BIOCHEMISTRY
including biophysical chemistry & molecular biology

Current Topic

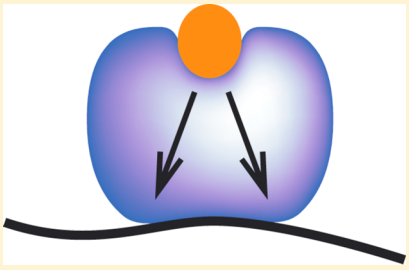
pubs.acs.org/biochemistry

Modulation of DNA Binding by Gene-Specific Transcription Factors

Robert F. Schleif*

Biology Department, Johns Hopkins University, 3400 North Charles Street, Baltimore, Maryland 21218, United States

ABSTRACT: The transcription of many genes, particularly in prokaryotes, is controlled by transcription factors whose activity can be modulated by controlling their DNA binding affinity. Understanding the molecular mechanisms by which DNA binding affinity is regulated is important, but because forming definitive conclusions usually requires detailed structural information in combination with data from extensive biophysical, biochemical, and sometimes genetic experiments, little is truly understood about this topic. This review describes the biological requirements placed upon DNA binding transcription factors and their consequent properties, particularly the ways that DNA binding affinity can be modulated and methods for its study. What is known and not known about the mechanisms modulating the DNA binding affinity of a number of prokaryotic transcription factors, including CAP and *lac* repressor, is provided.



Gene-specific transcription factors transduce changes in intracellular conditions to changes in transcription of a specific gene or class of genes. DNA binding transcription factors must possess three different activities. First, to direct their activities to specific genes, they must bind to specific DNA sequences that are associated with the gene(s) to be controlled. Second, a transcription factor must respond either to the appropriate intracellular condition, for example, the presence of a particular small molecule effector ligand, or to a signal transmitted to it, for example, by its becoming phosphorylated. Third, a transcription factor must be capable of controlling transcription. This can be as simple as sterically blocking access of RNA polymerase to a promoter, or more complicated as in affecting the activities of other transcription factors or in stimulating the binding, initiation, or elongation by RNA polymerase. A large class of transcription factors directly alter their DNA binding in response to the receipt of regulatory signals. The response of the *lac* operon repressor, LacI, to its inducers is one example. It is the binding or absence of binding of such a transcription factor to DNA in response to a signal that determines whether a gene is expressed. In a second class of transcription factors, a ligand or signal changes the factor's ability to affect transcription but does so without a substantial change in the factor's DNA binding affinity. The transcription factor known as CytR, a regulator of genes involved in nucleoside and deoxynucleoside uptake and metabolism, is an example of this second class.[1]

This review touches on a variety of issues related to transcription factors, but its central focus is the structural changes that alter a transcription factor's DNA binding affinities. Most of the transcription factors that have been sufficiently studied that we know something about their mechanism are bacterial. In particular, the CAP protein and the *lac* operon repressor, LacI, have been intensively studied for decades, and much is known about these two proteins. Even for these proteins, however, our understanding is incomplete.

Transcription factors that change their DNA binding in response to an effector are signal transducers because they convert changes in effector concentration to changes in transcription. They can also be considered to be allosterically regulated proteins because the binding of an effector at one site in the protein changes the affinity that a different (*allo-* for other) site in the protein binds DNA. A note on allostery. Many papers on allostery are concerned not with the mechanism of communication between the separated effector and active sites but with the cooperative or anticooperative binding of effector and substrate molecules that is often observed with multisubunit enzymes. The consequences, but not the underlying mechanical mechanisms, of postulating that all subunits of a multisubunit enzyme coordinately shift from one binding affinity to another are explored in the Monod, Wyman, Changeaux allostery model.[2] Analogously, the consequences of a subunit shifting its binding affinity and only influencing nearby subunits are developed in the Koshland, Nemethy, Filmer allostery model.[3] In general, the formal models of allostery occupy a position with respect to protein function and protein engineering like the position that thermodynamics occupies with respect to automobile engine design. A working engine will obey the laws of thermodynamics, but one cannot begin to design a working engine with a knowledge of only thermodynamics. This review focuses more on the nuts and bolts than on abstract theory.

GENERAL PRINCIPLES AVAILABLE FOR MODULATING DNA BINDING

Two ways that a small molecule ligand effector can modulate a protein's DNA binding affinity are by controlling the accessibility of the protein's DNA binding region(s) to contact the DNA target sequence(s) and by altering the intrinsic DNA

Received: July 19, 2013
Revised: August 20, 2013
Published: August 20, 2013

   dx.doi.org/10.1021/bi400968e | *Biochemistry* 2013, 52, 6755−6765

binding affinity (shape) of the DNA binding region(s). Ligands can drive these mechanisms directly, or indirectly by acting over a distance within the protein.

Many transcription factors are homo-oligomeric and simultaneously utilize two DNA binding domains to contact two similar DNA sites. While the DNA binding of such proteins is modulated by the mechanisms described above, one specific form of one of the mechanisms is available only to oligomeric proteins. It should be explicitly mentioned because of its mechanical simplicity and frequent occurrence in nature. An oligomeric protein that utilizes two DNA binding domains to contact simultaneously two DNA sites allows DNA binding affinity to be controlled by regulating the separation and/or relative orientation of the domains. If both of the DNA-contacting surfaces are correctly positioned to contact their DNA targets, the binding affinity can be high, but if both of the DNA-contacting surfaces cannot be properly oriented or positioned to simultaneously contact both DNA sites without the expenditure of significant amounts of free energy, the DNA binding affinity will be much lower. This is a consequence of the fact that any free energy that is required to position the domains is subtracted from the DNA interaction binding energy, leaving the protein less tightly bound than it is otherwise. If the domains are on long and flexible tethers, the free energy cost of confining them to the correct positions can be significant and is primarily entropic. Alternatively, if the DNA binding domains are on short tethers but are rigidly held in positions that prevent both from simultaneously contacting their binding sites, the positioning energy is the cost of distorting the protein or breaking the bonds that hold the domains in unfavorable binding positions. This is an enthalpic energy cost.

A third way that DNA binding affinity can be modulated is through regulation of oligomerization. The binding of a dimer to two DNA sites can be very much tighter than the binding of either monomer. Hence, if dimerization or oligomerization is controlled, then DNA binding can be controlled.

The reason for the very large potential increase in the binding affinity of an oligomeric state of a protein to two sites compared to the affinity of its monomeric state binding to a single site derives from the relative energetic costs and gains involved. When a monomer binds, some of the free energy of the protein−DNA interaction can be considered to be going into holding the monomer onto the DNA, and some of the free energy can be considered to be being used to overcome the translational and rotational freedom of motion of the monomer, that is, its translational and rotational entropy. When a homodimer binds to DNA, twice as much protein−DNA interaction energy can be available, but the entropic energy cost of freezing translation and rotation is not much more than that required for a monomer. Thus, the "extra" protein−DNA interaction free energy from the second subunit can go into holding the dimer on the DNA. Another way to view the situation is that if one subunit of a homodimer binds to its DNA site, the other subunit is automatically positioned to bind to its DNA site, and thus most its interaction energy can go into holding the dimer on DNA. These considerations give rise to a simple expression for the approximate maximal binding affinity of a dimer in terms of the monomer affinity as $K_D = 4K_M^2/C_{eff}$, where K_D and K_M are the dimer and monomer dissociation constants, respectively, C_{eff} is the effective concentration of the second monomer in the vicinity of its DNA binding site when the first monomer has bound to its site, and 4 is a statistical factor.[4−8] C_{eff} depends on the length and flexibility of the connection between the subunits.

Finally, a trivial way that binding of a transcription factor to DNA can be regulated is to control the factor's synthesis. For example, the intracellular level of SoxS, a positive acting transcriptional regulator of genes that provide protection against oxidative stress, is controlled by SoxR, a transcriptional regulator whose activity is regulated by redox potential.[9] It is unclear why SoxR cannot directly control the expression of the protective genes. Although the existence of a second regulation layer allows amplification, the activity of SoxS is not cooperative, and therefore, the signal is not sharpened the way it is in the eukaryotic phosphorylation MAP cascade.[10] Possibly a useful consequence of the SoxS layer of transcriptional control is to smooth and perhaps prolong the induction signals provided by SoxR.

