ELSEVIER

FEMS Microbiology Reviews 27 (2003) 131–143

FEMS MICROBIOLOGY Reviews

www.fems-microbiology.org

The SmtB/ArsR family of metalloregulatory transcriptional repressors: structural insights into prokaryotic metal resistance

Laura S. Busenlehner, Mario A. Pennella, David P. Giedroc *

Department of Biochemistry and Biophysics, Center for Advanced Biomolecular Research, 2128 TAMU, Texas A&M University, College Station, TX 77843-2128, USA

Received 9 October 2002; received in revised form 8 January 2003; accepted 29 January 2003

First published online 6 May 2003

Abstract

The SmtB/ArsR family of prokaryotic metalloregulatory transcriptional repressors represses the expression of operons linked to stress-inducing concentrations of di- and multivalent heavy metal ions. Derepression results from direct binding of metal ions by these homodimeric 'metal sensor' proteins. An evolutionary analysis, coupled with comparative structural and spectroscopic studies of six SmtB/ArsR family members, suggests a unifying 'theme and variations' model, in which individual members have evolved distinct metal selectivity profiles by alteration of one or both of two structurally distinct metal coordination sites. These two metal sites are designated α3N (or α3) and α5 (or α5C), named for the location of the metal binding ligands within the known or predicted secondary structure of individual family members. The α3N/α3 sensors, represented by *Staphylococcus aureus* pI258 CadC, *Listeria monocytogenes* CadC and *Escherichia coli* ArsR, form cysteine thiolate-rich coordination complexes (S_3 or S_4) with thiophilic heavy metal pollutants including Cd(II), Pb(II), Bi(III) and As(III) via inter-subunit coordination by ligands derived from the α3 helix and the N-terminal 'arm' (CadCs) or from the α3 helix only (ArsRs). The α5/α5C sensors *Synechococcus* SmtB, *Synechocystis* ZiaR, *S. aureus* CzrA, and *Mycobacterium tuberculosis* NmtR form metal complexes with biologically required metal ions Zn(II), Co(II) and Ni(II) characterized by four or more coordination bonds to a mixture of histidine and carboxylate ligands derived from the C-terminal α5 helices on opposite subunits. Direct binding of metal ions to either the α3N or α5 sites leads to strong, negative allosteric regulation of repressor operator/promoter binding affinity, consistent with a simple model for derepression. We hypothesize that distinct allosteric pathways for metal sensing have co-evolved with metal specificities of distinct α3N and α5 coordination complexes.

Keywords: Metalloregulation; Heavy metal resistance; SmtB; CadC; CzrA

Contents

* Corresponding author. Tel.: +1 (979) 845-4231; Fax: +1 (979) 845-4946. *E-mail address:* giedroc@tamu.edu (D.P. Giedroc).

0168-6445/03/$22.00
doi:10.1016/S0168-6445(03)00054-8

1. Introduction

Numerous cellular processes require metal ions as cofactors for enzymatic reactions or as structural components of proteins [1,2]. Prokaryotic cells actively maintain a beneficial intracellular concentration of these essential metal ions by delicately balancing the expression of proteins involved in specific metal uptake and export/storage [3,4]. However, even essential metal ions can be detrimental to cell viability if the intracellular, bioavailable concentration is in excess of normal, physiological levels [5]. In addition, some metal ions are toxic to bacterial cells at all concentrations, therefore detoxification and resistance systems that employ a variety of mechanisms to rid the cell of these potentially lethal toxins have evolved [6–8]. In most cases, the expression of such resistance systems is controlled at the level of transcription by metal sensor proteins that 'sense' specific metal ions via their direct coordination. It is the direct binding of inducing metal ions that allows for a change in the regulatory function of the metal sensor protein [6,9].

Resistance operons to heavy metal ions have been stably integrated into some bacterial chromosomes, but they can also be located on endogenous plasmids or transposons that often contain multiple resistance or detoxification operons [10–12]. It is interesting to note that some prokaryotes possess multiple, and even redundant, resistance systems for metal ions whether present on the chromosome or extrachromosomal DNA [8,13]. Usually, these operons encode resistance proteins such as metal-specific efflux pumps, membrane-bound transporters, metal reductases, cytoplasmic or periplasmic metal transport proteins, or

A

```
E. coli R773 ArsR                ----------------------------------------MLQLTPLQLFKNLSDETRLGIVLLLREMGELCVCD 35
E. coli R46 ArsR                 ----------------------------------------MPEIASLQLFKILSDETRLGIVLLLREMGELCVCD 35
E. coli ArsR                     ----------------------------------------MSFLLPIQLFKILADETRLGIVLLLSELGELCVCD 35
S. aureus pI258 ArsR             ----------------------------------------MSYKELSTILKILSDSSRLEILDLL-SCGELCACD 34
S. xylosis pSX267 ArsR           ----------------------------------------MSYKELSTILKVLSDPSRLEILDLL-SCGELCACD 34
B. subtilis ArsR                 ---------------------------------MDETKSELLRKYEQKFKALADQKRLEIMYELCQRGKTCVCD 41
S. aureus pI258 CadC             MKKK--------------DTCEIFCYDEEKVNRIQGDLQTV-DISGVSQILKAIADENRAKITYALCQDEELCVCD 41
S. maltophilia CadC              MKKK--------------DTCEIFCYDEEKVNRIQGDLQTV-DISGVSQMLKAIADENRAKITYALCQDEELCVCD 41
B. firmus CadC                   MNKK--------------DTCEIFCYDEEKVNRIQGDLKTI-DIVSVAQMLKAIADENRAKITYALCQDEESCVCD 41
B. halodurans CadC               MDEL--------------DTCDVFLYDKEKVQLVSEAMKGE-NFAKVADMFKALADETRVKIAYALLQQEELCVCD 41
G. stearothermophilus CadC       MAN---------------DVCEVFSINAEKVDAVRQRIN---EISGVEQLFKALADATRLKITYALMLEKELCVCD 58
L. monocytogenes CadC            MTV---------------DICEITCIDEEKVKRVKTGLETV-EVTTISQIFKILSDETRVKIVYALLTENELCVCD 60
L. innocua CadC                  MNN---------------EICEITCIHEDKVNRAKSKLANF-DTPSVSGFFKILSDENRLKIVHALVHEDELCVCD 60
S. aureus CadC                   MTK---------------DMCEVTYIHEDKVNRAKKDLAKQ-NPMDVAKVFKALSDDTRVKIAYVLSLEGELCVCD 60
Synechocystis sp. ZiaR           MSKSSLSKSQS-CQNEEMPLCDQPLVHLEQVRQVQPEVMSLDQAQQMAEFFSALADPSRLRLMSALAR-QELCVCD 74
Nostoc sp. ZiaR                  MNKHKKKQDLDLIQSSDTPTCDTHLVHLDNVRSSQAQILPTDKAQQMAEIFGVLADTNRIRLLSALAS-SELCVCD 75
Synechococcus sp. SmtB           MTKPVL-------QDGETVVCQGT--HAAIASELQAIAPEVAQSL--AEFFAVLADPNRLRLLSLLAR-SELCVGD 64
Nostoc sp. SmtB                  MIKN-------------HTNCTVEEINVSSITPLINEVLSTEKSQRMAEFFSFLGDANRLRILSFLAT-KELCVSD 62
D. radiodurans SmtB              MTVLDSK------PTEATDVCEINCVHPEAVAQARAALPGEAHTAAATTFFKVLGDTTRIKILYALFE-SELCVCD 69
T. tengcongensis SmtB            MKREDSK----------IPICEVSVIHEDIINKVKEAMPEETKLYDLAELFKALSDPTRVQILNLLQE-REMCVCD 65
L. innocua CzrA                  ----------------------------MENAKQTLDEETLFSVTQIFKALSDPTRVQILNLLQD-REYSVNE 44
L. monocytogenes CzrA            ----------------------------MENAKQPLDEETLFSVTQIFKALSDPTRIRILHLLSQ-GEYSVNE 44
B. subtilis CzrA                 -------------------------MTEFRETEQSAADLDEETLFLVAQTFKALSDPTRIRILHLLSQ-GEHAVNG 50
S. aureus CzrA                   ----------------------------MSEQYSEINTDTLERVTEIFKALGDYNRIRIMELLSV-SEASVGH 44
M. tuberculosis NmtR             ------------------------MGHGVEGRNRPSAPLDSQAAAQVASTLQALATPSRLMILTQLRN-GPLPVTD 51
                                                                                  *      *

E. coli R773 ArsR                LCMALDQSQPKISRHLAMLRESGILLDRKQGKWVHYRLS-PHIPSWAAQIIEQAWLSQQDDVQVIARKLASVNCSGSSKAVCI 117
E. coli R46 ArsR                 LCTALEQSQPKTSRHLAMLRESGLLLDRKQGKWVHYRLS-PHIPSWAALVIEQAWLSQQDDVQAIARKLASANCSGSGKAVCI 117
E. coli ArsR                     LCTALDQSQPKISRHLALLRESGLLLDRKQGKWVHYRLS-PHIPAWAAKIIDEAWRCEQEKVQAIVRNLARQNCSGDSK---- 113
S. aureus pI258 ArsR             LLEHFQFSQPTLSHHMKSLVDNELVTTRKDGNKHWYQLN-HAILDDIIQNLNIIINTSNQ---RCVCKNVKSGDC--------- 104
S. xylosis pSX267 ArsR           LLEHFQFSQPTLSHHMKSLVDNELVTTRKNGNKHMYQLN-HEFLDYINQNLDIINTSDQ---GCACKNMKSGEC--------- 104
B. subtilis ArsR                 LTEIFEVTQSKLSYHLKILLDANLITKETKGTWSYYDLNDEEVNGLLSEELCCIFRKKGEGD-CC------------------ 105
S. aureus pI258 CadC             IANILGVTIANASHHLRTLYKQGVVNFRKEGKLALYSLGDEHIRQIMMIALAHKKEVKVNV---------------------- 122
S. maltophilia CadC              IANILGVTIANASHHLRMLYKQGVVNFRKEGKLALYSLDDEHIGQIMMIALAHKKEVKVNV---------------------- 122
B. firmus CadC                   IANIIGITANASHHLRTLHKQGIVRYRKEGKLAFYSLDDEHIRQIMMIVLEHKKEVNVNV---------------------- 122
B. halodurans CadC               VAHIIGATTATASHHLRHLRNLGIAKRRKEGKLVFYSLDDDHVRTILSMAVVHGKEGKGNG---------------------- 122
G. stearothermophilus CadC       VATIIGCTVATASHHLRLLRDMGIAKHRKEGKLVFYSLKDDHVEQLVALALVHAKEGEENDGK---------------------- 121
L. monocytogenes CadC            LANIVEATVAATSHHLRFLKKQGIANYRKDGKLVYYSLANERVRDRIKLILLNFEGVGV---------------------- 119
L. innocua CadC                  IANIIDASVATTSHHLNSLKKLGVVDSHKDGKLVYYFIKNIKILNLMELGVNFKEEVLA---------------------- 119
S. aureus CadC                   VANIIESSTATASHHLRLLKNLGIAKYRKEGKLVYYSLDDEHVKQLVEKAFLHQREVASIG---------------------- 121
Synechocystis sp. ZiaR           LAAAMKVSESAVSHQLRILRSQRLVKYRRVGRNVYYSLADNHVMNLYREVADHLQESD---------------------- 132
Nostoc sp. ZiaR                  LAALTKMSESAVCHQLRLLKAMRLVSYRREGRNVYYSLADSHVINLYRSLVENNTYATGTG---------------------- 136
Synechococcus sp. SmtB           LAQAIGVSESAVSHQLRSLRNLRLVSYRKQGRHVYYQLQDHHIVALYQNALDHLQECR---------------------- 122
Nostoc sp. SmtB                  IATLLEMSESAVSHQLRNLRAMRLVSYRKQGRHVFYRLHDNHILELYQAVAEHLDEKD---------------------- 120
D. radiodurans SmtB              LAAVVGLSESAVSHQLRLLRTGRLVAFRKEGRVVYYRLADEHVTTLIRSALDHARE---------------------- 125
T. tengcongensis SmtB            IAALLGMSQSAVSHQLRILRQNKLVKFRKEGKVVYYSLADDHVKKIFEQGFDHINE---------------------- 121
L. innocua CzrA                  IARTLTFNQTTVSHQLRFLKNLRLVKSRREGTTIYYMQDDNHVLELLTQAIRHVHHH---------------------- 101
L. monocytogenes CzrA            IARTLAFNQTTVSHQLRFLKNLRLVKSRREGTTIYYMQDDNHVLELLTQAIRHVHHH---------------------- 101
B. subtilis CzrA                 IAEKLNLLQSTVSHQLRFLKNLRLVKSRREGTSIYYSPEDEHVLDVLQQMIHHTQHD---------------------- 107
S. aureus CzrA                   ISHQLNLSQSNVSHQLKLLKSVHLVKAKRQGQSMIYSLDDIHVATMLKQAIHHANHPKESGL---------------------- 106
M. tuberculosis NmtR             LAEAIGMEQSAVSHQLRVLRNLGLVVGDRAGRSIVYSLYDTHVAQLLDEAIYHSEHLHLGLSDRHPSAG-------------- 120
                                              *           *    *
```

