Transcriptional repressors

Charlene Kahler

Dept of Microbiology and Immunology
School of Biomedical, Biomolecular and Chemical Sciences
The University of Western Australia
Tel: (08) 9346 2058
Fax: (08) 9346 4519
E-mail: ckahler@cyllene.uwa.edu.au

Prokaryotes regulate cellular functions in response to environmental cues via signal transduction pathways. In principal, there are two thematic organisations of signal transduction proteins:

- One-component proteins, in which the input and output domains are physically linked. These are commonly called activators (see article by Schubert & Shearwin) and repressors (reviewed in this article).
- Two-component systems consisting of histidine kinases and response regulators (see article by Cheung & Rood).

A recent survey of 145 prokaryotic genomes has revealed the predominance of one-component systems in this domain of life and has provided the first indication of the diversity inherent in these signal transduction pathways [1].

As might be expected, the genome size, lifestyle and environment affect the distribution of one-component systems, including repressors, in microorganisms. Generally, free living organisms contain a larger number of regulators per 1000 proteins than do bacteria inhabiting a specific habitat (e.g. obligate human pathogen) [1, 2]. This distribution can be disturbed by the acquisition of regulatory proteins encoded on mobile DNA elements such as plasmids, transposons and integrons [2, 3].

Architecture of one-component repressor proteins

Each repressor protein consists of at least two recognisable domains – the output or DNA binding domain (DBD) containing a Helix-Turn-Helix (HTH) motif and the input or effector binding domain (EBD) [4]. Prokaryotic regulators have been subdivided into superfamilies based principally on the amino acid homology of the DBD, whilst sub-families are grouped according to the EBDs within each superfamily (Table 1).

EBDs of regulators share a common evolutionary origin with receptor and transport proteins which bind the same small molecules [12, 13]. The modular construction of signal transduction systems enables the optimisation of responses to the plethora of environmental signals encountered by a living organism [14].

As can be seen from Table 1, repressors control both fundamental aspects of the maintenance of microbial growth such as carbon, energy and nutrient metabolism, in addition to features commonly associated with pathogenesis such as antibiotic resistance, virulence and conjugation.

This essay will explain the basic concepts of transcriptional repression, mechanisms of DNA binding and mechanisms of repression/derepression.

Modes of transcriptional repression

The ability of a one-component regulator to act as a repressor or activator is principally determined by the location of the DNA binding site relative to the transcriptional start point in the promoter region [11].

There are three recognised modes of transcriptional repression (Figure 1). Depending on conditions within the cell (for example temperature and ionic strength), repressors can adopt multiple modes during transcriptional repression [15]:

Steric hindrance of RNA polymerase (RNAP) binding to the promoter

The DBD of the repressor binds to a site, termed the operator, located upstream of the open reading frame start codon, and inhibits RNAP binding. Steric hindrance occurs when the operator overlaps with those bases that RNAP binds to as a closed complex. The classic example of this mechanism is the lactose operon repressor, LacI. Alternatively, the location of the repressor operator in relation to both the RNAP binding site and the binding site(s) of accessory factors can prevent the interaction between the two, resulting in inhibition of transcription.

In addition, oligomerisation of the bound repressor proteins can result in complete occlusion of the RNAP binding site as has been observed for ferric uptake regulator, Fur [16].

Inhibition of RNAP transition from closed to open complex

In this instance, the repressor binds (either upstream from RNAP, or on the opposite face of the DNA helix) to a site that allows simultaneous binding of RNAP, and inhibits RNAP-mediated melting of the DNA strands at the initiation region. MarA, which regulates functions that include antibiotic resistance, persistence and survival in *E. coli*, represses transcription from a variety of promoters using this mechanism [17].

Inhibition of RNAP clearance from the promoter

In this model, the repressor allows simultaneous binding of RNAP, the formation of open complexes and abortive transcripts. However, RNAP escape is blocked either by contact with an upstream bound repressor or by collision with a downstream bound repressor.

As an example, IclR, the repressor of the acetate operon in *E. coli*, binds to an operator sequence upstream of the promoter and interacts with the α-subunit of the promoter bound RNAP, resulting in destabilisation and disassociation of the open complex [6].

