jmb Article

Indirect Readout of DNA Sequence by P22 Repressor: Roles of DNA and Protein Functional Groups in Modulating DNA Conformation

Lydia-Ann Harris[1], Derrick Watkins[2], Loren Dean Williams[2] and Gerald B. Koudelka[1]

*1 - **Department of Biological Sciences,** University at Buffalo, 607 Cooke Hall, Buffalo, NY 14260, USA*
*2 - **School of Chemistry and Biochemistry,** Georgia Institute of Technology, Atlanta, GA 30332-0400, USA*

***Correspondence to Gerald B. Koudelka:** koudelka@buffalo.edu*
http://dx.doi.org/10.1016/j.jmb.2012.10.008
Edited by B. Connolly

Abstract

The repressor of bacteriophage P22 (P22R) discriminates between its various DNA binding sites by sensing the identity of non-contacted base pairs at the center of its binding site. The "indirect readout" of these non-contacted bases is apparently based on DNA's sequence-dependent conformational preferences. The structures of P22R–DNA complexes indicate that the non-contacted base pairs at the center of the binding site are in the B′ state. This finding suggests that indirect readout and therefore binding site discrimination depend on P22R's ability to either sense and/or impose the B′ state on the non-contacted bases of its binding sites. We show here that the affinity of binding sites for P22R depends on the tendency of the central bases to assume the B′-DNA state. Furthermore, we identify functional groups in the minor groove of the non-contacted bases as the essential modulators of indirect readout by P22R. In P22R–DNA complexes, the negatively charged E44 and E48 residues are provocatively positioned near the negatively charged DNA phosphates of the non-contacted nucleotides. The close proximity of the negatively charged groups on protein and DNA suggests that electrostatics may play a key role in the indirect readout process. Changing either of two negatively charged residues to uncharged residues eliminates the ability of P22R to impose structural changes on DNA and to recognize non-contacted base sequence. These findings suggest that these negatively charged amino acids function to force the P22R-bound DNA into the B′ state and therefore play a key role in indirect readout by P22R.

Introduction

Most studies of DNA recognition by proteins primarily focus on direct sequence readout. Direct readout exploits chemically complimentary electrostatic and hydrophobic interactions between amino acid residues in the protein and the base pairs of the binding site. However, direct readout is only part of the mechanism used by proteins to recognize and/or discriminate between specific binding site(s).[1] It is now clear that non-contacted bases within the binding sites govern the ability of many different proteins to recognize their cognate DNA binding site(s).[2] Several lines of evidence suggest that such "indirect readout" of the sequence of these non-contacted bases is based on the different propensities of various DNA sequences to assume or to resist various altered conformation states. Therefore, these proteins distinguish between sequences by exploiting DNA sequence-dependent differences in conformational polymorphism and deformability.

The potential role of DNA conformation and deformability in sequence recognition by proteins was initially recognized by Koudelka *et al.*[3] and later put in a quantitative framework by Olson *et al.*[4] McFail-Isom *et al.* have described how DNA electrostatic potential and charged species can alter DNA conformation, especially minor groove width.[5] More recently, Rohs *et al.* have proposed that electrostatic potential and minor groove width are recognized by proteins.[6] Experimentally determined relationships between DNA flexibility and protein affinity have been

reported (e.g., Refs. 7 and 8). It is now known that proteins utilizing indirect readout are found in organisms arrayed in all domains of life (e.g., see Refs. 7–11). Despite the prevalence and biological importance of indirect readout, the mechanism by which DNA sequence governs DNA structure and flexibility is not well understood.

We are exploring DNA binding by bacteriophage P22 repressor protein to help illuminate the mechanism of indirect readout of DNA sequence by proteins. P22R is a DNA binding protein encoded by the c2 gene of the lambdoid bacteriophage P22. After infecting a host, this temperate lambdoid phage either grows lytically or forms a lysogenic prophage. By repressing the transcription of genes needed for lytic growth and stimulating transcription of genes needed for the maintenance of lysogeny, P22R serves as the master regulator of genes needed for establishment and maintenance of lysogeny. Our earlier results indicate that, in indirect readout, the affinity of P22 repressor for DNA is modulated by the sequence dependence of the free energy required to induce a specific DNA deformation.[12,13]

The bifunctional transcriptional regulatory activities of P22R require that it binds to and discriminates between six naturally occurring sites with different affinities. These six repressor binding sites are similar in sequence. The outermost bases are highly conserved between these sites, but the bases at the center of the binding site are not.[14] P22R makes direct base-specifying contacts with the conserved bases but makes no direct base-specifying contacts with the four central bases.[15] Despite the lack of direct repressor contacts, the affinity of natural and synthetic operators for P22R depends on the sequence of the central non-contacted bases.[14] Thus, P22R discriminates between sites by the indirect readout of four central non-contacted bases. This aspect of P22R's DNA sequence recognition is crucial to repressor's role as a regulator of bacteriophage gene expression.

The P22R binds DNA as homodimer. The DNA binding N-terminal domain of P22R contains a helix–turn–helix structural motif. Two “recognition” α-helices of this motif, one from each protein monomer, lie in successive major grooves on one face of the DNA. These contacted major grooves are separated by the four non-contacted bases, whose minor groove surfaces face P22R (Fig. 1). X-ray structural analysis of the P22R complexes showed that, regardless of its base sequence, the non-contacted DNA at the center of the binding site assumes the B′-DNA state,[15] with the balance of the bound DNA in the B form. B′-DNA differs from B-DNA in two significant ways: (1) at 2.2 Å, the minor groove of B′-DNA is nearly 10 Å narrower than that of B-DNA,[15,16] and (2) a multilayered solvent spine runs along the minor groove of B′-DNA. This spine is absent in B-DNA.