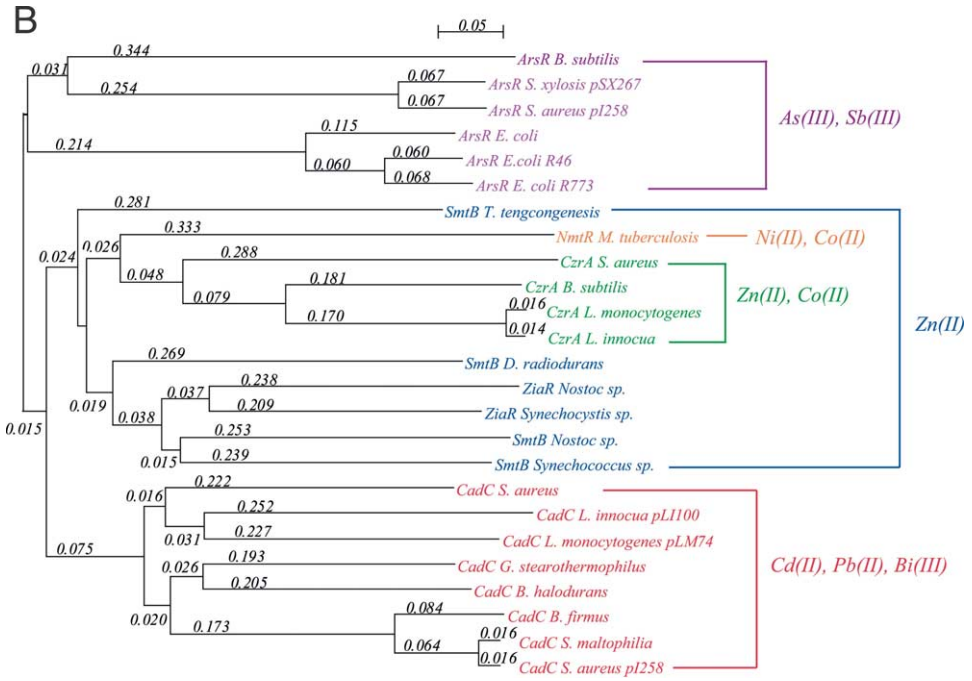

Fig. 1. Amino acid sequence alignment of the SmtB/ArsR family. A: Multiple sequence alignment of 25 SmtB/ArsR metalloregulatory transcriptional repressors generated using ClustalW (http://www.ebi.ac.uk/clustalw) [62]. Residues known or predicted to form the α3N metal site are shaded red and dark blue, while residues known or predicted to be metal ligands in the α5 site are shaded purple and pink (cf. Figs. 5 and 6). Invariant residues are denoted with an asterisk. Residues conserved only among α3(N) sensor proteins are shaded teal while those residues conserved within α5 sensors only are shaded orange (see text for details). B: A phylogenetic tree of the 25 sequences shown in A, created using the neighbor joining method and visualized using NJPlot [63]. The calculated distance between each pair of sequences was used to construct the phylogenetic tree which guides the final multiple sequence alignment. α3N and α5 sensors appear to cluster on separate nodes of the dendrogram and are linked by a common evolutionary ancestor.

metal-sequestering proteins. They also encode at least one *trans*-acting metal-responsive transcriptional regulator for the operon [11,14]. These transcriptional regulators can function as repressors or activators, and both types of sensor proteins have been observed in Gram-negative and Gram-positive bacteria. In some cases, the extent of amino acid sequence identity is so high as to suggest direct horizontal gene transfer from Gram-negative to Gram-positive bacteria [6,15].

The MerR and SmtB/ArsR families represent two general classes of transcriptional regulatory proteins that have endowed prokaryotes with the ability to respond to stress induced by heavy metal toxicity. MerR-like proteins generally function as repressors in the absence of metal ions and become activators upon metal binding, by driving a metal-induced DNA conformational switch that converts a sub-optimal promoter into a potent one [16]. Members of this family include MerR [17], ZntR [18] and CoaR [19], the mercury-, zinc- and cobalt-inducible transcriptional activators in *Escherichia coli* and *Synechocystis*, respectively. The family also includes the more recently described Cu(II)-activated *E. coli* CueR [20,21] and Pb(II) sensor *Ralstonia metallidurans* PbrR [22]. The SmtB/ArsR proteins, perhaps distantly related to the MerR family, function exclusively as transcriptional repressors and include *Synechococcus* sp. SmtB, *E. coli* ArsR, and *Staphylococcus aureus* pI258 CadC. In this case, the resistance operons are repressed when the apo-sensor proteins are specifically bound to the operator/promoter (O/P) DNA; metal binding strongly inhibits or negatively regulates DNA binding and the regulatory protein does not appear to actively participate in transcription (see below) [8,23,24]. This review will summarize recent biochemical and structural studies of the SmtB/ArsR family of metalloregulatory transcriptional repressors.

2. Overview of the SmtB/ArsR family

Individual members of the SmtB/ArsR family clearly derive from a common evolutionary origin, with members that exhibit known or postulated metal selectivities clustering on a rooted, phylogenetic tree (Fig. 1). Members appear to have evolved to sense distinct metal ions of various ionic radii, coordination geometry and valence (Fig. 2). One founding member, *Synechococcus* PCC 7942 SmtB, functions as a Zn(II)-responsive repressor that in the absence of metal ions represses transcription of the *smtA* gene, encoding a class II metallothionein (SmtA) involved in sequestering excess Zn(II) [25] (Fig. 2). Although SmtB also senses Co(II) and Cd(II) under some conditions in vivo [26] and in vitro [24], Zn(II) is the preferred metal ion effector [27]. A close homolog of *Synechococcus* sp. SmtB was subsequently identified in

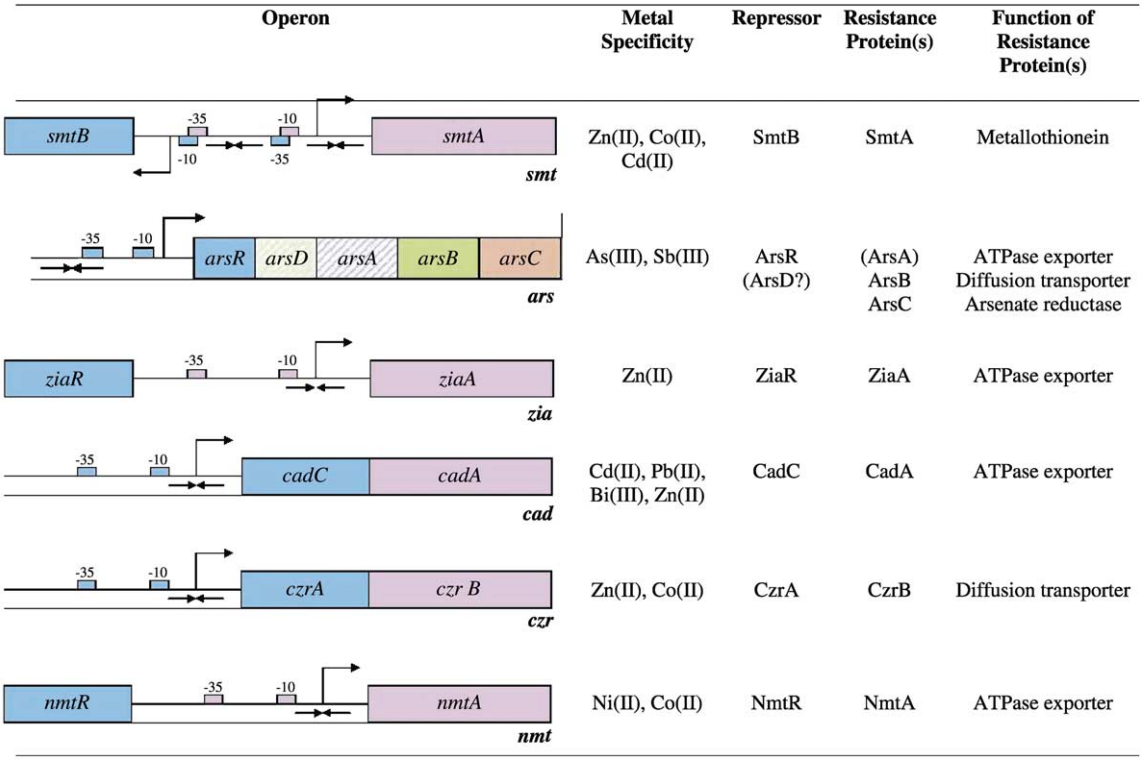

Fig. 2. Metal resistance operons regulated by SmtB/ArsR family transcriptional repressors. General operon structures which confer metal resistance in bacteria that are regulated in vivo by the indicated SmtB/ArsR repressors. In vivo metal specificity profile derived from [a][31], [b][28], [c][29], [d][27], [e][37], and [f][57]. Note that not all *ars* operons contain the *arsD* and *arsA* genes.

