Oxygen-regulated gene expression in *Escherichia coli*

The 1992 Marjory Stephenson Prize Lecture

(Delivered at the 122nd Ordinary Meeting of the Society for General Microbiology, 25 March 1992)

JOHN R. GUEST*

Krebs Institute for Biomolecular Research, Department of Molecular Biology and Biotechnology, University of Sheffield, P.O. Box 594, Firth Court, Western Bank, Sheffield S10 2UH, UK

I would like to give my very sincere thanks to the Society for inviting me to deliver the Marjory Stephenson Prize Lecture. I regard it as a very special honour and can hardly believe that I am a worthy recipient. Marjory Stephenson died in 1948 when I was still at school, but I later came to regard myself as one of her many scientific grandchildren. My descent is through my postgraduate supervisor, D. D. Woods (1912–1964), who had been one of her students and had delivered the first Marjory Stephenson Memorial Lecture in 1953 (Woods, 1953). Marjory Stephenson laid the foundations of Chemical Microbiology in the English-speaking world and, as outlined in her published work and three editions of her monograph *Bacterial Metabolism* (Stephenson, 1949), she went as far as exploring metabolic regulation and adaptive enzyme synthesis in the period immediately prior to the discovery of the double helix and the widespread use of bacterial genetics. D. D. Woods related how he was first attracted to chemical microbiology as a schoolboy, when he listened to a radio broadcast entitled 'How Microbes Live or Some Aspects of Bacterial Physiology' given by Marjory Stephenson in 1930. I have no such reminiscences but as a biochemistry student in Leeds and Oxford during the 1950s, I was much influenced by several of her former students and her many disciples.

I have a long-standing interest in the genetic aspects of central metabolic pathways in *Escherichia coli* and the way these pathways are affected by the presence or absence of oxygen. Like many facultative anaerobes, *E. coli* can grow under aerobic and anaerobic conditions, deriving energy from a variety of respiratory processes or from fermentation (Fig. 1). Respiration involves membrane-associated proton-translocating electron transport

chains in which utilizable energy is generated by coupling substrate oxidation to the reduction of an electron acceptor. Under aerobic conditions the aerobic respiratory chain is used with oxygen as the electron acceptor but under anaerobic conditions, oxygen can be replaced by alternative electron acceptors such as fumarate, nitrate, nitrite, dimethylsulphoxide (DMSO) or trimethylamine-*N*-oxide (TMNO), and the corresponding anaerobic respiratory chain is induced. Alternatively, in the absence of an exogenous electron acceptor, energy has to be derived by fermentation. This involves a redox-balanced dismutation of the carbon source, substrate-level phosphorylation and the formation of a variety of fermentation products: acetate, ethanol, formate, H₂ and CO₂, plus a little lactate and succinate. The switch between different metabolic modes is accompanied by fairly dramatic changes in enzyme synthesis (including synthesis of the citric acid cycle enzymes) and we now know that there are both global and specific regulatory mechanisms which respond to the availability of oxygen and the alternative electron acceptors. What's more, in a given environment, these regulatory mechanisms ensure that the most energetically favourable process is used. So, oxygen is used in preference to nitrate, nitrate to fumarate, and fumarate to the endogenously generated electron acceptors of fermentation.

In this lecture I propose to discuss two inter-related aspects of oxygen-regulated gene expression. The first concerns the genetic organization of the citric acid cycle genes of *E. coli* with particular reference to the effects of anaerobiosis and to recent studies on fumarase and aconitase. The second concerns the properties of the transcriptional regulators ArcA and FNR, which control two networks of oxygen-regulated gene expression. I should stress this is a personal account rather than a comprehensive review, and from the outset I want to

* Tel. (0742) 768555; fax (0742) 728697;
e-mail M11JRG @ UK.AC.SHEFFIELD.PRIMEA.

make a collective acknowledgment to the invaluable contributions made by colleagues too numerous to mention individually.

Genetic organization of the citric acid cycle of *E. coli*

The citric acid cycle (CAC) catalyses the total oxidation of pyruvate (or more precisely, acetyl units) and it is a major energy-generating pathway in aerobic heterotrophs and an important source of precursors for cellular biosynthesis (Fig. 1). The genes encoding all of the CAC enzymes of *E. coli* have been cloned, sequenced and located in the linkage map (Fig. 2). There is a major cluster at 16.3 min encoding citrate synthase (CS), succinate dehydrogenase (SDH), the specific components of the 2-oxoglutarate dehydrogenase (ODH) complex (E1o and E2o) and succinyl-CoA synthetase (SCS). Then there is a smaller cluster at 2.8 min encoding the specific components of the pyruvate dehydrogenase (PDH) complex (E1p and E2p) and lipoamide dehydrogenase (E3), the common component of PDH and ODH complexes and the glycine cleavage system. The other genes are scattered about the linkage map. Many of these genes were cloned and sequenced in Sheffield and in most cases their nucleotide sequences provided the first primary structures for the corresponding enzymes, regardless of source. Collecting CAC genes was not a primary goal, but some of our laboratories were used by Krebs in the 1930s and 40s and I sometimes wonder whether the course of our work was subconsciously influenced by our historic setting.