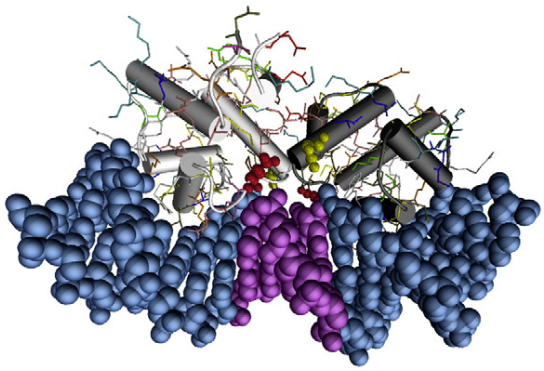

Fig. 1. Global view of the P22R–NTD–DNA complex. Amino acids in the two DNA binding P22 repressor subunits are colored by type. The secondary structure of each subunit is highlighted schematically in light and dark gray, respectively. Binding site DNA is represented in space filling with the central non-contacted bases highlighted in purple and the contacted bases colored light blue. Amino acids E44 and E48 (one from each subunit) are represented in ball and stick and colored red and yellow, respectively.

Based on our earlier findings,[12,14,15,17,18] we hypothesized that the indirect readout mechanism in P22R involves the induction of B′ state in the non-contacted bases of an operator. To test this hypothesis, we varied the composition of functional groups present on the minor groove surface of the non-contacted bases, changes that would influence the ability of these bases to assume a B′-DNA form. We found that the affinity of a given binding site for P22R correlates with the propensity of a given sequence to assume the B′-DNA form. Our findings also indicate that the 2-NH_2 group in the minor groove is the primary mediator of indirect readout by P22R. We also examined how P22R senses DNA conformation. Our data show that the direct juxtaposition of negatively charged amino acid residues with the DNA phosphates attached to the non-contacted bases is required for indirect readout of DNA sequence by P22R. Spectroscopic studies indicate that the negative charges are crucial to the ability of P22R to induce the formation of B′-DNA in the non-contacted region of the binding site. Together, our results suggest that the mechanism of indirect readout and thus binding site discrimination by P22R involves repressor-mediated introduction of the B′ state in the non-contacted region.

Results

Indirect readout of DNA sequence is governed by the minor groove functional groups

In P22R–DNA complexes, the non-contacted region of the binding site adopts the B′ state, regardless of the

sequence of bases in this region.[12,15] The B′ state is native to A-tract DNA[19–21] and can be induced in A/T-rich sequences.[22,23] However, G/C-containing DNAs were not found in the B′ state prior to the work performed by Watkins *et al.* where it was shown that P22R "trapped" the DNA in the B′ state.[24] Based on the crystal structures and results of earlier binding and footprinting studies,[14,17] we hypothesize that the DNA affinity of P22R depends in part on the ease with which the non-contacted bases can be induced to assume the B′ state.

To test this model, we determined the affinities of P22R for synthetic operator sites that differ only at non-contacted positions 9 and 10 (Table 1), where DNA^{9T} contains a TpA base step. Consistent with previous results,[14] P22R binds to DNA^{9T} with ~11-fold higher affinity than it does to DNA^{9C}, which bears a CpG step at positions 9 and 10. To identify the features of C•G base pairs that are responsible for the lower affinity of DNA^{9C} for P22R, we symmetrically substituted the guanine at the central positions in DNA^{9C} with inosine. This substitution gives the DNA$^{9C/I}$ sequence. Inosine (I) resembles guanine in the major groove but lacks a $-NH_2$ group at the 2-position. The floor of the DNA$^{9C/I}$ minor groove surface is identical with that of DNA^{9T}. P22R binds to DNA$^{9C/I}$ with an affinity nearly 15-fold higher than its affinity for DNA^{9C} (Table 1). P22R's affinity for DNA$^{9C/I}$ is nearly identical with that for DNA^{9T}. These findings show that the presence of the 2-NH_2 group on the minor groove floor is substantially responsible for destabilizing the P22R–DNA^{9C} complex, relative to the P22R–DNA^{9T} complex.

To determine if functional groups in the non-contacted major groove of C•G *versus* T•A base pairs play a role in modulating the affinity of DNA for P22R, we symmetrically substituted the thymine at positions 9 and 10 with a uracil creating the (DNA^{9U}) sequence. This modification allows us to determine if the methyl group in the major groove of an A•T base pair plays a role in the indirect readout by P22R. In contrast to the effect of the minor groove changes, the affinity of P22R for DNA^{9U} does not differ significantly ($p<0.005$) from its affinity for DNA^{9T} (Table 1). These results also show that the major groove functional groups on DNA bases in the central non-contacted positions in the binding site have little role in indirect readout of DNA sequence by P22R. Together with results obtained with DNA$^{9C/I}$, these findings show that the 2-NH_2 group is the primary determinant of indirect readout of DNA sequence by P22R.

B′-DNA is the native state of A-tract DNA, and outside of P22R–DNA complexes, only A/T-containing DNA sequences have been found to form B′-DNA. The only G/C-containing DNA found in the B′ state is that located in the non-contacted region of the P22R–DNA complex.[12] Consistent with the idea that the propensity of a particular sequence to form B′-DNA forms the basis for indirect readout of DNA sequence by P22R, we find that P22R binds an operator containing a central A-tract DNA^{9A} repressor with highest affinity (Table 1). The order of preference of P22R for the various DNAs examined is DNA^{9A} ≥ DNA$^{9C/I}$ ≥ DNA^{9U} ≥ DNA^{9T} ⋙ DNA^{9C}. Since DNA^{9A}, DNA^{9U}, DNA^{9T} and DNA$^{9C/I}$ would be anticipated to readily assume B′ state, but DNA^{9C} would not, our results are consistent with the suggestion that a binding site's affinity for P22R is governed by the ability of the non-contacted bases to assume the B′ state.

