Structure and function of the LysR-type transcriptional regulator (LTTR) family proteins

Sarah E. Maddocks¹ and Petra C. F. Oyston²

¹Department of Oral and Dental Science, University of Bristol, Lower Maudlin Street, Bristol BS1 2LY, UK
²Biomedical Sciences, DSTL Porton Down, Salisbury SP4 0JQ, UK

The LysR family of transcriptional regulators represents the most abundant type of transcriptional regulator in the prokaryotic kingdom. Members of this family have a conserved structure with an N-terminal DNA-binding helix–turn–helix motif and a C-terminal co-inducer-binding domain. Despite considerable conservation both structurally and functionally, LysR-type transcriptional regulators (LTTRs) regulate a diverse set of genes, including those involved in virulence, metabolism, quorum sensing and motility. Numerous structural and transcriptional studies of members of the LTTR family are helping to unravel a compelling paradigm that has evolved from the original observations and conclusions that were made about this family of transcriptional regulators.

Introduction

The LysR-type transcriptional regulator (LTTR) family is a well-characterized group of transcriptional regulators. They are highly conserved and ubiquitous amongst bacteria, with functional orthologues identified in archaea and eukaryotic organisms (Pérez-Rueda & Collado-Vides, 2001; Sun & Klein, 2004; Stec et al., 2006). The LTTR family was formally documented by Henikoff et al. (1988), who concluded that there were at least nine functionally similar transcriptional regulatory proteins (identified in Escherichia coli, Salmonella enterica serovar Typhimurium, Rhizobium spp. and Enterobacter cloacae) which on the basis of sequence similarity and DNA-binding-domain (DBD) conservation, could be distinguished as a related group of bacterial transcriptional regulators. Extensive amino acid and dot-matrix comparisons assisted in the identification of additional, putative LTTRs, expanding the family considerably. LysR, the transcriptional activator of lysA (encoding diaminopimelate decarboxylase, an enzyme that catalyses the decarboxylation of diaminopimelate to produce lysine), had been the subject of considerable study at the molecular level and was the best characterized of the group at this time, hence becoming the family namesake (Stragier et al., 1983; Stragier & Patte, 1983). Since then, numerous LTTRs have been identified, and this family of regulators is continually increasing in size. Currently it comprises the largest known family of prokaryotic DNA-binding proteins, with 800 members identified on the basis of their amino acid sequence (Schell, 1993).

Originally LTTRs were described as transcriptional activators of a single divergently transcribed gene, which exhibited negative autoregulation (Lindquist et al., 1989; Schell, 1993; Parsek et al., 1994a). Extensive research has now led to them being regarded as global transcriptional regulators, acting as either activators or repressors of single or operonic genes; they are often divergently transcribed but can be located elsewhere on the bacterial chromosome (Heroven & Dersch, 2006; Hernández-Lucas et al., 2008). Co-inducers are recognized as being important for the function of LTTRs and often appear to contribute to a feedback loop in which a product or intermediate of a given metabolic/synthesis pathway (usually activated by an LTTR) acts as the co-inducer necessary for transcriptional activation or repression (Fig. 1) (Celis, 1999; van Keulen et al., 2003; Picossi et al., 2007).

The conservation of LTTRs within the genomes of extremely diverse bacteria means that they have evolved a regulatory role over genes with similarly diverse functions, whose products can be involved in metabolism, cell division, quorum sensing, virulence, motility, nitrogen fixation, oxidative stress responses, toxin production, attachment and secretion, to name a few (Table 1) (Kovacikova & Skorupski, 1999; Deghmane et al., 2000, 2002; Cao et al., 2001; Kim et al., 2004; Russell et al., 2004; Byrne et al., 2007; Lu et al., 2007; Sperandio et al., 2007). This review aims to bring together the increasing body of knowledge concerning the structure, functions and molecular genetics that is helping to unravel the paradigm of the largest group of transcriptional regulators identified within the prokaryotic kingdom.

Abbreviations: ABS, activation binding site; DBD, DNA-binding domain; HTH, helix–turn–helix; wHTH, winged-HTH; LTTR, LysR-type transcriptional regulator; RBS, regulatory binding site.
The origin and evolution of LTTRs

LTTRs are thought to be evolutionarily distinct and to have arisen in bacteria; strong evidence suggests that they can be acquired by horizontal transfer (discussed below). The common evolutionary descent of LTTRs is strongly implied from the study of amino acid and DNA sequence similarity, which suggests considerable structural and functional homology. Orthologues of LTTRs are present in numerous species of bacteria and have retained a conserved structure and function. Multiple paralogues of LTTRs can be present within a given genome; these are likely to have arisen by gene duplication. Subsequent evolutionary pressures and genetic divergence have led to the emergence of groups of orthologous paralogues of LTTRs. These remain structurally and functionally similar but have diverged to govern distinct regulons that often exhibit little or no cross-talk. Examples include the Nod and RubisCO subfamilies of LTTRs that are found conserved amongst numerous bacteria and are discussed later in the review.

The helix–turn–helix (HTH) DNA-binding domain

Despite the size of the LTTR family and the diverse function of LTTR-regulated genes, important structural regions remain highly conserved. LTTRs comprise approximately 330 amino acids; at the C terminus is a co-factor-binding domain and at the N terminus is a helix–turn–helix (HTH) motif, which provides a means of binding to DNA (Fig. 2). The HTH motif is present in all LTTRs and approximately 95% of all prokaryotic DNA-binding proteins. This far exceeds the number of other DNA-binding motifs, which include the helix–loop–helix, zinc-finger or β-sheet-anti-parallel domain, which make up the remaining 5% (Pérez-Rueda & Collado-Vides, 2001; Huffman & Brennan, 2002; Aravind et al., 2005). The ‘ancestral’ HTH motif comprises a three-helical bundle with an open conformation. The second and third helices of the bundle interact with DNA, the third being inserted into the major groove of the DNA double helix (Brennan & Matthews, 1989; Huffman & Brennan, 2002; Aravind et al., 2005). This so-called universal common ancestor has given rise to a number of variations that still carry out the same regulatory function; these include the winged-helix variety (of which the LysR family is a member) which possesses a β-pleated sheet hairpin between the second and third helix, the ribbon helix–helix structure and the tetra-helical structure (Fig. 3).