Synechocystis PCC 6803 and termed ZiaR [28]. *Synechocystis* sp. ZiaR shares ~50% sequence identity with SmtB and regulates the expression of the divergent *zia* operon which encodes ZiaR and ZiaA, a P-type ATPase metal efflux pump (Fig. 2). ZiaR appears to be a highly Zn(II)-specific sensor [28]. *S. aureus* also contains a chromosomal Zn(II)/Co(II)-specific metalloregulatory transcriptional repressor, CzrA (also designated ZntR)[1], that negatively regulates the expression of the *czr* operon (Fig. 2) which encodes, in addition to CzrA, a membrane-bound Zn(II) transporter of the cation diffusion facilitator family, CzrB (ZntA)[1] [29,30]. CzrA has the highest degree of sequence identity to SmtB, consistent with the fact that Zn(II) is the strongest inducer of the *smt*, *czr* and *zia* operons [28,30,31]. A Ni(II)/Co(II)-responsive repressor termed NmtR from *Mycobacterium tuberculosis* has also recently been identified [27,32]. The *nmt* operon contains the *nmtA* gene whose protein product shares sequence similarity with a Co(II)-exporting ATPase exporter, CoaT from the cyanobacterium *Synechocystis*; NmtA is thought to perform the same function in *M. tuberculosis* (Fig. 2) [19,27].

The other founding member of this family is the plasmid- or chromosomally encoded ArsR that acts as the arsenic/antimony-responsive repressor of the *ars* operon in *E. coli* [33], and other bacteria (Fig. 2) [34]. The *ars* operon from most bacteria encodes an arsenate reductase, as well as proteins required for metal ion extrusion [35]. *S. aureus* also contains CadC, encoded either on the multiple resistance plasmid pI258 or integrated into the bacterial chromosome in some strains [36,37]. Plasmid pI258-encoded CadC has been extensively studied and has been shown to bind thiophilic, toxic metal ions including Cd(II), Pb(II) and Bi(III), as well as Zn(II) in vivo and in vitro [37–40]. CadC regulates the expression of the *cad* operon which encodes a membrane-bound Cd(II)/Pb(II)-specific P-type ATPase efflux pump (Fig. 2) [37].

3. Structural studies of SmtB/ArsR family members

The 2.2 Å resolution X-ray crystallographic structure of dimeric, apo-SmtB revealed that SmtB is an elongated dimer with a two-fold axis of symmetry consisting of five α-helices and two β-strands arranged into an α1-α2-α3-αR-β1-β2-α5-fold [41]. The primary subunit interface is formed between the two N-terminal α1 C-terminal α5 helices (Fig. 3) [41]. The helix-turn-helix domain (α3-turn-

[1] Two groups originally cloned and functionally characterized a Zn(II)/Co(II)-specific operon from two similar strains of *S. aureus*. One group designated the operon *czr* (**c**hromosome-determined **z**inc-**r**esponsible operon) and the open reading frames *czrA* and *czrB* [29]. The other group designated the operon *znt* (**Zn t**oxicity) and the open reading frames *zntA* and *zntB* [30]. We have adopted the *czr* nomenclature [29] to avoid any confusion with the already characterized *E. coli znt* operon [18].

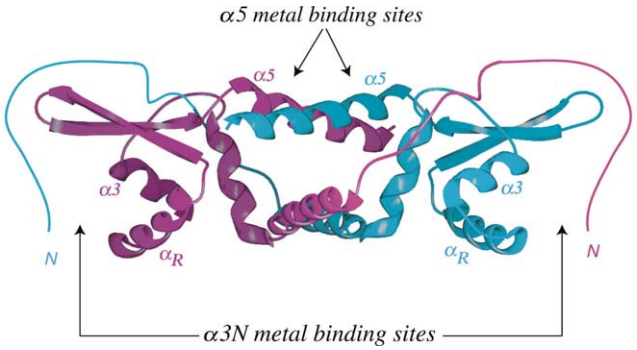

Fig. 3. Structure of Apo-SmtB. A ribbon representation of the structure of homodimeric apo-SmtB determined crystallographically with one monomer shaded purple and the other blue [41]. Each SmtB monomer adopts an αααααββα fold and contains a helix-turn-helix DNA binding domain of the winged helix family (α3-turn-αR). Four metal binding sites per dimer were inferred from difference electron density maps calculated after soaking these crystals in mercuric acetate [41]. Two symmetry-related inter-subunit sites were found to bridge the α5 helices (involving Asp104, His106, His117′, Glu120′), while two additional symmetry-related sites were observed near the α3 helices (involving Cys61, Asp64, His97, H_2O) [41]. Note: These Hg(II) complexes were characterized by very low occupancies, long coordination bonds, and irregular coordination geometries [41]. The N-terminal residues 1–24 were not observed crystallographically and were added manually.

αR) has a strong structural resemblance to other bacterial transcriptional regulators including CAP (catabolite activator protein) [42] and DtxR (the Fe(III)-regulated diphtheria toxin repressor) [43], and even MerR, the prototypical Hg(II)-responsive transcriptional regulator of the MerR family [17]. This DNA binding domain, particularly the sequence of the proposed DNA recognition α-helix (αR), is also highly conserved throughout the SmtB/ArsR family and is one of the distinguishing characteristics that define membership (cf. Fig. 1).

In an attempt to identify the metal binding site(s) of SmtB, a mercuric acetate soak of the apo-SmtB crystals was performed [41]. Despite the very low occupancies of the Hg(II) binding site(s) and the presence of some very unusual metal–ligand bond lengths and coordination geometries, two potential metal binding sites were mapped as difference electron density in $F-F_o$ maps prior to dissolution of the apo-SmtB crystal. One Hg(II) ion was proposed to be ligated by Cys61 and Asp64 in the α3 helix, His97 from β2, and potentially a water molecule. A second Hg(II) ion was coordinated via an intermolecular four-coordinate complex by Asp104 and His106 from one monomer and His117′ and Glu120′ from the other monomer [41]. Although inconclusive as to the structure of the coordination chelates formed with the native metal ion Zn(II), these studies provided an important structural framework with which to design experiments and interpret subsequent metal binding studies (see below).

Given the high degree of sequence similarity (25–50% identity) and the lack of a need to introduce significant gaps in a multiple sequence alignment (Fig. 1), threading of the amino acid sequence of *E. coli* R773 ArsR [44], *S. aureus* pI258 CadC [40] and *M. tuberculosis* NmtR [27] through the structure of apo-SmtB [41] has allowed the generation of robust, albeit low resolution, models of these homodimeric repressors. True atomic detail structures of ArsR, CadC and NmtR have yet to be reported. However, recent nuclear magnetic resonance studies of *S. aureus* CzrA reveal that the secondary structure of each subunit of the CzrA homodimer is exactly as anticipated from the sequence alignment with SmtB (Fig. 4A); furthermore, the secondary structure of Zn(II)-CzrA is, to a first approximation, identical to that of the apoprotein (Fig. 4B) (M. Pennella, C. Eicken, J. Sacchettini and

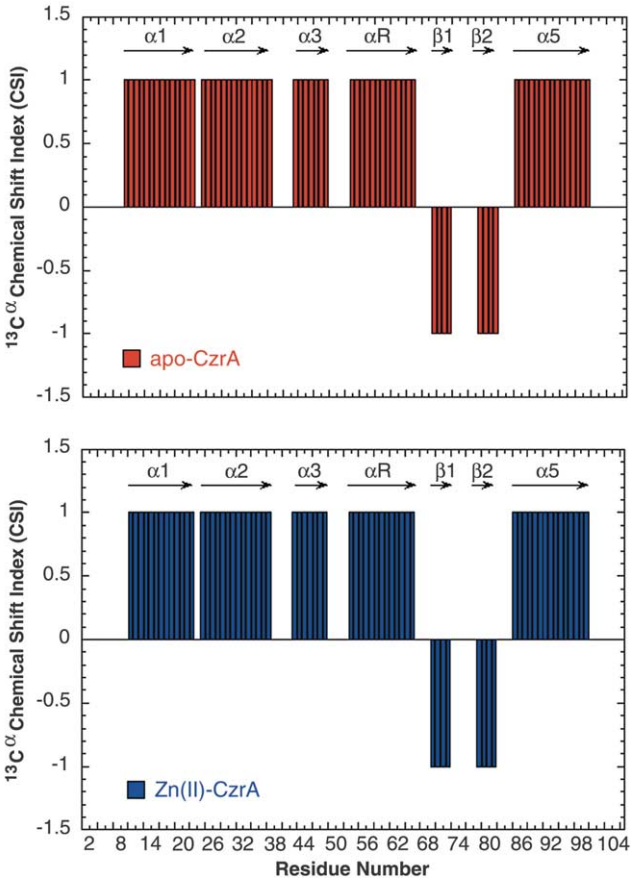

Fig. 4. Secondary structure analysis of apo- and Zn(II)-CzrA as determined by chemical shift indexing (M. Pennella, C. Eicken, J. Sacchettini and D.P. Giedroc, manuscript in preparation). A: $^{13}C\alpha$ chemical shift index (CSI) value vs. residue number for apo-CzrA. Chemical shift indexing compares the value of the observed $^{13}C\alpha$ chemical shift with those values reported for residues which are known to reside in regions of α-helix, β-sheet and random coil. Those residues in apo-CzrA characterized by an α-helical $^{13}C\alpha$ chemical shift are assigned a CSI value of +1, those with β-sheet $^{13}C\alpha$ chemical shifts a value of −1, while a CSI value of 0 is assigned to those residues with random coil $^{13}C\alpha$ chemical shifts [64]. Apo-CzrA has the same secondary structure as apo-SmtB, αααααββα. B: $^{13}C\alpha$ CSI vs. residue number for Zn(II)-CzrA. Apo- and Zn(II)-CzrA homodimers are characterized by the same monomeric secondary structure. Conditions: 0.8 mM ^{13}C, ^{15}N CzrA (±1 equiv. Zn^{2+}), 10 mM d^{18}-HEPES pH 6.0, 0.1 M NaCl, 10% D_2O.

