SYNOPSIS

Bacterial exporter proteins mediate the export of a wide variety of compounds and their activities are largely thought to play an adaptive role in mitigating the detrimental effects on bacterial growth caused by the presence of biotic stresses such as those imposed by antibiotics, heavy metals and other toxic compounds, in their natural environments. While it is easy to come to terms with the existence of proteins that mediate export of compounds described above, the presence of specific export systems for cellular metabolites such as sugars and amino acids is enigmatic. For example, *Escherichia coli* encodes multiple proteins for the export of sugars such as glucose and lactose (Liu et al 1999) and for arabinose (Koita and Rao 2012). In addition occurrence of proteins mediating export of multiple amino acids has been reported in *E. coli* (Hori et al 2011, Doroshenko et al 2007, Franke et al 2003, Dassler et al 2000). While the physiological basis for the existence of amino acid exporters is not clear, their occurrence is relatively widespread in bacteria (Burkovski and Krämer 2002, Marin and Krämer 2007). It is thought that an amino acid exporter may serve to contribute to the fitness of an organism during conditions of metabolic imbalance resulting from excessive levels of its amino acid substrate in the cytoplasm (Burkovski and Krämer 2002, Marin and Krämer 2007). This is of particular relevance where the cytoplasm possesses a limited capacity for the catabolism of the exported amino acid (Burkovski and Krämer 2002, Erdmann et al 1993). An alternative view is that export of the amino acid is purely coincidental, with the actual substrate being a naturally occurring structurally related antimetabolite present in the environment. For example, one mechanism mediating resistance of *E. coli* to the plant derived antimetabolite canavanine (CAN), an L-arginine (Arg) analogue, present in bulk soil, involves its export by ArgO, the ortholog of the LysE basic amino acid exporter of *Corynebacterium glutamicum*, following its uptake (Nandineni and Gowrishankar 2004). For the case of sugar exporters it is believed that their function may
serve to mitigate cellular toxicity arising under conditions of excessive accumulation of phosphorylated sugar intermediates (Koita and Rao 2012). Harnessing bacterial amino acid export coupled with appropriate metabolic engineering has found widespread application in the commercial production of amino acids (reviewed in Burkovski and Krämer 2002, Marin and Krämer 2007).

From the Laboratory of Bacterial Genetics, based mainly on physiological and genetic studies, ArgO was shown to mediate export of Arg and its toxic analogue CAN (Nandineni and Gowrishankar 2004). These studies also showed that its transcriptional activation and repression, respectively by Arg and L-lysine (Lys) is mediated by the regulator ArgP. Accordingly argO and argP mutants are CAN supersensitive (CANSS). As described earlier ArgO is an ortholog of LysE and its transcriptional regulator ArgP is also an ortholog of LysG that mediates transcriptional regulation of lysE (Bellmann et al 2001). Both ArgP and LysG are bracketed as members of the family of LysR-type transcriptional regulators (LTTRs) (Maddocks and Oyston, 2008) and ArgO and LysE are member proteins of the LysE family (Vrljic et al 1999). Given the significant identities and similarities between ArgO and LysE and ArgP and LysG, ArgP-argO of E. coli and LysG-lysE from C. glutamicum are thought to constitute orthologous protein-gene pairs (Marbaniang and Gowrishankar 2012). Despite the overall similarity between the aforementioned protein-gene pairs there are some differences. ArgO ordinarily mediates export of Arg (Nandineni and Gowrishankar 2004) whereas LysE mediates export of both Arg and Lys (Vrljic et al 1996, Bellmann et al 2001). The most intriguing difference between the protein-gene pairs, is seen at the level of transcriptional regulation mediated by the two regulators ArgP and LysG, on their transcriptional targets argO and lysE respectively. Transcription of lysE is activated by both Arg and Lys serving as coeffectors of LysG (Bellmann et al 2001), whereas argO expression is activated by Arg and repressed by Lys (Nandineni and Gowrishankar 2004) with repression by Lys occurring via
the formation in the presence of Lys (but not Arg) of an ArgP-Esig70 (sigma 70 bearing RNA polymerase) complex, reversibly trapped at the step of promoter clearance, following open complex formation (Laishram and Gowrishankar 2007). ArgP binds to the argO promoter/operator region even in the absence of Arg and Lys however Arg and Lys mediate recruitment of Esig70 to the argO promoter (Laishram and Gowrishankar 2007).