Most HTH-containing transcriptional regulators fall into two distinct groups, transcriptional activators or repressors. Transcriptional activators characteristically have the HTH located at the C terminus, whereas transcriptional repressors have the HTH at the N terminus (Pérez-Rueda & Collado-Vides, 2000). The LTTRs form a unique group, and have been termed dual regulators, in which the HTH is located 20–90 amino acids from the N terminus, regardless of whether the LTTR is activating or repressing the transcription of itself or the gene(s) it is regulating (Fig. 4)

A comprehensive phylogenetic tree compiled from amino acid sequence alignments inferred three putative subgroups of LTTRs (Schlaman et al., 1992b). Whether these subgroups are likely to be ‘true’ subgroups is uncertain given the reliance upon amino acid sequence alone as a basis; the regulators associated with each of the three groups have no particular defining factor, they do not necessarily regulate the same target genes (or those with a similar function) and they do not have the same co-inducer or origin.

Horizontal transfer of LTTRs

LTTRs are found throughout the different subdivisions of proteobacteria, with the majority represented in the α and γ subdivisions. Far fewer LTTRs have been identified for the β subdivision and the Gram-positives, and none have been identified in the δ subdivision (Schell, 1993). This is unlikely to be a true representation of the distribution of LTTRs amongst the different subdivisions and most likely reflects the extent of genetic characterization of members of
<table>
<thead>
<tr>
<th>LTTR</th>
<th>Regulation</th>
<th>Target gene function</th>
<th>Co-factor</th>
<th>Origin</th>
<th>Subdivision*</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>AlsR</td>
<td>Activator (local)</td>
<td>Acetoin synthesis</td>
<td>Acetate</td>
<td>Bacillus subtilis</td>
<td>Gm +</td>
<td>Renna et al. (1993)</td>
</tr>
<tr>
<td>AmpR</td>
<td>Activator (local)</td>
<td>β-Lactamase</td>
<td>Unknown</td>
<td>Rhodobacter capsulatus</td>
<td>x</td>
<td>Bartowsky &amp; Normark (1993)</td>
</tr>
<tr>
<td>BlaA</td>
<td>Activator (local)</td>
<td>β-Lactamase</td>
<td>Unknown</td>
<td>Acinetobacter spp.</td>
<td>γ</td>
<td>Collier et al. (1998)</td>
</tr>
<tr>
<td>CatM</td>
<td>Activator (global)</td>
<td>Aromatic compound degradation</td>
<td>cis,cis-Muconate</td>
<td>Acinetobacter calcoaceticus</td>
<td>γ</td>
<td>Ezezika et al. (2006)</td>
</tr>
<tr>
<td>CatR</td>
<td>Repressor (local)</td>
<td>Aromatic compound degradation</td>
<td>cis,cis-Muconate</td>
<td>Pseudomonas putida</td>
<td>γ</td>
<td>Chugani et al. (1998)</td>
</tr>
<tr>
<td>CbbR</td>
<td>Activator (global)</td>
<td>Carbon dioxide fixation</td>
<td>Auxotrophic growth conditions</td>
<td>Xanthobacter flavus</td>
<td>x</td>
<td>van Keulen et al. (2003)</td>
</tr>
<tr>
<td>ChiR</td>
<td>Activator (local)</td>
<td>Chitin binding/chitinase</td>
<td>Unknown</td>
<td>Serratia marcescens</td>
<td>γ</td>
<td>Suzuki et al. (2001)</td>
</tr>
<tr>
<td>CidR</td>
<td>Activator (local)</td>
<td>cid operon (murein hydrolase regulation)</td>
<td>Acetate</td>
<td>Staphylococcus spp.</td>
<td>Gm +</td>
<td>Yang et al. (2005)</td>
</tr>
<tr>
<td>CilR</td>
<td>Activator (local)</td>
<td>Chlorocatechol catabolism</td>
<td>cis,cis-Muconate</td>
<td>Neisseria meningitidis</td>
<td>γ</td>
<td>Dehghani et al. (1991)</td>
</tr>
<tr>
<td>CynR</td>
<td>Activator (local)</td>
<td>Cyanate detoxification</td>
<td>Cyanate</td>
<td>Escherichia coli</td>
<td>γ</td>
<td>Sung &amp; Fuchs (1992)</td>
</tr>
<tr>
<td>CysB</td>
<td>Activator (global)</td>
<td>Cysteine biosynthesis</td>
<td>N-Acetylserine</td>
<td>Salmonella enterica serovar</td>
<td>γ</td>
<td>van der Ploeg et al. (1997)</td>
</tr>
<tr>
<td>CysL</td>
<td>Activator (global)</td>
<td>Sulphite reductase</td>
<td>Sulphur source dependent</td>
<td>Bacillus subtilis</td>
<td>Gm +</td>
<td>Guillouard et al. (2002)</td>
</tr>
<tr>
<td>GltC</td>
<td>Activator (local)</td>
<td>Glutamate synthase</td>
<td>Bacillus subtilis</td>
<td>Gm +</td>
<td></td>
<td>Picossi et al. (2007)</td>
</tr>
<tr>
<td>HupR</td>
<td>Activator (global)</td>
<td>Haem uptake</td>
<td>Vibrio vulniicus</td>
<td>γ</td>
<td></td>
<td>Litwin &amp; Quackenbush (2001)</td>
</tr>
<tr>
<td>HvrB</td>
<td>Activator (global)</td>
<td>S-Adenosyl-l-homocysteine hydrolase expression</td>
<td>Light sensitivity</td>
<td>Rhodobacter capsulatus</td>
<td>x</td>
<td>Buggy et al. (1994)</td>
</tr>
<tr>
<td>IlvR</td>
<td>Activator (local)</td>
<td>Isoleucine/valine biosynthesis</td>
<td>Unknown</td>
<td>Caulobacter crescentus</td>
<td>x</td>
<td>Malakooti &amp; Ely (1994)</td>
</tr>
<tr>
<td>IlvY</td>
<td>Activator (local)</td>
<td>Isoleucine/valine biosynthesis</td>
<td>2-Acetolactate or 2-acetoxybutyrate</td>
<td>Escherichia coli</td>
<td>γ</td>
<td>Wek &amp; Hatfield (1988)</td>
</tr>
<tr>
<td>LrhA</td>
<td>Activator (global)</td>
<td>Flagella, motility and chemotaxis</td>
<td>Unknown</td>
<td>Escherichia coli</td>
<td>γ</td>
<td>Lehn et al. (2002)</td>
</tr>
<tr>
<td>LysR</td>
<td>Activator (local)</td>
<td>Lysine biosynthesis</td>
<td>Diaminopimelate</td>
<td>Escherichia coli</td>
<td>γ</td>
<td>Stragier et al. (1983)</td>
</tr>
<tr>
<td>MdcR</td>
<td>Activator (local)</td>
<td>Malonate catabolism</td>
<td>Malonate</td>
<td>Klebsiella pneumoniae</td>
<td>γ</td>
<td>Peng et al. (1999)</td>
</tr>
<tr>
<td>MetR</td>
<td>Activator (global)</td>
<td>Methionine and cysteine transport/biosynthesis</td>
<td>Homocysteine</td>
<td>Streptococcus spp.</td>
<td>Gm +</td>
<td>Kovaleva &amp; Gelfand (2007)</td>
</tr>
<tr>
<td>MleR</td>
<td>Activator (local)</td>
<td>Malolactic enzyme</td>
<td>Unknown</td>
<td>Lactococcus lactis</td>
<td>Gm +</td>
<td>Renault et al. (1989)</td>
</tr>
<tr>
<td>MtaR</td>
<td>Activator (global)</td>
<td>Methionine transport</td>
<td>Unknown</td>
<td>Group B streptococci</td>
<td>Gm +</td>
<td>Shesler et al. (2003)</td>
</tr>
</tbody>
</table>
Table 1. cont.

