

Allosteric Inhibition of a Zinc-Sensing Transcriptional Repressor: Insights into the Arsenic Repressor (ArsR) Family

Gregory C. Campanello¹, Zhen Ma¹, Nicholas E. Grosseohme², Alfredo J. Guerra¹, Brian P. Ward¹, Richard D. DiMarchi¹, Yuzhen Ye³, Charles E. Dann III¹ and David P. Giedroc¹

1 - Department of Chemistry, Indiana University, Bloomington, IN 47405–7102, USA

2 - Department of Chemistry, Physics and Geology, Winthrop University, Rock Hill, SC 29733, USA

3 - School of Informatics and Computing, Indiana University, Bloomington, IN 47405, USA

Correspondence to David P. Giedroc: giedroc@indiana.edu

<http://dx.doi.org/10.1016/j.jmb.2013.01.018>

Edited by C. Kalodimos

Abstract

The molecular basis of allosteric regulation remains a subject of intense interest. *Staphylococcus aureus* CzrA is a member of the ubiquitous arsenic repressor (ArsR) family of bacterial homodimeric metal-sensing proteins and has emerged as a model system for understanding allosteric regulation of operator DNA binding by transition metal ions. Using unnatural amino acid substitution and a standard linkage analysis, we show that a His97' NH^{ε2}...O=C His67 quaternary structural hydrogen bond is an energetically significant contributor to the magnitude of the allosteric coupling free energy, ΔG_c . A “cavity” introduced just beneath this hydrogen bond in V66A/L68V CzrA results in a significant reduction in regulation by Zn(II) despite adopting a wild-type global structure and Zn(II) binding and DNA binding affinities only minimally affected from wild type. The energetics of Zn(II) binding and heterotropic coupling free energies (ΔH_c , $-T\Delta S_c$) of the double mutant are also radically altered and suggest that increased internal dynamics leads to poorer allosteric negative regulation in V66A/L68V CzrA. A statistical coupling analysis of 3000 ArsR proteins reveals a sector that links the DNA-binding determinants and the $\alpha 5$ Zn(II)-sensing sites through V66/L68 in CzrA. We propose that distinct regulatory sites uniquely characteristic of individual ArsR proteins result from evolution of distinct connectivities to this sector, each capable of driving the same biological outcome, transcriptional derepression.

© 2013 Elsevier Ltd. All rights reserved.

Introduction

Allosteric regulation is ubiquitous in biology and governs processes as diverse as signal transduction, enzymatic activity and metabolic flux, transcriptional regulation and protein degradation. Allostery is a thermodynamic phenomenon^{1,2} defined by the free-energy differences of various allosteric states. Allosteric coupling is a quantitative measure of the degree to which two ligand binding sites, for example, an allosteric effector *versus* a substrate site in an enzyme, functionally interact, and thus, it fundamentally defines biological regulation by small molecules or protein ligands in the cell. Although

allostery is clearly widespread and impacts virtually every cellular process,³ the underlying mechanisms remain elusive.

Allostery is strongly historically rooted in the static structures of oligomeric or multisubunit proteins, from which evolved concerted⁴ and sequential⁵ models of homotropic (same ligand) and heterotropic (different ligand) cooperativity. This simple picture of allostery has changed with our ability to measure residue-specific backbone and side-chain internal dynamics over a wide range of amplitudes and timescales.⁶ These thermal motions are an intrinsic property of a protein that collectively define the conformational ensemble^{7,8} and thus can be

harnessed by an allosteric ligand to shift the populations of states within the ensemble, that is, “remodel the energy landscape”.³ This may be particularly true in “dynamically driven” allostery,^{2,9} where homotropic allostery is controlled by thermal

fluctuations that occur in the absence of a large change in the average structure of the protein. This perspective of allostery is analogous to folding funnels that describe the energy landscape of protein folding. In fact, folding and allostery might be considered “two sides of the same coin”,¹⁰ since allosteric coupling is often propagated through the protein interior, through the hydrophobic core of the molecule perhaps via distinct burial modes¹¹ or via folding–unfolding equilibria present in the native-state ensemble.¹²

We have employed bacterial metalloregulatory proteins to elucidate the rules by which a specific metal ion governs allosteric activation or inhibition of operator DNA binding.^{13–17} The arsenic repressor (ArsR) family of transcriptional repressors^{18,19} is the largest family of metallosensors and conservatively numbers over 3000 members²⁰ with nearly every bacterial genome encoding at least one.²¹ The actinomycetes *Mycobacterium tuberculosis* and *Streptomyces* ssp. encode more than 10, each of which must properly function in a common cytoplasm.^{17,22} In this family of proteins, metal sites with distinct coordination chemistries and metal specificities have evolved in different places on what is essentially an unchanging, single-domain, winged helical N-(α 0)- α 1- α 2- α 3- α R- β 1- β 2- α 5-C scaffold, with each site designated by the secondary structural element from which metal coordinating residues derive, for example, α 5 or α 3N.²³ This same structural scaffold is now also known to accommodate reversible thiol-disulfide exchange as an allosteric modulator of DNA binding activity.^{24,25}

The zinc-sensing repressor *Staphylococcus aureus* CzrA regulates the expression of zinc efflux transporter²⁶ and has served, with cyanobacterial SmtB, as the prototypical α 5-subgroup ArsR family

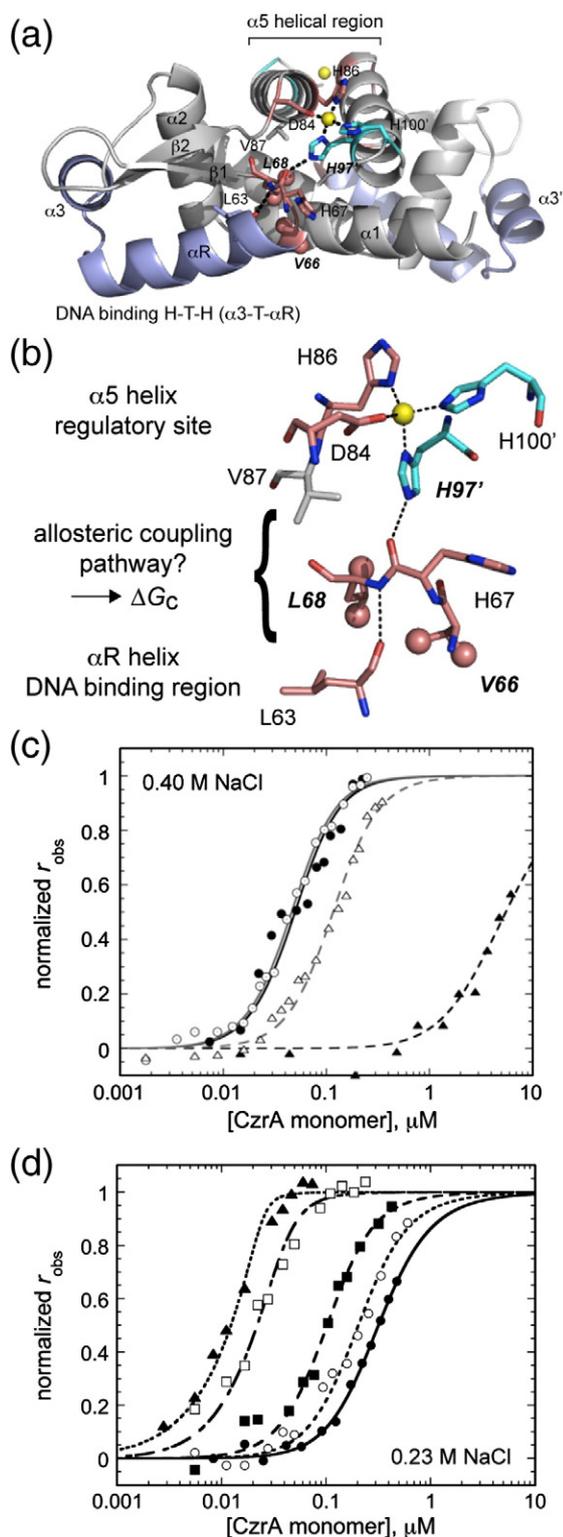


Fig. 1. Methyllhistidine-substituted (H97MeH) CzrA and cavity mutant CzrAs show poor allosteric inhibition of DNA binding upon Zn(II) binding. (a) Ribbon representation of the CzrA dimer illustrating the relative disposition of the allosteric Zn(II) sites [Zn(II) ion, yellow spheres] and the DNA binding helix–turn–helix (H–T–H) domain shaded in light blue. Residues of interest in this study are shown in stick representation. (b) Proposed allosteric coupling pathway that links each of the two ligand binding sites.^{23,27} (c) Normalized fluorescence-anisotropy-based DNA binding isotherms of H96C (closed symbols) and H97MeH (open symbols) CzrAs acquired in the absence (circles) and presence (triangles) of 10 μ M Zn(II). Curves represent the best fit using a 1:1 dimer:DNA binding model with the binding parameters compiled in Table 1. (d) Representative normalized fluorescence-anisotropy-based DNA binding curves for Zn(II)-saturated wild-type (closed circles), L68V (open circles), L68A (closed squares), V66A (open squares) and V66A/L68V (triangles) CzrAs. Parameter values are compiled in Table 2. Conditions: 10 mM Hepes, 0.23 M NaCl, 2.0 μ M ZnSO₄, pH 7.0, 25.0 °C.

Table 1. Summary of fitting parameters for H96C and H97MeH CzrAs in comparison with wild-type CzrA

CzrA	K_{Zn1} (M^{-1})	K_{Zn2} (M^{-1})	K_{apo} (M^{-1})	K_{Zn} (M^{-1})	ΔG_c (kcal/mol)
Wild type ^a	$2.5 (\pm 0.3) \times 10^{12}$	$3.4 (\pm 0.6) \times 10^{10}$	$2.7 (\pm 0.5) \times 10^{10}$	$5.7 (\pm 1.2) \times 10^5$	$6.3 (\pm 0.2)$
H96C	$\geq 10^9$	$\geq 10^9$	$2.9 (\pm 0.5) \times 10^9$	$7.4 (\pm 0.1) \times 10^5$	$4.9 (\pm 0.1)$
H97MeH ^b	$\geq 10^9$	$\geq 10^9$	$3.4 (\pm 0.3) \times 10^9$	$5.0 (\pm 0.7) \times 10^8$	$1.1 (\pm 0.1)$
	$8.0 (\pm 6.0) \times 10^{12}$	$3.0 (\pm 1.0) \times 10^{12}$			

^a Determined by UV–Vis and fluorescence spectroscopies adapted from Refs. 28 and 29.

^b Average values of two independent determinations using different concentrations of H97MeH CzrA and mag-fura-2 (top entry) or quin-2 (bottom entry) as previously described.^{28,30} Conditions: 10 mM Hepes, 0.40 M NaCl, pH 7.0.

repressor (Fig. 1a).^{15,17,31–33} We have previously shown that Zn(II)-mediated quenching of the conformational dynamics of CzrA is a key feature of negative heterotropic allosteric coupling.^{34,35} In this work, we show that the integrity of an interprotomer side chain–main hydrogen bond originating with its nonligating N^{ε2} face of a histidine ligand to the Zn(II) ion (Fig. 1a and b) is an energetically important contributor to allostery in physically connecting the zinc binding sites to the winged helical DNA binding domain. The magnitude of ΔG_c is strongly modulated by introduction of a methyl substituent on the N^{ε2} face of His97 or a “cavity” in side-chain packing³⁶ in the vicinity of this hydrogen bond, with only minor effects on zinc binding or apoprotein DNA binding affinities. A statistical coupling analysis (SCA) of ArsR family proteins^{37,38} is consistent with the idea that a common feature of allostery in this family of proteins is governed by concerted movement of the winged helical domain that pivots upon metal binding to distinct sites on the dimer that collectively stabilize the low-affinity DNA-binding state.