E44 and E48 mutations in P22 repressor cause a loss of non-contacted base specificity

Our previous findings suggest that P22R binding induces the DNA at the center of P22R binding site to assume the B′ state. The crystal structure of the P22R–DNA^{9T} and P22R–DNA^{9C} complexes[12,15] shows that the glutamates at positions 44 and 48 are positioned near the DNA phosphates of the non-contacted nucleotides (Fig. 1). We hypothesize that charge repulsion between the closely juxtaposed negatively charged amino acids and DNA phosphates forces the non-contacted minor groove to narrow thereby helping to impose the B′ state on the

Table 1. Relative affinities of wild-type P22 repressor protein for the central sequence variant synthetic P22 operators

Position	1	2	3	4	5	6	7	8	9		10	11	12	13	14	15	16	17	18	Relative K_D (K_D^{9x}/K_D^{9T})
9T	A	T	T	T	A	A	G	A	T	•	A	T	C	T	T	A	A	A	T	1.0±0.2
9C	A	T	T	T	A	A	G	A	C	•	G	T	C	T	T	A	A	A	T	10.6±0.3
9C/I	A	T	T	T	A	A	G	A	C	•	I	T	C	T	T	A	A	A	T	0.7±0.3
9U	A	T	T	T	A	A	G	A	U	•	A	T	C	T	T	A	A	A	T	0.8±0.3
9A	A	T	T	T	A	A	G	A	A	•	A	T	C	T	T	A	A	A	T	0.6±0.2

The sequences of the synthetic sites are shown along one strand of DNA. The center of symmetry of all the sequences lies between bases 9 and 10 and is highlighted by the *filled circle*. Binding affinities are given as the relative dissociation constant, that is, dissociation constant complex between P22 repressor and central sequence variant operators (K_D^{9x}) divided by the dissociation constant of the P22 repressor–9T complex (K_D^{9T}) plus/minus the standard deviation. The apparent dissociation constant (K_D) is defined as the concentration of P22 repressor monomers needed to half-maximally occupy a binding site. Under the conditions of these experiments, P22 repressor would be anticipated to be 100% monomeric; hence, any differences in apparent K_D reflect changes in the stability of the P22 repressor–DNA complexes. The protein concentrations used were corrected for activity. The value of the apparent K_D of the P22 repressor–9T binding site complex is 14.3×10^{-9} M.

Table 2. Affinities of wild-type and mutant P22 repressors for the 9T and 9C operators

Protein	K_D (nM)		
	DNA^{9T}	DNA^{9C}	K_D^{9C}/K_D^{9T}
Wild type	14.3±0.8	151.8±28	10.6
E44A	7.5±1.48	15.7±4.3	2.1
E44Q	1.2±0.6	3.0±1.2	2.5
E44D	13.9±1.6	>173	>12.5
E48A	2.7±0.7	7.8±1.1	2.8
E48Q	2.9±1.6	6.5±2.7	2.2

Affinities of each protein for these operators are given as the dissociation constants (K_D, nM) plus/minus the standard deviation. The symbol (>) in front of K_D indicates that the P22 repressor variant binds to that operator only nonspecifically. Also shown is the ratio (K_D^{9C}/K_D^{9T}) of the dissociation constants of each protein for 9T and 9C.

DNA in this region. To test this idea, we separately changed the identity of the amino acids at positions 44 and 48 and compared the ability of the wild-type and mutant repressors to bind to and discriminate between DNA^{9T} and DNA^{9C}.

As opposed to the strong preference of wild-type P22R for a binding site with a central sequence containing a TpA step over one containing a CpG (Table 2), P22R bearing an E44A substitution has a strongly reduced ability to discriminate between DNA^{9T} and DNA^{9C}. The E44A change both shortens the side chain and eliminates the charge on the amino acid at position 44 (Table 2). Similarly, changing E44 to Q, whose side chain is approximately the same length as that of glutamate but is also uncharged under these conditions, also reduces the ability of P22R to discriminate between DNA^{9T} and DNA^{9C}. As compared to the ~11-fold difference in affinity of wild-type P22R for DNA^{9T} and DNA^{9C}, the E44Q mutant P22R binds DNA^{9T} with only 2.5-fold higher affinity than it does to DNA^{9C} (Table 2). The finding that both E44A and E44Q mutations reduce P22R's ability to distinguish between DNA^{9T} and DNA^{9C} is consistent with the idea that a negatively charged amino acid at position 44 contributes to indirect readout.

To confirm this conclusion, we substituted E44 with aspartate (D), a charge conservative mutation. The E44D mutant P22R binds DNA^{9T} with an affinity similar to that of wild-type P22R but binds so poorly to DNA^{9C} that its affinity for this site is close to that for nonspecific DNA, that is, the affinity of E44D for DNA^{9T} is ≥12.5-fold higher than its affinity for DNA^{9C} (Table 2). Taken together, these findings confirm the suggestion that a negative charge on the amino acid at position 44 is partially responsible for P22R's preference for DNA^{9T} over DNA^{9C}, a preference that is the hallmark of indirect readout by P22R.

The results obtained with the E48A and E48Q mutant P22 repressors mirror the findings we obtained with E44 mutant P22 repressor proteins (Table 2). Specifically, P22R E48A binds DNA^{9T} with only a 2.8-fold higher affinity than it does to DNA^{9C}. E48Q binds DNA^{9T} with only a 2.2-fold higher affinity than it does to DNA^{9C}. Unfortunately, we were unable to express the E48D protein in our system, and thus, we could not examine the effect of a charge conservative mutation at position 48 on non-contacted base preference. Nonetheless, up to 5-fold reduction in the ability of E48A and E48Q mutant P22R to discriminate between DNA^{9T} and DNA^{9C} also indicates that the negative charge on the residue at position 48 also is important in allowing P22 repressor to discriminate between various in non-contacted base sequences.