METHODS FOR DETERMINING MECHANISMS THAT ALTER DNA BINDING

To begin to understand a transcription factor's allosteric mechanism for regulating DNA binding, we usually need detailed structural information in the absence and presence of the ligand effector, that is, in the apo and holo states of the protein. Although, detailed structural information is necessary, usually it is insufficient, and sometimes incorrect, or misleading. In X-ray diffraction studies, lattice interactions in a crystal can distort a protein or trap a protein in a misleading or dead-end conformation. In addition, the use of mutant proteins or incomplete proteins can also produce misrepresentative structures.

A hindrance to the universal use of crystal structures is the fact that often an interesting protein will fail to crystallize. Therefore, nuclear magnetic resonance (NMR) structural studies are sometimes undertaken. One impediment to their wider use, however, is the difficulty of applying NMR to large proteins. Another is that the protein must remain soluble at high concentrations. Finally, NMR measurements provide short distance constraints, but the determination of global structure utilizing only short distance constraints can be difficult. In brief, sometimes it is difficult to obtain essential high-resolution structural information.

Structural information alone is insufficient to verify a mechanism because affinities or changed affinities of DNA binding generally cannot be calculated directly from structure and compared to experimental values. Later we will see an example in which the proposed structural change responsible for the proposed change in DNA binding affinity is clearly inadequate.

Another approach for deducing mechanism relies on mutants. Deductions of allosteric mechanisms, however, based on genetic data alone have usually been unsuccessful because activity data for a collection of mutants are almost always consistent with many different mechanisms. On the other hand, genetic data can be most helpful in suggesting particular variant proteins to use in structural studies or to use in a physical experiment designed to test a specific mechanism.

As mentioned above, a change in the binding of the effector molecule to an oligomeric transcription factor may lead to alteration in the separation or orientation of the protein's two DNA binding domains so that they no longer match the two half-site target DNA binding sites. This can be called extrinsic control. Alternatively, the change in effector binding could change the shape of the DNA binding domains or their

accessibility to the DNA, thereby changing their intrinsic DNA binding affinity. This can be called intrinsic control.

Intrinsic and extrinsic control can be distinguished with the flexible linker method. In this, the two DNA half-sites of a protein's binding target are flexibly connected by a 20−50-base stretch of single-stranded DNA[11,12] (Figure 1). The flexibility

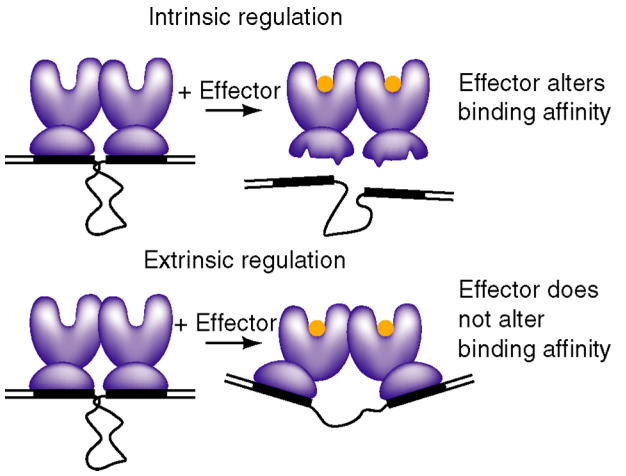


Figure 1. Intrinsic and extrinsic regulation of DNA binding affinity by a protein in which the small molecule effector normally decreases DNA binding affinity.

of the single-stranded connector allows the two DNA half-sites to easily position and orient to accommodate a transcription factor's DNA binding domains, even though they may be mispositioned for binding to the normal DNA binding site. That is, a transcription factor that utilizes extrinsic control will bind to the flexible linker DNA with nearly the same affinity in the presence and absence of the effector. On the other hand, if the factor modulates the intrinsic DNA binding affinity of its DNA binding domains, this modulation will be retained for the flexible linker DNA.

■ ENERGETIC AND KINETIC CONSIDERATIONS

If a transcription factor binds to DNA with an equilibrium dissociation constant of 10^{-11} M, the free energy change of the binding is $\Delta G = -RT \ln K_{eq} = 14.9$ kcal/mol. Suppose that binding of a ligand effector to a protein reduces its DNA binding affinity by a factor of 100. As a result, the free energy change of the binding reaction becomes 12.2 kcal/mol. The change in the free energy of the transcription factor's DNA binding has to come from the free energy change of the effector binding to the transcription factor. Thus, if the binding of the effector to a transcription factor changes the protein's DNA binding affinity, then the binding of the factor to DNA reciprocally changes the affinity of the DNA-bound transcription factor for the effector.

The changes in transcription factor and ligand binding affinities follow from analysis of the four possible binding reactions and their dissociation constants as shown in the box diagram (Figure 2). The relationship among the four dissociation constants is also useful when the experimental determination of one of them is difficult. Then, determination of the other three provides the missing fourth.

The binding of a monomeric transcription factor, P, to DNA, D, is represented by the two vertical reactions. That on the left is binding of the free protein to DNA and that on the right by ligand-bound protein. The horizontal reactions represent ligand binding. If ligand-bound protein binds DNA 100 times as tightly as the apoprotein, $K_2 = 0.01K_4$. Then by the relationship among the K values, K_3 must be $0.01K_1$, which is to say that the ligand binds to the DNA-bound complex 100 times as tightly as it binds to the free protein. Conceptually, this means that some of the energy of binding of the ligand to the free protein is used to shift the protein to the conformation that binds DNA more tightly. If, however, the apoprotein is bound to DNA, it is already in this conformation, and should an effector ligand bind to the protein−DNA complex, none of the ligand's binding energy needs to be used for shifting the protein's conformation; therefore, it binds more tightly than to the free protein.

As an aside, it should be noted that direct measurement of the equilibrium dissociation constant of a protein from DNA can be difficult, and published values can sometimes seriously be in error. The difficulty in taking measurements is a consequence of the fact that the transcription factors of some regulated systems must bind to their target DNA sites with sufficient affinity that, *in vivo*, most of the target DNA sites are occupied with high probability despite the fact that a cell may contain only a few to a few hundred molecules of transcription factor. This means that generally the dissociation constant of the transcription factor from its specific DNA binding site ranges from 10^{-9} to 10^{-13} M. Particularly when its value is $\leq 10^{-12}$, a dissociation constant can be difficult to measure. The problem is that determination of a dissociation constant requires knowing the concentrations free and DNA-bound protein and free DNA in an equilibrium mixture. Often, this necessitates that two of the three species be present and measurable at concentrations near or below the value of the dissociation constant.

The biological reason for tight binding to DNA can be seen by considering an idealized example of a repressor. If present at one molecule per bacterial cell, its molar concentration would be $\sim 10^{-9}$ M. If at such a concentration it is required that its target operator be free of repressor only 0.1% the time, as is the case in the *Escherichia coli lac* operon, whose uninduced basal level is 1/1000 of the fully induced level, then the equilibrium dissociation constant for the operator would need to be no greater than 10^{-12} M, and in fact, binding would have to be still tighter to compensate for molecules of repressor momentarily bound to the vast number of nonspecific sites in the genome. Of course, the requirement for such tight binding could be

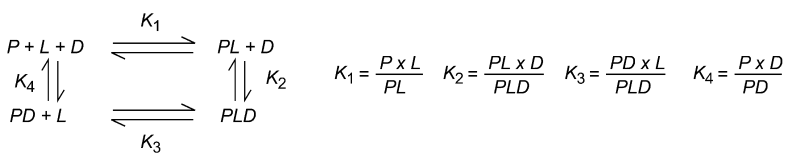


Figure 2. Linkage diagram in which P is a protein that binds a molecule of ligand effector L and D is DNA. The four equilibrium dissociation constants are written out, and it can be verified that they obey the relationship $K_1K_2 = K_3K_4$.

dx.doi.org/10.1021/bi400968e | *Biochemistry* 2013, 52, 6755−6765

ameliorated by substantially increasing the number of repressor molecules per cell, but increasing the repressor concentration by very much defeats the value of having a regulated system at all.