Results

Native chemical ligation and atom substitution in CzrA

Previous computational and structural studies of CzrA are consistent with a model in which an intersubunit His97 NH^{ε2}•••O=C His67' hydrogen bond (see Fig. 1a and b) plays an important role in linking the zinc binding and DNA binding sites in CzrA.^{28,32,35} To determine the magnitude of the degree to which this hydrogen bond contributes to allostery, we used an approach previously used in the Cu(I) sensor, *M. tuberculosis* CsoR,³⁹ to uniquely introduce a 1-methyl His residue (MeH) in place of His97. We did this using a semisynthetic intein-fusion-based strategy in which residues 1–95 of CzrA were expressed as a self-cleaving intein fusion as uniformly ¹⁵N-labeled, with residues 96–106 prepared as a synthetic peptide via solid phase peptide synthesis. This peptide incorporated a nonnative N-terminal H96C substitution to facilitate ligation to the 1–95 fragment⁴⁰ and a 1-Me-His residue at position 97. Ligation in the presence of

100 mM 2-mercaptoethanesulfonate and stepwise refolding yielded H96C/His97MeH CzrA, denoted simply as H97MeH CzrA (Supplementary Fig. 1). Functional properties of H97MeH CzrA were then compared to the parent H96C CzrA prepared by conventional site-directed mutagenesis. Examination of ¹H–¹⁵N heteronuclear single quantum coherence (HSQC) spectra of H96C and H97MeH CzrAs in the absence and presence of saturating Zn(II) reveals that both proteins adopt a wild-type fold and that the small chemical shift differences observed between the parent H96C and H97MeH CzrAs in the apo-state ($\Delta\text{ppm} \leq 0.2$ ppm) persist in the Zn₂ state (Supplementary Fig. 2a and b). Critically, Zn(II) induces the same large chemical shift perturbations upon binding to H97MeH CzrA as found in the parent H96C CzrA and, thus, is capable of undergoing wild-type-like Zn(II)-dependent conformational switching in the absence of DNA (Supplementary Fig. 2c–f). There is no evidence of recruitment of the nonnative Cys96 into the first coordination shell as revealed by inspection of the absorption spectrum of Co₂-H96C CzrA (Supplementary Fig. 2g). The Zn(II) binding affinity as measured in a chelator competition experiment for H97MeH CzrA is wild type like as well, with $K_{Zn} \approx 10^{12} M^{-1}$ (Table 1 and Supplementary Fig. 3).

The allosteric coupling free energy (ΔG_c) is a quantitative measure of the degree to which the binding of one ligand influences (positively or negatively) the binding of another ligand.¹ ΔG_c can be measured here by comparing the DNA binding affinities in the apo-bound (K_{apo}) and Zn(II)-bound (K_{Zn}) states from $\Delta G_c = -RT \ln(K_{Zn}/K_{apo})$. Application of a dimer linkage model to extract ΔG_c for wild-type CzrA reveals a value of ≈ 5 to 6 kcal mol⁻¹ at pH 7.0, 0.4 M NaCl, 25.0 °C assuming a K_{dimer} of $\approx 10^5 M^{-1}$ estimated by sedimentation equilibrium ultracentrifugation.^{28,29,33,34} To determine ΔG_c for H97MeH CzrA, we measured DNA binding affinities of H96C and H97MeH CzrAs in the presence and absence of Zn(II) (Fig. 1c and Table 1). These data reveal that apo-H96C and apo-H97MeH CzrAs bind with the same affinity to a 28-bp DNA harboring a single *czr* operator (CzrO), although with somewhat lower affinity than wild-type CzrA (Table 1). More importantly, the Zn₂ forms have vastly different affinities, with H96C CzrA strongly allosterically inhibited as expected ($\Delta G_c = 4.9$ kcal mol⁻¹), while

the binding affinity of Zn_2 H97MeH CzrA is only reduced ≈ 7 -fold relative to apo-H97MeH CzrA, corresponding to a ΔG_c of only $1.1 \text{ kcal mol}^{-1}$ (Fig. 1c). Thus, although the methyl substituent has little or no effect on conformational switching in the absence of DNA, this substitution reduces the free energy of allosteric negative regulation of DNA binding to just $\approx 20\%$ of the total.

The Zn_2 -CzrA-CzrO ternary complex adopts a hybrid conformation

In order to identify additional residues of CzrA important for allostery, we compared 1H - ^{15}N transverse relaxation optimized spectroscopy (TROSY) spectra of CzrA in the $Zn(II)$ -bound and *czr* operator (CzrO) DNA-bound allosteric end states^{15,41} with that obtained for a ternary complex formed with both negatively competing ligands bound (Supplementary Fig. 4a). Although the resonance linewidths are broad as might be expected for what is essentially a transiently formed intermediate in transcriptional derepression, the spectrum of the ternary complex appears to show three sets of resonances when compared to the component singly ligated states (Fig. 2). These include residues with cross-peaks

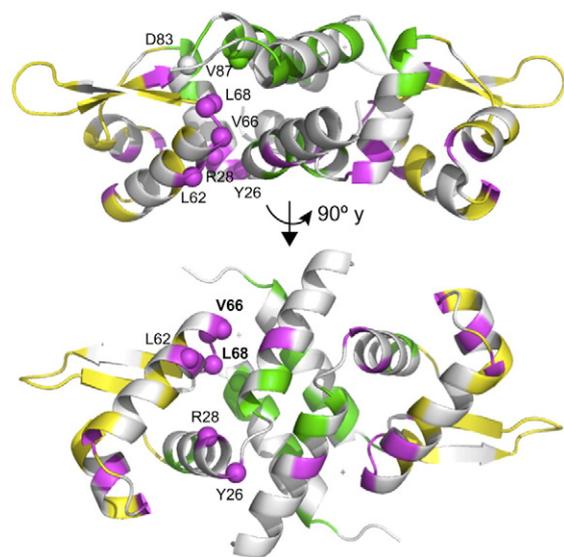


Fig. 2. Two ribbon representation views of CzrA in which the backbone amide resonances derived from an 1H - ^{15}N TROSY spectrum (see Supplementary Fig. 4a) of ternary CzrA• Zn_2 •DNA complex mimic those found in the apo-CzrA-DNA state (shaded yellow) or in Zn_2 CzrA (shaded green) or are found in a magnetically unique environment in each of the three states (shaded magenta). Backbone resonance assignments of the CzrA• Zn_2 •DNA complex were obtained by inspection and included $\approx 50\%$ coverage of the molecule (unassigned resonances shaded gray). Residues subjected to mutagenesis are indicated with C^α atoms represented as spheres.

most similar to the CzrO-bound state [localized to the $\alpha 3$ and DNA recognition (αR) helices and the β -wing; shaded yellow in Fig. 2], those most similar to the $Zn(II)$ -bound state (primarily confined to the $\alpha 1$ and $\alpha 5$ region; shaded green in Fig. 2) and a set of resonances that appear to reside in distinct chemical environments in each of the three allosteric states (shaded magenta in Fig. 2). Residues in this latter set include L62 and V66–L68 in the αR - $\beta 1$ loop region, as well as additional residues that are disproportionately localized in the more peripheral winged helical regions of the CzrA. The methyl ^{13}C and 1H resonances of V66 are also most different in the component singly ligated states (Supplementary Fig. 4b). We hypothesized that residues within this latter group might play an important role in energetically linking the two ligand binding sites⁴² and therefore targeted a subset of these residues for mutational analysis (Y26, R28, L62, V66 and L68, highlighted in Fig. 2) compared to two additional control residues not clearly in this group (D83 and V87).

Coupling free-energy analysis of CzrA mutants identifies an allosteric pathway in CzrA

We used the general approach outlined above and carried out DNA binding experiments for the apo- and $Zn(II)$ -bound mutants (Fig. 1c) and determined K_{Zn} and ΔG_c for each (Table 2). Mutant CzrAs with perturbed communication between the two ligand binding sites will show increased DNA binding affinity in the presence of $Zn(II)$ compared to wild-type CzrA (Fig. 1d). Y26F CzrA exhibits wild-type ΔG_c , while R28Q and L62V CzrAs could not be characterized as a result of a misfolding (R28Q) or weak DNA binding activity in the apo-state (L62V) (Supplementary Table 1). In contrast, Zn_2 V66A CzrA shows a very large coupling defect, binding 460-fold more tightly than Zn_2 wild-type CzrA to the CzrO DNA at 0.23 M NaCl, pH 7.0, while an L68V mutation is only modestly perturbed (Fig. 1d and Table 2). The V66/L68V double mutant CzrA has a dramatic influence on the magnitude of ΔG_c , binding DNA $\approx 12,000$ -fold more tightly than $Zn(II)$ -bound wild-type CzrA, which corresponds to a ΔG_c of $+1.1 \text{ kcal mol}^{-1}$, some $\approx 6.5 \text{ kcal mol}^{-1}$ less than wild-type CzrA (Fig. 1c and Table 2). All mutants are dimeric on the basis of gel-filtration chromatography, bind $Zn(II)$ with at or near wild-type binding affinity and retain a binding stoichiometry of 2 (see below) characterized by modest negative cooperativity of zinc binding (Table 2). The double mutant binds DNA with a similar [NaCl] dependence, SK_{obs} (Supplementary Table 2 and Supplementary Fig. 7) indicative of little or no change in the DNA binding interfacial region.⁴³ This value of SK_{obs} allowed us to obtain the binding affinity of all apoproteins at 0.23 M NaCl via linear extrapolation from conditions under which K_{obs}

Table 2. Binding parameters and coupling free energies for wild-type and selected mutant CzrAs

CzrA	$K_{Zn1} (M^{-1}) \times 10^{11}$	$K_{Zn2} (M^{-1}) \times 10^{11}$	$K_{apo} (M^{-1}) \times 10^{11a}$	$K_{Zn} (M^{-1}) \times 10^{7b}$	ΔG_c (kcal mol ⁻¹) ^c
Wild type	5.8 (± 2.1) ^d	0.48 (± 0.02) ^d	1.7 (± 0.6)	3.5 (± 0.6)	+7.6 (± 0.6)
V66A	2.6 (± 0.4) ^d	0.42 (± 0.02) ^d	0.31 (± 0.05)	1600 (± 300)	+3.7 (± 0.6)
L68V	3.0 (± 0.3) ^d	0.32 (± 0.01) ^d	0.14 (± 0.03)	7.8 (± 1.6)	+6.8 (± 0.5)
L68A	$\geq 0.01^e$	$\geq 0.01^e$	ND	30 (± 8)	+6.0 (± 0.6)
V66A/L68V	2.5 (± 0.3) ^d	0.30 (± 0.04) ^d	0.34 (± 0.07)	42,000 ($\pm 23,000$)	+1.1 (± 0.6)
V87A	$\geq 0.01^e$	$\geq 0.01^e$	ND	6.5 (± 0.9)	+6.9 (± 0.5)

^a Determined using DNA-binding fluorescence anisotropy at 0.40 M NaCl, 25.0 °C.