<table>
<thead>
<tr>
<th>LTTR</th>
<th>Regulation</th>
<th>Target gene function</th>
<th>Co-factor</th>
<th>Origin</th>
<th>Subdivision</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>MvrR</td>
<td>Activator (global)</td>
<td>Pathogenicity regulator</td>
<td>4-Hydroxy-2-heptylquinone</td>
<td><em>Pseudomonas aeruginosa</em></td>
<td>γ</td>
<td>Cao et al. (2001)</td>
</tr>
<tr>
<td>NagR</td>
<td>Activator (local)</td>
<td>Naphthalene catabolism</td>
<td>Salicylate</td>
<td><em>Ralstonia eutropha</em></td>
<td>α</td>
<td>Jones et al. (2003)</td>
</tr>
<tr>
<td>NahR</td>
<td>Activator (local)</td>
<td>Naphthalene/salicylate catabolism</td>
<td>Salicylate</td>
<td>NAH7 plasmid of <em>Pseudomonas putida</em></td>
<td>NA</td>
<td>Park et al. (2002)</td>
</tr>
<tr>
<td>NhaR</td>
<td>Activator (local)</td>
<td>Na⁺/H⁺ antiporter</td>
<td>Na⁺</td>
<td><em>Escherichia coli</em></td>
<td>γ</td>
<td>Dover &amp; Padan (2001)</td>
</tr>
<tr>
<td>NocR</td>
<td>Activator (local)</td>
<td>Nopaline catabolism</td>
<td>Octopine</td>
<td>Ti plasmids of <em>Agrobacterium tumefaciens</em></td>
<td>NA</td>
<td>von Lintig et al. (1994)</td>
</tr>
<tr>
<td>NodD</td>
<td>Activator (global)</td>
<td>Nitrogen fixation/symbiosis</td>
<td>Flavonoids</td>
<td><em>Rhizobium</em> spp.</td>
<td>β</td>
<td>Schlaman et al. (1992b)</td>
</tr>
<tr>
<td>OccR</td>
<td>Activator (local)</td>
<td>Octopine catabolism</td>
<td>Octopine</td>
<td>Ti plasmids of <em>Agrobacterium tumefaciens</em></td>
<td>NA</td>
<td>Habeeb et al. (1991)</td>
</tr>
<tr>
<td>OxyR</td>
<td>Activator (global)</td>
<td>Oxidative stress response (H₂O₂)</td>
<td>Redox changes</td>
<td><em>Salmonella enterica</em> serovar Typhimurium</td>
<td>γ</td>
<td>Farr &amp; Kogoma (1991)</td>
</tr>
<tr>
<td>PhcA</td>
<td>Activator (global)</td>
<td>Virulence regulator</td>
<td>Unknown</td>
<td><em>Pseudomonas solanacearum</em></td>
<td>γ</td>
<td>Brumbley et al. (1993)</td>
</tr>
<tr>
<td>QseA</td>
<td>Activator (global)</td>
<td>Quorum sensing</td>
<td>Unknown</td>
<td>Enteropathogenic/enterohemorrhagic <em>Escherichia coli</em></td>
<td>γ</td>
<td>Sperandio et al. (2002)</td>
</tr>
<tr>
<td>RbcR</td>
<td>Activator (local)</td>
<td>Carbon dioxide fixing</td>
<td>Unknown</td>
<td><em>Chromatium vinosum, Thiobacillus ferrooxidans</em></td>
<td>γβ</td>
<td>Viale et al. (1991)</td>
</tr>
<tr>
<td>RovM</td>
<td>Repressor (global)</td>
<td>Invasion/motility, virulence</td>
<td>Unknown</td>
<td><em>Yersinia pestis</em></td>
<td>γ</td>
<td>Heroen et al. (2007)</td>
</tr>
<tr>
<td>SpvR</td>
<td>Activator (local)</td>
<td>Virulence factor synthesis</td>
<td>Unknown</td>
<td><em>Salmonella</em> spp. virulence plasmids</td>
<td>NA</td>
<td>Sheehan &amp; Dorman (1998)</td>
</tr>
<tr>
<td>SyrM</td>
<td>Activator (global)</td>
<td>Exopolysaccharide synthesis</td>
<td>Unknown</td>
<td><em>Rhizobium meliloti</em></td>
<td>β</td>
<td>Barnett et al. (1998)</td>
</tr>
<tr>
<td>TcbR</td>
<td>Activator (local)</td>
<td>Chlorocatechol metabolism</td>
<td>2-Chloromuconic acid</td>
<td><em>Pseudomonas</em> spp. plasmid J51</td>
<td>NA</td>
<td>van der Meer et al. (1991)</td>
</tr>
<tr>
<td>ToxR</td>
<td>Activator (local)</td>
<td>Quorum sensing</td>
<td>Toxoflavin</td>
<td><em>Burkholderia glumae</em></td>
<td>β</td>
<td>Kim et al. (2004)</td>
</tr>
<tr>
<td>YofA</td>
<td>Activator (local)</td>
<td>Cell division</td>
<td>Unknown</td>
<td><em>Bacillus subtilis</em></td>
<td>Gm⁺</td>
<td>Lu et al. (2007)</td>
</tr>
<tr>
<td>YtxR</td>
<td>Activator (local/global)</td>
<td>ADP-ribosyl-transferase toxin</td>
<td>Unknown</td>
<td><em>Yersinia enterocolitica</em></td>
<td>γ</td>
<td>Axler-Diperte et al. (2006)</td>
</tr>
</tbody>
</table>