Because of the difficulties of measuring equilibrium DNA binding constants, often, the considerably simpler measurement of the dissociation rate of a protein from DNA is made instead. Such measurements are sensible because the well-documented measurements of the forward rate constants for binding to DNA with and without ligand effector of LacI,[13,14] TetR,[15] and AraC[16] show them to be relatively constant at values close to the diffusion-limited rate of 2×10^9 M^{-1} s^{-1}. (A recent claim that the inducer of the *lac* operon primarily affects the repressor–operator association rate[17] disagrees with earlier findings and does not appear to be supported by the references cited on its behalf.[18,19]) Thus, because the equilibrium dissociation constant K_{eq} equals the ratio of the association and dissociation rate constants, k_1 and k_{-1}, respectively, measurements of relative values of k_{-1} provide relative values of K_{eq}.

Many transcription factors are oligomeric and can bind more than one molecule of ligand. The relevant binding equilibria for such a dimer are shown in Figure 3. It is sometimes possible to

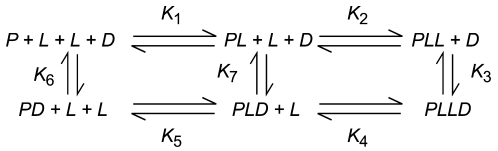


Figure 3. Linkage diagram in which a dimeric protein can bind two molecules of the effector. If all the K values are equilibrium dissociation constants, the following relationships are true for the two small cycles and the outer cycle: $K_1K_7 = K_5K_6$, $K_2K_3 = K_4K_7$, and $K_1K_2K_3 = K_4K_5K_6$. The third relationship follows directly from the first two and therefore is not independent.

determine whether it is the binding of the first ligand effector molecule or the binding of the second or later molecule that produces the major shift in the factor. In studying such situations, it is first necessary to determine the ligand dissociation constants by fitting experimental data to the binding equation that matches the physical situation. For example, in the case of a dimeric protein whose concentration is given by P and a ligand whose concentration is given by L, the fraction of dimers that have bound one or two molecules of ligand is $(L^2 + K_2L)/(K_1K_2 + K_2L + L^2)$, where K_1 and K_2 are the dissociation constants for the binding of the first and second ligand molecules, respectively. One might at first expect that the binding of the ligand is cooperative if $K_2 < K_1$ and anticooperative if $K_2 > K_1$. This is not quite the case, however. The fact that there are two subunits to which the first ligand molecule can bind and also two subunits of a doubly bound dimer from which a ligand can dissociate introduces two factors of 2 into the analysis. In the case of complete independence in the binding of the first and second molecules of ligand, that is, no cooperativity or anticooperativity, K_1 will be $^1/_4K_2$. Hence, an experimental finding that K_2 is $0.2K_1$ would indicate only very weak cooperativity in the binding of the ligand.

Suppose, as explained in the above, we can obtain the ratio of protein–DNA equilibrium dissociation constants K_6 and K_3 from the protein–DNA dissociation rates, and the four effector–protein equilibrium dissociation constants from measurements like fluorescence, equilibrium dialysis, or NMR. K_1 and K_2 can be separately determined if they are sufficiently different from one another. The same holds true for K_4 and K_5. Suppose that the binding of the ligand increases the DNA binding affinity of the protein by a factor of 100, that is, that $K_3/K_6 = 0.01$. If all of this effect comes from the binding of the first molecule of effector, then K_5/K_6 must be 0.01, and $K_2 = K_4$. Because all of the effect caused by ligand binding comes from the binding of the first molecule, it, and not the second effector molecule, is generating the change in the protein that facilitates DNA binding. Thus, the second effector molecule binds much more tightly than the first, and experimentally, we would observe cooperativity in the binding of the ligand to the free protein. Conversely, if we observe anticooperativity in the binding of ligand, we can conclude that it is the binding of the second molecule of effector that drives the protein into the tight DNA binding conformation.

CAP PROTEIN

CAP protein helps regulate more than 100 operons in *E. coli*[20,21] where, in most cases, its binding to specific DNA sequences near promoters assists the binding of RNA polymerase or stimulates the transition of RNA polymerase to an open complex capable of transcription. Many laboratories have studied CAP and its regulation for many years. Thus, an enormous amount of physiological, genetic, and biophysical data has been collected. The DNA binding properties of the dimeric CAP change dramatically upon its binding of cAMP. The main questions to be discussed here are the structural changes that occur within CAP in response to the binding of cAMP and whether they are sufficient to explain the changes in the protein's DNA binding properties.

Flexible linker experiments like those described in the preceding section show that CAP utilizes an intrinsic mechanism. That is, in the absence of cAMP, the protein's DNA binding domains cannot bind DNA. Either they are too distorted to bind DNA, or they are held in orientations in which DNA contact is not possible.[12] In other words, the weak DNA binding that is seen in the absence of cAMP is not a result of mispositioned DNA binding domains or a result of the domains being on such long and flexible tethers that the protein binds weakly. The binding of cAMP increases the extent of DNA binding by CAP, and the binding of cAMP is anticooperative.[22] As shown in the previous section, this indicates that the main conformational change in CAP that facilitates tight DNA binding occurs upon binding of the second molecule of cAMP.

Determination of the atomic details of the mechanisms ultimately responsible for the change in the DNA binding affinity of CAP requires high-quality structural information of both the apo and cAMP-bound forms of the protein. Incomplete or slightly incorrect structural information without supporting biochemical or biophysical evidence can be misleading. For example, the first determined structure of cAMP-CAP suggested to some that it might bind to left helical DNA.[23] Later, a more detailed X-ray structure of cAMP-CAP[24] correctly suggested that interactions critical to the allosteric change induced by cAMP involved hydrogen bonds between the adenine ring of bound cAMP and residues close to the hinge between the cAMP binding domain and the DNA binding domain of the protein but did not suggest additional details crucial to the identification of the actual mechanical details by which DNA binding affinity is modulated.