^b Determined using DNA-binding fluorescence anisotropy at 0.23 M NaCl, 25.0 °C. The numbers in parentheses reflect the standard error of the fitted parameters for a representative experiment of n independent experiments ($n \geq 2$).

^c $\Delta G_c = -RT \ln(K_{Zn}/K_{apo})$, determined at 0.23 M NaCl. See **Methods** for details on error propagation to obtain the standard error on ΔG_c .

^d Determined using ITC at 0.40 M NaCl, pH 7.0, from multiple experiments. The numbers in parentheses represent the SD of the mean values determined from three experiments.

^e Lower limit obtained from a mag-fura-2 competition experiment.

could be measured (Supplementary Fig. 7), thus allowing resolution of ΔG_c under these conditions.

Although a number of other single-site mutant CzrAs were characterized (Supplementary Table 1), the V66A substitution was found to be the single most detrimental substitution. For example, V66Q and V66Q/H67G CzrAs, the latter designed to mimic the Cd(II)/Pb(II) sensor CadC,⁴⁴ have physical properties indistinguishable from that of wild-type CzrA. This “cavity” defect is also specific for V66 since substitution of another Val with Ala in the same region (V87A CzrA; see Fig. 1a and b) shows a near wild-type-like ΔG_c (Table 2 and Supplementary Table 1). V66 and L68 may also function cooperatively since $\Delta \Delta G_c$ for V66A/L68V CzrA is ≈ 1.8 kcal mol⁻¹ larger than the sum of the component single-site V66A and L68V mutations, although this difference may be just inside statistical significance ($\Delta \Delta G_c = 1.8 \pm 1.1$ kcal mol⁻¹). Additionally, an L68A mutation decreases the coupling energy further than L68V, consistent with the “cavity” defect hypothesis. V66 and L68 are found in the loop between α R helix and the β -wing, physically interact and point toward the protomer core directly beneath the H97’–H67/L68–L63 hydrogen bonding network (*vide infra*). Thus, perturbation of the protein core near the hydrogen bond network results in a substantial disruption of communication between the two ligand binding sites, in much the same way as introduction of 1-methyl substitution on the N^{E2} face of His97 (Fig. 1c).

Wild-type and V66A/L68V CzrAs have identical Zn(II)-bound crystal structures

In order to determine the structural origin of this compromised allosteric linkage in V66A/L68V CzrA, we solved the crystal structure of Zn₂ V66A/L68V CzrA to 2.0 Å resolution. The global structures of wild-type³² and V66A/L68V CzrAs are essentially identical, with an r.m.s.d. of 0.38 Å over 185 C^α atoms (Fig. 3a and Supplementary Table 3 for

structure statistics); in addition, the first coordination shell around the Zn(II) ion and the integrity of the hydrogen bonding pathway is intact and nearly indistinguishable in the double mutant (Fig. 3b). Likewise, examination of an ¹H–¹⁵N HSQC spectrum of the Zn(II)-bound double mutant reveals largely local perturbations of the structure immediately around the site of the substitution relative to Zn₂ wild-type CzrA (Supplementary Fig. 5). These structural studies are fully consistent with very similar zinc and apoprotein DNA binding affinities of this mutant relative to wild-type CzrA (Table 2). However, closer inspection reveals the presence of a significant cavity indicative of poorer packing in the protein core for V66A/L68V CzrA relative to wild-type CzrA (Fig. 3c and d). We hypothesize that this poorer packing directly controls the magnitude of ΔG_c .

Energetics of Zn(II) binding to wild-type versus V66A/L68V CzrAs and other cavity mutants

We next carried out a series of isothermal titration calorimetry (ITC) experiments in order to determine if the poorer packing of the double mutant becomes manifest in the underlying energetics of Zn(II) binding to the dimer.³⁴ Here, we took advantage of the fact that the zinc binding affinity and structure of the first coordination sphere in the mutant are identical with that of wild-type CzrA (Fig. 3). Zn(II) binding experiments with wild-type CzrA give thermodynamic values comparable to those previously reported, although not corrected here for linkage to ligand deprotonation upon metal binding since this contribution will be identical in all cases (Fig. 4a and b and Supplementary Table 4).³⁴ Comparison of V66A, L68V and V66A/L68V mutant CzrAs shows that these proteins bind two equivalents of Zn(II) per dimer with high affinity and measurable negative homotropic cooperativity, resulting in nearly identical free energies of Zn(II) binding (ΔG^1) (Fig. 4a and b and Table 2). This is consistent with the fact that all mutants are known or expected to have substantially

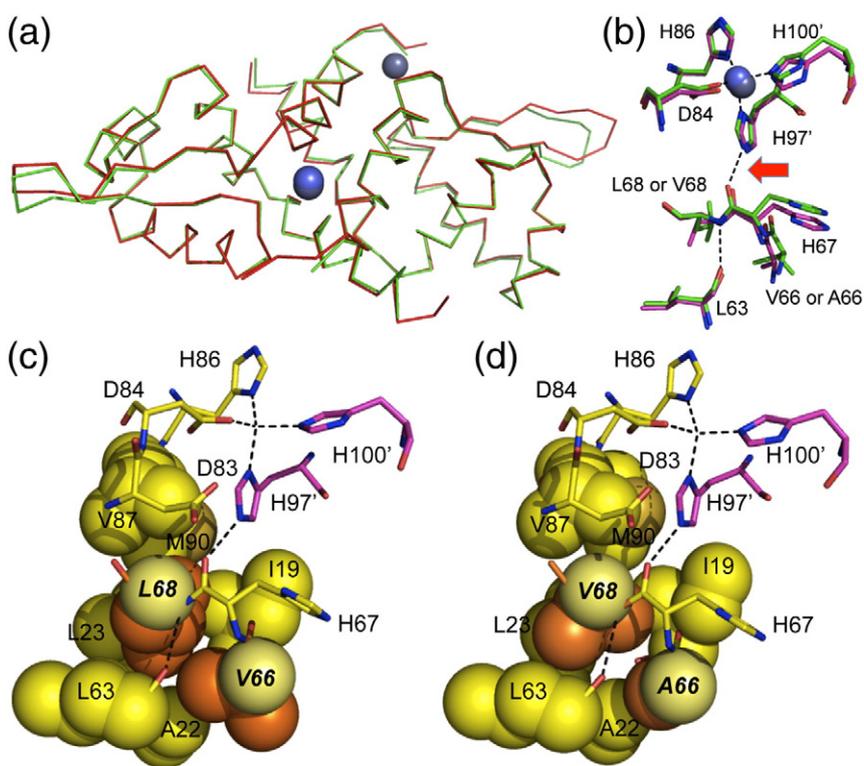


Fig. 3. The structures of Zn(II)-bound wild-type and V66A/L68V CzrAs are globally identical. (a) Global C α wireframe superposition of wild-type (green) and V66A/L68V (red) CzrAs, with the positions of the zinc atoms shown in slate or darker slate. (b) Detailed superposition of the first and selected second coordination shell region of wild-type (green) and V66A/L68V (magenta) CzrAs. The interprotomer H97'–H67 (highlighted by the red arrow) and L68–Leu63 hydrogen bonds are indicated by the dashed lines. (c and d) Spacefilling representations of the van der Waals packing region in the vicinity of these hydrogen bonds in Zn₂ wild-type CzrA (c) and V66A/L68V CzrA (d). Side chains of residues 66 and 68 and C α atoms are shaded orange and pale yellow, respectively.

identical first coordination shells (Fig. 3) and the effect of solvent release from the metal will be identical in each case.

Strikingly, the underlying energetics reveal that V66A/L68V CzrA has a significantly smaller enthalpy of Zn(II) binding, ΔH^{\ddagger} , than wild-type CzrA (Fig. 4a and b). This smaller enthalpy change is nearly precisely compensated by a more favorable entropy term for Zn(II) binding to V66A/L68V CzrA ($-T\Delta S^{\ddagger}$). This result is as anticipated for a cavity mutant CzrA containing fewer van der Waals contacts in the protein core (Fig. 3d) resulting in increased internal dynamics (Fig. 4b and Supplementary Table 4). The same trend is observed for each of the two-component single mutants, with the effect of the single V66A substitution larger than that of the L68V substitution (Fig. 4b and Supplementary Table 4).

We next examined the energetics of Zn(II) binding to the CzrA dimer–DNA complex formed by wild-type *versus* V66A/L68V CzrAs (Fig. 4c and Supplementary Table 4). Strikingly, the ΔH contribution to Zn(II) binding to the V66A/L68V complex is easily distinguished from that of the wild-type CzrA–DNA complex, with the two isotherms of nearly opposite sign (Fig. 4c). Propagating these energetics of Zn(II) binding to obtain ΔH_c^i and $-T\Delta S_c^i$ for each i th zinc binding step reveals a less positive ΔH_c and a less positive ΔS_c , manifested largely in the second zinc binding step, that is, in ΔH_c^2 and $-T\Delta S_c^2$, as expected if the distinct energetics of the Zn(II) binding isotherms (Fig. 4a–c) propagate to the energetics of

heterotropic coupling (Fig. 4d). These findings reveal that poorer side-chain packing observed crystallographically (Fig. 3d) and implied by the underlying energetics of Zn(II) binding to V66A/L68V CzrA relative to wild-type CzrA (Fig. 4) directly impact the magnitude and underlying energetics (ΔH_c , ΔS_c) of ΔG_c (Table 2). Since a major contributor to zinc-dependent allosteric inhibition of DNA binding is global “stiffening” of the dimer, much of which also occurs on the second Zn(II) binding step,^{34,35} decreased quenching of the conformational dynamics associated with the allosterically inhibited Zn(II)-bound state functions to reduce the effectiveness of negative allosteric regulation by Zn(II) on V66A/L68V CzrA function.