*Subdivisions of proteobacteria indicated by α, β, γ; Gm +, Gram positive; NA, not applicable.*
these subdivisions. The genes encoding LTTRs have a characteristically high G+C content, due to the distinct Lys/Arg ratio that is common to LTTR proteins (Henikoff et al., 1988; Viale et al., 1991). A number of LTTRs are found on transmissible regions of DNA, and the distinctive G+C percentage has enabled LTTRs that have been acquired by horizontal transfer to be identified within the genomes of many bacteria.

A well-documented example in which an LTTR has been acquired by horizontal transfer relates to the genes encoding ribulose-1,5-bisphosphate carboxylase/oxygenase (RubisCO) in the non-sulphur purple photosynthetic encoding ribulose-1,5-bisphosphate carboxylase/oxygenase acquired by horizontal transfer relates to the genes genomes of many bacteria. (RubisCO) in the non-sulphur purple photosynthetic encoding ribulose-1,5-bisphosphate carboxylase/oxygenase acquired by horizontal transfer relates to the genes A well-documented example in which an LTTR has been acquired by horizontal transfer to be identified within the genomes of many bacteria.

Comparisons of the form I complex between Rhodobacter spp. have two forms of RubisCO (form I and form II). Comparisons of the form I complex between Rhodobacter capsulatus and R. sphaeroides provided evidence that they were more divergent than previously anticipated (Paoli et al., 1998; Horken & Tabita, 1999). Phylogenetic studies confirmed that the form I complex of R. capsulatus was more closely related to the 'green-type' RubisCO group, associated with x/β/y chemooautotrophic proteobacteria, and green algae, than the 'red-type' found in x/β bacteria and the plastids of red algae. The genes encoding the form I RubisCO complex are operonic (cbbLSQ) and have a divergently transcribed LTTR (CbbR) that activates transcription in response to light intensity and CO₂ concentration (Gibson & Tabita, 1993). Molecular analysis of R. capsulatus indicated that CbbR had been acquired by horizontal transfer with the cbb operon. Form II was also shown to have its own endogenous CbbR divergently transcribed from the cbbM gene. The regulators are currently referred to as CbbR₁ and CbbR₂ (Paoli et al., 1998). A classical LTTR box (TTA-N₇/₈-TAA) is found upstream of both cbbLSQ and cbbM. R. sphaeroides has only one CbbR, which globally regulates both form I and form II (Smith & Tabita, 2002; Dubbs & Tabita, 2003; Dubbs et al., 2004). The additional level of regulation conferred on form I of R. capsulatus allows independent regulation of the two operons; the advantage of this, and whether each regulator responds to a different environmental signal, remains to be elucidated.

This is not the only example of the co-acquisition of LTTRs and their associated genes. Numerous LTTR-regulated virulence factors and antibiotic-resistance factors have been identified as having been acquired by horizontal transfer. These include SpvR of Salmonella spp., which regulates a four-gene operon that is carried on a 90 kbp virulence plasmid. The products of the spv operon have a role in bacterial dissemination from the Peyer’s patches to the liver and spleen (Caldwell & Gulig, 1991; Sheehan & Dorman, 1998). Additionally, the acquisition of antibiotic resistance in Pseudomonas aeruginosa is reliant upon the expression of a metallo-β-lactamase, which is regulated by a divergently transcribed LTTR (Toleman et al., 2002). These LTTRs specifically regulate the genes they are transferred with and do not tend to be global transcriptional regulators.

Structure and function of LTTRs

The role of the C-terminal co-inducer-binding domain

Studies of amino acid composition and secondary structure have helped to identify many LTTRs; residues 20–80 are the most highly conserved and are directly involved with DNA interaction at the major groove. Conversely, there is relatively little conservation at the amino acid level for the C terminus of LTTRs. This region comprises two distinct x/β subdomains (RD1 and RD2) which are connected by two cross-over regions that form a hinge or cleft, which is
likely to accommodate the co-inducer (Stec et al., 2006). Mutagenesis studies have identified a region that can span between residues 95 and 210 of the C-terminal domain, which forms a co-inducer-binding cleft (Burn et al., 1989; Cebolla et al., 1997; Jørgensen & Dandanell, 1999). This hinge-region/cleft appears to be present in the C-terminal region of all LTTRs that have been studied on a structural basis. The co-inducer-binding domain is joined to the HTH by another hinge region.

A conformational change to the tertiary structure upon co-inducer binding has been related to the differential binding ability of LTTRs. Mutagenesis studies similar to those that identified the co-inducer-binding cleft identified an approximately 70–80 amino acid region in the C-terminal domain that also plays a role in DNA binding. Amino acid substitutions between residues 225 and 290 abrogate the co-inducer-dependent state of LTTRs and have an effect on the binding capability of the protein. Studies undertaken with NahR established the co-inducer-binding domain to be in the region of 268 amino acids, with a DNA–protein interaction site at residue 169 (Huang & Schell, 1991; Schell et al., 1990; McIver et al., 1989; Storz et al., 1990; Bartowsky & Normark, 1991; Renna et al., 1993; Colyer & Kredich, 1994, 1996). This co-inducer-independent state mimics LTTR-DNA binding in the absence of a co-inducer and effects the transcriptional activator/repressor properties of the LTTR.