Role of negative charge at position 44 in altering P22R binding site structure

Indirect readout of the non-contacted bases by P22R is apparently governed by induction of the B′-DNA state (Table 1). Our results show that substituting the negatively charged E44 and E48 with neutral amino acids essentially eliminates the ability of P22R to discriminate between DNA^{9T} and DNA^{9C}. The reduced ability of these P22R mutants at positions 44 and 48 to discriminate between DNA^{9T} and DNA^{9C} suggests that these mutant repressors bind to B-DNA rather than to B′-DNA. To test this hypothesis, we used circular dichroism (CD) to assess the changes in DNA conformation upon formation of DNA complexes with wild-type, E44A and E44D mutant P22R. Protein absorbance dominates the far-UV (180 m to 260 nm) CD spectra, but in the near-UV range (260–300 nm), DNA absorbance dominates the CD profile.[25,26] In mixtures of protein and DNA, the observed spectra are additive; therefore, the protein spectrum can be subtracted from that of the complex leaving the spectrum of bound DNA. CD is a useful probe for assessing DNA structural changes. DNA CD depends on the stacking of the bases.[27] The CD intensity of DNA in the 260- to 300-nm range is extremely sensitive to propeller twist[28] and helical twist[29] and hence can readily monitor the deformations that result from a B→B′-DNA transition.

The spectra of the unbound DNA^{9T} and DNA^{9C} differ slightly, but overall, both resemble that of B-DNA, displaying a positive peak between 260 and 280 nm and a negative peak around 245 nm (Fig. 2a). As indicated by the negative bands at 190 nm and 210–222 nm, the spectra of the wild-type, E44A and E44D P22R, the secondary structure of each of these proteins is composed of a combination of α-helix and β-sheet (Fig. 2b). This observation fits with the expectation that the predominant secondary structure in the N-terminal domain is α-helix and that of the C-terminal domain is β-sheet.[30–32] As judged by the similarity of the CD spectra of the wild-type P22R, E44A and E44D, the

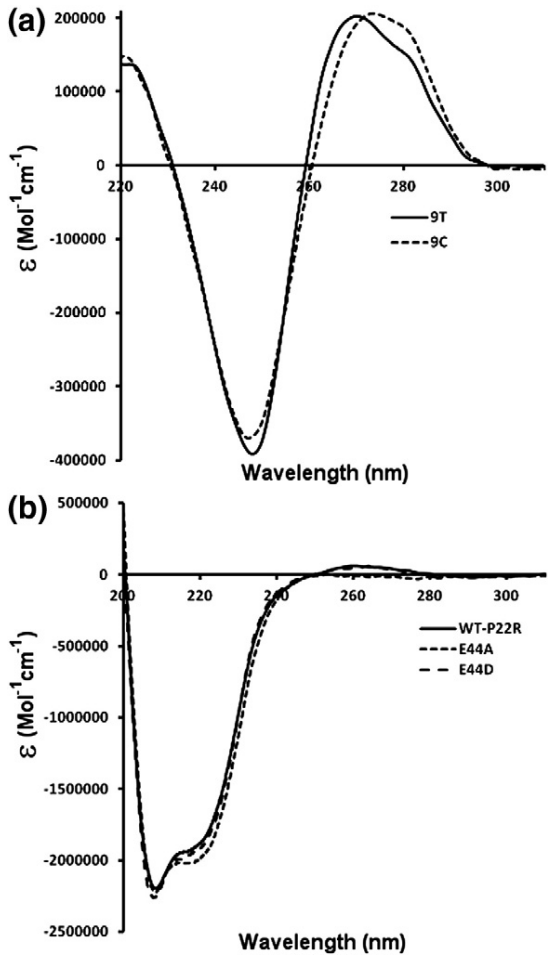

Fig. 2. CD spectra of (a) unbound 9T and 9C DNAs and (b) wild-type and E44x mutant P22 repressor proteins. Spectra were recorded as described in Materials and Methods. In (a), the spectrum of 9T is given as a continuous line and that for 9C is given as a broken line. In (b), the spectra of wild-type P22R, E44A and E44D are given as a continuous line, a broken line and a dashed line, respectively.

mutations apparently do not perturb the secondary structure content of P22R.

We analyzed the near-UV (250–300 nm) CD spectra of DNA^{9T} and DNA^{9C} in complex with P22R and compared it to the spectra of the respective unbound DNA (Fig. 3). Regardless of central sequence, wild-type P22R binding to DNA causes an increase in amplitude of the positive CD peak between 260 and 280 nm (Fig. 3a and b). The binding of P22R to DNA^{9T} increases the CD amplitude at ~280 nm from 20,000 deg M^{-1} cm^{-1} to 27,000 deg M^{-1} cm^{-1}. A smaller increase in CD amplitude of the positive CD peak between 260 and 280 nm is seen when P22R is added to DNA^{9C}. In contrast, the CD spectrum of DNA that does not contain a P22R binding site (i.e., "nonspecific DNA") is unaffected by the addition of P22R (Fig. 3c). This finding indicates that increases in peak amplitudes seen in mixtures of P22R with DNA^{9T} and DNA^{9C} are due to P22R-induced DNA structural changes that result from specific protein–DNA interaction.

