, yeiJ* and *yeiM* are over-expressed in the host strain. In panel A, over expression of XapB (symbol BI) causes a significant increase in the uptake of adenosine by the cells with transport reaching saturation point under the specific condition of this experiment as described above, after 2 minutes. In panel B, the over expression of putative nucleoside transporter YegT (symbol TI) enhances uptake of

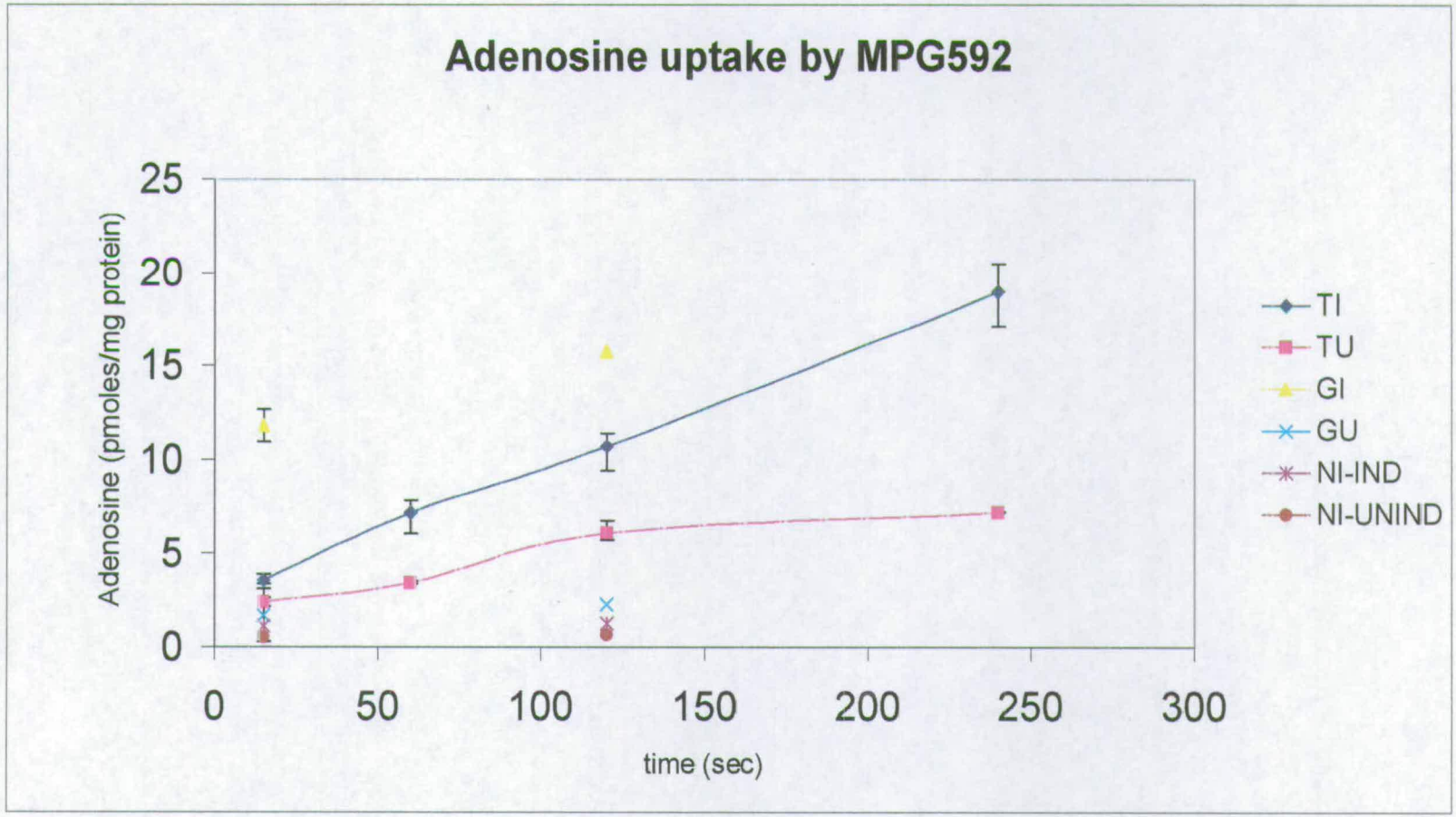
adenosine by the cells. After 4 minutes the cells are still transporting adenosine without reaching a saturation point. This reflects that the steasy state as well as the calculation of V_{max} for this transporter can be found at a time point beyond 4 minutes or that the substrate is the limiting factor at this concentration. Panel C shows the effect of over expression of YeiM on the uptake of adenosine by the cell. Saturation is reached before 2 minutes. Finally panel D shows uptake of adenosine by the cells after over expression of putative nucleoside transporter YeiJ. Overall all transporters seem to transport adenosine well. The initial rate of transport for YegT is lower than the rest of them, although clearly above background levels reaching the same levels of uptake as the rest of the transporters at 4 minutes without reaching saturation point.

Figure 4.2 (A, B, C, D): Time courses of radiolabelled adenosine uptake in *E. coli* MPG591, MPG592, MPG593 and MPG594 (BLR cells with pMRB, pMRT, pMRM and pMRJ respectively). 2,8-^{3}H Adenosine (50μM, 1μCi/ml) uptake was determined at 25°C at 15 seconds, 1, 2 and 4 minutes. BI, TI, MI and JI indicate induced BLR cells with plasmids pMRB, pMRT, pMRM or pMRJ respectively. BU, TU, MU and JU represent uninduced cells containing the respective plasmids. A plasmid containing *nupG*, known to transport nucleosides when induced (GI), is included as a positive control. The same plasmid uninduced (GU) is included as a negative control. Plasmids containing no insert, induced (NI-IND) and uninduced (NI-UNIND), were used to provide further control for background levels of uptake. The values shown on the graphs are means of triplicate determinations of uptake. Error bars represent the standard deviation from the means and they are not shown where they are smaller in size than the size of the symbols.

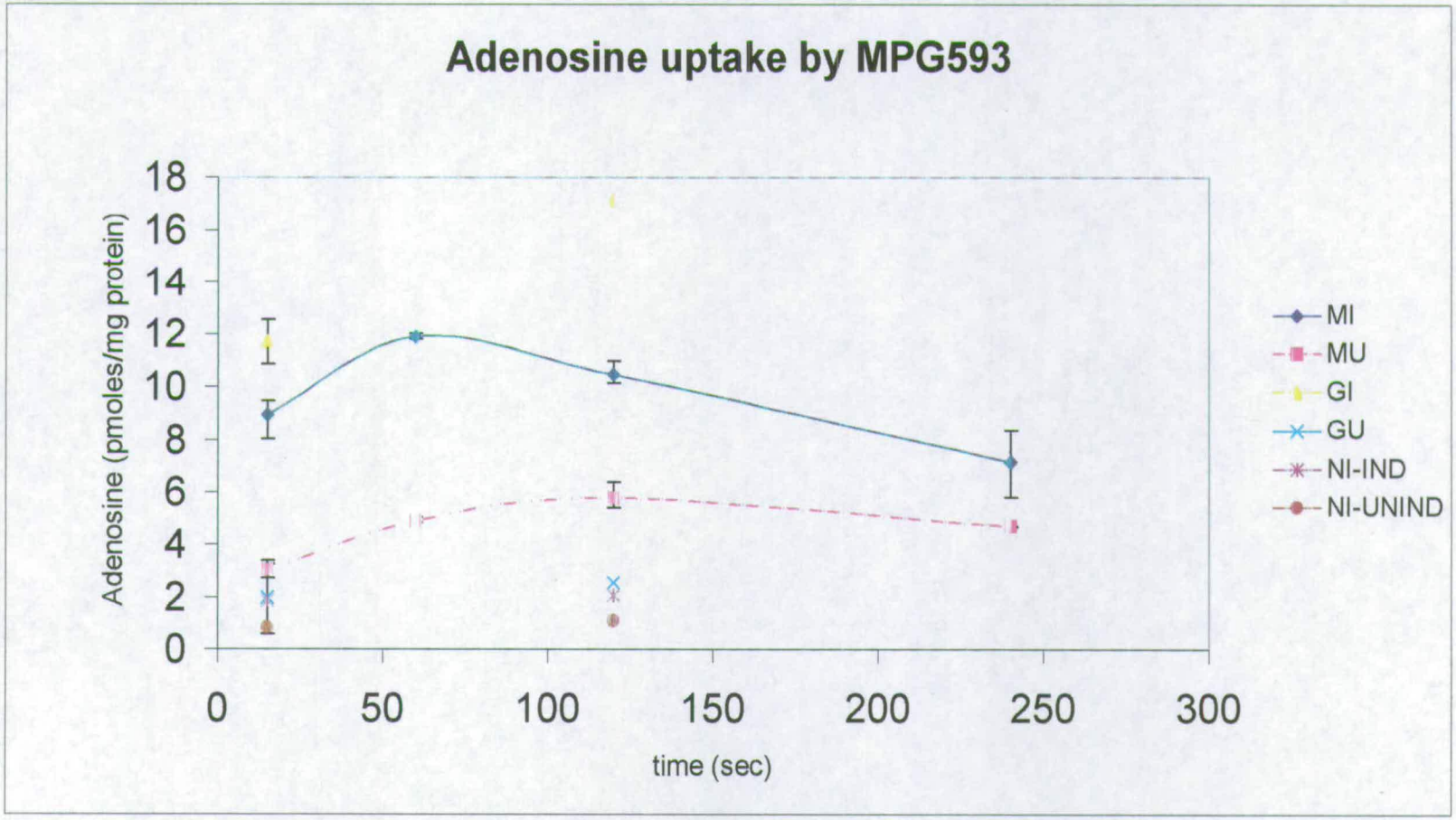
KEY: MPG591 overexpressing *xapB*
MPG592 overexpressing *yegT*
MPG593 overexpressing *yeiM*
MPG594 overexpressing *yeiJ*


Adenosine uptake by MPG591
Adenosine (pmoles/mg protein)
18
16
14
12
10
8
6
4
2
0
0
50
100
150
200
250
300
time (sec)
BI
BU
GI
GU
NI-IND
NI-UNIND

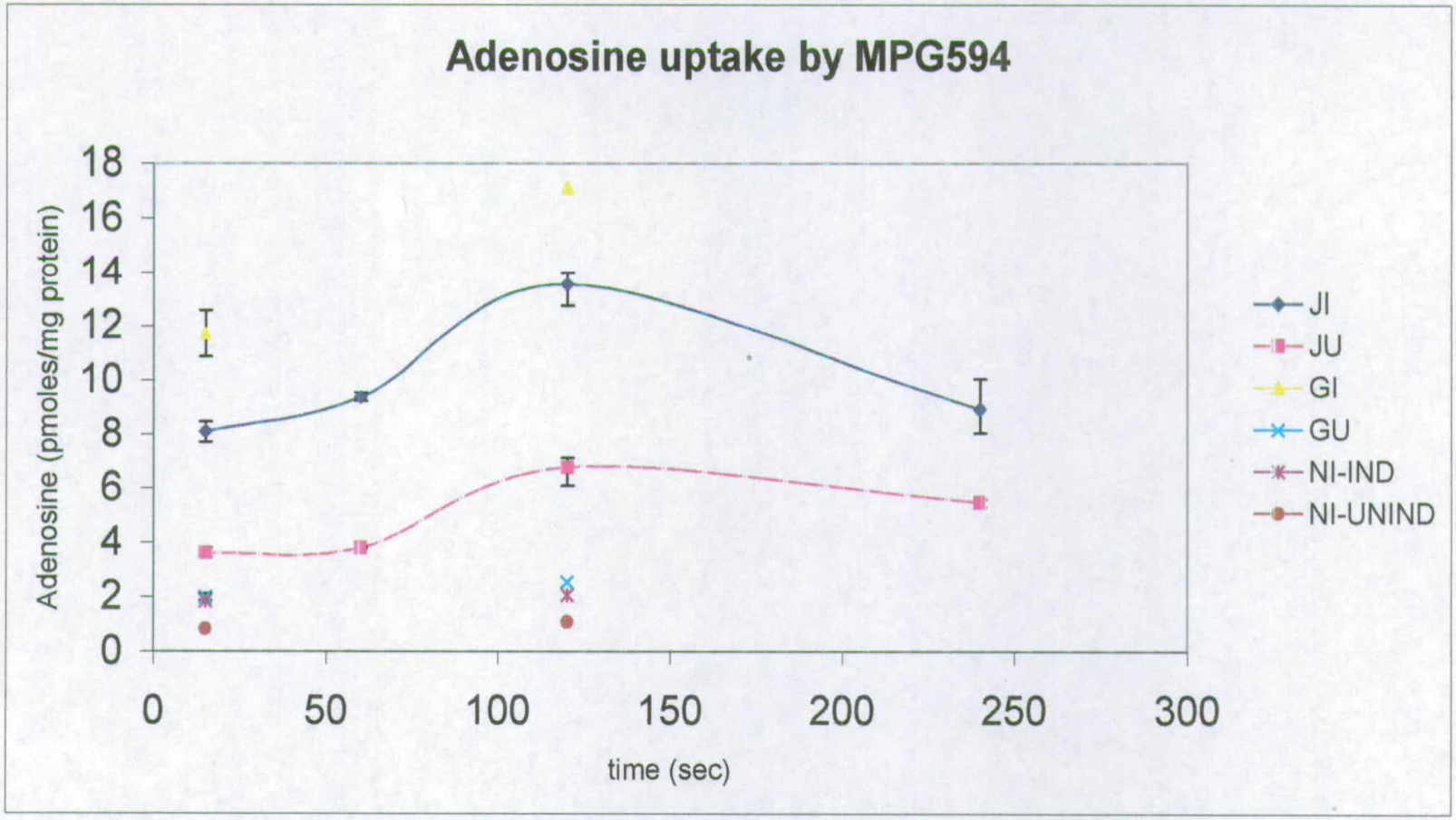
A.


Adenosine uptake by MPG592
Adenosine (pmoles/mg protein)
25
20
15
10
5
0
0
50
100
150
200
250
300
time (sec)
TI
TU
GI
GU
NI-IND
NI-UNIND

B.


Adenosine uptake by MPG593
Adenosine (pmoles/mg protein)
18
16
14
12
10
8
6
4
2
0
0
50
100
150
200
250
300
time (sec)
MI
MU
GI
GU
NI-IND
NI-UNIND

C.


Adenosine uptake by MPG594
Adenosine (pmoles/mg protein)
18
16
14
12
10
8
6
4
2
0
0
50
100
150
200
250
300
time (sec)
JI
JU
GI
GU
NI-IND
NI-UNIND

D.

4.4 Analysis of kinetic parameters

The time-dependent studies confirmed that adenosine is indeed a substrate for each of the transporters. These studies were followed by a series of experiments aimed at the determination of the kinetic parameters for nucleoside uptake by each of the gene products. By varying the concentrations of the substrate it is possible to determine the saturation parameter for XapB, YegT, YeiM and YeiJ.

As in the previous studies these experiments used *E. coli* BLR cells which were transformed with pBAD plasmids carrying the ORFs of the targeted genes as described above. Cell suspensions for the assays were prepared as described previously in the above section in terms of the setting up of the starter cultures, culture induction with arabinose, and cell washing and their adjustment to an OD_{680} of 1.8-2.2.

Because in these experiments the estimation of kinetic parameters of transport were investigated, a range of concentrations several times higher than the final concentration of 50μM which was used in the initial assay (Figure 4.2) were prepared. The substrate used for these experiments was again adenosine prepared to a range of concentrations and radiolabelled to 1μCi/ml with 2-8 ^{3}H adenosine. Cells were mixed with glycerol to a final concentration of 20mM at 25 ° C and then were reenergized by bubbling with air for 3 minutes as described before. Uptake was measured over a period of 15 seconds and cells were mixed with adenosine at one of six different final concentrations of 7.5, 10, 25, 50, 100 and 200μM. Radioactivity was measured on filters by liquid scintillation counting. More details are given in the Materials and Methods section 2.2.4 of Chapter 2.

Standards were included in duplicate (radioactive adenosine only) as well as duplicates of filter blanks as previously. *nupG* cloned in vector pGJL25, a pTTQ18 derivative under a controllable IPTG-induced promoter, was used as a positive control in order to ensure that the assay was working. Also, plasmids without inserts were used as insert

independent controls, with and without arabinose induction, to deduce the background levels of expression and consequent uptake of the substrate, as before.

Figure 4.3 (A-D) shows curves for concentration dependence for adenosine for the four putative nucleoside transporters XapB, YegT, YeiM and YeiJ, measured by initial rates of uptake within 15 seconds. Apparent Michaelis-Menten constant (Km) values for adenosine transport for the four genes are shown in table 4.1. Km was calculated by the formula

$$V = \frac{V_{max}\ [S]}{Km + [S]}$$

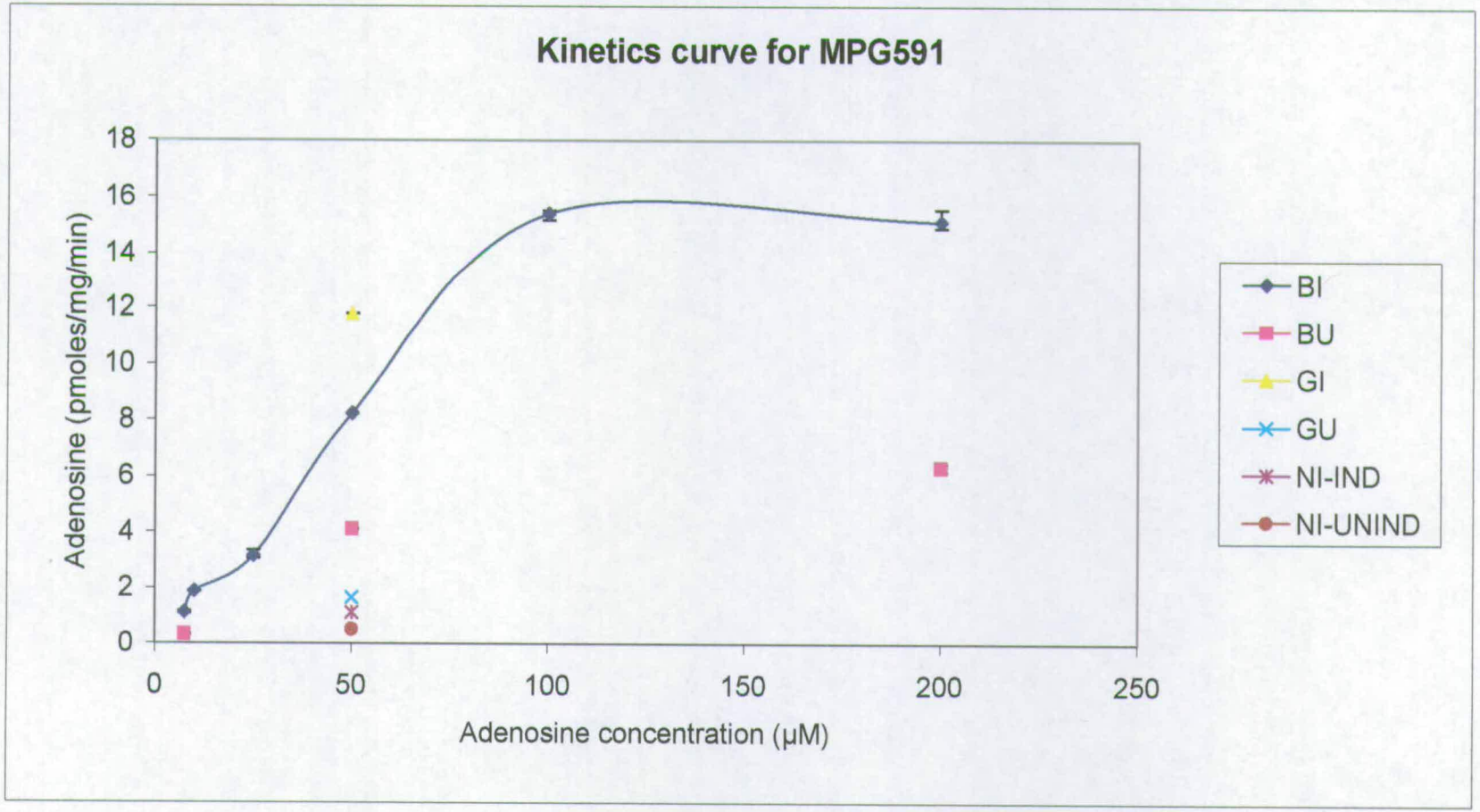
In this formula and for determination of the Km values at saturation point, [S] is the substrate concentration and Vo equals half the maximum velocity of substrate saturation for a specific transporter. So Km is the concentration of the substrate at which $V = V_{max}/2$.

Because visual determination of the V_{max} and Km from the Michaelis-Menten kinetics curves can be inaccurate, enzyme kinetics software was used for precise determination. The Km values determined via the software program GraphPad Prism 4 are summarised in Table 4.1. The Km values were determined from graphs based on the data set presented in figure 4.3 which are summarised in Appendix 2. These include the Lineweaver-Burk plots from which Km values are derived. As seen the Km values determined by the program are 97.1, 65, 160.1 and 453.9μM for transporters XapB, YegT, YeiM and YeiJ respectively. Although these are higher values than the values of 0.3-1μM determined for NupC- and NupG- mediated cytidine transport (Munch-Petersen *et al,* 1979 and Komatsu and Tanaka, 1973), they seem to be in the same range as found for previous experiments in *E. coli* for XapB (Nørholm and Dandanell, 2001), as well as for transporters in other organisms such as the Pacific hagfish *hf*CNT1 (Yao *et*

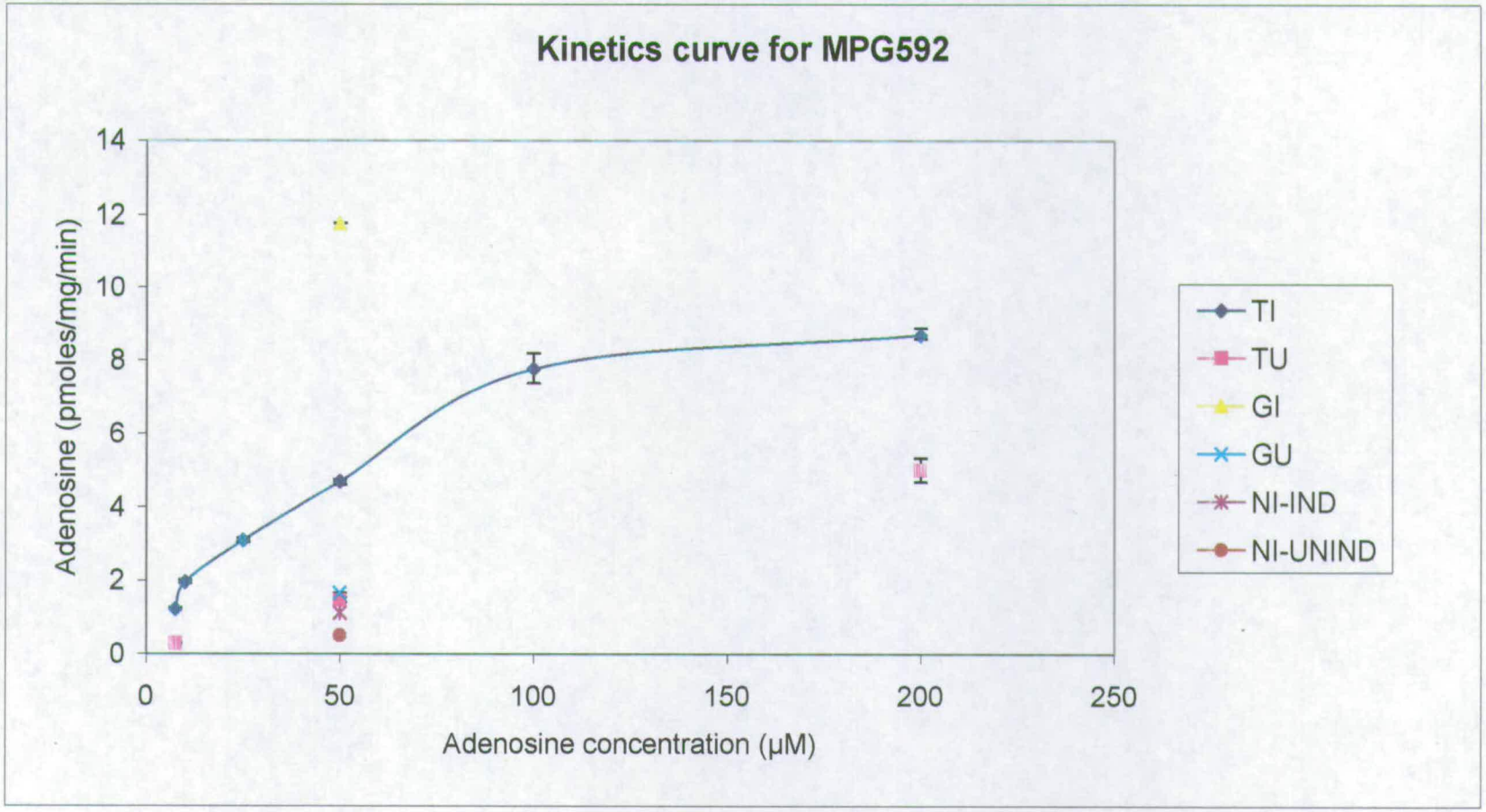
al, 2002), rat jejunal epithelium rCNT1 (Huang *et al,* 1994) or rat jejunal/kidney rCNT1 and rCNT2 (Yao *et al,* 1996), and human CNT1, 2 and 3 (Lostao *et al,* 2000, Wang *et al,* 1997, Ritzel *et al,* 2001a and b).

From figure 4.3 it can be seen that adenosine uptake by XapB, YegT and YeiM reached saturation at specific concentrations of adenosine whilst YeiJ did not reach saturation even at the maximum concentration of adenosine used of 200μM. This is also reflected by the high Km values for this transporter. This is also the case with a number of other transporters that have appeared in the literature, such as the hagfish transporter *hf*CNT1, which reaches saturation of thymidine and inosine at a concentration of more than 400 μM (Yao *et al,* 2002). However routinely, *E. coli* nucleoside transporters are tested with physiological nucleosides at concentrations of a maximum of 50μM. Since the transporters studied here are uncharacterised, it might be anticipated that there would be a need to experiment with nucleoside concentrations over a wide range. Equally it may be that these transporters have evolved to recognise substrates other than adenosine and thus, exhibit different kinetic parameters.

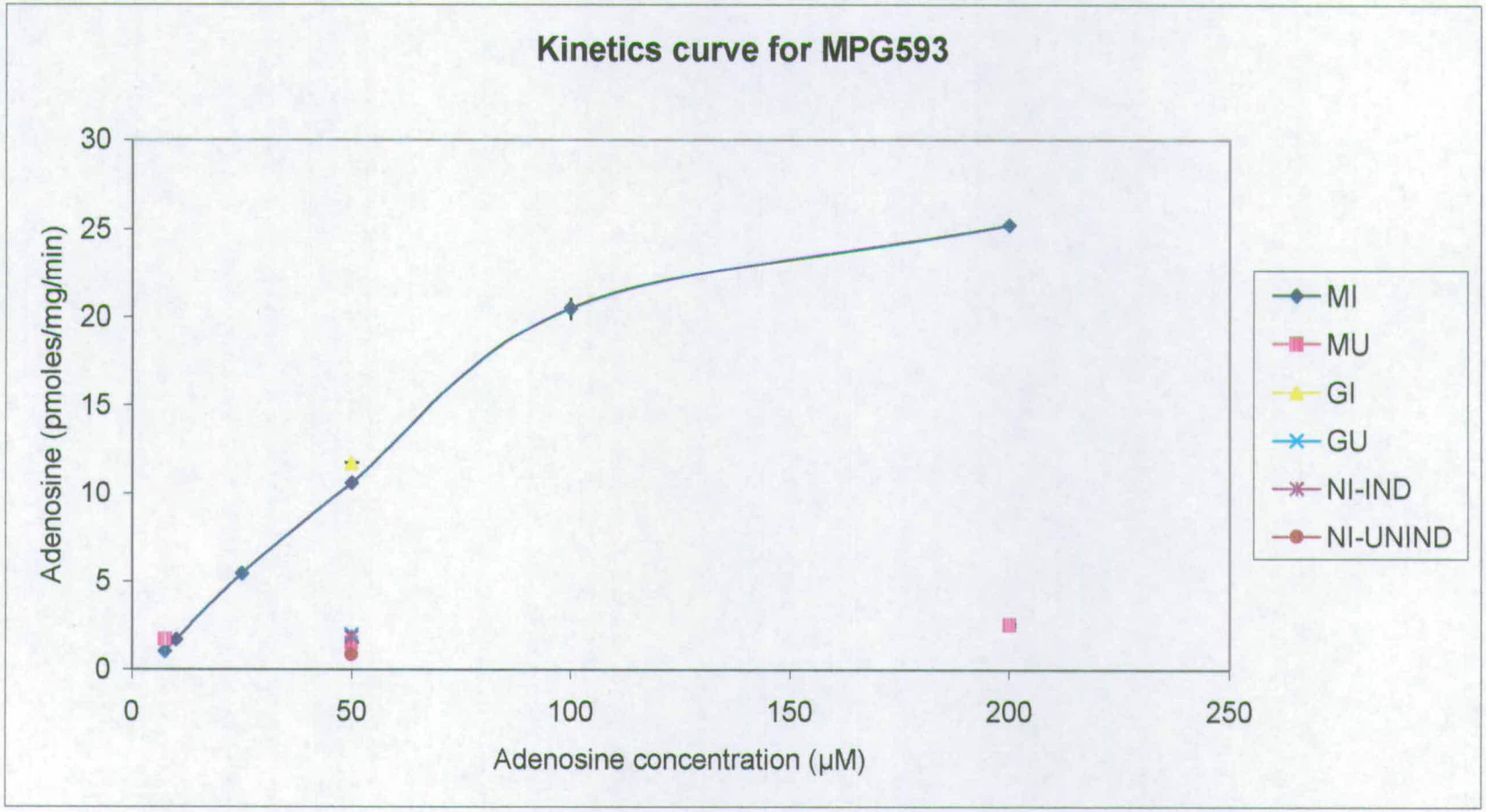
Figure 4.3 (A, B, C, D): Concentration dependent profiles of radiolabelled adenosine uptake in *E. coli* MPG591, MPG592, MPG593 and MPG594 (BLR cells with pMRB, pMRT, pMRM and pMRJ respectively). Final concentration of radiolabelled [2,8- ^{3}H] adenosine spread in a range of 7.5, 10, 25, 50, 100 and 200μM and uptake was determined at 25°C at 15 seconds. BI, TI, MI and JI indicate induced BLR cells with plasmids pMRB, pMRT, pMRM or pMRJ respectively. BU, TU, MU and JU represent uninduced cells containing the respective plasmids. A plasmid containing *nupG*, known to transport nucleosides when induced (GI) is included as a positive control. The same plasmid uninduced (GU) is included as a negative control. Plasmids containing no insert induced (NI-IND) and uninduced (NI-UNIND) were used to provide further control for background levels of uptake. The values shown on the graphs are means of three experimental repeats. Error bars represent the standard deviation for the means and they are not shown where they are smaller in size than the size of the symbols


Kinetics curve for MPG591
Adenosine (pmoles/mg/min)
18
16
14
12
10
8
6
4
2
0
0
50
100
150
200
250
Adenosine concentration (μM)
BI
BU
GI
GU
NI-IND
NI-UNIND

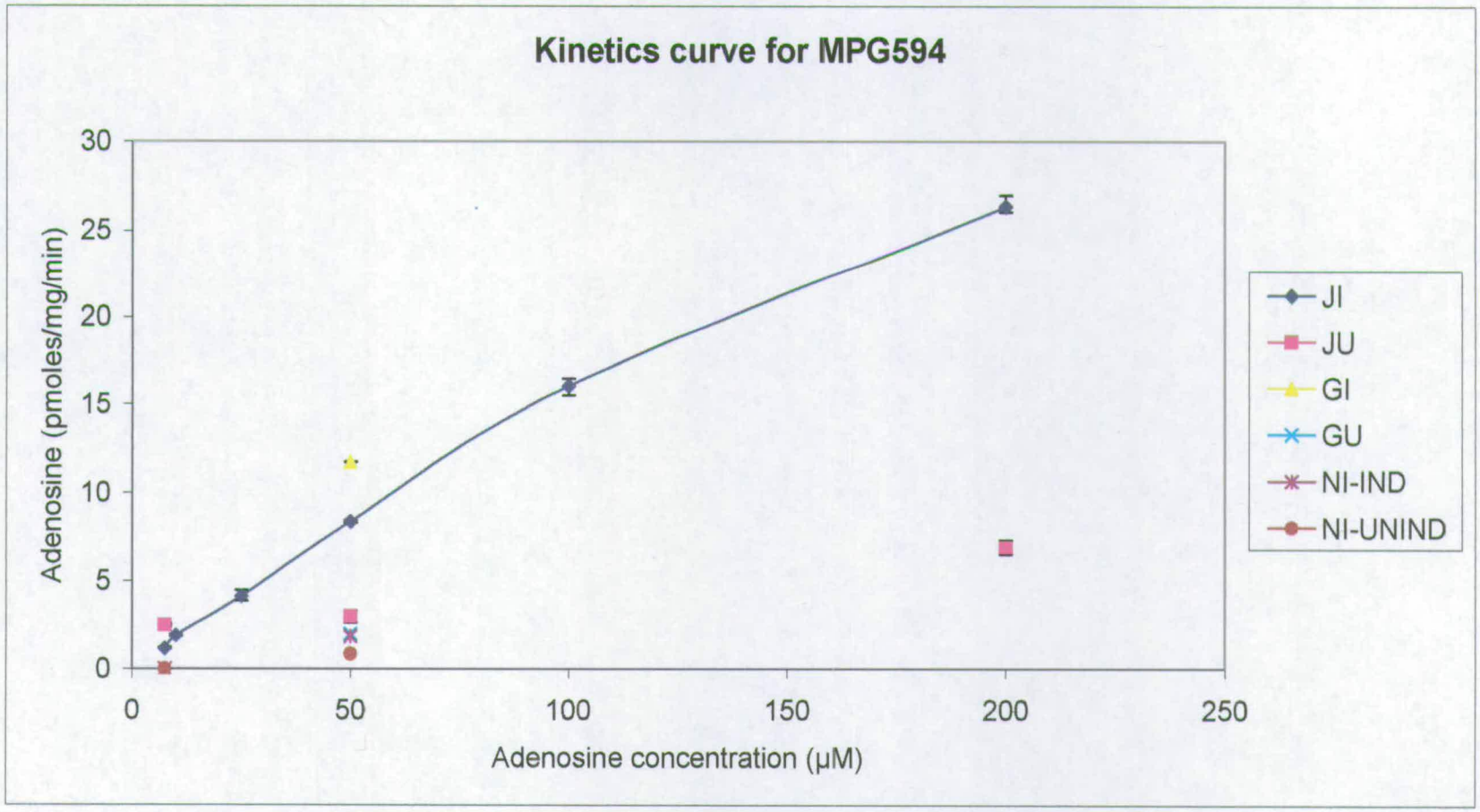
A.


Kinetics curve for MPG592
Adenosine (pmoles/mg/min)
0
2
4
6
8
10
12
14
0
50
100
150
200
250
Adenosine concentration (μM)
TI
TU
GI
GU
NI-IND
NI-UNIND

B.


Kinetics curve for MPG593
Adenosine (pmoles/mg/min)
30
25
20
15
10
5
0
0
50
100
150
200
250
Adenosine concentration (μM)
MI
MU
GI
GU
NI-IND
NI-UNIND

C.


Kinetics curve for MPG594
Adenosine (pmoles/mg/min)
30
25
20
15
10
5
0
0
50
100
150
200
250
Adenosine concentration (μM)
JI
JU
GI
GU
NI-IND
NI-UNIND

D..

(Putative) Nucleoside Transporter	Substrate	Expression system	Apparent Km (μM)	Reference
XapB	Xanthosine	Whole *E. coli* cells	136	Nørholm & Dandanell, 2001
NupC	Adenosine	*Xenopus* oocytes	1.6	Loewen *et al*, 2003
rCNT1	Adenosine	*Xenopus* oocytes	26	Yao *et al*, 1996
rCNT1	Uridine	*Xenopus* oocytes	37	Huang *et al*, 1994
hfCNT1	Uridine	*Xenopus* oocytes	10	Yao *et al*, 2002
hfCNT1	Inosine	*Xenopus* oocytes	35	Yao *et al*, 2002
hfCNT1	Thymidine	*Xenopus* oocytes	45	Yao *et al*, 2002
hCNT1	Cytidine	*Xenopus* oocytes	140	Lostao *et al*, 2000
hCNT2	Uridine	*Xenopus* oocytes	116	Wang *et al*, 1997
hCNT2	Adenosine	*Xenopus* oocytes	8	Ritzel *et al*, 1998
hCNT3	Adenosine	*Xenopus* oocytes	15.1	Ritzel *et al*, 2001a
NupC	Cytidine	*E. coli* membrane vesicles	0.3-1	Munch-Petersen, 1979
NupG		Whole *E. coli* cells		Komatsu & Tanaka, 1973
YegT	Adenosine	Whole *E. coli* cells	65	This project
XapB	Adenosine	Whole *E. coli* cells	97.09	This project
YeiM	Adenosine	Whole *E. coli* cells	160.1	This project
YeiJ	Adenosine	Whole *E. coli* cells	453.9	This project

Table 4.1: Kinetic parameters for *E. coli* and mammalian nucleoside transporters with a variety of substrates. Included are Km values as determined from data shown in figure 4.3 A-D and Lineweaver-Burk plots in Appendix 2. For gene *yeiJ*, the Km is higher than the rest and the graph did not plateau within the range of adenosine concentration used in these experiments. As seen from the graph, variables in Km determination include transporter to be studied, substrate, and expression system.

4.5 Characterisation of the specificity of *E. coli* nucleoside transport systems

Initially cells containing the plasmids bearing the ORFs from genes *xapB, yegT, yeiM* and *yeiJ* were tested for uptake of radiolabelled adenosine. These assays showed that the genes' products are able to utilise adenosine above background levels. To further investigate the functions of those proteins as nucleoside transporters and discover if the putative or unknown nucleoside transporters have specificity for particular nucleosides, inhibition of adenosine uptake by competition with other purine and pyrimidine nucleosides was examined with *E. coli* BLR cells, transformed with plasmids pMRT, pMRB, pMRJ and pMRM.

As in the previous studies these experiments used *E. coli* BLR cells which were transformed with pBAD plasmids carrying the ORFs of the target genes as described above. Cell suspensions for the assays were prepared as described previously in the above section in terms of the setting up of the starter cultures, culture induction (including a control of uninduced cells) with arabinose, cell washing and adjustment to an OD_{680} of 1.8-2.2. The main substrate used in these experiments was adenosine prepared to a final concentration of 50μM and labelled to 1μCi/ml with 2-8 ^{3}H adenosine. Specificity of the nucleoside transport systems was measured by comparing the uptake of this radiolabelled adenosine at this specific concentration against an excess of a range of nucleosides which were prepared to a final concentration of 1mM. These samples of nucleoside competitors were not radiolabelled. The cells were mixed with glycerol to a final concentration of 20mM at 25 °C and then they were re-energized by bubbling with air for 3 minutes as described before. Adenosine of 2mM, with a radioactive concentration of 10μCi/ml, was used and at the appropriate time point of 15 seconds, the radiolabelled adenosine was added to the cell sample so as to yield a final concentration of 50μM/ 1μCi in the assay. Also samples of induced cells containing the competitor nucleoside to a final concentration of 0.5mM were treated exactly the same way and the impact of the competitor's presence on adenosine uptake was measured. Because XapB- and YegT-mediated adenosine transport inhibition is not profoundly

shown after 15 seconds, due to the high number of the uptake of the positive control the inhibitory effect of various nucleosides was studied as well after two minutes of incubation.

From figure 4.4 B it can be seen that adenosine transport via XapB was strongly inhibited by inosine, showing a 78% decrease in adenosine transport after addition of inosine in the assay. Thymidine and uridine addition in the assay caused decrease of more than 30% in adenosine uptake (35% and 39% respectively). In addition guanosine and cytidine decreased adenosine uptake in the assay to a smaller extent than the previously mentioned nucleosides (19% and 12% decrease). In other words XapB seems to be able to support uptake of almost all tested nucleosides, suggesting that it might have wide substrate specificity, like *nupG*, although the affinities for specific nucleosides vary.

Similarly adenosine transport via YegT (figure 4.4 D) was inhibited by all nucleosides tested, which also reflects that the protein has the ability to utilise a wide range of substrates. In particular adenosine transport is inhibited by as much as 70% by cytidine and 60% by uridine. Thymidine or guanosine addition in the assay decreased adenosine transport via YegT by more than 50% (59% and 52% respectively). Finally inosine decreased adenosine transport by 27%.

Adenosine transport via YeiM (figure 4.4 E) was inhibited strongly by cytidine and uridine (75% and 66%) while guanosine addition in the assay also caused a decrease in adenosine transport albeit smaller. In particular guanosine addition caused a decrease of 27% while inosine resulted in a mere 7% decrease. Thymidine addition had no effect in the adenosine uptake via YeiM.

Finally adenosine transport via transporter YeiJ (figure 4.4 F) was inhibited by more than 60% by cytidine and uridine (65% and 61% respectively). Inosine had a very small effect on adenosine transport via YeiJ, with a recorded decrease in transport of 7%.

Guanosine addition in the assay had no effect at all on adenosine uptake. Unusually the addition of thymidine in the assay caused an enhanced uptake of adenosine by YeiJ by more than 50%. This was not recorded for any other putative or known nucleoside transporter studied in this project. Table 4.2 summarises the effect of addition of various nucleosides on adenosine uptake assay by transporters XapB, YegT, YeiM, and YeiJ.

More specifically transport of adenosine by XapB was strongly inhibited by inosine. This is an indication of the specificity of the transporter. It is now known that the substrate of *xapB* is xanthosine. Interestingly the *E. coli* IMP dehydrogenase *guaB* catalyses the conversion of inosine to xanthosine. IMP dehydrogenases are also associated with cell proliferation and they are possible targets for cancer chemotherapy in other species.