SCA of ArsR family repressors

Having determined that V66 and L68 may function cooperatively in controlling the magnitude of ΔG_c in CzrA, we then asked if these two residues are evolutionarily pairwise coupled in ArsR family repressors. To address this, we carried out a multiple sequence alignment-based SCA^{37,38} of 3000 ArsR family repressors (Fig. 5) and mapped the results of this analysis onto the structure of Zn(II)-bound CzrA (Fig. 6a). We find that V66 and L68 are not strongly conserved, nor are they strongly evolutionarily coupled (see Fig. 5). Despite this, V66 and L68 appear to physically connect the $\alpha 5$ allosteric sites with a contiguous network of coupled residues, or

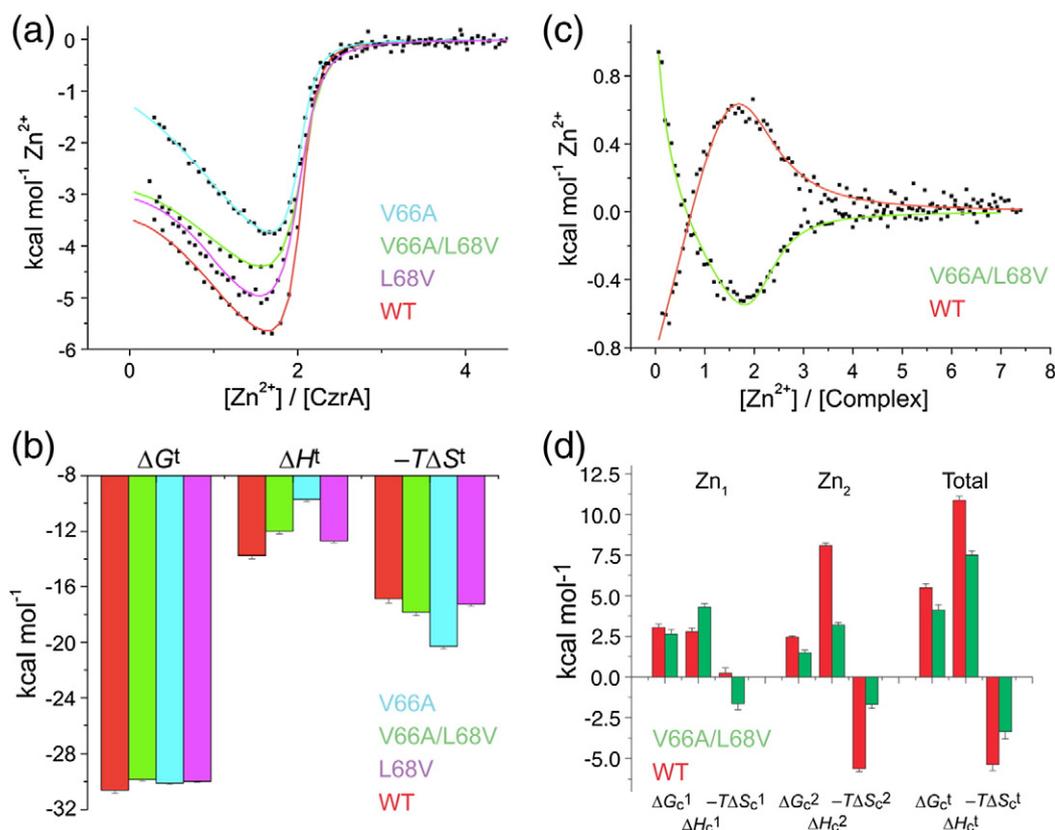


Fig. 4. Analysis by ITC of Zn(II) binding to various apo-CzrAs and the apo-CzrA–DNA complex. (a) Representative titrations of Zn(II) binding to wild-type (WT), V66A, L68V and V66A/L68V CzrA dimers. (b) Thermodynamic parameters obtained for Zn(II) binding to CzrAs obtained from experiments such as those shown in (a). t , total, obtained by summing parameters obtained for the first (Zn_1) and second (Zn_2) Zn(II) binding steps. (c) Representative titrations of Zn(II) into CzrO complexes formed by wild-type (WT) and V66A/L68V CzrAs. These titrations correspond to the “top” and “bottom” of the heterotropic coupling equilibrium that defines this system.³⁴ (d) Graphical illustration of the coupling energetics of wild-type *versus* V66A/L68V CzrAs derived from the first (Zn_1) and second (Zn_2) binding steps. t , total. Conditions: 38 μ M dimer, 3 mM NTA as a zinc competitor or 38 μ M dimer–DNA complex, with 1 mM NTA as competitor in 10 mM Hepes, 0.4 M NaCl, pH 7.0, 25.0 °C.

sector, that extends from the $\alpha 1$ helical region to the $\alpha 3$ - αR - β -wing region, largely along the DNA-binding interface, that collectively encircles key DNA residues on the αR helix (S57 and H58),¹⁵ which themselves do not covary (Fig. 5a). Residues in the $\alpha 5$ helix do not strongly covary either, as expected for a family of repressors that respond to a range of metal and nonmetal effectors that bind to distinct sites.^{21,23} Indeed, analysis of the $\alpha 5$ helical region reveals a near complete absence of interacting residues (Fig. 6a and b). This sequence-based identification of a sector residues bears strong similarity to the subset of “hybrid-state” and “DNA-binding-state” residues determined by analysis of the experimental ¹H–¹⁵N TROSY spectrum of Zn₂-CzrA•CzrO (Fig. 2).

This sector may allow the more peripheral winged helical region to move in a concerted fashion with respect to the $\alpha 1$ – $\alpha 5$ core in response to inducer recognition to distinct sites on the ArsR scaffold that

“moves” and/or “stiffens” the DNA-binding interface that ultimately inhibits DNA binding (Fig. 6c and d). Consistent with this, a number of residues in CzrA in or near this sector exhibit dynamical quenching upon Zn(II) binding.¹⁵ Previous studies of other ArsR family sensors are also consistent with this model. Pb(II)/Cd(II)-sensing CadCs and canonical As(III)/Sb(III)-sensing ArsRs employ a pair of metal-binding Cys residues arranged as a CXC motif in the $\alpha 3$ helix, the C-terminal of which aligns with strongly coupled residue G43 (Fig. 5c) and is a known allosteric ligand.⁴⁴ This metal site is surrounded by sector residues (Fig. 6c). The $\alpha 4$ Cd(II)/Pb(II) sensor *M. tuberculosis* CmtR binds metal ions to a pair of cysteines in a CXXXC motif, the C-terminal of which aligns precisely with V66 in CzrA.^{22,45,46} This metal site is also in physical contact with sector residues (Fig. 6d), and Cd(II) binding to CmtR has also been reported to quench the internal dynamics of CmtR particularly across

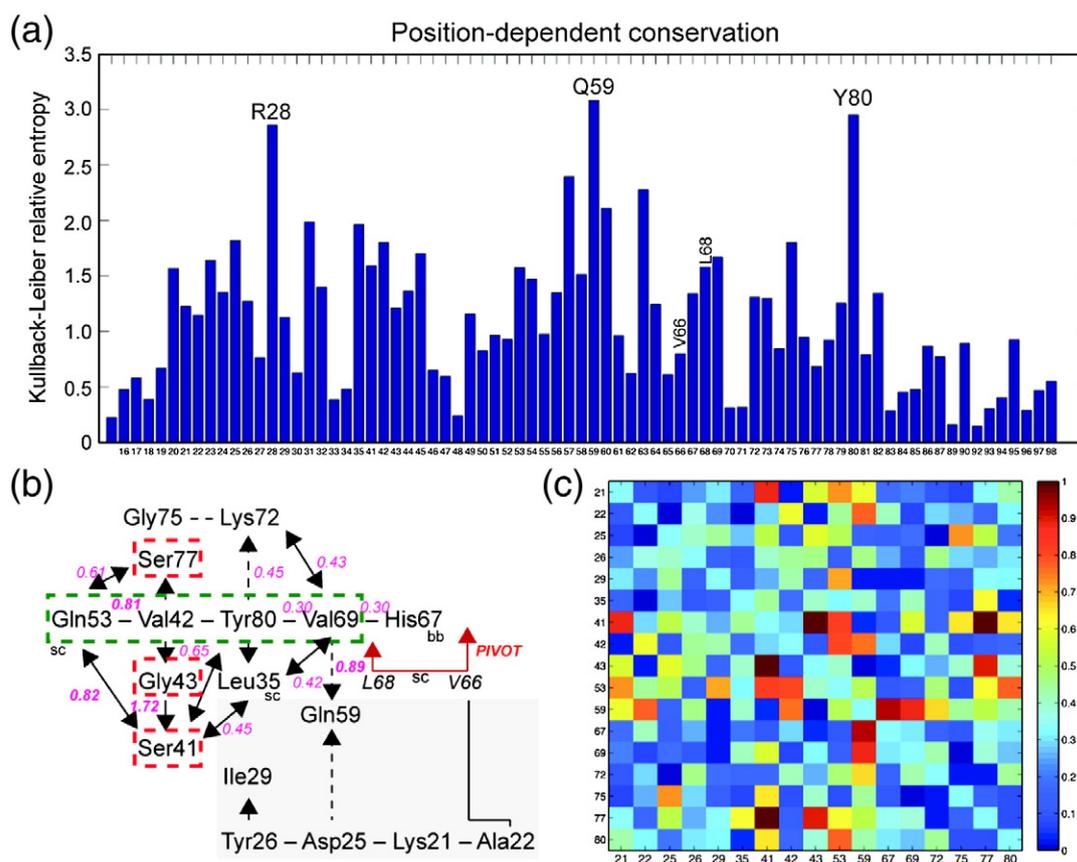


Fig. 5. SCA of ArsR family repressors. (a) Residue-specific conservation plotted as relative entropy of residue positions 15–35, 41–87, 89–90 and 92–98 that correspond to CzrA residue positions. Gaps correspond to gaps in the multiple sequence alignment. Selected residues with the highest conservation are highlighted with CzrA residue type and number. (b) Schematic illustration of residues in physical contact arranged as two parts of the same network of 17 coupled residues (upper left; lower right, shaded gray), with selected pairwise statistical coupling energies shown in pink (see Methods). Three key residues that are strongly pairwise coupled to one another are shown boxed in red. The green box represents four residues in direct physical contact that connect β -wing and more peripheral regions of the DNA binding site to the pivot point defined by V66/L68. (c) Normalized heat map of all (17 \times 17) significantly pairwise coupled interactions scaled from 0 (no coupling) to 1 (strong coupling). The diagonal is false-colored light blue and represents the mean value of the correlation map. Residues pairs 41/43 and 41/77 exhibit the strongest pairwise couplings. It is interesting to note that the CXC Cys pair in Pb(II)/Cd(II)-sensing CadCs (which correspond to residue positions 41/43 in CzrA) are ligands to the Cd/Pb ion.⁴⁴

the dimer interface.⁴⁶ Finally, in the related nickel sensor, *M. tuberculosis* NmtR, Ni(II) binding to the $\alpha 5$ sites induces conformational exchange broadening to distal residues corresponding to Q53, H67 and G75 each of which are part of the deduced sector (Fig. 5b and c).¹⁷ These data suggest that a region defined by allosteric residues Val66 and Leu68 in ArsR family homodimers has evolved as an allosteric “hot-spot” on which the DNA binding domain is capable of pivoting or damping in response to the binding of an allosteric metal ion (Fig. 6e).