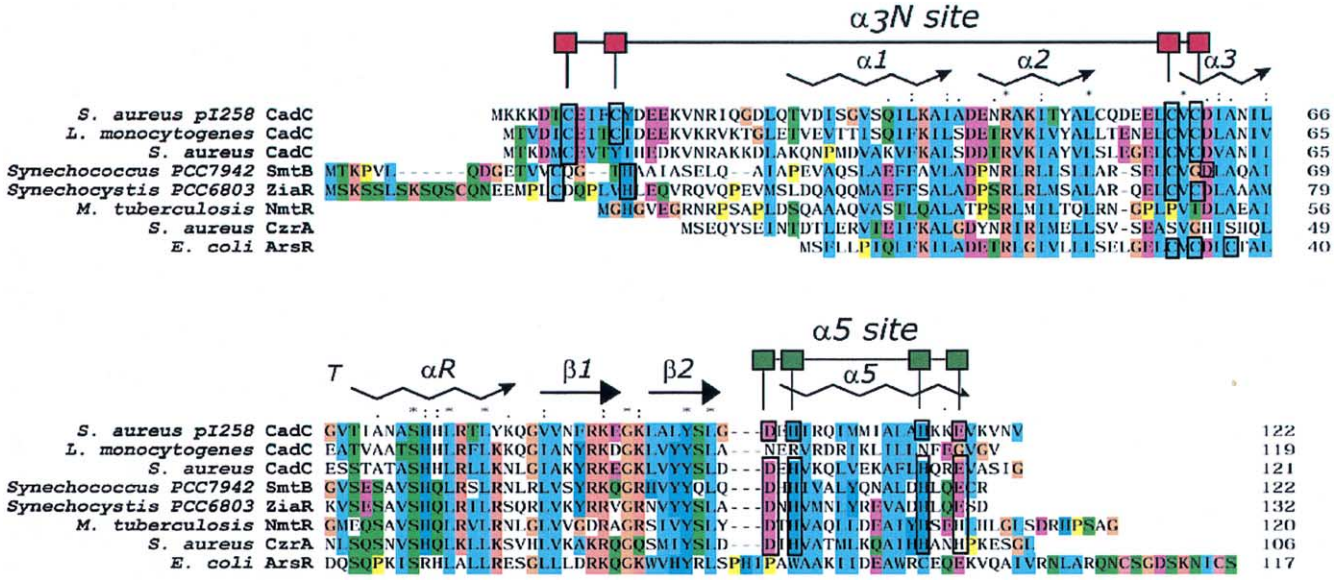

Fig. 5. Proposed α3N and α5 metal binding sites for functionally characterized SmtB/ArsR proteins. A multiple sequence alignment of several functionally characterized SmtB/ArsR proteins generated by ClustalW [62] (Swiss-Prot, TrEMBL, or GenBank accession numbers are in parentheses): *S. aureus* pI258 CadC (P20047), *L. monocytogenes* CadC (Q56405), *S. aureus* CadC (P37374), *Synechococcus* PCC 7942 SmtB (P30340), *Synechocystis* PCC 6803 ZiaR (Q55940), *M. tuberculosis* NmtR (NP_218261.1), *S. aureus* CzrA (O85142), and *E. coli* R773 ArsR (P15905). The proposed α3N and α5 metal binding sites are denoted with the secondary structure of apo-SmtB [41] indicated above the alignment, with the conserved EL**C**V(**C**/*G*)D metal binding box located in the α3 helix. Conserved metal binding residues are boxed and the amino acids are color-coded as follows: orange-basic; purple-acidic; green-polar neutral; blue-ionizable aromatic; light blue-non-polar aliphatic; yellow-proline; brown-glycine; black-divergent. In both α3N and α5 sites, the N-terminal ligand pair is known or predicted to be derived from one monomer, with the C-terminal ligand pair derived from the other monomer within the homodimeric protein to create an inter-subunit metal binding site. The α3N metal binding ligands are indicated with red boxes in the secondary sequence designation, while the α5 ligands are indicated with green boxes.

D.P. Giedroc, manuscript in preparation). This experiment strongly suggests that metal binding to apo-CzrA results in a change in tertiary or quaternary structure. Interestingly, significant chemical shift changes localize largely to the α5 helix. These studies suggest that a similar tertiary and quaternary structural fold characterizes individual members of the entire SmtB/ArsR family.

4. Identification of two distinct metal-sensing sites within a conserved structural scaffold

A highly conserved EL**C**V(**C**/*G*)D motif termed the 'metal binding box' was initially identified in members of the SmtB/ArsR family (Fig. 5) [9]. This motif was proposed to contain residues involved in metal coordination and therefore directly involved in metal ion sensing. In fact, several repressor proteins in this family do utilize ligands from this 'box' including ArsR and CadC proteins; however, it is now known that this parallel cannot be made to all other members of the family [39,45,46]. The X-ray structure of apo-SmtB revealed that the EL**C**V(**C**/*G*)D motif is present in the α3 helix, as part of the projected α3-turn-αR DNA binding motif (Fig. 5) [41]. This supported the hypothesis that this sequence was required for metal ion sensing by the direct binding of metal ions.

Early reports on ArsR corroborated this hypothesis since substitution of one or both cysteines with non-metal-liganding residues in the ^{30}EL**C**V**C**D^{35} motif inhibited the ability of arsenate salts to dissociate ArsR from the *ars* O/P [46]. X-ray absorption spectroscopy of As(III)-ArsR revealed that As(III) was coordinated via three cysteine thiolate ligands within the putative α3 helix (Fig. 5), two of which were clearly derived from the ^{30}EL**C**V**C**D^{35} motif [46]. The crystallographic studies of SmtB subsequently showed that one Hg(II) ion was close to Cys61 and Asp64 in the ^{59}EL**C**V*G*D^{64} sequence, analogous to the ^{30}EL**C**V**C**D^{35} sequence of ArsR (Fig. 3) [41]. Surprisingly, Cys61 of SmtB was previously found *not* to be essential for Zn(II) sensing in vivo, in contrast to Cys32 of ArsR [46,47]. A second Hg(II) ion was crystallographically identified near Asp104 and His106 from the α5 helix of one monomer and His117′ and Glu120′ from the α5 helix of the other monomer, as well as the symmetry-related site on the dimer (Fig. 5) [41]. This was the first structural evidence that two distinct metal binding sites might exist within the family. Mutagenesis of His105 and His106, together, in SmtB had earlier been shown to inhibit Zn(II) sensing in vivo, suggesting that if Zn(II) and other inducing ions bound to apo-SmtB like Hg(II), the metal site across the α5 helices may be more important for metal sensing by SmtB, in contrast to ArsR [47]. Note that the ligands associated with the putative α5 helices are not conserved in any ArsR (cf. Fig. 1).

Recent comparative spectroscopic and functional studies of highly purified wild-type and ligand substitution

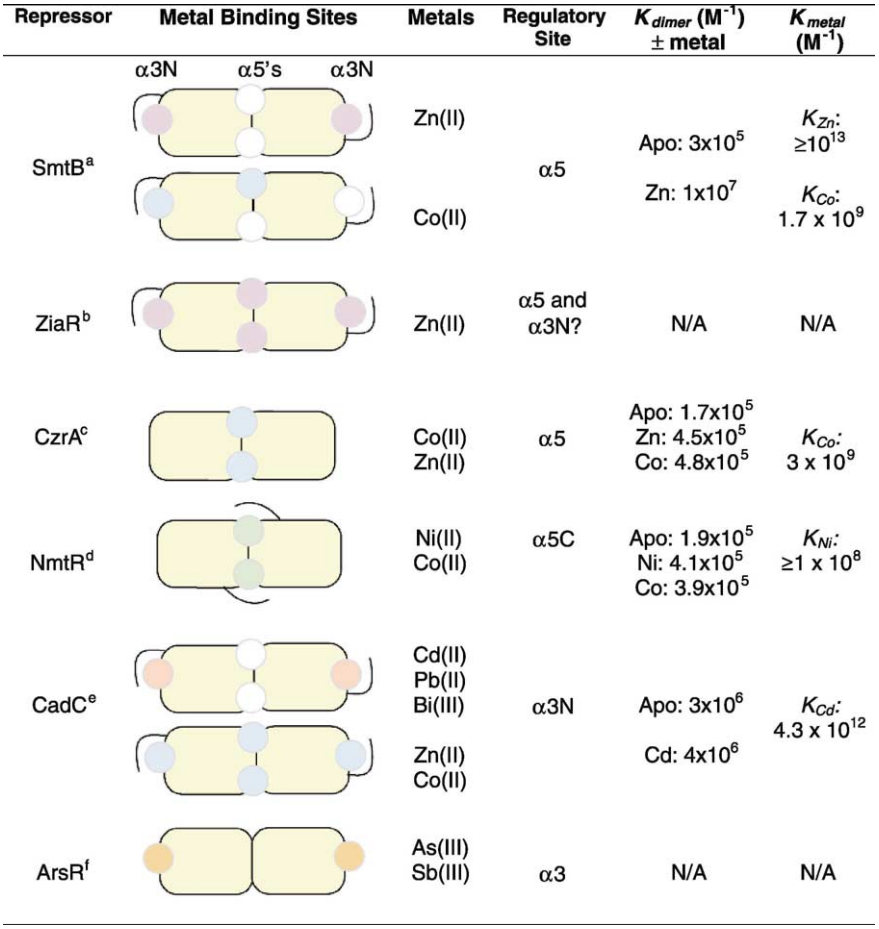

Fig. 6. Summary of metal binding sites and thermodynamic parameters of SmtB/ArsR proteins. Cartoon representations of functionally characterized SmtB/ArsR repressors depicting a 'theme and variations' model for metal binding ([a][24,48,55]; [b][28]; [c]M. Pennella, C. Eicken, J. Sacchettini and D.P. Giedroc, manuscript in preparation; [d][27,32]; [e][23,39]; [f][46,57]). Metal ions, denoted as spheres, are color-coded: blue-Co(II); purple-Zn(II); green-Ni(II); pink-Cd(II)/Pb(II); orange-As(III). The α3N site is composed of ligands derived from the N-terminus and α3 helix, while the α5 site is composed of ligands derived exclusively from the α5 helix. For CzrA and NmtR, no α3N site is present, while for NmtR only, the α5 metal binding site is proposed to contain two additional ligands derived from a short C-terminal extension of the α5 helix (α5C) [27,32]. For ArsR proteins, the α5 site is absent altogether as is the N-terminal 'arm'; this metal binding site is denoted α3. These proteins are weakly dissociable homodimers (K_{dimer}). Metal binding to the α5 site has a measurable effect on the dimerization equilibrium (K_{dimer}+metal). All SmtB/ArsR repressors bind metal ions with high affinity at equilibrium (K_{metal}).

