

Synopsis

The present thesis describes work undertaken to determine the role of the ArgP protein in *E. coli* physiology. Primarily, the project was initiated because of two previous findings from this laboratory, namely (i) the *gltBD argP* double mutant is (osmo)sensitive to hyperosmotic stress induced by high concentrations of NaCl in the growth medium (Nandineni *et al.*, 2004); and (ii) ArgP is a transcriptional regulator of the arginine (Arg) exporter, *argO* (Nandineni and Gowrishankar, 2004; Laishram and Gowrishankar, 2007).

Bacteria, when under hyperosmotic stress, accumulate glutamate (Glu) intracellularly through increased synthesis; Glu serves as a counter-ion for K^+ , which also accumulates under these conditions (Csonka, 1989; Csonka and Epstein, 1996). In *E. coli*, *gltBD* encodes Glu synthase, which catalyzes one of the two pathways of Glu synthesis. The second pathway is catalyzed by Glu dehydrogenase (GDH) encoded by *gdhA*. The osmosensitive *gltBD argP* strain showed reduced intracellular levels of Glu and GDH compared to a *gltBD* single mutant in both low and high osmolarity conditions (Nandineni *et al.*, 2004). Therefore, it was suggested that the osmosensitivity of the *gltBD argP* strain is because of limited Glu and that ArgP has a role in osmoregulation through regulation of *gdhA* (Nandineni *et al.*, 2004).

ArgO (previously designated as YggA) is one of the few amino acid exporters that have been reported in *E. coli*. It is an Arg exporter and shares highest sequence similarity with LysE of *Corynebacterium glutamicum* which mediates both Arg and lysine (Lys) export. LysE is under transcriptional control of LysG, a protein that shares sequence similarity with *E. coli* ArgP. Both Arg and Lys, through LysG, serve as inducers of *lysE* expression (Bellmann *et al.*, 2001). *In vivo* studies with *argO-lac* fusions have established that *argO* is also under the transcriptional control of ArgP and that its expression is induced by Arg as well as by its toxic analog canavanine (CAN). However, unlike the situation with *C. glutamicum* LysE, *argO* expression is repressed in the presence of Lys. Dominant gain-of-function mutations in *argP* (*argP^d*), that confer elevated, Arg independent *argO* expression and thereby a CAN resistant phenotype (CAN^r) were also identified (Nandineni and Gowrishankar, 2004). *In vitro* studies showed that ArgP in presence of either co-effectors, Arg or Lys, can recruit RNA polymerase (RNAP) to the *argO* promoter. This then presents two different scenarios, productive transcription in the former but a trapped RNAP and no transcription in the

latter (Laishram and Gowrishankar, 2007).

Additionally, there have been various reports from other laboratories on ArgP function in *E. coli*. Arthur Kornberg's group previously described ArgP as an inhibitor of chromosomal initiation (IciA). IciA was shown to bind the AT-rich, conserved 13-mers at *oriC* preventing opening of this region and thereby blocking chromosomal initiation of replication (Hwang and Kornberg, 1990, 1992; Thony *et al.*, 1991). Subsequently, IciA was reported to be a transcriptional activator of *dnaA* and *nrdA* (Lee *et al.*, 1997; Han *et al.*, 1998), genes known to have roles in chromosomal replication and DNA synthesis respectively. Interestingly, however, the *iciA* mutant shows no growth or replication related phenotype (Thony *et al.*, 1991). In yet other studies, IciA is reported as a nucleoid protein (Azam and Ishihama, 1999) that exhibits affinity for AT-rich and curved DNA sequences (Wei and Bernander, 1996; Azam and Ishihama, 1999). Recently ArgP was described as a transcriptional activator of the genes *dapB* (encoding an enzyme of the Lys biosynthetic pathway) in *E. coli* and *gdhA* in *Klebsiella aerogenes*. The co-effector Arg does not have any effect on expression of these genes; Lys represses expression and this is through its prevention of ArgP binding to the operator (Bouvier *et al.*, 2008; Goss *et al.*, 2008). Accordingly, this thesis describes studies that were undertaken to understand ArgP's role in *E. coli* physiology and the mechanism of its function.