In *E. coli* the CAC is an inducible pathway. It is most highly induced under conditions that make the greatest demands on its dual catabolic and anabolic functions, e.g. during aerobic growth on a non-fermentable substrate in a medium lacking preformed biosynthetic intermediates. Under these conditions most substrates, including glucose at non-repressing concentrations, are oxidized to CO₂ via the CAC and yield an abundance of reducing equivalents to energize the proton-translocation machinery. The cycle is subject to catabolite repression (glucose-mediated repression), anaerobic repression and end-product repression (amino acid repression). Thus under anaerobic conditions, or even under aerobic conditions in the presence of excess glucose, the cycle is transformed into its branched or non-cyclic form (Fig. 1). Carbon then flows at a much reduced rate through an oxidative route leading to 2-oxoglutarate, and through two reductive routes leading via fumarate to the minor fermentation product succinate. These routes satisfy the biosynthetic functions of the CAC and also provide fumarate as a useful anaerobic electron acceptor. In consequence, the carbon flows to the fermentation products (acetate, ethanol, and formate

Fig. 2. Circular linkage map of *E. coli* showing the locations of citric acid cycle and related genes (outside) and transcriptional regulator genes (inside).

which is interconvertible with H₂ and CO₂, plus a little lactate and succinate), and the recovery of utilizable energy is correspondingly low. Early studies showed that most of the CAC enzymes are repressed by glucose and further repressed by anaerobiosis (Amarasingham & Davis, 1965; Gray *et al.*, 1966). SDH and fumarase (FUM) were notable exceptions but it is now known that their repression is masked by the induction of analogous but differentially regulated enzymes (see below). The ODH complex is very severely repressed and the PDH complex is partially repressed and inhibited, whereas pyruvate formate lyase (PFL) is induced and activated.

Two structurally related pairs of enzymes in the dicarboxylate sector of the CAC

I was originally interested in the genetic organization of enzyme complexes that catalyse closely related reactions, e.g. the oxidative decarboxylation of pyruvate and 2-oxoglutarate and the interconversion of succinate and fumarate. At the outset I would never have predicted the astounding pace of molecular-genetic innovation nor that I would see the detailed molecular characterizations and structure-function comparisons that later emerged. Genetic studies with the PDH and ODH complexes established that the lipoamide dehydrogenase subunits (E3) are encoded by a single gene (*lpd*), and that this gene is regulated from the *ace* promoter as if it were the distal gene of the *aceEF-lpd* operon, when the PDH complex is synthesized, and it is also co-regulated with the *sucABCD* operon via its own promoter, when the ODH complex is synthesized (Spencer & Guest, 1985). A different situation exists for the interconversion of succinate and

fumarate. Here our studies helped to establish that there are two differentially regulated operons encoding independent aerobic and anaerobic enzymes (Wood *et al.*, 1984). SDH is synthesized aerobically to participate in the CAC, donating electrons to the aerobic respiratory chain (Fig. 1). However, SDH is repressed by glucose and anaerobiosis, being replaced by fumarate reductase (FRD), an analogous membrane-bound flavoprotein complex, which functions in the reductive branch of the anaerobic CAC and as the terminal reductase of a proton-translocating anaerobic respiratory pathway (Fig. 1).