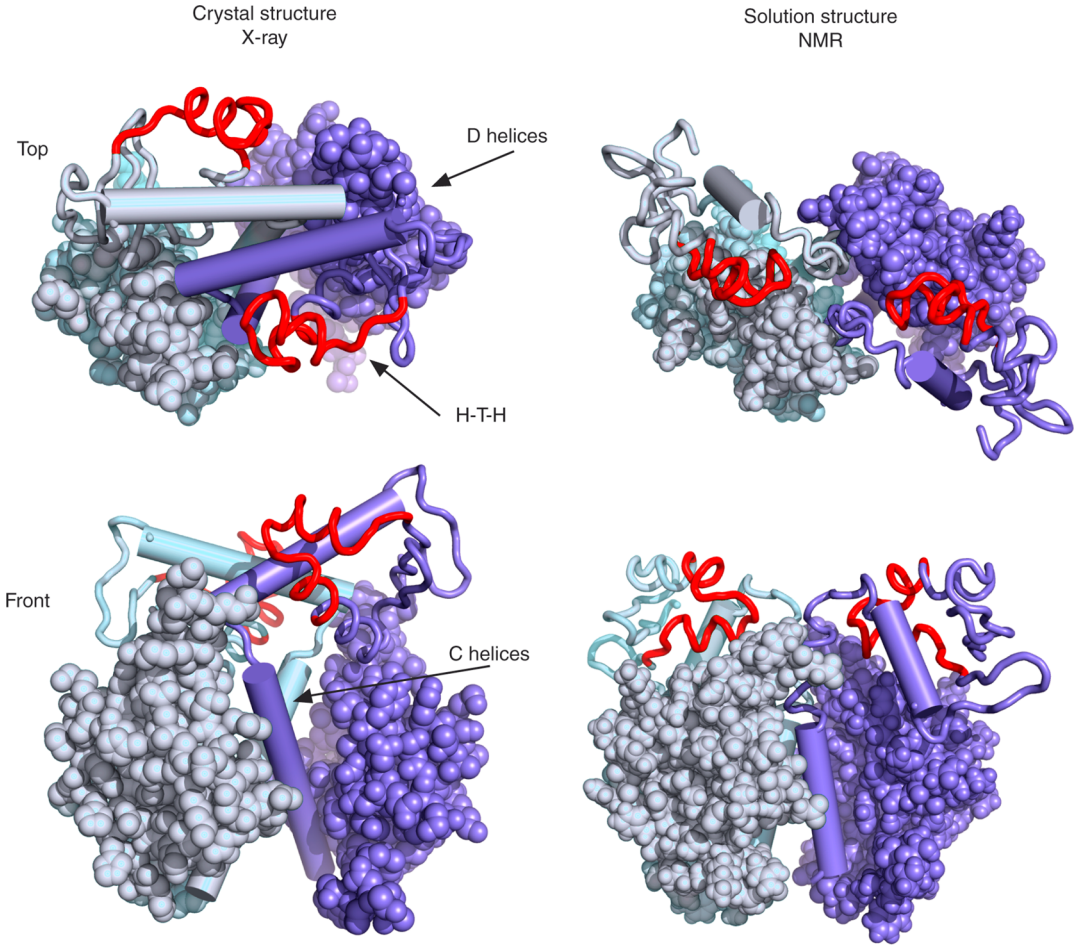


Figure 4. Apo-CAP structure as determined by X-ray diffraction [Protein Data Bank (PDB) entry 3FWE, left] and by NMR (PDB entry 2WC2, right). Only the C and D helices are shown as cylinders. The DNA-contacting H-T-H motif regions are colored red.

Although the structure of CAP in solution containing bound cAMP was published in 1981,[23] it was not until 2009 that the structure of apoCAP was published. At this time, both a crystal structure and a solution structure were presented.[25,26] The two structures are similar in many respects; however, they differ in critical details, and unfortunately, the two structures lead to the prediction of different mechanisms for the cAMP-induced change in DNA binding affinity. The flexible linker experiment described above and several additional pieces of evidence described below suggest that the mechanism suggested by the crystal structure provides the more correct explanation for the cAMP-induced change in the DNA binding affinity of CAP.

The early determinations of the structure of cAMP-CAP in the absence of DNA showed that the main dimerizing element of the protein is helix C that forms a parallel coiled coil with helix C′ on the other subunit.[23] In cAMP-CAP, helix C spans residues 110−136 and critical hydrogen bonds are formed between the adenine ring of each molecule of cAMP bound to one subunit and residue 127 of helix C of the same subunit and residue 128 of helix C′ of the other subunit. The importance of these residues to the protein's response to cAMP was later verified biochemically.[27] The subsequent determination of the structure(s) of apo-CAP showed (Figure 4), indeed, without the stabilization provided by helix−cAMP interactions, what had been the terminal part of the C helices becomes random coil. In the apo-CAP structures, the C helices terminate before residue 127 rather than extending to residue 136 as they do in cAMP-CAP. In the holostructure, residues 137 and 138 constitute an interdomain linker, and helix D in the DNA binding domain begins at residue 139. The next two helices, a helix−turn−helix (H−T−H) motif, constitute the DNA contacting region of CAP. Both the X-ray crystal structure and the NMR solution structure of apo-CAP differ from that of cAMP-CAP in the region of the end of the C helices and the beginning of the D helices. In the X-ray structure, the end points of both the C and D helices change compared to those of cAMP-CAP. The C helices are shorter, extending only to residue 130, and the D helices began at residue 135 rather than residue 139.[25] These structural changes allow the extended D helices of the two DNA binding domains to interact with each other and form a rather stable structure in which the DNA-contacting surfaces of the H-T-H regions[28] are not sufficiently exposed to allow the domains to bind normally to DNA. Thus, Sharma et al. consider that the major impediment to DNA binding of apo-CAP is the breaking of the interaction between the D helices.[25]

In contrast to what is seen in the X-ray structure, in the NMR structure of apo-CAP, only helix C is observed to change.[26] Instead of ending at residue 136 as in cAMP-CAP or at residue 130 as in the X-ray-determined structure of apo-CAP, in the NMR structure, helix C is seen to end at residue 125. Thus, in the NMR-determined structure, the interdomain linker becomes 13 residues long. Because the DNA binding domains do not directly interact as they do in the X-ray structure, the

dx.doi.org/10.1021/bi400968e | *Biochemistry* 2013, 52, 6755−6765

authors imply that the relatively long length of the interdomain linker provides so much more conformational freedom to the positioning of the DNA binding domains that the apoprotein loses its ability to bind tightly to its DNA binding site.

We can make a rough estimation of the magnitude of the decrease in DNA binding affinity produced by increasing the length of the interdomain linker from 2 to 13 residues. This can be done by considering the binding of the second DNA binding domain to its DNA half-site after the first has bound to its half-site as discussed earlier. This affinity is inversely proportional to the effective concentration of the second DNA binding domain in the vicinity of its binding site.[4−8] In the case of lengthening the linker from 2 to 13 residues, the DNA binding affinity will decrease to roughly $(2/13)^3 \approx 1/300$ of that of the apo state. At the *lac* operon binding site, the presence of cAMP increases the CAP binding affinity from 200- to 500-fold,[27] and thus, the NMR-based mechanism is compatible with the known energetics. We cannot tell whether the energetics are also compatible with the model derived from the X-ray structures. The free energy that is required to generate a 300-fold affinity change is 3.4 kcal/mol, which is compatible with the energy that it might cost to separate the D helices of the DNA binding domains. Until this energy can be measured or accurately calculated, we cannot be certain that the self-association of the D helices of the two DNA binding domains is primarily responsible for the low DNA binding affinity of apo-CAP.

On the basis of the X-ray structure, one could view apo-CAP as being actively prevented from binding to DNA, an enthalpic effect. On the basis of the NMR-determined structure, it would be the conformational freedom of the DNA binding domains provided by the long interdomain linker that would prevent binding of the apo form of CAP, an entropic effect. One might hope that the entropic and enthalpic mechanisms could be distinguishable by direct thermodynamic measurements. This is unlikely, however, because, in addition to the thermodynamics of positioning the DNA binding domains, significant additional entropy and enthalpy changes will be associated with the binding of cAMP, displacement of water from CAP, cAMP, and DNA, and the bending of the DNA. These would all tend to mask the "mechanistic" energetic costs.

Why, in fact, do the crystal structures and solution structures differ from each other? The main crystallographic work was done using a mutant that stabilized the apo structure.[25] Crystals of wild-type protein, although of lower resolution, revealed the same basic structure, so it seems unlikely that the crystal structure is the misleading one. As suggested by Sharma et al.,[25] possibly the NMR structure was misleading because a spin-label inserted to assist in determining the position of the DNA binding domains interfered with an important domain−domain interaction.

Did genetic studies provide critical insights for the determination of the allosteric mechanism of CAP? Unfortunately, they did not. For example, mutants of CAP that do not require the presence of cAMP for induction of the *lac* operon can relatively easily be isolated and were found at residues 72, 141, 142, and 144.[29] On the basis of the only structure of cAMP-CAP then available,[23] the authors proposed that in the absence of cAMP, the DNA binding domains of wild-type protein bound to the dimerization domains, leaving the protein able to bind only weakly to DNA. When cAMP bound, the DNA binding domains were postulated to be released from the dimerization domains and tight DNA binding ensued. The mutations at residues 141, 142, and 144 were postulated to weaken the interactions proposed between the DNA binding and dimerization domains, leaving the mutant proteins in a tight DNA binding state even in the absence of cAMP. It is not clear how the behavior of these mutations is explained in the NMR structure and mechanism of apo-CAP, but their behavior is simply explained in the X-ray structure and mechanism. They interfere with interactions between the D helices of the DNA binding domains that, in the absence of cAMP, hold the domains together in orientations that interfere with DNA binding.