To test the hypothesis that ArgP's role in osmoregulation is through regulation of GDH and hence Glu levels in the cell, the effect of ArgP on *gdhA-lac* expression *in vivo* and ArgP binding *in vitro* to the *gdhA* upstream regulatory region were determined. Three lines of evidence were obtained to substantiate the same hypothesis: (i) activation by ArgP of *gdhA-lac* transcription *in vivo* in a Lys-sensitive manner which clearly correlated with the manner of ArgP binding to the *cis* regulatory region *in vitro*; (ii) osmosensitive phenotype of two strains namely, $\Delta gltBD argP^+$ on supplementation with Lys, and $\Delta gltBD argP^d$ -P274S (harboring the *argP^d* allele that is $\Delta argP$ -like w.r.t. *gdhA-lac* expression); and (iii) the ability of multicopy *gdhA*⁺ to relieve the osmosensitive phenotype of the *gltBD argP* double mutant strain. Primer extension analysis and site directed mutagenesis also enabled the precise location of the functional and ArgP-regulated *gdhA* promoter corresponding to a transcription start site that was mapped at -63-bp upstream of the initiation codon of *gdhA*.

To investigate the role of ArgP as a transcriptional activator, data obtained from a whole genome differential gene expression microarray along with genes shortlisted using

a candidate gene approach (of genes repressed in the presence of Lys), were examined. Validation of ArgP regulation of the shortlisted genes was achieved through *in vivo* promoter-*lac* fusion experiments in *argP⁺*, *argP^d* and Δ *argP* strain backgrounds; electrophoretic mobility shift assays (EMSAs) of genes showing *in vivo* ArgP activation were also performed.

The microarray experiment revealed that *lysP*, the gene encoding a protein with the dual functions of a Lys-specific permease and a repressor of the Lys decarboxylase system, is a potential regulatory target of ArgP. The Δ *argP* mutation behaved like Δ *lysP* in conferring thialysine (toxic analog of Lys) resistance. Subsequent experiments showed that ArgP activates *lysP* by about 35-fold *in vivo* and that there is about 5-fold ArgP-dependent repression in presence of Lys. *In vitro* ArgP bound the *lysP* regulatory region, and similar to the findings at *gdhA* (and *dapB*), the binding affinity was diminished upon Lys addition. Furthermore, sequences of the *lysP* regulatory region between -114 and -76 (relative to the start site of transcription) were shown to carry important determinants for ArgP binding and regulation.

The Δ *argP* mutant did not phenocopy a Δ *lysP* mutant for expression of the Lys decarboxylase system (encoded by the *cadBA* operon) suggesting that the basal levels of LysP in a Δ *argP* strain is sufficient for the negative regulation of *cadBA* in absence of Lys.

Compared to the wild-type allele, the *argP^d* mutations showed varying *in vivo* effects on *lysP-lac* expression. A conspicuous difference was presented by *argP^d*-P274S, which showed much less activation than *argP⁺* and 3-fold Lys repression. Concomitantly, EMSAs showed no significant differences in binding affinities. On the other hand, differences in mobility of the protein-DNA complexes formed by native ArgP and the ArgP^d variants were seen, such that the ArgP^d-DNA complexes migrated faster than native ArgP-DNA complex. Such differences in mobility of protein-DNA complexes have been reported to be due to differences in DNA bending such that less bent DNA migrates faster (Wu and Crothers, 1984). It was thus proposed that the differences in ArgP^d regulation at *lysP* might be due to differences in bending of the regulatory DNA region.

Two other genes shortlisted from the microarray data as probable ArgP targets were *lysC* and *asd*, which encode enzymes of the Lys biosynthetic pathway. Using the candidate gene approach, *dapD* and *lysA*, also encoding enzymes of the Lys biosynthetic

pathway were further examined for ArgP regulation. *lysC*, *asd*, *dapD* and *lysA* were shown to be activated by ArgP to varying extents *in vivo* (2- to 4-fold). This activation was independent of the co-effector Arg but was Lys-sensitive. The *argP^d* variants used in this study behaved much like the *argP⁺* allele both for activation and Lys repression at these four loci. However, allele -P274S was again different from the others in that it behaved more like the Δ *argP* mutation; this effect was therefore strikingly different from its effect at *argO* where it activates by 270-fold. EMSA showed Lys-sensitive binding of ArgP to the regulatory regions of *lysC*, *asd*, *dapD* and *lysA*, suggesting that Lys repression is due to inability of the activator to bind to these operators. These results emphasized the fact that the mechanism of ArgP regulation at *argO* is fundamentally different from that at the other regulated loci. First, *argO* is the only gene that requires Arg as co-effector for its activation. Second, the -P274S variant of ArgP is by far the most effective for constitutive *argO* expression, whereas it behaves much like Δ *argP* for a majority of other target genes. Finally, ArgP binding to the *cis* regulatory region is Lys-insensitive at *argO* whereas it is Lys-sensitive at all other target regulatory loci.