These observations suggested that the adjacent step in the CAC might likewise be mediated by genetically distinct aerobic and anaerobic fumarases (Fig. 1). Fumarase catalyses the interconversion of fumarate and malate by the reversible addition of water across the fumarate double bond and we later found that *E. coli* possesses not one, nor even two, but three differentially regulated fumarase genes; *fumA*, *fumB* and *fumC* located at 35, 93 and 35 min, respectively (Fig. 2). Furthermore, two of the genes (*fumA* and *fumB*) specified enzymes belonging to a novel and hitherto unsuspected class of fumarase (Bell *et al.*, 1989*b*). They are virtually identical, dimeric ($2 \times 60\,000$) and oxygen-labile enzymes, each containing a [4Fe-4S] centre (Woods *et al.*, 1988). Studies with *fum-lacZ* fusions showed that FumA is the aerobic CAC enzyme, repressed by glucose and anaerobiosis (Woods & Guest, 1987). In contrast, FumB is essentially a malate dehydratase whose synthesis is unaffected by glucose and derepressed anaerobically so that it can function in the anaerobic CAC and in the conversion of malate to fumarate for use as a growth-supporting anaerobic electron acceptor. The third enzyme, FumC, is a typical oxygen-stable tetrameric fumarase ($4 \times 50\,000$); it is not related to FumA or FumB except for conservation in a 17-residue segment presumed to contain active-site residues, but it is 60% identical to the mammalian and yeast mitochondrial enzymes and a *Bacillus subtilis* fumarase. These are members of a wider family of enzymes that generate fumarate by trans-elimination reactions, including aspartase (38% identity), argininosuccinase (18%) and adenylosuccinase (19%). FumC is synthesized under aerobic and anaerobic conditions and it is the fumarase that is recovered from *E. coli* when no precautions are taken to prevent the inactivation of FumA and FumB. Whether any significance should be attached to the fact that *fumA* and *fumC* are adjacent genes is not known, nor can a specific function be assigned to FumC, except for some recent evidence suggesting that FumC is a member of the SoxRS regulon (Liochev & Fridovich, 1992). The function of this regulon is to maintain redox balance and normal metabolism under conditions of oxidative stress,

and replacing the unstable FumA by FumC would contribute to this.

In summary, it is clear that *E. coli* expresses two pairs of structurally analogous but differentially regulated CAC enzymes: FumA and SDH are the aerobic enzymes whereas FumB and FRD are their anaerobic counterparts (Fig. 1).

Aconitase

I would now like to describe some recent work on aconitase (ACN). This enzyme catalyses the interconversion of citrate and isocitrate, essentially by the reversible addition of water across the double bond of *cis*-aconitate to generate either citrate or isocitrate (Fig. 1). There were two reasons for looking at ACN. First, I anticipated that it might be structurally related to the new fumarases, because both types of enzyme contain [4Fe-4S] iron-sulphur centres and both catalyse analogous hydratase-dehydratase reactions. Second, the aconitase gene (*acn*) was the only CAC gene that had not been cloned or sequenced when we started, and no *acn* mutants of *E. coli* had ever been characterized.

The cloning proved quite difficult but was eventually accomplished by, devising a procedure for isolating this unstable monomeric enzyme (M_r 97 000), raising specific antisera for screening λ -*E. coli* gene banks, and then locating and sequencing the *acn* gene from immunopositive phages (Prodromou *et al.*, 1991, 1992). This work showed that apart from conserved motifs containing a potential active-site histidine residue and putative cysteine ligands for the Fe-S centre, there was no sustained sequence similarity between ACN and either type of fumarase. However, it did raise some interesting prospects for future work (see below).

During the course of our work, an X-ray structure and the amino acid sequence of porcine mitochondrial ACN were published (Robbins & Stout, 1989; Zheng *et al.*, 1990). The enzyme contains three closely packed domains separated from a fourth domain by a cleft leading to the active site and Fe-S centre. An alignment of the amino acid sequences of *E. coli* ACN and the porcine enzyme showed that they are 29% identical, 46% similar when conservatively-substituted residues are included (Fig. 3*a*). All but one of 20 active-site residues is conserved and it could be predicted that the two enzymes have similar structures. More interesting is the observation that the *E. coli* enzyme is 50% identical (67% similar) to a mammalian translational regulatory protein, the iron-responsive-element binding protein, IRE-BP (Rouault *et al.*, 1991; Hentze & Argos, 1991); see Fig. 3*b*. Under conditions of low iron, the IRE-BP binds to specific stem-loop structures (IRE) at the 5' or 3' ends of the relevant mRNAs, either inhibiting translation

Fig. 3. Comparison matrices for *E. coli* aconitase with (a) porcine mitochondrial aconitase and (b) human iron-responsive-element binding protein, IRE-BP. The DIAGON relationships are shown for a span of 23 residues and a double matching probability of <0.001 . The domain boundaries are derived from the structure of porcine aconitase.

(ferritin) or enhancing translation by stabilizing the mRNA (transferrin receptor). The observed similarities raise important questions concerning the functional identities of these catalytic and regulatory proteins. Indeed, it would now appear that IRE-BP may be a cytoplasmic aconitase (Kaptain *et al.*, 1991). However, it is not known whether *E. coli* ACN has an iron-responsive regulatory function. Such a relationship would provide a hitherto undetected link between iron availability and citric acid cycle activity. The isolation of the *acn* gene now offers an opportunity to investigate the properties of an *acn* mutant by replacing the chromosomal gene with an *in vitro* disrupted derivative. This could unmask the existence of a second ACN in *E. coli* and as outlined above, there are clear precedents for the occurrence of more than one form of some CAC enzymes in this organism. It is often thought that after 50 years there is little to learn about the CAC, but my experience suggests that it could still harbour some surprises.