What about physical studies performed after the determination of the structures of apo-CAP? Lee et al.[30] have measured the thermodynamics of cAMP binding and DNA binding, protease sensitivity, and Stokes radius of wild-type CAP and a number of residue 138 mutants, the N-terminal end of helix D in apo-CAP as determined by crystallography. They find a reasonable correlation between the propensity of a residue to N-cap an α helix and the tendency of CAP to be in the apo state, i.e., reduced affinity for the binding of the second molecule of cAMP and reduced DNA binding affinity. Although the authors do not explicitly state it, these results provide good biochemical data supporting the X-ray-determined structure and mechanism of CAP, and not the NMR-derived structure and mechanism.

One would expect that CAP from other bacteria would function via the same basic mechanism. Therefore, it is a bit of a surprise that the disposition of the DNA binding domains of unliganded CAP from *Mycobacterium tuberculosis*[31] is more similar to that of the NMR-determined structure of CAP than that of the X-ray-determined structure. Because of the significant structural differences between the two subunits in the structure, however, perhaps this structure should not be considered further without additional genetic and/or biophysical data.

LAC REPRESSOR

In the absence of the natural inducer, allolactose,[32] or the useful gratuitous artificial inducer, isopropyl thiogalactopyranoside (IPTG), the *E. coli lac* repressor binds to the nearly symmetric *lac* operator and represses transcription of the *lac* operon. The presence of an inducer weakens the affinity of the repressor for the *lac* DNA (containing the three operators O_1, O_2, and O_3) 1000-fold,[33] and correspondingly, *in vivo*, the inducer increases the level of expression of the *lacZYA* genes a ≥1000-fold.[34] As in the case of CAP, the *lac* repressor has been extensively studied for many years. Much information is known, but only a tiny fraction will be discussed here. Structurally and functionally, the tetrameric repressor consists of two dimer units. A single dimer unit contacts an operator using two monomers. Each monomer consists of a DNA binding domain, a ligand binding domain with two recognizable subdomains (the N-terminal and C-terminal subdomains), and a short region at the C-terminus that forms a four-strand helical bundle that holds two dimer units together to form a tetramer. The DNA binding domain consists of a hinge helix, another short helical region, and the well-known H−T−H motif (Figure 5).

The dimer unit of a repressor binds to an operator with the two hinge helices together contacting the minor groove at the center of the nearly symmetric operator and the two H−T−H motifs in the two adjacent major grooves of the operator half-sites.[35,36] Normally, the tetrameric *lac* repressor binds to two pairs of half-sites, the main operator, O_1, that is located at the *lac* promoter and one of two weaker binding pseudo-operators,

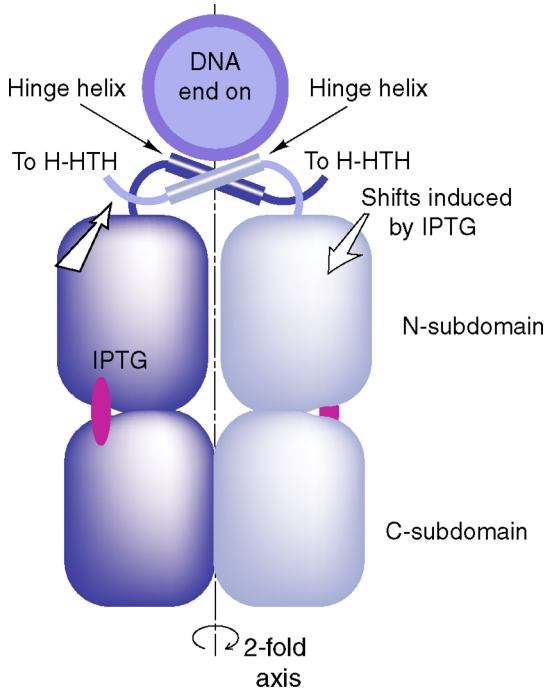


Figure 5. Hinge helices and subdomains of the *lac* repressor dimer. The H−T−H region and C-terminal tetramerizing helices are not represented. The C subdomains are bound to each other, and the binding of IPTG between the subdomains tips the N subdomain with respect to the C subdomain as shown by the arrows. The movement separates the hinge helices, leading to substantially weaker DNA binding. The drawing is based on PDB entries 1EFA and 2P9H.

O_2 and O_3, that are 82 bp upstream in the *lacI* gene and 410 bp downstream in the *lacZ* gene.[37,38] The repressor binds considerably more tightly to DNA when DNA looping by binding to two operators is possible compared to its DNA binding when only a single operator is available.[39] This is the case because the free energy available from binding to a pair of operators exceeds the energetic cost of bending the DNA into a loop plus the energy of simultaneously distorting the repressor to accommodate the partially looped DNA. When only a single *lac* operator is available, DNA looping is not possible, and only two of a repressor molecule's four subunits are involved in binding to the operator. Consequently, the DNA binding affinity is much lower, and the decrease in DNA binding affinity produced by the binding of inducer is also much reduced. This is reflected in the fact that the *in vivo* induction ratio when the pseudo-operators are absent is ∼50-fold.[40]

Inducer binding in the cleft between the N- and C-terminal subdomains tips one with respect to the other, 6−10°. This separates the hinge helices of the DNA binding domain, and they no longer fit well in the minor groove and can no longer interact with one another and mutually stabilize their helical structures. Without their mutual stabilizing interactions, as well as the loss of some interactions with the N subdomains, they unfold. Thus, the favored state becomes one in which the repressor is dissociated from DNA and the hinge helix is unfolded. This allosteric mechanism can be considered to be partly enthalpic, resulting from the loss of the hinge helix interactions in the minor groove, and partly entropic, resulting from the long effective linker between the core of the repressor and the headpiece containing the H−T−H motif.

The mechanism of the response of *lac* repressor is pleasing and plausible, but should it be considered as proven? It would be comforting to know that in the presence of IPTG, the fraction of dimers in repressor molecules containing a folded DNA binding domain capable of binding to the *lac* operator has been reduced by a factor of at least 50. As in the case of CAP, the mechanism indicated by the structures is compatible with the direction of the change in affinity induced by the ligand effector, but whether the structural change is sufficient to generate the magnitude of the observed change in the DNA binding affinity is yet to be demonstrated.

Do homologous transcription factors possess similar allosteric mechanisms for modulating their DNA binding? In addition to LacI (allolactose and IPTG), a number of LacI homologues and their allosteric effectors are known: PurR (hypoxanthine), GalR (galactose), GalS (galactose), FruR (fructose 1-phosphate), TreR (trehalose 6-phosphate), RbsR (ribose), CytR (cytidine), CelR (cellobiose), and AscG (unknown). Their sequences are all sufficiently similar to that of LacI that they most certainly possess the same basic tertiary structure.[41] This allowed the construction and characterization of chimeras consisting of the DNA binding domain from LacI fused to the cores of the other repressors.[42] The abilities of these proteins to bind to the *lac* operator and repress the expression of *lacZ* could then be easily measured. All except the CytR and AscG chimeras changed their DNA binding appropriately in response to the correct allosteric effector. CytR is different from the rest of its family members. In response to its effector ligand, cytidine, CytR does not alter its DNA binding; rather, it changes its ability to interact with CAP, which under some conditions binds on either side of DNA-bound CytR.[28,43] Therefore, it is not surprising that the LacI−CytR chimera did not change its DNA binding affinity in response to cytidine. Because the effector of AscG is unknown, no change in DNA binding could be tested. Despite the commonality in the mechanism of response to effectors, additional protein-specific interactions between the DNA binding and ligand binding domains must be involved. This is seen by the fact that the tightness of DNA binding and the magnitude of the allosteric effects were also sensitive to the chimeric interface between the domains, that is, to the interactions between residues of the *lac* DNA binding domain and the fused allosteric core.