DNA binding mechanisms

In all of these models, the repressor binds to the operator sequence via the HTH of the DBD. Operator sequences consist of inverted repeats, termed half sites, which are separated by a spacer region. Composition of the inverted repeats (that is, the number of mismatches with the consensus sequence) and the spacing both dictate the binding avidity of the repressor to the operator.

One of the α-helices in the HTH motif serves as the recognition helix or sequence reading helix which inserts into the major groove of the DNA helix of the operator half site, thus making multiple specific interactions with the DNA through the amino acid side chains in this part of the repressor. The second helix of the HTH motif primarily supports the first helix via hydrophobic interactions with the DNA sequence.

Repressor proteins bind to two half sites of the operator on both DNA strands as a dimer, with many repressors binding as multimers of dimers (Figure 2). Although TetR and QacR both belong to the same superfamily and contain the same structural DBD, they do not share conserved operator sequences, and the interactions of the HTH motif with the operators involve different amino acid side chains resulting in completely distinct mechanisms of DNA binding [2].

These differences can be measured by the degree to which the DNA bends in the presence of the repressor bound to the operator site: TetR induces a 17 degree bend towards the protein to optimise the HTH position relative to the operator half sites, whilst QacR widens the major groove of the entire operator binding site, resulting in only a 3 degree bend in the DNA. As a result, both repressors bind to their cognate operator sites with equal avidity but use different mechanisms to do so.

How do repressors switch between repression and derepression?

The ability of the repressor to bind DNA is effectively controlled by whether or not the EBD is occupied by the cognate small molecule ligand. Detailed structural

Table 1. Superfamilies of transcriptional repressors.

Superfamily* & reference	Some regulated functions	DNA binding domain (DBD)/position †	Effector binding domains (EBDs)
TetR [2]	Biosynthesis of antibiotics, efflux pumps, osmotic stress and many more	Tetra helical bundle/ C-terminal	Variable containing many EBD superfamilies
LacI [4]	Carbon source utilisation	HTH/ N-terminal	PBP-I (periplasmic binding protein type-I domain) which bind small molecule ligands
ArsR/SmtB [5]	Metal resistance	3 stranded winged HTH/Central	Conserved
IclR Δ [6]	Carbon metabolism, efflux pumps, quorum sensing	HTH/N-terminal	GAF domain which binds ligands such as cNMPs, tetrapyrroles and formate ⌘
DeoR [4]	Sugar metabolism	Winged HTH/ N-terminal	DeoR-C which binds sugars, shares a common protein fold with phosphosugar isomerases such as ribose-5-phosphate isomerase
GntR [7]	General metabolism, conjugation	Winged HTH with a C-terminal helix/ N-terminal	Four major subfamilies (FadR, HutC, MocR, Ytr), one example containing PBP1
AsnC/Lrp Δ Ω [8]	Amino acid biosynthesis	Winged helix/ N-terminal	Amino acid sensing RRM-like and ACT folds
MarR Δ [9]	Multiple antibiotic resistance, virulence, response to environmental stress	Winged HTH with a C-terminal helix/	Sugar kinase domain and others
Crp/Fnr Δ [10]	Global responses, catabolite repression and anaerobiosis	Winged HTH/ C-terminal	cNMP binding domain
DxtR/Fur [5]	Metal uptake	Winged HTH/ N-terminal	DtxR-N which binds metal ions

* The majority of regulators in these superfamilies act as repressors. However, superfamilies such as LysR, AraC/XylS, LuxR, MarR, NtrC and OmpR which predominantly contain activators, also contain regulators that function as repressors [11].

† Each family is characterised by the type of HTH motif and the location of the DBD within the protein. There are three major HTH families based upon the type of folds they are comprised of: (a) variants of a simple three helical bundle (such as the tetrahelical bundle), (b) variants of the winged HTH domain (such as 3 stranded winged HTH and winged HTH with a C-terminal helix) and (c) a group of highly modified variants [4].

Δ IclR family also contains regulatory proteins that act as activators or are bifunctional [6]. The AsnC, MarR and Crp families also contain regulators that function as activators.

⌘ GAF = cGMP phosphodiesterase, adenylate cyclase, FhlA domain.

Ω This group has been recently been re-named as the feast/famine regulatory proteins (FFPR) [8].

analyses of apo (unoccupied) and bound forms of the metal-responsive transcriptional regulators from the ArsR/SmtB and DtxR/Fur families has elegantly demonstrated how ligand binding, in this case metals, governs the tertiary structure of the DBD [5].