The effect of E44D repressor binding on the CD spectra of DNA^{9T} and DNA^{9C} is similar to that seen with addition of wild-type P22R (Fig. 3d–f). That is, there is an increase in the amplitude of the positive peak between 260 and 280 nm in E44D-bound DNA^{9T} and DNA^{9C} that is not observed when this protein is mixed with nonspecific DNA. Thus, E44D, like wild-type P22R, induces a DNA conformational change when it binds. In contrast to wild-type and E44D repressors, the binding of E44A mutant repressor, a protein that shows reduced discrimination between non-contacted base sequences, induces a smaller increase in the amplitude-positive, 260- to 280-nm CD peak of bound DNA (Fig. 3g and h). This observation indicates that E44A binding to DNA only negligibly affects DNA conformation.

Binding of P22R, E44 and E48 mutants to O_R1 and O_R2 natural sites: the role of charge in repressor's discrimination between naturally occurring operators

The negatively charged residues at positions 44 and 48 are necessary to allow P22R to discriminate between non-contacted base sequence variants DNA^{9T} and DNA^{9C} (Table 1). Since the sequences of naturally occurring P22R binding sites primarily differ in their non-contacted regions,[14] we hypothesized that residues 44 and 48 also have a role in allowing P22R to discriminate between its naturally occurring sites. We tested this hypothesis by comparing the relative affinities of wild type, E44A, E44Q and E44D for P22 O_R1 and P22 O_R2. Wild-type P22R binds to O_R1 with >16-fold higher affinity than it does to O_R2 (Table 3). In contrast, both E44A and E44Q mutant repressors have a reduced ability to distinguish between O_R1 and O_R2. The E44A mutation reduces that ability of P22R to distinguish between these sites by nearly 5-fold; that is, E44A binds O_R1 with a 3.7-fold higher affinity that it does to O_R2 (Table 3). Similarly, the E44Q mutant protein binds O_R1 with only an ~5-fold higher affinity that it does to O_R2 (Table 3). The reduced ability of the E44A and E44Q mutant proteins to distinguish between O_R1 and O_R2 mirrors the results obtained with these proteins using the non-contacted base-variant DNA^{9T} and DNA^{9C} sites (Table 2). E44D binds to O_R1 with similar affinity as does wild-type P22R. However, E44D binds poorly to O_R2 and its affinity for this site is close to that for nonspecific DNA. Hence, the affinity of E44D for O_R1 is ≥7-fold higher than its affinity for O_R2. This weak binding of E44D to O_R2 implies that E44D, like the wild-type

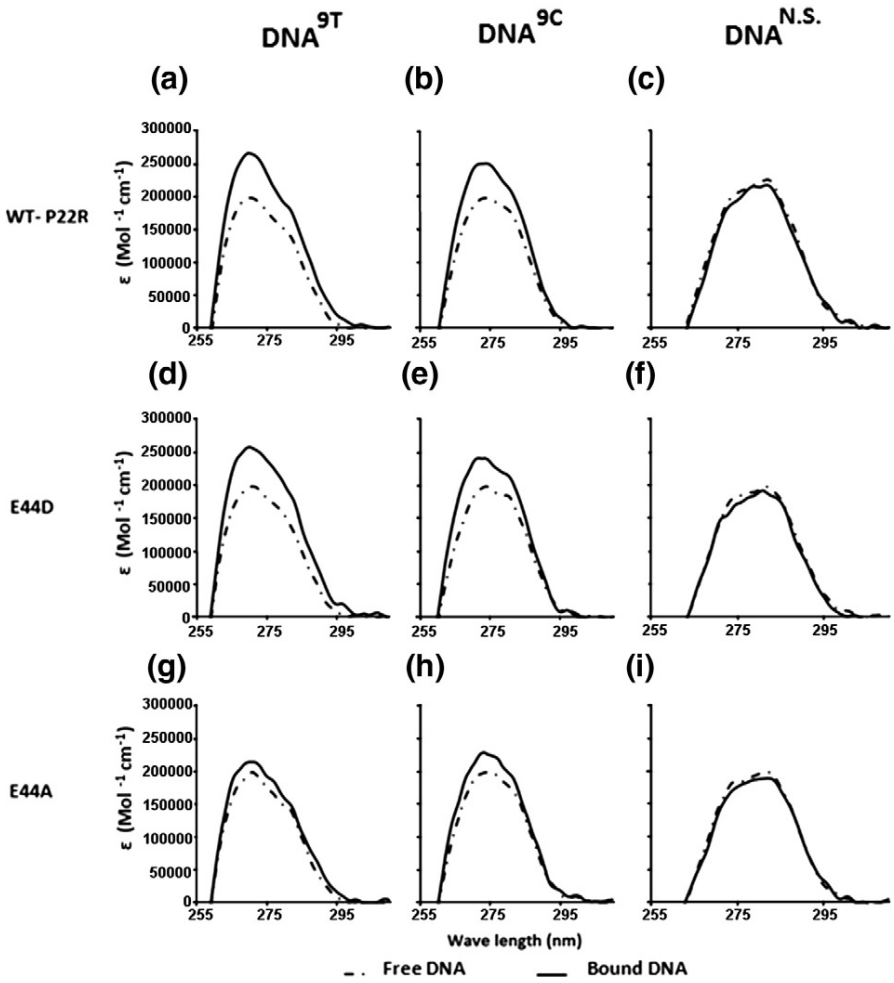

Fig. 3. CD spectra of unbound DNA and P22R–DNA complexes. Shown are the CD spectra of unbound 9T, 9C and nonspecific DNA in the absence or presence of wild-type (a–c), E44D (d–f) and E44A (g–i) P22 repressor proteins between 255 and 300 nm. Spectra were recorded and processed as described in Materials and Methods. In each panel, the spectrum of the unbound DNA spectrum is represented by a broken line and the bound DNA spectra are represented by a continuous line. Bound and unbound CD spectra of DNA^{9T} are shown in the first column, followed by spectra of DNA^{9C}, and the third column shows the spectra of nonspecific DNA in the presence and absence of P22 repressor proteins.

repressor, can discriminate between naturally occurring sites with a preference for O_R1 over O_R2. Taken together, these findings indicate that, irrespective of surrounding sequence context, the negative charge on E44 underlies the ability of P22R to discriminate between binding sites bearing different non-contacted base sequences. These results also confirm that non-contacted bases direct the ability of P22R to discriminate between its naturally occurring sites.