■ TETR REPRESSOR

The homodimeric TetR repressor controls its own expression and that of a membrane-bound tetracycline exporter, TetA, in response to the presence of Mg-tetracycline (MgTc). X-ray crystallographic structures of TetR, MgTc-TetR, and *tetO*-TetR[44−46] together appear to suggest a mechanism by which the binding of tetracycline to TetR can reduce the protein's affinity for *tetO*. In each subunit, one end of the four-ring tetracycline structure is held to the rigid base structure of the protein (Figure 6). Mg^{2+} is chelated between the two middle rings of tetracycline. The Mg^{2+} interacts with a histidine residue on helix 4 and a threonine residue on helix 6, the combined effect of which is to draw helix 4 toward the tetracycline. That is, helix 6 partially unwinds, and helix 4 rotates by 5° about a point near one end. This then moves the other end by ∼1.5 Å in a swinging motion (Figure 6). As this end is connected to the DNA binding domain, which consists of a helix and the oft-seen H−T−H motif, the DNA binding motif is also shifted by ∼1.5 Å. Analogous movements in the other subunit similarly

dx.doi.org/10.1021/bi400968e | *Biochemistry* 2013, 52, 6755−6765

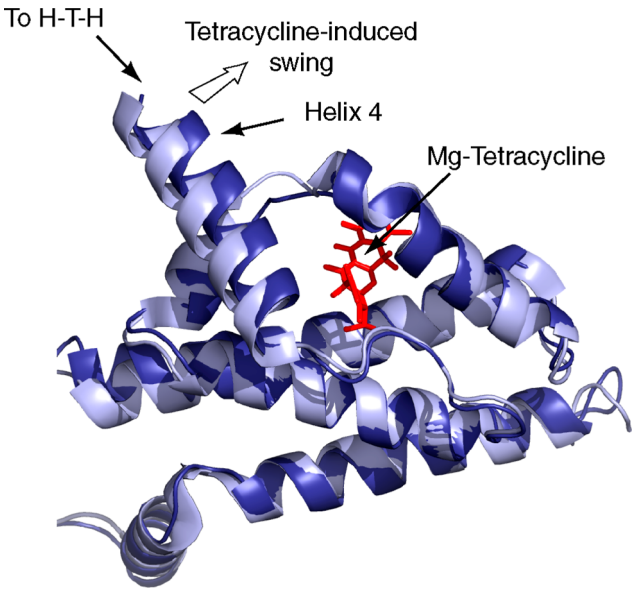


Figure 6. Core of one monomer of the dimeric TetR repressor with the apo form colored light blue (PDB entry 1QPI) and bound MgTc colored dark blue (PDB entry 2TCT). The DNA binding domain is connected to helix 4. The bottom half of the structure forms a solid base that joins the base of the other subunit. In response to the binding of MgTc, helices 4 of the two subunits twist with respect to this base, altering the separation of the two H−T−H motifs by a total of 3 Å.

move the other H−T−H motif by 1.5 Å, producing a total shift in their separation by 3 Å, a distance that is presumed to be sufficient to substantially weaken the repressor's affinity for *tetO*. Independent evidence consistent with domain movement, the extrinsic mechanism, comes from flexible linker DNA experiments as described earlier that show that tetracycline changes the relative positions of the DNA binding domains of the *tet* repressor rather than changing the intrinsic DNA binding affinity of each DNA binding domain.[12]

On the basis of indirect data from methylation protection experiments and using extensive calculations to simulate a system with 9 reactants and 12 equilibrium constants, Lederer et al.[47] concluded that the binding of two molecules of MgTc to TetR reduces the affinity of the repressor for *tetO* by a factor of 10^9! This number probably is much too large because it implies that a free energy greater than $RT \ln(10^9) = 12.2$ kcal/mol would be required to distort TetR to undo the 3 Å mispositioning induced by the binding of MgTc. For comparison, the stability against unfolding most proteins is 5−10 kcal/mol. As a result of the energetic discrepancy, it seems doubtful that the 3 Å shift is the whole story and/or the reduction in DNA binding affinity induced by MgTc is much less than 10^9.

In bacteria, multidrug resistance is provided by systems that efflux a wide variety of drugs or toxic compounds. The expression of one such system in *Staphylococcus aureus* is under control of the QacR repressor protein, a homologue of TetR, that responds to a number of cationic lipophilic drugs. This then raises two questions. How can one protein specifically bind a number of different small molecules, and is the mechanism by which the small molecule effectors modulate the DNA binding affinity of QacR the same for all effectors and essentially the same as that of TetR for the binding of MgTc? While a large number of genetic and biophysical measurements on QacR have not been performed, we do have the crystal structure of the protein bound to DNA in the absence of effector and also the structure when not bound to DNA, but with several different effectors bound.[48] Apparently, the effector binding pocket consists of two partially overlapping binding sites. Although different effectors use mainly these two sites, each different effector binds somewhat differently from the others.

The variability in the binding of different effectors to QacR is similar to what has been observed in some other proteins. Analogues structurally similar to a binding molecule often are seen to bind in the active site, but often in a manner somewhat different from that of the binding molecule. It is as though some binding sites in proteins have evolved to possess multiple but overlapping binding sites, perhaps to increase binding rates or the apparent binding affinity.

Despite the variable binding of effectors in the binding site of QacR, the mechanism of the response to all is pretty much the same. The ligand binding pocket of QacR increases in volume. This is opposite to the change induced in TetR by the binding of tetracycline. In QacR, the binding pushes helix 4 outward, and as in CAP, the binding of the effector stabilizes the helical state of several residues, lengthening helix 5. The result is a translation and shift in the position of the DNA binding domain that increases the separation between the protein's two DNA binding domains by 11 Å.

TRPR REPRESSOR

The dimeric TrpR repressor controls its own expression and that of the enzymes required for the uptake and synthesis of tryptophan. When tryptophan is present at appropriate levels, it binds to TrpR, stimulating the repressor's binding to various *trp* operators and repressing transcription. TrpR consists of a rather solid core that is formed by the interlocked folding of the two subunits. Tryptophan binds between this core and the C-terminal DNA binding domain that, as in many other transcription factors, consists of a H−T−H motif. The binding of tryptophan pushes the two DNA binding domains outward into positions such that the pair can bind to adjacent major groove regions of the *trp* operator.[49] In addition to positioning the two H−T−H domains by the bound tryptophan, the heterocyclic amino nitrogen of the bound tryptophan corepressor is positioned to make a good hydrogen bond to a phosphate oxygen of the *trp* operator,[50] and binding experiments have shown that this interaction contributes significantly to operator binding by TrpR.[51] Few biophysical data appear to be available for comparison of experiments with predictions based on this mechanism.

OMPR AND LAMBDA PHAGE REPRESSOR

As mentioned earlier, cooperativity in DNA binding affinity can be generated by oligomerization. Thus, it is not surprising that transcription factor activity can be regulated by regulating oligomerization. OmpR, a transcriptional activator of genes affecting osmotic pressure in *E. coli*, is part of a two-component regulatory system. EnvZ phosphorylates OmpR in response to osomotic stress, and as a result of the phosphorylation, OmpR dimerization is stimulated, which greatly increases its DNA binding affinity.[52] The structural basis of the phosphorylation-induced dimerization is not, however, understood, although it is known that the dimerization interface includes a β-sheet region.[53] The dimerization reaction could occur in solution, to

dx.doi.org/10.1021/bi400968e | *Biochemistry* 2013, 52, 6755−6765

be followed by binding to DNA, or it could occur between monomers bound to DNA. The concentration of OmpR monomers, the equilibrium binding constant of monomers for binding to DNA, and the various rate constants all determine which of the dimerization reactions is more relevant *in vivo*. Likely, it is dimerization in solution.[54]

Lambda phage and some other lysogenic phage utilize a mechanism for modulating DNA binding that is the opposite of that used by OmpR. During the repair of DNA damage, such as that caused by UV irradiation, RecA protein polymerizes along one strand of melted DNA. Under these conditions, RecA stimulates an inherent self-cleavage activity of lambda repressor.[55] The repressor cleaves itself in an unstructured connector between the DNA binding domain and the dimerization domain.[56] As a result, the monomeric DNA binding domains bind to DNA with an affinity much weaker than that of uncleaved oligomeric repressor, repression ends, and the phage induces.