Several crystal structures of LTTRs have been resolved; these have focused primarily on the C-terminal domain, the HTH domain being particularly difficult to crystallize due to the high degree of flexibility found in the ‘wing’ region. The first crystal structures of the C-terminal domain were resolved with the co-inducer or a substitute at the co-inducer-binding cleft, and have highlighted a likeness to the interdomain fold and cleft found in the LacI repressor family (Muraoka et al., 2003a, b).

The co-factor binding domain has been well defined for both CatM and BenM, which are paralogous LTTRs found in Acinetobacter spp. CatM was initially identified as a repressor of the catBCIJFD operon encoding proteins required to convert benzoate into tricarboxylic acid cycle intermediates (Romero-Arroyo et al., 1995; Collier et al., 1998; Clark et al., 2003). It was later reclassified as a transcriptional activator and found to be part of a more complex regulatory network involving BenM. Both CatM and BenM activate the transcription of catBCIJFD but BenM additionally regulates expression of the benABCDE operon, which encodes proteins necessary for benzoate degradation (Collier et al., 1998; Ezezika et al., 2006, 2007).

A larger subgroup of LTTRs that are associated with degradation of catechols and chlorinated aromatics has emerged in which CatM and BenM are classified. The degradation of catechols results in the production of cis,cis-muconate, which is also the co-inducer for CatM and BenM. Structural studies identified the binding site for the

![Fig. 4. An amino acid sequence alignment of the N-terminal HTH-containing region (1 to between 60 and 85 amino acids) of LTTRs found in a group of bacteria, representative of the α, β, and γ subdivisions of proteobacteria and Gram-positive bacteria (constructed using vector NTI). The degree of conservation is indicated (green, high; yellow, moderate; red, weak). The secondary structure is shown above the alignment (cylinders, α-helices; arrows, β-strands).]
co-inducer lying between RD1 and RD2 of both LTTRs. RD1 and RD2 were shown to be connected by two hinge-like, antiparallel β-strands which provide flexibility to the protein, enabling the two domains to rotate relative to each other. The remaining C-terminal region was found to consist of α and β structures (nine α-helices and nine β-strands) with Rossmann-like folds (Neidle et al., 1989). BenM has a unique feature compared to other LTTRs in that it can bind to two different co-inducers. Benzoate binds a second region in BenM that is not present in CatM. This secondary site is located in a highly hydrophobic region of RD1 and alters the conformation of BenM once the co-inducer is bound. The altered conformation still enables cis,cis-muconate to bind at the primary binding site and produces a synergistic effect resulting in very high levels of transcriptional regulation. It is thought that occupation of the secondary site alters the salt bridges formed between glutamate residues and arginine residues within the primary binding site, producing an altered protein conformation but not affecting the capacity to bind cis,cis-muconate.

The C-terminal domain of a number of other LTTRs appears to be similar in structure to those of CatM and BenM. These include DntR isolated from Burkholderia spp., which regulates the expression of enzymes that are involved in catalysing the initial steps of the oxidative degradation of 2,4-dinitrotoluene (2,4-DNT) (Lönneberg et al., 2007). The crystal structure has been resolved to 2.6 Å (acetate at the co-inducer site) and 2.3 Å (thiocyanate at the co-inducer site) (Smirnova et al., 2004). These are not the physiological co-inducers for DntR, and studies have indicated that both sodium salicylate and 2,4-DNT are more likely to be the true co-inducers. The C-terminal region includes RD1 and RD2 domains joined by hinge regions found at residues 167–170 and 270–273. The co-inducer-binding cleft resides between the RD1 and RD2 regions, with a depth of 10 Å and diameter of 7 Å.

Full-length crystal structures of LTTRs and determination of the winged-HTH (wHTH) domain

The first full-length LTTR crystal structure to be resolved was CbnR. It is divergently transcribed from the cbnABCD operon, the products of which are involved in the degradation of chlorocatechols (Ogawa et al., 1999; Muraoaka et al., 2003a, b). CbnR was crystallized as a tetramer consisting of two dimers. Each dimer comprises one short-form subunit and one extended-form subunit, giving the tetrameric molecule an asymmetrical ellipsoidal shape (130 Å × 70 Å × 60 Å). Each subunit has two domains, a DBD (residues 1–58) and a regulatory domain (residues 88–294) joined by a linker region (residues 59–87). The subunits dimerize through an anti-parallel helix–helix interaction, and the dimers interact along a twofold axis. The resulting ellipsoid has a cavity of 30 Å × 15 Å × 10 Å that accommodates the co-inducer. The DBD lies in a V-shape at the base of the tetramer and consists of three α-helices and two β-strands that form a winged-HTH (Muraoka et al., 2003a, b). This conformation is very closely related to ModE of E. coli and has been used to model numerous HTH regions, including that of OxyR.

OxyR was first identified as a member of the LTTR family by Christman et al. (1989). The full-length structure of OxyR has been determined using ModE as a model and has provided information about the wHTH region, which appears to be similar to that of the iron-responsive regulator, DtxR (Zaim & Kierzek, 2003). It is located in the N-terminal domain, as is the case for all other LTTRs, and is attached to a long α-helical backbone. The recognition helix of the HTH appears longer than that of other HTH regions and has been described as a helix–loop–helix with the ‘winged’ portion likely to interact with the phosphate backbone or minor groove of the double helix. OxyR does not respond to a classical co-inducer, but relies upon a redox change to alter its conformation and DNA affinity. Specifically, it senses H2O2 and is activated through the formation of a transient disulphide bond. The presence or absence of the disulphide bond affects the oligomerization state of OxyR. In the reduced form OxyR appears to be dimeric, only occupying two DNA-binding regions; when oxidized it binds to four regions, corresponding to a tetrameric structure (Kullik et al., 1995a, b). The oxidation of OxyR also influences co-operative interaction with RNA polymerase at the promoter region, thus initiating transcription (Zaim & Kierzek, 2003). Determining the structure of the regulatory domain revealed two domains (corresponding to RD1 and RD2) in which the redox-active cysteines are found.