The reaction of inosine conversion to xanthosine is shown below:

Inosine 5-phosphate + NAD+ + H_2O = Xanthosine 5-phosphate + NADH

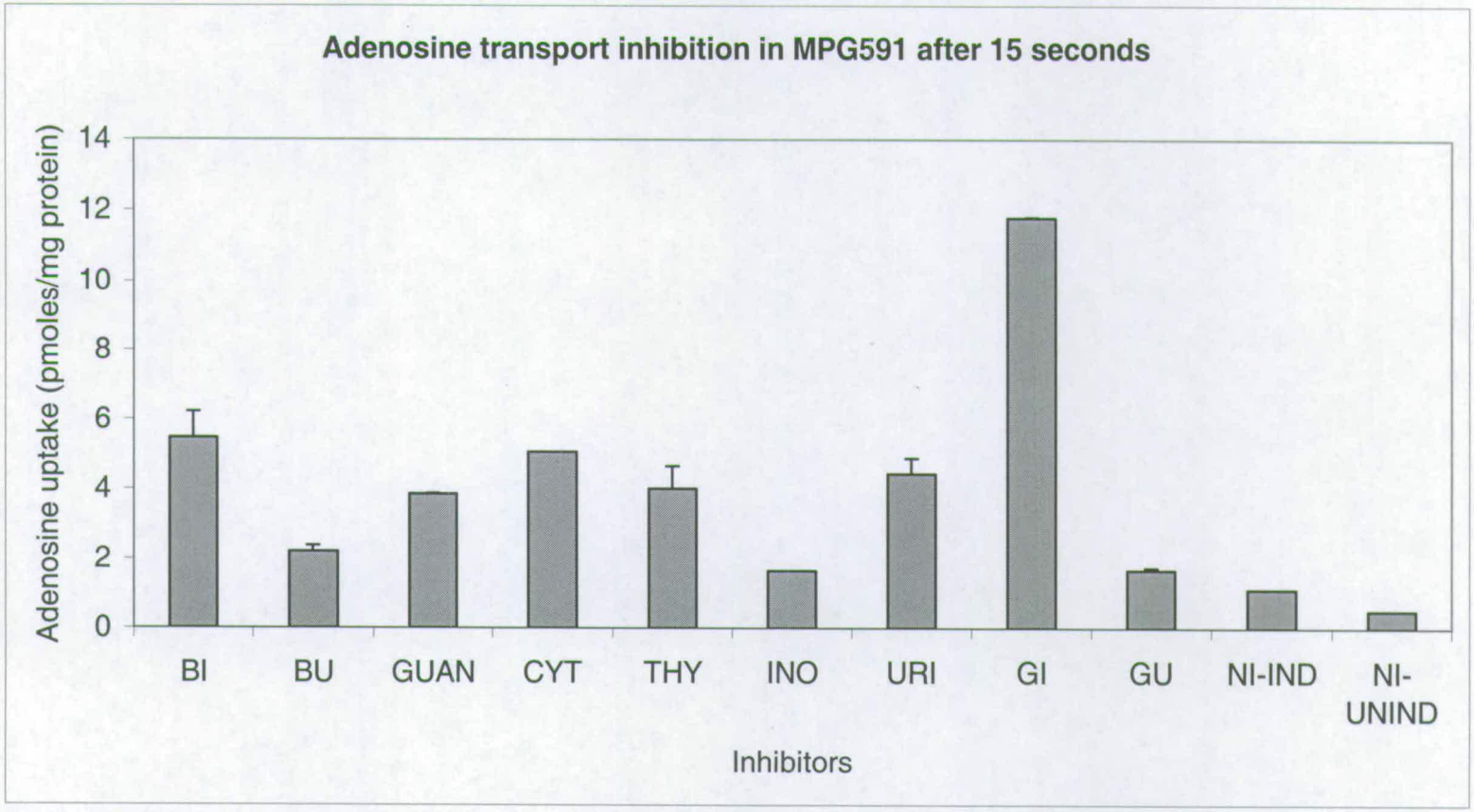
Also transport of adenosine via YegT is most strongly inhibited by cytidine and uridine. Both these substrates are pyrimidines so this might suggest that while *yegT* is indeed related to *nupG* through homology, the preferred substrates for it might be one of these two nucleosides.

Adenosine transport in the cases of YeiM and YeiJ is strongly inhibited by pyrimidines cytidine and uridine. This is in line with the fact that both genes are paralogues of *nupC* which is a strong permease for pyrimidines. However both transporters are unaffected by thymidine competition while guanosine, in the case of YeiM, seems to have an inhibitory effect.

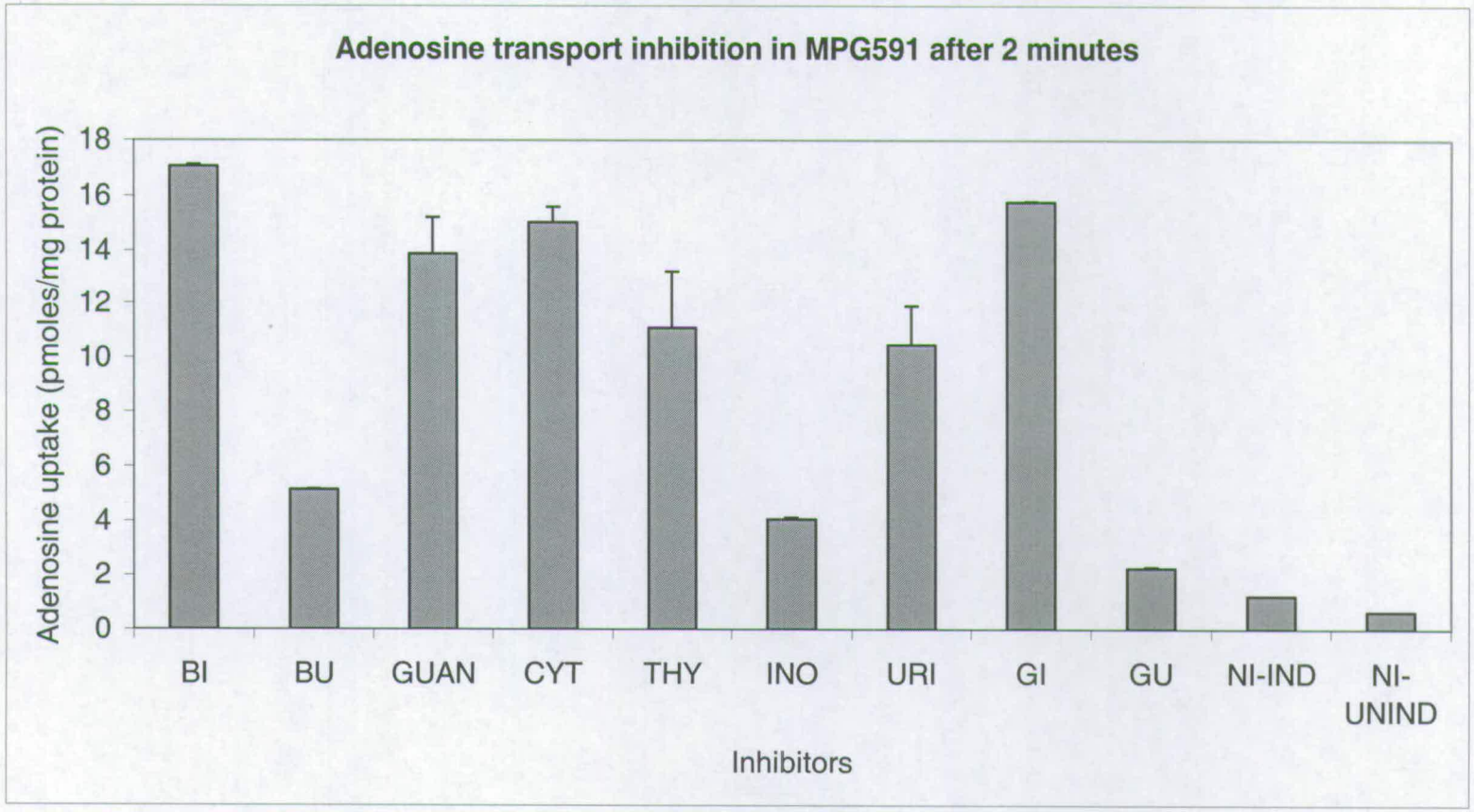
Furthermore a paradox appears in figure 4.4 F where t after addition of thymidine the transport of adenosine seems to increase instead of either decrease or remain the same.

This could either mean that addition of thymidine indeed enhanced adenosine transport via YeiJ or that this was an atypical experimental error. Further repetition of this experiment is required to clarify the situation.

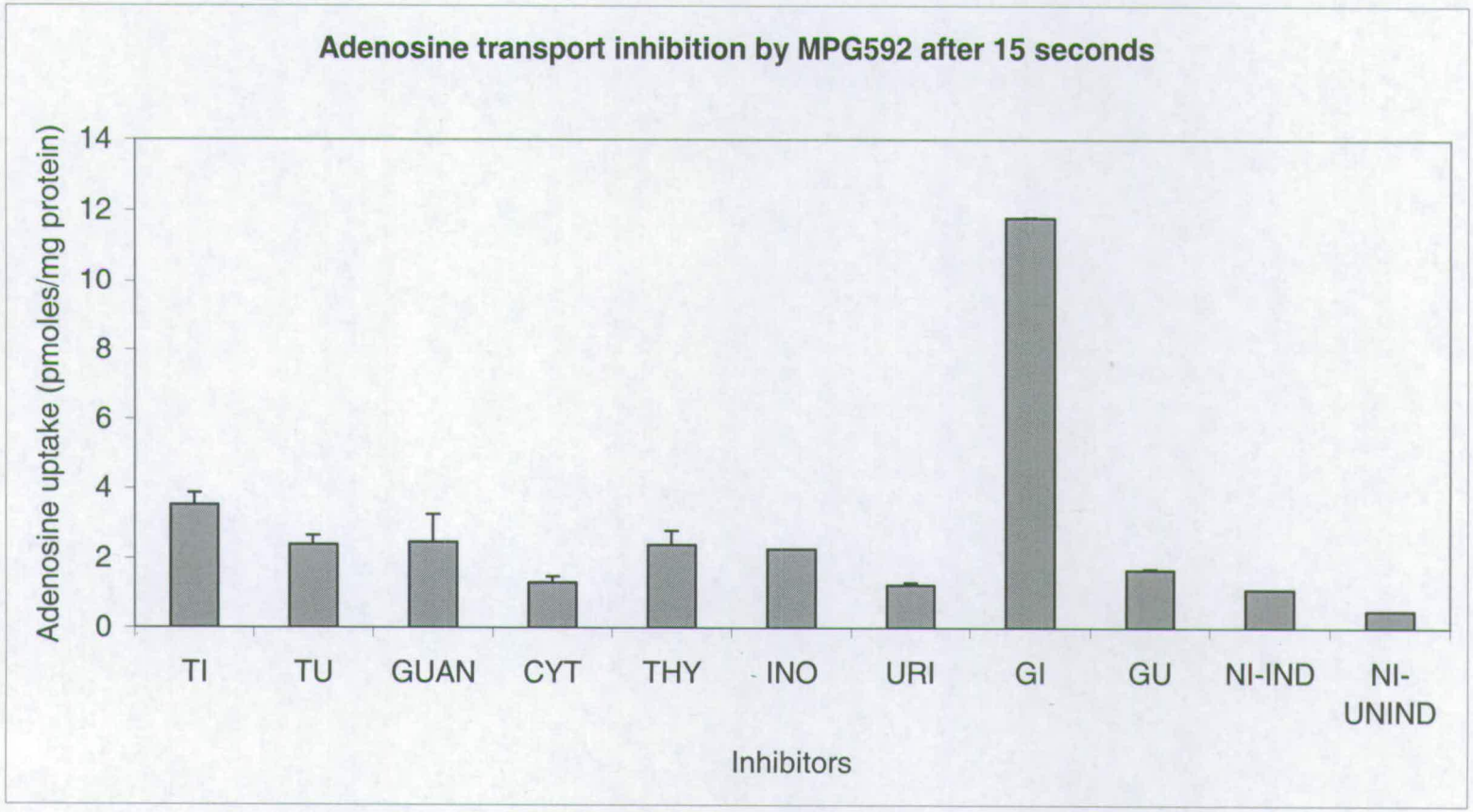
Figure 4.4 (A-F): Inhibition of [2, 8- ^{3}H] adenosine uptake by BLR cells transformed with pMRB, pMRT, pMRM and pMRJ (strains MPG591, MPG592, MPG293 and MPG594respectively). Adenosine was prepared at 50μM, 1μCi/ml and uptake was determined at 25°C 15 seconds after addition of the labelled adenosine in the assay. BI, TI, MI and JI stand for induced BLR cells with plasmids pMRT, pMRB, pMRJ and pMRM respectively while BU, TU, MU and JU are uninduced cells containing the respective plasmids without the addition of arabinose. A plasmid containing *nupG*, known to transport nucleosides when induced (GI) is included as a positive control. The same plasmid uninduced (GU) is included as a negative control. Plasmids containing no insert induced (NI-IND) and uninduced (NI-UNIND) were used to provide further control for background levels of uptake. The values shown on the graphs are means of triplicate determinations of uptake. Error bars represent the standard deviation for the means and they are not shown where they are smaller in size than the size of the symbols Nucleosides guanosine (GUAN), cytidine (CYT), thymidine (THY), inosine (INO) and uridine (URI) were used to challenge the adenosine uptake to a final concentration of 0.5mM in the assay. For transporters XapB and YegT there are graphs shown (panels B and D) that show the effect of the added nucleoside on adenosine uptake after two minutes as well as after 15 seconds (panels A and C respectively for XapB and YegT). The reason for that is that due to the very high uptake of the positive control (GI) in both cases (panels A and C respectively for XapB and YegT) so the effect of the individual nucleosides, although present, is not as apparent as after two minutes.


Adenosine transport inhibition in MPG591 after 15 seconds
Adenosine uptake (pmoles/mg protein)
14
12
10
8
6
4
2
0
BI
BU
GUAN
CYT
THY
INO
URI
GI
GU
NI-IND
NI-UNIND
Inhibitors

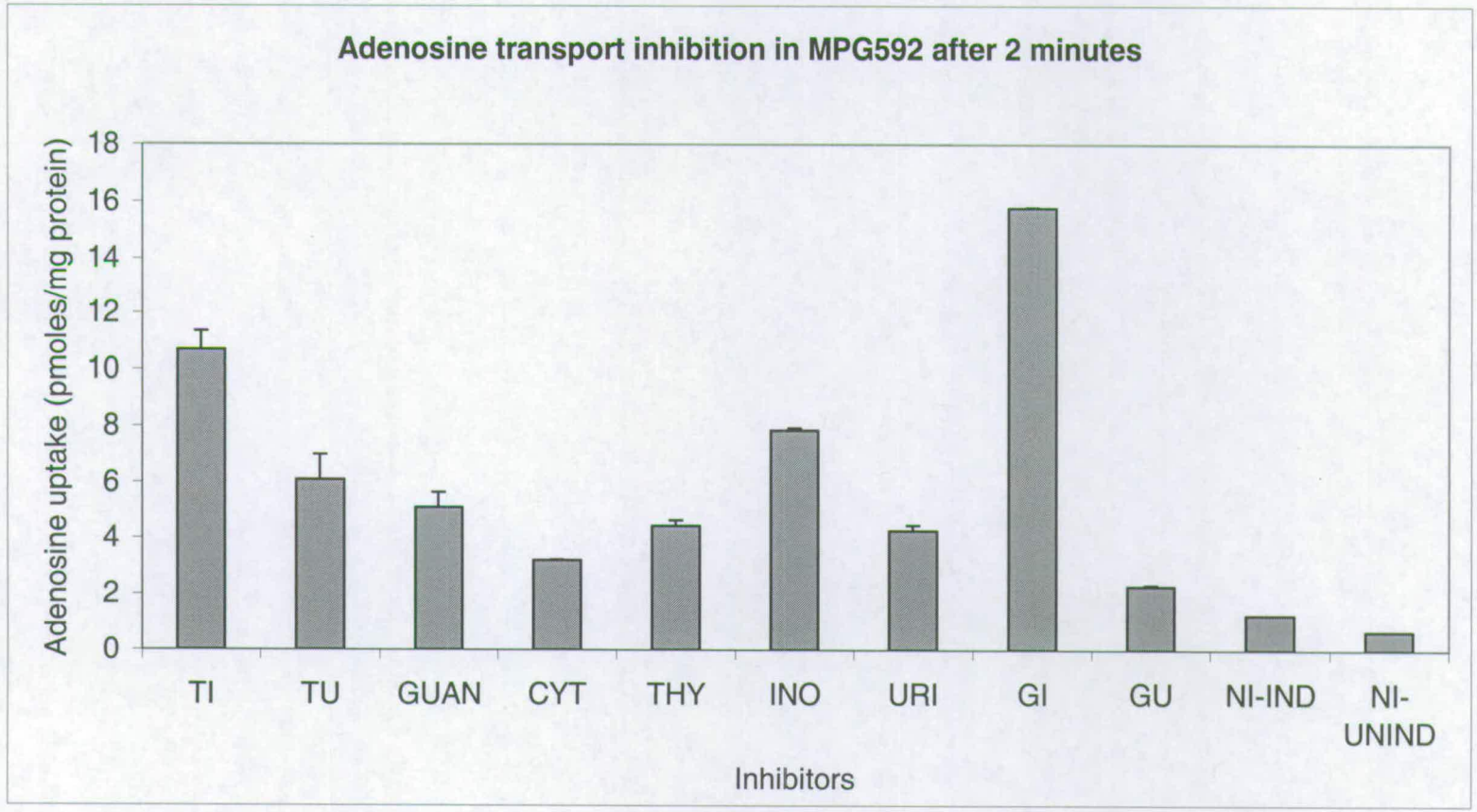
A.


Adenosine transport inhibition in MPG591 after 2 minutes
Adenosine uptake (pmoles/mg protein)
18
16
14
12
10
8
6
4
2
0
BI
BU
GUAN
CYT
THY
INO
URI
GI
GU
NI-IND
NI-UNIND
Inhibitors

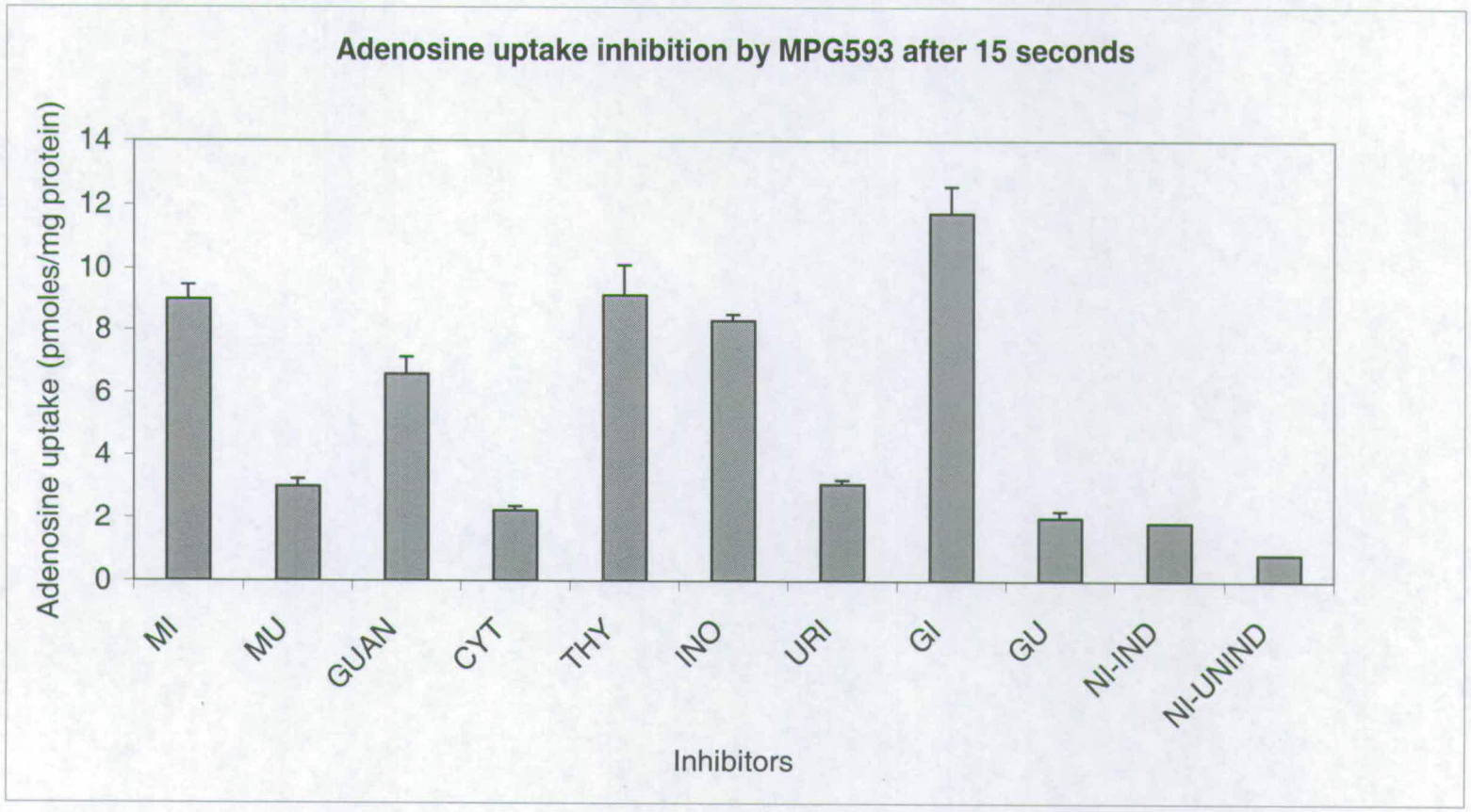
B.


Adenosine transport inhibition by MPG592 after 15 seconds
Adenosine uptake (pmoles/mg protein)
14
12
10
8
6
4
2
0
TI
TU
GUAN
CYT
THY
INO
URI
GI
GU
NI-IND
NI-UNIND
Inhibitors

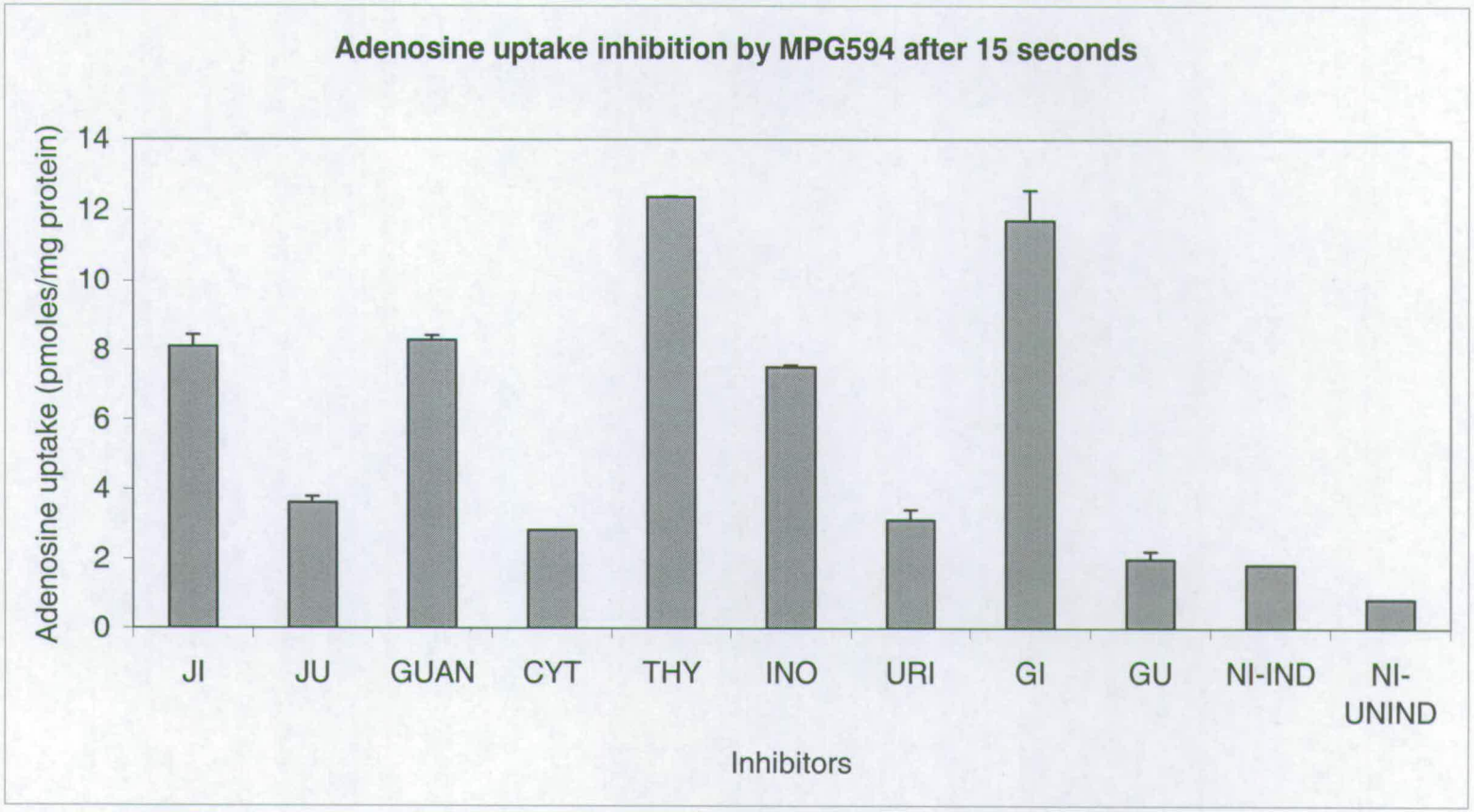
C.


Adenosine transport inhibition in MPG592 after 2 minutes
Adenosine uptake (pmoles/mg protein)
18
16
14
12
10
8
6
4
2
0
TI
TU
GUAN
CYT
THY
INO
URI
GI
GU
NI-IND
NI-UNIND
Inhibitors

D.


Adenosine uptake inhibition by MPG593 after 15 seconds
Adenosine uptake (pmoles/mg protein)
0
2
4
6
8
10
12
14
MI
MU
GUAN
CYT
THY
INO
URI
GI
GU
NI-IND
NI-UNIND
Inhibitors

E.


Adenosine uptake inhibition by MPG594 after 15 seconds
Adenosine uptake (pmoles/mg protein)
0
2
4
6
8
10
12
14
JI
JU
GUAN
CYT
THY
INO
URI
GI
GU
NI-IND
NI-UNIND
Inhibitors

F.

	Purines		Pyrimidines		
Nucleoside	**Inosine**	**Guanosine**	**Thymidine**	**Cytidine**	**Uridine**
Overexpressed transporter					
XapB	↓↓	↓	↓	—	—
YegT	↓	↓	↓	↓↓	↓↓
YeiM	—	↓	—	↓↓	↓↓
YeiJ	—	—	↑	↓↓	↓↓

Table 4.2: Summary of the effect of nucleosides on adenosine transport. Adenosine was prepared at 50μM, 1μCi/ml, and uptake was determined at 25°C after 15 seconds. Overexpressed transporters refer to the gene cloned in plasmids pMRB, pMRT, pMRM and pMRJ. This table summarises key results obtained from nucleoside inhibition tests on the strains as shown in figure 4.4 A-F.

KEY:
- ↓↓ Strong inhibition of adenosine transport (more than 50%)
- ↓ Poor inhibition of adenosine transport (less than 50%)
- ↑ Increased adenosine transport
- — No effect on adenosine transport

4.6 Utilisation of nucleoside as sole C-source

E. coli is known to be able to scavenge nucleosides from its environment and utilise them as sole carbon sources for growth. This is true for most wild type *E. coli* strains and single enzymes have been reported to catalyze the cleavage of adenosine, guanosine and inosine (Jensen and Nygaard, 1975). To further explore the ability of XapB, YeiM, YeiJ and YegT to transport nucleosides, the capacity of cells carrying one of the genes to grow on nucleoside as the sole carbon source was tested.

The monogene strains MPG578, MPG579, MPG580, MPG581, MPG582, MPG583 encoding functional genes *xapB, yegT nupG, yeiM, nupC,* and *yeiJ* respectively and in addition MPG584 (null strain) and wild type *E. coli* MG1655, were grown in minimal media containing one of a range of nucleosides at a final concentration of 5mM. Cultures were prepared from single colonies in a bijou bottle containing minimal media with glucose as carbon source and incubated overnight at 37°C. These overnight cultures were then used to inoculate fresh minimal media in bijou bottles containing nucleosides as the sole carbon sources, which were left to grow at 37°C for 2 hours. 200µl aliquots were then transferred in triplicate into a 96 well plate and were placed into an ELISA reader in Edinburgh University's COIL facilities choosing the settings so that the reader (ELISA microplate reader, Alpha Diagnotics) would take measurements at OD_{600} every 20 minutes for an appropriate number of hours (usually 14 to 27 hours). The 96 well plates were incubated at room temperature without shaking. The readings were then obtained using KC5 software (KCSoftwares). For more details on the specific components of the media refer to the Materials and Methods chapter.

This experiment allowed a high throughput approach to testing a number of stains in combination with a number of media containing different nucleosides. This experiment also allowed analysis of whether the null strain is indeed negative for transport of nucleosides under the examined conditions. This provides a basis for using this strain for further experiments where uptake of toxic nucleosides such as AZT could be examined.

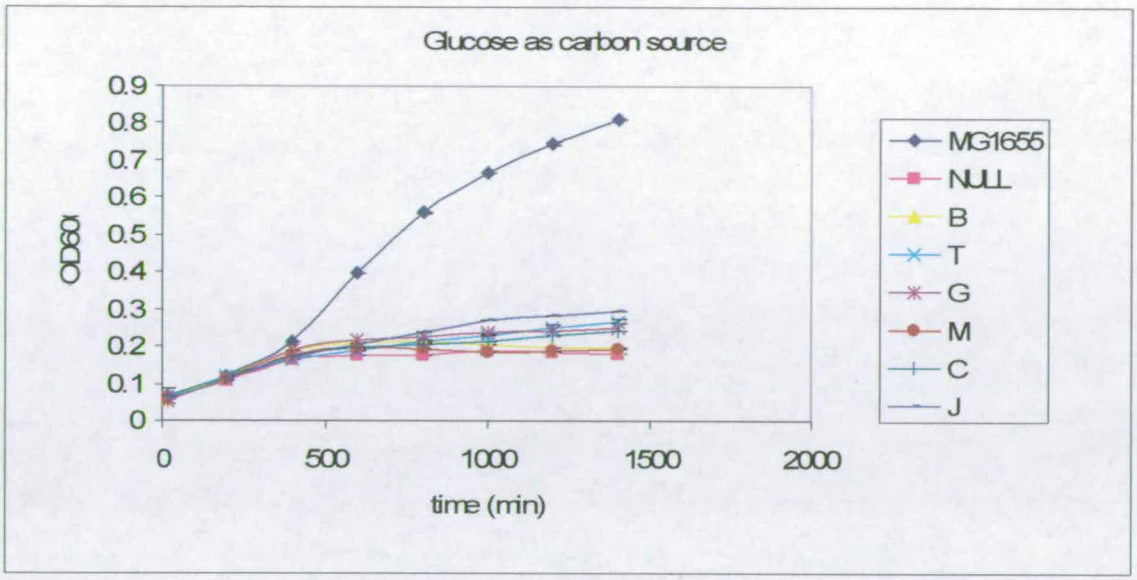
As controls and to allow comparison of the resulting data, the monogene strains, the wild type and the null strain were also grown under the same conditions of minimal media but with glucose as the sole carbon source (0.26% w/v). The nucleosides tested as carbon sources in the minimal media were inosine, adenosine, thymidine, uridine, and cytidine. Also xanthosine was tested but was suspended in minimal media at a pH of 6.6 due to solubility problems at neutral pH. An extra control was included of minimal media with glucose at pH 6.6 to check growth inhibition due to the pH rather than the carbon source. Finally, due to solubility issues guanosine was tested only to a final concentration of 1mM which was not high enough to support growth (data not shown).

As seen from the graphs in figure 4.5A, all strains including the null grew in minimal media with glucose as the carbon source. Notably wild type strain *E. coli* MG1655 grew much better than the mutated strains. The non-mutated strain MG1655 also grew albeit slowly in minimal media with each of the nucleoside supplements examined as sole carbon source. On the other hand the null strain seemed unable to take up nucleosides from the media and utilise them as carbon source for growth under the conditions examined. This probably reflects the inability of the null strain to transport nucleosides.

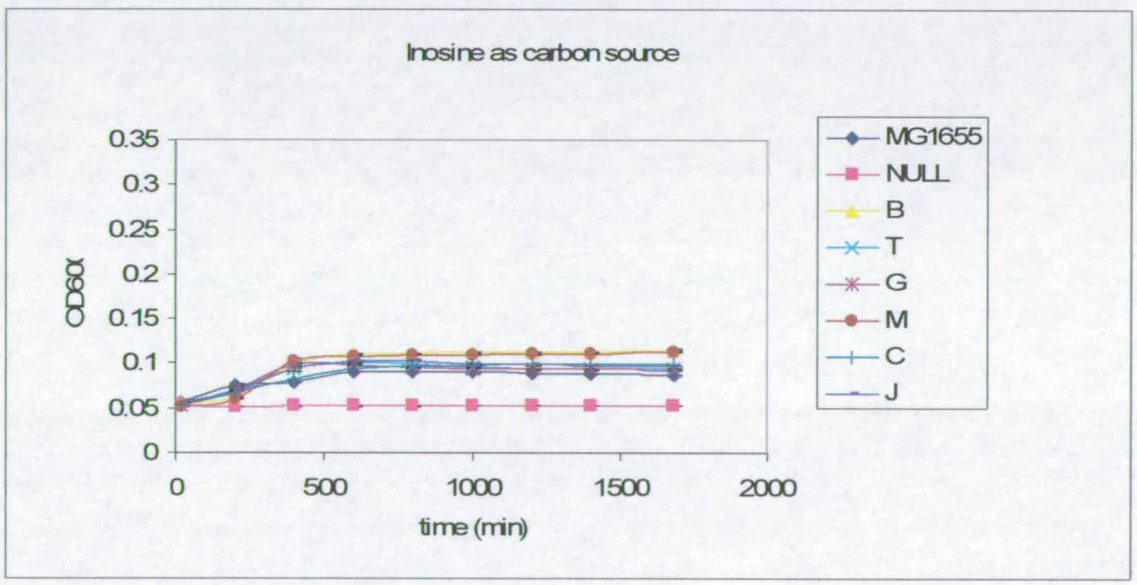
The transporter NupC was found to support growth when cytidine or xanthosine were used as a carbon source in the media however its ability to transport any of the other nucleosides tested was not demonstrated in this work. The NupC monogene strain would be expected to grow better on cytidine, however there may be regulation issues under these conditions of growth that cannot be controlled for in this study. NupG demonstrated its ability to allow uptake of a variety of nucleosides as sole carbon source since the monogene strain MPG580 was also able to grow using adenosine, thymidine, cytidine, uridine, and more poorly, inosine. Its ability to utilise cytidine as a carbon source is in line with the information gathered by other researchers meaning that NupC is able to utilise cytidine while NupG has broader specificity and it can transport both cytidine and guanosine (Munch-Petersen *et al,* 1979).

As far as the putative transporters are concerned, adenosine as a sole carbon source supported growth of all monogene strains. Adenosine use as sole carbon source was good when transported by nucleoside transporters XapB, NupG and YeiM. Thymidine as a carbon source supported the growth of the monogene strain containing the YeiM transporter, while cytidine only supported growth for NupC and NupG as mentioned above. Uridine as sole carbon source exhibited the ability to support growth of the monogene strain harbouring NupG at levels as good as the MG1655 strain, however none of the other monogene strains exhibited prominent growth. Furthermore inosine as carbon source gave poor growth for all strains. However MG1655 and monogene strains harbouring XapB and YeiM did grow slightly better than the rest of the strains. The monogene strain harbouring XapB grew well in minimal medium with xanthosine as a carbon source despite the fact that the medium pH was lowered to 6.6, as opposed to the neutral pH of 7 used in all the other minimal medium-nucleoside combinations, due to solubility problems. It was important to include an extra control of the strains growing in minimal medium with glucose as the sole carbon source and the pH lowered to 6.6 for this experiment (figure 4.5 panel G) so as to allow direct conclusions to be made for the effect of the lowered pH. In fact all monogene strains grew well when xanthosine was the sole carbon source, particularly those with transporters YeiJ and YegT. A summary of the findings of this study can be seen in Table 4.3.

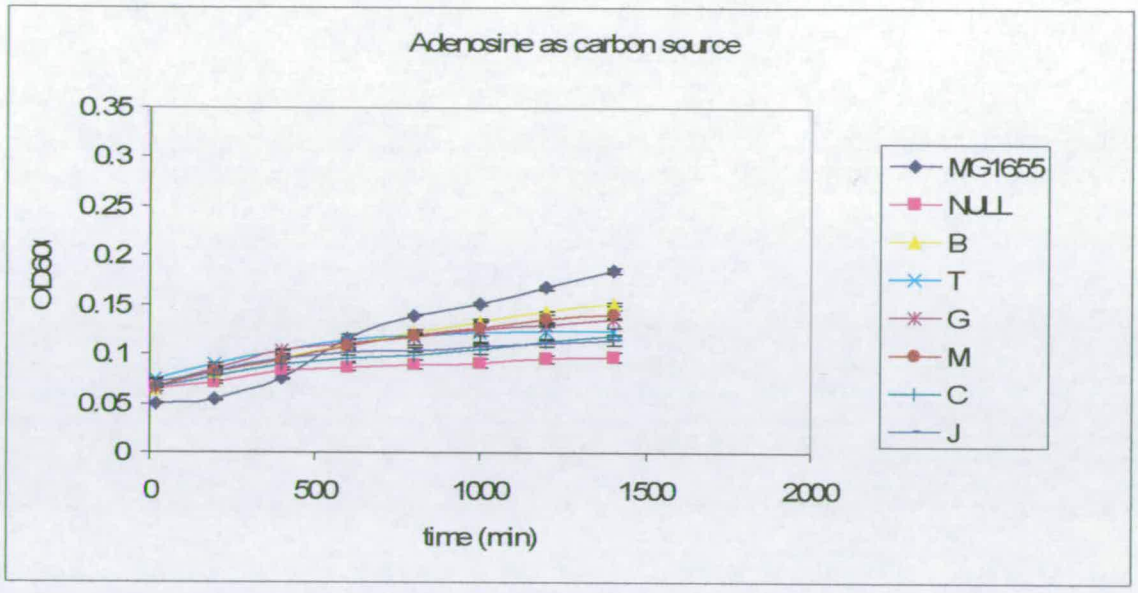
Figure 4.5 (A-I): Minimal media with nucleoside as carbon source growth curves for MG1655, the null, and monogene strains. MG1655, MPG584 (null) and monogene strains MPG578, MPG579, MPG580, MPG582, MPG581 and MPG583, the null and non-mutated parental strain MG1655 were grown in minimal media with various nucleosides as the carbon source to a final concentration of 5mM. All strains were also grown in minimal media with glucose as the carbon source as a positive control. In the case of xanthosine due to solubility problems, the pH had to be kept slightly lower than normal, decreasing to pH 6.6. For this reason an extra control was used, growing all strains in minimal media with glucose as the carbon source and adjusting the pH at 6.6. In the figure legends the initials B, T, G, M, C, J symbolise monogene strains with genes *xapB*, *yegT*, *nupG*, *yeiM*, *nupC* and *yeiJ* intact respectively (monogene strains MPG578, MPG579, MPG580, MPG581, MPG582, MPG 583).


Glucose as carbon source
OD600
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
0
500
1000
1500
2000
time (min)
MG1655
NULL
B
T
G
M
C
J

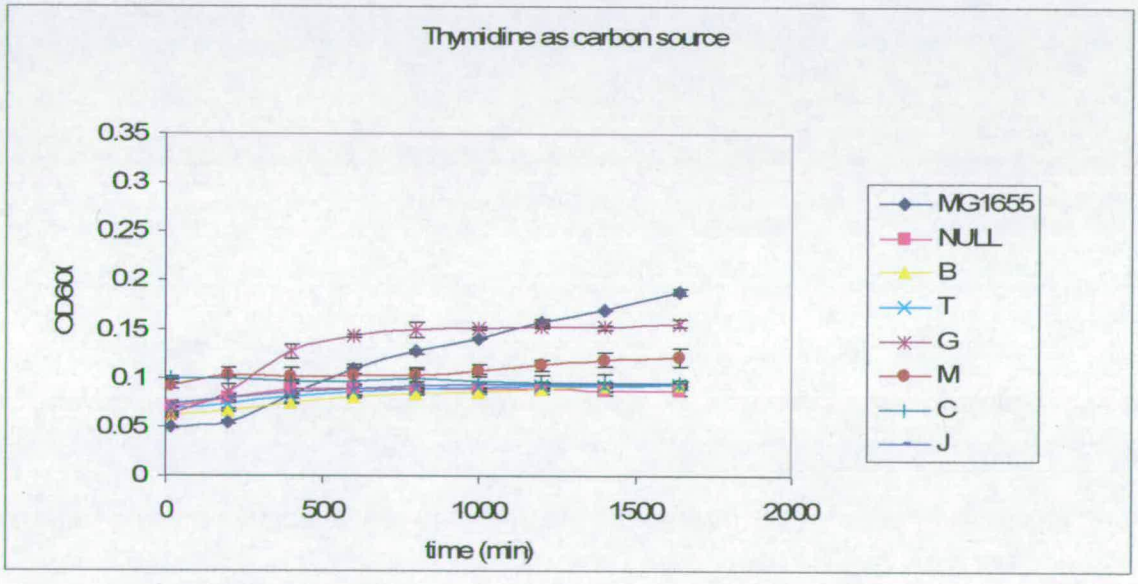
A.