mutants of *Synechococcus* SmtB, *S. aureus* pI258 CadC, *Synechocystis* ZiaR, *S. aureus* CzrA (M. Pennella, C. Eicken, J. Sacchettini and D.P. Giedroc, manuscript in preparation) and *M. tuberculosis* NmtR are consistent with a unifying 'theme and variations' model for metal binding by this family of proteins. These solution studies clearly identify two distinct metal binding sites within the homodimeric repressor, designated the α3N (or α3 in the case of ArsR) and α5 (or α5C in the case of NmtR) metal binding sites (Fig. 5, see also Fig. 3) [24,27,28,32,39,48]. As the designation implies, the α3N site is composed of three or four cysteine thiolate ligands, two of which are derived from the **CXC** sequence within the α3 helix metal binding box and with one or two cysteine ligands derived from the N-terminal region. This N-terminal region of the protein was not observed crystallographically for apo-SmtB and these ligands were not proposed for the Hg(II) ion in this site (Fig. 3) [41]. This high affinity, thiolate-containing site is the regulatory site for the Cd(II)/Pb(II) sensor protein CadC and is characterized by a preference for larger, more thiophilic metal ions [23,38–40]. Recent experiments suggest that the N-terminal thiolate ligands are derived from one monomer and the α3 thiolate ligands are derived from the opposite monomer within the homodimer to create an inter-subunit α3N site, which is likely due to the inherent flexibility of the N-terminal arm [40,49]. ArsRs lack an N-terminal 'arm' and coordinate the smaller, thiophilic ions such as As(III) and Sb(III) with three clustered cysteines within the α3 helix, although just two of these are absolutely invariant and appear necessary and sufficient for As(III) sensing in vivo [46].

The α5 metal site consists of four metal ligands derived exclusively from the C-terminal α5 helices, forming a tetrahedral or distorted tetrahedral metal complex across the dimerization interface, as originally hypothesized from the crystallographic studies of SmtB (Fig. 3) [41]. The Zn(II)

sensors SmtB and CzrA utilize this site within the homodimer for strong negative regulation of O/P binding (see below). Like the α3N site, the α5 metal binding site is characterized by an extremely high affinity for metals at equilibrium, but shows a clear preference for biologically required 'borderline' hard/soft metal ions such as Zn(II), Co(II) and Ni(II) [27,48,50]. Recent functional and spectroscopic findings reveal that the Ni(II)/Co(II) sensing site for NmtR differs from the Zn(II) sensing site for CzrA and SmtB in that one or two additional N/O ligands are recruited into the chelate from a short C-terminal extension of the α5 helix (Fig. 5) [27,32]. This forms an α5C site with an increased coordination number (5 or 6), consistent with simple expectations from examination of the most common coordination geometries for Zn(II) vs. Ni(II) in the protein database [51].

Another interesting finding from these studies is that some SmtB/ArsR family members possess exclusively the α3/α3N site (e.g., all ArsRs, *Listeria monocytogenes* CadC), others contain only the α5/α5C site (e.g., CzrA, NmtR) and still others contain both metal binding sites (e.g., most CadCs, SmtB and ZiaR) (Figs. 5 and 6). It is unclear what role, if any, the 'additional' metal binding sites play for proteins that contain both the α3N and α5 metal sites, but thus far, the general trend is that only one appears essential for allosteric metalloregulation in vitro and metal sensing in vivo (Fig. 6) [24,39,49]. The possible exception to this is *Synechocystis* ZiaR since Zn(II) and Co(II) binding experiments reveal that both metal sites can be simultaneously occupied to form what appear to be S_3(N/O) α3N and $(N/O)_4$ α5 coordination complexes (Fig. 7) (M.L. VanZile and D.P. Giedroc, unpublished results). Unlike the case for SmtB, however, substitution of the Cys71/Cys73 pair, proposed ligands to the α3N metal ion, or His116, a proposed α5 ligand (cf. Fig. 1), to non-liganding amino acids abolished Zn(II) sensing by ZiaR in *Synechocystis* [28].

Another 'variation' is the finding that Co(II) appears to bind to the *Synechococcus* SmtB homodimer in such a way that one Co(II) ion is bound in one α5 site and the other to an α3N site with an overall stoichiometry of two metal ions per dimer, with the remaining two symmetry-related α5 and α3N sites on the dimer unoccupied (Fig. 6) [48]. However, when the primary α3N cysteine ligands Cys14 and Cys61 are mutated, Co(II) preferentially binds at the inter-subunit α5 site with the recruitment of Cys121. *S*-Methylated SmtB also coordinates Co(II) exclusively at the α5 sites and rapidly equilibrates between the α3N and α5 metal binding sites upon reduction of the RS–S–CH_3 mixed disulfide linkages [48]. It is not yet known if Zn(II) also partitions between the α3N and α5 sites (Fig. 6), but on thermodynamic grounds, the Zn(II) binding affinity for wild-type SmtB is at least 20-fold greater than any SmtB mutant which binds Zn(II) in the α5 sites exclusively. Therefore, the presence of an intact α3N metal binding site must enhance the overall Zn(II) affinity [48].

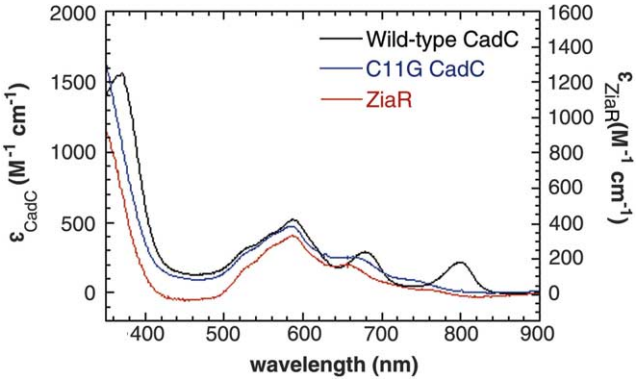

Fig. 7. Visible absorption spectra of Co(II)-substituted *Synechocystis* ZiaR (red line), *S. aureus* pI258 CadC (black line), and a Cys11→Gly variant of pI258 CadC (blue line). The absorption spectrum of wild-type *S. aureus* pI258 CadC is a superposition of Co(II) complexes formed at the α3N and α5 metal binding sites (see Fig. 5) [39]. The lowest energy transitions (600–800 nm) are suggestive of a strongly distorted tetrathiolate α3N Co(II) complex. The spectrum for C11G-CadC is also characterized by Co(II) binding to both the α3N and α5 sites, except that the strong blue shift in the low energy transitions suggests that just three cysteine thiolate–Co(II) coordination bonds are formed in the mutant α3N Co(II) complex, i.e., Cys11 is a ligand [39]. The good correspondence of the Co(II) absorption spectra of *Synechocystis* ZiaR and C11G-CadC suggests that ZiaR coordinates the α3N Co(II) in a S_3(N/O) complex and the α5 Co(II) in a $(N/O)_4$ complex. This is consistent with the multiple sequence alignment (cf. Fig. 1) which shows that ZiaR contains just three of the conserved α3N cysteines, with the Cys11 equivalent absent (cf. Fig. 5). Conditions: 100–150 μM protein, 2 equiv. Co(II) per monomer, 5 mM MES pH 7.0, 0.2 M NaCl.

5. Characteristics of O/P sequences

The structures of all metal-inducible operons contain one or two imperfect 12-2-12 inverted repeats, generally located near or overlapping the transcriptional start site of the metal-regulated gene (Figs. 2 and 8). The *smt* operon is unique in that it contains two imperfect, inverted 12-2-12 repeats of similar sequence termed 'S2/S1' and 'S4/S3'. In addition, the sequence of each repeat is also conserved for other native metal ion sensors of the SmtB/ArsR family (Fig. 8). The S2/S1 repeat is required for full Zn(II) responsiveness of *smtA* expression from the *smtA* O/P in vivo and is also required for transcriptional regulation by SmtB, while the S4/S3 repeat has little if any effect on the regulation of *smtA* expression [31,47]. The *Synechocystis zia* O/P region is nearly identical in organization and sequence to that of the *smtA* O/P, containing a single conserved 12-2-12 inverted repeat located between the −10 hexamer and the translational start site of *ziaA*, with the gene encoding the repressor, *ziaR*, divergently transcribed (Figs. 2 and 8) [28].

The *czr* operon of *S. aureus* confers resistance to Zn(II) and Co(II) ions in vivo and contains two adjacent genes designated *czrA* and *czrB* (Fig. 2). The *czrA* and *czrB* genes are co-transcribed, much like the *cadC*/*cadA* genes in *S. aureus* [30]. The level of the *czrA*/*B* transcript is significantly increased when the cells are exposed to

Zn(II) and Co(II) [29,30]. DNase I footprinting mapped the CzrA binding site within the *czr* O/P to approximately 49 nucleotides located between the −10 sequence and the *czrAB* translational start site, a region which contains an imperfect 12-2-12 inverted repeat highly similar to that of the *smtA* O/P (Fig. 8) [52]. Like the *smtA*, *zia* and *czr* O/Ps, the *M. tuberculosis nmt* O/P contains a single, α5 sensor consensus 12-2-12 inverted repeat with the gene encoding *nmtR* divergently transcribed (Figs. 2 and 8); purified NmtR has been shown to bind tightly to an oligonucleotide containing this sequence [27,32].