FNR and oxygen-regulated gene expression

In recent years transcriptional regulators have been identified for two networks of oxygen-regulated gene expression. These are the ArcAB and FNR systems and their relationship to the foregoing discussion stems from the fact that each was discovered as a result of studies on the aerobic and anaerobic interconversion of succinate and fumarate; more precisely, the anaerobic repression of SDH and the anaerobic induction of FRD.

The ArcAB System

The ArcAB system was characterized by E. C. C. Lin and co-workers who isolated two classes of mutant that failed to repress SDH under anaerobic conditions (Iuchi

& Lin, 1988; Iuchi *et al.*, 1990; Lin & Iuchi, 1991). In fact, these mutants were unable to repress a whole host of aerobic functions, including the citric acid cycle, the glyoxylate cycle, fatty acid degradation and cytochrome *o*, the major aerobic cytochrome oxidase (Fig. 4). The corresponding genes were designated *arcA* and *arcB* to denote their lack of aerobic respiration control (Fig. 2) and their sequences showed that they encode the sensor and regulator components of a two-component signal-transducing system (Fig. 4). By analogy with other such systems (Stock *et al.*, 1989), it can be predicted that ArcB is a membrane-bound sensor which detects anaerobic stress in the aerobic respiratory chain, possibly through an increase in the ratio of ubiquinol to ubiquinone. It can likewise be predicted that ArcB is a histidine protein kinase, which autophosphorylates with ATP or acetyl phosphate in response to oxygen-limitation, and then by transphosphorylation converts the cytoplasmic regulator, ArcA, into its active phosphorylated form (Fig. 4). So, ArcA is the anaerobic repressor for a variety of aerobic functions. It is also an anaerobic transcriptional activator for cytochrome *d*, the oxidase that is expressed under micro-aerobic conditions. The *arcA* gene had previously been denoted by symbols relating to other features of its pleiotropic mutant phenotype, e.g. increased sensitivity to toluidine blue (*dye*) and altered sex factor regulation (*sfrA*) or control of pilus expression (*cpxC*), and it was as *dye* that its product was first identified as the regulator of a two-component system (Drury & Buxton, 1985). In controlling conjugal fertility, ArcA is converted into a specific activator for sex factor functions by interaction with another membrane-bound sensor, CpxA, which seems to detect opportunities for mating. Thus it would appear that ArcA can regulate two independent groups of target genes in response to signals generated by different sensors.

Fig. 4. The ArcAB two-component signal-transducing system for aerobic respiration control, and representative target functions that are repressed (–) or activated (+). From Spiro & Guest (1991, *Trends in Biochemical Sciences* **16**, 310–314); reproduced with permission.

Fig. 5. Speculative mechanism for the anaerobic formation of active FNR, and a selection of target functions that are activated (+) or repressed (–). From Spiro & Guest (1991, *Trends in Biochemical Sciences* **16**, 310–314); reproduced with permission.

The FNR system

The FNR system was identified in studies with pleiotropic mutants that cannot use fumarate, nitrate and nitrite as terminal electron acceptors in anaerobic respiration (Spiro & Guest, 1990; Unden & Trageser, 1991). The corresponding gene is designated *fnr* to denote the defects in *f*umarate and *n*itrate *r*eduction (Fig. 2). It is now clear that under anaerobic conditions the *fnr* gene product, FNR, is converted into a transcriptional activator for a host of anaerobic func-

tions including the terminal reductases of anaerobic respiration, pyruvate formate lyase, FumB, aspartase and cytochrome *d* (Fig. 5). It also functions as the anaerobic repressor of several aerobic activities including NADH dehydrogenase II and cytochrome *o*, and it also represses its own synthesis. Interestingly, cytochrome *o* is repressed by ArcA and FNR whereas cytochrome *d* is activated by ArcA and FNR; the latter occurs in a mutually competitive manner which fine-tunes cytochrome *d* expression under micro-aerobic conditions.

Thus, FNR represents the master-switch which ensures that aerobic respiration is used in preference to anaerobic respiratory metabolism or fermentation, simply because important anaerobic genes are not expressed unless FNR is in its active (anaerobic) form. Under anaerobic conditions, further metabolic preferences are established by other regulators such as the NarXL system. This is another two-component signal-transducing system in which the sensor (NarX) detects nitrate and the regulator (NarL) induces nitrate reduction and represses the reduction of both fumarate and trimethylamine-*N*-oxide (Lin & Iuchi, 1991).

I now want to outline what is known about the mode of action of FNR, particularly how it senses and responds to anaerobiosis and how it recognizes its target genes.