SUMMARY

The mechanisms by which CAP, LacI, TetR, and TrpR modulate DNA binding affinity in response to the binding of small molecule effectors exhibit important similarities and important differences. It is not surprising that helix−coil transitions play an important role (CAP, LacI, and TetR). The differences in stiffness and the dimensions between the two structural states are dramatic, and the cooperative nature of forming an α helix makes the helix−coil transition a useful bistable switch. All of the proteins considered here utilize the H−T−H DNA binding motif. It will be interesting to see whether and how other DNA binding motifs are used in ligand-modulated DNA binding proteins. For CAP, LacI, and TrpR, the interactions between the protein and ligand effector that initiate the conformational changes are reasonably clear. In CAP, interactions between the adenine of cAMP and residues of a region of quasi-stable helix stabilize the helical form. In LacI, inducer binding appears to wedge a cleft more open, and in TrpR, the corepressor both contacts DNA directly and prevents the DNA binding domains from occupying positions incompatible with contacting both halves of the *trp* operator. The role of the ligand in TetR is less transparent.

A deficiency in all the proposed mechanisms for altering DNA binding affinity of transcription factors is that we cannot yet test whether they lead to the appropriate changes in DNA binding affinity. Merely seeing the DNA binding domains positioned such that correct contacts with DNA are impossible or positioned such that only one of two DNA binding domains can contact a protein's two DNA half-sites does not show that the protein would bind significantly less tightly than if the DNA binding domains were perfectly positioned. If the cost in the free energy of correctly positioning the binding domains is low, the protein will bind nearly as tightly as would one in which the domains were correctly positioned. As discussed earlier, what is relevant is the energetic costs of correctly positioning the DNA binding domains. This cost can be primarily enthalpic if bonds must be made or broken, or it can be primarily entropic if moving the domains around is easy; however, there are many different positions possible and only a few that are compatible with binding. The insufficiencies of the structural models for predicting DNA binding affinity changes are not unique to this class of protein. They arise whenever the energetics of conformational changes must be extracted from a knowledge of structure only. Before changes in DNA binding affinity can be significantly understood, it will be necessary to be able to calculate accurately the free energy difference between the DNA binding state and the nonbinding state(s) of a protein. While it appears that we understand the principles of such calculations, their complexity necessitates the use of approximations that significantly limit accuracy.[57] Thus, it is necessary at present to be cautious in attempting to infer mechanism from structures alone.

AUTHOR INFORMATION

Corresponding Author

*E-mail: rfschleif@jhu.edu. Phone: (410) 516-5206.

Funding

This work supported by National Science Foundation Grant 1021031.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

I thank Michael Rodgers for many discussions on the topics covered in this review and for comments on the manuscript.

REFERENCES

(1) Kallipolitis, B., Norregaard-Madsen, M., and Valentin-Hansen, P. (1997) Protein-protein communication: Structural model of the repression complex formed by Cytr and the global regulator CRP. *Cell 89*, 1101−1109.

(2) Monod, J., Wyman, J., and Changeux, J. (1965) On the nature of allosteric transitions: A plausible model. *J. Mol. Biol. 12*, 88−118.

(3) Koshland, D. E., Jr., Nemethy, G., and Filmer, D. (1966) Comparison of experimental binding data and theoretical models in proteins containing subunits. *Biochemistry 5*, 365−385.

(4) Frato, K. E., and Schleif, R. F. (2009) A DNA-assisted binding assay for weak protein-protein interactions. *J. Mol. Biol. 394*, 805−814.

(5) Crothers, D. M., and Metzger, H. (1972) The influence of polyvalency on the binding properties of antibodies. *Immunochemistry 9*, 341−357.

(6) Klemm, J. D., and Pabo, C. O. (1996) Oct-1 POU domain-DNA interactions: Cooperative binding of isolated subdomains and effects of covalent linkage. *Genes Dev. 10*, 27−36.

(7) Schleif, R., and Wolberger, C. (2004) Arm-domain interactions can provide high binding cooperativity. *Protein Sci. 13*, 2829−2831.

(8) Timmes, A., Rodgers, M., and Schleif, R. (2004) Biochemical and physiological properties of the DNA binding domain of AraC protein. *J. Mol. Biol. 340*, 731−738.

(9) Nunoshiba, T., Hidalgo, E., Amabile Cuevas, C., and Demple, B. (1992) Two-stage control of an oxidative stress regulon: The *Escherichia coli* SoxR protein triggers redox-inducible expression of the *soxS* regulatory gene. *J. Bacteriol. 174*, 6054−6060.

(10) Kholodenko, B. N., and Birtwistle, M. R. (2009) Four-dimensional dynamics of MAPK information-processing systems. *Wiley Interdiscip. Rev.: Syst. Biol. Med. 1*, 28−44.

(11) Harmer, T., Wu, M., and Schleif, R. (2001) The role of rigidity in DNA looping-unlooping by AraC. *Proc. Natl. Acad. Sci. U.S.A. 98*, 427−431.

(12) Seedorff, J. E., Rodgers, M. E., and Schleif, R. (2009) Opposite allosteric mechanisms in TetR and CAP. *Protein Sci. 18*, 775−781.

(13) Riggs, A. D., Newby, R. F., and Bourgeois, S. (1970) Lac repressor−operator interaction: II. Effect of galactosides and other ligands. *J. Mol. Biol. 51*, 303−314.

(14) Jobe, A., and Bourgeois, S. (1972) The Lac repressor-operator interaction: VII. A repressor with unique binding properties: The X86 repressor. *J. Mol. Biol. 72*, 139−152.

(15) Kleinschmidt, C., Tovar, K., Hillen, W., and Porschke, D. (1988) Dynamics of repressor-operator recognition: The Tn10-encoded tetracycline resistance control. *Biochemistry 27*, 1094−1104.