Discussion

The P22R protein recognizes specific DNA sequences in part by directly contacting highly conserved base pairs in the outer regions of the binding sites. P22R discriminates between similar naturally occurring and synthetic sites by indirect readout, that is, sensing the sequence of non-contacted bases at the center of the binding site.[13] Indirect readout relies on complex relationships between DNA sequence, conformation, deformation, solvation and electrostatics. *A priori*, DNA sequence could indirectly affect DNA recognition by any or all of three non-mutually exclusive strategies: (1) altering the flexibility of the DNA, modulating the ease with which the DNA can be distorted into the proper state for complex formation; (2) altering the average structure of the DNA helix, whereby an unbound DNA whose conformation resembles that found in the protein–DNA complex will bind with higher affinity than one

Table 3. Affinities of wild-type and mutant P22 repressors for the naturally occurring O_R1 and O_R2 P22 operators

Protein	O_R1	O_R2	K_D^{OR2}/K_D^{OR1}
Wild type	19.6±5.6	320±51.8	16.4
E44A	14.3±3.2	52.9±13.2	3.7
E44Q	1.4±0.1	7.0±2.0	5.2
E44D	24.9±1.9	>175	>7

Position	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
O_R1	A	T	T	A	A	A	G	A	A	C	A	C	T	T	A	A	A	T
O_R2	A	C	T	A	A	A	A	G	A	G	T	C	T	T	T	A	G	T

The sequences of the O_R1 and O_R2 operator sites are shown along one strand of DNA. Affinities of each protein for these operators are given as the dissociation constants (K_D, nM) plus/minus the standard deviation. The symbol (>) in front of K_D indicates that the P22 repressor variant binds to that operator only nonspecifically Also shown is the relative K_D (i.e., the ratio K_D^{OR2}/K_D^{OR1}) of the dissociation constants of each protein for O_R1 and O_R2.

whose unbound structure is dissimilar; or (3) changing the conformation of protein–DNA complex, causing alterations in the geometry and/or strength of direct protein–DNA contacts. The first two of these strategies infer that protein binding alters DNA conformation. In these cases, the differential affinities of various DNA binding sites for a cognate protein involve DNA deformation. Alternatively, the last strategy suggests that the affinities of different binding sites for a cognate protein do not depend on protein-induced DNA deformation.

Crystallographic data[12,15,33] and spectroscopic analysis (Fig. 3) show that, despite differing in their non-contacted base sequence and hence complex stability, the conformation of all P22R–DNA complexes are essentially identical. The similarities of DNA conformation, protein positioning and protein–DNA contact distances between these P22R–DNA complexes eliminate a central role for direct readout, suggesting that non-contacted bases influence protein–DNA complex stability by altering the global conformation of the P22R–DNA complex or geometry of one or more P22R–DNA contacts.

Without knowing the three-dimensional structures of the unbound forms of the various P22R binding sites studied here, it is difficult to determine what role, if any, DNA sequence-dependent differences in the structure of non-contacted bases in unbound DNA may play in determining the affinity of a binding site P22R. Regardless, our studies do indicate that the affinity of P22R for a particular binding site is determined, at least in part, by DNA sequence-dependent differences in the free energy required to change the conformation of the non-contacted bases at the center of the binding site from the B form to the B′ state. The results in Table 1 show that P22R has highest affinity for binding sites bearing an ApA step at its center, intermediate affinity for sites bearing a TpA step at these positions and lowest affinity for a site bearing a CpG step at its non-contacted positions. A-tracts, which contain contiguous $^{5'}$ApA$^{3'}$ steps, spontaneously assume the B′ form.[19–21] $^{5'}$TpA$^{3'}$ steps provide a barrier to formation of the B′ form, which can be overcome by environment. That is, $^{5'}$TpA$^{3'}$ can be induced to the B′ form by crystal lattice effects[22] and some minor groove binders,[23] as well as by protein binding.[12,15] C•G base pairs appear to provide the largest barrier to the transition from B form to B′ form. C•G base pairs are found in the B′ form only in complexes of certain protein–DNA complexes.[12,15]

Based on the known structural preferences of DNA sequences, the non-contacted regions of the various P22R binding sites studied here are not likely to be in the B′ state in the absence of P22R. P22R binding causes an increase in the long-wave UV CD amplitude of the binding sites. The observed increases in CD amplitude cannot be attributed to a change in protein conformation or the influence of protein spectra. This is because, in the 250- to 310-nm region, the contribution of protein absorbance to the CD spectrum is orders of magnitude smaller than that of DNA (Fig. 2). Moreover, the increase in peak amplitude was not observed in the spectra of nonspecific DNA in the presence of wild-type and mutant repressors. Thus, the changes in the bound DNA spectra are due to the specific imposition of DNA conformational changes by wild-type and mutant P22 repressors. Hence, the CD results show that P22R–DNA complex formation is accompanied by changes in the conformation of DNA. Thus, regardless of whether non-contacted bases indirectly modulate P22R's DNA affinity by altering the flexibility of DNA, and/or altering the structure of the unbound DNA, it is apparent that indirect readout by P22R relies, at least in part, on P22R-induced DNA distortion. However, the CD signature of B′-DNA is as yet unknown. Therefore, from the CD spectra alone, we cannot definitively conclude that the protein-induced changes in CD spectrum indicate the presence of B′ conformation. Nonetheless, the absence of B′ structure in unbound DNAs and the presence of this structure in P22R–DNA complexes argue that CD is monitoring the induction of the B′ state in the non-contacted bases.