Discussion

In this work, we deconstruct the heterotropic allosteric linkage free energy into component

thermodynamic driving forces in the paradigm ArsR-family zinc-sensing repressor *S. aureus* CzrA. A comprehensive analysis of heterotropic allostery involves structural, dynamical and thermodynamic interrogation of all four allosteric states of a two-ligand system, which in this case includes the ligand-free (apo) state, the two singly ligated states (Zn_2 -CzrA, CzrA•DNA) and the doubly ligated (Zn_2 •CzrA)•DNA state and coupled equilibria between states as a function of ligand activity.⁴² High-resolution crystallographic structures of apo-CzrA and Zn_2 CzrA are available, as solution structural models of Zn_2 and DNA-bound CzrAs.^{15,32,35} A major conclusion from the previous work is that zinc binding narrows the conformational ensemble and “freezes out” a low-affinity DNA conformation that is unable to undergo the large quaternary structural

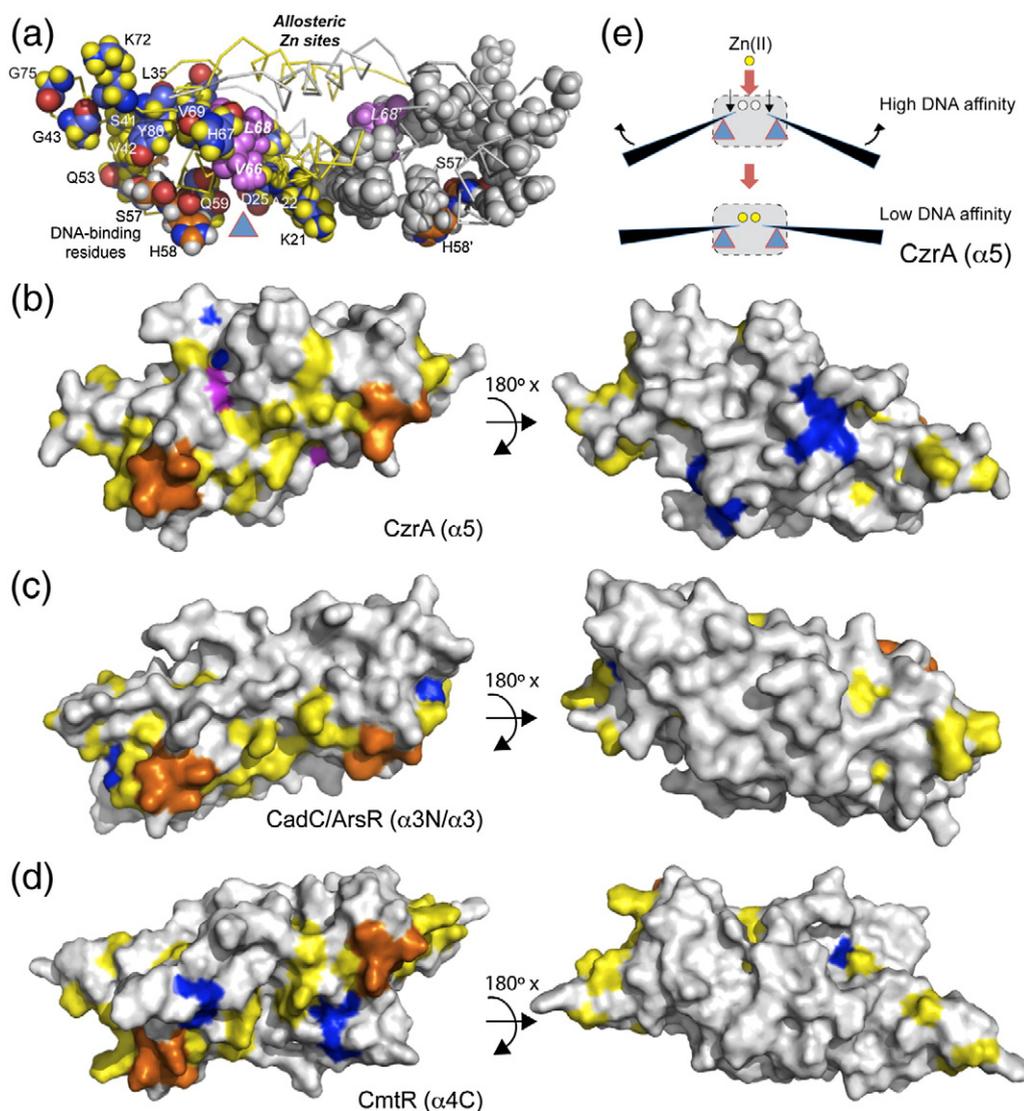


Fig. 6. Structural representation of an SCA of *ArsR* family repressor reveals a coupled network connected to an allosteric "hot-spot". (a) Network of coupled residues shown in spacefill on both protomers (CPK coloring shown in left protomer with protons shaded yellow and carbons slate; gray on right protomer). Key DNA binding residues (S57 and H58)¹⁵ shaded CPK with protons white and carbon atoms shaded orange; V66 and L68 define an allosteric "hot-spot" with all atoms shaded magenta. *CzrA* is shown in its open or "flat" allosterically inhibited Zn(II)-bound low DNA binding affinity state. (b–d) Surface representations of (b) *CzrA* (1R1V); (c) *CadC/ArsR* (1U2W) and (d) *CmtR* (2JSC) highlighting a sector of co-evolving residues (shaded yellow), the allosteric metal site chelates (shaded blue), major energetic determinants of the DNA binding domain (shaded orange; corresponding to residues 54–58 in *CzrA*)¹⁵ and the V66/L68 in *CzrA* (shaded magenta). View from the DNA-binding interface is shown on the left, with a view from the top of the molecule shown on the right. (e) Schematic rendering of the conformational transition from a high to a low DNA binding conformation with the sector schematized as a lever, and the V66/L68 "pivot" point indicated by the triangle with the $\alpha 5$ metal sites shown as circles.

change required to form a high-affinity complex with the DNA operator.¹⁵ Similar conclusions have been reached for the Cd(II)/Pb(II)-selective *ArsR* family sensor, *M. tuberculosis* *CmtR*.⁴⁶ Here, we use the NMR fingerprint region of the doubly ligated ternary complex to identify candidate residues that control the magnitude of the heterotropic allosteric coupling

free energy, ΔG_c , and in so doing uncover an allosteric hot-spot in *CzrA*.

As in other studies of protein–ligand interactions, for example, CAP•cAMP₂•DNA⁴⁷ and calmodulin–peptide complexes,⁴⁸ we find that Zn(II) binds to allosteric mutants of *CzrA* with similar free energies, but very different component enthalpies and entropies of

binding and thus represent clear examples of enthalpy–entropy compensation in biomolecular interactions.⁴⁹ The striking departure reported here is that enthalpy–entropy compensation of ligand Zn(II) binding leads to an *uncoupling* of Zn(II) and DNA binding to varying degrees, resulting in a *loss* in biological regulation. This is true for the two cases discussed here (methyl substitution and cavity mutants), in which the functional defect lies squarely in the magnitude of ΔG_c with little or no effect on the binding affinity of Zn(II) or operator DNA, thus allowing resolution of underlying entropic and enthalpic driving forces to the coupling free energy (ΔH_c , $-T\Delta S_c$) (Fig. 4). Since unfavorable entropic driving force is a major contributor to negative allosteric regulation of DNA binding,^{34,35} we propose that this contribution is reduced by introduction of poor side-chain packing (as reported by decreased ΔH_c) and concomitant increased mobility (increased $-T\Delta S_c$) in the immediate vicinity of a key hydrogen bonding interaction, which itself plays a major role in controlling the magnitude of ΔG_c (Fig. 4).

It is formally possible that some of the CzrA mutants characterized here perturb the dimer stability (K_{dimer}) of CzrA. V66 is not part of the dimer interface but packs against the $\alpha 1$ helix of the same protomer, which forms the $\alpha 1$ - $\alpha 1'$ - $\alpha 5$ - $\alpha 5'$ helical dimer core (Fig. 2). If so, this will have an impact on our determination of ΔG_c , since the DNA binding constants K_{apo} and K_{Zn} are determined under conditions of strong linkage to the monomer–dimer equilibrium using fluorescence anisotropy methods (see Supplementary Fig. 6). Although the zinc dependence of K_{dimer} is known for wild-type CzrA³³ (see [Methods](#)) and we did not determine this for mutant CzrAs, we consider significant perturbation of K_{dimer} unlikely for several reasons. First, the affinity of the apo-CzrAs for DNA are all within a factor of 10 of wild-type CzrA, and even the most strongly allosterically perturbed double mutant, V66A/L68V, is within a factor of ≈ 5 of wild type and is characterized by a very similar salt dependence of K_{apo} (Supplementary Fig. 7). If there were a large change in dimer stability, this would result in a far more greatly perturbed K_{apo} and the mutants would not chromatograph as dimers by gel-filtration chromatography (see [Methods](#)). To significantly impact our resolution of ΔG_c , then zinc binding would have to *stabilize* the mutant dimers far more strongly than the wild-type dimer (simulations suggest by $>10^5$ -fold), which would then lead to tighter DNA binding and weaker apparent allosteric negative regulation by zinc. This possibility seems remote given that the dimer interface is basically unchanged in V66A/L68V CzrA (Fig. 3) and a *smaller* rather than larger ΔH component is associated with zinc binding to the mutants relative to wild-type CzrA, with similar K_{Zn} (Table 2). In any case, even some perturbation of

the monomer–dimer equilibrium would formally remain an allosteric effect and naturally derives from the fact that cooperativity and folding are intimately interconnected.¹⁰

Two limiting models have been put forth in an effort to explain the physicochemical linkage of two ligand binding sites in classical heterotropic allostery in structural or dynamical terms. These models differ on the presence^{50–52} or absence^{53,54} of a preferred or dominant pathway⁵⁵ of allosteric connectivity between ligand binding sites. In CzrA, we show here that allostery in CzrA hinges on the integrity of a key interprotomer side chain–main chain hydrogen bond.^{51,55} The coordinate covalent nature of transition metal–ligand bonds may well establish strong directionality into this allosteric network. Blocking formation of this hydrogen bond chemically or introduction of a cavity (but not a larger side chain) just below this hydrogen bond quantitatively and specifically reduces the magnitude of ΔG_c . This latter effect is selective and focused on V66. However, it is important to recognize that ΔG_c is not zero in any CzrA mutant, consistent with the idea that otherwise non-dominant pathways can make a contribution when a major one is disrupted.⁵⁵

Our elucidation of what would appear to be a compact pathway of allosteric communication between Zn(II) and DNA binding sites in CzrA stabilized by a cooperative network of van der Waals interactions involving V66 and L68 does not become manifest in a standard pairwise covariation or SCA carried out on a large family of ArsR family sensors (Figs. 5 and 6), as had been found in previous systems examined by SCA.^{38,56,57} This is perhaps not so surprising given the involvement of both main chain and side atoms in allosteric hydrogen bonding in CzrA (Fig. 1b), the small subset of interactions that control much of the magnitude of ΔG_c , and the breadth of distinct subfamilies of sensors with different regulatory ligand binding sites and specificities.²³ The major finding from this analysis is the identification of a sector of interconnected residues that immediately suggests a way in which the more peripheral elements of the ArsR fold that control DNA binding affinity, that is, the recognition helices and β -wing, move in a concerted fashion (Fig. 6a). The allosteric “hot-spot” identified here centered on V66 and L68 then simply defines a pivot point on which the entire DNA-binding interface is remodeled upon metal binding (Fig. 6b–d), rather than an allosteric pathway.⁵⁷ Rapid evolution of new inducer specificity on this simple scaffold would then become a matter of evolving distinct connectivities to this sector of co-evolving residues (Fig. 6b–d), the end result of which is to reposition the DNA-binding levers and/or quench the dynamics in a way that ultimately stabilizes a low-affinity “open” DNA binding conformation. The recent structures of two

ArsR family proteins that are known or projected to exploit thiol-disulfide chemistry to stabilize “closed” and “open” conformations,^{15,35} for example, *Bacillus subtilis* HypR²⁴ and *Xylella fastidiosa* BigR,²⁵ are generally consistent with this model as a common feature that underlies ArsR family protein allosteric function.