Inosine as carbon source
OD600
0.35
0.3
0.25
0.2
0.15
0.1
0.05
0
0
500
1000
1500
2000
time (min)
MG1655
NULL
B
T
G
M
C
J

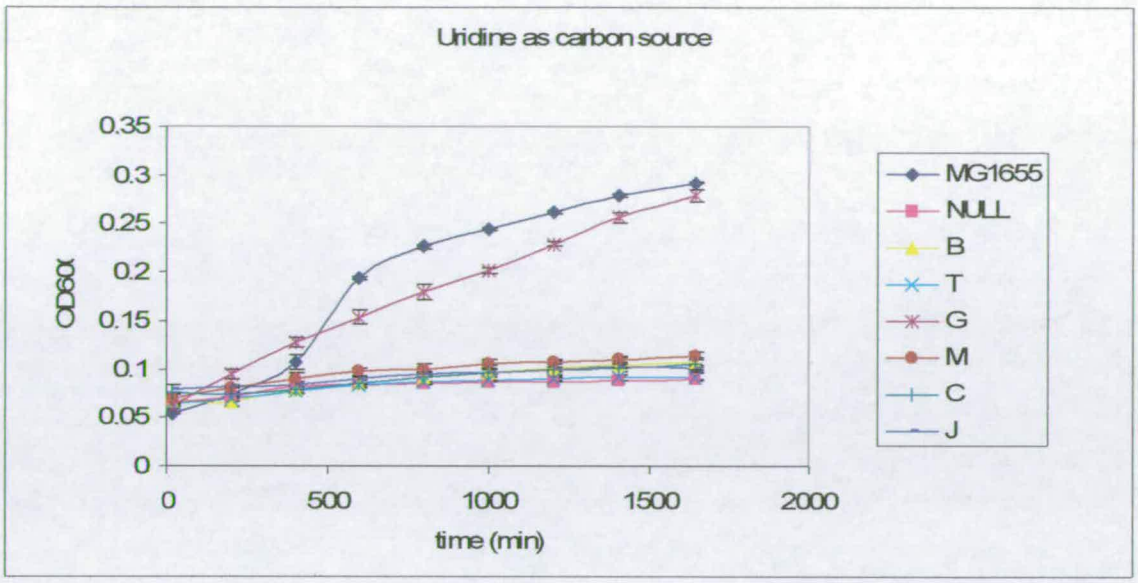
B.


Adenosine as carbon source
OD600
0.35
0.3
0.25
0.2
0.15
0.1
0.05
0
0
500
1000
1500
2000
time (min)
MG1655
NULL
B
T
G
M
C
J

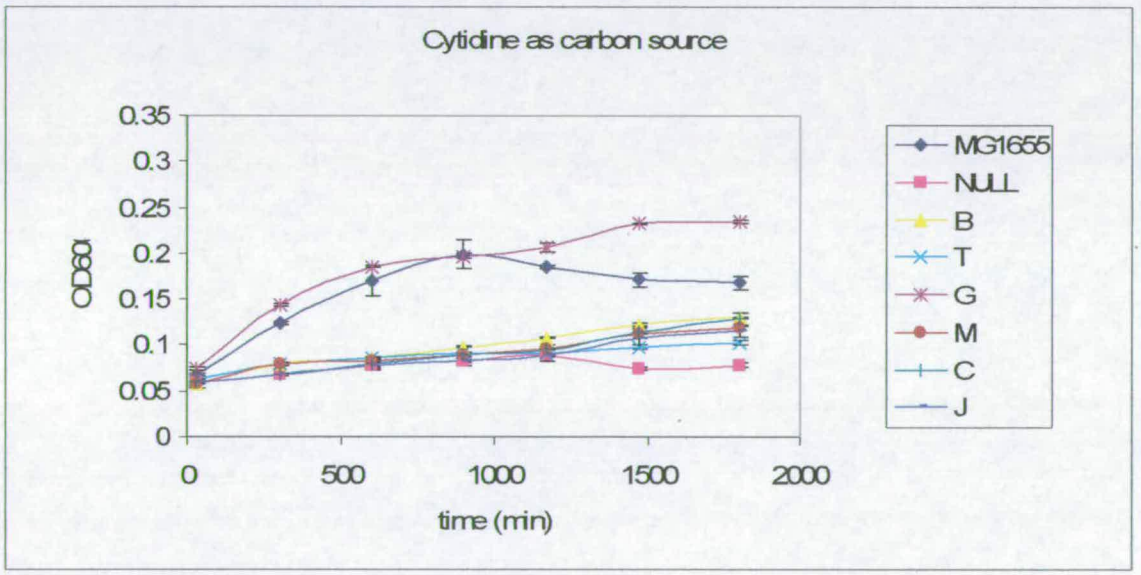
C.


Thymidine as carbon source
OD600
0.35
0.3
0.25
0.2
0.15
0.1
0.05
0
0
500
1000
1500
2000
time (min)
MG1655
NULL
B
T
G
M
C
J

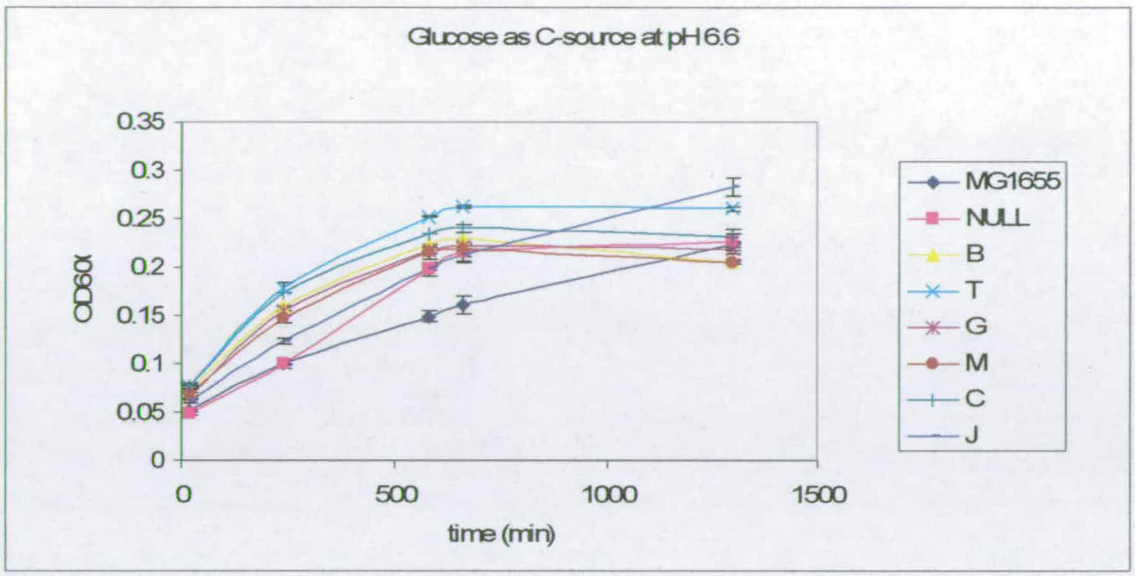
D.


Uridine as carbon source
OD60
0.35
0.3
0.25
0.2
0.15
0.1
0.05
0
0
500
1000
1500
2000
time (min)
MG1655
NULL
B
T
G
M
C
J

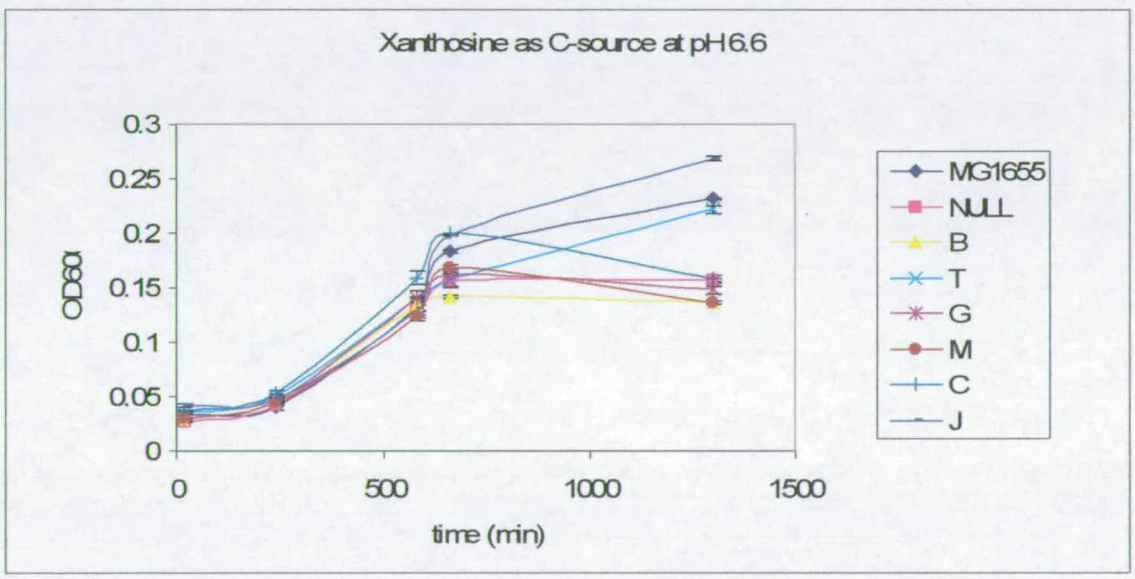
E.


Cytidine as carbon source
OD60
0.35
0.3
0.25
0.2
0.15
0.1
0.05
0
0
500
1000
1500
2000
time (min)
MG1655
NULL
B
T
G
M
C
J

F.


Glucose as C-source at pH6.6
OD60
0.35
0.3
0.25
0.2
0.15
0.1
0.05
0
0
500
1000
1500
time (min)
MG1655
NULL
B
T
G
M
C
J

G.


Xanthosine as C-source at pH6.6
OD60
0.3
0.25
0.2
0.15
0.1
0.05
0
0
500
1000
1500
time (min)
MG1655
NULL
B
T
G
M
C
J

H.

E. coli strain	MG1655	NULL	B	T	G	M	C	J
Carbon source								
Glucose	*****	****	****	*****	****	*****	****	****
Adenosine	****	**	***	**	***	***	**	**
Thymidine	****	-	-	-	***	**	-	-
Cytidine	****	-	-	-	****	-	**	-
Uridine	*****	-	-	-	*****	**	-	**
Inosine	**	-	**	-	-	**	-	-
Xanthosine	*****	***	***	****	*****	***	****	*****

Table 4.3 Summary of growth of monogene, null and MG1655 strains in minimal media with nucleosides as sole carbon sources. The bacteria were grown in minimal media with glucose (0.26% v/v) or nucleosides as sole carbon sources in a 96 well microplate, at room temperature without shaking for approximately 27 hours at final concentration of 5mM. This table summarises key results obtained from growth patterns of the strains as shown in figure 4.5 A-H.

KEY: ***** Excellent growth
**** Good growth
** * Weak growth
** Poor growth
- No growth

Initials B, T, G, M, C, J symbolise monogene strains with genes *xapB*, *yegT*, *nupG*, *yeiM*, *nupC* and *yeiJ* intact respectively

4.7 Discussion

Plasmids with the unknown or putative nucleoside transporter genes *xapB, yeiM, yeiJ* and *yegT* were constructed in order to perform experiments in an attempt to gain knowledge of their specificity for nucleoside transport in *E. coli* strains carrying these plasmids were used in radiolabelled nucleoside assays which give information on uptake, kinetic features and specificity of the transporters. Furthermore the monogene strains, as well as the null and an *E. coli* strain in which the genes encoding the putative transporters had been specifically deleted, were used in experiments where nucleoside utilisation as the sole carbon source for growth was examined.

Nucleoside uptake above background was shown for strains expressing all four unknown or putative nucleoside transporters with use of radiolabelled adenosine as a substrate. This indicated that all of these transporters are in fact able to transport adenosine, although other conditions would need to be examined to determine whether it is the preferred or natural principal nucleoside substrate.

Km values were determined as 97.1, 65, 160.1 and 453.9μM for transporters XapB, YegT, YeiM and YeiJ respectively. For transporter YeiJ the Km given is only estimated since the saturation point was not reached in this case. Therefore it can be assumed that the constant for YeiJ will be within the same order of magnitude as for the transporters but higher than the others. Two types of transporters can be envisaged, one with high affinity usually with a medium V_{max} for scavenging and a second one with low affinity but with a high V_{max} for transporting in an environment of substrate excess. It is unclear which category applies for YeiJ, whilst the other three transporters are likely to represent the first category.

The Km previously calculated for related transporters NupC and NupG was between 0.3μM in whole cells (Komatsu and Tanaka, 1973) and the Km for transport of cytidine and deoxycytidine in membrane vesicles was estimated to be between 0.5 and 0.8μM

(Munch-Petersen *et al,* 1979). Also the Km of NupC as determined with adenosine as a substrate in *Xenopus* oocytes was found to be 1.3μM (Loewen *et al,* 2003). Additionally Nørholm and Dandanell (2001) determined the Km of XapB to be in the same order of magnitude as mammalian transporters. Km for XapB was found to be 136μM in whole cell preparations with xanthosine as substrate.

Similarly the Km values determined for the putative transporters in this study are in the same order of magnitude as mammalian transporter kinetics determined in other studies (table 4.1). Specifically the rat jejunal epithelium nucleoside transporter rCNT1 was found to have a Km of 37μM with uridine as a substrate (Huang *et al,* 1994) and rat jejunal/ kidney nucleoside transporters rCNT1 and rCNT2, with adenosine as a substrate displayed a Km of 26μM (Yao *et al,* 1996). Human transporters hCNT1, hCNT2 and hCNT3 have been tested with a variety of substrates obtaining substrate and kinetic information. Namely the Km value for hCNT1 with cytidine as substrate has been shown to be around 140μM (Lostao *et al,* 2000) hCNT2 has been shown to have a Km of 8 and 116 μM with substrates adenosine and uridine respectively (Cass *et al,* 1998, Wang *et al,* 1997) and hCNT3 has been tested with a variety of substrates in a study by Ritzel *et al* (2001a, b) determining Km values ranging from 15.1μM (with adenosine as substrate) to 52.5μM (with Inosine as substrate). Also from work done on the hagfish nucleoside transporter *hf*NCT1 with a variety of substrates, Km values were determined to be 10μM, 35μM and 45μM for substrates uridine, inosine and thymidine respectively (Yao *et al,* 2002). A good understanding of all CNT and ENT Km values, including work yet to be completed, as well as the most recent updates is given in a review by Kong *et al* (2004).

More information on reciprocal kinetics and the behaviour of the transporters with other substrates would be of interest in order to investigate if other substrates are inhibited by nucleosides added in the cell suspension and check if transporters such as YeiM and YeiJ have higher affinities and different kinetic patterns when the substrate is a pyrimidine (Carter *et al,* 1995). It is of note that YeiM and YeiJ show homology to

NupC and the CNT family. However whilst NupC has preference for pyrimidines, this is not so for all members of the CNT family (Cass *et al*, 1998).

Adenosine transport by each one of the four transporters was found to be inhibited in experiments involving a series of nucleosides. In particular adenosine transport via XapB was strongly inhibited by inosine (more than a 70% decrease in adenosine transport), while YegT- mediated adenosine transport was greatly affected by cytidine and uridine (70 and 60% decrease in adenosine transport respectively). Adenosine transport via YegT and XapB was inhibited by all tested nucleosides indicating that products of *yegT* and *xapB*, which are *nupG* paralogues, have broad substrate specificity. Adenosine transport by transporters YeiM and YeiJ was inhibited by cytidine and uridine by more than 60%. *yeiM* and *yeiJ* are paralogues of *nupC* which encodes a permease with preference for pyrimidines (Nørholm and Dandanell, 2001). Therefore the fact that YeiM- and YeiJ- mediated adenosine transport were mostly inhibited by pyrimidine nucleosides is an indication of the fact that they might have greater affinity for pyrimidines as substrates.

To add information on these findings it would be suggested that the transporters' ability to transport nucleosides should be tested by using other nucleosides as substrates to establish a more refined hierarchy of substrate specificity.

From the experiments where nucleosides were used as the sole carbon source for growth with minimal media it can be deduced that all the examined genes are able to scavenge nucleosides from their external environment. Specifically the monogene strain containing functional gene *xapB* grow poorly on minimal medium with inosine as the sole carbon source, however better than other strains. The competition experiments showed that XapB-mediated transport of adenosine was greatly reduced by the presence of inosine. The monogene strain MPG578 harbouring XapB grew well in minimal medium with xanthosine as a carbon source despite the fact that the pH was lowered to 6.6, however the strain harbouring YegT grew better than the XapB monogene strain.

Monogene strain MPG579 containing functional gene *yegT* grew poorly on most minimal medium-nucleoside combinations. This might be because *yegT* is under different regulatory control than *xapB* or because the two encoded proteins although appearing to have similar characteristics and specificities are in fact different in their specificities, affinities and optimal nucleoside concentration in which they concentrate nucleosides in the cell. Also monogene strains with functional *xapB* or *yegT* grew in minimal adenosine medium.

The monogene strain with *yeiM* as the sole functional nucleoside transporter gene was able to grow on minimal adenosine, uridine and thymidine. Finally monogene strain with *yeiJ* intact grew on uridine. These results are in line with the findings from the competition experiments in which YeiM and YeiJ favour transport of uridine.

E. coli cells are known to be able to scavenge nucleosides from their environment (Mygind and Munch-Petersen, 1975) and the fact that the monogene strains were able to grow to an extent under minimal conditions with nucleosides as a carbon source suggests that the genes being studied might play a role in the scavenging process of nucleosides. However such experiments would be difficult to repeat in other micro organisms such as *Bacillus subtilis* due to their slow catabolism (Schuch *et al,* 1999).

Some controversial results were obtained in the study of transporters XapB, YegT, YeiM and YeiJ. In particular in the case of YeiM thymidine and inosine supported poor growth of the monogene strain MPG581 however they did not inhibit adenosine uptake in the competition experiments. This could be because thymidine and inosine are able to support growth of the particular strain at the elevated concentration of 5mM but at the concentration of 0.5mM in the competition assay they do not have such an obvious effect. Generally in the case of YeiM it seems that it can be expressed as the sole transport system and it can support growth on all nucleosides except cytidine. However cytidine was the nucleoside that inhibited adenosine uptake more strongly in the competition assay. This could be because the high concentration of cytidine in the

growth experiments actually inhibits growth although it can be transported by YeiM. A different explanation is that for the transport of cytidine the other transport systems (for example NupC and / or some of the others) are necessary for the initial uptake of small amounts of cytidine and that after that full induction of the other systems (such as YeiM) can occur. Therefore in the competition experiments where YeiM is overexpressed but there are copies of all six nucleoside transporters cytidine is transported into the cell and inhibits adenosine uptake. Indeed cytidine did not support growth in any of the monogene strains except poor growth of MPG582 which has only NupC intact. The poor growth could be attributed to regulation issues which are not investigated in this work.

On the other hand high concentrations of cytidine reduced growth of wild type strain MG1655 as compared to growth of MG1655 in other nucleosides. This could be an indication that although the transporters are able to transport cytidine very efficiently, cytidine at 5mM is too high a concentration for the bacterial strain to grow normally.

In the case of YegT, although in competition experiments all tested nucleosides inhibited uptake of adenosine, the YegT monogene strain MPG579 only grew when xanthosine or adenosine were the carbon source. This could be explained by the fact that the lack of the other nucleoside transporters in the case of the monogene strain abolishes growth if one or some of the other transporters are necessary for uptake of the nucleosides. However it does seem that YegT is able to be expressed and concentrate adenosine and xanthosine so as to support growth. Thymidine did not support growth of strain MPG579 ($YegT^+$) or indeed monogene stains expressing any other nucleoside transporter except YeiM and NupG, however it did inhibit adenosine transport in the competition assays in strains overexpressing YegT and XapB. This could be because thymidine requires at least one of the two systems in order to initially transport thymidine into the cell and then express some of the other nucleoside systems. Thymidine also was shown to increase adenosine uptake when added in the competition assay when YeiJ was the nucleoside transporter overexpressed. This could mean that there are multiple binding epitopes for adenosine, and possibly other nucleosides that

were not determined here, that in the presence of thymidine are expressed and bind adenosine at higher levels.

Although not supporting growth, uridine also inhibited adenosine uptake in the cases of YegT and XapB being the nucleoside transporters. This is also another case in which more than one transporter may be required to initially allow uptake of uridine and then express other nucleoside transport systems such as YegT and XapB. Alternatively this could be an indication that at 5mM uridine can be toxic to cells expressing these systems because they concentrate thymidine at high quantities in the cell. The NupG monogene strain on the other hand was able to grow normally on uridine at similar high enough levels as the wild type strain MG1655. In recombinant *E. coli* cells NupG transported uridine with apparent Km of 20-30 μM showing a low affinity system (Xie *et al,* 2004). This is further indication that initially NupG takes up uridine and other uptake systems could be induced as a secondary mechanism to deal with high concentrations of uridine and then use it as an energy source if needed.

Finally inosine although inhibiting adenosine uptake at 0.5mM in the case of YegT being overexpressed did not support growth of the YegT monogene strain MPG579 in the growth experiments. However growth on inosine as a sole carbon source was not good as a whole for any of the strains including the wild type strain MG1655. This could be because inosine is transported into the cell but has a toxic effect on strains' growth at such high concentrations.

Therefore although the monogene strains provide a set of strains which are good working models because a clear result can be obtained on the individual role of each one of the transporters, the experiments are lacking insight in the effect the transporters have on each other. Also knowledge is necessary on the regulatory mechanisms of action of the transporters in order to further elucidate their role, and the impact different nutrients and substrates have on the expression of each one or groups of them. It should be also noted that experiments on utilisation of nucleosides as the sole carbon source were

performed on the monogene strains which have only a single copy of the gene as the gene is not a plasmid vector, and therefore the levels of expression are low thus giving a low signal to be measured in the experimental procedure

.

On the other hand uptake assays, kinetics assays and competition experiments were performed in strains which had one copy of the whole set of genes and additionally had one of the genes overexpressed. Although this experiment gives a strong sense of direction on the overexpressed gene and whether it can utilise a specific substrate, it is fundamentally different from attempting to fine-tune differences between the six genes that are studied, since natural cellular operations in terms of possible undetected nucleotide transport by other systems may occur in the background of the strains since the whole set of genes is present.

Nevertheless despite the special conditions of these experiments, the results obtained can be used as a frame of reference and as a preview of the fact that the transporters studied are indeed able to transport nucleosides, a fact which coupled with the findings of the radiolabelled nucleoside experiments reflects the strong prospect of these transporters being nucleoside permeases or having a key role in nucleoside transport.

Chapter 5

Development of a high throughput bioassay for nucleoside analogue assessment

5.1 Introduction

In the previous chapters it was described how nucleoside transporter negative strains for genes *nupC, nupG xapB, yeiM, yeiJ* and *yegT* were constructed. The construction process by which multi-knock out strains that are mutated for all but one of these genes, was described. These strains were used in studies to establish that genes *xapB, yeiM, yeiJ* and *yegT*, which encode unknown or putative nucleoside transporters are involved in nucleoside transport. Also, strains which are mutated for known genes *nupG* and *nupC*, were used in experiments to confirm their functions as nucleoside transporters.

The primary objective of the current chapter was to establish whether a bioassay could be developed which would have the potential to be used in a high throughput manner, to assess the effect of toxic nucleoside analogues in transport-negative strains. A bioassay which uses *E. coli* monogene strains could be easily and quickly manipulated as a first screening method to examine a range of newly developed nucleoside analogue drugs, whose function is not fully understood, for efficacy and toxicity levels.

Pharmaceutical companies and chemists have developed a large number of novel nucleoside analogues which can potentially be used as chemotherapeutic and antiviral agents. The drug development industry is fast moving and the use of high throughput processes has permitted medical chemists to prepare massive libraries of compounds that can be funnelled through high throughput assays to identify lead candidates for drug development (Gardner and Fernandes, 2004). However the approval of new chemotherapeutic agents is declining because selectivity and potency are requisite but not sufficient criteria for a successful drug candidate. Other criteria include validation of the biochemical target in the disease state, safety profiles, appropriate pharmacokinetic and metabolic properties of the candidate drug, and chemical characteristics of the compound that ensure the compounds can be synthesised and produced on a large scale, in an economical way (Gardner *et al,* 2004).

Novel high throughput methodologies of screening drugs have a central role in assessing the physical and chemical properties of candidate drugs that may become new therapeutic drugs. In order to facilitate high throughput screening there is a need for quick ways of assessing whether a new drug would be useful. One way of approaching this issue is by the development of a first screen which would determine rapidly whether any new drugs in question are effective and would qualify to move to the next level of assessment tests.

The null strain MPG584, which is a transport negative strain for nucleosides, was therefore examined as a potential candidate for the formation of a bioassay since it fulfils a number of requirements for test development. First of all it is a bacterial model, and therefore it can be used rapidly and in a high throughput manner to assess drug transport into the cell. Secondly, the apparent lack of nucleoside transporters in the null strain makes it a unique candidate to assess uptake of nucleoside analogues. The rationale behind this is that theoretically, this strain would not be able to actively transport any nucleosides or their toxic analogues. If a drug is able to pass through the membrane of the null strain it must have other mechanisms of transport e.g. passive diffusion or partitioning. Furthermore the isogenic monogene strains can be examined in parallel to narrow down the particular mode(s) of entry of specific nucleoside analogues and therefore provide some insight into the way these drugs are taken up by the cell.

In order to develop these strains as candidates for a high throughput test, a way of assessing the cells' viability was needed. The viability of the cell is a direct indication of the cell's ability to take up the toxic nucleosides.

Two forms of bioassay were examined for use in a high throughput manner using a 96 well plate. Firstly a bioassay based on bioluminescence as a direct indication of whether the cell was able to take up toxic nucleosides, such as AZT, was explored. Alternatively the possibility of using optical density, which is a direct indicator of cell abundance, was

examined to see if the toxic effect of an analogue could be measured reliably after administration.

A key value of using bioluminescence as opposed to optical density (OD) in the assay is that it reflects cell viability because it requires regeneration of a cofactor and so may make the bioassay more appealing for commercialisation. Bacterial luciferase is a popular choice as a reporter system because functional enzyme is created upon translation and the assay is quick, reliable and easy to perform. Furthermore it is well suited to laboratory automation and high throughput applications.

The *lux* system is a sensitive and self contained reporter system capable of monitoring transcriptional activity within living cells where the aldehyde substrate is produced constitutively in the cell. (Naylor, 1999). This system is made up of five key structural genes *luxCDABE* and is derived from the bioluminescence operon of the marine bacterium *V. fischeri.* The *luxCDE* genes encode an enzyme complex (fatty acid reductase, synthetase and transferase) that synthesises the substrate (an aldehyde) for luciferase. Then the *luxAB* genes encode the luciferase enzyme. The luciferase activity from *V. fischeri* utilises $FMNH_2$ and O_2 and is reported to be stable at temperatures up to approximately 33°C, above which the enzyme loses its activity (reviewed in D'Souza, 2001a and 2001b). For this reason all strains were grown at 30°C for OD and light emission work, even though MG1655 grows more rapidly at 37°C.

AZT was among the first of the dideoxynucleosides to demonstrate anti-HIV activity *in vitro* by inhibition of the viral reverse transcriptase (RT) and it has become a standard for anti-HIV therapy. AZT was first administered to individuals with severe HIV infection at the NCI and Duke University Medical Centre in 1985 (Yatvin, *et al,* 1999). These initial studies showed that the patients underwent immunological, virological and clinical improvement during therapy. AZT is therefore now a common pharmacological agent that is used in HIV monotherapy or in combination with other nucleoside analogues, such as ddC and ddI.

AZT was used as a model toxic nucleoside compound in the present study to give an idea of how strains might behave during the assay. Also AZT was used to assess the two different assay systems and help decipher which one is more reliable and effective.

In the following sections both bioluminescence and OD are analysed as ways to assess the impact of toxic nucleosides on cell growth, as an attempt to develop a bioassay.

5.2 Measuring bioluminescence in a growing culture for strains MPG560, MPG567 and monogene strains MPG566, MPG564, MPG565, MPG561, MPG562, MPG563

In order to understand light emission patterns in the strains used in this study the monogene strains as well as the null and non mutated strain MG1655 were used as template strains for the construction of a luminescence-based bioassay, so as to study the uptake of nucleoside analogues by the nucleoside transporter-encoding genes, in a high throughput manner. For that purpose strains MG1655 (wild type), MPG584 (null) and monogene strains MPG578 (*xapB*$^+$), MPG579 (*yegT*$^+$), MPG580(*nupG*$^+$), MPG582(*nupC*$^+$), MPG581(*yeiM*$^+$) and MPG583(*yeiJ*$^+$) were transformed with a plasmid that bears the *lux* operon from *V. fischeri*. This plasmid is called pucΔ320 (Shaw and Kado, 1986) and will express functional luciferase via the *lux* operon as long as the host strain grows at a temperature lower that 33°C. The plasmid-containing strains were named MPG560 (no nucleoside transporter mutations), MPG567 (null), MPG566 (*xapB*$^+$), MPG564(*yegT*$^+$), MPG565(*nupG*$^+$), MPG561(*nupC*$^+$), MPG562 (*yeiM*$^+$) and MPG563 (*yeiJ*$^+$) and were derived from MG1655, MPG584, MPG578, MPG579, MPG580, MPG582, MPG581and MPG583, respectively.

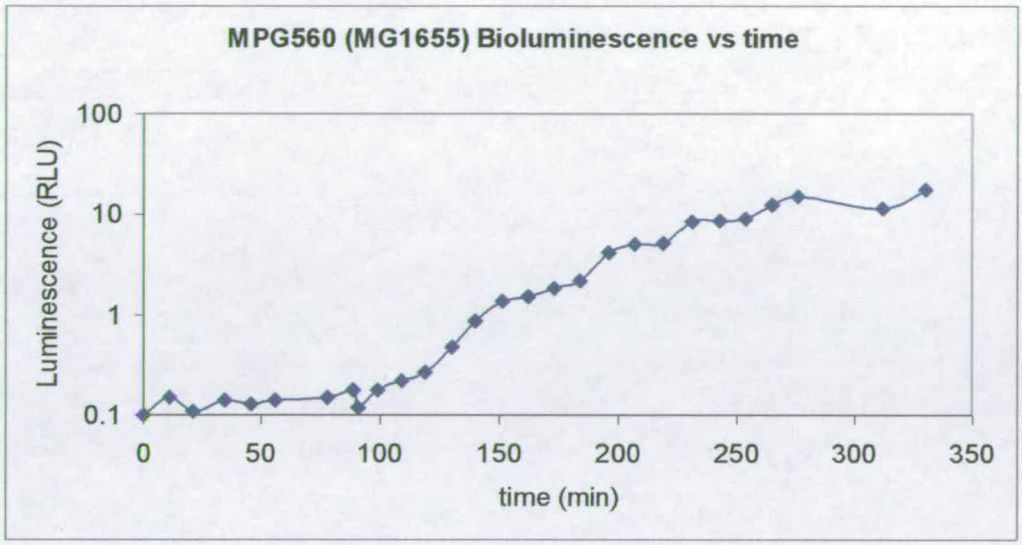
In order to evaluate light emission levels in these strains, 1 ml of over night cultures of MPG560, MPG567 and the monogene strains, MPG566, MPG564, MPG565, MPG561, MPG562, MPG563 were grown in 250ml flasks, containing 50ml of LB at 30°C, with vigorous shaking (1000 rpm). Light emission was measured by taking 50μl aliquots and measuring bioluminescence in a 96 well plate which was static and at room temperature. The sample was taken from the culture while it was shaking i.e. the culture remained shaking while the experiment was taking place, so as to ensure that the measurements taken were representative. Aliquots were taken roughly every 20 minutes and, as described in section 5.3, the growth of the strains was monitored by measuring the OD_{600} in the spectrophotometer at the same time. The values measured were then plotted in graphs as shown in figure 5.1 (A-H). This set of experiments gathered data to help

understand the bioluminescence patterns of the strains used. However for the purposes of a bioassay it was also necessary to measure the AZT effect on these strains in a high throughput manner, in a 96 well plate format-see section 5.4 of this chapter.

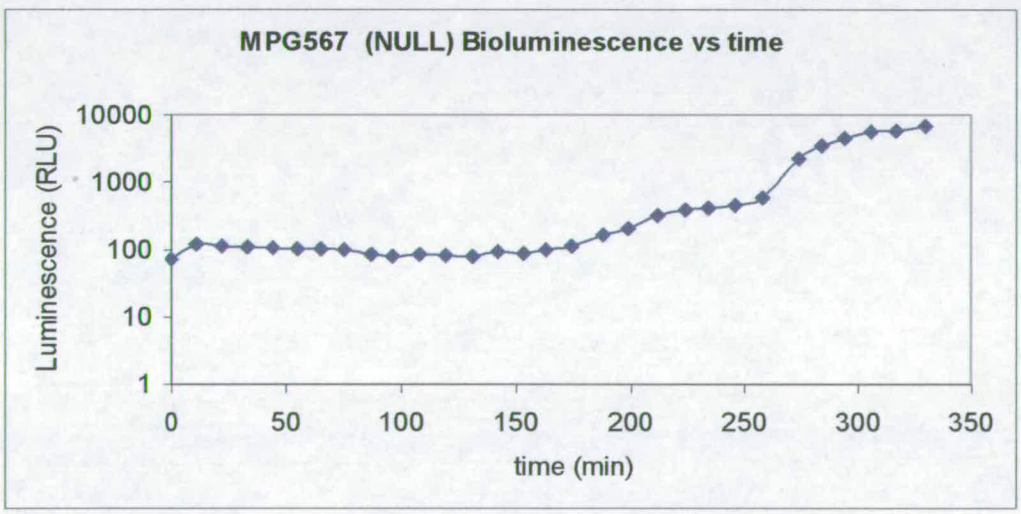
As seen from the graphs in figures 5.1 (A-H) the level of light emission from individual strains at a given OD is to a certain extent different. In other words while all strains, may reach a particular OD at a certain time point, the level of bioluminescence at this OD is not necessarily the same for each strain. The reason for this cannot be conclusively derived from the work presented here but it could be that while OD monitors cell concentration, light production reflects promoter activity and by extension, luciferase activity and cellular metabolism since O_2 and $FMNH_2$ are required. Therefore light emission is directly related to the metabolic rate of the cell, whereas OD is directly related to cell concentration. Thus inhibition of cellular activity may be mirrored in a reduction of bioluminescence but not necessarily of OD while any inhibition of cell growth is reflected accurately by OD but not necessarily by bioluminescence. In section 5.4 whether correlation exists between those two factors, is examined.

Figure 5.1 (A-H): Bioluminescence over time for strains MPG560, MPG567, MPG566, MPG564, MPG565, MPG561, MPG562, and MPG563.

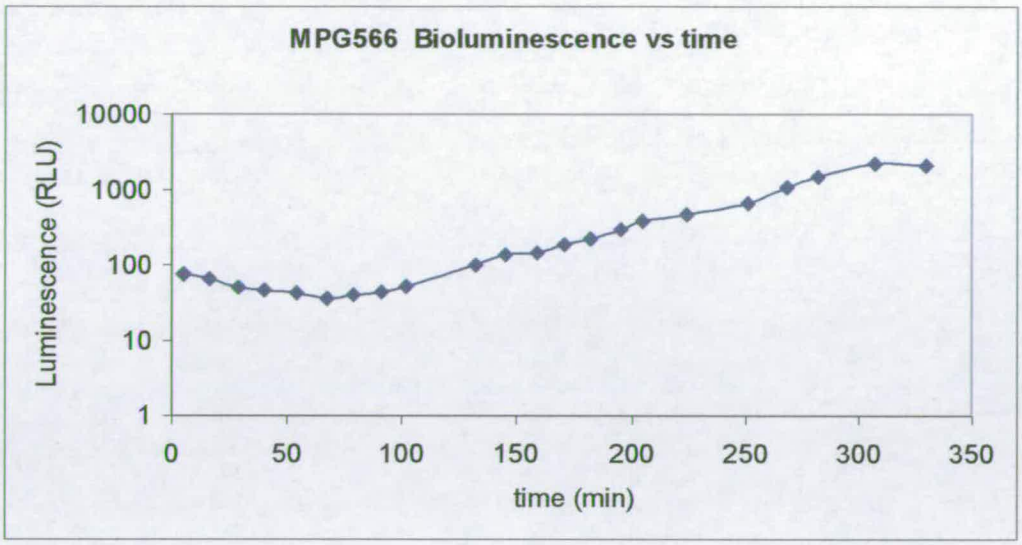
Strains MPG560, MPG567 and the monogene strains, MPG566, MPG564, MPG565, MPG561, MPG562, MPG563 were grown in 250ml flasks containing 50ml of LB, at 30°C, with vigorous shaking (1000 rpm). Bioluminescence was measured in a 96 well plate and is shown as relative light units (RLU), expressed in Logs. Aliquots were taken when the culture reached an OD_{600} of about 0.2 and were subsequently taken approximately every 20 minutes and these can be related to growth of the strains, as monitored by OD_{600} in the spectrophotometer (see Figure 5.2 A-I).


MPG560 (MG1655) Bioluminescence vs time
Luminescence (RLU)
100
10
1
0.1
0
50
100
150
200
250
300
350
time (min)

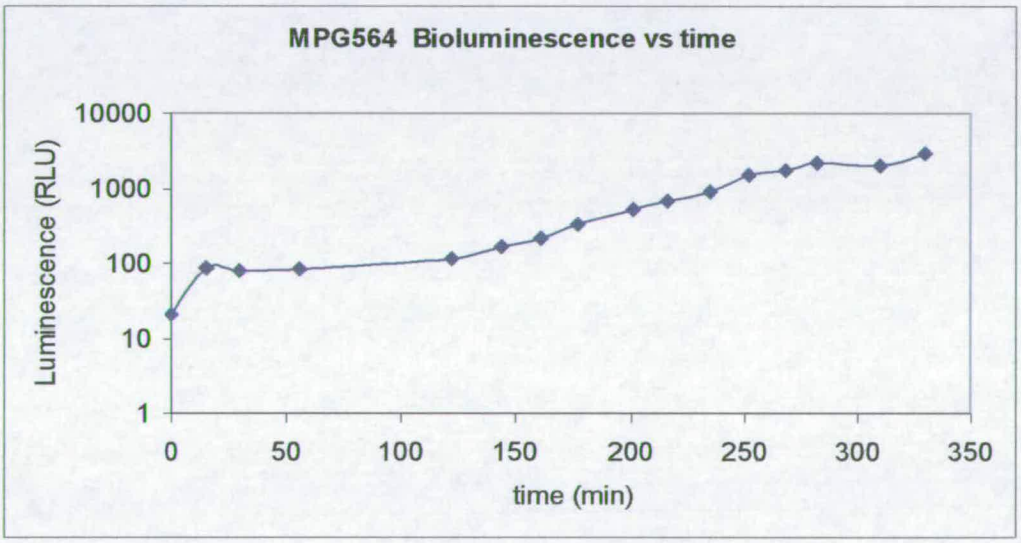
A.


MPG567 (NULL) Bioluminescence vs time
Luminescence (RLU)
10000
1000
100
10
1
0
50
100
150
200
250
300
350
time (min)

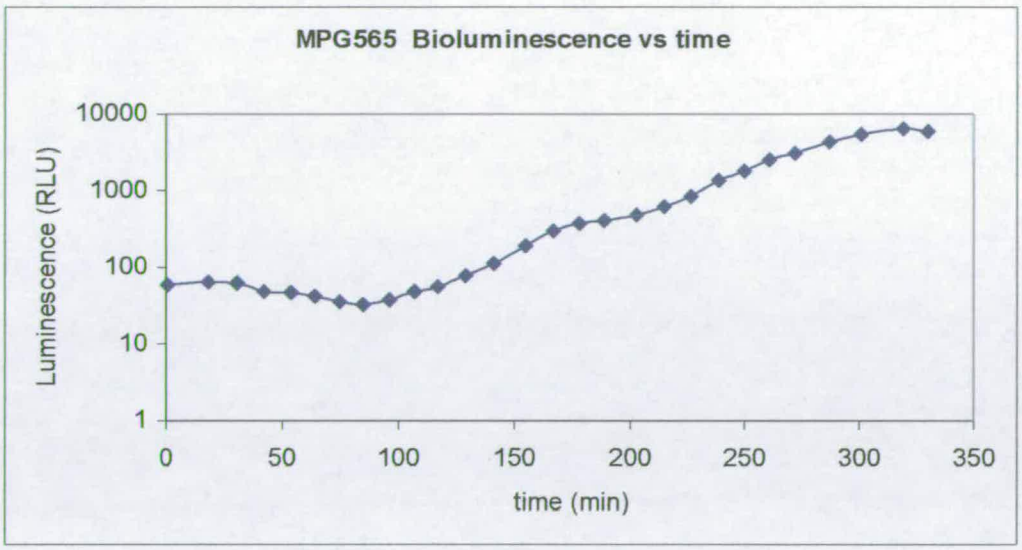
B.


MPG566 Bioluminescence vs time
Luminescence (RLU)
10000
1000
100
10
1
0
50
100
150
200
250
300
350
time (min)

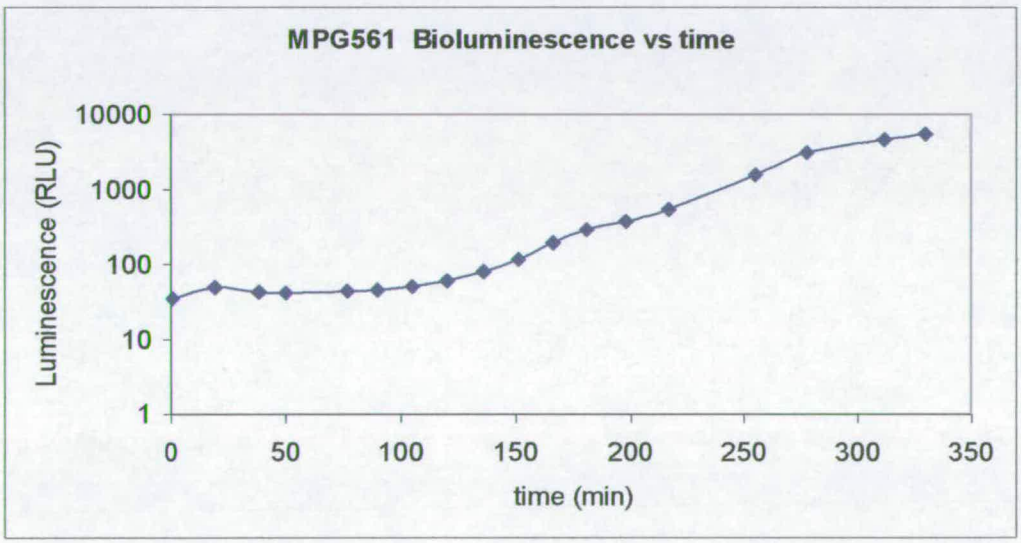
C.