The focus of this thesis is directed towards the study of processes of basic amino acid export and its regulation in *E. coli*. Accordingly, this thesis is divided into 5 chapters. In the introduction chapter the processes of microbial amino acid uptake, with respect to basic amino acid uptake and of amino acid export are discussed. Aspects related to sugar exporters, the MATE family of exporters and membrane protein topogenesis are also discussed. Chapters 2, 3 and 4 constitute the chapters of results and the relevant methodology of experimental procedures employed are described in each chapter. Chapter 5 summarizes the findings of this thesis and discusses future perspectives derived from this study.

At the time this work was initiated it was known that in *E. coli* ArgO mediated export of Arg but whether it could mediate Lys export was not known. Furthermore, it was also not clear whether *E. coli* possesses a separate exporter dedicated for Lys export. Chapter 2 describes the identification of the gene *ybjE* predicted to encode a cytoplasmic membrane protein and multiple lines of evidence obtained indicate that YbjE mediates Lys export (Pathania and Sardesai 2015). *ybjE* was isolated as a multicopy suppressor of the CAN<sup>SS</sup> phenotype of a strain lacking ArgO and a rationale for identification of *ybjE* is provided in the discussion section of chapter 2. Notwithstanding the manner of identification of *ybjE*, the Δ*ybjE* mutant did not display any discernible CAN<sup>SS</sup> phenotype, instead it was rendered hypersensitive to the Lys antimetabolite thialysine. In dipeptide feeding experiments the Δ*ybjE* mutant was found to be growth inhibited by the dipeptide L-lysylalanine (Lys-Ala) that correlated with elevated cellular Lys content. YbjE overproduction resulted in Lys excretion
and syntrophic cross-feeding of a Lys auxotroph, the latter providing a visual indication of YbjE mediated Lys export. Under conditions where expression of \textit{argO} was dissociated from the repressive effect of Lys on its expression, which occurs via the ArgP transcriptional factor (Nandineni and Gowrishankar 2004), the Lys export potential of ArgO was rendered apparent. Thus expression of \textit{argO} from the P\textsubscript{lac} promoter rescued the thialysine sensitive phenotype of the \textit{ΔybjE} mutant, indicating that ordinarily the Lys export potential of ArgO in \textit{E. coli} is rendered cryptic. In addition Lys export phenotype of \textit{argO} was also apparent in strains expressing Lys insensitive versions of ArgP (ArgP\textsuperscript{Ds}) that were previously shown to cause Lys insensitive constitutive high level expression of \textit{argO} (Nandineni and Gowrishankar 2004). Additional studies were performed to characterize the 5′ promoter/regulatory region of \textit{ybjE} which included delineation of the core promoter elements of \textit{ybjE} and the start of the \textit{ybjE} mRNA. Furthermore, extracellular Arg modestly repressed \textit{ybjE} transcription in an ArgR dependent manner and ArgR, the repressor of genes of the Arg biosynthesis regulon (Maas 1994, Glansdorff 1996), displayed Arg sensitive binding to the \textit{ybjE} promoter region \textit{in vitro}, with the ArgR binding site likely to exist in an overlap with the \textit{ybjE} promoter. It is speculated that ArgR at the \textit{ybjE} promoter is likely to exert its repressive effects in a manner analogous to its role as a repressor of genes of Arg biosynthesis regulon that is by occluding promoter binding of the RNA polymerase holoenzyme (Maas 1994, Glansdorff 1996). However subtle differences between the manner in which ArgR exerted its effects on \textit{ybjE} expression \textit{vis-à-vis} the repression it exerts on genes of the Arg biosynthesis regulon were observed that are described in the discussion section of chapter 2. It is suggested that the repression of the Arg exporter ArgO by Lys and of the Lys exporter YbjE by Arg in \textit{E. coli} is indicative of a mechanism for maintenance of an Arg/Lys balance. Thus in \textit{E. coli} two proteins ArgO and YbjE perform the task of separately exporting Arg and Lys respectively, which is distinct from that seen for \textit{C. glutamicum} where the ortholog of ArgO,
LysE mediates export of both Arg and Lys. Repression of argO transcription by Lys is thought to effect this separation. ybjE has thus been re-designated as lysO (for Lys outward permease). Lastly, results described in chapter 2 are consistent with reports by Ueda and coworkers who via a different approach have described the identification of YbjE as a Lys exporter, in multiple patents (Ueda et al 2009, Ueda et al 2009, Ueda et al 2013).