(16) Hendrickson, W., and Schleif, R. F. (1984) Regulation of the *Escherichia coli* L-arabinose operon studied by gel electrophoresis DNA binding assay. *J. Mol. Biol. 178*, 611−628.
(17) Lewis, M. (2013) Allostery and the lac operon. *J. Mol. Biol. 425*, 2309−2316.
(18) Riggs, A., Bourgeois, S., and Cohn, M. (1970) The lac repressor-operator interaction III. Kinetic studies. *J. Mol. Biol. 53*, 401−417.
(19) Swint-Kruse, L., Zhan, H., and Matthews, K. S. (2005) Integrated insights from simulation, experiment, and mutational analysis yield new details of LacI function. *Biochemistry 44*, 11201−11213.
(20) Busby, S., and Ebright, R. H. (1999) Transcription activation by catabolite activator protein (CAP). *J. Mol. Biol. 293*, 199−213.
(21) Gama-Castro, S., Salgado, H., Peralta-Gil, M., Santos-Zavaleta, A., Muñiz-Rascado, L., Solano-Lira, H., Jimenez-Jacinto, V., Weiss, V., García-Sotelo, J. S., López-Fuentes, A., Porrón-Sotelo, L., Alquicira-Hernández, S., Medina-Rivera, A., Martínez-Flores, I., Alquicira-Hernández, K., Martínez-Adame, R., Bonavides-Martínez, C., Miranda-Ríos, J., Huerta, A. M., Mendoza-Vargas, A., Collado-Torres, L., Taboada, B., Vega-Alvarado, L., Olvera, M., Olvera, L., Grande, R., Morett, E., and Collado-Vides, J. (2011) RegulonDB Version 7.0: Transcriptional Regulation of *Escherichia coli* K-12 Integrated within Genetic Sensory Response Units (Gensor Units). *Nucleic Acids Res. 39*, D98−D105.
(22) Popovych, N., Sun, S., Ebright, R. H., and Kalodimos, C. G. (2006) Dynamically driven protein allostery. *Nat. Struct. Mol. Biol. 13*, 831−838.
(23) McKay, D. B., and Steitz, T. A. (1981) Structure of catabolite gene activator protein at 2.9 Å resolution suggests binding to left-handed B-DNA. *Nature 290*, 744−749.
(24) Passner, J. M., Schultz, S. C., and Steitz, T. A. (2000) Modeling the cAMP-induced allosteric transition using the crystal structure of CAP-cAMP at 2.1 Å resolution. *J. Mol. Biol. 304*, 847−859.
(25) Sharma, H., Yu, S., Kong, J., Wang, J., and Steitz, T. A. (2009) Structure of apo-CAP reveals that large conformational changes are necessary for DNA binding. *Proc. Natl. Acad. Sci. U.S.A. 106*, 16604−16609.
(26) Popovych, N., Tzeng, S., Tonelli, M., Ebright, R. H., and Kalodimos, C. G. (2009) Structural basis for cAMP-mediated allosteric control of the catabolite activator protein. *Proc. Natl. Acad. Sci. U.S.A. 106*, 6927−6932.
(27) Leu, S., Baker, C. H., Lee, E. J., and Harman, J. G. (1999) Position 127 amino acid substitutions affect the formation of CRP:CAMP:LacP complexes but not CRP:CAMP:RNA polymerase complexes at LacP. *Biochemistry 38*, 6222−6230.
(28) Schultz, S. C., Shields, G. C., and Steitz, T. A. (1991) Crystal structure of a CAP-DNA complex: The DNA is bent by 90°. *Science 253*, 1001−1007.
(29) Garges, S., and Adhya, S. (1985) Sites of allosteric shift in the structure of the cyclic AMP receptor protein. *Cell 41*, 745−751.
(30) Yu, S., Maillard, R. A., Gribenko, A. V., and Lee, J. C. (2012) The N-terminal capping propensities of the D-helix modulate the allosteric activation of the *Escherichia coli* cAMP receptor protein. *J. Biol. Chem. 287*, 39402−39411.
(31) Gallagher, D. T., Smith, N., Kim, S., Robinson, H., and Reddy, P. T. (2009) Profound asymmetry in the structure of the cAMP-free cAMP receptor protein (CRP) from *Mycobacterium tuberculosis*. *J. Biol. Chem. 284*, 8228−8232.
(32) Jobe, A., and Bourgeois, S. (1972) Lac repressor-operator interaction. VI. The natural inducer of the *lac* operon. *J. Mol. Biol. 69*, 397−408.
(33) Barkley, M. D., Riggs, A. D., Jobe, A., and Burgeois, S. (1975) Interaction of effecting ligands with *lac* repressor and repressor-operator complex. *Biochemistry 14*, 1700−1712.
(34) Baumberg, S., and Roberts, M. (1984) Anomalous expression of the *E. coli* Lac operon in *Proteus mirabilis*. *Mol. Gen. Genet. 198*, 166−171.
(35) Bell, C. E., and Lewis, M. (2000) A closer view of the conformation of the Lac repressor bound to operator. *Nat. Struct. Biol. 7*, 209−214.
(36) Bell, C. E., and Lewis, M. (2001) The Lac repressor: A second generation of structural and functional studies. *Curr. Opin. Struct. Biol. 11*, 19−25.
(37) Pfahl, M., Gulde, V., and Bourgeois, S. (1979) "Second" and "Third Operator" of the Lac operon: An investigation of their role in the regulatory mechanism. *J. Mol. Biol. 127*, 339−344.
(38) Gilbert, W., Gralla, J., Majors, J., and Maxam, A. (1975) Specific binding of CAP factor to Lac promoter DNAP. In *Protein-Ligand Interactions* (Sund, H. H., and Blauer, G., Eds.) pp 193−210, Walter de Gryuter, Berlin.
(39) Whitson, P. A., Hsieh, W. T., Wells, R. D., and Matthews, K. S. (1987) Influence of supercoiling and sequence context on operator DNA binding with lac repressor. *J. Biol. Chem. 262*, 14592−14599.
(40) Glascock, C. B., and Weickert, M. J. (1998) Using chromosomal Laciq1 to control expression of genes on high-copy-number plasmids in *Escherichia coli*. *Gene 223*, 221−231.
(41) Sander, C., and Schneider, R. (1991) Database of homology-derived protein structures and the structural meaning of sequence alignment. *Proteins 9*, 56−68.
(42) Meinhardt, S., Manley, M. W., Becker, N. A., Hessman, J. A., Maher, L. J., and Swint-Kruse, L. (2012) Novel insights from hybrid LacI/GalR proteins: Family-wide functional attributes and biologically significant variation in transcription repression. *Nucleic Acids Res. 40*, 11139−11154.
(43) Pedersen, H., Søgaard-Andersen, L., Holst, B., and Valentin-Hansen, P. (1991) Heterologous cooperativity in *Escherichia coli*. The CytR repressor both contacts DNA and the cAMP receptor protein when binding to the deoP2 promoter. *J. Biol. Chem. 266*, 17804−17808.
(44) Orth, P., Schnappinger, D., Hillen, W., Saenger, W., and Hinrichs, W. (2000) Structural basis of gene regulation by the tetracycline inducible tet repressor-operator system. *Nat. Struct. Biol. 7*, 215−219.
(45) Hinrichs, W., Kisker, C., Düvel, M., Müller, A., Tovar, K., Hillen, W., and Saenger, W. (1994) Structure of the Tet repressor-tetracycline complex and regulation of antibiotic resistance. *Science 264*, 418−420.
(46) Orth, P., Cordes, F., Schnappinger, D., Hillen, W., Saenger, W., and Hinrichs, W. (1998) Conformational changes of the Tet repressor induced by tetracycline trapping. *J. Mol. Biol. 279*, 439−447.
(47) Lederer, T., Takahashi, M., and Hillen, W. (1995) Thermodynamic analysis of tetracycline-mediated induction of Tet repressor by a quantitative methylation protection assay. *Anal. Biochem. 232*, 190−196.
(48) Schumacher, M. A., Miller, M. C., Grkovic, S., Brown, M. H., Skurray, R. A., and Brennan, R. G. (2001) Structural mechanisms of QacR induction and multidrug recognition. *Science 294*, 2158−2163.
(49) Zhang, R., Joachimiak, A., Lawson, C. L., Schevitz, R. W., Otwinowski, Z., and Sigler, P. B. (1987) The crystal structure of trp aporepressor at 1.8 Å shows how binding tryptophan enhances DNA affinity. *Nature 327*, 591−597.
(50) Otwinowski, Z., Schevitz, R. W., Zhang, R., Lawson, C. L., Joachimiak, A., Marmorstein, R. Q., Luisi, B. F., and Sigler, P. B. (1988) Crystal structure of Trp represser/operator complex at atomic resolution. *Nature 335*, 321−329.
(51) Marmorstein, R. Q., and Sigler, P. B. (1989) Stereochemical effects of L-tryptophan and its analogs on Trp repressor's affinity for operator-DNA. *J. Biol. Chem. 264*, 9149−9154.
(52) Head, C. G., Tardy, A., and Kenney, L. J. (1998) Relative binding affinities of OmpR and OmpR-phosphate at the ompF and ompC regulatory sites. *J. Mol. Biol. 281*, 857−870.
(53) Maris, A. E., Walthers, D., Mattison, K., Byers, N., and Kenney, L. J. (2005) The response regulator OmpR oligomerizes via β-sheets to form head-to-head dimers. *J. Mol. Biol. 350*, 843−856.
(54) Barbieri, C. M., Wu, T., and Stock, A. M. (2013) Comprehensive analysis of OmpR phosphorylation, dimerization,

and DNA binding supports a canonical model for activation. *J. Mol. Biol. 425*, 1612−1626.
(55) Little, J. W. (1984) Autodigestion of LexA and phage lambda repressors. *Proc. Natl. Acad. Sci. U.S.A. 81*, 1375−1379.
(56) Sauer, R. T., Ross, M. J., and Ptashne, M. (1982) Cleavage of the lambda and P22 repressors by recA protein. *J. Biol. Chem. 257*, 4458−4462.
(57) Jiao, D., Golubkov, P. A., Darden, T. A., and Ren, P. (2008) Calculation of protein-ligand binding free energy by using a polarizable potential. *Proc. Natl. Acad. Sci. U.S.A. 105*, 6290−6295.