MPG564 Bioluminescence vs time
Luminescence (RLU)
10000
1000
100
10
1
0
50
100
150
200
250
300
350
time (min)

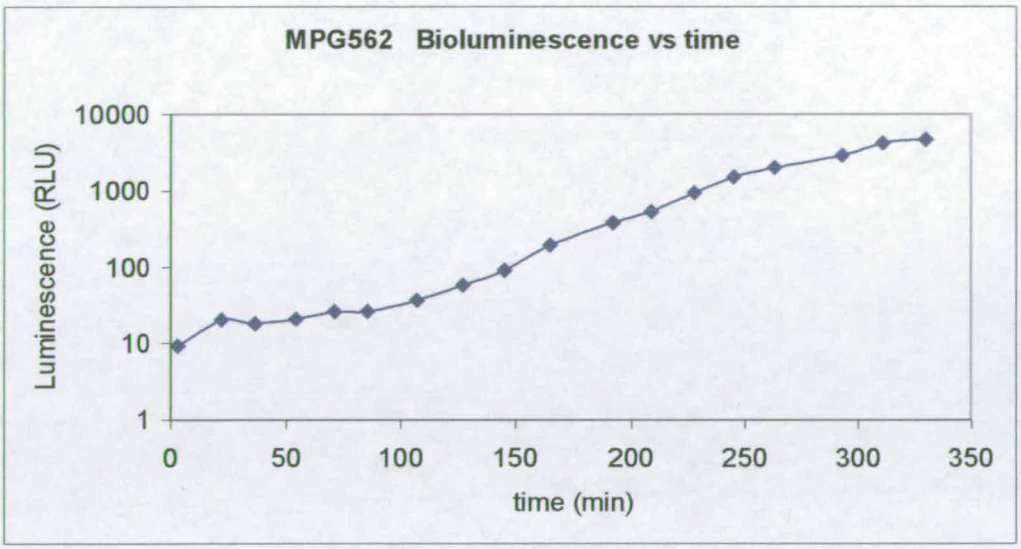
D.


MPG565 Bioluminescence vs time
Luminescence (RLU)
10000
1000
100
10
1
0
50
100
150
200
250
300
350
time (min)

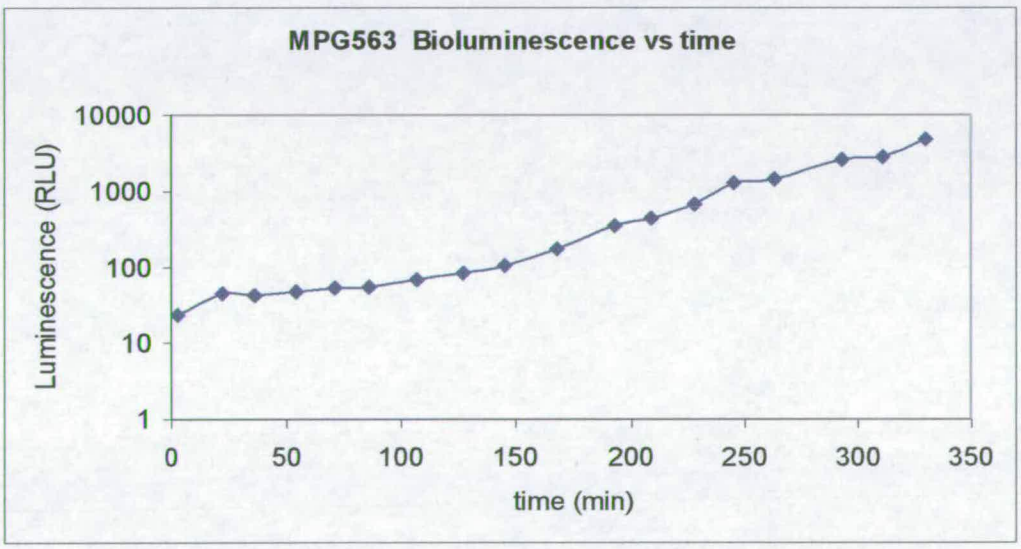
E.


MPG561 Bioluminescence vs time
Luminescence (RLU)
10000
1000
100
10
1
0
50
100
150
200
250
300
350
time (min)

F.


MPG562 Bioluminescence vs time
Luminescence (RLU)
10000
1000
100
10
1
0
50
100
150
200
250
300
350
time (min)

G.


MPG563 Bioluminescence vs time
Luminescence (RLU)
10000
1000
100
10
1
0
50
100
150
200
250
300
350
time (min)

H.

5.3 Growth patterns of strains measured by OD

In order to understand how growth was affected after mutations of the parental strain MG1655, strains MPG560, MPG567 and monogene stains MPG566, MPG564, MPG565, MPG561, MPG562, MPG563 were grown in 250ml flasks containing 50ml of LB at 30°C, with vigorous shaking (1000 rpm). Use of strains carrying the plasmid allowed direct comparison between the data from luminescence and OD.

As mentioned in section 5.2, when samples were taken for luminescence measurements they were also examined by absorbance. Thus, the growth of the strains was monitored by taking 1ml aliquots from the shaking culture approximately every twenty minutes and measuring the OD_{600} in the spectrophotometer. The values measured were then plotted - see figures 5.2 (A-H).

As can be seen in figure 5.2 panel A, MPG560 gave a smooth biphasic growth line that conforms to the growth patterns, which are typical of MG1655 (Daniels, 1990). However, the differences in rates between the two phases was relatively small and so, for the purposes of this work all growth curves, namely the MG1655 growth curve, as well as growth curves on the MG1655-derived mutants that are the monogene and null strains, were considered to be monophasic. The growth of MPG560 was monitored in two different experiments (Figure 5.2 A). Mean values for the data from each experiment 1 and 2 for MPG560 were plotted with error bars showing maximum deviation from the mean, so as to establish the reproducibility of the results. Based on the finding that reproducibility was good, only one set of data was collected for other strains and this is considered representative (see figures 5.1C-H).

From these experiments it can be seen that the mutant strains grew in a normal fashion (Figures 5.2 A-I). Calculations of doubling time based on the graphs in figures 5.2 (A) show that MPG560 in experiment 1 has a doubling time of 56 minutes whereas in experiment 2 the same strain, under the same conditions of growth and the same

monitoring conditions of OD has a doubling time of about 55 minutes. This difference can then be considered to reflect experimental variation. The rest of the strains grew at similar rates to the parental strain. Doubling times were measured during the mid exponential growth phase for all strains at OD_{600} between 0.2 and 0.4. (see table 5.1) The doubling times are summarised in table 5.2.

In particular the doubling times calculated for each strain were 52.6 minutes for MPG560 and MPG567, 62.5 minutes for MPG566, 50 minutes for MPG564, 51.3 minutes for MPG565, 47.6 minutes for MPG561, 54.1 minutes for MPG562 and 58.8 minutes for MPG563. The maximum variation between the longest and shortest doubling time is 14 minutes between strains MPG566 and MPG561. This is a 15% variation factor in the total doubling time between the two outermost strains. Therefore the variation percentage between any other two strains is smaller than 15%.

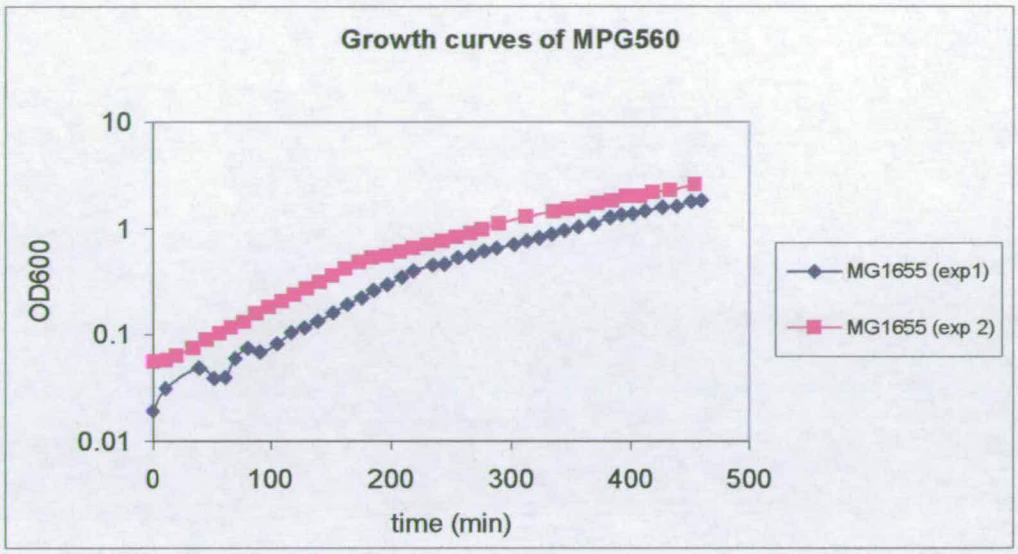
These findings give us some more understanding of the mutant strains and can be used as a basis for bioassay development which will use the growth patterns of the strains to study how the loss of nucleoside transporter genes in these strains affect the growth and survival of the cell after the addition of toxic nucleoside substances, such as AZT.

Figure 5.2 A-I: Growth curves of strains MPG560, MPG567 and MPG566, MPG564, MPG565, MPG561, MPG562, MPG563 as measured over time through OD_{600}

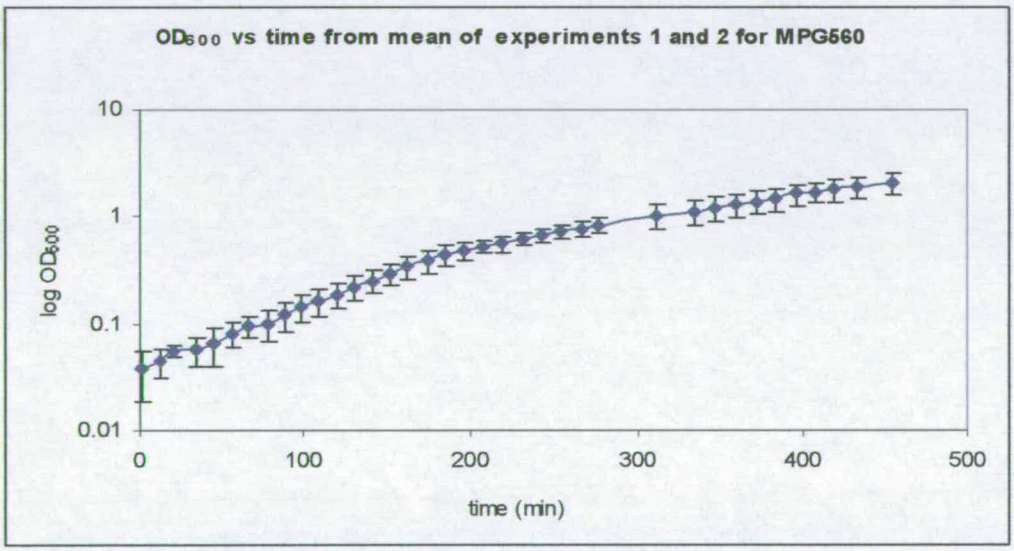
A. Growth curves of *E. coli* MPG560 in two different experiments (exp1 and exp2). OD_{600} was measured over time under the same conditions (LB medium, aerobic conditions, at 30°C).

B. Comparison of the two experiments as shown above in panel A. Mean values of the two experiments were calculated and error bars were fitted to demonstrate low variation from the mean and therefore similarity between the two experiments.

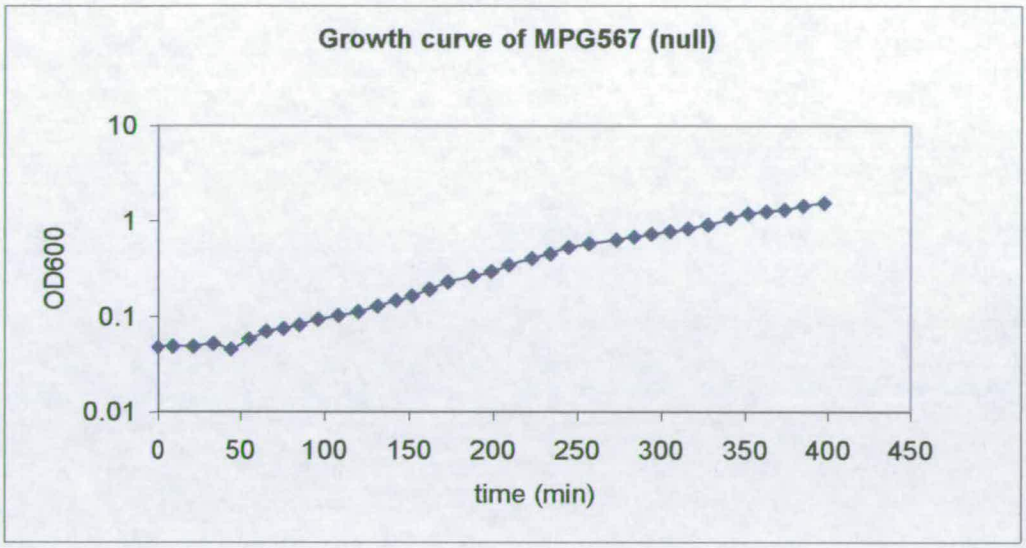
C- I. Growth curves of *E. coli* MPG567 (null), and monogene strains MPG566, MPG564, MPG565, MPG561, MPG562, MPG563 as measured using spectroscopy (OD_{600}) over time. All strains grew in aerobic conditions in rich LB medium at 30°C. The doubling time for each strain was calculated at mid exponential growth (average of three calculations at OD_{600} 0.2-0.5). A summary of the doubling time values is shown in Table 5.1.


Growth curves of MPG560
OD600
10
1
0.1
0.01
0
100
200
300
400
500
time (min)
MG1655 (exp1)
MG1655 (exp 2)

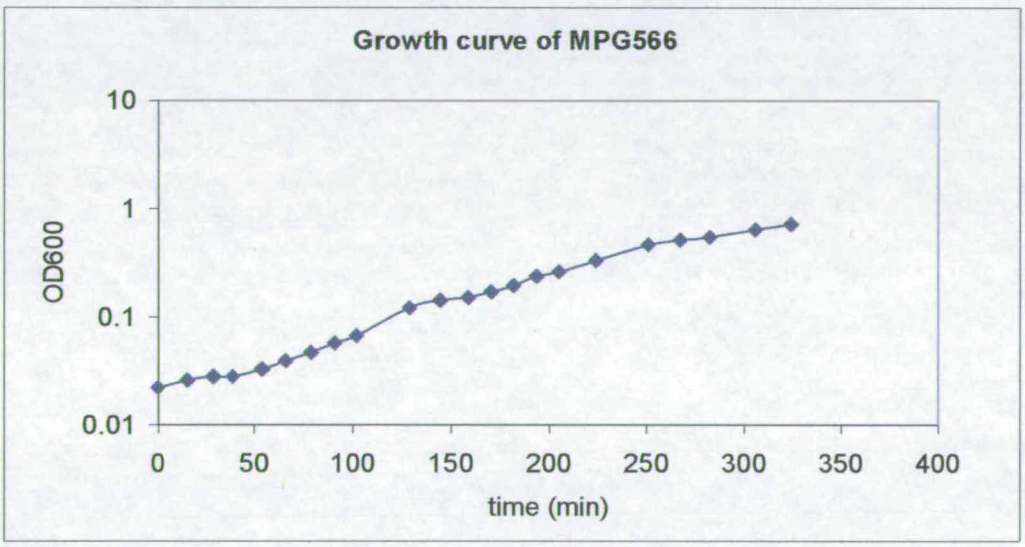
A


OD600 vs time from mean of experiments 1 and 2 for MPG560
log OD600
10
1
0.1
0.01
0
100
200
300
400
500
time (min)

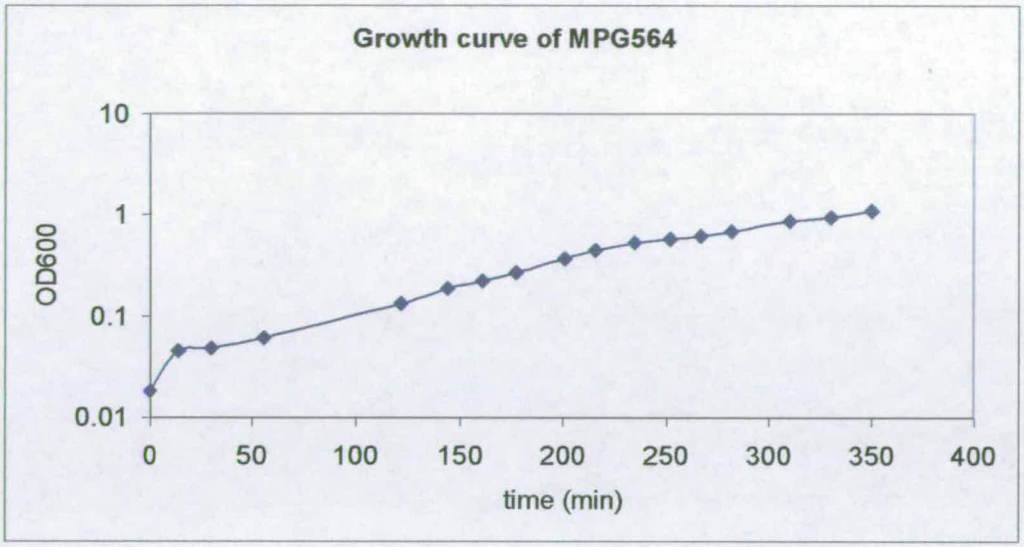
B.


Growth curve of MPG567 (null)
OD600
10
1
0.1
0.01
0
50
100
150
200
250
300
350
400
450
time (min)

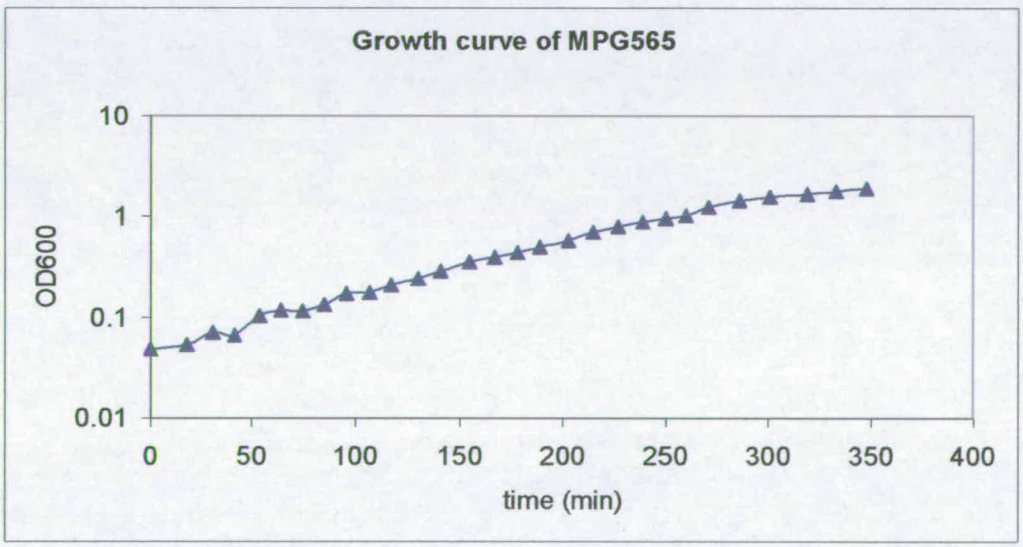
C.


Growth curve of MPG566
OD600
10
1
0.1
0.01
0
50
100
150
200
250
300
350
400
time (min)

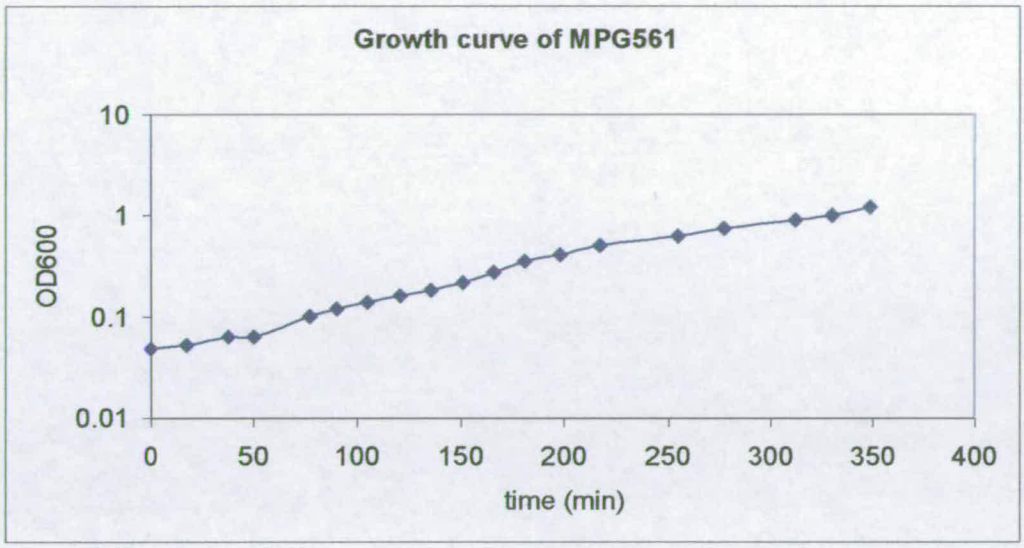
D.


Growth curve of MPG564
OD600
10
1
0.1
0.01
0
50
100
150
200
250
300
350
400
time (min)

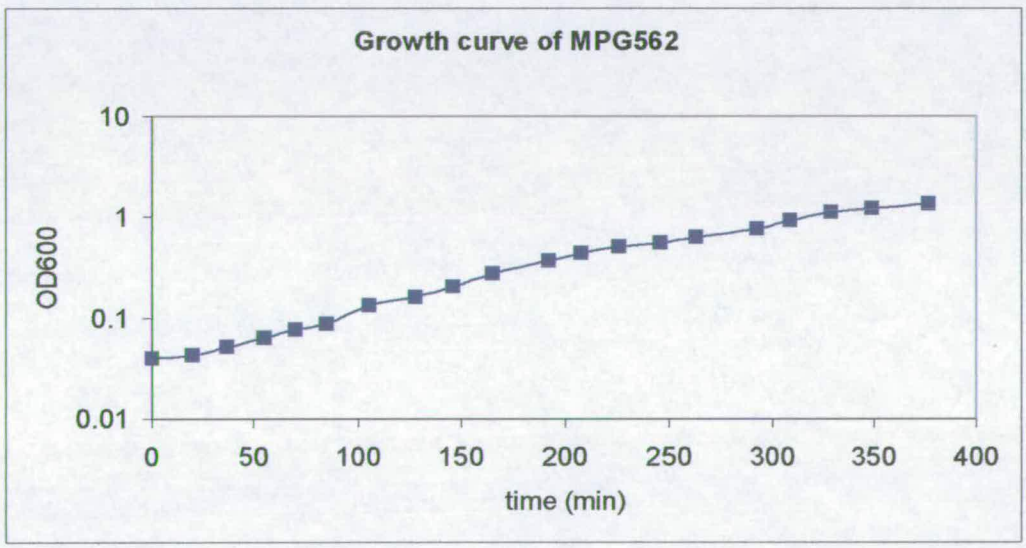
E.


Growth curve of MPG565
OD600
10
1
0.1
0.01
0
50
100
150
200
250
300
350
400
time (min)

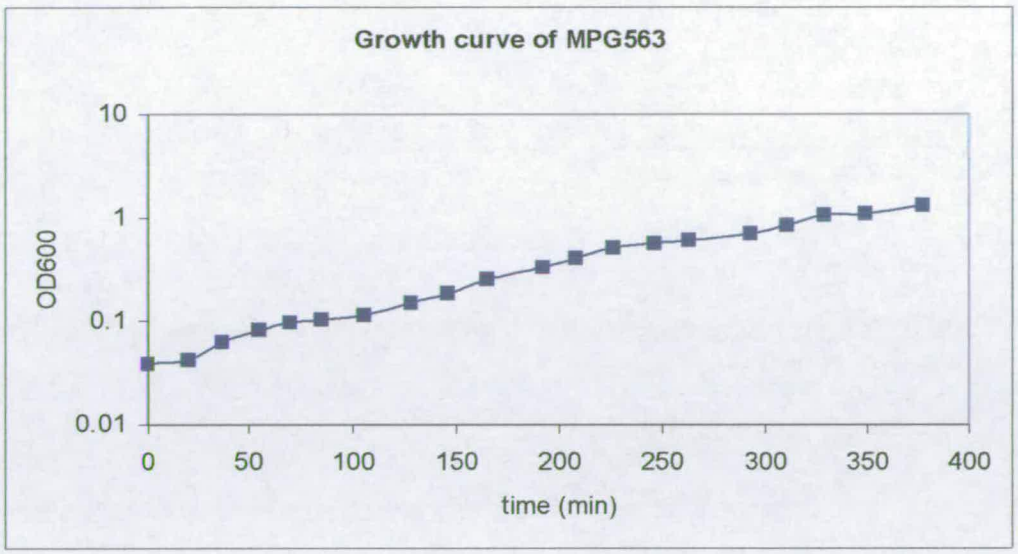
F.


Growth curve of MPG561
OD600
10
1
0.1
0.01
0
50
100
150
200
250
300
350
400
time (min)

G.


Growth curve of MPG562
OD600
10
1
0.1
0.01
0
50
100
150
200
250
300
350
400
time (min)

H.


Growth curve of MPG563
10
1
0.1
0.01
OD600
0
50
100
150
200
250
300
350
400
time (min)

I.

Strain	FORMULA	OD_{600} = 0.2	OD_{600} = 0.4	Doubling time
MPG560	y = 0.0038x - 0.4412	168.7	221.4	52.6
MPG567	y = 0.0038x - 0.4603	173.8	226.4	52.6
MPG566	y = 0.0032x - 0.3792	181.0	243.5	62.5
MPG564	y = 0.004x - 0.4302	157.6	207.6	50.0
MPG565	y = 0.0039x - 0.2618	118.4	169.7	51.3
MPG562	y = 0.0042x - 0.4159	146.6	194.3	47.6
MPG561	y = 0.0037x - 0.3457	147.5	201.5	54.1
MPG563	y = 0.0034x - 0.3159	151.7	210.5	58.8

Table 5.1: Calculations of doubling time for strains MPG560, MPG567 and MPG566, MPG564, MPG565, MPG561, MPG562, MPG563.

All calculations are based on data obtained from graphs on figure 5.1 (A-H). Data between OD_{600} of 0.2 and 0.4 were plotted and a linear trend line was fitted. The formula describing the trendline was used to calculate the time at which an OD_{600} of 0.2 is achieved for each strain. Also the time at which OD_{600} of 0.4 was calculated in the same manner. The time difference between OD_{600} of 0.2 and 0.4 is the doubling time for each strain. Analysis of the calculations of the data are presented on the table

5.4 Relationship between luminescence and OD and their potential use as reporters for bacterial activity and transport of nucleoside analogues

It was important to be able to correlate bacterial growth with light emission in order to assess whether each bioassay is reproducible and accurate in terms of effect of addition of toxic substances in a normal bacterial culture. It was also important to determine that the drug was given in the correct phase of growth. As mentioned in section 5.2 previously, the cultures had been examined by placing the 1 ml aliquot in a cuvette for OD_{600} measurements and 50µl of the same sample in a 96 well plate for luminescence measurement. By taking readings from the same sample it was ensured that the time points at which OD and luminescence were measured were as close to being the same as possible.

OD monitors cell concentration and can be used to observe the state of a growing culture, whereas in the current systems light production reflects promoter activity, luciferase activity and availability of substrate $FMNH_2$, O_2 and aldehyde. As a reminder the reaction that summarises the key events that take place in light production in this system is shown below.

$$FMNH_2 + O_2 + RCHO \rightarrow FMN + R\text{-}COOH + H_2O + \underline{LIGHT}$$

Thus, light emission is directly related to the metabolic rate of the cell, and any inhibition of cellular activity is mirrored in a reduction of bioluminescence production.

If OD is to be substituted by the light production measurement and thereby show that the assays measuring OD and luminescence can be comparable, it has to be shown that there is a direct and reproducible relationship. In this pilot study growth rate as well as light emission were compared for reproducibility.

In order to evaluate the relationship, firstly the logarithmic values of OD were plotted against time from two experiments performed as seen previously in figures 5.1 A and B. In graph 5.3 the two experiments are shown as EXP1 and EXP2. As seen there are two areas on each graph that the characteristic biphasic growth patterns of MG1655 are evident and in both graphs they have a linear relationship. However, since the slopes differ marginally they are considered being monophasic for the current study. This relationship corresponds to the time points of 46-385 minutes for experiment 1 and 81-439 minutes for experiment 2. Data acquired before and after these points were not taken into consideration in this analysis. In other words a window in bacterial growth was chosen for comparison which reflected exponential growth. Thus, overall data acquired over the course of about 5.6 hours for experiment 1 and 5.9 hours for experiment 2 were taken under consideration for the analysis.

In figure 5.4 it is shown that the logarithmic values of luminescence and time from both experiments have a linear relationship. However based on the formulae that describe the trend lines fitted in the graph, it can be seen that the Log of OD and the Log of luminescence evolve with different rates.

To investigate further the similarly of Log (OD) and Log (luminescence), and whether they are proportional, the two parameters were plotted as shown in figure 5.5. It is shown from this graph that there is an area of linearity between the two parameters when plotted on a log scale, that have a relationship to the power of 1.96. This area was identified for these two experiments and lies between the 46-385 minutes of growth for experiment 1 and between the 81 and 439 minutes of growth for experiment 2. Below there is a mathematical outline of the power relationship between Log (OD) and Log (luminescence).

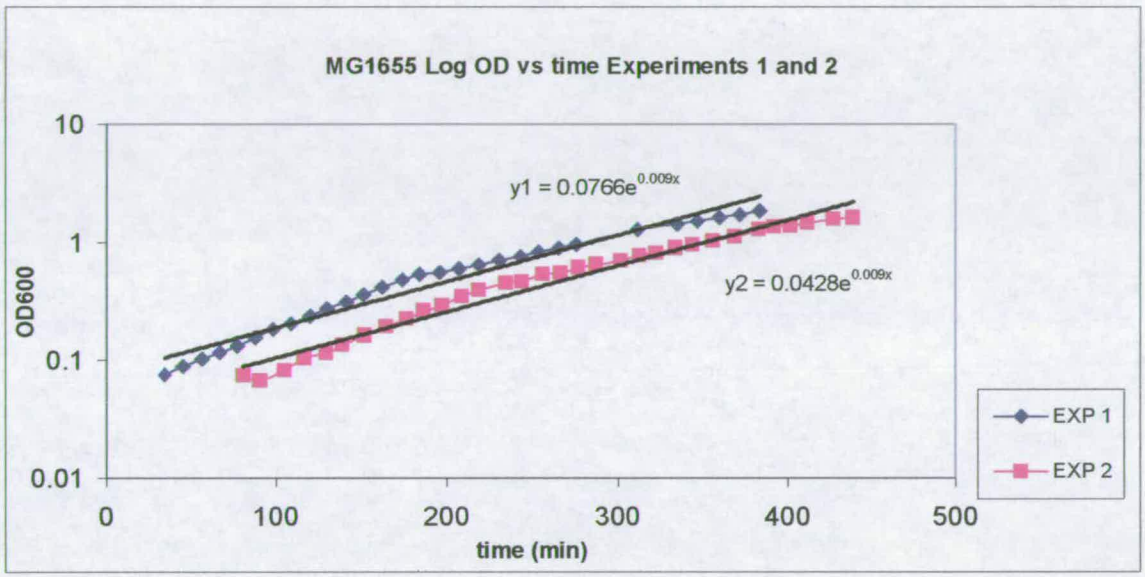


Figure 5.3: *E. coli* MG1655 (MPG560) growth curves based on OD_{600} for two experiments (1 and 2). Experiment 1 (blue diamond) and experiment 2 (pink square) were performed exactly under the same conditions as described in text. Trend lines shown as black lines were fitted for both experiments. As seen from the graph the pairs of trend lines of the two experiments show a proportional relationship between Log of OD and time in both cases. This relationship corresponds to the time points of 46-385 minutes for experiment 1 and 81-439 minutes for experiment 2. Data acquired before and after these points were not taken into consideration in this analysis.

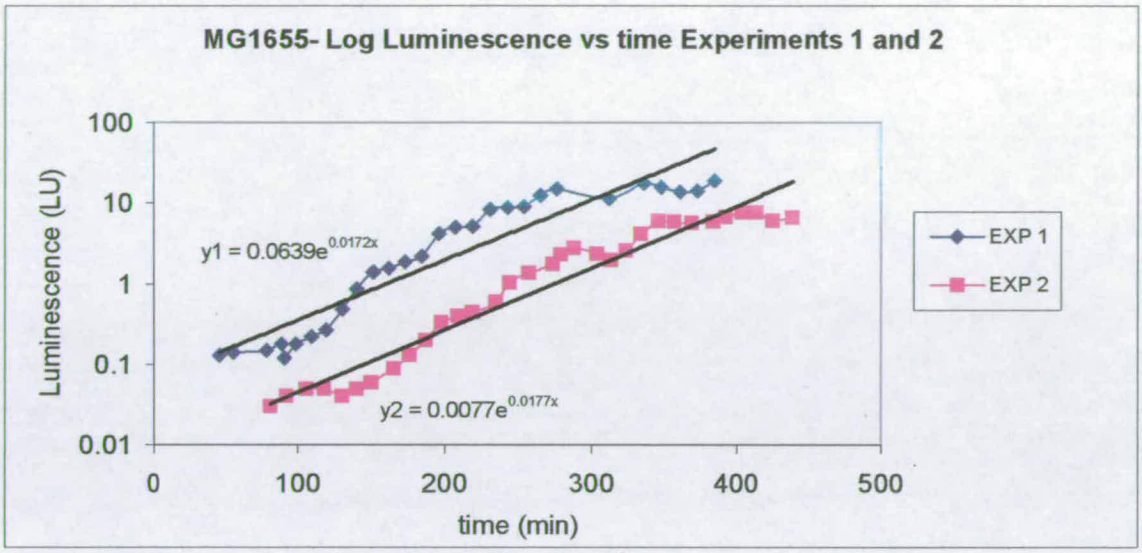


Figure 5.4: *E. coli* MG1655 (MPG560) growth curves based on luminescence for two experiments (1 and 2).

Experiment 1 (blue diamond) and experiment 2 (pink square) were performed exactly under the same conditions as described in the text (5.3). Trend lines were fitted for both experiments and they are shown as black lines. As seen from the graph the trend lines of the two experiments show a proportional relationship between Log of luminescence and time in both cases. This relationship corresponds to the time points of 46-385 minutes for experiment 1 and 81-439 minutes for experiment 2. Data acquired before and after these points were not taken into consideration in this analysis.

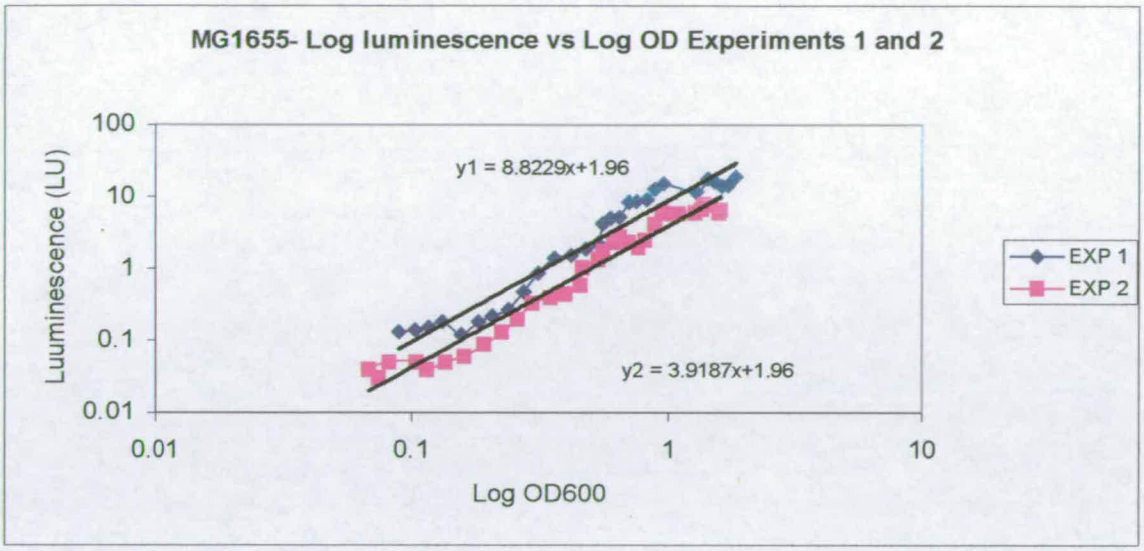


Figure 5.5: *E. coli* MG1655 (MPG560) Logarithm of luminescence versus logarithm of OD_{600} for two experiments (1 and 2).

Experiment 1 (blue diamond) and experiment 2 (pink square) were performed exactly under the same conditions as described in the text (5.3. Trend lines were fitted for both experiments and they are shown as black lines As seen from the graph the trend lines of the two experiments shows a power relationship between Log of luminescence and Log of OD in both cases. This is reflected by the formulae derived from the graph shown below:

y(exper1)= 8.8229 X+1.9609

y(exper2)= 3.9187 X+1.968 (for the log values)

It should be kept in mind that this relationship corresponds to the time points of 46-385 minutes for experiment 1 and 81 –439 minutes for experiment 2. Data acquired before and after these points were not taken into consideration in this analysis

In the investigation of whether Log (luminescence) and Log (OD.) are proportional from figure 5.5 it is shown that the relationship of Log of OD and Log of luminescence is described by the following formula for a straight line as given by Excel.

Log (Lum) = p × Log(OD) + q , where p reflects the gradient of the slope and q is a constant.

Therefore to simplify the formula, the logarithms describing the two factors investigated (OD and luminescence) can be rearranged through power regression as follows.

$\text{Log (Lum)} = \text{Log}(\text{OD}^{p}) + q$ **or** $\text{Log (Lum)} = p\text{Log (OD)} + q$

and so

$10^{\text{Log(Lum)}} = 10^{[\text{Log(ODp)} + q]}$

Therefore

$\text{Lum} = [10^{\text{Log(ODp)}}] \times [10^{q}]$

So finally the two factors without the logarithms, are shown as:

$\text{Lum} = \text{OD}^{p} \times [10^{q}]$

Thus, from the analysis it can be concluded that there is no direct linear relationship between light emission and OD within the whole experimental time examined. This is not surprising since on the principle, two parameters that are not evolving at the same speed should not have a linear relationship. However it is shown that there is an exponential relationship between OD and time and an exponential (power) relationship between light emission and OD.

According to the power regression, the coefficient is 1.96 which gives Lum = $OD^{1.96}$. This is only true for part of the experimental time used for the two experiments. More specifically, the OD based growth curves in these experiments ran for seven hours reaching an OD_{600} of 2.57 and 1.78 in experiments 1 and 2 respectively. But the analysis accounts for data gathered in a timeframe of about 6 hours at which point OD_{600} reached a maximum of 1.84 and 1.604 in experiments 1 and 2 respectively.

Nevertheless, from a practical and commercial point of view it was interesting to investigate the effect of AZT in the null and monogene strains as well as MG1655 as a control in a high throughput manner, and therefore the effects of AZT were examined in a 96 well plate.

5.5 Measuring AZT dose dependence effect on MG1655 (MPG560) growth via bioluminescence

As described above preliminary experiments were performed in order to understand trends of light emission in *E. coli* and the nucleoside transporter mutants when strains were grown in rich media in aerobic conditions. To move to a high throughput experimental approach and investigate the effect of AZT treatment on the strains MPG584, MPG578, MPG579, MPG580, MPG582, MPG581 and MPG583 it was necessary to firstly investigate the wild type strain MG1655. MG1655 was transformed with plasmid pucΔ320. The resulting strain, MPG560, was grown in rich media with vigorous shaking at 30°C until an OD_{600} of 0.2 was reached. At this point a 50μl aliquot was transferred to a 96 well microplate and the strain was left to grow statically, at room temperature in the luminometer. Light emission was then measured in the luminometer, every 10 minutes for two and half hours.

Additionally 50μl aliquots of the culture described above were used to investigate the best conditions at which the effect of AZT in this bioluminescence-based assay would be demonstrable. For this, the effect of a variety of AZT concentrations were examined on MPG560 as measured with luminescence. The final concentrations used were 40μM, 4μM and 0.4μM of AZT.

In figure 5.1A light emission is measured over time from wild type strain MPG560. Light emission was measured for two and a half hours at room temperature without shaking. The average of light emission for the two aliquots was plotted in figure5.6A and after an initial adaptation lag phase of about 70 minutes the strain began to grow and emit light.

In figure 5.6B the effect of AZT on light emission was examined for the same strain over 150 minutes of incubation in the luminometer. For this 50μl aliquots were

transferred in duplicate to the 96 well plate which contained AZT, to final concentrations of 40μM, 4μM and 0.4μM. Figure 5.6B shows the results from two different samples for each of the different concentrations used. As seen from the graph there is an inverse dose-dependent relationship between concentration of AZT and light emission. In particular, in 0.4μM AZT (final concentration) after 150 minutes of incubation in the plate, light emission from MPG560 decreased by a little more than 50% compared to the control. In contrast a 100 fold increase of AZT concentration (from 0.4μM to 40μM) decreased light emission by about 97 %. For this reason the concentration of AZT that was chosen for the future assay experiments was 40μM.