A comparison of the two families is very interesting since ligand binding results in functionally opposite outcomes – the metallated form of ArsR/SmtB releases the operator resulting in derepression, whilst the metallated form of DxtR/Fur binds the operator, resulting in repression. The structure of Zn(II) bound SmtB reveals that the metal binding site in the α5 helix and the DBD are interconnected via a hydrogen bonded network which determines the quaternary structural conformation of the repressor. The presence of Zn(II) therefore controls a conformational switch that compacts the homodimer and changes the relative

Figure 1. Interactions between transcriptional repressor proteins, promoter regions, operator sequences and RNA polymerase (RNAP). The RNAP is shown as a multi-subunit enzyme containing the ß/ß', α-CTD and α-NTD subunits bound to a sigma factor (σ). The location of the operator sequence (green box) and repressor dimers (purple boxes) relative to the consensus –10 and –35 promoter sequences (blue boxes) are indicated in each model.

(A) Normal association of the RNAP with the promoter consensus sequences which initiates transcription from a transcription start point (red arrow).

(B) Steric hindrance of RNA polymerase binding to the promoter.

(C) and (D) Inhibition of RNA polymerase transition from closed to open complex.

(E) and (F) Inhibition of RNA polymerase clearance from the promoter; see text for details (adapted [7]).

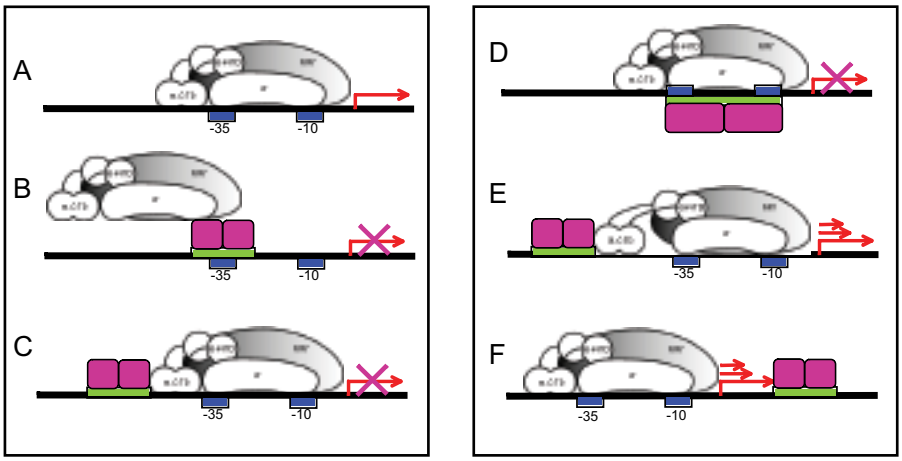
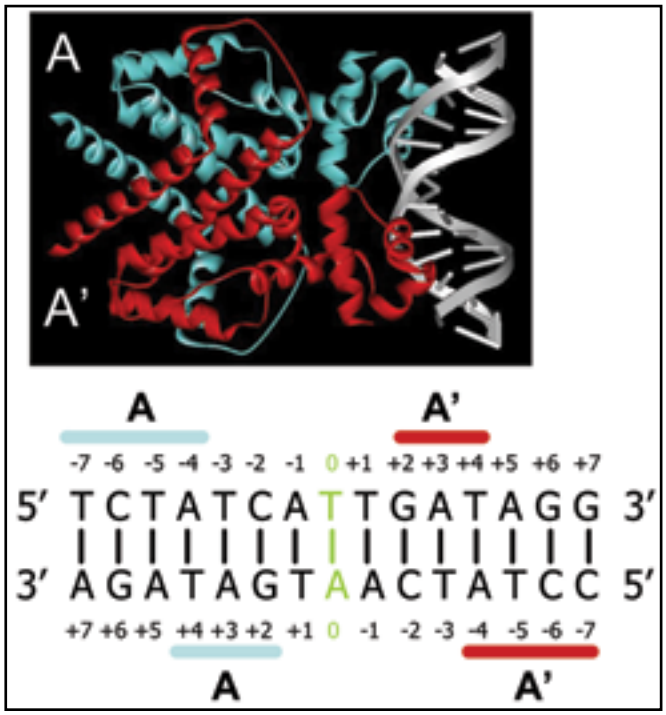