An alternative view is that this sector comprises a classically defined allosteric network that links the energetically important but more peripheral components of the DNA-binding interface, for example, the N-terminus of the α R helix and β -wing tips¹⁵ to the core of the dimer defined largely by the α 1– α 5 helical bundle, to achieve a quaternary structural conformational change. Evolution of distinct allosteric sites in the ArsR family would then be possible provided only that these sites physically connect to any point along this sector.⁵⁵ In this view, the allosteric hot-spot defined here becomes more of an extension of the allosteric ligand binding site itself (Fig. 6b) that mediates a physical connectivity to a preexisting allosteric network. This view is perhaps more consistent with the spectral characteristics of the Zn₂-CzrA•CzrO complex (Fig. 2), which reveal that candidate allosteric residues are interspersed with DNA-binding-like residues or, alternatively, serve to physically connect the Zn(II) and DNA binding sites on the dimer. The overrepresentation of these residues at the termini of secondary structural elements (see Fig. 2) may help drive this connectivity within the sector in a concerted manner.

The compact, essentially single-domain architecture of CzrA and related ArsR proteins¹⁷ does not allow us to distinguish between these two general views. On the other hand, a survey of known, structurally defined regulatory sites on the ubiquitous ArsR repressor scaffold^{21,23} suggests an overrepresentation of allosteric sites positioned roughly on opposite ends of the sector defined here (Fig. 6), sometimes including determinants from both ends.^{17,46,58} Evolution of distinct metal coordination chemistries and chemical reactivities characteristic of each functional ArsR subfamily would then result in new biological specificities.

Methods

Protein production

Construction and purification of H97MeH CzrA using native chemical ligation

The C-terminal peptide of CzrA (residues 96–106, H96 substituted with Cys) was synthesized with incorporation of MeH (1-methylhistidine) obtained as a Boc-derivative (Bachem, Torrance, CA) at residue 97 using Boc-based solid phase peptide synthesis. The DNA sequence encoding the N-terminal peptide (residues 1–95) was

cloned into the pTXB1 vector (New England Biolabs) between NdeI and SpeI restriction sites in-frame to a C-terminal intein fusion. The CzrA 1-95-intein fusion was expressed in *Escherichia coli* BL21(DE3) and, after sonication in Buffer C [25 mM Tris, 0.5 M NaCl and 2 mM tris(2-carboxyethyl)phosphine (TCEP), pH 8.0], was found to remain in the low speed lysis pellet. This pellet was then resuspended in Buffer C containing 7 M urea and refolded by stepwise decreasing the urea concentration in Buffer C. The resultant soluble fraction of CzrA 1-95-intein was cleaved with the addition of 100 mM mercaptoethanesulfonate 2-mercaptoethanesulfonate (Sigma, St. Louis, MO) with CzrA 1-95-thioester further purified on a C18 reverse phase column by running a 0–75% acetonitrile gradient in 0.1% trifluoroacetic acid. Fractions containing CzrA 1–95 thioester were pooled and concentrated to ~1 mL and ligated to the C-terminal peptide using conditions analogous to those as previously described.³⁹ The resultant H96C/H97MeH CzrA (denoted simply as H97MeH CzrA) was further purified on a μ RP (GE Healthcare, Piscataway, NJ) reverse phase column under denaturing conditions and finally refolded into Buffer P by stepwise increasing pH (10 mM Hepes and 0.4 M NaCl, pH 7.0) with 1 mM TCEP.

Purification of mutant CzrAs

Overexpression plasmids encoding mutant *S. aureus* CzrAs were constructed by site-directed PCR-based QuikChange mutagenesis using pET3a-CzrA as template²⁸ with plasmid integrity verified using DNA sequencing. The proteins were expressed in *E. coli* BL21(DE3) at 37 °C on M9 minimum medium containing 100 mg/mL ampicillin supplemented with ¹⁵NH₄Cl as the sole nitrogen source¹⁵ or on LB medium containing 100 mg/mL ampicillin and purified using published procedures.^{28,33} For H96C CzrA, 2 mM dithiothreitol was added to all the buffers used during the purification. Purified H96C CzrA was extensively dialyzed anaerobically against Buffer P (10 mM Hepes and 0.4 M NaCl, pH 7.0). The protein concentration was determined using $\epsilon_{280\text{nm}}=4470 \text{ M}^{-1} \text{ cm}^{-1}$, and the mole equivalent of free reduced thiol was determined by the dithionitrobenzoic acid assay to be 0.9 (1.0 expected). All other mutant CzrAs were purified using the same purification as wild-type CzrA,³³ dialyzed extensively and confirmed to contain less than 0.05 mol equivalents of Zn(II) by atomic absorption spectroscopy. All chromatographed as dimers by gel-filtration chromatography.

Co(II) and Zn(II) binding to H96C and H97MeH CzrAs

All metal binding experiments were conducted on a Hewlett-Packard model 8452A spectrophotometer. CoCl₂ titrations with 100 μ M CzrA monomer (50 μ M dimer) were carried out in Buffer P anaerobically as previously described.^{33,59} Two zinc chelator indicator dyes were used for zinc competition experiments: Quin-2 ($K_{\text{Zn}}=2.7 \times 10^{11} \text{ M}^{-1}$ at pH 7.0 and 25 °C⁶⁰) and mag-fura-2 ($K_{\text{Zn}}=5.0 \times 10^7 \text{ M}^{-1}$).⁶¹ For mag-fura-2 Zn(II) titrations, ZnSO₄ was titrated into a mixture of 2.4 μ M mag-fura-2 and 1.7 μ M CzrA monomer in Buffer P containing 0.1 mM

TCEP. The excitation spectrum from 265 to 455 nm with $\lambda_{em}=497$ nm was measured after each i th addition. Fluorescence intensities at 325 and 379 nm were plotted against total Zn(II) concentration and the data were simultaneously fitted to a simple competition model using Dynafit⁶² as previously described.⁶³ Quin-2 experiments were carried as described previously³⁰ with ZnSO₄ titrated into a mixture of 1.7 μ M CzrA monomer and 1.5 μ M quin-2 in Buffer P containing 0.1 mM TCEP. The presence of TCEP does not interfere with Zn(II) binding in these experiments.⁶⁴

Zinc binding experiments for all other mutant CzrAs

Briefly, mag-fura-2 Zn(II) binding competition assays were carried out in 10 mM Hepes, 400 mM NaCl (pH 7.0), 10–15 μ M protein dimer and 10–15 μ M mag-fura-2, using $K_{Zn}=5.0 \times 10^7$ M⁻¹.³³ Under these conditions, mag-fura-2 shows no Zn(II) binding competition with CzrA, and therefore only a lower limit is reported. Co(II) was also titrated into all CzrAs until metal binding saturation and the optical spectrum was recorded to confirm the presence of a tetrahedral metal coordination environment.

Fluorescence-anisotropy-based DNA binding experiments and ΔG_c calculations

The DNA binding affinities of H96C and H97MeH were measured in 10 mM Hepes, 0.40 M NaCl and 2 mM DTT (pH 7.0) with 10 μ M Zn(II) or 1 mM ethylenediaminetetraacetic acid present in the solution. We used 4 nM fluorescein-labeled 28-bp native CzrO DNA operator with the 12-2-12 inverted repeat underlined (5'-FL-TAATATAT-GAACAAATATTCA GATGAAA-3') (FL, fluorescein). For the apo-CzrA–DNA binding experiments, the data were fit with Dynafit⁶² using a dimer linkage 1:1 dimer:DNA binding model¹⁵ with the dimerization constant fixed at $K_{dimer}=1.7 \times 10^5$ M⁻¹.³³ The initial anisotropy was fixed to the measured value (r_0) for the free DNA, with $r_{complex}$, the anisotropy of the protein–DNA complex, optimized during a nonlinear least squares fit (see Supplementary Fig. 6 for Dynafit⁶² script file).^{29,33} For other mutant CzrAs, the same 28-bp oligonucleotide was used at 10 nM concentration with the data fit in the same way.³³ All Zn(II)-bound experiments employed protein stocks that were preloaded 1:1 with Zn(II) as titrant, with an additional 3 μ M Zn(II) in the fluorescence cuvette in a buffer containing 10 mM Hepes and 0.40 M NaCl (pH 7.0) unless otherwise noted, and data were fit in the same manner with $K_{dimer}=4.5 \times 10^5$ M⁻¹.³³ ΔG_c was calculated from $\Delta G_c = -RT \ln(K_{Zn}/K_{apo})$ for wild-type, H96C and H97MeH CzrAs at 0.40 M NaCl (see Table 1), with the error in ΔG_c propagated from the square root of the sum of the squares of the standard deviation (SD) of the mean value of K_{apo} and K_{Zn} obtained from two or more experiments. For all other CzrA mutants, K_{apo} at 0.23 M NaCl was obtained via linear extrapolation⁴³ of a plot of $\log K_{apo}$ versus \log [NaCl] for wild-type and V66A/L68V CzrAs with K_{apo} measured at various [NaCl] (Supplementary Table 2 and Supplementary Fig. 7). This gave K_{apo} values of 1.3×10^{13} M⁻¹ and 2.9×10^{12} M⁻¹ for wild-type and V66A/L68V CzrAs, respectively, at 0.23 M NaCl. K_{Zn} was measured by direct titration at 0.23 M NaCl (see Fig. 1d) and ΔG_c was then calculated from, $\Delta G_c = -RT \ln(K_{Zn}/K_{apo})$. In order to esti-

mate the associated error on ΔG_c , we calculated minimal and maximal ΔG_c values, $\Delta G_{c,min}$ and $\Delta G_{c,max}$, respectively, by first determining $K_{apo,min}$ and $K_{apo,max}$ by incorporating the associated errors on the slope and y-intercept obtained from the linear fit of $\log K_{apo}$ versus \log [NaCl] for wild-type and V66A/L68V CzrAs (see Supplementary Fig. 7). $K_{Zn,min}$ and $K_{Zn,max}$ were determined by using the standard error of the fit giving $\Delta G_{c,min} = -RT \ln(K_{Zn,max}/K_{apo,min})$ and $\Delta G_{c,max} = -RT \ln(K_{Zn,min}/K_{apo,max})$. The standard error reported for ΔG_c is the average error of the difference of $\Delta G_{c,max}$ minus ΔG_c and ΔG_c minus $\Delta G_{c,min}$ for wild-type and V66A/L68V CzrAs. For all other single mutants, K_{apo} was calculated as the average value of the extrapolated values for K_{apo} for wild-type and V66A/L68V CzrAs at 0.23 M NaCl (8.0×10^{12} M⁻¹). ΔG_c and the associated error was then calculated in the same manner, except $K_{apo,min} = K_{apo,V66A/L68V}$ and $K_{apo,max} = K_{apo,WT}$.