(i) *Sensing and response to anaerobiosis.* The first significant clues about the role of the *fnr* gene came from its nucleotide sequence. This established the resemblance between FNR and CRP or CAP, the cAMP receptor protein or catabolite gene activator protein (Shaw *et al.*, 1983). CRP is the transcriptional regulator which mediates catabolite repression, ensuring that glucose is used in preference to lactose and other carbohydrates. The cell's readiness to express the genes that are subject to catabolite repression is signalled by cAMP, which induces specific binding between CRP dimers and target promoters and initiates transcription (Fig. 6). By analogy it was predicted that FNR activates the transcription of anaerobic genes in response to an unknown co-effector or stimulus that signals oxygen starvation. Unfortunately, FNR proved to be a rather troublesome relative of CRP and until recently progress has been rather slow.

Sequence comparison and secondary structure prediction indicates that all of the structural elements of CRP are retained in FNR, including the helix-turn-helix motif in the DNA-binding domain and the eight β -strands in the nucleotide-binding or sensory domain (Fig. 7). However, there are some notable differences. First, there is a cysteine-rich N-terminal extension in FNR, which has no counterpart in CRP. It contains four of the five cysteine residues in FNR and all but the most N-terminal cysteine is essential. Second, the residues that contact cAMP are not conserved in FNR and there is no evidence that cAMP serves as a co-effector for FNR. Third, purified FNR is monomeric, unlike CRP which is dimeric. And fourth, FNR binds up to one atom of iron per monomer and the iron content is inversely related to the sulphhydryl content, indicating that cysteine residues provide the iron ligands. So, our current working hypothesis is that the cysteine-rich N-terminal segment forms part of a metal-binding redox-sensing domain, and that a conformational change accompanying the reduc-

Fig. 6. Summary scheme for transcriptional activation by CRP or FNR at the promoters of respective target genes in response to signals communicating glucose or oxygen limitation.

tion of the cysteine-bound iron cofactor converts FNR into the active form (Fig. 5). This model infers that FNR is its own sensor. It would also appear that FNR detects the cellular redox-state rather than oxygen *per se*, because the switch can be mediated by ferricyanide under anaerobic conditions (Unden & Trageser, 1991). It is also very significant that mutants (FNR*) which can activate the nitrate reductase expression under aerobic conditions have been isolated (Kiley & Reznikoff, 1991). The corresponding mutations affect residues either in the proposed redox-sensing domain or the putative dimer interface, as if they fix the molecule in its transcriptionally active conformation. Other switching mechanisms including the uptake and release of iron, the modification of one or more of the essential sulphhydryl groups, a monomer-dimer transition, and the interaction with an unidentified redox-signalling co-effector have not been ruled out, so there is still a lot to learn about this system.

(ii) *DNA-binding and the molecular basis of target recognition.* A consensus for the FNR-binding site was originally identified by comparing the promoters of FNR-regulated genes, and this has now been amply confirmed by mutation and studies with synthetic sites. The sequence closely resembles the CRP-binding site, the major difference being the presence of a well-conserved TTGA core motif in each FNR half-site compared with GTGA in the CRP half-sites (Fig. 7).

FNR: A-A-**TTGAT**----**ATCAAT**---
CRP: AA-**TGTGA**-----**TCACA**-TT

In fact, natural and synthetic FNR-sites can be converted into CRP-sites and vice versa, simply by interchanging T/G and A/C at the critical position in each half-site (Bell *et al.*, 1989a; Zhang & Ebright, 1990).

Fig. 7. Predicted structure of the FNR monomer based on the known structure of CRP. The core motifs in FNR- and CRP-binding half-sites are indicated, and the specificity-conferring interactions between base-pairs in these sites and residues in the DNA-recognition helices of the regulators are arrowed.

Where studied, the promoters of genes that are activated by FNR have no -35 sequence but they have an FNR-site centred at -39 to -49 relative to the transcription start-point, whereas genes that are repressed by FNR have a site which either overlaps the -10 sequence and $+1$ startpoint (*fnr*) or is just upstream of the -35 sequence (*ndh*). Attempts to demonstrate site-specific DNA-binding by gel-retardation (band-shift) have so far failed, but purified monomeric FNR was shown to bind specifically to the predicted sites in positively and negatively regulated genes, in DNAaseI footprinting studies; it was also apparent from the size of the protected region that two FNR monomers are bound (Green *et al.*, 1991; Sharrocks *et al.*, 1991). In the same studies, purified FNR was shown to activate and repress transcription *in vitro*, whereas a mutant protein substituted at an essential cysteine codon (C122A) would not. These *in vitro* experiments represent important milestones because they provide tangible support for the original prediction that FNR and CRP are structurally and functionally related transcriptional regulators. Further support comes from the use of site-directed mutagenesis to explore the molecular basis of DNA recognition and to interconvert the binding-specificities of FNR and CRP. The DNA-binding specificity depends on interactions between the side-chains of the