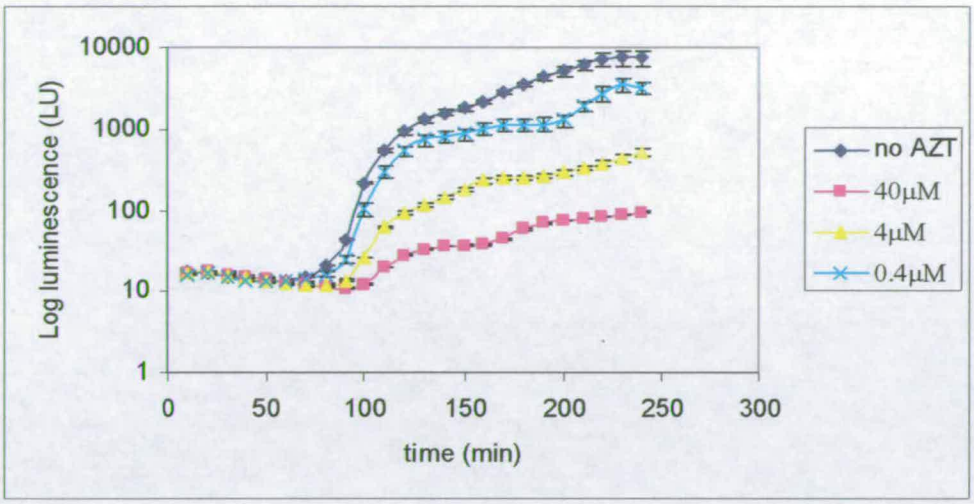


Figure 5.6 : Light emission patterns of strain MPG560 and the effect of different doses of AZT. Strain MG1655 was transformed with plasmid pucΔ320. The new strain known as MPG560 was grown in LB medium, with shaking at 30°C. Error bars represent standard deviation from the mean. When the strain reached OD_{600} of 0.2, a 50μl aliquot was removed from the liquid culture and was placed in a 96 well plate to measure light emission over a period of 150 minutes (no AZT). Also 50μl aliquots were placed in the 96 well plate with AZT at final concentrations of 40μM, 4μM and 0.4μM. All aliquots were left to incubate for 150 minutes to see the toxic effect of AZT on the strains and decide which concentration was most suitable for the bioassay experiments. As seen from the graph, at 40μM AZT is toxic to MPG560 to levels that inhibit light emission by more than 95%.

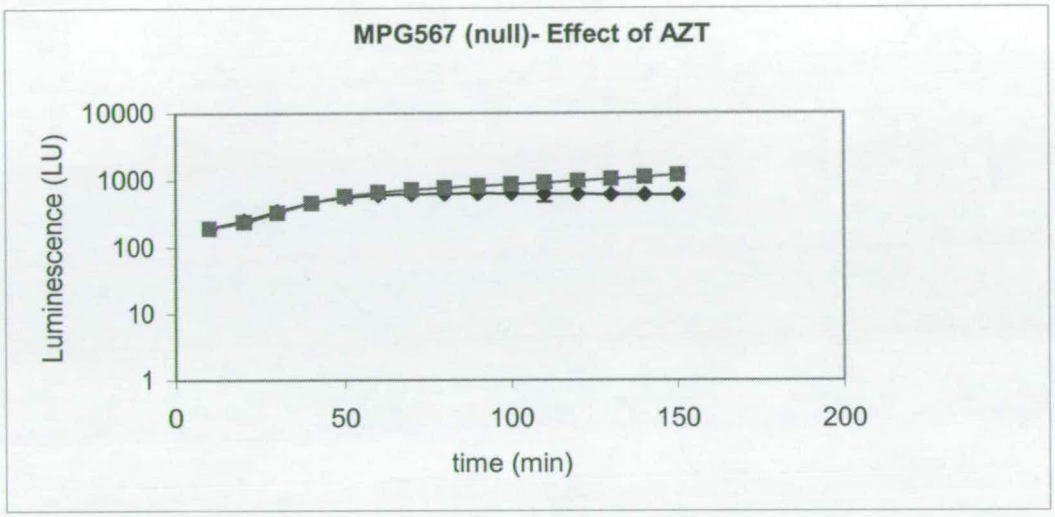
5.6 Measuring AZT effect on luminescence from strains MPG567, MPG566, MPG564, MPG565, MPG561, MPG562, MPG563

The effect of AZT on the null and monogene strains was next examined. For this the strains were transformed with plasmid pucΔ320 resulting in MPG567, MPG566, MPG564, MPG565, MPG561, MPG562, MPG563 and were grown in rich media with vigorous shaking at 30°C until an OD_{600} of 0.2 was reached. At this point 50μl aliquots were transferred to a 96 well microplate in the presence and absence of 40μM of AZT per well, and the strains were left to grow at room temperature in the luminometer, in triplicate. Light emission was measured by the luminometer every 10 minutes for two and half hours as determined in the pilot study for MPG560.

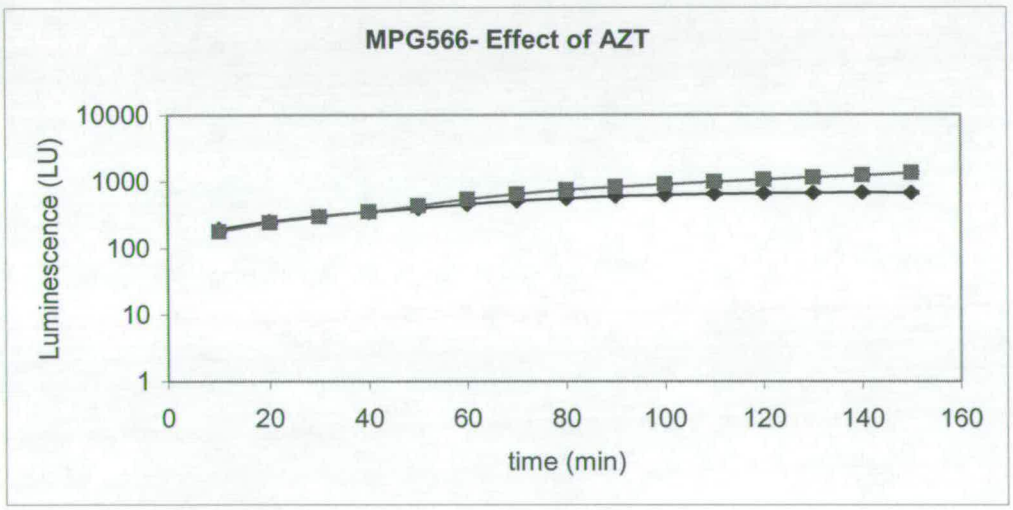
In figure 5.7 (A-G) the effect of AZT on light emission from each one of the strains can be seen. After 150 minutes of incubation of the strains in the luminometer there was a drop in light emission from all strains treated with AZT, including MPG567 which has none of the studied nucleoside transporter genes in its chromosome. In particular, as measured at 150 minutes, bioluminescence dropped by 49% in MPG567 and MPG566 when treated with AZT. This drop in light emission was 46% with MPG564 treated with AZT, 48% for MPG565, 57% for MPG561, 50% for MPG562 and 44% for MPG563. Overall the bioluminescence decreased by about 50% in all monogene strains as well as in the null, as opposed to more than a 97% decrease for wild type strain MPG560 at the same AZT concentration.

It was of note that the effect of all six mutations in strain MPG567 did not completely abolish the strain's ability to transport AZT. Furthermore it seems clear there is an accumulative though not additive effect of transporter mutations. However it is not possible to specify accurately the effect of mutations since all monogene strains have quite similar patterns of light emission after AZT treatment. Of course this may reflect that each transporter is capable of transporting AZT.

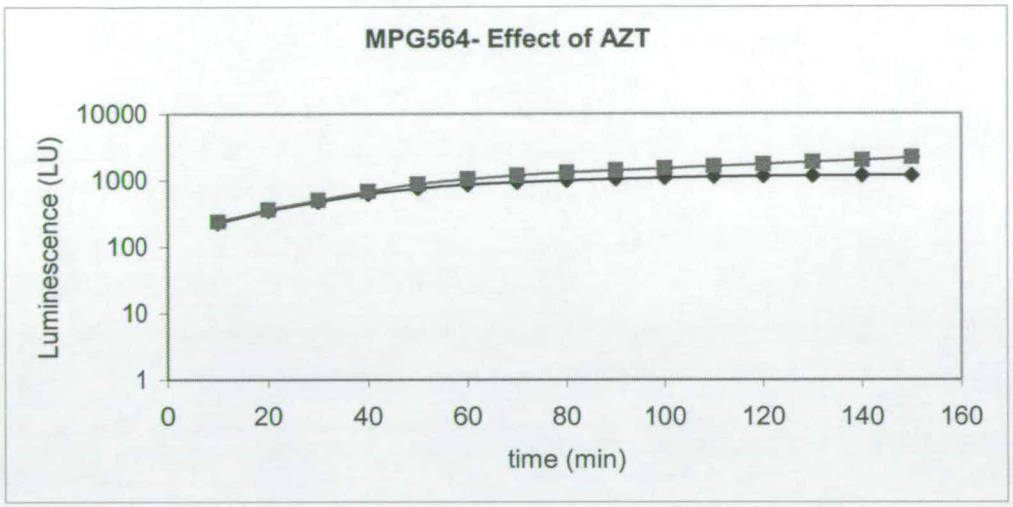
Figure 5.7 Effect of AZT on strains MPG567, MPG566, MPG564, MPG565, MPG561, MPG562, MPG563 as measured by bioluminescence in a 96 well plate. Strains MPG584, MPG578, MPG579, MPG580, MPG582, MPG581 and MPG583, were transformed with plasmid pucΔ320. The new strains known as MPG567, MPG566, MPG564, MPG565, MPG561, MPG562, MPG563, respectively, were grown in LB medium, with shaking at 30°C. When the strains reached OD_{600} of 0.2, three 50μl aliquots were removed from each liquid culture and were placed in a 96 well plate to measure light emission. Also three 50μl aliquots were placed in the 96 well plate with AZT at a final concentration of 40μM. All aliquots were left to incubate for 150 minutes. On the graph solid squares represent no addition of AZT while solid diamond represents AZT treatment to a final concentration of 40μM.


MPG567 (null)- Effect of AZT
Luminescence (LU)
10000
1000
100
10
1
0
50
100
150
200
time (min)

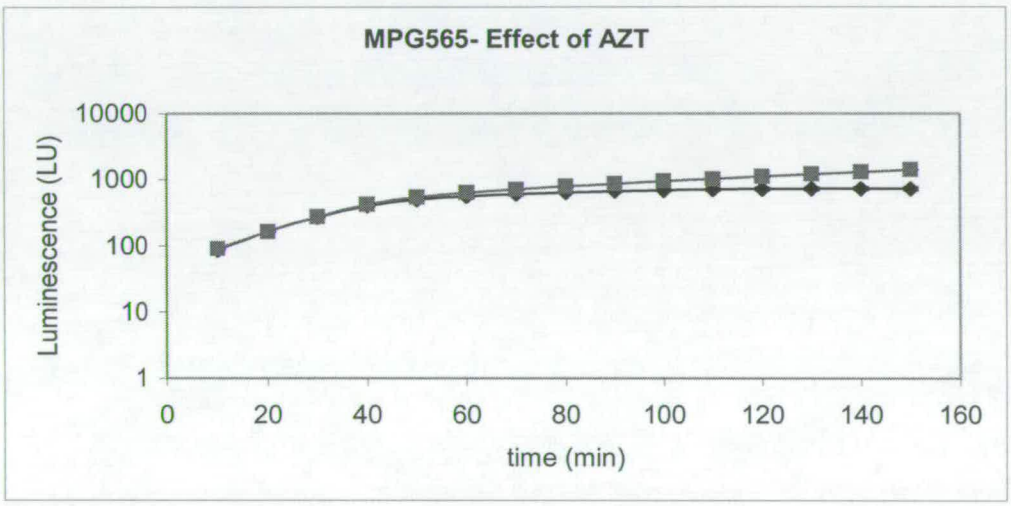
A.


MPG566- Effect of AZT
Luminescence (LU)
10000
1000
100
10
1
0
20
40
60
80
100
120
140
160
time (min)

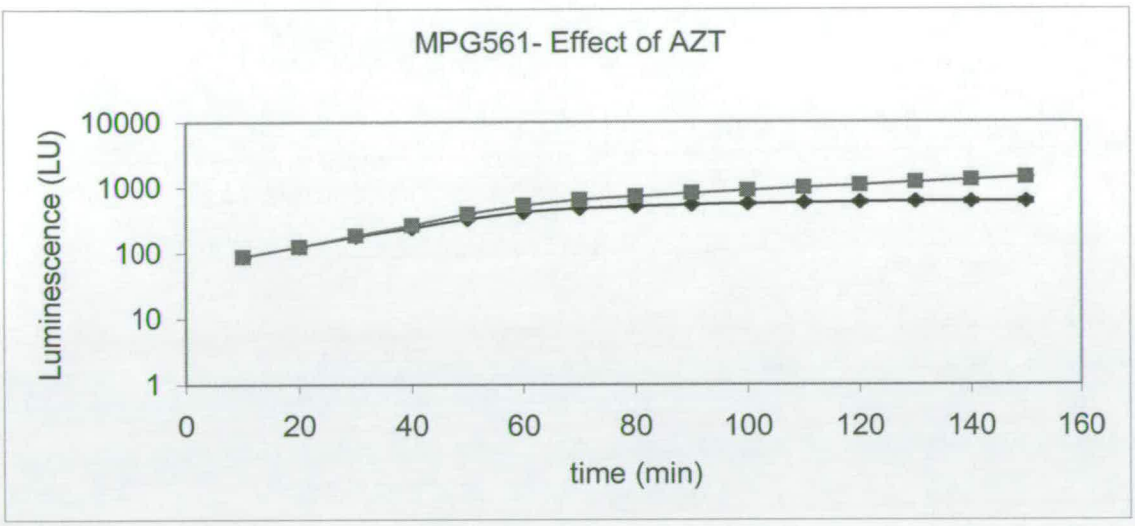
B.


MPG564- Effect of AZT
Luminescence (LU)
10000
1000
100
10
1
0
20
40
60
80
100
120
140
160
time (min)

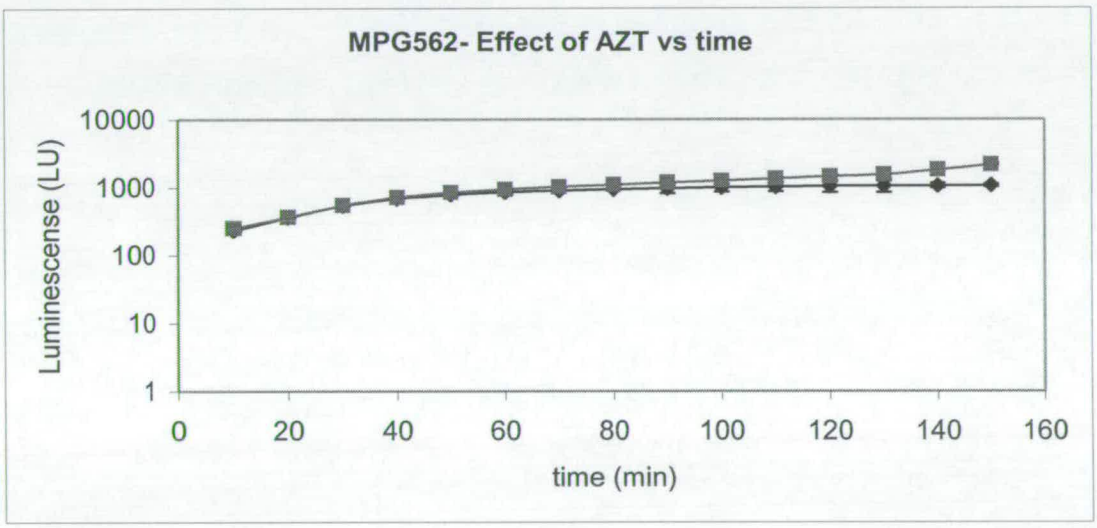
C


MPG565- Effect of AZT
Luminescence (LU)
10000
1000
100
10
1
0
20
40
60
80
100
120
140
160
time (min)

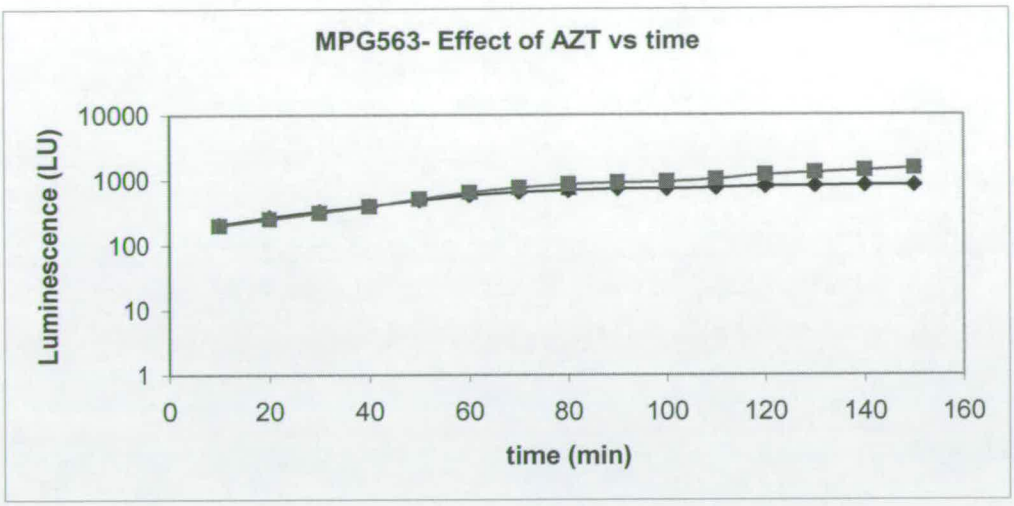
D.


MPG561- Effect of AZT
Luminescence (LU)
10000
1000
100
10
1
0
20
40
60
80
100
120
140
160
time (min)

E.


MPG562- Effect of AZT vs time
Luminescense (LU)
10000
1000
100
10
1
0
20
40
60
80
100
120
140
160
time (min)

F.


MPG563- Effect of AZT vs time
Luminescence (LU)
10000
1000
100
10
1
0
20
40
60
80
100
120
140
160
time (min)

G.

5.7 Discussion

The aim of this experimental work was to develop a high through-put assay to be used as a first screening method that can be used to help design more efficient candidate nucleoside analogues. A small number of potential drug candidates would then be screened for selectivity and potency and scientists could use this test to develop a broader and deeper understanding of properties of molecules that are being evaluated (Gardner *et al,* 2004).

The mutants constructed as described in chapter 3 provided a promising opportunity since they had all but one of the known and putative nucleoside transporters knocked out. Furthermore the monogene strains could provide more detailed information on the specificity of each transporter for such candidate drugs.

Two bioassays were performed to assess the best approach to monitor such cell activity, change in OD_{600} or light emission. OD measures bacterial cell density and allows accurate conclusions on how a factor such as a toxic substance affects culture growth. In contrast light is a signal of promoter activity and also cell viability since it requires $FMNH_2$.

A comparison between the methods was also undertaken. For OD and light to relate, the two factors must evolve in the same way. No linear relationship was found between those two parameters from preliminary experiments. However a power relationship was identified between the two measured parameters of OD and luminescence, as was shown in figures 5.4 and 5.5.

Light emission demonstrated two basic problems. First of all, it shows a level of fluctuation a bit (to the contrary OD curves are smooth) and secondly it is dependent on expression of *luxCDABE* from P_{tac} and therefore may not always reflect cell density accurately. Fluctuation in light emission is common in most systems using *lux* as their

reporter (Baumstark-Khan *et al,* 2003). Also, it was found that light emission levels are not consistent for different strains at the same OD levels, meaning that it is not possible to directly compare between strains.

Nevertheless, light emission is a more attractive reporter in the context of commercialising a bioassay for use in a rapid and high throughput manner.

Change in OD, as a way of analysing the effect of toxic nucleosides on the non-mutated and the mutated cells from this work did not display excessive sensitivity since at 40µl of AZT, OD_{600} measurements showed complete inhibition of cell growth for all strains, regardless of whether the strain was mutated or not (data not shown).

It is possible that an unidentified nucleoside transporter exists in the null strain which transports ATZ. However this seems unlikely since, from previous experiments it was seen that the null mutant does not grow well in minimal media with a nucleoside as a carbon source as seen in figures 4.5 (A-I). The addition of AZT under such growth conditions proves lethal, an indication that AZT is transported into the cell by other transporters in high enough levels so as to cause cell death. These transporters could be passive diffusion transporters of the inner membrane, however the only known passive diffusion transporter in *E. coli* is glycerol transporter GlpF which is unlikely to be able to transport AZT. On the other hand if AZT is broken down into its sugar and base moieties and then transported, the toxic azido group which is on the sugar would require transport through deoxyribose sugar transporters. Alternatively it is possible that a lipophilic compound such as AZT which has a high oil/water partition coefficient might diffuse fast through the membrane via partitioning (Albert Adrien, Selective Toxicity).

The nucleoside transporter deletions seemed to have a cumulative effect on diminishing AZT transport as illustrated by the fact that *E. coli* without any mutations shows a greater drop in light emission by more than 97% when AZT is added to the media whereas AZT addition to the monogene mutants resulted in a smaller decrease in light

emission (about 50%) (See figures 5.6B and 5.7). Nevertheless this work cannot formally exclude the fact that maybe more nucleoside specific transporters may exist in the inner membrane of *E. coli*. Also, as mentioned above the fact the bioluminescence levels were not constant between the different strains brings about the problem of not being able to directly compare between strains. This is why effects of AZT on strains are given in percentage drop of light emission when compared as opposed to RLU.

Although nucleoside analogues resemble naturally occurring nucleosides they are transported into the cell by passive diffusion (Li and Chan 1999). Intestinal transport studies have shown that AZT as well as other nucleoside analogues such as ddI are transported by non-facilitated membrane diffusion via paracellular pathways (Park and Mitra, 1992). The permeation of AZT was studied through erythrocyte and lymphocyte membranes and the conclusion was that that the non facilitated diffusion of this agent can be attributed to the lack of 3'-hydroxyl moiety and the relatively high partition coefficient (Domin *et al,* 1988). Therefore the cell death that is shown in the experimental work carried out in this study might be a direct effect of the transport of AZT through the cell membrane via non facilitated diffusion receptors such as the outer membrane Omp porins or via facilitated diffusion via Tsx.

On the other hand it has been found that both bacterial *nupC* as well as *h*CNTs are able to transport AZT (Loewen *et al,* 2003). These experiments were based on over expression of the associated nucleoside transporter gene and looking for an increase in uptake of substrate, which in these experiments was AZT. Although this shows that these transporters are able to transport AZT the work that is presented here shows that lack of the genes responsible for nucleoside transport do not have the opposite effect i.e. it does not prevent the cell from transporting AZT into the cytoplasm in sufficient amounts for it to become lethal.

Overall it appears that the bioassays could still be useful for studying nucleoside analogues that use specialised transporters to enter the cells such as gemcitabine. In this

regard it is useful to note that whilst the knocking out genes does not prevent AZT from entering the cell, the overexpression of *nupG* has been shown to increase AZT uptake by the cell (Xie *et al,* 2004).

Bioassays used traditionally commonly involve some type of measurement of plaque reduction, such as a plaque inhibition test. Also compounds are assayed for biological activity based on the effect of the compound on cell number (Krenitsky *et al,* 1983). The cells used by Krenitsky *et al* (1983) were human D-98 cells (HeLa), mouse L-cells, as well as gram positive and gram negative bacterial cells. Chemo sensitivity tests of compound toxicity on cells viability as measured by OD have also been described extensively for a variety of nucleoside analogues such as troxacitabine, a deoxycytidine analogue with antiviral activity in human malignancies (Gourdeau *et al,* 2001) or gemcitabine (Sandler and Ettinger, 1999).

In particular, Skehan *et al,* (1990) developed a rapid, sensitive, and inexpensive method for measuring the cellular protein content of adherent and suspension cultures in 96-well microtiter plates. The method is suitable for ordinary laboratory purposes and for very large-scale applications, such as the National Cancer Institute's disease-oriented in vitro anticancer-drug discovery screen, which requires the use of several million culture wells per year. This method is known as the SRP method. Cultures fixed with trichloroacetic acid are stained with sulforhodamine B (SRB). Unbound dye is removed and protein-bound dye is extracted for determination of optical density in a computer-interfaced, 96-well microtiter plate reader. The sensitivity of the SRB assay compared favourably with sensitivities of several fluorescence assays and was superior to those of both the Lowry and Bradford assays and to those of 20 other visible dyes. The advantages of the SRB assay are that it provides a colorimetric end point that is non-destructive, indefinitely stable, and visible to the naked eye. Furthermore, it provides a sensitive measure of drug-induced cytotoxicity, and it is well suited to high-volume, automated drug screening. SRB fluoresces strongly with laser excitation at 488 nm and can be measured quantitatively at the single-cell level by static fluorescence cytometry.

The same year, The National Cancer Institute (NCI) implemented a large-scale in vitro drug-screening program that required a very efficient automated assay of drug effects on tumour cell viability or growth. Many laboratories worldwide had adopted a microculture assay based on metabolic reduction of 3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide (MTT). However, because of certain technical advantages the use of the protein-binding dye sulforhodamine B (SRB) in a large-scale screening application was followed (Rubinstein *et al,* 1990)

The MTT and SRB assays were each used to test 197 compounds, on simultaneous days, against up to 38 human tumour cell lines representing seven major tumour categories. On subsequent days, 38 compounds were retested with the SRB assay and 25 compounds were retested with the MTT assay. For each of these three comparisons, the differences between the two assays for cell survival and accuracy were measured. The results indicated that under the experimental conditions used and within the limits of the data analyses, the assays perform similarly. Because the SRB assay has practical advantages for large-scale screening, however, it has been adopted for routine use in the NCI in vitro antitumour screen (Rubinstein *et al,* 1990)

Within this work, the bioassay has to do with a bacterial model-based assay, which makes the assay easier to use than animal cell-based assays, due to the practicality of handling *E. coli* in the laboratory, as well as the fast results acquired from this kind of work. The added benefit of having an assay which is based on luminescence makes the bacterial model an even better candidate.

However within the context of the present work, improvements in the way the bioassays were tested could be performed. Specifically to analyse the specificities of the nucleoside transporters studied here for particular analogues such as AZT, it would be possible to over-express each one of the candidate genes in a *E. coli* strain such as BLR under an inducible promoter such as P_{tac} and perform an uptake study and kinetics, using radiolabelled AZT as the substrate. Such transport profiles could potentially give

information on the ability and affinity of each one of the gene products for AZT transport. A range of other nucleoside analogues could be used in similar experiments to characterise the genes further and to get a better understanding of the nucleoside transporters' specificity.

Chapter 6

Discussion and conclusions

6.1 Evolution of multiple coexisting transporters

Six different transporters have been described in this work which catalyse nucleoside transport in the inner membrane of *E. coli*. Elsewhere, three types of human nucleoside transporters have been extensively studied and another two groups are being examined (Kong *et al,* 2004). Generally groups of seemingly similar proteins which serve the same purpose have arisen as a direct result of environmental or genetic pressure. For example sugar transport is served in the cell by a large group of transporters each one dedicated for particular sugars which may be utilised by cells in different environments or at different stages of the cell life.

Transport proteins translocate their solutes either by channel- type mechanisms or by carrier-type mechanisms, and a close relationship between these two kinetically distinguishable types of transport mechanisms has been proposed. A preformed channel may be a structural requirement for carrier-mediated transport. Introduction of specific residues in the channel, allowing for stereospecific solute recognition as well as conformational fluctuation between two alternative states, with the solute-binding site alternating between orientations facing inward versus outward, may be all that is required to convert a channel into a carrier (Tam and Saier, 1993).

Distinctive structural features, as well as recognition of the fact that such patterns may have evolved independently many times during evolutionary history to give transport proteins of the channel or carrier type, leads to the possibility that specific structural motifs are particularly well suited for the formation of transmembrane solute channels (Saier 1994).

Substrate specificity within any one family (or superfamily) of porters correlates to a remarkable degree with sequence conservation. This generalization appears valid for functionally diverse proteins with a range of specificities from simple ions, sugars, amino acids, to oligosaccharides and peptides and on up to complex carbohydrates, proteins, and lipophilic substances. The nature of the protein seems to be immaterial to the conclusion that the degree of sequence relatedness generally correlates with

the degree of functional similarity, particularly regarding solute binding specificity (Saier 1994, Saier, 2003).

Second, functionally and structurally important regions in proteins can be identified on the basis of sequence comparisons for members of homologous families, provided that these functional and structural features are common to the different members of the family. These regions can be identified on the basis of sequence conservation. The nature of the residues conserved within a specific region of a family of proteins may provide hints about whether a region of conservation plays a structural or a functional role, and specific, currently recognized motifs, if present, can provide clear indications of function. In most cases, however, site-specific mutagenic and biochemical analyses are required to establish the specific function involved (Saier, 1990).

The ancient protein transmembrane systems were considered to be simple ones, with two or three transmembrane segments (TMS) that could form flexible oligomeric channels. These provided simple transmembrane diffusion functions with low degrees of substrate selectivity. A need for higher degrees of specificity, for coupling two or more transport processes (i.e. co-transport, antiport) and for energy coupling for solute accumulation or expulsion provided the driving force for carrier evolution. The formation of stereospecific binding sites, strict stoichiometric recognition and conformational coupling either within the transporter polypeptide chain itself or with other superimposed energy-coupling subunits were all required (Saier, 2003).

It is thought that such requirements resulted, first, in the evolutionary appearance of more constrained channels that could assume carrier functions, secondly, in the appearance of larger, more complex obligatory secondary carriers that could no longer catalyse passive diffusion and, thirdly, in the emergence of primary active transporters and group translocators with superimposed energy-coupling subunits.

Thus, the ancestral precursors of all these transporter types would appear to be simple peptide channels. The pathways most frequently taken were evidently tandem

intragenic duplications giving rise to larger helical bundles that had the potential to form discrete stereospecific intramembranous substrate binding sites. They could also be constrained for coupling to other transport processes and, through conformational coupling, they were subject to control by a superimposed primary energy-yielding process such as ATP hydrolysis. It is assumed that none of these requirements could be satisfied by the ancestral one, two or three TMS polypeptides that served as the ancestral precursor as no such carrier has yet been documented in the scientific literature. Thus, the primary advantages provided by intragenic duplication or triplication to generate large transmembrane proteins over oligomerization of smaller peptides were (i) specificity; (ii) stoichiometric transport coupling; and (iii) control by energy expenditure (Saier, 2003).

These last two processes allow accumulation of specific solutes within or expulsion of solutes from cells, processes that simple channels are incapable of catalysing (Saier, 2003). Multiple approaches should allow definition of the requirements for flexibility and rigidity in the construction and evolutionary modification of transport proteins.

Acimovic and Coe (2002) investigated the evolutionary history of the ENT family of transporters in which they propose a model in which four regions of high conservation were identified and they point to the evolutionary history of these transporters from prokaryotes to humans. However, based on analyses in combination with previous functional characterisation of these proteins (Bremer *et al,* 1990; Nieweg and Bremer 1997), they proposed that the Tsx family members are homologues of the eukaryotic ENTs. The Tsx proteins have been identified in a number of gram-negative bacteria (Nieweg and Bremer 1997). These bacteria possess an outer membrane, the thickness of which is similar to that of other biological membranes. However, this outer membrane differs from eukaryotic membranes, in that it consists of an asymmetrical bilayer of phospholipid and lipopolysaccharide, producing a somewhat different context for transport proteins.

Tsx proteins have been shown to transport nucleosides and deoxynucleosides (used as nutrients) by facilitative diffusion from the extracellular environment, where they are present at submicromolar concentrations

However Tsx proteins were found unable to transport nucleobases or phosphorylated derivatives of nucleosides like ENTs. Moreover predicted topology of Tsx is an integral, multitransmembrane protein with a large periplasmic loop in the central portion of the protein which is suggestive of the prototypic ENT structure. Most strikingly, there is both sequence homology and conservation of specific residues in Tsx and eukaryotic ENTs (Acimovic and Coe, 2002) However from this work it was not possible to determine, if Tsx is mechanistically homologous to eukaryotic ENTs.

The reasons underlying the absence of ENT homologues in the nonproteobacteria and archaea is not clear. In the same study, Acimovic and Coe hypothesized that ENTs could be a late acquisition by this prokaryotic lineage or the other lineages may have ''lost'' their ENTs, relying instead on proton-dependent nucleoside transport mediated by concentrative nucleoside transporter (CNT) homologues (e.g., NupC in *E. coli*; Craig, *et al,* 1994). CNTs are present in the inner membrane of prokaryotes, including the gram-negative bacteria, suggesting that the ENTs are only required for the transport of nucleosides across the outer membrane of this group and therefore would not necessarily be present or functional in other prokaryotes.

6.1.1 Nucleoside transporters and their diversity as candidates for clinical research

Nucleosides transported across cell membranes have been shown to have a central role in various functions of the cell, such as being used as intermediates of nucleic acid production of nucleic acid and of cofactors, such as NAD, thiamine and folic acid and, as a direct source of carbon (Neuhard and Nygaard, 1987) as well as having important roles in human physiology.

Transport across human, animal, fungal and bacterial cell membranes has been studied intensively and proteins mediating nucleoside transport are classified into two groups: one group catalyses facilitated diffusion down a concentration gradient (ENT, equilibrative nucleoside transporters), whereas the members of the other group (CNT, concentrative nucleoside transporters) are able to transport nucleosides against an existing concentration gradient.

Nucleoside transport systems have been studied well in bacteria and in particular in *E. coli*. Nucleosides are transported across the inner membrane of *E. coli* by at least two separate energy-driven systems, known as NupC and NupG. These systems are encoded by the genes *nupC* and *nupG*, respectively. Studies have reported that NupC is unable to transport guanosine but is able to transport other common nucleosides and their deoxyribonucleoside derivatives. NupG on the other hand is thought to be able to transport all common nucleosides and their deoxyribonucleoside derivatives and it is able to take up the antibiotic showdomycin (Craig *et al,* 1994).

In *E. coli* NupC and NupG are proton gradient-driven systems, allowing the accumulation of nucleosides in the cell even when the concentration in the external environment is very low. In these systems the flow of protons into the cell provides the necessary energy for the nucleosides to be taken up by the cell. This is a distinction from the concentrative mammalian nucleoside transport systems that are driven by sodium ion gradients (Baldwin *et al,* 1999, Cass et al, 1999). CNT1 was the first mammalian homologue of NupC to be identified. Interestingly, CNT1 is able

to transport the nucleoside analogue AZT, used to combat HIV infection. CNT1 and NupC share substrate specificity and they have amino acid sequence identity of 27%, particularly in their C-terminal region, indicating that this probably contains the substrate binding region (Huang *et al,* 1994).

With the completion of the *E. coli* genome project paralogues of NupC and NupG have been found designated as YeiM and YeiJ, and XapB and YegT respectively. (Maurice P Gallagher-unpublished data). It is worth noting that XapB has been proposed to be involved in xanthosine catabolism (Seeger, *et al,* 1995).

Natural nucleosides can be structurally modified to generate derivatives which are pharmacologically active. By retaining most of the metabolic properties of the parent compounds they can be transported into the cell and be metabolised. They are able to interfere with nucleic acid synthesis and therefore inhibit proliferation events in the cases of cancers or resistance to virus replication in virally infected cells (Baldwin *et al,* 1999, Damaraju *et al,* 2003).

The potential of nucleoside analogues to be incorporated into nucleic acids by the DNA repair machinery makes them interesting candidates for combination with DNA damaging agents. Once incorporated, nucleoside analogues are fairly resistant to repair excision and cause irreversible damage to the cell (Sommadossi, 1998).

Knowledge gained from the identification and molecular characterisation of human nucleoside transporters is an important step in future developments in the use of nucleoside drugs in cancer chemotherapy as well as viral and parasitic infections (Dwyer and Kesson, 1997, Dighiero, 1994, Parker *et al,* 2000).

Generally speaking nucleoside analogues can become more effective by expanding knowledge on the way they enter the cell so as to improve selectivity and prevent uninfected cells from uptake of toxic compounds.

The different kinetic properties of nucleoside transporters might also be a key way of designing chemotherapeutic schedules "tailored" to the specific transporters present in neoplastic cells. Protection of normal cells from the toxic effects of nucleoside analogues might also be enhanced by the co-administration of inhibitors that block ENT but not CNT mediated transport.

There is potential in detecting transporters or toxicity levels of the nucleoside analogues to be used, through assays. Such assays should prove helpful in identifying patients for whom therapy with a particular nucleoside analogue will fail. That is why there is a bigger need for development not only of new molecules which are candidates as therapeutic agents but also, of detection methods which verify which drug is active and hopefully by which transporter activity it is taken up by cells. This project focused on the function and specific characteristics of nucleoside transporters in *E. coli*.

6.1.2 Novel nucleoside transporters in *E. coli* and their properties

To functionally characterise the putative nucleoside transporter- encoding genes *xapB, yegT, yeiM* and *yeiJ* and study whether they are involved in nucleoside transport as the *nupG* and *nupC* products are, mutants were constructed which had the coding sequence of these genes deleted from their chromosome. These mutants had five of the genes deleted, leaving one of the genes intact and functional in the *E. coli* chromosome. Also a mutant was constructed that had all six genes deleted from the chromosome, known as the null mutant. This way the mutants could be used in nucleoside uptake experiments in order to find out more about specificity and characteristics.

To delete the coding sequences of the genes from the chromosome the technique used was allelic exchange via linear recombination of the gene sequence and a replacement cassette that also had the option to be excised from the chromosome, leaving a single scar area in place of the former gene coding sequence and the whole of the gene sequence deleted. The linear recombination method was described by Murphy (1998) and two groups have developed his findings further by developing systems that allow accurate gene deletions (Yu *et al,* 2000, Datsenko and Wanner,2000). The two groups in 2000 each suggested ways of increasing accuracy, efficiency and speed with which gene deletions can occur. In this project a combination of the two methods was followed, by using a highly recombinogenic strain that harbours the λ red gam recombination system (Yu *et al,* 2000) while the cassette used for initial gene replacement with a selectable marker and subsequent excision was based on the plasmid and findings of Datsenko and Wanner (2000). Expression of the Exo and Bet functions from the chromosome as opposed to from a plasmid has been shown to increase the efficiency of linear transformation (Murphy, 1998).

The mutants acquired from the linear recombination experiments and other strains with high copy numbers of the ORFs of the genes (complementation strains) where used to investigate the possibility that genes *xapB, yegT, yeiM* and *yeiJ* were

centrally involved in nucleoside transport in *E. coli*. The experiments performed using these mutants were nucleoside uptake experiments, as well as more specific experiments to determine the gene specificity through competition experiments, and kinetic studies to work out specific affinities of each gene product with regard to nucleoside transport in the cell.

Genes *xapB*, *yegT*, *yeiM* and *yeiJ* were assessed in a variety of ways to determine whether they encode permeases and whether permease activity is related to nucleoside transport. It was of interest to see if these putative nucleoside transporters had similar functions to the established nucleoside transporters NupC and NupG which have been shown to have a central role in nucleoside transport across the inner membrane of *E. coli*. A variety of experiments were used to assess the function of the genes, namely radiolabelled nucleoside transport experiments as well as growth experiments where the mutants were grown under minimal conditions with a set of different physiological nucleosides as the sole carbon source. Uptake of the nucleoside by the monogene strains would show that the remaining functional gene product has an affinity for the nucleoside and is able to transport it into the cell.

E. coli BLR cells were transformed with high copy number plasmids, engineered with one of the ORFs allowing over expression of the gene product once the plasmids were induced. These strains were used in nucleoside uptake experiments, during which radio labelled adenosine was used as substrate. The labelled adenosine was monitored for time-dependent uptake by the *E. coli* cells that were over expressing just one of the putative nucleoside transporters.

Kinetic parameters of radioactive adenosine uptake were also determined by concentration-dependent experiments, so as to test for the way adenosine is taken up by the cell and have a better understanding of whether adenosine is a main substrate in nucleoside transport by the particular nucleoside transporter.

Finally adenosine uptake was challenged in competition experiments which were performed by over-saturating the adenosine uptake assay with other physiological nucleosides and monitoring the effect of each one on adenosine transporter.
The findings from these experiments established that all four putative nucleoside transporters XapB, YegT, YeiM and YeiJ were able to transport adenosine in the assays at similar levels as the positive control. Adenosine transport was inhibited strongly (more than 65%) in the cases of YegT, YeiM and YeiJ when cytidine was present in the assay, whilst NupC and NupG have both been shown in the past to be able to transport.

Some controversial results were obtained in the study of transporters XapB, YegT, YeiM and YeiJ. In particular in some instances, the transport assay showed that a nucleoside transported by a specific overexpressed transporter was able to inhibit adenosine uptake indicating a better affinity for this competing nucleoside. However, when the competing nucleoside was supplied as the sole carbon source in the monogene strain with the transporter in question intact, the bacterial strain was not able to grow. This was the case for example when transporter YeiM was overexpressed in the BLR strain, where cytidine was shown to be the nucleoside that mostly inhibited adenosine uptake in the competition assay. However cytidine did not support growth of the corresponding monogene strain with only YeiM intact, when supplied in minimal medium as the sole carbon source for growth. Similarly, although not supporting growth when used as the sole carbon source, uridine inhibited adenosine uptake when YegT and XapB were provided as the nucleoside transporters.

The cases of YegT and XapB and uridine may be instances in which more than one transporter is required to initially allow uptake of uridine in order to elicit expression of other nucleoside transport systems such as YegT and XapB. Alternatively this could be an indication that at 5mM, uridine can be toxic to cells expressing these systems because they concentrate thymidine at high quantities in the cell (Beskid *et al,* 1981, Nørholm and Dandanell, 2001). The NupG monogene strain on the other hand was able to grow normally on uridine as well as the wild type strain MG1655.

In recombinant *E. coli* cells NupG transported uridine with apparent Km of 20-30 μM, showing a high affinity system (Xie *et al,* 2004). This could be further indication that initially NupG takes up uridine and other uptake systems could be induced as a secondary mechanism to deal with high concentrations of uridine and then use it as an energy source, if needed

Inosine although inhibiting adenosine uptake at 0.5mM when YegT is overexpressed then did not support growth of the YegT monogene strain MPG579 in the growth experiments. However growth on inosine as a sole carbon source was not good as a whole for any of the strains including the wild type strain MG1655. This could be because the high concentration of inosine in the growth experiments actually inhibits growth although it can be transported by the transporters when present in lower concentrations in the inhibition assay. A different explanation is that the transport of a specific nucleoside by one or more of the other transport systems (for example NupC and / or some of the others in the case of cytidine) is necessary for the initial uptake of small amounts of the nucleoside and full induction of the other systems (such as YeiM) can then occur. Therefore in the competition experiments where YeiM is overexpressed but there are copies of all six nucleoside transporters in the chromosome of the *E. coli* BLR strain, the nucleoside (such as cytidine) is transported in the cell initially and inhibits adenosine uptake. Indeed cytidine did not support growth of any of the monogene strains except poor growth of MPG582 which has only NupC intact.

On the other hand high concentrations of cytidine resulted in slow growth of wild type strain MG1655 as compared to growth of MG1655 on glucose. This could be an indication that although the transporters are able to transport cytidine very efficiently, cytidine at 5mM is too high a concentration for the bacterial strain to grow optimally.