NMR spectroscopy

NMR spectra were acquired on a Varian VNMRs 600-MHz spectrometer equipped with a cryoprobe in the METACyt Biomolecular NMR Laboratory at Indiana University. NMR samples contained ≈ 0.25 mM ¹⁵N-labeled H96C CzrA or 0.07 mM ¹⁵N-labeled H97MeH CzrA in 10 mM *d*₁₃-4-morpholineethanesulfonic acid, 50 mM NaCl and 2 mM DTT with or without 1.1 monomer mole equivalents of Zn(II) added (pH 6.0). ¹H–¹⁵N HSQC or TROSY spectra were acquired at 40 °C as described previously.²⁸ All spectra were processed and analyzed using NMRPipe and Sparky^{65,66} with resonance assignments made by inspection. Uniformly ¹³C, ¹⁵N-labeled V66A/L68V CzrA was prepared as previously described for wild-type CzrA, and residue-specific backbone assignments of Zn₂ V66A/L68V CzrA were obtained in 10 mM *d*₁₃-4-morpholineethanesulfonic acid and 50 mM NaCl (pH 6.0) using standard triple resonance methods.^{17,67}

X-ray crystallography

V66A/L68V CzrA was extensively dialyzed into 10 mM Hepes and 50 mM NaCl (pH 7.0). The calculated protein concentration after dialysis was 460 μ M. The protein stock was loaded 1:1 with Zn(II) and crystallized under conditions used previously for Zn₂ wild-type CzrA³² of 100 mM Ches (pH 9.5), 200 mM NaCl and 10% polyethylene glycol 8000 (Wizard I, Emerald Biosystems) by hanging-drop diffusion at 20 °C. Diffraction data were collected at –160 °C on an R-AXIX IV+ detector at Indiana University. All data were processed with HKL2000,⁶⁸ and phase calculations were performed using the PHENIX AutoMR module.⁶⁹ The Zn(II)-bound wild-type CzrA structure was used as a molecular replacement model,³² and an initial refinement model was produced using the PHENIX AutoBuild module.⁶⁹ Model building was conducted using Coot⁷⁰ and subsequent refinement models in PHENIX (see Supplementary Table 3 for structure statistics).

Isothermal titration calorimetry

ITC experiments were carried out using a MicroCal VP-ITC calorimeter using 1.61 mM Zn(II) as titrant in the

syringe and solution conditions of 50 mM Hepes, 3.0 mM NTA [Zn(II) → CzrA] or 1.0 mM NTA [Zn(II) → CzrA•CzrO] as a Zn(II) competitor, 0.40 M NaCl (pH 7.0, 25.0 °C), 30–50 μM protein dimer or 38 μM complex. A self-complementary 28mer DNA was synthesized (MerMade 4) based on the native *czr* operator sequence (5'-TAACATATGAA-CATATGTTTCATATGTTA) annealed and purified as previously described.³⁴ CzrA•CzrO complex was formed by mixing 38 μM CzrA dimer and 41 μM CzrO. The raw ITC data were integrated, concentration normalized and plotted as heat *versus* metal:protein ratio using Origin(r). All data were fit using the sequential two-site model included in the data analysis software provided by MicroCal. NTA-independent binding constants were determined by using methods previously described.³⁴ The SD of the mean values from multiple experiments is given for all thermodynamic parameters.

Collection of ArsR sequences and the SCA

We started with sequences assigned to ArsR (PFAM accession number: PF01022) in the PFAM database (total 10,519 sequences) and selected ArsR sequences of typical lengths (90–140 aa) that match to PF01022 with significant *E*-value (<1e-10) for the SCA analysis. Selection resulted in 3000 nonredundant ArsR sequences. MUSCLE⁷¹ was used to prepare the multiple alignment of the selected ArsR sequences, which was used as input for the SCA Toolbox 5.0†. Details of the characteristics of this multiple sequence alignment are shown in Supplementary Fig. 8. The input multiple sequence alignment was truncated to sequence positions with % gap frequency no greater than 20% so that only largely non-gapped positions were used for the co-evolution analysis. Outputs from the SCA analysis include the conservation scores of different positions, measured as the Kullback–Leibler relative entropy, and a positional correlation matrix, which quantitatively indicates the correlated evolution of all pairs of positions in the alignment, with larger numbers indicating stronger coupling. The positional correlation matrix was further analyzed using the eigenvalue decomposition approach available in SCA toolbox. Examination of the top eigenmodes revealed a single sector⁷² in CzrA (106 residues) consisting of 17 co-evolving, physically connected, residues: K21, A22, D25, Y26, L29, L35, S41, V42, G43, Q53, Q59, H67, V69, K72, G75, S77 and Y80.

Accession codes

The coordinates and structure factors for Zn₂ V66A/L68V CzrA have been deposited in the Protein Data Bank under accession code 4GGG.

Acknowledgements

This work was supported by grants from the National Institutes of Health to D.P.G. (GM042569) and C.E.D.,III (GM094472). We thank Drs. A. Arunkumar, X. Kong and D. Ma for help in acquiring some of the NMR spectra presented here and Dr.

Randy Arnold for assistance in analyzing tandem mass spectrometry data.

Supplementary Data

Supplementary data to this article can be found online at <http://dx.doi.org/10.1016/j.jmb.2013.01.018>

Received 12 December 2012;

Received in revised form 13 January 2013;

Accepted 16 January 2013

Available online 23 January 2013

Keywords:

allostery;
zinc sensor;
metal homeostasis;
statistical coupling analysis;
ArsR

† http://systems.swmed.edu/rr_lab

Present address: Z. Ma, Department of Microbiology, Cornell University, Ithaca, NY 14853, USA

Abbreviations used:

HSQC, heteronuclear single quantum coherence;
TROSY, transverse relaxation optimized spectroscopy;
ITC, isothermal titration calorimetry; SCA, statistical coupling analysis; TCEP, tris(2-carboxyethyl)phosphine.

References

- Reinhart, G. D. (2004). Quantitative analysis and interpretation of allosteric behavior. *Methods Enzymol.* **380**, 187–203.
- Toncrova, H. & McLeish, T. C. (2010). Substrate-modulated thermal fluctuations affect long-range allosteric signaling in protein homodimers: exemplified in CAP. *Biophys. J.* **98**, 2317–2326.
- Smock, R. G. & Gierasch, L. M. (2009). Sending signals dynamically. *Science*, **324**, 198–203.
- Monod, J., Wyman, J. & Changeux, J. P. (1965). On the nature of allosteric transitions: a plausible model. *J. Mol. Biol.* **12**, 88–118.
- Koshland, D. E., Jr., Nemethy, G. & Filmer, D. (1966). Comparison of experimental binding data and theoretical models in proteins containing subunits. *Biochemistry*, **5**, 365–385.
- Henzler-Wildman, K. A., Lei, M., Thai, V., Kerns, S. J., Karplus, M. & Kern, D. (2007). A hierarchy of timescales in protein dynamics is linked to enzyme catalysis. *Nature*, **450**, 913–916.
- Henzler-Wildman, K. A., Thai, V., Lei, M., Ott, M., Wolf-Watz, M., Fenn, T. *et al.* (2007). Intrinsic motions along an enzymatic reaction trajectory. *Nature*, **450**, 838–844.
- Henzler-Wildman, K. & Kern, D. (2007). Dynamic personalities of proteins. *Nature*, **450**, 964–972.
- Popovych, N., Sun, S., Ebright, R. H. & Kalodimos, C. G. (2006). Dynamically driven protein allostery. *Nat. Struct. Mol. Biol.* **13**, 831–838.

10. Luque, I., Leavitt, S. A. & Freire, E. (2002). The linkage between protein folding and functional cooperativity: two sides of the same coin? *Annu. Rev. Biophys. Biomol. Struct.* **31**, 235–256.
11. England, J. L. (2011). Allostery in protein domains reflects a balance of steric and hydrophobic effects. *Structure*, **19**, 967–975.
12. Schrank, T. P., Bolen, D. W. & Hilser, V. J. (2009). Rational modulation of conformational fluctuations in adenylate kinase reveals a local unfolding mechanism for allostery and functional adaptation in proteins. *Proc. Natl Acad. Sci. USA*, **106**, 16984–16989.
13. Liu, T., Ramesh, A., Ma, Z., Ward, S. K., Zhang, L., George, G. N. *et al.* (2007). CsoR is a novel *Mycobacterium tuberculosis* copper-sensing transcriptional regulator. *Nat. Chem. Biol.* **3**, 60–68.
14. Ma, Z., Cowart, D. M., Scott, R. A. & Giedroc, D. P. (2009). Molecular insights into the metal selectivity of the copper(I)-sensing repressor CsoR from *Bacillus subtilis*. *Biochemistry*, **48**, 3325–3334.
15. Arunkumar, A. I., Campanello, G. C. & Giedroc, D. P. (2009). Solution structure of a paradigm ArsR family zinc sensor in the DNA-bound state. *Proc. Natl Acad. Sci. USA*, **106**, 18177–18182.
16. Guerra, A. J., Dann, C. E. & Giedroc, D. P. (2011). Crystal structure of the zinc-dependent MarR family transcriptional regulator AdcR in the Zn(II)-bound state. *J. Am. Chem. Soc.* **133**, 19614–19617.
17. Lee, C. W., Chakravorty, D. K., Chang, F. M., Reyes-Caballero, H., Ye, Y., Merz, K. M., Jr. *et al.* (2012). Solution structure of *Mycobacterium tuberculosis* NmtR in the apo state: insights into Ni(II)-mediated allostery. *Biochemistry*, **51**, 2619–2629.
18. Shi, W., Wu, J. & Rosen, B. P. (1994). Identification of a putative metal binding site in a new family of metalloregulatory proteins. *J. Biol. Chem.* **269**, 19826–19829.
19. Morby, A. P., Turner, J. S., Huckle, J. W. & 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.
20. Campbell, D. R., Chapman, K. E., Waldron, K. J., Tottey, S., Kendall, S., Cavallaro, G. *et al.* (2007). Mycobacterial cells have dual nickel-cobalt sensors: sequence relationships and metal sites of metal-responsive repressors are not congruent. *J. Biol. Chem.* **282**, 32298–32310.
21. Osman, D. & Cavet, J. S. (2010). Bacterial metal-sensing proteins exemplified by ArsR-SmtB family repressors. *Nat. Prod. Rep.* **27**, 668–680.
22. Wang, Y., Kendall, J., Cavet, J. S. & Giedroc, D. P. (2010). Elucidation of the functional metal binding profile of a Cd(II)/Pb(II) sensor CmtR(Sc) from *Streptomyces coelicolor*. *Biochemistry*, **49**, 6617–6626.
23. Ma, Z., Jacobsen, F. E. & Giedroc, D. P. (2009). Coordination chemistry of bacterial metal transport and sensing. *Chem. Rev.* **109**, 4644–4681.
24. Palm, G. J., Khanh Chi, B., Waack, P., Gronau, K., Becher, D., Albrecht, D. *et al.* (2012). Structural insights into the redox-switch mechanism of the MarR/DUF24-type regulator HypR. *Nucleic Acids Res.* **40**, 4178–4192.
25. Guimaraes, B. G., Barbosa, R. L., Soprano, A. S., Campos, B. M., de Souza, T. A., Tonoli, C. C. *et al.* (2011). Plant pathogenic bacteria utilize biofilm growth-associated repressor (BigR), a novel winged-helix redox switch, to control hydrogen sulfide detoxification under hypoxia. *J. Biol. Chem.* **286**, 26148–26157.
26. Kuroda, M., Hayashi, H. & Ohta, T. (1999). Chromosome-determined zinc-responsive operon *czr* in *Staphylococcus aureus* strain 912. *Microbiol. Immunol.* **43**, 115–125.
27. Giedroc, D. P. & Arunkumar, A. I. (2007). Metal sensor proteins: nature's metalloregulated allosteric switches. *Dalton Trans.* **29**, 3107–3120.
28. Pennella, M. A., Arunkumar, A. I. & Giedroc, D. P. (2006). Individual metal ligands play distinct functional roles in the zinc sensor *Staphylococcus aureus* CzrA. *J. Mol. Biol.* **356**, 1124–1136.
29. Lee, S., Arunkumar, A. I., Chen, X. & Giedroc, D. P. (2006). Structural insights into homo- and heterotropic allosteric coupling in the zinc sensor *S. aureus* CzrA from covalently fused dimers. *J. Am. Chem. Soc.* **128**, 1937–1947.
30. Reyes-Caballero, H., Guerra, A. J., Jacobsen, F. E., Kazmierczak, K. M., Cowart, D., Koppolu, U. M. *et al.* (2010). The metalloregulatory zinc site in *Streptococcus pneumoniae* AdcR, a zinc-activated MarR family repressor. *J. Mol. Biol.* **403**, 197–216.
31. VanZile, M. L., Chen, X. & Giedroc, D. P. (2002). Structural characterization of distinct α 3N and α 5 metal sites in the cyanobacterial zinc sensor SmtB. *Biochemistry*, **41**, 9765–9775.
32. Eicken, C., Pennella, M. A., Chen, X., Koshlap, K. M., VanZile, M. L., Sacchettini, J. C. *et al.* (2003). A metal-ligand-mediated intersubunit allosteric switch in related SmtB/ArsR zinc sensor proteins. *J. Mol. Biol.* **333**, 683–695.
33. Pennella, M. A., Shokes, J. E., Cospers, N. J., Scott, R. A. & Giedroc, D. P. (2003). Structural elements of metal selectivity in metal sensor proteins. *Proc. Natl Acad. Sci. USA*, **100**, 3713–3718.
34. Grosseohme, N. E. & Giedroc, D. P. (2009). Energetics of allosteric negative coupling in the zinc sensor *S. aureus* CzrA. *J. Am. Chem. Soc.* **131**, 17860–17870.
35. Chakravorty, D. K., Wang, B., Lee, C. W., Giedroc, D. P. & Merz, K. M., Jr. (2012). Simulations of allosteric motions in the zinc sensor CzrA. *J. Am. Chem. Soc.* **134**, 3367–3376.
36. Mulder, F. A., Hon, B., Mittermaier, A., Dahlquist, F. W. & Kay, L. E. (2002). Slow internal dynamics in proteins: application of NMR relaxation dispersion spectroscopy to methyl groups in a cavity mutant of T4 lysozyme. *J. Am. Chem. Soc.* **124**, 1443–1451.
37. Lockless, S. W. & Ranganathan, R. (1999). Evolutionarily conserved pathways of energetic connectivity in protein families. *Science*, **286**, 295–299.
38. Suel, G. M., Lockless, S. W., Wall, M. A. & Ranganathan, R. (2003). Evolutionarily conserved networks of residues mediate allosteric communication in proteins. *Nat. Struct. Biol.* **10**, 59–69.
39. Ma, Z., Cowart, D. M., Ward, B. P., Arnold, R. J., DiMarchi, R. D., Zhang, L. *et al.* (2009). Unnatural amino acid substitution as a probe of the allosteric coupling pathway in a mycobacterial Cu(I) sensor. *J. Am. Chem. Soc.* **131**, 18044–18045.
40. Hackeng, T. M., Griffin, J. H. & Dawson, P. E. (1999). Protein synthesis by native chemical ligation: expanded