Attempts have been made to crystallize other LTTRs, but the insolubility associated with the wHTH domain has meant that often only truncated forms can be resolved. CysB and CblR are closely related LTTRs found in Klebsiella spp., E. coli and Pseudomonas spp. as regulators of sulphate starvation inducible genes (Delic-Attree et al., 1997; Verschuuren et al., 2001; Jovanovic et al., 2003). CblR is part of the CysB regulon and exhibits 41% identity at the amino acid level. Despite there being no structural data concerning the DBDs of these proteins both have been shown to be tetrameric and ellipsoidal in shape (van der Ploeg et al., 1997; Lochowska et al., 2004). Their tetrameric nature allows these proteins to span a large region of DNA, causing it to bend (Hryniewicz & Kedrick, 1994). This feature is apparent in numerous LTTRs and is believed to be a result of the V-shape in which the DBD resides.

Despite limited structural information for the HTH region of LTTRs, mutational studies have highlighted a number of important, conserved residues, which appear to be required for DNA binding. If these residues are altered, the LTTR in question loses its ability to bind DNA. For OxyR the mutations that cause this phenotype are T31M and S33N; for CysB, S34R; for GcvA, S38P and for NahR, R43H (Kullik et al., 1995a, b; Lochowska et al., 2004). These
observations are assisting in the prediction of complete structural data.

**LTTR transcriptional regulation**

**The LTTR box**

There are multiple binding sites within the intergenic region between an LTTR and its associated gene/operon (or upstream of distant LTTR-regulated genes). Broadly they bind at −35 to +20 bp (regulatory binding site, RBS, and autoregulatory site) and −40 to −20 bp (activation binding site, ABS) with hypersensitivity (associated with DNA bending) at −55 bp (Belitsky *et al.*, 1995; Lochowska *et al.*, 2001; Porrua *et al.*, 2007). However, binding sites as far away as −218 bp with respect to the promoter region as well as internal binding sites (+350 bp; IBS), have been identified (Wilson *et al.*, 1995; Viswanathan *et al.*, 2007). These distinct binding sites were identified by virtue of DNA-footprinting, DNase I-protection studies and mutagenesis.

A palindromic DNA sequence has been identified to which LTTRs are known to bind; this is often found to form part of an imperfect, dyadic region. The LTTR box was identified first in *Rhizobium* spp. as an interrupted palindrome with the sequence ATC-N₁₁-GAT, between −20 and −75 bp upstream of the *nod* gene, and was referred to as the ‘Nod-box’ (Goethals *et al.*, 1992). From this the generally accepted LTTR box was identified. It consists of the sequence T-N₁₁-A, but can vary in both base pair composition and length; it is present at the RBS but not the ABS site (Parsek *et al.*, 1994b). The apo-form and co-inducer-bound LTTR differ in their affinity for the LTTR box, which may affect preferential binding at the RBS or ABS sites; this can result in DNA bending and can affect the interaction with RNA polymerase at the promoter region as described below.

**DNA bending**

LTTRs are known to be functionally active as tetramers, and as such have been shown to ‘protect’ large regions of DNA (between 50 and 60 bp) by DNase I protection assay (Muraoka *et al.*, 2003a, b). This large region of protection is consistent with the observation that LTTRs bind at multiple locations in the promoter region. The affinity of LTTR for each distinct binding region is determined by the co-inducer; the apo-form of the protein will often only bind to RBSs, the ABS sites only being occupied once the co-inducer is bound to the protein (Fig. 5) (Tropel & van der Meer, 2004).

---

**Fig. 5.** A schematic representation of the role of DNA bending and LTTR binding at the ABS and RBS sites in LTTR-dependent transcriptional activation. (1) Shows one LTTR dimer bound at the RBS and a second dimer bound at the ABS. (2) Shows protein–protein interaction between the two LTTR dimers leading to oligomerization to form a tetrameric protein. The protein interactions cause the DNA to bend. (3) Shows the binding of RNA polymerase at the promoter region of the target gene, but no transcriptional activation by the LTTR tetramer in the absence of a co-inducer binding to it. (4) Shows the co-inducer binding to the LTTR tetramer and the DNA bend relaxing; the LTTR tetramer is consequently brought into contact with the RNA polymerase at the promoter site of the target gene, and this activates transcription of the target gene.
DNA bending in itself is an important factor in terms of protein interaction, and in the case of LTTRs two dimeric proteins located at the ABS and RBS come into contact to form a tetrameric, active structure as a direct consequence of DNA bending; this allows the formation of a higher-order complex involving RNA polymerase, thus initiating transcription. The position of the LTTR-binding region and the extent of DNA bending appear to have no correlation to whether the LTTR is behaving as a transcriptional activator or repressor.

The pattern of binding of LTTRs to the 15–17 bp palidromic RBS region is in keeping with the type of binding observed for dimeric proteins. The tetrameric nature of active LTTRs and their ability to protect large regions of DNA implies that DNA bending is an important contributor to LTTR-dependent transcriptional regulation and oligomerization. Transcriptional regulators that rely upon DNA bending to provide additional levels of regulation frequently enhance transcription by increasing the likelihood of regulator–RNA polymerase interactions. Some well-documented examples include Fis, H-NS, HU and IHF; regions where DNA bending occurs most readily are often A/T rich, a common motif being recognized as CA\textsubscript{56}T (Goosen & van de Putte, 1995; Martin & Rosner, 1997; Pérez-Martín & de Lorenzo, 1997; Huo et al., 2006).

The studies of several LTTRs have indicated that they can cause DNA to bend between 50° and 100° and that the degree of DNA bending is determined by the presence or absence of the co-inducer. Generally the presence of a co-inducer bound to an LTTR relaxes the degree of DNA bending from as little as 9° to as much as 50° (van Keulen et al., 1998). The relaxation of DNA bending correlates with a shift in DNA protection to encompass a smaller area and appears to be paramount for transcriptional activation or repression (Fig. 5).