amino acids on the solvent-exposed face of the DNA-recognition helix and the nucleotide base-pairs in the target core-motifs located in the major grooves of the DNA duplex (Fig. 7). In CRP, R180 and E181 make important contacts with the two GC base-pairs in the core motif. The residues on the DNA-binding face of FNR can be predicted from comparisons with CRP, and using site-directed mutagenesis, E209 and S212 proved to be essential (Spiro *et al.*, 1990). This indicates that the E→GC interaction is conserved in FNR and CRP and that the discriminatory contacts are provided by S212 and R180 interacting with the crucial TA or GC base-pairs in the respective FNR and CRP sites (Fig. 7).

This analysis has been taken a stage further by constructing regulators in which the binding specificities were altered by interconverting the three unique residues in the DNA-recognition helices of FNR (G, S212, V208) and CRP (K, G184, R180) as they appear in Fig. 7. This produced an altered FNR which allows good activation of CRP-dependent genes in response to anaerobiosis, and an altered CRP which activates FNR-dependent genes, albeit weakly (Spiro & Guest, 1987; Spiro *et al.*, 1990). Strains containing these genes should confuse oxygen starvation with glucose starvation or vice versa and this possibility is now being investigated to see whether the ensuing metabolic insanity might be

Fig. 8. The CRP-FNR family of regulatory proteins. Diagrammatic alignments for CRP and FNR showing the essential cysteine residues (C) and relevant residues in the helix-turn-helix motifs (open boxes). Related proteins are tabulated to indicate: sequence identity with FNR (%); function regulated; presence of conserved cysteine residues (Cys₄); putative target sequences, CRP-like (GTGA) or FNR-like (TTGA); and conservation of key residues in the DNA-recognition helices. See Spiro & Guest (1991) for references.

exploitable, e.g. to induce aerobic nitrate respiration or to diversify anaerobic carbon metabolism in the presence of glucose.

Global regulators and overlapping regulatory networks

Oxygen controls several important physiological processes in bacteria, notably nitrogen fixation, denitrification and photosynthesis (in the purple non-sulphur bacteria), as well as respiration and fermentation. There is also increasing evidence that the two types of regulatory system described here for controlling oxygen-regulated gene expression are not confined to *E. coli*. Indeed, the ArcAB system belongs to a large family of two-component systems which regulate gene expression in response to a diverse range of environmental stimuli. Likewise, FNR is a member of the growing CRP-FNR family of regulatory proteins, and it would appear that, at least in Gram-negative bacteria, this type of regulator has been recruited to control a variety of physiological functions (Fig. 8). Thus: FixK and FnrN are involved in the regulation of nitrogen fixation in *Rhizobium meliloti* and *Rhizobium leguminosarum* (respectively); HlyX probably regulates haemolysin synthesis in the swine pathogen *Actinobacillus pleuropneumoniae*; and ANR regulates anaerobic arginine and nitrate reduction in *Pseudomonas aeruginosa*. All except CRP and FixK retain the four essential cysteine residues found in FNR, suggesting that they are redox-responsive regulators. Furthermore, all except CRP seem to be associated with FNR-like targets, and consistent with the binding interactions deduced above, their DNA-recognition helices retain the conserved glutamate residue and the discriminatory serine residue. It is also significant that database searches reveal potential FNR-sites in the promoters of

genes from diverse species. These include species that are not known to contain a homologue of FNR, but they are generally competent in anaerobic respiration (*Proteus vulgaris*, *Alcaligenes faecalis*, *Alcaligenes denitrificans*), denitrification (*Pseudomonas stutzeri*), nitrogen fixation (*Azorhizobium caulinodans*, *Bradyrhizobium japonicum*) or bacterial photosynthesis (*Rhodobacter sphaeroides*). There is still much to learn about the molecular basis of redox-sensing and transcription control in the ArcAB and FNR systems, and it is conceivable that further systems will emerge as related physiological switches are investigated in more detail.