The results here show convincingly that differential distortion propensities of various DNA sequences are important contributors to indirect readout. However, several mechanistic questions remain. Does P22R trap a minor population of the DNA that is in the altered state, or does it bind to the canonical population and induce the altered state. In the former scenario, the dissociation constant is driven by the equilibrium distribution of DNA between conformations bearing B′-DNA or B-DNA structure in the non-contacted region. If so, once bound to DNA, repressor maintains the B′ form of the non-contacted

DNA using the E44 and E48 guardrails and discriminates based on ease with which the imposed structure is maintained in the bound DNA. In the latter scenario, the affinity of P22R for a particular site depends on the ease with which it can impose a favorable binding structure in the DNA.

Several lines of evidence support the induction model. P22R has some affinity for canonical DNA (nonspecific binding). The results here suggest that electrostatic charge repulsion in the bound state could play a key role in induction of the altered state. First, the three-dimensional structures of the P22R–DNA complexes show that the negatively charged E44 and E48 residues are positioned provocatively near the phosphate atoms of the non-contacted nucleotides[12,15] (Fig. 1). P22R molecules bearing uncharged amino acids at either of these positions do not induce the altered state and discriminate less well between the central sequence variant DNA^{9T} and DNA^{9C} sites, as well as between the naturally occurring O_R1 and O_R2 sites, than does the wild-type or E44D mutant P22 repressors. Second, the change in CD intensity near 275 nm upon by DNA binding of wild-type and E44D mutant P22 repressors to DNA^{9T} or DNA^{9C} (Fig. 3) is substantially larger than that caused by DNA binding of the E44A (Fig. 3) and E48A (data not shown) mutant P22 repressors. The intensity of the transition in the long-wave UV portion of DNA's CD spectrum monitors changes in mutual orientation of neighboring bases,[34] especially changes in DNA twist[29] and propeller twist.[28] Hence, the CD results indicate that DNA binding by P22R containing negatively charged residues at positions 44 and 48 induce a much larger change in DNA conformation than do those containing neutral residues at these positions. Together, the results of our binding and spectroscopic studies show that indirect readout by P22R requires DNA distortion that is mediated by negatively charged residues at positions 44 and 48. We suggest that E44 and E48 residues function as "guardrails" that use like-charge repulsion to induce minor groove narrowing and the B′ state. In this way, E44 and E48 allow P22 repressor to discriminate between various central sequences by assessing the ease of forming and/or maintaining the B′ state.

The finding that the juxtaposition of negatively charged residues with the DNA phosphate backbone plays a large role in P22R's sequence recognition mechanism prompted us to ask two questions. First, how frequently negatively charged glutamates are found near the DNA backbone? Second, is the close approach of glutamates and DNA phosphates associated with alterations in minor groove conformation? A search of protein–DNA complexes in the Protein Data Bank showed that approximately 100 glutamate residues are found within 4 Å of a DNA phosphate.[35] These complexes include a number of specific DNA binding transcriptional regulators and several proteins that bind DNA in a non-sequence-specific but structure-specific fashion, including the nucleosome. In addition to solved protein–DNA complexes, models of complexes between several other DNA binding proteins that are structurally similar to P22R and model-built DNA also apparently have negatively charged residues closely juxtaposed with DNA phosphates. Hence, the juxtaposition of negatively charged residues with DNA phosphates is a relatively common feature of protein–DNA complexes.

Inspection of several solved protein–DNA complex structures revealed that, in a number of these complexes, the closely juxtaposed glutamates and DNA phosphates occur in regions of where the DNA minor groove is acutely narrowed. For example, the minor groove directly adjacent to the pair of glutamates flanking the central minor groove of the Mec I repressor–DNA complex (1SAX) is as narrow as that found in the P22R–DNA complexes.[36] A similarly narrowed minor groove adjacent to the pair of glutamates is also found in the related Bla I repressor–DNA complex (1XSD).[37] Interestingly, the sequence of the DNA in the narrow minor groove region in 1XSD is G/C rich. However, the structural data do not allow the solvent organization to be resolved and hence we cannot discern whether this DNA region is in B′ state.

Model building studies indicate that a very narrow groove in the non-contacted base region of P22R–DNA complexes is required to bring the two halves of the binding site into proper alignment with the P22R. That is, the overall three-dimensional shape of the binding site, which is largely determined by the assumption of the B′ state by the non-contacted bases, allows each operator half-site to make optimal contacts with each monomer of the bound P22R dimer. Similar to our findings,[3,14,15] it has been suggested that many other proteins may use DNA shape recognition to distinguish between subtly different DNA sequences.[1] As opposed to the repulsive interactions between negatively charged residues in the protein and DNA phosphates found with P22R and other proteins, Rohs *et al.* conclude that shape recognition by many proteins involves attractive electrostatic interactions between electronegative DNA minor grooves and positively charged arginine residue(s) positioned near the mouth of a narrowed minor groove.[6] Similar to what is seen with P22R, Rohs *et al.* propose that presence of A/T-rich sequences in the minor groove facilitates narrowing of the groove, which in turn enhances its electronegativity and thereby strengthens the interaction with arginine.[6]

Both our analyses[12] and those of others[38–40] demonstrate that sequence-dependent modulation of minor groove geometry DNA is the major contributor in DNA shape recognition. That is,

regardless of whether minor groove recognition is mediated by the juxtaposition of positively or negatively charged amino acids with DNA phosphates, sequence discrimination is ultimately governed by the propensity of a given DNA sequence to assume a particular minor groove geometry. This assertion is consistent with the well-known finding that minor groove width, as assessed by the accessibility of this groove to various nucleases, varies predictably with DNA sequence.[39,41–43]