Thereafter, the previously reported findings of ArgP activation of *dnaA* and *nrda* were reexamined. The results indicated that there are no differences in *argP⁺*, *argP^d* and Δ *argP* for *dnaA-lac* or *nrda-lac* expression suggesting that ArgP does not regulate their transcription *in vivo*. However, in conformity with reports from other groups, *in vitro* binding of ArgP to the regulatory regions of *dnaA* and *nrda* with moderate affinity was observed. This binding was Lys-insensitive.

The role of ArgP in the expression of the arginine uptake genes (*artP*, *artJ*, *hisJ* and *argT*) was also assessed. Results showed that ArgP does not regulate Arg import in *E. coli*. Similar to findings at *dnaA* and *nrda*, ArgP showed weak affinity, Lys-insensitive binding to the regulatory regions of genes encoding Arg uptake.

To examine if ArgP binds DNA non-specifically, an internal (from the coding region) *lacZ* DNA fragment was used in EMSA reactions. Negligible binding was noticed, lending support to the conclusion that ArgP does not bind DNA non-specifically. These studies permitted the conclusion that with the exception of *argO* (which showed Lys-insensitive, high affinity binding), a distinguishing feature between *in vivo* ArgP regulated and non-regulated genes was Lys-sensitive binding *in vitro*.

As previously mentioned, the protein-gene pairs ArgP-*argO* of *E. coli* and LysG-*lysE* of *C. glutamicum* are orthologous. Whereas LysE is an exporter of Arg and Lys

whose expression is induced by Arg, Lys, or histidine (His), ArgO exports Arg alone, and its expression is activated by Arg but not Lys or His. In an attempt to understand the inter-relationship between these orthologous pairs, transcriptional cross-regulation studies were performed. Consequently, reconstitution of activation of *lysE* by LysG in the presence of its co-effectors in *E. coli* was achieved. It was shown that neither ArgP nor LysG can regulate expression of the noncognate orthologous target. Of several ArgP^d variants, some namely ArgP^d-P274S, -S94L and to a lesser extent, -P108S activated *lysE* expression in *E. coli*. However, the individual activating effects of LysG and ArgP^d on *lysE* were mutually extinguished when both proteins were coexpressed in Arg- or His-supplemented cultures. In comparison with native ArgP, the active ArgP^d variants exhibited higher affinity of binding to the *lysE* regulatory region and less DNA bending at both *argO* and *lysE*. These findings led to the conclusions that the transcription factor LysG from a Gram-positive bacterium, *C. glutamicum*, is able to engage appropriately with the RNAP from a Gram-negative bacterium, *E. coli*, for activation of its cognate target *lysE in vivo* and that single-amino-acid-substitution variants of ArgP can also activate the distantly orthologous target *lysE*, but by a subtly different mechanism that renders them non-interchangeable with LysG.

The chapter-wise organization of this thesis describing the studies and findings summarized above is as follows:

Chapter 1 introduces topics related to this thesis and is organized into 4 sections describing, physiology of Arg and Lys metabolism in *E. coli*, ArgP as a transcriptional regulator in *E. coli*, LysG-*lysE* of *C. glutamicum* and finally the objectives addressed are listed.

Chapter 2 provides a description of experimental materials and methods used in this study.

Chapter 3 describes experiments performed to examine ArgP's role in osmoregulation through its regulation of *gdhA*.

Chapter 4 details results from the microarray experiment and ArgP regulation of the Lys-specific permease, *lysP*. The entire list of genes that showed at least 2-fold activation or 2-fold repression by ArgP^(d) in the microarray experiment are also listed in Appendix I and Appendix II respectively.

Chapter 5 is an account of ArgP regulation of genes *lysC*, *asd*, *dapD* and *lysA* each encoding enzymes of the Lys biosynthetic pathway.

Chapter 6 revisits the previous findings of ArgP regulation of genes involved in DNA

replication or metabolism and Arg uptake. Results as presented in this chapter suggest that ArgP is a non-canonical transcriptional regulator that binds upstream regulatory sites without a regulatory outcome.

Chapter 7 details the transcriptional cross-regulation studies between orthologous regulator-target gene pairs ArgP-*argO* of *E. coli* and LysG-*lysE* of *C. glutamicum*.

Chapter 8 provides a brief statement of the conclusions from this study and the future perspectives.