The *cad* operon contains two co-transcribed genes that encode CadC and CadA from one promoter (Fig. 2) [37]. In vivo, CadC repression from the *cad* O/P could be alleviated by the addition of Cd(II), Pb(II), Bi(III), and Zn(II) salts to the growth medium [37,53]. As measured by DNase I footprinting, the binding of CadC protects a 20-bp region between the −10 hexamer and the translational start site of the *cad* O/P [38]. This protected area is located around a 6-2-6 core hyphenated inverted repeat which is part of an extended 12-2-12 imperfect repeat, similar to that of the *smt*, *zia*, *czr*, and *nmt* operons (Fig. 8). The *ars* operon of *E. coli* also contains an imperfect 12-2-12 repeat, more closely similar to the *cad* O/P sequence than to the *smtA*, *zia*, *czr* and *nmt* repeats (Fig. 8) [54].

6. Stoichiometry of SmtB/ArsR repressor–O/P binding

Equilibrium analytical ultracentrifugation studies reveal that SmtB/ArsR repressors are weakly dissociable homodimers, in both the metal-bound and metal-free states (K_{dimer}, Fig. 6) [23,32,39,55], with the monomer–dimer equilibrium linked to site-specific DNA binding equilibria. SmtB is thus far unique in that a monomer–dimer–tetramer equilibrium better describes the self-association of this protein [55]. The monomer is assumed not to have appreciable DNA binding activity, although an assembly or cooperative, monomer binding model cannot yet be rigorously ruled out. It has been shown, however, that apo-SmtB binds weakly to an oligonucleotide containing one-half of an inverted repeat in vitro [24], which gives rise to the weak repression observed in vivo [47].

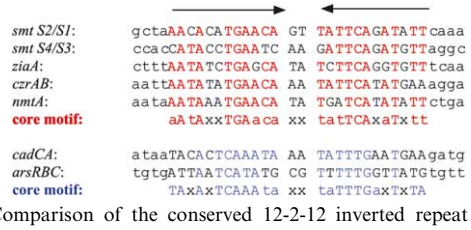

Fig. 8. Comparison of the conserved 12-2-12 inverted repeat of SmtB/ArsR-regulated O/Ps. Alignment of the imperfect, hyphenated 12-2-12 inverted repeats from the *Synechococcus smt* O/P (S2/S1, S4/S3), *Synechocystis zia* O/P, *S. aureus czr* O/P, *M. tuberculosis nmt* O/P, *S. aureus* pI258 *cad* O/P, and *E. coli* R773 *ars* O/P suggests distinct consensus core motifs for the α5 and α3N sensors. The apparent divergence may be directly linked to the differences in amino acid sequences of the DNA recognition α-helices (cf. Fig. 5) [24].

The stoichiometry of O/P binding by metal sensor protein dimers remains the subject of ongoing investigation. From the simple finding that both the SmtB homodimer and the inverted 12-2-12 repeat are approximately two-fold symmetric, the expectation was that a single homodimer would bind to a single inverted repeat, with the DNA recognition helices (αR) within the α3-T-αR motif interacting with successive major grooves. A molecular modeling experiment, however, suggested that if this were the case, a significant conformational change in the DNA (underwinding and/or bending toward the minor groove) would have to occur to accommodate the binding of a single dimer to a single 12-2-12 inverted repeat [41]. Recent experiments with SmtB reveal that the stoichiometry of binding is two tight-binding SmtB dimers ($K_a \geq 10^9$ M^{-1} at pH 7.4, 0.15 M NaCl) to a 40-bp oligonucleotide containing a single S1/S2 12-2-12 inverted repeat as determined by fluorescence anisotropy, analytical equilibrium ultracentrifugation and isothermal titration calorimetry (L.S. Busenlehner and D.P. Giedroc, unpublished results); an additional dimer was also found to bind with modest affinity under these conditions ($K_a \geq 10^7$ M^{-1}) [24,56]. The physical arrangement of these two dimers on a single 12-2-12 inverted repeat is not known, but several models have been proposed [24,56]. Interestingly, the full ∼100-bp O/P region containing both S2/S1 and S4/S3 inverted repeats was found to bind just two SmtB dimers, rather than the four expected from the results obtained for each inverted repeat separately. This finding suggests that the full two-site *smt* O/P may form a looped complex, stabilized by dimer–dimer interactions [56].

Homodimeric CzrA also forms multimeric complexes with a 48-bp *czr* O/P-containing oligonucleotide with a limiting stoichiometry of approx. four dimers, with two again bound very tightly, like SmtB (pH 7.0, 0.4 M NaCl, 25°C). This finding is consistent with the extensive region of DNase I protection observed by the binding of CzrA to the *czr* O/P [52]. Although the binding stoichiometry of ZiaR has not been similarly investigated, evidence for multimeric protein–DNA complexes was obtained from gel mobility shift assays in the presence of the metal ion chelator, 1,10-phenanthroline [28].

Fluorescence anisotropy titrations reveal that a single CadC homodimer binds to a 34-bp *cad* O/P-containing oligonucleotide with high affinity, even at the high monovalent salt concentration ($K_a = 1.1 \times 10^9$ M^{-1}; 0.4 M NaCl, pH 7.0, 25°C) and that CadC monomer–dimer equilibrium (Fig. 6) is linked to DNA binding under these conditions [23]. However, if the NaCl concentration is reduced, a second CadC dimer does indeed bind, albeit with an estimated affinity of $K_a < 10^6$ M^{-1} [39]. Consistent with this, gel mobility shift experiments reveal two distinct CadC–DNA complexes, with the slowest migrating complex significantly more populated at higher concentrations

of CadC [38]. The stoichiometry of ArsR–DNA binding equilibria has not been similarly investigated. However, DNase I footprinting reveals that both *E. coli* and plasmid R773-encoded ArsR protect a region of ~33 bp overlapping the −35 element in the *ars* O/P, while gel mobility shift experiments suggest that both ArsRs form one resolvable protein–DNA complex [54,57].

7. Negative regulation of repressor–O/P binding by inducing metals

The direct binding of metal ions to the regulatory (α3N/α3 or α5/α5C) sites of individual SmtB/ArsR repressors (Fig. 6) results in strong inhibition or negative allosteric regulation of specific O/P DNA binding. This finding is consistent with a simple model of derepression, in which metal binding by the sensor protein weakens the DNA binding affinity significantly, such that RNA polymerase can load and initiate transcription of the operon. Quantitative estimates of the extent of regulation reveal that for SmtB, the binding affinity of both high affinity dimers for a single 12-2-12 repeat is reduced by ~1000-fold upon coordination of a single equivalent of Zn(II), Cd(II) or Co(II) to the monomeric protein. These metals also dissociate preformed apo-SmtB–*smt* O/P complexes to a level expected from equilibrium binding isotherms[2] [24]. The same findings characterize the interaction of CzrA with the *czr* O/P as determined by fluorescence anisotropy [32] and DNase I footprinting experiments [52]. Interestingly, gel mobility shift experiments seemed to indicate that Zn(II) is incapable of dissociating *Synechocystis* ZiaR from the *zia* O/P in vitro; however, the observed enhancement of DNA binding in the presence of metal chelator suggests that Zn(II) and other metals do indeed bind to ZiaR and modulate its affinity for the *zia* O/P [28].

The O/P binding properties of other members of the SmtB/ArsR family, including the α5C sensor NmtR and the α3N sensor CadC, are also strongly negatively regulated by metal binding. For example, Pb(II)-, Cd(II)-, Zn(II)-, and Bi(III)-bound CadC proteins have significantly decreased affinities (340-, 230-, 185-, and 170-fold decrease respectively) for the DNA compared to the apoprotein [23,39,40]; addition of these metal ions to the preformed apo-CadC–DNA complex also results in CadC dissociation from the DNA (M.L. VanZile and D.P. Giedroc, unpublished results). The binding of Cd(II), Pb(II) and Zn(II) to *both* α3N metal binding sites in dimeric CadC is required for full negative regulation of DNA binding, with the α5 site fully dispensable in vitro [39] and in vivo [49]. *M. tuberculosis* NmtR is extremely interesting since although this Ni(II)/Co(II) repressor binds Zn(II) nearly as avidly as Ni(II), Zn(II) is not as potent an allosteric effector of DNA binding as Ni(II) [27,32]. This finding suggests that metal ion selectivity of individual SmtB/ArsR family members might operate at multiple levels of regulation.

In essentially all cases examined, amino acid substitutions placed into the first coordination shell of metal ligands in any SmtB/ArsR repressor give rise to a repressor that maintains high affinity for the O/P in the absence of metal ions in vitro and in vivo, but is totally or partially insensitive to negative regulation of DNA binding by metal ions [24,27,28,39,45]. The simplest interpretation of such a finding is that these mutant repressors no longer bind metal ions. However, in some cases, these metal–ligand mutant repressors do maintain a high equilibrium affinity for metal ions, but simply form non-native coordination complexes which are functionally compromised [39,48]. Such an approach can, in fact, be used to estimate the extent to which individual metal coordination bonds drive allosteric regulation of DNA binding. For example, characterization of a collection of Cys→(Gly/Ser) mutants of *S. aureus* pI258 CadC has revealed that all mutant apo-CadCs maintain a high affinity for the *cad* O/P in vitro [39] and function as active repressors in vivo [45]. All bind metals tightly, but form coordination complexes that are spectroscopically distinct from that of wild-type CadC [39,40]. However, the functional importance of individual metal–ligand bonds was not found to be identical, with the rank order of each α3N cysteine in mediating the allosteric response by Cd(II), Pb(II) and Bi(III) found to be Cys60 ≥ Cys7 > Cys58 ≫ Cys11, with C60G CadC completely non-functional [39,40,45].

In other cases, first coordination shell ligand substitutions may simply destabilize the site to a degree that metals no longer bind with measurable affinity. This has been shown to occur in H106Q SmtB, where Zn(II) and Co(II) binding to this mutant α5 site cannot be detected; as expected, this leads to a complete loss of metal-mediated dissociation from the DNA in vitro [24,48], consistent with the finding that His106 is obligatory for zinc sensing by SmtB in vivo [47]. Which mechanism is operative in other metal ligand mutant SmtB/ArsR repressors shown to be defective in metal sensing in vivo has not yet been investigated, but should prove to be a valuable approach for defining the structural basis of metal-mediated signal transduction in these proteins.