The role of DNA bending in LTTR-mediated transcriptional regulation is well illustrated by studies undertaken with OccR. This regulator is found on the Ti plasmid of Agrobacterium tumefaciens and regulates genes required for octopine catabolism. DNA-footprinting has shown that when bound to DNA, OccR occupies a region spanning between −80 and −28 bp upstream of the transcriptional start site of the genes it regulates. When the co-inducer octopine is present, OccR spans a shorter region of between −80 and −38 bp (Akakura & Winans, 2002a, b). This coincides with a change in DNA bending from a high-angle bend (in the absence of co-inducer) to a low-angle bend (co-inducer dependent). DNA bending is a common feature of prokaryotic transcriptional regulation and is dependent upon the multimeric nature of transcriptional regulators. Numerous LTTRs have been shown to induce DNA bending; often the intergenic region to which an LTTR binds possesses a hypersensitive region at approximately 50–55 bp upstream of the regulated gene and is the point at which a DNA bend is induced (Hryniewicz & Kredich, 1994; Ogawa et al., 1999). Many LTTR-binding sites may be employed: for OccR five have been identified, two of which have a characteristic LTTR-box and lie on the same face of the DNA helix.

**LTTR autoregulation**

Much research has focused on the regulation of genes constituting a given LTTR regulon; less has been undertaken to elucidate the autoregulatory function. The RBS region of genes divergently transcribed from their LTTR has been implicated as a possible autoregulatory site. The RBS characteristically contains an LTTR-box, suggesting that this recognition sequence is necessary for autoregulation. DNA-footprinting assays imply that LTTR binding at the RBS is consistent with the pattern often observed for dimeric proteins and that the apo-form of LTTR binds the region with a greater affinity than the co-inducer-bound tetrameric form. Taken together these data suggest that as an autoregulator LTTR might function as a dimer in a co-inducer-independent manner. The tetrameric form might be necessary only for transcriptional activation of divergent genes, and dependent upon the co-inducer. The presence or absence of a co-inducer might influence the multimeric state of the LTTR, which undergoes a conformational change when co-inducer is inserted into the binding cleft. The formation of a tetramer may be dependent upon this conformational change; however, no evidence exists to date to support this assertion. Where genes are not divergently transcribed the mechanism of autoregulation is even less clear. No single LTTR of this nature has been extensively studied at the genetic level to determine whether the LTTR-box is present in the upstream intergenic region, and no footprinting assays have been undertaken to ascertain whether LTTR proteins bind upstream of any given LTTR gene.

**‘Classical’ LTTR regulation – transcriptional activation and negative autoregulation**

IlvY is regarded as a prototypical LTTR protein-regulated system, and has been best studied in E. coli and Salmonella spp. (Blazey & Burns, 1980; Rhee et al., 1999). It forms part of a two-gene operon with ilvC; the two genes are transcribed from overlapping divergent promoters. IlvY exhibits classical LTTR-like regulation whereby expression of ilvY is negatively autoregulated by IlvY, and expression of ilvC is activated by IlvY. IlvC is an acetyl-CoA-acetyl-CoA isomerase, and the second enzyme of the parallel biosynthetic pathway for L-valine and L-isoleucine (Biel & Umbarger, 1981). The substrates for acetyl-CoA-acetyl-CoA isomerase are α-acetolactate and α-acetoxybutyric acid; both of these substrates are the co-inducers for IlvY and are necessary for transcriptional activation of ilvC, thus forming a feedback loop (Blazey & Burns, 1980). This type of regulation coupled with a feedback loop is commonly observed for LTTRs and associated LTTR-regulated genes. Classical regulation is also exhibited by LTTRs of Gram-positive bacteria. CidR has been characterized in...
Staphylococcus spp. and also identified in Bacillus anthracis (Yang et al., 2005; Ahn et al., 2006). It is divergently transcribed from cidABC and transcriptional activation is dependent upon acetate acid produced from the metabolism of glucose, as a co-inducer (Yang et al., 2005). Negative autoregulation of CidR is believed to be co-inducer independent.

**LTTR transcriptional repressors**

CcpC is a member of a novel subgroup of LTTRs that act as transcriptional repressors. It has been identified in Bacillus subtilis and is regarded as a member of the LTTR family by virtue of extensive amino acid sequence similarity; it has been shown to interact with regions of DNA possessing an LTTR-box (Jourlin-Castelli et al., 2000). It is not divergently transcribed and is a global negative regulator of the genes encoding enzymes involved in the tricarboxylic acid cycle. Two well-studied CcpC regulated genes are citB (aconitase) and citZ (citrate synthase). CcpC binds at the −66 and −27 regions to repress transcription, and in the presence of citrate as a co-inducer is seen to derepress expression of citB and citZ (Kim et al., 2002, 2003). CcpC-binding regions have been identified upstream of ccpC and are referred to as Box I and Box II. Negative autoregulation appears to depend solely upon binding to Box I, with no defined role for Box II.

Numerous other transcriptional repressors of the LTTR family exist and it is becoming apparent that the distinction between transcriptional activator and transcriptional repressor is more complex than previously thought. GltC has been regarded as a classical LTTR in B. subtilis and is regarded as a member of the LTTR family by virtue of extensive amino acid sequence similarity; it has been shown to interact with regions of DNA possessing an LTTR-box (Jourlin-Castelli et al., 2000). It is not divergently transcribed and is a global negative regulator of the genes encoding enzymes involved in the tricarboxylic acid cycle. Two well-studied CcpC regulated genes are citB (aconitase) and citZ (citrate synthase). CcpC binds at the −66 and −27 regions to repress transcription, and in the presence of citrate as a co-inducer is seen to derepress expression of citB and citZ (Kim et al., 2002, 2003). CcpC-binding regions have been identified upstream of ccpC and are referred to as Box I and Box II. Negative autoregulation appears to depend solely upon binding to Box I, with no defined role for Box II.