In conclusion, it is now clear that the pattern of metabolism adopted by an organism is the product of a complex network of interacting and overlapping regulatory processes. A small sector of the overall scheme illustrating different tiers of regulation is shown in Fig. 9. We once thought solely in terms of genes (regulons) or operons controlled by specific regulators and their co-effectors, but we must now include the global regulators and their families of target regulons and operons. These super-families have been usefully termed 'modulons' (Iuchi & Lin, 1988). So, there are modulons for ArcA, FNR, CRP, NarL (nitrate), SoxR (superoxide) and Fur (iron uptake), and different genes can belong to more than one modulon. Membership of a modulon facilitates integration into the diverse network of gene expression associated with a specific metabolic mode or physiological state, and membership of several modulons controls the expression of individual genes to that required in different environments. Thus, the CAC genes belong to the ArcA and CRP modulons, cytochrome *o* gene belongs to the ArcA, FNR and CRP modulons, and cytochrome *d* gene to the ArcA and FNR modulons. The NarL modulon contains the fumarate and nitrate

Fig. 9. Overlapping network of super-families (modulons) containing operons and genes controlled by different global regulators. The regulatory proteins can either activate (+) or repress (−) the transcription of genes within specific modulons. Adapted from Lin & Iuchi (1991).

reductase genes and lies entirely within the FNR modulon, where it maintains the anaerobic preference for nitrate before fumarate. Then at yet another level, gene expression is affected by DNA-supercoiling, which in turn responds to anaerobiosis and osmolarity, but this is beyond the scope of my lecture.

I am more than grateful to numerous colleagues and co-workers who have made invaluable contributions to the work described and I count myself extremely fortunate to have had the freedom to study topics of my own choice. I am also grateful to the editor of Elsevier Trends Journals for permission to reproduce Figs 1, 4 and 5.

References

- AMARASINGHAM, C. D. & DAVIS, B. D. (1965). Regulation of α -ketoglutarate dehydrogenase formation in *Escherichia coli*. *Journal of Biological Chemistry* **240**, 3664–3668.
- BELL, A. I., GASTON, K. L. & BUSBY, S. J. W. (1989a). Cloning of binding sequences for the *Escherichia coli* transcriptional activators FNR and CRP, location of bases involved in discrimination between FNR and CRP. *Nucleic Acids Research* **17**, 3865–3874.
- BELL, P. J., ANDREWS, S. C., SIVAK, M. N. & GUEST, J. R. (1989b). Nucleotide sequence of the FNR-regulated fumarase gene (*fumB*) of *Escherichia coli* K-12. *Journal of Bacteriology* **171**, 3494–3503.
- DRURY, L. S. & BUXTON, R. S. (1985). DNA sequence analysis of the *dye* gene of *Escherichia coli* reveals amino acid homology between the Dye and OmpR proteins. *Journal of Biological Chemistry* **260**, 4236–4242.
- GRAY, C. T., WIMPENNY, J. W. T. & MOSSMAN, M. R. (1966). Regulation of metabolism in facultative bacteria. II. Effects of aerobiosis, anaerobiosis and nutrition on the formation of Krebs cycle enzymes in *Escherichia coli*. *Biochimica et Biophysica Acta* **117**, 33–41.
- GREEN, J., TRAGESER, M., SIX, S., UNDEN, G. & GUEST, J. R. (1991). Characterization of the FNR protein of *Escherichia coli*, an iron-binding transcriptional regulator. *Proceedings of the Royal Society of London B244*, 137–144.
- HENTZE, M. W. & ARGOS, P. (1991). Homology between IRE-BP, a regulatory RNA-binding protein, aconitase, and isopropylmalate isomerase. *Nucleic Acids Research* **19**, 1739–1740.
- IUCHI, S. & LIN, E. C. C. (1988). *ArcA* (*dye*), a global regulatory gene in *Escherichia coli* mediating repression of enzymes in aerobic pathways. *Proceedings of the National Academy of Sciences of the United States of America* **85**, 1888–1892.
- IUCHI, S., MATSUDA, Z., FUJIWARA, T. & LIN, E. C. C. (1990). The *arcB* gene of *Escherichia coli* encodes a sensor-regulator protein for anaerobic repression of the *arc* modulon. *Molecular Microbiology* **4**, 715–727.
- KAPTAIN, S., DOWNEY, W. E., TANG, C., PHILPOTT, C., HAILE, D., ORLOFF, D. G., HARFORD, J. B., ROUAULT, T. A. & KLAUSNER, R. D. (1991). A regulated RNA binding protein also possesses aconitase activity. *Proceedings of the National Academy of Sciences of the United States of America* **88**, 10109–10113.
- KILEY, P. J. & REZNIKOFF, W. S. (1991). Fnr mutants that activate gene expression in the presence of oxygen. *Journal of Bacteriology* **173**, 16–22.
- LIN, E. C. C. & IUCHI, S. (1991). Regulation of gene expression in fermentative and respiratory systems in *Escherichia coli* and related bacteria. *Annual Review of Genetics* **25**, 361–387.
- LIOCHEV, S. I. & FRIDOVICH, I. (1992). Fumarase C, the stable fumarase of *Escherichia coli*, is controlled by the *soxRS* regulon. *Proceedings of the National Academy of Sciences of the United States of America* **89**, 5892–5896.
- PRODROMOU, C., HAYNES, M. J. & GUEST, J. R. (1991). The aconitase of *Escherichia coli*: purification of the enzyme and molecular cloning and map location of the gene (*acn*). *Journal of General Microbiology* **137**, 2505–2515.
- PRODROMOU, C., ARTYMIUK, P. J. & GUEST, J. R. (1992). The aconitase of *Escherichia coli*. Nucleotide sequence of the aconitase gene and amino acid sequence similarity with mitochondrial aconitases, the iron-responsive-element-binding protein and isopropylmalate isomerases. *European Journal of Biochemistry* **204**, 599–609.