8. Identification of new SmtB-like metal sensors and the molecular basis for allosteric regulation

A series of BLAST searches of the genomic sequence database using each of the six SmtB/ArsR proteins discussed here identifies >250 open reading frames derived from a wide spectrum of prokaryotic genomes (cf. Fig. 1).

[2] One report [56] suggests that Zn(II) does not dissociate the complex formed between SmtB and a 100-bp oligonucleotide containing the S1/S2 and S3/S4 sites; possible reasons for this have been discussed [24].

Although it is not yet possible to predict the metal ion selectivity of each putative SmtB/ArsR repressor from the amino acid sequence alone, thiolate-rich α3N and/or C-terminal α5 sites can be readily identified in ~15% of these sequences. It is notable that the *M. tuberculosis* genome encodes 10 putative SmtB/ArsR family members [27], most of which do not precisely fit the α3N and α5 metal binding site signatures documented here. In addition, the Hg(II) sensor MerR from *Streptomyces lividans* 1326 has been classified as a SmtB/ArsR protein on the basis of sequence analysis and the finding that Hg(II) mediates strong negative regulation of O/P binding, like other SmtB/ArsR repressors [58]. *S. lividans* MerR contains clustered cysteine residues in the N-terminal region, as well as in the putative α3 and α5 helices. A future challenge will be to identify the metal selectivities of each of these novel repressors.

A detailed multiple sequence alignment, carried out in the context of the high resolution structures of one or more apo- and metallated SmtB/ArsR repressors, using for example an evolutionary trace method, will reveal not only those residues that are common to all family members (these presumably play a role in maintaining tertiary or quaternary structure or are conserved as part of the αR DNA recognition helix), but other key residues outside those that are required to bind metal that co-vary with the evolution of new metal binding or functional specificities, e.g., α3N vs. α5 sensors (cf. Fig. 1) [59,60]. These will surely include those residues required to transduce the metal-mediated allosteric response. For example, several residues conserved uniquely in the α5 sensors are found primarily in the αR and α5 helices, as well as the β-strands, some of which may be ideal candidates for transducing the allosteric response in this subgroup. Conversely, for the α3N sensor CadC, a cluster of α3N-only residues is conserved in the N-terminal 'arm', while others are found in the α3 and αR helices (cf. Fig. 1).

Evolutionarily, at least one of the α5 sensors (*Synechococcus* SmtB) has apparently lost the α3N allosteric pathway even though an intact metal binding site remains; the opposite situation holds true for most CadCs (cf. Fig. 1). It may be possible to generate a novel α3N sensor by reintroducing an α3N pathway into an α5 sensor with a few key amino acid substitutions. Of course, the allosteric pathway of communication between metal and DNA binding sites (if one truly exists) is likely to be more complex than this. Detailed comparisons of the structures and dynamics of several α5 and α3N sensors, in the presence and absence of bound metal ions, will help to identify and refine the distinct allosteric pathways for each repressor subgroup.

9. Concluding remarks

The prokaryotic SmtB/ArsR family now includes not only the Zn(II)- and As(III)-specific transcriptional repressors originally characterized, but also those which confer resistance to Cd(II), Pb(II), Bi(III), Co(II), and Ni(II) (Fig. 2). A phylogenetic analysis of a subset of SmtB/ArsR repressor sequences clearly shows that, thus far, the metal ion sensors that respond to biologically required metal ions cluster on a distinct branch of the dendrogram and may have evolved later than those which confer resistance to heavy metal pollutants, likely as the result of environmental stress (cf. Fig. 1B). Both the α3N and α5 metal binding sites are relatively solvent-exposed and are characterized by extremely high equilibrium affinities for metal ions (Fig. 6). As argued for *E. coli* ZntR and Zur, this makes it unlikely that these metal sensor proteins function under thermodynamic control, where the metal binding affinity of the sensor protein is appropriately tuned to $1/[\text{metal}]_{\text{free}}$ such that a successful stress response can be mounted [4]. It seems more likely that these repressors function under kinetic control, in which the relative rates of metal–ligand exchange between low molecular mass intracellular thiolates, e.g., cysteine and glutathione, and metal sensor proteins regulate the transcriptional response. Detailed investigation of the kinetics of metal binding and metal-induced dissociation of O/P-bound sensor proteins in the presence and absence of low molecular mass thiolates will be required to gain more insight into the mechanism of transcriptional control by SmtB/ArsR sensor proteins.

Finally, the overriding evidence is that simple expectations from coordination chemistry of different metal ions are largely met when the structure of the specific metalloregulatory chelates are examined in detail. For example, the α3N site of pI258-encoded CadC adopts a trigonal S_3 structure when bound to Pb(II), but forms four-coordinate S_4 complexes with Cd(II) and Bi(III) [23,39,40], exactly as expected from studies with low molecular mass multi-dentate thiolate-containing ligands [61]. Likewise, the α5 site of the Zn(II) and Zn(II)/Co(II) sensors is four-coordinate and (distorted) tetrahedral [48], while the α5C site for the Ni(II)/Co(II) sensor NmtR is characterized by a higher coordination number (5 or 6) complex [27,32]. However, recent studies with *M. tuberculosis* NmtR and *Synechococcus* SmtB cast doubt on the proposal that affinity and measured specificity of metal coordination complexes in vitro are the sole determinants for functional metal selectivity in vivo [27]. NmtR binds Co(II) far more weakly than SmtB (cf. Fig. 6), but functions as a clear Co(II) sensor in an *smt*$^-$ strain of *Synechococcus*; strikingly, SmtB does not sense Co(II) in the same strain [27]. These and other findings argue that the biological response toward toxic metal ions will be strongly modulated by the nature of the intracellular environment, i.e., how individual metals are trafficked and compartmentalized in different cells, as well as the presence and relative efficiency of metal import and efflux systems. Studies along these

lines should prove a fruitful avenue for future investigation.

Acknowledgements

Work on this project in the authors' laboratory was supported by grants from the NIH (GM42569) and the Robert A. Welch Foundation (A-1295).

References

[1] McCall, K.A., Huang, C. and Fierke, C.A. (2000) Function and mechanism of zinc metalloenzymes. J. Nutr. 130, 1437–1446.
[2] Rees, D.C. (2002) Great metalloclusters in enzymology. Annu. Rev. Biochem. 71, 221–246.
[3] Wunderli-Ye, H. and Solioz, M. (1999) Copper homeostasis in *Enterococcus hirae*. Adv. Exp. Med. Biol. 448, 255–264.
[4] Outten, C.E. and O'Halloran, T.V. (2001) Femtomolar sensitivity of metalloregulatory proteins controlling zinc homeostasis. Science 292, 2488–2492.
[5] Touati, D. (2000) Iron and oxidative stress in bacteria. Arch. Biochem. Biophys. 373, 1–6.
[6] O'Halloran, T.V. (1993) Transition metals in control of gene expression. Science 261, 715–725.
[7] O'Halloran, T.V. (1989) Interrelations among metal ions, enzymes, and gene expression. In: Metal Ions in Biological Systems (Sigel, H. and Sigel, A., Eds.), Vol. 25, pp. 105–146. Marcel Dekker, New York.
[8] Silver, S. (1996) Bacterial resistances to toxic metal ions – a review. Gene 179, 9–19.
[9] Shi, W., Wu, J. and Rosen, B.P. (1994) Identification of a putative metal binding site in a new family of metalloregulatory proteins. J. Biol. Chem. 269, 19826–19829.
[10] Taghavi, S., Mergeay, M., Nies, D. and van der Lelie, D. (1997) *Alcaligenes eutrophus* as a model system for bacterial interactions with heavy metals in the environment. Res. Microbiol. 148, 536–551.
[11] Silver, S. and Phung, L.T. (1996) Bacterial heavy metal resistance: new surprises. Annu. Rev. Microbiol. 50, 753–789.
[12] Silver, S. (1992) Plasmid-determined metal resistance mechanisms: range and overview. Plasmid 27, 1–3.
[13] Meng, W., Cavet, J.S. and Robinson, N.J. (2001) SmtB/ArsR transcriptional regulators: Action and specificity. J. Exp. Biol. 52, 71.
[14] Rosen, B.P. (1996) Bacterial resistance to heavy metals and metalloids. J. Biol. Inorg. Chem. 1, 273–277.
[15] Alonso, A., Sanchez, P. and Martinez, J.L. (2000) *Stenotrophomonas maltophilia* D457R contains a cluster of genes from gram-positive bacteria involved in antibiotic and heavy metal resistance. Antimicrob. Agents Chemother. 44, 1778–1782.
[16] Ansari, A.Z., Bradner, J.E. and O'Halloran, T.V. (1995) DNA-bend modulation in a repressor-to-activator switching mechanism. Nature 364, 371–375.
[17] Shewchuk, L.M., Verdine, G.L. and Walsh, C.T. (1989) Transcriptional switching by the metalloregulatory MerR protein: initial characterization of DNA and mercury(II) binding activities. Biochemistry 28, 2331–2339.
[18] Outten, C.E., Outten, F.W. and O'Halloran, T.V. (1999) DNA distortion mechanism for transcriptional activation by ZntR, a Zn(II)-responsive MerR homologue in *Escherichia coli*. J. Biol. Chem. 274, 37517–37524.
[19] Rutherford, J.C., Cavet, J.S. and Robinson, N.J. (1999) Cobalt-dependent transcriptional switching by a dual-effector MerR-like protein regulates a cobalt-exporting variant CPx-type ATPase. J. Biol. Chem. 274, 25827–25832.
[20] Outten, F.W., Outten, C.E., Hale, J. and O'Halloran, T.V. (2000) Transcriptional activation of an *Escherichia coli* copper efflux regulon by the chromosomal MerR homologue, CueR. J. Biol. Chem. 275, 31024–31029.
[21] Stoyanov, J.V., Hobman, J.L. and Brown, N.L. (2001) CueR (YbbI) of *Escherichia coli* is a MerR family regulator controlling expression of the copper exporter CopA. Mol. Microbiol. 39, 502–511.
[22] Borremans, B., Hobman, J.L., Provoost, A., Brown, N.L. and van Der Lelie, D. (2001) Cloning and functional analysis of the *pbr* lead resistance determinant of *Ralstonia metallidurans* CH34. J. Bacteriol. 183, 5651–5658.
[23] Busenlehner, L.S., Cosper, N.J., Scott, R.A., Rosen, B.P., Wong, M.D. and Giedroc, D.P. (2001) Spectroscopic properties of the metalloregulatory Cd(II) and Pb(II) sites of *S. aureus* pI258 CadC. Biochemistry 40, 4426–4436.
[24] VanZile, M.L., Chen, X. and Giedroc, D.P. (2002) Allosteric negative regulation of *smt* O/P binding of the zinc sensor, SmtB, by metal ions: a coupled equilibrium analysis. Biochemistry 41, 9776–9786.
[25] Morby, A.P., Turner, J.S., Huckle, J.W. and Robinson, N.J. (1993) SmtB is a metal-dependent repressor of the cyanobacterial metallothionein gene *smtA*: identification of a Zn inhibited DNA-protein complex. Nucleic Acids Res. 21, 921–925.
[26] Huckle, J.W., Morby, A.P., Turner, J.S. and Robinson, N.J. (1993) Isolation of a prokaryotic metallothionein locus and analysis of transcriptional control by trace metal ions. Mol. Microbiol. 7, 177–187.
[27] Cavet, J.S., Meng, W., Pennella, M.A., Appelhoff, R.J., Giedroc, D.P. and Robinson, N.J. (2002) A nickel-cobalt sensing ArsR-SmtB family repressor: contributions of cytosol and effector binding sites to metal selectivity. J. Biol. Chem. 277, 38441–38448.
[28] Thelwell, C., Robinson, N.J. and Turner-Cavet, J.S. (1998) An SmtB-like repressor from *Synechocystis* PCC 6803 regulates a zinc exporter. Proc. Natl. Acad. Sci. USA 95, 10728–10733.
[29] Kuroda, M., Hayashi, H. and Ohta, T. (1999) Chromosome-determined zinc-responsible operon *czr* in *Staphylococcus aureus* strain 912. Microbiol. Immunol. 43, 115–125.
[30] Xiong, A. and Jayaswal, R.K. (1998) Molecular characterization of a chromosomal determinant conferring resistance to zinc and cobalt ions in *Staphylococcus aureus*. J. Bacteriol. 180, 4024–4029.
[31] Erbe, J.L., Taylor, K.B. and Hall, L.M. (1995) Metalloregulation of the cyanobacterial *smt* locus: identification of SmtB binding sites and direct interaction with metals. Nucleic Acids Res. 23, 2472–2478.
[32] Pennella, M.A., Shokes, J.E., Cosper, N.J., Scott, R.A. and Giedroc, D.P. (2003) Structural elements of metal selectivity in metal sensor proteins. Proc. Natl. Acad. Sci. USA 100, 3713–3718.
[33] Wu, J. and Rosen, B.P. (1991) The ArsR protein is a *trans*-acting regulatory protein. Mol. Microbiol. 5, 1331–1336.
[34] Ji, G. and Silver, S. (1992) Regulation and expression of the arsenic resistance operon from *Staphylococcus aureus* plasmid pI258. J. Bacteriol. 174, 3684–3694.
[35] Xu, C., Zhou, T., Kuroda, M. and Rosen, B.P. (1998) Metalloid resistance mechanisms in prokaryotes. J. Biochem. 123, 16–23.
[36] Smith, K. and Novick, R.P. (1972) Genetic studies on plasmid-linked cadmium resistance in *Staphylococcus aureus*. J. Bacteriol. 112, 761–772.
[37] Yoon, K.P., Misra, T.K. and Silver, S. (1991) Regulation of the *cadA* cadmium resistance determinant of *Staphylococcus aureus* plasmid pI258. J. Bacteriol. 173, 7643–7649.
[38] Endo, G. and Silver, S. (1995) CadC, the transcriptional regulatory protein of the cadmium resistance system of *Staphylococcus aureus* plasmid pI258. J. Bacteriol. 177, 4437–4441.
[39] Busenlehner, L.S., Weng, T.C., Penner-Hahn, J.E. and Giedroc, D.P. (2002) Elucidation of primary (α3N) and vestigial (α5) heavy metal binding sites in *Staphylococcus aureus* pI258 CadC: evolutionary implications for metal ion selectivity of ArsR/SmtB metal sensor proteins. J. Mol. Biol. 319, 685–701.