ArgO and LysE are members of the LysE superfamily that is widely distributed, with orthologs found in eubacterial and archeal genomes with some genomes such as that of M. tuberculosis H37Rv, bearing the potential to encode multiple paralogs of LysE (Vrljic 1999, Tsu and Saier 2015). These proteins contain on an average 200 to 220 amino acids, are largely thought to mediate amino acid export and predictions of their topology are supportive of a 6 transmembrane helical arrangement. While LysE and ArgO remain the best functionally characterized members of the LysE family there is an absence of structural information pertaining to them with no structures reported for any member. As an initial step towards determination of the molecular basis of Arg export by ArgO a detailed analysis of its topology in E. coli has been undertaken and the results of this study are described in chapter 3. Using a combination of cysteine accessibility in situ and alkaline phosphatase fusion reporters, a topological model for ArgO has been proposed. Significant concordance in the topological assignments of amino acyl residues of ArgO was obtained between the two techniques used to evaluate the topology of ArgO. These studies indicated that ArgO assumes an N_{In}-C_{Out} configuration potentially forming a five transmembrane helix bundle flanked by an N-terminal cytoplasmic domain (NTD) comprising roughly its first 38 amino acyl residues and a short C-terminal periplasmic region (CTR). Deletion and mutagenesis studies indicated that the CTR but not the NTD is dispensable for ArgO function in vivo and that a pair of conserved aspartate residues, located near the opposing edges of the cytoplasmic membrane may play a pivotal role in facilitating transmembrane Arg flux. Given the importance of the
aspartate pair and its conservation amongst orthologs of ArgO, it is proposed that they may be positioned in the conduit for Arg translocation in ArgO, facilitating transmembrane export of a cationic amino acid like Arg. Additional studies on amino acid substitutions that impair ArgO function in vivo and their derivatives bearing compensatory amino acid alterations, implicate a role for interhelical interactions in the Arg export mechanism. The performed experimentation suggests that in vivo ArgO may exist as a monomer, highlighting thus the requirement for intramolecular interactions in ArgO as opposed to interactions across multiple ArgO monomers, in the formation of an Arg translocating conduit. Lastly, studies in chapter 3 indicate that ArgO possesses a membrane topology that is distinct from that reported for its ortholog LysE with the distinction restricted to the topological arrangement of the proximal one third portions of the two exporters. Studies on the topology of ArgO have been submitted for publication.