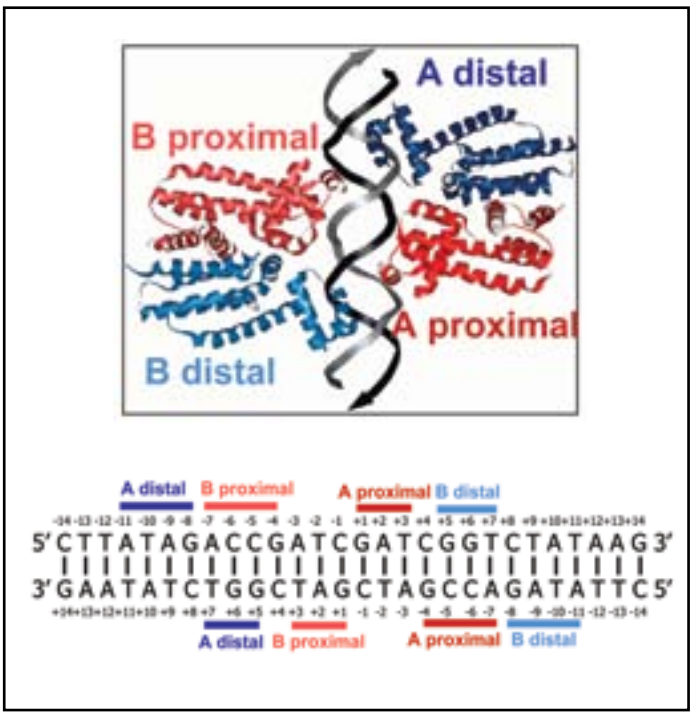
Figure 2. Two members of the TetR superfamily, TetR and QacR, bind cognate operators using different mechanisms. Each panel contains a ribbon diagram showing the relative position of the repressor protein subunits (colour coded to distinguish each subunit) with the DNA helix (from www.bactregulators.org). Below this panel is the nucleotide sequence of the operator aligned with coloured horizontal bars indicating those regions of the operator sequences which are in direct contact with the amino acid side chains of the sequence reading helix for each repressor.

A



Panel A. TetR dimer binds to one face of a palindromic operator sequence which overlaps the promoter sequence, thereby sterically hindering access to this site by RNA polymerase.

B



Panel B. QacR tetramer binds to both faces of an inverted repeat sequence downstream of the promoter and prevents RNA polymerase transition from the closed to open form (adapted [2]).

dispositions of the sequence reading helices, moving them out of register with respect to the major groove of the duplex DNA [5], resulting in release from the operator and derepression of the promoter. For members of the DxtR family, occupancy of the metal binding site stabilises the dimer by moving the sequence reading helices by 3-4 Å which improves the fit with the successive major grooves of the DNA [5], thus resulting in repression.

Recently, a second, and to date unique, mechanism of derepression has been noted for the *E. coli* global repressor Mlc [18].

Mlc is an unusual one-component repressor which is found in the two-component system family, NtrC/XylR, whose members predominantly function as activators. Although Mlc contains three structural domains, the protein consists of two functional units – an N-terminal DBD unit and a C-terminal EBD consisting of glucokinase folds characteristic of the ROK family [19].

In *E. coli* Mlc represses the expression of the glucose-specific transporter, PtsG. When glucose is absent, PtsG is phosphorylated on the EIIB domain and Mlc represses a global regulon including *ptsG*. However, in the presence of glucose, phosphorylation of PtsG is lost, and Mlc is sequestered to the inner membrane via a physical interaction with PtsG [18].

Somewhat surprisingly, ligand binding to the Mlc EBD has not been established, and the current model for derepression relies on the conformational changes of the repressor when bound to the operator sequence or unphosphorylated PtsG [18].

Future directions

In recent years, one of the most intriguing discoveries in transcriptional regulation was made in the plant pathogen *Agrobacterium tumefaciens*. This organism uses a conjugal transfer system to export T-DNA to the host plant cell, the first step in establishing a symbiotic relationship between the bacterium and the plant host. This process is controlled by the bacterial chromosomally encoded repressor, Ros, which controls the expression of the virulence genes involved in conjugal transfer. Ros is most closely related to eukaryote zinc finger proteins from the TFIIIA family [20] and represents the first prokaryote repressor which has a demonstrated function in a eukaryotic environment, in this instance, the control of the biosynthesis of cytokinin in the host plant cell [20]. It remains to be seen whether other bacterial pathogens, in particular, obligate human pathogens, utilise a similar strategy for regulating the host cell environment.