[40] Busenlehner, L.S., Apuy, J.L. and Giedroc, D.P. (2002) Characterization of a metalloregulatory bismuth(III) site in *Staphylococcus aureus* pI258 CadC repressor. J. Biol. Inorg. Chem. 7, 551–559.
[41] Cook, W.J., Kar, S.R., Taylor, K.B. and Hall, L.M. (1998) Crystal structure of the cyanobacterial metallothionein repressor SmtB: A model for metalloregulatory proteins. J. Mol. Biol. 275, 337–346.
[42] Schultz, S.C., Shields, G.C. and Steitz, T.A. (1991) Crystal structure of a CAP-DNA complex: the DNA is bent by 90 degrees. Science 253, 1001–1007.
[43] Pohl, E., Holmes, R.K. and Hol, W.G.J. (1999) Crystal structure of a cobalt-activated diphtheria toxin repressor-DNA complex reveals a metal-binding SH3-like domain. J. Mol. Biol. 292, 653–667.
[44] Xu, C. and Rosen, B.P. (1999) Metalloregulation of soft metal resistance pumps. In: Metals and Genetics (Sarkar, B., Ed.), pp. 5–19. Plenum Press, New York.
[45] Sun, Y., Wong, M.D. and Rosen, B.P. (2001) Role of cysteinyl residues in sensing Pb(II), Cd(II), and Zn(II) by the plasmid pI258 CadC repressor. J. Biol. Chem. 276, 14955–14960.
[46] Shi, W., Dong, J., Scott, R.A., Ksenzenko, M.Y. and Rosen, B.P. (1996) The role of arsenic-thiol interactions in metalloregulation of the *ars* operon. J. Biol. Chem. 271, 9291–9297.
[47] Turner, J.S., Glands, P.D., Samson, A.C.R. and Robinson, N.J. (1996) Zn^{2+}-sensing by the cyanobacterial metallothionein repressor SmtB: different motifs mediate metal-induced protein-DNA dissociation. Nucleic Acids Res. 19, 3714–3721.
[48] VanZile, M.L., Chen, X. and Giedroc, D.P. (2002) Structural characterization of distinct α3N and α5 metal sites in the cyanobacterial zinc sensor SmtB. Biochemistry 41, 9765–9775.
[49] Wong, M.D., Lin, Y.F. and Rosen, B.P. (2002) The soft metal ion binding sites in the *Staphylococcus aureus* pI258 CadC Cd(II)/Pb(II)/Zn(II)-responsive repressor are formed between subunits of the homodimer. J. Biol. Chem. 277, 40930–40936.
[50] VanZile, M.L., Cosper, N.J., Scott, R.A. and Giedroc, D.P. (2000) The zinc metalloregulatory protein *Synechococcus* PCC7942 SmtB binds a single zinc ion per monomer with high affinity in a tetrahedral coordination geometry. Biochemistry 38, 11818–11829.
[51] Rulisek, L. and Vondrasek, J. (1998) Coordination geometries of selected transition metal ions (Co^{2+}, Ni^{2+}, Cu^{2+}, Zn^{2+}, Cd^{2+}, and Hg^{2+}) in metalloproteins. J. Inorg. Biochem. 71, 115–127.
[52] Singh, V.K., Xiong, A., Usgaard, T.R., Chakrabarti, S., Deora, R., Misra, T.K. and Jayaswal, R.K. (1999) ZntR is an autoregulatory protein and negatively regulates the chromosomal zinc resistance operon *znt* of *Staphylococcus aureus*. Mol. Microbiol. 33, 200–207.
[53] Rensing, C., Sun, Y., Mitra, B. and Rosen, B.P. (1998) Pb(II)-translocating P-type ATPases. J. Biol. Chem. 273, 32614–32617.
[54] Xu, C., Shi, W. and Rosen, B.P. (1996) The chromosomal *arsR* gene of *Escherichia coli* encodes a *trans*-acting metalloregulatory protein. J. Biol. Chem. 271, 2427–2432.
[55] Kar, S.R., Adams, A.C., Lebowitz, J., Taylor, K.B. and Hall, L.M. (1997) The cyanobacterial repressor SmtB is predominantly a dimer and binds two Zn^{2+} ions per subunit. Biochemistry 36, 15343–15348.
[56] Kar, S.R., Lebowitz, J., Blume, S., Taylor, K.B. and Hall, L.M. (2001) SmtB-DNA and protein-protein interactions in the formation of the cyanobacterial metallothionein repression complex: Zn^{2+} does not dissociate the protein-DNA complex in vitro. Biochemistry 40, 13378–13389.
[57] Wu, J. and Rosen, B.P. (1993) Metalloregulated expression of the *ars* operon. J. Biol. Chem. 268, 52–58.
[58] Brunker, P., Rother, D., Sedlmeier, R., Klein, J., Mattes, R. and Altenbuchner, J. (1996) Regulation of the operon responsible for broad-spectrum mercury resistance in *Streptomyces lividans* 1326. Mol. Gen. Genet. 251, 307–315.
[59] Lichtarge, O., Bourne, H.R. and Cohen, F.E. (1996) An evolutionary trace method defines binding surfaces common to protein families. J. Mol. Biol. 257, 342–358.
[60] Madabushi, S., Yao, H., Marsh, M., Kristensen, D.M., Philippi, A., Sowa, M.E. and Lichtarge, O. (2002) Structural clusters of evolutionary trace residues are statistically significant and common in proteins. J. Mol. Biol. 316, 139–154.
[61] Bridgewater, B.M. and Parkin, G. (2000) Lead poisoning and the inactivation of 5-aminolevulinate dehydratase as modeled by the tris(2-mercapto-1-phenylimidazolyl)hydroborato lead complex, $\{[Tm^{Ph}]Pb\}[C10_4]$. J. Am. Chem. Soc. 122, 7140–7141.
[62] Thompson, J.D., Gibson, T.J., Plewniak, F., Jeanmougin, F. and Higgins, D.G. (1997) The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Res. 25, 4876–4882.
[63] Persson, B. (2000) Bioinformatics in protein analysis. EXS 88, 215–231.
[64] Wishart, D.S. and Sykes, B.D. (1994) Chemical shifts as a tool for structure determination. Methods Enzymol. 239, 363–392.