Dipeptides provide a means to increase the intracellular level of an amino acid as their entry into the cytoplasm occurs via the peptide uptake systems. Following their uptake, they are catabolised into their constituent amino acids which then leads to increase in the intracellular levels of the two amino acids of the dipeptide. The Lys-Ala dipeptide has previously been employed to increase the intracellular levels of Lys in *C. glutamicum* (Erdmann et al 1993). In studies described in chapter 2 it was observed that the ΔybjE (ΔlysO) mutant was found to be growth inhibited by the dipeptide Lys-Ala with growth inhibition correlating with elevated cellular Lys content. This phenotype of the ΔybjE mutant is consistent with the functional assignment for YbjE as a Lys exporter. In these studies unexpectedly the argO mutant was not growth inhibited by the presence of the L-arginylalanine (Arg-Ala) dipeptide in the medium which was surprising since in *C. glutamicum* the ΔlysE mutant is rendered sensitive to both Arg and Lys containing dipeptides (Vrljic et al 1996, Bellmann et al 2001). This observation indicated that *E. coli* appeared to
bear a mechanism(s), to alleviate the potential growth inhibitory effects of elevated cytoplasmic Arg levels following uptake of Arg containing dipeptides (Pathania and Sardesai 2015). The studies described in chapter 4 are directed towards understanding the genetic basis of the Arg-Ala resistant (RA\textsuperscript{R}) phenotype of the \textit{argO} mutant. It was hypothesized that another protein perhaps another Arg exporter in \textit{E. coli} was responsible for the RA\textsuperscript{R} phenotype. Towards its identification an ampicillin enrichment strategy was employed to isolate mutant derivatives of the \textit{argO} strain bearing transposon insertions conferring an Arg-Ala sensitive (RA\textsuperscript{S}) phenotype. Further studies showed that the multiple transposon insertions in \textit{ydhE}, encoding a membrane protein belonging to the MATE (multidrug and toxic compound extrusion) family of exporter proteins (Brown et al 1999), conferred an RA\textsuperscript{S} phenotype in the \textit{argO} mutant. The RA\textsuperscript{S} phenotype correlated with absence of YdhE, as a clean deletion of \textit{ydhE} also led the RA\textsuperscript{S} phenotype in a strain lacking ArgO. In addition expression of an ortholog of YdhE, NorM from \textit{V. cholerae} (Singh et al 2006, He et al 2010) was shown to complement the RA\textsuperscript{S} phenotype. Upon closer examination it was observed that to a large extent the RA\textsuperscript{S} phenotype resulted from absence of YdhE and that the \textit{argO} mutation accentuated the RA\textsuperscript{S} phenotype. Furthermore, absence of YdhE did not lead to a discernible CAN\textsuperscript{SS} phenotype implying that YdhE may not play any role in mediating Arg export. Circumstantial evidence indicated that the RA\textsuperscript{S} phenotype of the \textit{argO ydhE} double mutant does not occur due to elevated intracellular levels of Arg but is specific to the presence of Arg-Ala in the medium as the \textit{argO ydhE} double mutant was not inhibited by the presence of the L-alanylarginine (Ala-Arg) dipeptide in the medium. In addition, it was found that the RA\textsuperscript{S} phenotype could be alleviated by the presence of 20 amino acids in the medium. In order to understand the physiological defect in the \textit{argO ydhE} double mutant causal to its RA\textsuperscript{S} phenotype, suppressor studies were performed which showed that a variety of recessive genetic lesions in \textit{tppB} encoding the di-tripeptide permease suppressed the RA\textsuperscript{S} phenotype.
The property of TppB to mediate preferential uptake of dipeptides bearing a positively charged amino acyl R group (side chain of an amino acid) (Harder et al 2008), provides a rationale to account for the suppression of the RA$^S$ phenotype by mutations in $tppB$. Studies in this chapter describe a new phenotype for a null mutation in $ydhE$. Based on these studies it is suggested that YdhE may mediate export of Arg-Ala and that ArgO may also contribute to the export. Furthermore, it is speculated that Arg-Ala may serve as a proxy for an as yet unknown, naturally occurring substrate for YdhE (and ArgO), probably an antimicrobial compound. The physiological defect causal to the RA$^S$ phenotype of the $argO$ $ydhE$ double mutant remains to be determined, as also the basis for resistance of the $argO$ mutant of $E. coli$ to Arg containing dipeptides.

Chapter 5 reviews and summarizes the work done in this thesis and, possible future experimentation in relation to the current work is outlined.