References

1. Ulrich LE, Koonin EV & Zhulin IB. One-component systems dominate signal transduction in prokaryotes. *Trends Microbiol* 2005; 13:52-6.
2. Ramos JL *et al.* The TetR family of transcriptional repressors. *Microbiol Mol Biol Rev* 2005; 69:326-56.
3. Ochman H, Lawrence JG & Groisman EA. Lateral gene transfer and the nature of bacterial innovation. *Nature* 2000; 405:299-304.
4. Aravind L, Anantharaman V, Balaji S, Babu MM & Iyer LM. The many faces of the helix-turn-helix domain: transcription regulation and beyond. *FEMS Microbiol Rev* 2005; 29:231-62.
5. Pennella MA & Giedroc DP. Structural determinants of metal selectivity in prokaryotic metal-responsive transcriptional regulators. *Biometals* 2005; 18:413-28.
6. Molina-Henares AJ, Krell T, Eugenia Guazzaroni M, Segura A & Ramos JL. Members of the IclR family of bacterial transcriptional regulators function as activators and/or repressors. *FEMS Microbiol Rev* 2006; 30:157-86.
7. Rigali S, Derouaux A, Giannotta F & Dusart J. Subdivision of the helix-turn-helix GntR family of bacterial regulators in the FadR, HutC, MocR and YtrA subfamilies. *J Biol Chem* 2002; 277:12507-15.
8. Yokoyama K, Ishijima SA, Clowney L, Koike H, Aramaki H, Tanaka C, Makino K & Suzuki M. Feast/famine regulatory proteins (FFRPs): *Escherichia coli* Lrp, AsnC and related archaeal transcription factors. *FEMS Microbiol Rev* 2006; 30:89-108.
9. Wilkinson SP & Grove A. Ligand-responsive transcriptional regulation by members of the MarR family of winged helix proteins. *Curr Issues Mol Biol* 2006; 8:51-62.
10. Korner H, Sofia HJ & Zumft WG. Phylogeny of the bacterial superfamily of Crp-Fnr transcription regulators: exploiting the metabolic spectrum by controlling alternative gene programs. *FEMS Microbiol Rev* 2003; 27:559-92.
11. Madan Babu M & Teichmann SA. Functional determinants of transcription factors in *Escherichia coli:* protein families and binding sites. *Trends Genet* 2003; 19:75-9.
12. Anantharaman V, Koonin EV & Aravind L. Regulatory potential, phyletic distribution and evolution of ancient, intracellular small-molecule-binding domains. *J Mol Biol* 2001; 307:1271-92.
13. Tam R & Saier MH Jr. Structural, functional, and evolutionary relationships among extracellular solute-binding receptors of bacteria. *Microbiol Rev* 1993; 57:320-46.
14. McAdams HH, Srinivasan B & Arkin AP. The evolution of genetic regulatory systems in bacteria. *Nat Rev Genet* 2004; 5:169-78.
15. Rojo F. Mechanisms of transcriptional repression. *Curr Opin Microbiol* 2001; 4:145-51.
16. Escolar L, Perez-Martin J & de Lorenzo V. Evidence of an unusually long operator for the fur repressor in the aerobactin promoter of *Escherichia coli*. *J Biol Chem* 2000; 275:24709-14.
17. Schneiders T & Levy SB. MarA-mediated transcriptional repression of the rob promoter. *J Biol Chem* 2006; 281:10049-55.
18. Bohm A & Boos W. Gene regulation in prokaryotes by subcellular relocalization of transcription factors. *Curr Opin Microbiol* 2004; 7:151-6.
19. Schiefner A, Gerber K, Seitz S, Welte W, Diederichs K & Boos W. The crystal structure of Mlc, a global regulator of sugar metabolism in *Escherichia coli*. *J Biol Chem* 2005; 280:29073-9.
20. Bouhouche N, Syvanen M & Kado CI. The origin of prokaryotic C2H2 zinc finger regulators. *Trends Microbiol* 2000; 8:77-81.