- scope by using straightforward methodology. *Proc. Natl Acad. Sci. USA*, **96**, 10068–10073.
41. Arunkumar, A. I., Pennella, M. A., Kong, X. & Giedroc, D. P. (2007). Resonance assignments of the metal sensor protein CzrA in the apo-, Zn²⁺- and DNA-bound (42 kD) states. *Biomol. NMR Assignments*, **1**, 99–101.
 42. Fenton, A. W. (2008). Allostery: an illustrated definition for the “second secret of life”. *Trends Biochem. Sci.* **33**, 420–425.
 43. Record, M. T., Jr, Ha, J. H. & Fisher, M. A. (1991). Analysis of equilibrium and kinetic measurements to determine thermodynamic origins of stability and specificity and mechanism of formation of site-specific complexes between proteins and helical DNA. *Methods Enzymol.* **208**, 291–343.
 44. Busenlehner, L. S., Weng, T. C., Penner-Hahn, J. E. & Giedroc, D. P. (2002). Elucidation of primary (a3N) and vestigial (a5) heavy metal-binding sites in *Staphylococcus aureus* pl258 CadC: evolutionary implications for metal ion selectivity of ArsR/SmtB metal sensor proteins. *J. Mol. Biol.* **319**, 685–701.
 45. Wang, Y., Hemmingsen, L. & Giedroc, D. P. (2005). Structural and functional characterization of *Mycobacterium tuberculosis* CmtR, a PblI/CdII-sensing SmtB/ArsR metalloregulatory repressor. *Biochemistry*, **44**, 8976–8988.
 46. Banci, L., Bertini, I., Cantini, F., Ciofi-Baffoni, S., Cavet, J. S., Dennison, C. *et al.* (2007). NMR structural analysis of cadmium sensing by winged helix repressor CmtR. *J. Biol. Chem.* **282**, 30181–30188.
 47. Tzeng, S. R. & Kalodimos, C. G. (2009). Dynamic activation of an allosteric regulatory protein. *Nature*, **462**, 368–372.
 48. Frederick, K. K., Marlow, M. S., Valentine, K. G. & Wand, A. J. (2007). Conformational entropy in molecular recognition by proteins. *Nature*, **448**, 325–329.
 49. Olsson, T. S., Ladbury, J. E., Pitt, W. R. & Williams, M. A. (2011). Extent of enthalpy–entropy compensation in protein–ligand interactions. *Protein Sci.* **20**, 1607–1618.
 50. Diehl, C., Engstrom, O., Delaine, T., Hakansson, M., Genheden, S., Modig, K. *et al.* (2010). Protein flexibility and conformational entropy in ligand design targeting the carbohydrate recognition domain of galectin-3. *J. Am. Chem. Soc.* **132**, 14577–14589.
 51. Datta, D., Scheer, J. M., Romanowski, M. J. & Wells, J. A. (2008). An allosteric circuit in caspase-1. *J. Mol. Biol.* **381**, 1157–1167.
 52. Shen, A., Lupardus, P. J., Gersch, M. M., Puri, A. W., Albrow, V. E., Garcia, K. C. *et al.* (2011). Defining an allosteric circuit in the cysteine protease domain of *Clostridium difficile* toxins. *Nat. Struct. Mol. Biol.* **18**, 364–371.
 53. Wrabl, J. O., Gu, J., Liu, T., Schrank, T. P., Whitten, S. T. & Hilser, V. J. (2011). The role of protein conformational fluctuations in allostery, function, and evolution. *Biophys. Chem.* **159**, 129–141.
 54. Hilser, V. J., Wrabl, J. O. & Motlagh, H. N. (2012). Structural and energetic basis of allostery. *Annu. Rev. Biophys.* **41**, 585–609.
 55. del Sol, A., Tsai, C. J., Ma, B. & Nussinov, R. (2009). The origin of allosteric functional modulation: multiple pre-existing pathways. *Structure*, **17**, 1042–1050.
 56. Hatley, M. E., Lockless, S. W., Gibson, S. K., Gilman, A. G. & Ranganathan, R. (2003). Allosteric determinants in guanine nucleotide-binding proteins. *Proc. Natl Acad. Sci. USA*, **100**, 14445–14450.
 57. Smock, R. G., Rivoire, O., Russ, W. P., Swain, J. F., Leibler, S., Ranganathan, R. *et al.* (2010). An interdomain sector mediating allostery in Hsp70 molecular chaperones. *Mol. Syst. Biol.* **6**, 414.
 58. Ye, J., Kandedgedara, A., Martin, P. & Rosen, B. P. (2005). Crystal structure of the *Staphylococcus aureus* pl258 CadC Cd(II)/Pb(II)/Zn(II)-responsive repressor. *J. Bacteriol.* **187**, 4214–4221.
 59. VanZile, M. L., Chen, X. & 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.
 60. Jefferson, J. R., Hunt, J. B. & Ginsburg, A. (1990). Characterization of indo-1 and quin-2 as spectroscopic probes for Zn²⁺–protein interactions. *Anal. Biochem.* **187**, 328–336.
 61. Walkup, G. K. & Imperiali, B. (1997). Fluorescent chemosensors for divalent zinc based on zinc finger domains. Enhanced oxidative stability, metal binding affinity, and structural and functional characterization. *J. Am. Chem. Soc.* **119**, 3443–3450.
 62. Kuzmic, P. (1996). Program DYNAFIT for the analysis of enzyme kinetic data: application to HIV proteinase. *Anal. Biochem.* **237**, 260–273.
 63. Lee, Y. H., Dorwart, M. R., Hazlett, K. R., Deka, R. K., Norgard, M. V., Radolf, J. D. *et al.* (2002). The crystal structure of Zn(II)-free *Treponema pallidum* TroA, a periplasmic metal-binding protein, reveals a closed conformation. *J. Bacteriol.* **184**, 2300–2304.
 64. Ma, Z., Gabriel, S. E. & Helmann, J. D. (2011). Sequential binding and sensing of Zn(II) by *Bacillus subtilis* Zur. *Nucleic Acids Res.* **39**, 9130–9138.
 65. Delaglio, F., Grzesiek, S., Vuister, G. W., Zhu, G., Pfeifer, J. & Bax, A. (1995). NMRPipe: a multi-dimensional spectral processing system based on UNIX pipes. *J. Biomol. NMR*, **6**, 277–293.
 66. Goddard, T. D. and Kneller, D. G. Sparky 3, University of California, San Francisco.
 67. Lee, C. W. & Giedroc, D. P. (2012). ¹H, ¹³C, and ¹⁵N resonance assignments of NmtR, a Ni(II)/Co(II) metalloregulatory protein of *Mycobacterium tuberculosis*. *Biomol. NMR Assignments*, [Epub ahead of print].
 68. Otwinowski, Z. (1993). Oscillation data reduction program. In *Proceedings of the CCP4 Study Weekend: Data Collection and Processing* (Sawyer, L., Isaacs, N. & Bailey, S., eds), pp. 56–62, SERC Daresbury Laboratory, Warrington, UK.
 69. Adams, P. D., Afonine, P. V., Bunkoczi, G., Chen, V. B., Davis, I. W., Echols, N. *et al.* (2010). PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **66**, 213–221.
 70. Emsley, P. & Cowtan, K. (2004). Coot: model-building tools for molecular graphics. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **60**, 2126–2132.
 71. Edgar, R. C. (2004). MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* **32**, 1792–1797.
 72. Halabi, N., Rivoire, O., Leibler, S. & Ranganathan, R. (2009). Protein sectors: evolutionary units of three-dimensional structure. *Cell*, **138**, 774–786.