Positive autoregulation

An additional class of LTTRs that act as transcriptional repressors or activators have been identified that positively autoregulate. LrhA is an example of this type of LTTR and was first identified in E. coli. Other members of this subgroup include HexA and PecT of Erwinia spp. (Gibson & Silhavy, 1999). These three regulators control the expression of genes required for flagellation, motility and chemotaxis. They negatively regulate the expression of other transcriptional regulators that are involved in a global, complex regulatory network. The environmental stimuli for this group of regulators remains undefined and currently no co-factor has been identified. Despite the method of transcriptional regulation appearing to be quite different from that of the classical LTTRs, amino acid sequence identities place this subgroup within the LTTR family.

YtxR is another positive autoregulatory LTTR. It activates the expression of ytxAB, which is an operon found in ‘American’ strains of Yersinia enterocolitica and encodes heat-labile, ADP-ribosylating toxin. The precise regulatory mechanism of YtxR remains to be elucidated; preliminary analysis has suggested that it is a co-inducer-independent global regulator (Axler-Diperte et al., 2006).

A role for LTTRs in therapeutics, diagnostics and vaccine development

**Vaccines**

Global transcriptional regulators are fundamental tools for the study of virulence, disease progression, bacterial growth and metabolism. The transcriptomic approach applied to the study of global regulators has wide applications for vaccine development, diagnostics and therapeutics. The importance of global transcriptional regulators in the development of attenuated and protective vaccine strains is highlighted by comparative genomics of Mycobacterium bovis and the derivative BCG vaccine. BCG has been used since 1921; extensive passage under different laboratory conditions led to the emergence of numerous daughter strains with differing efficacy (Keller et al., 2008; Ritz et al., 2008). The efficacy of the modern ‘BCG vaccine’, encompassing data for all of the different daughter strains, is thought to be between 0 and 80% (Chen et al., 2007).
Large polymorphisms exist in the genome of BCG strains; in particular there have been four major deletion events and two duplications (Leung et al., 2008). Two of the better-documented deletions are of region of difference 1 (RD1) and RD2. The loss of RD1 appears to have been a critical event for attenuation of *M. bovis* given that all of the vaccine strains have this deletion (Ritz et al., 2008). Other mutations outside of RD1 include changes in the expression of certain transcriptional regulators. A point mutation in the CRP-FNR-like regulator (M₃7₀₀) alters global gene expression and attenuates the strain. This mutation has been traced back to post-1924 and affects the HTH domain of the regulatory protein (Spreadbury et al., 2005). This mutation is likely to be present in many of the daughter derivative strains that were passaged independently by different laboratories prior to the development of a seed-stock in the 1960s.

The post-1927 deletion of RD2 also led to the loss of the two-component sensor–regulator system PhoPQ which has a significant impact on attenuation. PhoPQ is also found in *S. enterica* serovar Typhimurium, where it is part of a complex regulatory network. Null mutations in either *phoP* or *phoQ* result in strains that are attenuated in mice and humans, unable to survive in macrophages and show sensitivity to cationic antimicrobial peptides. Despite this attenuation *phoP* mutants also exhibit increased antigen presentation compared to the wild-type as a consequence of an inability to modify the LPS (Gunn & Miller, 1996; Groisman et al., 1997). This is the result of a complex regulatory network involving additional sensor–regulator proteins PmrAB, which require PhoP. Similar attenuation is observed for *M. bovis*, in which the RD2 region is intact but the *phoPQ* genes have been disrupted. Preliminary studies using guinea pigs have shown superior efficacy compared to the current BCG strains.

Transcriptomic studies involving LTTRs could help to identify bacterial strains that are attenuated in a defined genetic background with the potential for use as a vaccine candidate. Due to the diverse variety of bacterial species in which the LTTRs are found, identification of an LTTR that is attenuated and protective in a given bacterium could have wider applications for a number of other bacteria with orthologous LTTRs.

**Therapeutics and diagnostics**

Studies regarding the role of global regulators for therapeutic applications can be illustrated by the extensive investigations of the MarR family of global transcriptional regulators. These are wHTH-containing transcriptional regulators found throughout the prokaryotes and archaea. They regulate genes responsible for multidrug resistance in numerous bacteria, including *E. coli*, *P. aeruginosa* and *B. subtilis*. If it is possible to identify MarR-like transcriptional regulators within a distinct species of bacteria they could serve as indicators of multidrug-resistant strains, with applications for rapid screening of clinical isolates (Wilkinson & Grove, 2006). Some LTTRs, including CysB and AmpR, regulate genes that confer resistance to novobiocin and β-lactamases, respectively, and could similarly be used as markers for antibiotic-resistant strains.

Studies in which research has focused upon the regulon of a given transcriptional regulator in terms of identifying therapeutic or diagnostic candidates include analysis of the transcriptome of *Streptococcus pyogenes* during mouse soft tissue infection and in cynomolgus macaques presenting with *Strep. pyogenes* pharyngitis (Virtaneva et al., 2005; Graham et al., 2006). This has provided insights into clinical disease progression and has identified numerous transcriptional regulators that are affiliated with the progressing stages of GAS infection. Global comparisons with a *Strep. pyogenes* in which the transcriptional regulator CovR is knocked out have identified many putative virulence factors and other transcriptional regulators controlled by CovR. This has helped to ascertain whether any of the CovR-regulated genes might provide vaccine or diagnostic candidates. Given that the LTTRs are the most abundant regulator found within the genomes of bacteria (http://www.era7.com/ExtraTrain), focusing upon identifying important members of the regulons that might contribute to virulence and pathogenesis seems paramount, and could provide new insights into disease progression and vaccine development.

**Concluding remarks**

The LysR family encompasses a huge number of transcriptional regulators, many more than the nine originally identified by Henikoff et al. (1988). As more research into prokaryotic transcriptional regulation continues to be undertaken, there is no doubt that more LTTRs will be identified and this family will continue to grow in size. The emergence of new subgroups of the LTTR family is likely and will shed light upon the evolutionary divergence of LTTRs away from the ancestral prototypical transcriptional activator. Progress in terms of the understanding of DNA binding by LTTRs is rapid and there is much that can be built upon the information that has been gathered so far. More complete structural data combined with a broader understanding of the role of co-inducers in transcriptional regulation will help to elucidate an increasingly complex paradigm for LTTR-dependent regulation.

**References**