Thymidine did not support growth of strain MPG579 (*yegT* monogene strain) or indeed, monogene stains expressing any other nucleoside transporter except YeiM and NupG. However it did inhibit adenosine transport in the competition assays in

strains overexpressing YegT and XapB. This could be because thymidine requires at least one copy of one of the two systems in order to initially transport thymidine into the cell and then express some of the other nucleoside systems. In the case of YegT, in competition experiments all tested nucleosides inhibited uptake of adenosine. However, the YegT monogene strain MPG579 only grew when xanthosine or adenosine were the carbon source. This could be explained by the fact that the absence of the other nucleoside transporters in the monogene strain abolished growth because one or some of the other transporters are necessary for uptake of the nucleosides. However it does seem that YegT is able to be expressed and concentrate adenosine and xanthosine so as to support growth when these nucleosides were provided as the sole carbon source.

Another controversial result was shown in cases where a nucleoside supported growth of a monogene strain, when used as the sole carbon source, but did not have an inhibitory effect on adenosine uptake when the specific transporter was overexpressed. Generally in the case of YeiM it seems that it can be expressed as the sole transport system and it can support growth on all nucleosides except cytidine and uridine. Thymidine and inosine supported poor growth of the NupC monogene strain MPG581 however they did not inhibit adenosine uptake in the competition experiments. This could be because thymidine and inosine are able to support growth of the particular strain at the elevated concentration of 5mM but at the concentration of 0.5mM in the competition assay they do not have such an obvious effect.

Thymidine also was shown to increase adenosine uptake when added in the competition assay when YeiJ was the nucleoside transporter overexpressed. This could mean that there are multiple binding epitopes for adenosine, and possibly other nucleoside transporters that were not determined here, that in the presence of thymidine are expressed and bind adenosine at higher levels.

A similar insight into this phenomenon has been described for the regulator CytR. The central role of cooperativity in regulation is highlighted by the fact that binding of the inducer, cytidine, to CytR is coupled to CytR-CRP cooperativity and this

underlies the mechanism for induction. Similar interactions at the different promoters of the CytR regulon coordinate expression of the transport proteins and enzymes required for nucleoside catabolism but also provide differential expression of these genes. Gavigan and colleagues (Gavigan *et al,* 1999) showed that CytR binds specifically to multiple sites in the *E.coli deoP* promoter, thereby providing competition for CRP binding to CRP operator site 1 (CRP1) and CRP2 as well as cooperativity.

The monogene strains provide a set of strains which are good working models because a clear result is obtained on the individual role of each one of the transporters. However the experiments do not provide insight into the effect the transporters have on each other, or knowledge on the regulatory mechanisms of action of the transporters which control their expression, and the impact different nutrients and substrates have on expression.

Experiments involving the monogene strains and the use of nucleosides as the sole carbon source, have the limitation that the levels of expression are low thus, giving a low measurable signal.

Uptake assays, kinetic assays and competition experiments were performed on strains which had one copy of the whole set of genes and additionally had one of the genes overexpressed and, although such experiments give a strong indication on the utilisation of a specific substrate by the overexpressed gene, this work is fundamentally different from attempting to detect differences between the six monogene strains that are studied, since nucleoside transport by other systems may occur in the background of the strain since the whole set of genes is present as single copies on the chromosome.

Nevertheless despite the special conditions of these experiments, the results obtained show that these genes do indeed encode nucleoside transporters, strongly reflected by the radiolabelled nucleoside experiments and carbon source experiments.

From the kinetic analysis of the transporters it was possible to determine Km values for the fours transporters with adenosine as substrate. The Km values as detailed in chapter 4 were in the same order of magnitude as mammalian nucleoside transporters and that shown for xanthosine uptake by XapB (Huang *et al,* 1994, Ritzel *et al,* 1998, Nørholm and Dandanell, 2001, Toan *et al,* 2003). Although this is encouraging, in order to establish further profiles on substrate specificities for the unknown transport proteins, it is not possible to draw direct comparisons with work performed in the current study and work cited in this thesis, since each research group has used different expression vectors (namely *E. coli* whole cells, or vesicles, *Xenopus* oocytes or yeast cells), different substrates (a variety of nucleosides such as cytidine, guanosine, xanthosine etc) or simply different transporters, as their main focus.

6.1.3 Implementing the nucleoside transporter-negative model in a nucleoside analogue test bioassay

Following the work on gene characterisation on a functional level it was interesting to test if the mutant strains were able to transport nucleoside analogues. AZT was used as a model substrate since it has been studied before and it is a well known compound used to combat HIV infections. The rationale of this work was to try to develop a high throughput assay that could be used as a first screening method to determine the way by which various nucleoside analogues are entering the cell.

Pharmaceutical companies dedicate a number of resources as well as time to develop assays which can be used to assess the potential of newly discovered drugs and in particular nucleoside analogues. A large number of novel nucleoside analogues which can potentially be used as chemotherapeutic and antiviral agents are developed every year.

The industry of drug development is fast moving and the use of high throughput processes has permitted medical chemists to prepare massive libraries of compounds that can be funnelled through high throughput assays to identify lead candidates for drug development (Gardner *et al,* 2004). However the approval of new chemotherapeutic agents is declining because selectivity and potency are requisite but not sufficient criteria for a successful drug candidate. Other criteria include validation of the biochemical target in the disease state, safety profiles, appropriate pharmacokinetic and metabolic properties of the candidate drug and chemical characteristics of the compound that ensure the compounds can be synthesised and produced in a large scale in an economical way (Smith *et al,* 2001).

Traditionally viral bioassays involve the measurement of plaque reduction, via plaque inhibition tests. Also the effect of the compound is assayed based on cell numbers as a direct indication of the compound's biological activity (Krenitsky *et al,* 1983). Chemo-sensitivity tests of compound toxicity on cells' viability are measured by OD and they have been reported in literature extensively for a variety of antiviral

nucleoside analogues such as troxacitabine (Gourdeau *et al,* 2001) or gemcitabine (Skehan *et al,* 1990 Sandler and Ettinger, 1999).

Skehan *et al,* (1990) developed the SRP method for measuring the cellular protein content of adherent and suspension cultures in 96-well microtiter plates. The method is suitable for ordinary laboratory purposes and for very large-scale applications, such as the National Cancer Institute's disease-oriented in vitro anticancer-drug discovery screen, which requires the use of several million culture wells per year.

The National Cancer Institute implemented a large-scale *in vitro* drug-screening program in 1990 that required a very efficient automated assay of drug effects on tumour cell viability or growth and microculture assay based on metabolic reduction of MTT was developed. Because the SRB assay has practical advantages for large-scale screening, however, it has been adopted for routine use in the NCI in vitro antitumour screen (Rubinstein *et al,* 1990).

Novel high throughput methodologies find application in the building of physical and chemical properties of candidate drugs, taking them from interesting compounds to therapeutic drug candidates. However there is a need for quick ways of assessing whether a new drug would be useful on a clinical level and a way of coming closer to this issue is by the development of a first screen which would determine whether the new drug in question is effective or not. The method should give a reliable result quickly and would qualify a compound to move to the next level of assessment tests (Skehan *et al,* 1990).

The null strain which is a transport negative strain for nucleosides is therefore a suitable candidate for the development of this bioassay in the field of nucleoside analogue development research since it fulfils a number of requirements for the test development.

First of all it is a bacterial model, and therefore it can be used rapidly and in a high throughput manner to assess drug transport through the bacterial membrane and into

the cell. A bacterial model as an option for this kind of test means that it can be easy to work with in the lab, and it can be used in a high throughput manner to assess a large number of potential nucleoside drug analogues at the same time. Secondly, the lack of nucleoside transporters in the null strain makes it a unique candidate to assess uptake of nucleoside analogues. The rationale behind this is that theoretically this strain would not be able to actively transport any nucleosides or their toxic analogues. Therefore if a drug is able to pass through the membrane of the null strain it must have mechanisms of transport other than via the studied bacterial nucleoside transporters.

Also the monogene strains can be used in the same test to estimate and assess the possible ways of entry of specific nucleoside analogues and therefore provide some insight to the way these drugs are taken up by the cell. However in order to use these strains and, in order to make them attractive candidates for a quick test of this sort, a way of assessing the cells viability was needed. The viability of the cell is a direct indication of the cell's ability to take up the toxic nucleosides.

Summarising the findings of this work on this pilot bioassay two bioassays were performed to assess the best approach to monitor cell activity, via OD and light emission. OD measures cell density and allows accurate conclusions on how a factor such as a toxic substance affects culture growth. Light is a signal of promoter activity and also cell viability since it requires $FMNH_2$. Although light emission demonstrated limitations as compared to OD, regarding fluctuation and its dependence on expression of *luxCDABE* from P_{tac} it is a more attractive reporter in the context of commercialising a bioassay for use in a rapid and high throughput manner. The initial step towards such an assay which was defined in this work was that mutations in nucleoside transport systems have an effect on growth under certain conditions which was demonstrated as AZT inhibited light emission in all strains, noticeably in all mutated strains as compared to the wild type strain.

6.2 Nucleoside transporters and their diverse role in species

Collectively the information gathered can constitute a picture in which cellular nucleoside transport in many species from *Homo sapiens* to *E. coli* is supported by a number of different transporters. However the variation of factors studied in different groups causes some limitations in direct comparisons of results for different transporters belonging to different species (Hansen *et al,* 1987).

In human tissues, five Na+-dependent nucleoside transport subtypes (N1-N5) have been recognized. To date, genes encoding N1-N3 transporters have been cloned and belong to a CNT gene family (Chapter 4, table 4.1). Genes encoding N4 and N5 activities have not been identified (Kong *et al,* 2004).

In other enteric bacteria there are also multiple nucleoside transporters. For example in *Salmonella typhimurium* LT2 there are reports of nucleoside transporter-encoding genes *nupC, nupG, xapB* and *yegT* (Zalkin and Nygaard, 1996, Hansen and Dandanell, 2005), while in three species of *Salmonella* (namely *Salmonella typhimurium, Salmonella typhi* and *Salmonella paratyphi*) genes *yeiM* and *yeiJ* are absent (http://www.ncbi.nlm.nih.gov/entrez). Also the virulent *E. coli* strain O157 has been shown to lack the gene *xapB* (http://www.ncbi.nlm.nih.gov/entrez).

In *Bacillus subtilis* there are four known transporters involved in nucleoside transport encoded by *nupC, yxjA* (now known as *nupG*), *ydhL* (now known as *pbuE*) and *pbuG*. *nupC* has been shown to be involved in pyrimidine nucleoside transport and is regulated by DeoR (Saxild *et al,* 1996), while *nupG* was shown to be involved in purine nucleoside transport (Johansen *et al,* 2003). The expression of *nupG* is co-regulated by a common mechanism that regulates the expression of genes *ydhL* and *pbuG*. Hypoxanthine and guanine negatively regulate the operons or genes, except *ydhL* which is regulated positively (Johansen *et al,* 2003). Two genes, *yutK* and *ydcT* are also thought to be involved in nucleoside transport, *yutK* being annotated as a nucleotide co-transporter (Helmann *et al,* 2001).

It therefore seems that in all studied organisms ranging from bacteria to mammals there is more than one transporter for entry of nucleosides into the cells. There are conserved pathways which are found in all species such as the purine specific nucleoside transporters and other pathways of nucleoside transport that are unique to a species or that have been obliterated from the genome if there was no biological significance for the specific transporter to be there.

6.2.1 Comparative genomics of *E. coli* novel nucleoside transporters and human nucleoside transporters

In the quest for understanding the nucleoside transport systems analysed in this project, primary structure can be indicative of specificity and therefore roles of each one of the transporters. In table 6.1 there is a summary of the identity in amino acid sequence between the *E. coli* nucleoside transporters studied here and human nucleoside transporters. It is striking that YeiM and YeiJ have identity in amino acid sequence of more that 25% and in many cases more than 30%. Some of the most striking percentage identities are shown in figure 6.1 below, namely the identity of amino acid residues between YeiM and hCNT3 and YeiJ and hCNT1.

Mammalian CNT1 proteins have about 650 residues and exhibit more than 64% sequence identity to the CNT2 proteins (about 660 residues). They show no similarity to either equilibrative transporters or other sodium-dependent transporters. Homologues of the human CNT family are present in a diverse spectrum of living organisms ranging from bacteria to other mammals. The best studied example in other organisms is NupC of *E. coli* homologies of which are shown in figure 6.2 against human, rat, rabbit, pig, fish and *C. elegans* CNTs. As seen from the alignments there are areas of high conservation of all species as well as areas of high conservation among bacterial nucleoside transporters and human CNTs.

This is indicative of evolutionary descent of these transporters as well as their central role in cell metabolism.

	NupG	XapB	YegT	YeiJ	YeiM
hCNT1	12.7%	12.4%	13.3%	30.1%	31.1%
hCNT2	13.3%	10.9%	15.4%	29.4%	27.7%
hCNT3	12.5%	12.8%	11%	31.9%	32.2%

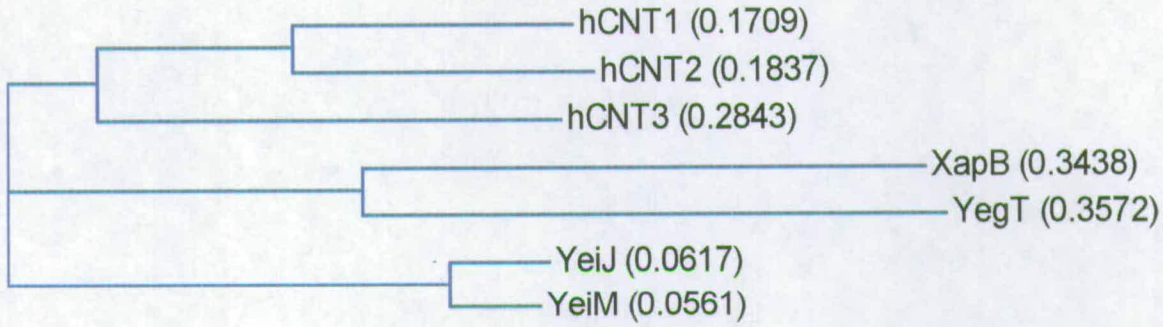
Table 6.1: Percentage identity between bacterial nucleoside transporters and human nucleoside transporters.

Figure 6.1: Relation of *E. coli* nucleoside transporter genes to human nucleoside transporter genes. A.) tree showing phylogenetic relation between genes *xapB* and *yegT* (paralogues of *nupG*), *yeiM* and *yeiJ* (paralogues of *nupC*) and *hCNT1* and *hCNT2*. **B.)** Amino acid sequence comparison between *hCNT3* and *yeiM* **C.)** Amino acid sequence identity between *hCNT2* and *YegT* and **D.)** Amino acid sequence identity between *hCNT3* and *YeiJ*.

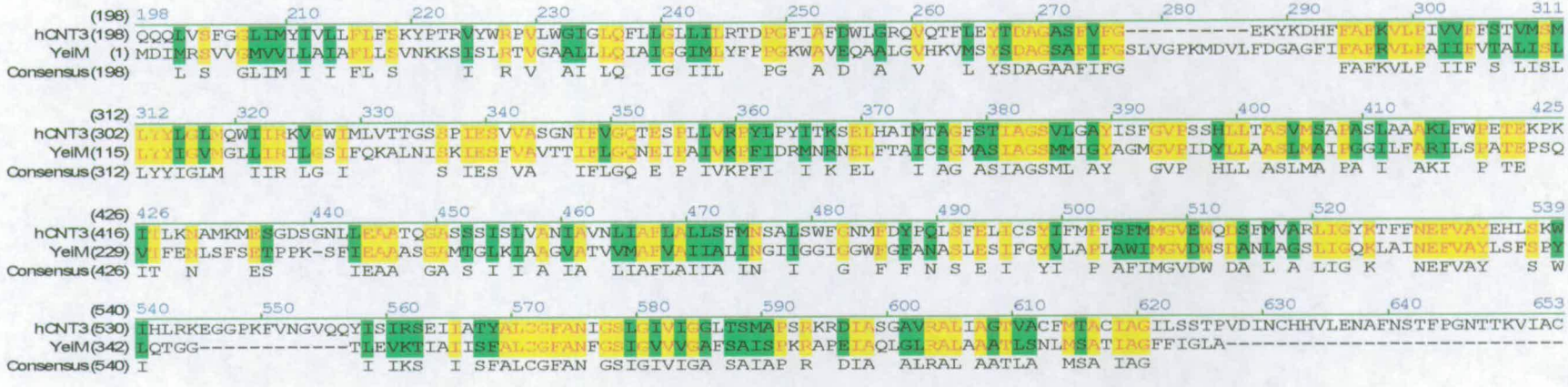
KEY:

Yellow= identical aa

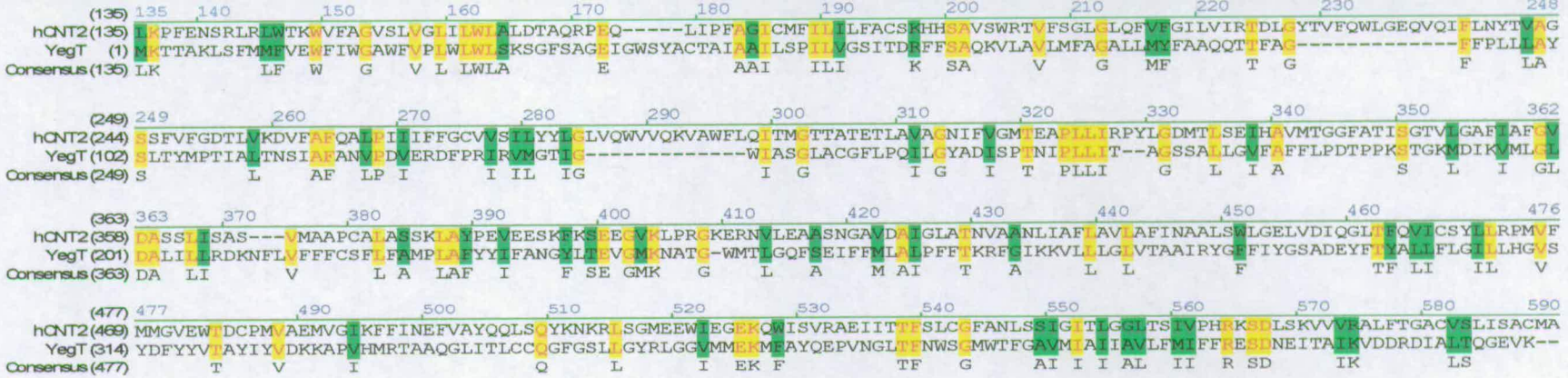
Green= conservative aa


hCNT1 (0.1709)
hCNT2 (0.1837)
hCNT3 (0.2843)
XapB (0.3438)
YegT (0.3572)
YeiJ (0.0617)
YeiM (0.0561)

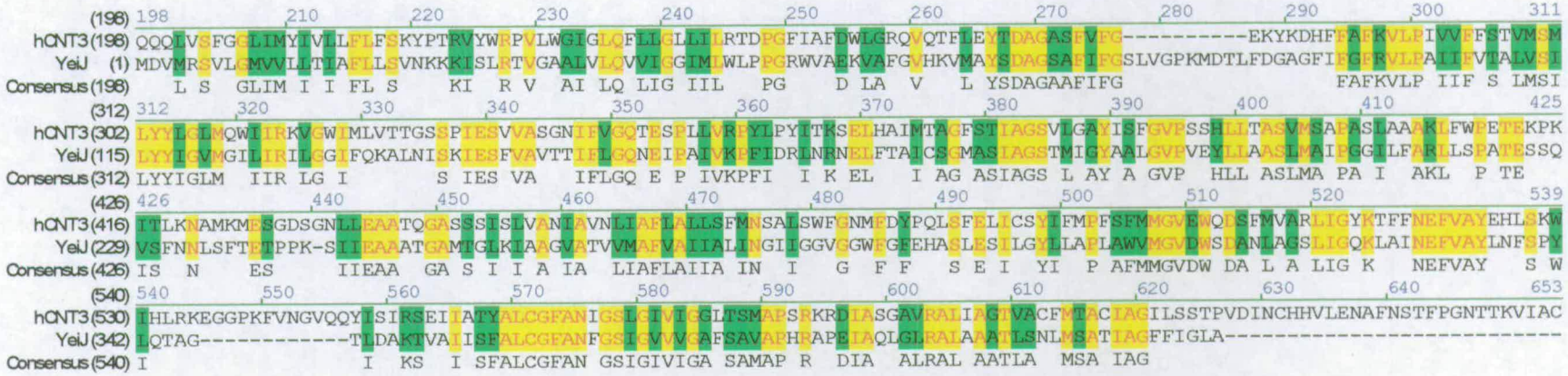
A.


(198) 198 210 220 230 240 250 260 270 280 290 300 311
hCNT3(198) QQQLVSFGGLIMYIVLLFLFSKYPTRVYWRPVLWGIGLQFLLGLLILRTDPGFIAFDWLGRQVQTFLEYTDAGASFVPG----------EKYKDHFFAFKVLPIVVFFSTVMSM
YeiM (1) MDIMRSVVGMVVLLIAIAFLLSVNKKSISLRTVGAALLLQIAIGGIMLYFPPGKWAVEQAALGVHKVMSYSDAGSAFIPGSLVGPKMDVLFDGAGFIFAFRVLPAIIFVIALISL
Consensus(198) L S GLIM I I FL S I R V AI LQ IG IIL PG A D A V L YSDAGAAFIFG FAFKVLP IIF S LISL
(312) 312 320 330 340 350 360 370 380 390 400 410 425
hCNT3(302) LIYLLLYQWIIRKVGWIMLVTTGSSPIESVVASGNIPVGQTESPLLVRFYLPYITKSPIHAINTAGFSTIAGSVLGAYISFGVPSSHLLTASVMSADASLAAAKLFWPETEKPK
YeiM(115) LGYIGVVGLLIRILGSIFQKALNISKIESFVAVTTIFLGGNEIPATVKFIDRMNRNEIFTAICSGMASIAGSMMIGYAGMGVPIDYLLAASLMAIPGGILFARILSPATEPSQ
Consensus(312) LYYIGLM IIR LG I S IES VA IFLGQ E P IVKPFI I K EL I AG ASIAGSML AY GVP HLL ASLMA PA I AKI P TE
(426) 426 440 450 460 470 480 490 500 510 520 539
hCNT3(416) IILKNAMKMESGDSGNLLEAATQGASSSISIVANIAVNLIAFLALLSFMNSALSWFGNMFDYPQLSFELICSYIFMPFSFMMGVEWQDSFMVARLIGYKTFFNEFVAYEHLSKW
YeiM(229) VIFENLSFSPTPPK-SFIEAAASGAMTGLKIAAGVATVVMAFVAIIALINGIIGGIGGWFGFANASLESIFGYVLAPLAWIMGVDWSIANLAGSLIGQKLAINEFVAYLSFSPY
Consensus(426) IT N ES IEAA GA S I I A IA LIAFLAIIA IN I G F F N S E I YI P AFIMGVDW DA L A LIG K NEFVAY S W
(540) 540 550 560 570 580 590 600 610 620 630 640 653
hCNT3(530) IHLRKEGGPKFVNGVQQYISIRSEIIATYALCGFANIGSLGIVIGGLTSMAPSRKRDLASGAVRALIAGTVACFMTACIAGILSSTPVDINCHHVLENAFNSTFPGNTTKVIAC
YeiM(342) LQTGG----------TLEVKTIAIISFALCGFANFGSIGVVVGAFSAISPKRAPEISQLGLRALAAATLSNLMSATIAGFFIGLA----------
Consensus(540) I I IKS I SFALCGFAN GSIGIVIGA SAIAP R DIA ALRAL AATLA MSA IAG

B.



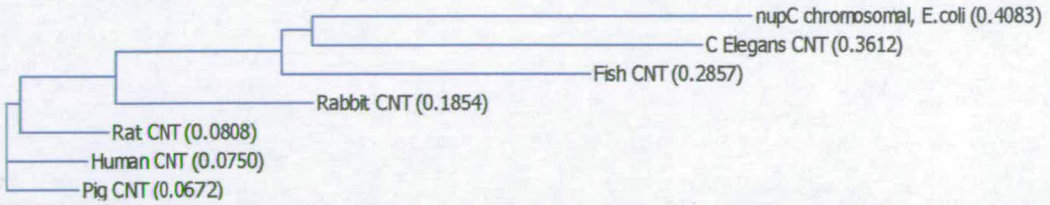
C.



D.

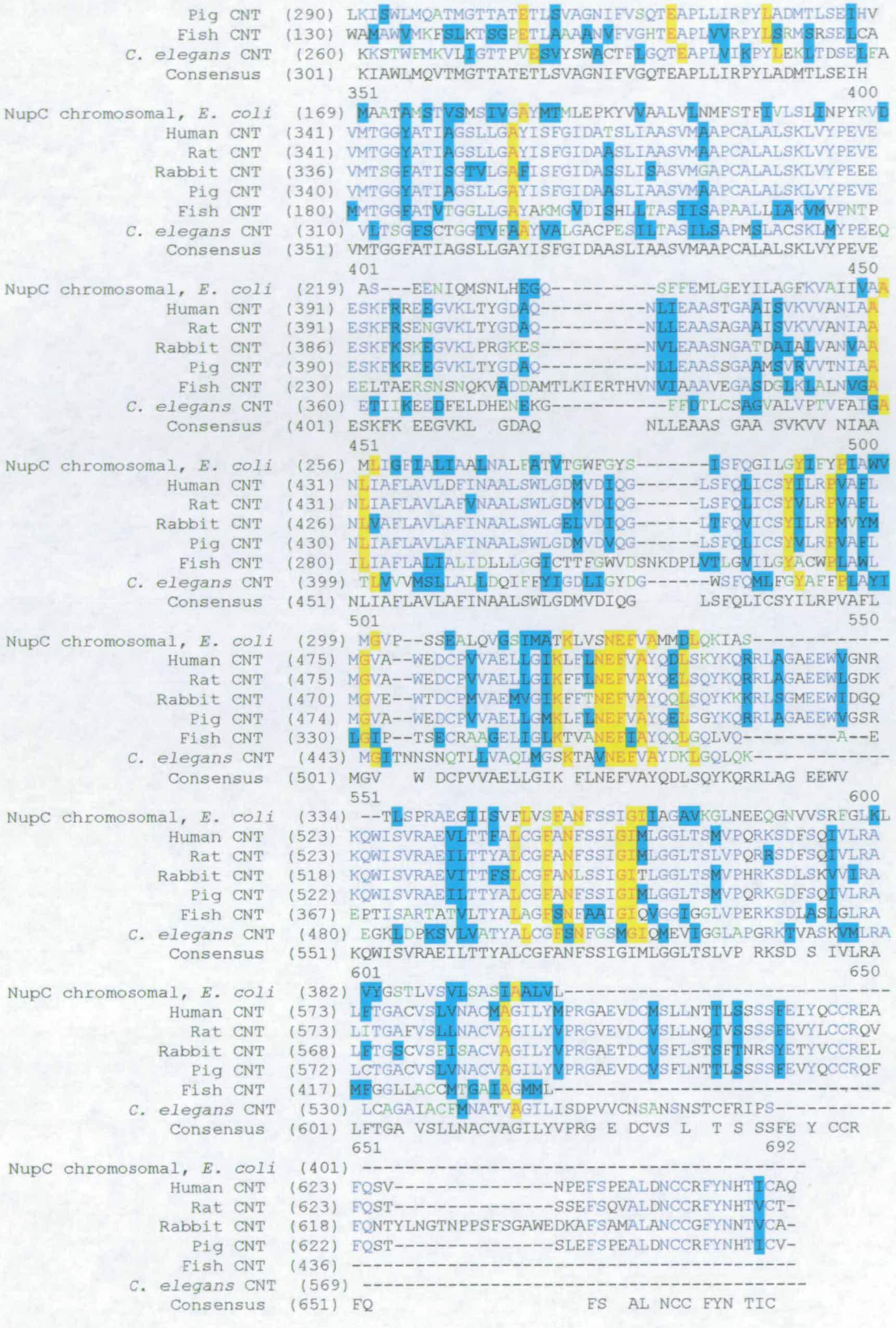
Figure 6.2 Relation of *E. coli* nucleoside transporter NupC to CNT1 transporter from human, rat, rabbit, pig, fish and *C. elegans*. A.) tree showing phylogenetic relation between CNT genes of different species. **B.)** Amino acid sequence identity between CNT proteins and NupC.

KEY:
Yellow: identical aa
Blue: conservative aa



A).

```
NupC chromosomal, E. coli     (1) ----------------------------------------------------
                Human CNT     (1) MENDPSRRRESISLTPVAKGLENMGADFLESLEEGQLPRSDLSPAEIRSS
                  Rat CNT     (1) MADNTQRQRESISLTPMAHGLENMGAEFLESMEEGRLPHSHSSLPEGEGG
               Rabbit CNT     (1) --MEKASGRKSLALSTAENGIENAG---LELTEEGINSEQTRRMEVQGHS
                  Pig CNT     (1) MEDNTPRQRDPISLTSVANGLENMGAELLESLEEGRAPGSDSSPAEVGGG
                 Fish CNT     (1) ----------------------------------------------------
           C. elegans CNT     (1) ----------------------------------------MVLNKKVHQ
                Consensus     (1)         R  ISLT    GLEN G   LE  EEG
                                  51                                              100
NupC chromosomal, E. coli     (1) ----------------------------------------------------
                Human CNT    (51) WSEAAPKPFSRWRNLQPALRARSFCREHMQLFRWIGTGLLCTGLSAFLLV
                  Rat CNT    (51) LNKAERKAFSRWRSLQPTVQARSFCREHRQLFGWICKGLLSTACLGFLMV
               Rabbit CNT    (46) LSDDVRPATHQRSYLQPLTKARTFCQRHASLFKKILLGLLCLAYAAYFLA
                  Pig CNT    (51) WSKAGPEHLGR-RSLQPALRVRRFCREHTQLFRWICTGLLCTAFAAFLLI
                 Fish CNT     (1) ----------------------------------------------------
           C. elegans CNT    (10) NKEKNLISVTKSSLEHNANFYQRIYQKICPFIGPAAIVAFLLVYHGYLIA
                Consensus    (51)          R   LQP    R FC  H  LF  I  GLL  A  AFLL
                                 101                                             150
NupC chromosomal, E. coli     (1) ----------------------------------------------------
                Human CNT   (101) ACLLDFQRALALFVLTCVVLTFLGHRLLKRLLGPKLRRFL---KPQGHPR
                  Rat CNT   (101) ACLLDLQRALALLIITCVVLVFLAYDLLKRLLGSKLRRCV---KFQGHSC
               Rabbit CNT    (96) ACILDFQRALALFVITCLVILVLLLHFLKKFLGKKLTRCL---KPFKNSQ
                  Pig CNT   (100) ACLLDFQRALALFVLFCVVLFFLAHSLLKRLLGPKLLRCV---KPLRHPC
                 Fish CNT     (1) ----------------------------------------------------
           C. elegans CNT    (60) AGIHNYQKASPLIYVTLFFWLCFVINFIVGTKNFKVVYADFAGKLNDFAK
                Consensus   (101) ACLLDFQRALAL VITCVVL  L    LKR LG KL R L   K
                                 151                                             200
NupC chromosomal, E. coli     (1) ----------------------------MDRVLHFVLALAVVAILALLVSS
                Human CNT   (148) LLLWFKRGLALAAFLGLVLWLSLDTSQRPEQLVSFAGICVFIALLFACSK
                  Rat CNT   (148) LSLWLKRGLALAAGVGLILWLSLDTAQRPEQLVSFAGICVFLVLLFAGSK
               Rabbit CNT   (143) LRLWIKRVFAGVSLVGLILWLALDTAQRPEQLISFAGICMFVLILFACSK
                  Pig CNT   (147) LNLWFKRGLALAAFLGLVLWLVLDTAQRPEQLVSFGGICVFILLLFAGSK
                 Fish CNT     (1) ----------------------MPQLNEIFPRLVCGLGIVVFVTLAWSIST
           C. elegans CNT    (110) RRAFVPITFKLAFAGALFAYTIVESLSEPTRLTGFGGYVFFVIFMVVFSN
                Consensus   (151) L LW KR  ALAA LGLILWL LDTAQRPEQLVSFAGICVFVILLFA SK
                                 201                                             250
NupC chromosomal, E. coli    (24) DRKKIRIRYVIQLLVIEVLLAWFFLNSDVGLGFVKGFSEMFEKLLGFANE
                Human CNT   (198) HHCAVSWRAVSWGLGLQFVLGLLVIRTEPGFIAFEWLGEQIRIFLSYTKA
                  Rat CNT   (198) HHRAVSWRAVSWGLGLQFVLGLFVIRTEPGFIAFQWLGDQIQVFLSYTEA
               Rabbit CNT   (193) HHSAVSWRTVFWGLGLQFVFGLLVIRTDPGFIAFQWLGDQVQIFLAYTVA
                  Pig CNT   (197) HHRAVSWRAVSWGLGLQFALGLFVIRTEPGFIAFQWLGDQIQIFLSYTEA
                 Fish CNT    (30) DRRQFPWRIVIGGLLLQFTLAVLVLQTDLGQKTFKHIGDGFQKLMSTVDA
           C. elegans CNT   (160) RPRKINWNIVTSALIFHYCLALIVLKWPTGRWFFDQLSQLIIGLLEYAQV
                Consensus   (201) HHRAVSWR VSWGLGLQFVLGLLVIRTDPGFIAF WLGDQIQIFLSYT A
                                 251                                             300
NupC chromosomal, E. coli    (74) GTNFVFGSMNDQGL-----AFFFLKVLCPIVFISALIGILQHIRVLPVII
                Human CNT   (248) GSSFVFGEALVK-------DVFAFQVLPIIVFFSCVISVLYHVGLMQWVI
                  Rat CNT   (248) GSSFVFGEALVK-------DVFAFQVLPIIIFFSCVMSVLYYLGLMQWVI
               Rabbit CNT   (243) GSSFVLGDTLVN-------DVFAFQSLPIIIFFGCVMSILYYLGLVQWVV
                  Pig CNT   (247) GSSFVFGEALVK-------DVFAFQVLPIIVFFSCAMSVLYYVGLMQWVI
                 Fish CNT    (80) GSGFLFSTGANNPLGDSLLATFAFGVLPTVIFFSSLMSVLYHLGVMQRIV
           C. elegans CNT   (210) GAKFVFGFIAGPPNICDLAPVFIFTSLQTLIYFSAVVALFYFGIIQIAL
                Consensus   (251) GSSFVFGEALV        DVFAFQVLPIIIFFSCVMSVLYYLGLMQWVI
                                 301                                             350
NupC chromosomal, E. coli   (119) RAIGFLLSKVNGMGKLESFNAVSSLIILGQSENFIAYKDILGKISRNRMYT
                Human CNT   (291) LKIAWLMQVTMGTTATETLSVAGNIFVSQTEAPLLIRPYLADMTLSEVHV
                  Rat CNT   (291) LKIAWLMQVTMGTSATETLSVAGNIFVSQTEAPLLIRPYLADMTLSEVHV
               Rabbit CNT   (286) QKIAWFLQVTMRTTATETLAVAGNIFVGMTEAPLLIRPYLADITLSEIHA
```


Pig CNT (290) LKISWLMQATMGTTATETLSVAGNIFVSQTEAPLLIRPYLADMTLSEIHV
Fish CNT (130) WAMAWVMKFSLKTSGPETLAAAANVFVGHTEAPLVVRPYLSRMSRSELCA
C. elegans CNT (260) KKSTWFMKVLIGTTPVESVYSWACTFLGQTEAPLVIKPYLEKLTDSELFA
Consensus (301) KIAWLMQVTMGTTATETLSVAGNIFVGQTEAPLLIRPYLADMTLSEIH
351 400
NupC chromosomal, E. coli (169) MAATAMSIVSMSIVGAYMTMLEPKYVVAALVLNMFSTFIVLSLINPYRVL
Human CNT (341) VMTGGYATIAGSLLGAYISFGIDATSLIAASVMAAPCALALSKLVYPEVE
Rat CNT (341) VMTGGYATIAGSLLGAYISFGIDAASLIAASVMAAPCALALSKLVYPEVE
Rabbit CNT (336) VMTSGFATISGTVLGAFISFGIDASSLISASVMGAPCALALSKLVYPEEE
Pig CNT (340) VMTGGYATIAGSLLGAYISFGIDAASLIAASVMAAPCALALSKLVYPEVE
Fish CNT (180) MMTGGFATVTGGLLGAYAKMGVDISHLLTASIISAPAALLIAKVMVPNTP
C. elegans CNT (310) VLTSGFSCTGGTVFAAYVALGACPESIITASILSAPMSLACSKLMYPEEQ
Consensus (351) VMTGGFATIAGSLLGAYISFGIDAASLIAASVMAAPCALALSKLVYPEVE
401 450
NupC chromosomal, E. coli (219) AS---EENIQMSNLHEGQ----------SFFEMLGEYILAGFKVAIIVAA
Human CNT (391) ESKFRREEGVKLTYGDAQ----------NLIEAASTGAAISVKVVANIAA
Rat CNT (391) ESKFRSENGVKLTYGDAQ----------NLLEAASAGAAISVKVVANIAA
Rabbit CNT (386) ESKFKSKEGVKLPRGKES----------NVLEAASNGATDAIAIVANVAA
Pig CNT (390) ESKFKREEGVKLTYGDAQ----------NLLEAASSGAAMSVRVVTNIAA
Fish CNT (230) EELTAERSNSNQKVADDAMTLKIERTHVNVIAAAVEGASDGLKLALNVGA
C. elegans CNT (360) ETIIKEEDFELDHENEKG----------FFDTLCSAGVALVPTVFAIGA
Consensus (401) ESKFK EEGVKL GDAQ NLLEAAS GAA SVKVV NIAA
451 500
NupC chromosomal, E. coli (256) MLIGPLALIAALNALFATVTGWFGYS-------ISFQGILGYIFYPIAWV
Human CNT (431) NLIAFLAVLDFINAALSWLGDMVDIQG------LSFQLICSYILRPVAFL
Rat CNT (431) NLIAFLAVLAFVNAALSWLGDMVDIQG------LSFQLICSYVLRPVAFL
Rabbit CNT (426) NLVAFLAVLAFINAALSWLGELVDIQG------LTFQVICSYILRPMVYM
Pig CNT (430) NLIAFLAVLAFINAALSWLGDMVDVQG------LSFQLICSYVLRPVAFL
Fish CNT (280) ILIAFLALLIALIDLLLGGICTTFGWVDSNKDPLVTLGVILGYACWPLAWL
C. elegans CNT (399) TLVVVMSLLALLDQIFFYIGDLIGYDG------WSFQMLFGYAFFPLAYI
Consensus (451) NLIAFLAVLAFINAALSWLGDMVDIQG LSFQLICSYILRPVAFL
501 550
NupC chromosomal, E. coli (299) MGVP--SSEALQVGSIMATKLVSNEFVAMMDLQKIAS----------
Human CNT (475) MGVA--WEDCPVVAELLGIKLFLNEFVAYQDLSKYKQRRLAGAEEWVGNR
Rat CNT (475) MGVA--WEDCPVVAELLGIKFFLNEFVAYQELSQYKQRRLAGAEEWLGDK
Rabbit CNT (470) MGVE--WTDCPMVAEMVGIKFFTNEFVAYQQLSQYKKKRLSGMEEWIDGQ
Pig CNT (474) MGVA--WEDCPVVAELLGMKLFLNEFVAYQELSGYKQRRLAGAEEWVGSR
Fish CNT (330) LGIP--TSECRAAGELIGLKTVANEFIAYQQLGQLVQ---------A--E
C. elegans CNT (443) MGITNNSNQTLIVAQLMGSKTAVNEFVAYDKLGQLQK----------
Consensus (501) MGV W DCPVVAELLGIK FLNEFVAYQDLSQYKQRRLAG EEWV
551 600
NupC chromosomal, E. coli (334) --TLSPRAEGIISVFLVSFANFSSIGIIAGAVKGLNEEQGNVVSRFGLKL
Human CNT (523) KQWISVRAEVLTTFALCGFANFSSIGIMLGGLTSMVPQRKSDFSQIVLRA
Rat CNT (523) KQWISVRAEILTTYALCGFANFSSIGIMLGGLTSLVPQRRSDFSQIVLRA
Rabbit CNT (518) KQWISVRAEVITTFSLCGFANLSSIGITLGGLTSMVPHRKSDLSKVVIRA
Pig CNT (522) KQWISVRAEILTTYALCGFANFSSIGIMLGGLTSMVPQRKGDFSQIVLRA
Fish CNT (367) EPTISARTATVLTYALAGFSNFAAIGIQVGGIGGLVPERKSDLASLGLRA
C. elegans CNT (480) EGKLDPKSVLVATYALCGFSNFGSMGIQMEVIGGLAPGRKTVASKVMLRA
Consensus (551) KQWISVRAEILTTYALCGFANFSSIGIMLGGLTSLVP RKSD S IVLRA
601 650
NupC chromosomal, E. coli (382) VYGSTLVSVLSASIAALVL----------
Human CNT (573) LFTGACVSLVNACMAGILYMPRGAEVDCMSLLNTTLSSSSFEIYQCCREA
Rat CNT (573) LITGAFVSLLNACVAGILYVPRGVEVDCVSLLNQTVSSSSFEVYLCCRQV
Rabbit CNT (568) LFTGSCVSFISACVAGILYVPRGAETDCVSFLSTSFTNRSYETYVCCREL
Pig CNT (572) LCTGACVSLVNACVAGILYVPRGAEVDCVSFLNTTLSSSSFEVYQCCRQF
Fish CNT (417) MFGGLLACCMTGALAGMML----------
C. elegans CNT (530) LCAGAIACFMNATVAGILISDPVVCNSANSNSTCFRIPS----------
Consensus (601) LFTGA VSLLNACVAGILYVPRG E DCVS L T S SSFE Y CCR
651 692
NupC chromosomal, E. coli (401) ----------
Human CNT (623) FQSV----------NPEFSPEALDNCCRFYNHTICAQ
Rat CNT (623) FQST----------SSEFSQVALDNCCRFYNHTVCT-
Rabbit CNT (618) FQNTYLNGTNPPSFSGAWEDKAFSAMALANCCGFYNNTVCA-
Pig CNT (622) FQST----------SLEFSPEALDNCCRFYNHTICV-
Fish CNT (436) ----------
C. elegans CNT (569) ----------
Consensus (651) FQ FS AL NCC FYN TIC

B).