- ROBBINS, A. H. & STOUT, C. D. (1989). The structure of aconitase. *Proteins* **5**, 289–312.
- ROUAULT, T. A., STOUT, C. D., KAPTAIN, S., HARFORD, J. B. & KLAUSNER, R. D. (1991). Structural relationship between an iron-regulated RNA-binding protein (IRE-BP) and aconitase: functional implications. *Cell* **64**, 881–883.
- SHARROCKS, A. D., GREEN, J. & GUEST, J. R. (1991). FNR activates and represses transcription *in vitro*. *Proceedings of the Royal Society of London* **B245**, 219–226.
- SHAW, D. J., RICE, D. W. & GUEST, J. R. (1983). Homology between CAP and Fnr, a regulator of anaerobic respiration in *Escherichia coli*. *Journal of Molecular Biology* **166**, 241–247.
- SPENCER, M. E. & GUEST, J. R. (1985). Transcription analysis of the *sucAB*, *aceEF* and *lpd* genes of *Escherichia coli*. *Molecular and General Genetics* **200**, 145–154.
- SPIRO, S. & GUEST, J. R. (1987). Activation of the *lac* operon of *Escherichia coli* by a mutant FNR protein. *Molecular Microbiology* **1**, 553–558.
- SPIRO, S. & GUEST, J. R. (1990). FNR and its role in oxygen-regulated gene expression in *Escherichia coli*. *FEMS Microbiology Reviews* **75**, 399–428.
- SPIRO, S. & GUEST, J. R. (1991). Adaptive responses to oxygen limitation in *Escherichia coli*. *Trends in Biochemical Sciences* **16**, 310–314.
- SPIRO, S., GASTON, K. L., BELL, A. I., ROBERTS, R. E., BUSBY, S. J. W. & GUEST, J. R. (1990). Interconversion of the DNA-binding specificities of two related transcription regulators, CRP and FNR. *Molecular Microbiology* **4**, 1831–1838.
- STEPHENSON, M. (1949). *Bacterial Metabolism*, 3rd edn. London: Longman.
- STOCK, J., NINFA, A. J. & STOCK, A. M. (1989). Protein phosphorylation and regulation of adaptive responses in bacteria. *Microbiological Reviews* **53**, 450–490.
- UNDEN, G. & TRAGESER, M. (1991). Oxygen regulated gene expression in *Escherichia coli*: control of anaerobic respiration by the FNR protein. *Antonie van Leeuwenhoek* **59**, 65–76.
- WOOD, D., DARLISON, M. G., WILDE, R. J. & GUEST, J. R. (1984). Nucleotide sequence encoding the flavoprotein and hydrophobic subunits of the succinate dehydrogenase of *Escherichia coli*. *Biochemical Journal* **222**, 519–534.
- WOODS, D. D. (1953). The integration of research on the nutrition and metabolism of micro-organisms. *Journal of General Microbiology* **9**, 151–173.
- WOODS, S. A. & GUEST, J. R. (1987). Differential roles of the *Escherichia coli* fumarases and *fnr*-dependent expression of fumarase B and aspartase. *FEMS Microbiology Letters* **48**, 219–224.
- WOODS, S. A., SCHWARTZBACH, S. D. & GUEST, J. R. (1988). Two biochemically-distinct classes of fumarase in *Escherichia coli*. *Biochimica et Biophysica Acta* **954**, 14–26.
- ZHANG, X. & EBRIGHT, R. H. (1990). Substitution of 2 base pairs (1 base pair per DNA half-site) within the *Escherichia coli lac* promoter DNA site for catabolite gene activator protein places the *lac* promoter in the FNR regulon. *Journal of Biological Chemistry* **265**, 12400–12403.
- ZHENG, L., ANDREWS, P. C., HERMODSON, M. A., DIXON, J. E. & ZALKIN, H. (1990). Cloning and structural characterization of porcine heart aconitase. *Journal of Biological Chemistry* **265**, 2814–2821.