6.3 Elucidation of the role of the novel nucleoside transporters in *E. coli*

A possible explanation of the role of genes such as *yeiM* and *yeiJ* in nucleoside transport and in relation to their presence and role in different species can be given by the study of four fast growing bacteria by Karlin and Mrazek (2000) and Karlin *et al,* (2001). In a study of predicted highly expressed (PHX) genes for the completely sequenced genomes of the four fast-growing bacteria *Escherichia coli, Haemophilus influenzae, Vibrio cholerae,* and *Bacillus subtilis*, Karlin *et al,* (2001) report the finding of two open reading frames (*yeiM* and *yeiJ*) of unknown function which they hypothesise may encode nucleoside transporters, which are highly expressed in *V. cholerae.* The homologue of *yeiM* was also highly expressed in *H. influenzae* but was not PHX in *E. coli* and *B. subtilis*. In the top 20 PHX genes, Karlin *et al,* (2001) isolated the ORFs of *E. coli yeiM* and *yeiJ*, *H. influenzae* HI0519, and *B. subtilis yutK*.

In the above study a gene was considered to be PHX if its codon frequencies were close to those of the ribosomal proteins, major translation/ transcription processing factors, and chaperone/degradation standards but strongly deviant from the average gene codon frequencies. PHX genes identified by their codon usage frequencies among prokaryotic genomes commonly include those for ribosomal proteins, major transcription/translation processing factors (several occurring in multiple copies), and major chaperone/degradation proteins. Also PHX genes generally include those encoding enzymes of essential energy metabolism pathways of glycolysis, pyruvate oxidation, and respiration (aerobic and anaerobic), genes of fatty acid biosynthesis, and the principal genes of amino acid and nucleotide biosynthesis. Gene classes generally not PHX include most repair protein genes, virtually all vitamin biosynthesis genes, genes of two-component sensor systems, most genes encoding regulatory proteins, and most genes expressed in stationary phase or during starvation.

Some insight into why *yeiM* and *yeiJ* are potentially highly expressed in *V. cholerae* but not in *E. coli* and *B. subtilis*, and what role *yeiM* and *yeiJ* could have in

nucleoside transport could be answered by looking at the effect of nucleoside transport on the ability of *V. cholerae* to form biofilms. *V. cholerae* is the aetiological agent of Asiatic cholera, and is also a natural inhabitant of aquatic environments (Faruque *et al,* 1998). In habitats of this kind, *V. cholerae* has been hypothesized to switch between the smooth and rugose colony morphotypes, a capacity that was proposed by White in 1938 to contribute to its environmental survival (White, 1938). HapR, a positive regulator of *hap* haemaglutinin/protease (HA/P) expression, represses the rugose colony morphotype in the smooth colonial variant.

Recent studies have shown that HapR negatively regulates biofilm formation through its role in a quorum-sensing signal transduction system (Yildiz *et al,* 2004). The products of HapR regulated genes are involved in the catabolism of nucleotides for use as an energy and carbon source as well as for the synthesis of nucleotides (Zalkin and Nygaard, 1996). Expression of these genes is under the control of the CytR repressor, DeoR repressor and cAMP–CRP complex in *E. coli*, the same systems used for control of nucleoside transport (Zalkin and Nygaard, 1996)

6.3.1 Relation of regulatory patterns for novel nucleoside transporters

Findings show that CytR acts as a repressor of exopolysaccharide biosynthesis and, in turn, biofilm formation in *V. cholerae* (Haugo and Watnick. 2002). As the CytR mutant exhibits an increased capacity to take up nucleosides, it was suggested that an increase in intracellular cytidine levels is used as a signal for switching to the biofilm growth mode from the planktonic one (Zalkin and Nygaard, 1996). Alternatively, increased nucleoside levels in the cell may lead to an increase in the formation of activated sugar nucleotide intermediates, which would in turn increase exopolysaccharide synthesis. As a result, the increased production of exopolysaccharide could lead to an increased capacity to attach to surfaces and form biofilm matrix material. HapR is part of the quorum-sensing system of *V. cholerae*, and its expression is repressed by the action of LuxO (Zhu *et al,* 2002). LuxO is an

RpoN-dependent response regulator, its activity is controlled by the phosphorylation state of the protein and, when LuxO is phosphorylated at low cell density, it prevents *hapR* transcription (Klose *et al,* 1998; Zhu *et al,* 2002). In contrast, at high cell density, LuxO is dephosphorylated, expression of *hapR* is increased, and HapR downregulates the expression of pathogenesis genes through its action on *aph*A (Kovacikova and Skorupski 2002). Therefore the roles of genes *yeiM* and *yeiJ* could be to encode transporters of nucleosides such as cytidine or one of its derivatives under the regulation of CytR, which aid quorum sensing.

Although the mechanism proposed for *V. cholerae* for biofilm formation and quorum sensing might not be the same for *E. coli* the role of the gene may be the same in the different species. Therefore *yeiM* and *yeiJ* might have a role in quorum sensing in *E. coli* as well and data presented in this thesis support their involvement in nucleoside uptake with a possible down-regulatory control by CytR, where nucleosides or derivatives are sensed as indicators of cell density.

The different modes of regulation as well as the secondary roles in which nucleoside transport is implicated in survival of the bacterial strains, might be ways of explaining the reason why bacteria and other organisms have a set of more than one transport system of entry for nucleosides. It may be simpler for cells to sense a unique derivative, rather than an abundant component which can be transported into the cells by several routes.

A large number of CRP binding sites clustered in groups were found in all genes of interest to this project, indicating that the genes are using CRP as transcription factor similarly to NupC and NupG. The large number of recurring *crp* sites in the sequences upstream of the genes as well as within the gene ORFs indicate that the genes may be upregulated the same way as NupC and NupG.

crp is the site where cAMP-CRP binds and where RNA polymerase starts the transcription of the adjacent operon. CRP works on several operons encoding mainly catabolic pathways. *crp* seems to be the binding site for genes *yeiM, yeiJ, yegT* as

determined by analysis using the software available in http://www.gene-regulation.com/cgi-bin/pub/programs/match/bin/match.cgi.

CRP may have evolved to assist auxotrophic organisms which cannot control the carbon source that they will use as energy source. *E. coli* can be exposed to a number of different energy sources depending on the host environment. *E. coli* has two main habitats, the gut of animals and humans and environmental habitats such as soil, water, food etc. Gene regulation in *E. coli*, and more specifically of nucleoside transport has been shown to be controlled by negative and positive regulators. Negative control systems may have evolved from a system which once functioned constitutively and now works as a fail-safe mechanism and an environmentally responsive way of controlling gene expression.

Further clues to the regulatory mechanism of the genes of interest in this project are revealed by a study which concentrated on the isolation of a potential new family member of the CRP-FNR family of transcription regulators. The CRP-FNR family of transcription regulators contains at least 70 members that respond to a variety of metabolic and environmental signals to control a diverse range of physiological functions in an equally diverse range of bacteria (Spiro, 1994; Guest *et al,* 1996; Vollack *et al,* 1999). The family can be divided into three broad phylogenetically distinct groups designated CRP, NtcA and FNR, although there is considerable variation within the groups and also some overlap between the groups with respect to functional specificities for signal recognition, DNA-binding, and RNA polymerase interaction (where known). Analysis from 38 members of the CRP-FNR family indicated that YeiL is a distant relative of the CRP group (Anjum *et al*, 2000). The *yeiL* gene is located in the *nfo-fruA* region of the *E. coli* linkage map at 48-57 min where it is flanked by a cluster of unidentified reading frames, (URFs) (Anjum *et al*, 2000).

Putative functions have been assigned to the products of these URFs by similarity searches. They include two potential guanosine kinases, YeiC and YeiI, which are 38% identical (61% similar), and two potential pyrimidine nucleoside transport

proteins, YeiJ and YeiM, which are 88% identical. There is also a potential nucleoside hydrolase, YeiK, which is 40% identical to the inosine-uridine nucleoside hydrolase of *Crithidia fasciculata* (Degano *et al,* 1996). 26 of the 27 residues surrounding the catalytic site of this trypanosomal purine salvage pathway enzyme YbeK are conserved in YeiK. Thus it would appear that several genes involved in nucleoside metabolism are located in the *nfo-fruA* region of the *E. coli* chromosome together with a novel paralogue of CRP and FNR.

The *yeiL* gene of *E. coli* was predicted to encode a new and distinct member of the CRP-FNR family of transcriptional regulators that might have an iron- sulphur centre and a reversible intra- or inter-molecular disulphide bond. (Anjum *et al,* 2000)

The *yeiL* gene encodes the third of three CRP-FNR proteins found in *E. coli*. The YeiL factor is associated with homologous proteins from *Bacillus anthracis*. YeiL homologues branch deeply and their position may have to be reconsidered in an expanded database. YeiL is of interest because of its complex mode of regulation, distinct properties of a cofactor-binding site, and suggested functional role in stationary growth phase (Lange and Hengge-Aronis, 1991). YeiL conserves principal structural elements of CRP and FNR such as the sensory domain, the K-helix for the dimer interface, and the DNA recognition helix of FNR.

However, the specific residues in the helix for DNA binding are not conserved. A target gene for YeiL is not known. The regulator is unlikely to bind cAMP, as only one residue is conserved of six involved in nucleotide binding.

Transcription of *yeiL* increases several-fold in *E. coli* during the stationary growth phase under aerobic conditions indicating a role in starvation survival. Its expression is autoregulated, influenced by FNR, and also dependent on RpoS, the major factor for gene activation in stationary cells. FNR exerts a repressive effect on *yeiL* expression. It is thought that YeiL may play a role in enhancing the use of nucleobases during nitrogen starvation (Anjum *et al,* .2000).

6.3.2 *E. coli* adaptability and survival in diverse conditions

E. coli has been shown in many cases to be adaptable enough and to have sufficient systems for uptake and metabolism, that it can colonise successfully a variety of environments. For example in the mid-1970s, Chapman and coworkers found that most enterics could use aromatic compounds, with the ability to utilise hydroxyphenylacetic acid (HPA) being the most widespread. It was the work of Cooper and Skinner (1980) on the ability of *E. coli* to mineralise 3- and 4-hydroxyphenylacetic acids (3HPA and 4HPA) that delineated for the first time a complete catabolic pathway in enteric bacteria. Later, Burlingame and Chapman (1983) reported that many laboratory strains and clinical isolates of *E. coli* can catabolise a range of aromatic compounds.

The intestine of warm-blooded animals, the primary habitat of *E. coli*, contains some 400 to 500 different bacterial species, with *E. coli* being the most abundant (making up about 1% of the total faecal bacterial flora). The success of *E. coli* in the gut ecosystem is thought to reflect its ability to occupy different ecological niches. Thus, since *E. coli* grows both anaerobically and aerobically, it is able to colonize intestinal habitats in which oxygen offers some ecological advantage. Such a habitat could be one in close proximity to epithelial cells, where oxygen molecules might pass from the blood through the epithelium, to the microbes attached to it (Neidhardt, 1996). By assimilating such molecules, *E. coli* may be important in developing and maintaining the oxygen-free conditions and low oxidation-reduction potential which favours strict anaerobes in the large intestine. While embedded within the mucus layer overlying intestinal epithelial cells, *E. coli* grows with a generation time of 40 to 80 min. In contrast, the population of *E. coli* cells in the ceacal luminal contents is essentially static with respect to growth and excretion in the faeces (Blum *et al,* 1996).

Although *E. coli* is a highly successful commensal of the intestines of warm-blooded animals, some strains are pathogens of the enteric, urinary, pulmonary, and nervous systems. This facultative anaerobe must be able to survive and grow outside the

animal host to effect successful inter-host spread. *E. coli* is not famous for extra-corporeal existence, but it nonetheless shares with its soil-inhabiting relatives, such as *Klebsiella* species, the ability to thrive under a wide range of physical and chemical conditions, including adaptation for survival over long periods of non-growth (Diaz *et al,* 2001, Daniels 1990).

Soil, water, sediment, and perhaps food are other habitats of *E. coli*, and the bacterium might spend comparable times in each of these main habitats (Savageau, 1983). While pollution from human and animal sources may be the most important source of *E. coli* in the environment, the fact that this bacterium was found in pristine tropical waters, where it remained physiologically active and grew at rates dependent on nutrient levels, suggests that it can be a natural inhabitant in these environments and that it may be part of a previously established community. *E. coli* can also replicate and survive in some soil protozoa. Since protozoa are widely distributed in soils and effluents, they may also constitute an environmental reservoir for transmission of this enterobacterium. Thus, unlike host-specific or obligate parasites, *E. coli* is a highly adaptable micro organism with an extensive repertoire of metabolic and regulatory genes that may facilitate the colonization of widely different environments (Peekhaus and. Conway. 1998).

Aromatic compounds are highly abundant in soil and water, and therefore it is obvious that they can constitute a normal carbon source for *E. coli* when this bacterium reaches its extra-intestinal habitat. Although it is still not known which substrates *E. coli* grows on in the large intestine and which pathways provide it with the metabolic advantage necessary for it to compete with the many other bacteria with which it shares this habitat, it is likely that aromatic compounds can also be a frequent carbon source for *E. coli* in the animal gut. Plant constituents are the major sources of aromatic compounds in the gastrointestinal tract. Minor sources of aromatic compounds in the human gut arise from some steroids, drugs and food constituents (additives, colorants, and contaminants) (Diaz *et al,* 2001).

Estimates suggest that between 3 and 25 g of protein and peptides enter the large bowel every day from the diet, as well as from endogenous sources such as host tissues, bacterial debris, pancreatic enzymes, and other secretions. In the large gut, these substances are depolymerized by a mixture of pancreatic endopeptidases and bacterial proteases and peptidases. The resulting short peptides and amino acids then become available for fermentation by many intestinal anaerobes such as *Bacteroides*, *Lactobacillus*, *Bifidobacterium*, *Clostridium*, and *Streptococcus* species, generating a wide range of phenolic and indolic compounds in a series of deamination, transamination, decarboxylation and dehydrogenation reactions. Phenol, *p*-cresol, HPA, hydroxyphenylpropionic acid (HPP), and hydroxybenzoic acid are the principal products of tyrosine fermentation in the human large intestine, while PA, phenylpropionic acid (PP), and benzoic acid are produced from phenylalanine. PP may, however, also be formed from tyrosine (Diaz *et al,* 1998, Diaz *et al,* 2001).

One cluster of genes, the *hca* cluster has been identified to be involved in PP metabolism. At the 5' end of the *hca* cluster is located a gene, *hcaT* which encodes a protein with similarity to MFS proteins and it is a predicted transporter. Interestingly, the predicted permease HcaT (379 aa) is smaller than the majority of MFS proteins (about 400 aa), and some common amino acid sequences that characterise the members of this superfamily are not found in the primary structure of HcaT. However the HcaT protein shows the highest similarity to nucleoside transport proteins such as NupG and XapB of *E. coli* (Seeger *et al,* 1995).The regulator of the cluster called *hcaR* is of the LysR type of regulators (Diaz *et al,* 1998, Diaz *et al,* 2001). To date, the HcaT protein has been classified at the Transport Commission website (http://www.biology.ucsd.edu/msaier/transport/), as the only representative of the phenylpropionic permease family of MFS proteins. However, an experimental demonstration that this protein is involved in the uptake of PP in *E. coli* as opposed to nucleoside for example is still required.

The aromatic compounds generated by the intestinal anaerobes meet a variety of fates in the body. Thus, they may be detoxified by glucuronide or sulfate conjugation or may remain unabsorbed and be voided in the faeces and urine. They can also be

completely broken down under local aerobic conditions in the large intestine, as a result of the action of facultative anaerobes such as *E. coli*. A second major source of aromatic compounds in the animal gut involves the dietary plant constituents such as ferulic and caffeic acid, which result in HPP acids as well as different flavonoid glycosides that are ingested in daily quantities of 1 to 2 g by humans. A collaborative bacterial catabolism, i.e., syntrophic interactions among microorganisms, of flavonoids in the human gut has been suggested (Peppercorn, and Goldman. 1971). Thus, a number of obligately anaerobic bacteria from the human intestinal flora, e.g., *Bacteroides* species, are capable of cleaving the glycosidic bond of flavonoids, generating the corresponding aglycones which are then subject to ring cleavage by different bacteria.

6.3.3 Quorum sensing and nucleoside metabolism

Quorum sensing is the regulation of gene expression by producing and responding to secreted autoinducers (AIs) whose concentrations reflect the population density (Bassler, 1999), and it exists commonly in bacteria. Gram-negative bacteria use acylated homoserine lactones as AIs, and gram-positive bacteria use oligopeptides. When the cell density is high, the binding of AIs to cell receptors regulates gene expression for a variety of phenotypes, such as production of virulence factors protein production, bioluminescence, biofilm formation, and plasmid conjugation (Ren *et al*, 2004). Generally, each bacterial species uses its own signal; however, a common AI-2 signal has been discovered for inter-species communication. Recent studies have found that *E. coli* O157:H7 uses AI-2 to control the expression of virulence factors, type III secretion, chemotaxis, flagellar synthesis, and motility and that *E. coli* K-12 uses AI-2 to control chemotaxis, motility, and flagellar synthesis. In addition, *E. coli* RP37 uses AI-2 to control cell aggregation (Sperandio *et al*, 1999). In *Salmonella enterica* serovar Typhimurium, AI-2 concentrations are maximal in mid-exponential- phase growth, and it is degraded in the stationary phase by an unknown mechanism. *E. coli* DH5 does not produce AI-2 due to a 60-amino-acid

deletion stemming from a 1-bp deletion that results in early truncation of *luxS* (formerly *ygaG*) (Ren *et al*, 2004).

With the ability to quantify the expression of all genes in one organism whose genome has been sequenced, DNA microarrays have been successfully used to study bacterial responses to different stimuli such as heat shock and other stresses, quorum sensing, anaerobic metabolism, sporulation, and biofilm formation.

The Gram-negative quorum-sensing systems were first discovered and extensively characterized in the marine *Vibrios*. While species-specific quorum sensing apparently allows recognition of self in a mixed population, it seems likely that in such situations, bacteria also need a mechanism or mechanisms to detect the presence of other species. Additionally, it is conceivable that it is useful for bacteria to have the ability to calculate the ratio of self to others in mixed populations, and in turn, to specifically modulate behaviour based on fluctuations in this ratio. This cell-to-cell signalling in prokaryotes that leads to co-ordinated behaviour has been termed quorum sensing. This type of signalling can have profound impacts on microbial community structure and host-microbe interactions (Fuqua *et al,* 1994).

Evidence for the existence of these types of complex mechanisms of cell-cell communication came first from studies of the Gram-negative bioluminescent shrimp pathogen *Vibrio harveyi*. Genetic analysis of the *V. harveyi* quorum-sensing circuit revealed that this bacterium uses two different autoinducer signals to regulate light production and a number of other target outputs (Nealson *et al,* 1970). In this regulatory circuit, it appears that one signal is used for intra-species communication and the second signal is used for inter-species communication. Genetic and biochemical evidence suggests that an AHL-type autoinducer termed AI-1 mediates intra-species communication and AI-2 mediates inter-species cell-cell communication (Federle and Bassler, 2003). AI-1 is detected by the two-component hybrid sensor kinase–response regulator protein LuxN (Bassler *et al,* 1993). The second *V. harveyi* signal, AI-2, is synthesized by an enzyme called LuxS (Schauder *et al,* 2001, Surette *et al,* 1999).

Detection of AI-2 requires two proteins: LuxP and LuxQ. LuxP is a soluble periplasmic protein, resembling sugar binding proteins such as RsbB of *E. coli*, which recognises ribose, part of the building block of nucleosides (Jun et al, 2002). LuxP is the AI-2–binding protein, and the LuxP–AI-2 complex interacts with the second protein required for detection, LuxQ. LuxQ is similar to LuxN and is a hybrid two-component protein containing both sensor kinase and response regulator modules. Sensory information from both LuxN and LuxPQ is transduced to the phosphotransferase protein LuxU, and LuxU transmits the signal to the downstream response regulator LuxO. In the *V. harveyi* quorum-sensing circuit, under conditions of low cell density (i.e., in the absence of autoinducers), LuxN and LuxQ act as kinases. Analogous to all other two-component sensors, they autophosphorylate on conserved histidine residues. Phosphate is sequentially transferred to the conserved aspartate residue on the response regulator modules of the hybrid proteins, and then to a conserved histidine on LuxU, and finally to the conserved aspartate on LuxO. Phospho-LuxO is indirectly responsible for repression of the *luxCDABE* operon, which encodes the luciferase enzymes necessary for light production (Fuqua *et al,* 1994).

At high cell density (i.e., when the autoinducers are present), interactions of the signals with LuxN and LuxPQ cause the sensors to switch from kinase mode to phosphatase mode. Phosphate flows backward through the circuit traveling via the same histidine and aspartate residues used in the forward direction. The net result is the dephosphorylation and inactivation of LuxO, leading to derepression of *luxCDABE*. Under this condition, a transcriptional activator called LuxR (not similar to other LuxR-type proteins) induces transcription of *luxCDABE* and light production. Production and detection of AI-1 appear to be unique to *V. harveyi* and the closely related species *Vibrio parahaemolyticus.*

Thus in *V. harveyi,* a sensory system responds to the presence of AI-2 by regulating the expression of luciferase. Presumably, *V. harveyi* could respond to either the AI-2 synthesized by its own species or the collective concentration of AI-2 synthesized by an entire community of different species. *E. coli* and *Salmonella* also produce AI-2,

and this production is maximal at low pH, high osmolarity and in the presence of a preferred carbon source such as glucose (Surette *et al,* 1999; 1999; DeLisa *et al,* 2001).

However, until recently, it was not clear whether organisms outside the genus *Vibrio*, including *Escherichia* and *Salmonella*, encode a sensory apparatus for detecting and responding to AI-2. Some components corresponding to those of the *Vibrio* systems are present in the classical genetic model organisms *E. coli* and *Salmonella enterica.* Both organisms encode a signal receptor of the LuxR family, SdiA, but not a corresponding signal-generating enzyme. Instead, SdiA of *Salmonella* detects and responds to signals generated only by other microbial species. On the other hand, *E. coli* and *Salmonella* encode the signal-generating component of a second system (a LuxS homologue that generates AI-2), but the sensory apparatus for AI-2 differs substantially from the *Vibrio* system (Ren *et al,* 2004).

Gene *luxS* is present in the genomes of a wide variety of Gram-negative and Gram positive bacteria including *E. coli, Salmonella typhi, Salmonella typhimurium.* Every bacterium whose genome contains a functional *luxS* gene has been shown to be capable of producing an activity that can be detected by an AI-2–specific *V. harveyi* reporter strain. Together, these two findings raise the possibility that unlike AHL- and AIP-mediated species-specific communication, AI-2 is a more universal signal that could promote interspecies bacterial communication. Furthermore, the finding that bacteria detect and integrate information supplied by multiple chemical signals with different origins suggests that the bacteria use a combinatorial approach to distinguish between one another that is reminiscent of the molecular mechanism underlying odour detection and differentiation in higher organisms. It is also possible that bacteria possessing multiple AHL or AIP circuits use combinatorial means to process the information provided by these intraspecies-specific signals (Sperandio *et al,* 2001, Reading and Sperandio, 2006).

One apparent complication is that *E. coli* and *S. enterica* have what appears to be an incomplete set of quorum-sensing components. Receptors have not yet been

identified for some of the putative signalling molecules and, conversely, these organisms have a receptor, SdiA, for which they fail to synthesize a signal ligand (Ahmer, 2004). The discovery of the first component led to premature claims of quorum sensing by these organisms. A second complication is that some of the putative quorum-sensing systems are entangled with metabolism. The degree of metabolism versus signalling in these systems is controversial and will take time to dissect fully (Reading and Sperandio, 2006).

DNA microarrays were used to study *E. coli* DH5 gene expression with and without the addition of *E. coli*-conditioned medium. The effect of stationary-phase quorum-sensing signals on AI-2 synthesis in the exponential phase was also investigated by using *E. coli* K-12 (De Lisa *et al,* 2001). This was the first study to investigate the effect of *E. coli* stationary-phase quorum-sensing signals on *E. coli* global gene expression by using DNA microarrays and to show that stationary-phase quorum-sensing signals repress AI-2 concentrations in the exponential phase. In total, three genes were induced and five genes including *nupC*, were repressed by both fresh and autoclaved stationary-phase supernatants. Furthermore, supernatant from *E. coli* DH5 stationary culture was found to repress *E. coli* K-12 concentrations in liquid cutures by 4.8-fold, suggesting that an additional quorum-sensing system in *E. coli* exists and that gene expression is controlled as a network with different signals working at different growth stages (Ren *et al,* 2004). Also *yeiK*, which was reported in this study to be a potential nucleoside hydrolase, was identified as one of the unknown genes that are related to quorum sensing (De Lisa *et al,* 2001).

Significantly the synthesis of YeiK (a potential nucleoside hydrolase) has been shown to have a regulatory link with YeiL. Studies have shown that the *yeiK* and *yeiL* genes are divergently transcribed from overlapping promoters. The complete genome of *Escherichia coli* (Blattner *et al*, 1997) was shown to encode a third member of the CRP-FNR family (YeiL). A BLAST analysis showed that YeiL is more closely related to members of the CRP-FNR family (16-22% identical, 42-45% similar) than to other known proteins. A phylogenetic tree derived from 38 members of the CRP-FNR family indicated that YeiL is a distant relative of the CRP group

(Gostick *et al*, 1999). As mentioned earlier, the *yeiL* gene is located in the *nfo-fruA* region of the *E. coli* linkage map where it is flanked by a cluster of unidentified reading frames (URFs). Putative functions have been assigned to the products of these URFs by similarity searches and they include the two potential guanosine kinases, YeiC and YeiI and the two potential nucleoside transport proteins studied here, YeiJ and YeiM (Anjum *et al*, 2000).

In a study of compromised formation of biofilms in *E. coli* after the effect of ursolic acid, *yeiE* was found to be among the genes induced. (Ren *et al,* 2005). Other genes induced were motility, chemotaxis and stress response genes. *yeiL,* the gene predicted to encode a CRP-like transcription regulator during stationary phase (Anjum *et al,* 2000) may function as a post exponential phase nitrogen starvation regulator in the *yei* cluster of genes. Therefore YeiM and YeiJ may be transporters of nucleosides under nitrogen starvation or under conditions which threaten the biofilm formation ability of *E. coli.*

6.4 Conclusion

In brief the sum of the experimental work in this project led to findings that show that gene products XapB, YegT, YeiM, and YeiJ can function as nucleoside permeases. This was proven by the use of multiple mutants in experiments involving growth on nucleoside as a sole carbon source. Preliminary experiments also gave an insight into developing a high throughput bioassay which could be developed further in order to produce a useful tool for nucleoside drug analogue assessment.

In future it would be useful to obtain more information on substrate specificity of these transporters with the hope to establish in detail specific transport patterns of the different permeases. The reason why there are six, and possibly more than six, nucleoside transport systems could be further elucidated from experiments showing substrate specificities and different affinities for substrate concentrations. Also regulation of the genes is an area that would shed light into the use of these transporters and could potentially answer when and under which specific conditions these transporters are being used for nucleoside uptake.

Nucleoside transporters can be important in determining intracellular bioavailability and disposition of therapeutic nucleoside analogues. Knowledge of such transporters is central in the evaluation and forecast of the kinetics, effectiveness and toxicities of nucleoside-derived therapeutic agents. In recent years, significant progress has been made in studies of nucleoside transporters (Kong *et al,* 2004). The cloning of major mammalian and other ENT and CNT transporters has opened the door for the study of structure, function, regulation, tissue and cellular localization of nucleoside transporters. Knowledge obtained from this research, together with detailed substrate and kinetic information, will allow the prediction of *in vivo* disposition and targeting of therapeutic nucleoside analogues.

Further structure-function analysis coupled with molecular simulation assists in the design of nucleoside drugs with improved pharmacokinetic and pharmacodynamic properties. It is not clear from the substrate specificities determined from

experimental data and from amino acid sequence identities from the alignments shown in figure 6.1 if the multiple bacterial nucleoside transport systems have direct analogies with the multiple recognised human nucleoside transport systems. However it is clear that an organism apparently as simple as *E. coli* has evolved a number of pathways with which it can adapt in numerous environments and can survive effectively, and this adaptability includes nucleoside transport. Knowledge gained from such organisms can be used to elucidate complex matters and problems such as disease therapy. Importantly it seems that organisms evolve alternative mechanisms of action so as to ensure their survival. The more important a biological pathway is to the survival of the organism, the larger the number of safety switches the organism will develop to ensure it is never deficient in such pathways.

"Everything should be made as simple as possible. But not simpler"

Albert Einstein

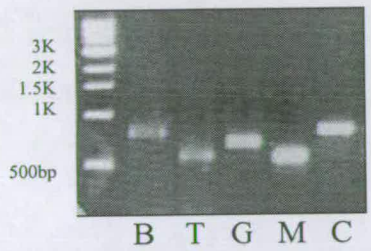
APPENDIX 1

Monogene strains as checked by ethidium bromide stained agarose gel electrophoresis for correct gene deletions and cassette insertions. DNA band sizes can be found in tables 3.1 and 3.2.

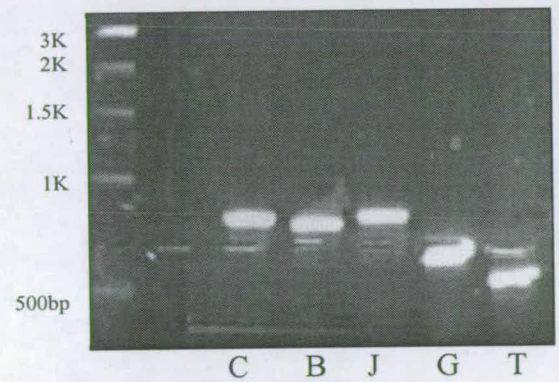
KEY:

Initials B, T, G, M, C, J stand for deleted genes *xapB, yegT, nupG, yeiM, nupC, yeiJ.* J::kan , M:: kan signifies a monogene strain in which the cassette has replaced genes *yeiJ* and *yeiM.*

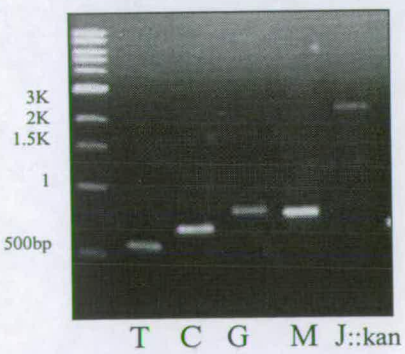
i) MPG583 (MG1655/Δ*xapB, yegT, nupG, yeiM, nupC*) yeiJ+



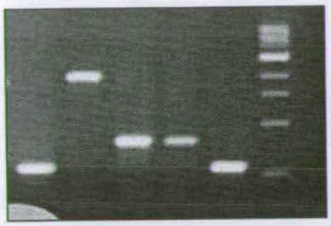
ii) MPG581 (MG1655/Δ*yegT, nupC, xapB, yeiJ, nupG*) yeiM+



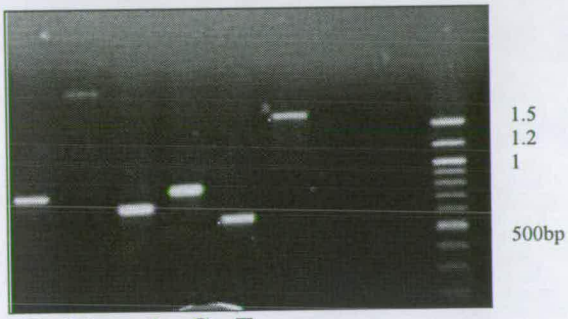
iii) MPG578 (MG1655/Δ*yegT, nupC, nupG, yeiM, yeiJ*) xapB+



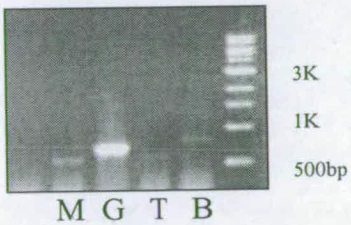
iv) MPG579 (MG1655/Δ*xapB, nupC, nupG, yeiM, yeiJ*) yegT+

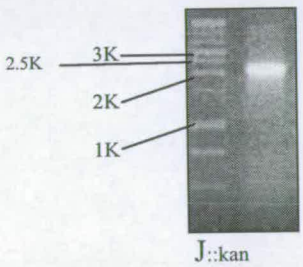


v) MPG580 (MG1655/Δ*yegT, nupC, xapB, yeiJ, yeiM*) nupG+



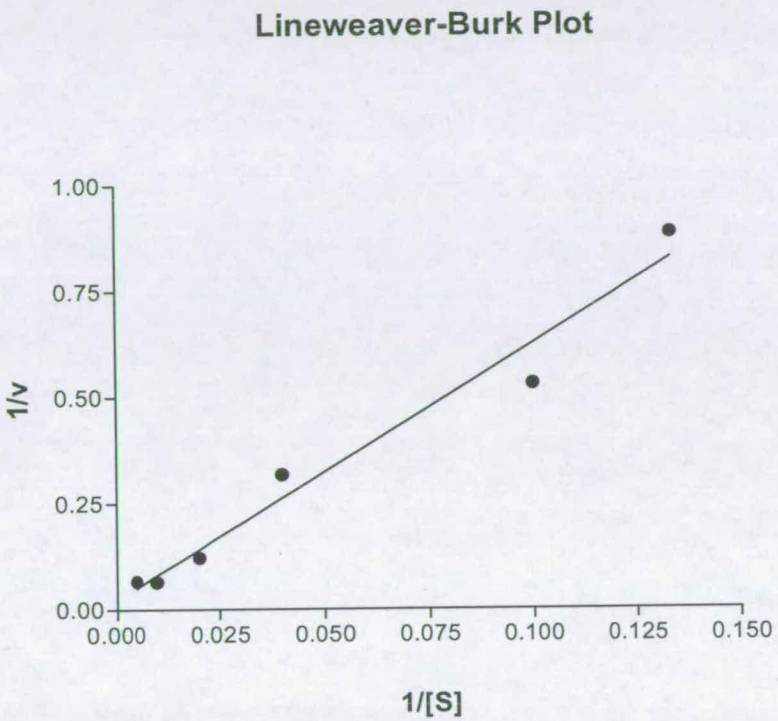
vi) MPG582 (MG1655/Δ*xapB, yegT, nupG, yeiM, yeiJ*) nupC+



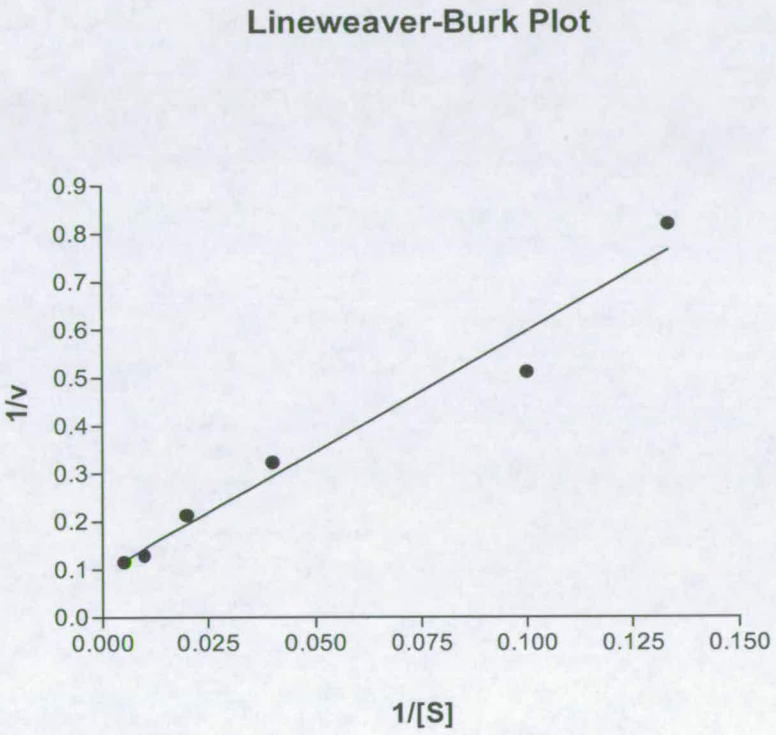


APPENDIX 2

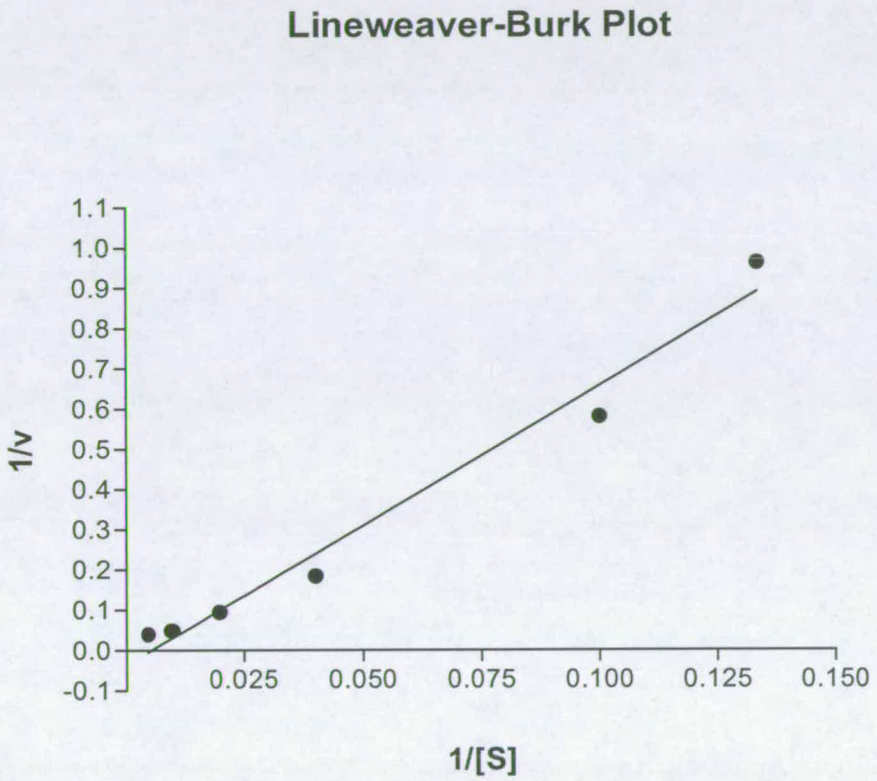
Lineweaver-Burk plots from which Km values are derived. The Km values determined via the software program GraphPad Prism 4 are 97.1, 65, 160.1 and 453.9μM for strains MPG591, MPG592, MPG593 and MPG594 which over express the nucleoside transporters XapB, YegT, YeiM and YeiJ respectively. Vmax values are also determined by the software and are summarised below each graph.



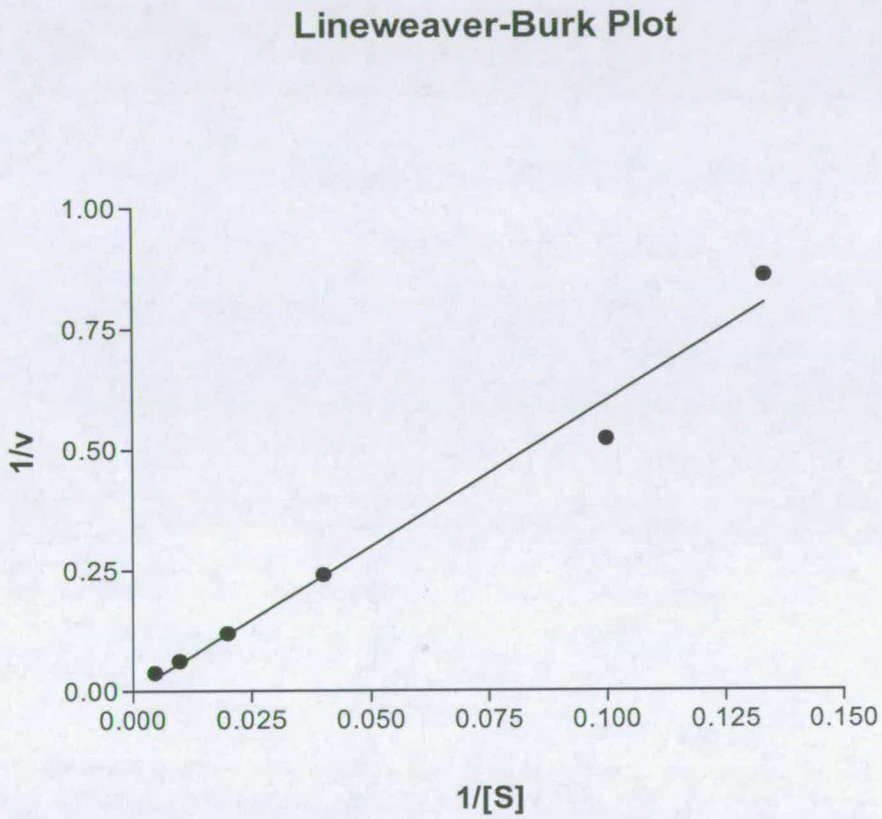
VMAX 24.62
Km 97.09
STRAIN: MPG591



VMAX 11.82
Km 65.02
STRAIN: MPG592



VMAX 47.15
Km 160.1
STRAIN: **MPG593**



VMAX 86.56
Km 453.9
STRAIN: **MPG594**

References

Abrahams JP, Leslie AGW, Lutter R and Walker JE (1994) Structure at 28Å resolution of Fi-ATPase from bovine mitochondria Nature, 370:621-628

Abramson J, Smirnova I, Kasho V, Verner G, Kaback HR (2003) So Iwata Structure and Mechanism of the Lactose Permease of *Escherichia coli* Science 1 301 (5633):610 – 615

Acimovic Y, Coe IR (2002) Molecular evolution of the equilibrative nucleoside transporter family: identification of novel family members in prokaryotes and eukaryotes Mol Biol Evol19 (12):2199-210

Ahmer BM (2004) Cell-to-cell signaling in *Escherichia coli* and *Salmonella enterica* Mol Microbiol 52(4):933-45

Albert A (1952) Selective toxicity Lect Sci Basis Med 2, 14-28

Alocilja EC, Radke SM (2003) Market analysis of biosensors for food safety Biosens Bioelectron 18(5-6):841-6

Ames GF (1986) Bacterial periplasmic transport systems: structure, mechanism, and evolution Annu Rev Biochem:55:397-425 Review

Anjum MF, Green J, Guest JR (2000) YeiL, the third member of the CRP-FNR family in *Escherichia coli* Microbiology 146 (12) :3157-70

Anthony BF, Hill HR (1988) Gram-positive bacteria: an overview and summary of session Rev Infect Dis 10 Suppl 2:S345-50 Review

Baldwin SA (1990) Molecular mechanisms of sugar transport across mammalian and microbial cell membranes Biotechnol Appl Biochem 12(5):512-6 Review

Baldwin, SA, Mackey, JR, Cass, CE & Young, JD (1999) Nucleoside transporters: molecular biology and implications for therapeutic development Molecular Medicine Today 5, 216-224

Barile M, Valenti D, Quagliariello E, Passarella S (1998) Mitochondria as cell targets of AZT (zidovudine) Gen Pharmacol 31(4), 531-8

Bassler BL (1999) How bacteria talk to each other: regulation of gene expression by quorum sensingCurr Opin Microbiol 2(6):582-7

Bassler BL, Wright M, Showalter RE, Silverman MR (1993) Intercellular signalling in *Vibrio harveyi*: sequence and function of genes regulating expression of luminescence Mol Microbiol 9(4):773-86

Baty D, Pattus F, Parker M, Benedetti H, Frenette M, Bourdineaud JP, Cavard D, Knibiehler M, Lazdunski C (1990) Uptake across the cell envelope and insertion into the inner membrane of ion channel-forming colicins in *E. coli* Biochimie 72(2-3):123-30

Baumstark-Khan C, Heilmann J, Rink H (2003) Induction and repair of DNA strand breaks in bovine lens epithelial cells after high LET irradiation Adv Space Res 31(6):1583-91

Becker Wayne M, (2000) The World of the cell, The Benjamin/Cummings Publishing Company, Menlo Park, CA,

Benos D, Deamer D, Kleinzeller A, Fambrough D (1999) membrane permeability: 100 years since Ernest Overton, Academic Press

Beskid G, Eskin B, Cleeland R, Siebelist J, Cappetta A, Hill AD, Geiger RH (1981) Antibacterial activity of 2',3'-dideoxyadenosine *in vivo* and *in vitro* Antimicrob Agents Chemother 19(3):424-8

Blattner FR, Plunkett G 3rd, Bloch CA, Perna NT, Burland V, Riley M, Collado-Vides J, Glasner JD, Rode CK, Mayhew GF, Gregor J, Davis NW, Kirkpatrick HA, Goeden MA, Rose DJ, Mau B, Shao Y (1997) The complete genome sequence of *Escherichia coli* K-12 Science Sep 5:277(5331):1453-74

Blomfeld IC, Vaughn V, Rest RF, Eisenstein BI (1991) Allelic exchange in *Escherichia coli* using *Bacillus subtilis sacB* gene and a temperature-sensitive pSC101 replicon, Mol Microbiol, 5, 1447-1457

Blum G, Muhldorfer I, Kuhnert P, Frey J, and Hacker J (1996) Comparative methodology to investigate the presence of *Escherichia coli* K-12 strains in environmental and human stool samples FEMS Microbiol Lett 143:77–82

Boldt R, Zrenner R (2003) Purine and pyrimidine biosynthesis in higher plants Physiol Plant 117, 297-304

Bremer E, Middendorf A, Martinussen J, Valentin-Hansen P (1990) Analysis of the *tsx* gene, which encodes a nucleoside-specific channel-forming protein (Tsx) in the outer membrane of *Escherichia coli* Gene 96:59-65

Burlingame R, Chapman PJ (1983) Catabolism of phenylpropionic acid and its 3-hydroxy derivative by *Escherichia coli* J Bacteriol 155(1):113-21

Busch W, Saier MH, Jr (2002) The transporter classification (TC) system, 2002 Crit Rev Biochem Mol Biol 37, 287-337

Carter DM, Radding CM (1971) The role of exonuclease and beta protein of phage lambda in genetic recombination II Substrate specificity and the mode of action of lambda exonuclease Biol Chem 246(8):2502-12

Carter NS, Berger BJ, Fairlamb AH (1995) Uptake of diamidine drugs by the P2 nucleoside transporter in melarsen-sensitive and -resistant *Trypanosoma brucei brucei* J Biol Chem 270(47):28153-7

Carter NS, Fairlamb AH (1993) Arsenical-resistant trypanosomes lack an unusual adenosine transporter Nature 361(6408):173-6

Carter NS, Landfear SM, Ullman B (2001) Nucleoside transporters of parasitic protozoa Trends Parasitol 17(3):142-5

Cass CE, Young JD, Baldwin SA (1998) Recent advances in the molecular biology of nucleoside transporters of mammalian cells Biochem Cell Biol 76(5):761-70

Cass CE, Young JD, Baldwin SA, Cabrita MA, Graham KA, Griffiths M, Jennings LL, Mackey JR, Ng AM, Ritzel MW (1999) Nucleoside transporters of mammalian cells Pharm Biotechnol 12, 313–352

Centelles JJ, Cascante M, Canela EI, Franco R (1992) A model for adenosine transport and metabolism Biochem J 287(Pt 2):461-72

Chan B, Weidemaier K, Yip WT, Barbara PF, Musier-Forsyth K (1999) Intra-tRNA distance measurements for nucleocapsid protein dependent tRNA unwinding during priming of HIV reverse transcription PNAS 19:96(2):459-64

Chen Y, Narendra U, Iype LE, Cox MM, Rice PA (2000) Crystal structure of a Flp recombinase-Holliday junction complex: assembly of an active oligomer by helix swapping Mol Cell 6(4):885-97

Cherepanov PP, Wachernagel W (1995) Gene disruption in *Escherichia coli*: TcR and KmR cassettes with the option of Flp- catalyzed excision of the antibiotic - resistance determinant Gene 158, 9-14

Cooper RA, Skinner MA (1980) Catabolism of 3- and 4-hydroxyphenylacetate by the 3,4-dihydroxyphenylacetate pathway in *Escherichia coli* J Bacteriol 143(1):302-6

Corbett William P and Catlin B Wesley (1968) Galactosidase Activity of Lactose-positive *Neisseria* J Bacteriol 95(1) 52–57

Craig JE, Zhang Y, Gallagher MP (1994) Cloning of the *nupC* gene of *Escherichia coli* encoding a nucleoside transport system, and identification of an adjacent insertion element, Mol Microbiol 11(6) 1159-68

Crawford CR, Cass CE, Young JD, Belt JA (1998) Stable expression of a recombinant sodium-dependent, pyrimidine-selective nucleoside transporter (CNT1) in a transport-deficient mouse leukaemia cell line Biochem Cell Biol 76(5), 843-51

Daniels (1990) The complete AvrII restriction map of the *Escherichia coli* genome and comparisons of several laboratory strains Nucleic Acids Res 18(9): 2649–2651

Damaraju VL, Damaraju S, Young JD, Baldwin SA, Mackey J, Sawyer MB, Cass CE (2003) Nucleoside anticancer drugs: the role of nucleoside transporters in resistance to cancer chemotherapy Oncogene 22(47):7524-36

Dandanell G, Szczepanowski RH, Kierdaszuk B, Shugar D, Bochtler M (2005) *Escherichia coli* purine nucleoside phosphorylase II, the product of the *xapA* gene J Mol Biol 348(1):113-25

Danielli, J F and Davson, H (1935) A contribution to the theory of permeability of thin films J Cell Comp Physiol 5:495-508

Datsenko KA, Wanner BL (2000) One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products PNAS 97: 6640-6645

Dean M (2005) The genetics of ATP-binding cassette transporters. Methods Enzymol 400:409-29

Decottignies, A Goffeau, A (1997) Complete inventory of the yeast ABC proteins Nat Genet 15, 137-145

Degano M, Gopaul DN, Scapin G, Schramm VL, Sacchettini JC (1996) Three-dimensional structure of the inosine-uridine nucleoside N-ribohydrolase from *Crithidia fasciculata* Biochemistry 14 (19):5971-81

DeLisa MP, Wu CF, Wang L, Valdes JJ, Bentley WE (2001) DNA microarray-based identification of genes controlled by autoinducer 2-stimulated quorum sensing in *Escherichia coli* J Bacteriol 183(18):5239-47

Diaz E, Ferrandez A, Garcia JL (1998) Characterization of the *hca* cluster encoding the dioxygenolytic pathway for initial catabolism of 3-phenylpropionic acid in *Escherichia coli* K-12 J Bacteriol 180(11):2915-23

Diaz E, Ferrandez A, Prieto MA, Garcia JL (2001) Biodegradation of aromatic compounds by *Escherichia coli* Microbiol Mol Biol Rev 65(4):523-69

Dighiero G (1994) Potential immunological action of purine nucleoside analogues Drugs 47 Suppl 6:57-62

Domin BA, Mahony WB, Zimmerman TP (1988) Purine nucleobase transport in human erythrocytes Reinvestigation with a novel "inhibitor-stop" assay J Biol Chem 263(19):9276-8

Doyle DA (2004) Structural themes in ion channels Eur Biophys J 33(3):175-9

D'Souza SF (2001a) Immobilization and stabilization of biomaterials for biosensor applications Appl Biochem Biotechnol 96(1-3):225-38

D'Souza SF (2001b) Microbial biosensors Biosens Bioelectron 16(6):337-53

Dwyer DE, Kesson AM (1997) Advances in antiviral therapy Curr Opin Pediat 9(1):24-30

Faruque, S M, Albert, M J, Mekalanos, J J (1998) Epidemiology, Genetics, and Ecology of Toxigenic *Vibrio cholerae* Microbiol Mol Biol Rev 62: 1301-1314

Federle MJ, Bassler BL (2003) Interspecies communication in bacteria J Clin Invest 112(9):1291-9

Fuqua WC, Winans SC, Greenberg EP (1994) Quorum sensing in bacteria: the LuxR-LuxI family of cell density-responsive transcriptional regulators J Bacteriol 176(2):269-75

Galmarini CM, Mackey JR, Dumontet C (2002) Nucleoside analogues and nucleobases in cancer treatment Lancet Oncol 3 (7):415-24

Galmarini CM, Mackey JR, Dumontet C (2002) Nucleoside analogues and nucleobases in cancer treatment Lancet Oncol 3(7):415-24

Ganapathy, Leibach FH (1985) Is intestinal peptide transport energized by a proton gradient? Am J Physiol 249(2 Pt 1):G153-60

Ganapathy V, Leibach FH (1991) Proton-coupled solute transport in the animal cell plasma membrane Curr Opin Cell Biol. 1991 Aug;3(4):695-701.

Gardner SN, Fernandes M (2004) Cytostatic anticancer drug development J Exp Ther Onco l4(1):9-18

Gavigan SA, Nguyen T, Nguyen N, and Senear DF (1999) Role of Multiple CytR Binding Sites on Cooperativity, Competition, and Induction at the *Escherichia coli udp* Promoter J Biol Chem, 274(23): 16010-16019

Gorter, E and Grendel, F (1925) On bimolecular layers of lipoids on the chromocytes of the blood J Exp Med 41:439-443

Gostick, D O, Griffin, H G, Shearman, C A, Scott, C, Green, J, Gasson, M J & Guest, J R (1999) Two operons that encode FNR-like proteins in *Lactococcus lactis* Mol Microbiol 31, 1523-1535

Gourdeau H, Clarke ML, Ouellet F (2001) Mechanisms of uptake and resistance to troxacitabine, a novel deoxycytidine nucleoside analogue, in human leukemic and solid tumor cell lines Cancer Res 61:7217-24

Griffith DA, Jarvis SM (1996) Nucleoside and nucleobase transport systems of mammalian cells Biochim Biophys Acta 1286(3):153-81

Guest, J R, Green, J, Irvine, A S & Spiro, S (1996) The FNR modulon and FNR regulated gene expression In Regulation of Gene Expression in *Escherichia coli* , pp 317-342 Edited by E C C Lin & S A Lynch Austin, TX:R G Landes

Hamilton CM, Aldea BK, Washburn P, Babitzke P, Kushner SR (1989) New method for generating deletions and gene replacements in *Escherichia coli*, J Bacteriol 171: 4617-4622

Hamilton, SR, Yao, SYM, Ingram, JC, Hadden, DA, Ritzel, MW, Gallagher, MP, Henderson, PJF, Cass, CE, Young, JD & Baldwin, SA (2001) Subcellular distribution and membrane topology of the mammalian concentrative Na+-nucleoside cotransporter rCNT1 J Biol Chem 276: 27981-27988

Hanahan D (1985) Heritable formation of pancreatic beta-cell tumours in transgenic mice expressing recombinant insulin/simian virus 40 oncogenes Nature 15;315(6015):115-22.

Hansen MR, Dandanell G (2005) Purification and characterization of RihC, a xanthosine-inosine-uridine-adenosine-preferring hydrolase from *Salmonella enterica* serovar Typhimurium Biochim Biophys Acta 1723(1-3):55-62

Hansen Westh SE, Jensen N, and Munch-Petersen A (1987) Studies on the sequence and structure of the *Escherichia coli* K-12 *nupG* gene, encoding a nucleoside-transport system Eur J Biochem 168: 385-391

Haugo AJ, and Watnick PI (2002) *Vibrio cholerae* CytR is a repressor of biofilm development Mol Microbiol 45:471-483

Hediger MA (1994) Structure, function and evolution of solute transporters in prokaryotes and eukaryotes J Exp Biol 196: 15-49

Helmann JD, Winston Wu MF, Kobel PA, Gamo FJ, Wilson M, Morshedi MM, Navre M, and Paddon C (2001) Transcriptional Response of *Bacillus subtilis* to Heat Shock J Bacteriol 183(24): 7318–7328

Helmann JD, Wu MF, Kobel PA, Gamo FJ, Wilson M, Morshedi MM, Navre M, Paddon C (2001) Global transcriptional response of *Bacillus subtilis* to heat shock J Bacteriol 183(24):7318-28

Henderson PJ, Giddens RA, Jones-Mortimer MC (1977) Transport of galactose, glucose and their molecular analogues by Escherichia coli K12. Biochem J. 162(2):309-20

Hendler, R W (1971) Biological membrane ultrastructure Physiol Rev 51(1):66-97

Hernandez JA, Fischbarg J (2005) A general channel model accounts for channel, carrier, counter-transport and co-transport kinetics. J Membr Biol. 206(3):215-26

Hoang TT, Karkhoff-Schweizer RR, Kutchma AJ, Schweizer HP (1998) A broad-host-range Flp-FRT recombination system for site-specific excision of

chromosomally-located DNA sequences: application for isolation of unmarked *Pseudomonas aeruginosa* mutants Gene 28:212(1):77-86

httl://genomebiology.com/2001/2/9/research/0035

http://genolist.pasteur.fr/Colibri

http://www.biology.ucsd.edu/~msaiertransport/titlepage.html

http://www.gene-regulation.com/cgi-bin/pub/programs/match/bin/matchcgi

http://www.ncbi.nlm.nih.gov/entrez

Huang LC, Wood EA, Cox MM (1997) Convenient and reversible site-specific targeting of exogenous DNA into a bacterial chromosome by use of the FLP recombinase: the FLIRT system J Bacteriol, 179(19):6076-83

Huang, Q Q, S Y Yao, M W Ritzel, A R Paterson, C E Cass, and J D Young (1994) Cloning and functional expression of a complementary DNA encoding a mammalian nucleoside transport protein J Biol Chem 269, 17757-17760

Hyde RJ, Cass CE, Young JD, Baldwin SA (2001) The ENT family of eukaryote nucleoside and nucleobase transporters: recent advances in the investigation of

structure/function relationships and the identification of novel isoforms Mol Membr Biol 18(1):53-63.

Jap BK, Walian PJ (1996) Structure and functional mechanism of porins Physiol Rev 76(4):1073-88

Jensen KF, Nygaard (1975) P Purine nucleoside phosphorylase from *Escherichia coli* and *Salmonella typhimurium* Purification and some properties Eur J Biochem 51(1):253-65

Johansen LE, Nygaard P, Lassen C, Agersø Y, and Saxild HH (2003) Definition of a Second *Bacillus subtilis* pur Regulon Comprising the pur and *xpt-pbuX* Operons plus *pbuG, nupG* (*yxjA*), and *pbuE* (*ydhL*) Journal of Bacteriology 185(17): 5200-5209

Jørgensen C, Dandanell G (1999) Isolation and Characterization of Mutations in the *Escherichia coli* Regulatory Protein XapR Bacteriol 181(14):4397–4403

Josephsen J, Hammer-Jespersen K, Hansen TD. (1983) Mapping of the gene for cytidine deaminase (cdd) in *Escherichia coli* K-12: J Bacteriol 154(1):72-5

Jun Z, Miller MB, Vance RE, Dziejman M, Bassler BL, and Mekalanos JJ (2002) Quorum-sensing regulators control virulence gene expression in *Vibrio cholerae* PNAS 99:3129-3134

Kaback HR (1969) Regulation of Sugar Transport in Isolated Bacterial Membrane Preparations from *Escherichia coli* PNAS 63:724-731

Kaback HR, Wu J (1999) What To Do While Awaiting Crystals of a Membrane Transport Protein and Thereafter Accounts of Chemical Research 32:805-813

Kamboj RK, Jackson JF (1985) Pyrimidine nucleosides uptake by petunia pollen Plant Physiol 79:801-805

Karlin S, Mrazek (2000) J Predicted highly expressed genes of diverse prokaryotic genomes J Bacteriol 182(18):5238-50

Karlin S, Mrazek J, Campbell A, Kaiser D (2001) Characterizations of highly expressed genes of four fast-growing bacteria J Bacteriol 183(17):5025-40

Klose KE, Novik V, Mekalanos JJ (1998) Identification of multiple sigma54-dependent transcriptional activators in *Vibrio cholerae* J Bacteriol 180(19):5256-9

Komatsu,Y, and K Tanaka (1973) Deoxycytidine uptake by isolated membrane vesicles from *Escherichia coli* K 12 Biochim Biophys Acta 311:496-506

Kong W, Engel K and Wang J (2004) Mammalian Nucleoside Transporters Current drug metabolism 5:63-84

Kovacikova G, Skorupski K (2002) Regulation of virulence gene expression in *Vibrio cholerae* by quorum sensing: HapR functions at the *aphA* promoter Mol Microbiol 46(4):1135-47

Krenitsky TA, Freeman GA, Shaver SR, Beacham LM 3rd, Hurlbert S, Cohn NK, Elwell LP, Selway JW (1983) 3'-Amino-2',3'-dideoxyribonucleosides of some pyrimidines: synthesis and biological activities J Med Chem 26(6):891-5

Kuo MM, Haynes WJ, Loukin SH, Kung C, Saimi Y (2005) Prokaryotic K(+) channels: from crystal structures to diversity FEMS Microbiol Rev 29(5):961-85 Epub 2005 Jul 1

Lage H (2003) Molecular analysis of therapy resistance in gastric cancer Dig Dis 21(4) 326-38

Lange R, Hengge-Aronis R (1991) Identification of a central regulator of stationary-phase gene expression in *Escherichia coli* Mol Microbiol 5(1):49-59

Langmuir, I (1917) The constitution and fundamental properties of solids and liquids II Liquids J Am Chem Soc 39:1848-1906

Lee G, Kim MG, Yim JB, Hong SH. (2001) Alternative transcriptional initiation and splicing of mouse Lamc2 message Mol Cells 12(3):380-90

Leonard P, Hearty S, Brennan J, Dunne L, Quinn J, Chakraborty T, O'Kennedy R (2003) Advances in biosensors for detection of pathogens in food and water 32:3-13

Leveau JHJ, Lindow SE (2002) Bioreporters in microbial ecology Current opinion in microbiology 5:259-265

Li X, Chan WK (1999) Transport, metabolism and elimination mechanisms of anti-HIV agents Adv Drug Deliv Rev 18: 81-103

Link AJ, Phillips D, Church GM (1997) Methods for generating precise deletions and insertion in the genome of wild type *Escherichia coli*: Application to open reading frame characterization, J Bacteriol 179(20), 6228-6237

Linton KJ, Higgins CF (1998) The *Escherichia coli* ATP-binding cassette (ABC) proteins Mol Microbiol 28 (1): 5-13

Liu X, Xie W, Huang RH (2005) Structure-based design, synthesis, and in vitro assay of novel nucleoside analog inhibitors against HIV-1 reverse transcriptase Bioorg Med Chem Lett 15(16):3775-7

Loewen SK, Ng AML, Mohabir NN, Baldwin SA , Cass CE, Young JD (2003) Functional characterization of a H+/nucleoside co-transporter (CaCNT) from *Candida albicans*, a fungal member of the concentrative nucleoside transporter (CNT) family of membrane proteins Yeast 20 (8): 661-675

Loewen SK, Yao SYM, Slugoski MD, Mohabir NN, Turner RJ, Mackey JR, Weiner JH, Gallagher MP, Henderson PJF, Baldwin SA , Cass CE, Young JD (2004) Transport of physiological nucleosides and anti-viral and anti-neoplastic nucleoside drugs by recombinant *Escherichia coli* nucleoside-H+ cotransporter (NupC) produced in *Xenopus laevis* oocytes Molecular Membrane Biology 21 (1): 1-10

Loewen, SK, Ng, AL, Yao, SYM, Cass, CE, Baldwin, SA & Young, JD (1999) Identification of amino acid residues responsible for the pyrimidine and purine nucleoside specificities of human concentrative Na+ nucleoside cotransporters hCNT1 and hCNT2 J Biol Chem 274: 24475-24484

Lostao MP, Mata JF, Larrayoz IM, Inzillo SM, Casado FJ, Pastor-Anglada M (2000) Electrogenic uptake of nucleosides and nucleoside-derived drugs by the human nucleoside transporter 1 (hCNT1) expressed in *Xenopus laevis* oocytes FEBS Lett 481(2):137-40

Marger MD, Saier MH Jr (1993) A major superfamily of transmembrane facilitators that catalyse uniport, symport and antiport. Trends Biochem Sci 18(1):13-20

Masters M (1977) The frequency of P1 transduction of the genes of *Escherichia coli* as a function of chromosomal position: preferential transduction of the origin of replication Mol Gen Genet 155(2):197-202

Masters M (1987) Generalised transduction, *E coli Salmonella*: cellular and molecular biology, 2nd edition pp2421-2441

Meighen EA (1993) Bacterial bioluminescence: organization, L application of the lux genes FASEB 7:1016-1022

Meighen EA (1993) Bacterial bioluminescence: organization, regulation, and application of the *lux* genes FASEB 7:1016-22

Merlin C, Gardiner G, Durand S, Masters M (2002a) The *Escherichia coli metD* locus encodes an ABC transporter which includes *Abc* (*MetN*), *YaeE* (*MetI*), and *YaeC* (*MetQ*) J Bacteriol, 184(19):5513-7

Merlin C, McAteer S, Masters M (2002b) Tools for characterization of *Escherichia coli* genes of unknown function J Bacteriol 184(16):4573-81

Mubagwa K, Mullane K, Flameng W (1996) Role of adenosine in the heart and circulation Cardiovasc Res 32(5):797-813

Munch-Petersen A, Mygind B (1976) Nucleoside transport systems in *Escherichia coli* K12: specificity and regulation J Cell Physiol 89(4):551-9

Munch-Petersen, A, and B Mygind (1983) Transport of nucleic acid precursors, p 259-305 In A Munch-Petersen (ed), Metabolism of nucleotides, nucleosides and nucleobases in microorganisms Academic Press, London, England

Munch-Petersen, A, B Mygind, A Nicolaisen, and N J Pihl (1979) Nucleoside transport in cells and membrane vesicles from *Escherichia coli* K12 J Biol Chem 254:3730-3737

Murphy KC (1998) Use of bacteriophage l recombination functions to promote gene replacement in *Escherichia coli* J Bacteriol, 180(8), 2063-2071

Murphy KC, Campellone KG, Poteete AR (2000) PCR-mediated gene replacement in *Escherichia coli* Gene 246: 321-330

Mygind B, Munch-Petersen A (1975) Transport of pyrimidine nucleosides in cells of *Escherichia coli* K 12 Eur J Biochem 59(2):365-72

Namiesnik J, Zabiegala B, Kot-Wasik A, Partyka M, Wasik A (2005) Passive sampling and/or extraction techniques in environmental analysis: a review Anal Bioanalyt Chem 381(2):279-301

Naylor LH (1999) Reporter gene technology: the future looks bright Biochemical Pharmacology 58:749-757

Nealson KH, Platt T, Hastings JW (1970) Cellular control of the synthesis and activity of the bacterial luminescent system J Bacteriol 104(1):313-22

Neidhardt, F C (1996) The enteric bacterial cell and the age of bacteria, p 1–3 *In* F C Neidhardt, R Curtiss III, J L Ingraham, E C C Lin, K B Low, B Magasanik, W S Reznikoff, M Riley, M Schaechter, and H E Umbarger (ed), *Escherichia coli* and *Salmonella*: cellular and molecular biology, 2nd ed ASM Press, Washington, DC

Neuhard J, Nygaard P (1987) Purines and Pyrimidines In FC Neidhardt, JL Ingraham, BK Low, B Magasanik, M Schaechter, HE Umbarger, eds, *Escherichia coli* and *Salmonella typhimurium*: Cellular and Molecular Biology American Society for Microbiology, Washington, DC, pp 445-473

Newman BJ, Masters M (1980) The variation in frequency with which markers are transduced by phage P1 is primarily a result of discrimination during recombination Mol Gen Genet 180(3):585-9

Nieweg A, Bremer E (1997) The nucleoside-specific Tsx channel from the outer membrane of *Salmonella typhimurium*, *Klebsiella pneumoniae* and *Enterobacter*

aerogenes: functional characterization and DNA sequence analysis of the *tsx* genes Microbiology 143 (Pt 2):603-15

Nikaido H, Saier MH Jr (1992) Transport proteins in bacteria: common themes in their design Science 258(5084):936-42

Nikaido, Hiroshi (1994) Prevention of Drug Access to Bacterial Targets: Permeability Barriers and Active Efflux, Science 264: 382-387

Nørholm MHH, Dandanell G (2001) Specificity and Topology of the *Escherichia coli* Xanthosine Permease, a Representative of the NHS Subfamily of the Major Facilitator Superfamily J Bacteriol, 183(16):4900-4904

Pao SS, Paulsen IT, Saier MH Jr (1998) Major facilitator superfamily Microbiol Mol Biol Rev 62(1):1-34

Park GB, Mitra AK (1992) Mechanism and site dependency of intestinal mucosal transport and metabolism of thymidine analogues Pharm Res 9(3):326-31

Parker MD, Hyde RJ, Yao SY, McRobert L, Cass CE, Young JD, McConkey GA, Baldwin SA (2000) Identification of a nucleoside/nucleobase transporter from Plasmodium falciparum, a novel target for anti-malarial chemotherapy Biochem J 349(Pt 1):67-75

Parker WB, Cheng YC (1990) Metabolism and mechanism of action of 5-fluorouracil Pharmacol Ther 48(3):381-95

Pastor-Anglada M, Felipe A, Casado FJ (1998) Transport and mode of action of nucleoside derivatives used in chemical and antiviral therapies Trends Pharmacol Sci 19(10):424-30

Patching SG, Baldwin SA , Baldwin AD, Young JD, Gallagher MP, Henderson PJF, Herbert RB (2005) The nucleoside transport proteins, NupC and NupG, from *Escherichia coli*: specific structural motifs necessary for the binding of ligands Organic & Biomolecular Chemistry 3:462-470

Patel DH, Crawford CR, Naeve CW, Belt JA (2000) Cloning, genomic organization and chromosomal localization of the gene encoding the murine sodium-dependent, purine-selective, concentrative nucleoside transporter (CNT2) Gene 242(1-2):51-8

Peekhaus N and Conway T (1998) What's for dinner? Entner-Doudoroff metabolism in *Escherichia coli* J Bacteriol 180:3495–3502 281

Peppercorn, MA, and Goldman P (1971) Caffeic acid metabolism by bacteria of the human gastrointestinal tract J Bacteriol 108:996–1000

Pettitt AR (2003) Mechanism of action of purine analogues in chronic lymphocytic leukaemia Br J Haematol 121(5):692-702

Plunkett W, Huang P, Xu YZ, Heinemann V, Grunewald R, Gandhi V (1995) Gemcitabine: metabolism, mechanisms of action, and self-potentiation Semin Oncol 22(4 Suppl 11):3-10

Pósfai G, Kolisnychenko V, Bereczki Z, Blattner FR (1999) Markerless gene replacement in *Escherichia coli* stimulated b a double-strand break in the chromosome 27 (22):4409-4415

Poteete AR, Fenton AC, Murphy KC (1999) Roles of RuvC and RecG in phage lambda red-mediated recombination J Bacteriol 181(17):5402-8

Purohit PV, Rocke DM 2003 Discriminant models for high-throughput proteomics mass spectrometer data Proteomics 3(9):1699-703

Reading NC, Sperandio V (2006) Quorum sensing: the many languages of bacteria FEMS Microbiol Lett 254(1):1-11

Ren D, Bedzyk LA, Ye RW, Thomas SM, Wood TK (2004) Stationary-phase quorum-sensing signals affect autoinducer-2 and gene expression in *Escherichia coli* Appl Environ Microbiol 70(4):2038-43

Ren D, Zuo R, Gonzalez Barrios AF, Bedzyk LA, Eldridge GR, Pasmore ME, Wood TK (2005) Differential gene expression for investigation of *Escherichia coli* biofilm inhibition by plant extract ursolic acid Appl Environ Microbiol 71(7):4022-34

Reyrat JM, Pelicic V, Gicquel B, Rappuoli R (1998) Counterselectable markers: untapped tools for bacterial genetics and pathogenesis Infect Immun 66(9):4011-7

Ritzel MW, Yao SY, Ng AM, Mackey JR, Cass CE, Young JD (1998) Molecular cloning, functional expression and chromosomal localization of a cDNA encoding a human Na+/nucleoside cotransporter (hCNT2) selective for purine nucleosides and uridine. Mol Membr Biol. 15(4):203-11

Ritzel MWL, Ng AML, Yao SYM, Graham K, Loewen SK, Smith KM, Hyde RJ, Karpinski E, Cass CE, Baldwin SA, Young JD (2001a) Recent molecular advances in studies of the concentrative Na+-dependent nucleoside transporter (CNT) family: identification and characterization of novel human and mouse proteins (hCNT3 and mCNT3) broadly selective for purine and pyrimidine nucleosides (system cib) Molecular Membrane Biology 18 (1): 65-72

Ritzel, MWL, Ng, AML, Yao, SYM, Graham, K, Loewen, SK, Smith, KM, Ritzel, RG, Mowles, DA, Carpenter, P-A, Chen, X-Z, Karpinski, E, Hyde, RJ, Baldwin, SA, Cass, CE & Young, JD (2001b) Molecular identification and characterization of novel human and mouse concentrative Na+-nucleoside cotransporter proteins

(hCNT3 and mCNT3) broadly selective for purine and pyrimidine nucleosides (system cib) J Biol Chem 276, 2914-2927

Robertson, JD (1957) New observations on the ultrastructure of the membranes of frog peripheral nerve fibers J Biophys Biochem Cytol 3:1043-1047

Rosenblum LL, Patton G, Grigg AR, Frater AJ, Cain D, Erlwein O, Hill CL, Clarke JR, McClure MO (2001) Differential susceptibility of retroviruses to nucleoside analogues Antivir Chem Chemother 12(2):91-7

Rubinstein LV, Shoemaker RH, Paull KD, Simon RM, Tosini S, Skehan P, Scudiero DA, Monks A, Boyd MR (1990) Comparison of in vitro anticancer-drug-screening data generated with a tetrazolium assay versus a protein assay against a diverse panel of human tumor cell lines J Natl Cancer Inst 82(13):1113-8

Sadee W, Graul RC, Lee AY (1999) Classification of membrane transporters Pharm Biotechnol 12:29-58.

Sahin-Tóth M, Karlin A, and Kaback H R (2000) Unravelling the mechanism of the lactose permease of *Escherichia coli* PNAS 97: 10729-10732

Saier MH Jr (1990) Evolution of permease diversity and energy-coupling mechanisms: an introduction Res Microbiol 141(3):281-6

Saier MH Jr (2000) Families of proteins forming transmembrane channels J Membr Biol 175(3):165-80

Saier MH, Jr (1994) Computer-aided analyses of transport protein sequences: gleaning evidence concerning function, structure, biogenesis, and evolution Microbiol Revs 58:71-93

Saier, M H, Jr (2003) Tracing pathways of transport protein evolution Mol Microbiol 48: 1145–1156

Salton MR (1987) Bacterial membrane proteins Microbiol Sci 4(4):100-5

Sambrook, J, Fritsch EF, and Maniatis T (1989) Molecular cloning: a laboratory manual Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY

Sandler A, Ettinger DS (1999) Gemcitabine: Single-Agent and Combination Therapy in Non-Small Cell Lung Cancer The Oncologist 4(3):241-251

Savageau MA (1983) Regulation of differentiated cell-specific functions PNAS 80:1411–1415

Savageau, MA (1983) Regulation of differentiated cell-specific functions PNAS 80:1411–1415

Saxild, H H, Andersen, L N & Hammer, K (1996) dra–nupC–pdp operon of *Bacillus subtilis*: nucleotide sequence, induction by deoxyribonucleosides, and transcriptional regulation by the *deoR*-encoded DeoR repressor protein J Bacteriol 178:424-434

Schauder S, Shokat K, Surette MG, Bassler BL (2001) The LuxS family of bacterial autoinducers: biosynthesis of a novel quorum-sensing signal molecule Mol Microbiol 41(2):463-76

Schell MA (1993) Molecular biology of the LysR family of transcriptional regulators Annu Rev Microbiol 47:597–626

Schuch R, Garibian A, Saxild HH, Piggot PJ, Nygaard P (1999) Nucleosides as a carbon source in *Bacillus subtilis*: characterization of the drm-pupG operon Microbiology 145: 2957-2966

Seeger C, Poulsen C, Dandanell G (1995) Identification and characterization of genes (*xapA, xapB,* and *xapR*) involved in xanthosine catabolism in *Escherichia coli* J Bacteriol 177(19):5506–5516

Serres MH, Gopal S, Nahum LA, Pliang T, Gaasterland T, Riley M (2001) A functional update of *E coli* K-12 genome Genome Biol 2 (online)

Shaw JJ & Kado CI (1986) Development of a *Vibrio* bioluyminescence gene-set to monitor phytopathogenic bacteria during the ongoing disease process in a non-disruptive manner Biotechnology 4:560-563

Shryock JC, Belardinelli L (1997) Adenosine and adenosine receptors in the cardiovascular system: biochemistry, physiology, and pharmacology Am J Cardiol, 79:2-10

Singer SJ and Nicolson GL (1972) The fluid mosaic model of the structure of cell membranes Science 175:720-731

Skehan P, Storeng R, Scudiero D, Monks A, McMahon J, Vistica D, Warren JT, Bokesch H, Kenney S, Boyd MR (1990) New colorimetric cytotoxicity assay for anticancer-drug screening J Natl Cancer Inst 82(13):1107-12

Smith KM, Chen XZ, Cass CE, Baldwin SA , Karpinski E, Young JD (2004) Voltage clamp analysis of the presteady-state currents mediated by the human concentrative Na+-nucleoside cotransporter hCNT3 Biophysical Journal 86:262A-262A Part 2 Suppl S

Smith PG, Thomas HD, Barlow HC, Griffin RJ, Golding BT, Calvert AH, Newell DR, Curtin NJ (2001) *In vitro* and *in vivo* properties of novel nucleoside transport inhibitors with improved pharmacological properties that potentiate antifolate activity Clin Cancer Res 7 (7):2105-13

Sommadossi JP (1998) Cellular nucleoside pharmacokinetics and pharmacology: a potentially important determinant of antiretroviral efficacy AIDS 12 Suppl 3:S1-8

Sommadossi JP, Carlisle R, Zhou Z (1989) Cellular pharmacology of 3'-azido-3'-deoxythymidine with evidence of incorporation into DNA of human bone marrow cells Mol Pharmacol 36(1): 9-14

Sperandio V, Mellies JL, Nguyen W, Shin S, Kaper JB (1999) Quorum sensing controls expression of the type III secretion gene transcription and protein secretion in enterohemorrhagic and enteropathogenic *Escherichia coli* PNSA 96(26):15196-201

Sperandio V, Torres AG, Giron JA, Kaper JB (2001) Quorum sensing is a global regulatory mechanism in enterohemorrhagic *Escherichia coli* O157:H7 J Bacteriol 183(17):5187-97

Spiro S (1994) The FNR family of transcriptional regulators Antonie Leeuwenhoek 66: 23-36

Stewart GS, Williams P (1992) lux genes and the applications of bacterial bioluminescence J Gen Microbiol 138(7):1289-300

Studier FW, Rosenberg AH, Dunn JJ, Dubendorff JW (1990) Use of T7 RNA polymerase to direct expression of cloned genes Methods Enzymol 185:60-89

Surette MG, Miller MB, Bassler BL (1999) Quorum sensing in *Escherichia coli*, *Salmonella typhimurium*, and *Vibrio harveyi*: a new family of genes responsible for autoinducer production PNAS 96(4):1639-44

Tam R, Saier MH Jr (1993) Structural, functional, and evolutionary relationships among extracellular solute-binding receptors of bacteria Microbiol Rev 57(2):320-46

Theodoulou FL (2000) Plant ABC transporters Biochim Biophys Acta 1465:79-103

Toan SV, To KK, Leung GP, de Souza MO, Ward JL, Tse CM (2003) Genomic organization and functional characterization of the human concentrative nucleoside transporter-3 isoform (hCNT3) expressed in mammalian cells Pflugers Arch 447(2):195-204

Tribble G, Ahn YT, Lee J, Dandekar T, Jayaram M (2000) DNA recognition, strand selectivity, and cleavage mode during integrase family site-specific recombination J Biol Chem 275(29):22255-67

Vazquez-Ibar JL, Guan L, Svrakic M, Kaback HR (2003) Exploiting luminescence spectroscopy to elucidate the interaction between sugar and a tryptophan residue in the lactose permease of *Escherichia coli* PNAS 100(22):12706-11

Vesterlund S, Paltta J, Laukova A, Karp M, Ouwehand AC (2004) Rapid screening method for the detection of antimicrobial substances J Microbiol Methods 57(1):23-31

Vickers MF, Young JD, Baldwin SA, Ellison MJ, Cass CE (2001) Functional production of mammalian concentrative nucleoside transporters in *Saccharomyces cerevisiae* Molecular Membrane Biology 18 (1): 73-79

Vollack, KU, Hartig E, Korner H & Zumft WG (1999) Multiple transcription factors of the FNR family in denitrifying *Pseudomonas stutzeri*: characterisation of four fnr-like genes, regulatory responses and cognate metabolic processes Mol Microbiol 31: 1681-1694

Wall L, Rodriguez A, Meighen E (1986) Intersubunit transfer of fatty acyl groups during fatty acid reduction J Biol Chem 261(34):15981-8

Wang J, Su SF, Dresser MJ, Schaner ME, Washington CB, Giacomini KM (1997) Na(+)-dependent purine nucleoside transporter from human kidney: cloning and functional characterization Am J Physiol 273(6 Pt 2):F1058-65

White, P (1938) The rugose variant of *Vibrios* J Pathol 46:1-6

Wild J, Sektas M, Hradečná Z, Szybalski W (1998) Targeting and retrofitting pre-existing libraries of transposon insertions with FRT and *oriV* elements for *in-vivo* generation of large quantities of any genomic fragment Gene 223: 55-66

Williams P (1988) Role of the cell envelope in bacterial adaptation to growth *in vivo* in infections Biochimie 70(8):987-1011

Wohlhueter RM, McIvor RS, Plagemann PG (1980) Facilitated transport of uracil and 5-fluorouracil, and permeation of orotic acid into cultured mammalian cells J Cell Physiol 104 (3): 309-19

Wormit A, Traub M, Flörchinger M, Neuhaus HE, Möhlmann T (2004) Characterization of three novel members of the Arabidopsis thaliana equilibrative-nucleoside transporter (ENT) family J Biochem 383:19-26

Wu TT (1966) A model for three-point analysis of random general transduction Genetics 54: 405-410

Xiao G, Wang J, Tangen T, Giacomini KM (2001) A novel proton-dependent nucleoside transporter, CeCNT3, from *Caenorhabditis elegans* Mol Pharmacol 59(2):339-48

Xie H, Patching SG, Gallagher MP, Litherland GJ, Brough AR, Venter H, Yao SYM, Ng AML, Young JD, Herbert RB, Henderson PJF, Baldwin SA (2004) Purification and properties of the *Escherichia coli* nucleoside transporter NupG, a paradigm for a major facilitator transporter sub-family Molecular Membrane Biology 21:323-336

Yang, J, and Thomason DB (1993) An easily synthesized, photolyzable luciferase substrate for *in vivo* luciferase activity measurement BioTechniques 15:848-850

Yao SY, Cass CE, Young JD (1996) Transport of the antiviral nucleoside analogs 3'-azido-3'-deoxythymidine and 2',3'-dideoxycytidine by a recombinant nucleoside transporter (rCNT) expressed in *Xenopus laevis* oocytes Mol Pharmacol 50(2):388-93

Yao, SY, Ng, AM, Loewen, SK, Cass CE, Baldwin SA and Young, JD (2002) An ancient prevertebrate Na+-nucleoside cotransporter (hfCNT) from the Pacific hagfish (*Eptatretus stouti*) Am J Physiol Cell Physiol 283: C155-C168

Yatvin MB, Li W, Meredith MJ, Shenoy MA (1999) Improved uptake and retention of lipophilic prodrug to improve treatment of HIV Adv Drug Deliv Rev18:39(1-3):165-182

Yildiz FH, Liu XS, Heydorn A, Schoolnik GK (2004) Molecular analysis of rugosity in a *Vibrio cholerae* O1 El Tor phase variant Mol Microbiol 53(2):497-515

Yu D, Ellis HM, Lee EC, Jenkins NA, Copeland NG, Court DL (2000) An efficient recombination system for chromosome engineering in *Escherichia coli* PNAS 97: 5978-5983

Zalkin H, and Nygaard P (1996) Biosynthesis of purine nucleotides In *Escherichia coli* and *Salmonella*: Cellular and Molecular Biology, Vol I Neidhardt, NC (ed) Washington, DC: American Society for Microbiology Press, pp561–579

Zhu J, Miller MB, Vance RE, Dziejman M, Bassler BL, Mekalanos JJ. (2002) Quorum-sensing regulators control virulence gene expression in *Vibrio cholerae* PNSA 99 (5) 3129-34