Despite the common role of sequence in governing recognition of minor groove geometry in the two mechanisms of DNA shape recognition, there is a clear distinction between them regarding how DNA sequence influences protein–DNA affinity. The most straightforward illustration of these different roles can be made by comparing the effects of different non-contacted bases on proteins whose indirect readout mechanism involves repulsive interactions between protein residues and DNA phosphates (P22R) or attractive interactions between arginines and narrowed minor grooves (Hox family proteins).[44] In the case of P22R (and presumably analogous protein–DNA complexes), the conformation of DNA, protein positioning and protein–DNA contact distances in all these complexes are identical, regardless of the sequence of the "indirectly read" bases. However, in the case of ternary complexes between DNA, Hox and DNA-bound cofactor, variation of the sequence of the DNA in the indirectly recognized portion of the binding site significantly affects the precise positioning of the relevant arginine (and other) amino acids with respect to the minor groove.[44] Hence, in the P22R case, indirect readout apparently depends on DNA sequence-dependent differences in conformational polymorphism and deformability, whereas in the Hox case, indirect readout is driven by DNA sequence-dependent differences in one or more protein–DNA contacts. In the case of P22R, it is apparent that sequence-dependent differences in minor groove solvation play significant role in indirect readout by this protein. It remains to be determined whether differential minor groove solvation drives the sequence-dependent variation in Hox protein–DNA complex conformation.

Materials and Methods

Binding sites, purification and end-labeling

DNA oligonucleotides were obtained from Integrated DNA Technologies (Coralville, IA). The single-stranded DNA oligonucleotides were purified from denaturing polyacrylamide gel containing 7 M urea, 8% polyacrylamide and TBE [89 mM Tris (pH 7.5), 89 mM boric acid and 1 mM ethylenediaminetetraacetic acid (EDTA)] by crush and soak. After desalting on Sepharose G-10 columns (GE Healthcare Life Sciences, Piscataway, NJ) in TEN [10 mM Tris (pH 7.5), 50 mM NaCl and 1 mM EDTA], equimolar quantities of complementary single strands were annealed by heating to 80 °C for 2 min and slow-cooled to room temperature. Double-stranded DNAs were separated from single strands by gel electrophoresis on 8% polyacrylamide gels and purified as described above.

For filter binding experiments, 1 μg of DNA was 5′ end labeled with [γ-^{32}P]dATP (6000 Ci/mmol) (Perkin-Elmer, Boston, MA) in the presence of T4 polynucleotide kinase (Epicentre, Inc., Madison, WI). The resulting products were ethanol precipitated from a 0.5-M ammonium acetate solution and resuspended in TE [10 mM Tris (pH 7.5) and 1 mM EDTA]. The 9T-(3C) DNA fragment was prepared and purified as described in Ref. 18.

Preparation of P22R mutants

Plasmids directing overproduction of E44D, E44A, E44Q, E48A and E48Q were constructed by site-directed mutagenesis using the plasmid pTPΔ125[13,45] as a template. Primers used for mutagenesis were obtained from Integrated DNA Technologies. Wild-type and mutant P22 repressors were isolated from the X90 *Escherichia coli*[46] strain bearing the appropriate plasmids. All mutant proteins were purified as described by DeAnda *et al.*[45] The activity of all proteins was determined as described in Ref. 47, and dissociation constants were corrected for activity.

Filter binding assay

Filter binding was used to determine the affinities of wild-type and mutant repressors for the various DNA sites. These experiments were performed as described previously.[48] Briefly, we evaluated the binding of wild-type and mutant P22 repressors to DNA in a reaction buffer composed of 10 mM Tris, 150 mM NaCl, 0.1 mg/ml bovine serum albumin, 10 mM IPTG and 1 mM dithiothreitol. The concentration of DNA used in each experiment was $\leq$0.27 nM. After vacuum filtration through one layer each of nitrocellulose and diethylaminoethanol membrane, radioactivity corresponding to bound and unbound DNA was measured using a Storm imager (GE Healthcare Life Sciences) and quantified using Image Quant 5.0 software. Dissociation constants (K_D) were determined by nonlinear least-squares fitting of the filter binding data to a hyperbolic equation using Prism 3.0 software (GraphPad Software, Inc.). Each K_D was determined from five to eight replicate measurements.

CD spectroscopy

The conformational changes induced by wild-type repressors and E44A and E44D mutant P22 repressors in DNA fragments that are 26 base pairs long and DNA^{9T} and DNA^{9C} were monitored by CD spectroscopy. Spectra were recorded from 340 to 180 nm using a JASCO-715 at 22 °C in a 0.2-cm-pathlength cuvette. Data were collected every 0.1 nm, averaged over 10 scans and corrected for baseline. The spectra of either 6 μM DNA only (unbound DNA), 12 μM protein only or a mixture of 12 μM protein and 6 μM DNA dissolved in buffer 150 mM sodium phosphate,

pH 6.8, were monitored. Prior to CD measurements, protein and DNA were equilibrated, either separately or together, in buffer at 25 °C for 5 min. The protein-only spectra were subtracted from the protein–DNA complex spectra in order to determine the spectra of DNA within the complex (bound DNA). The data obtained for bound and unbound DNA were compared to determine the presence and extent of conformational changes induced in DNA by protein binding as indicated by differences in their observed spectra.

Acknowledgements

The work in this manuscript was supported by funds from the College of Arts and Sciences at the University at Buffalo.

Received 23 August 2012;
Received in revised form 9 October 2012;
Accepted 11 October 2012
Available online 17 October 2012

Keywords:
DNA binding;
DNA structure;
indirect readout;
bacteriophage

Present address: L. -A. Harris, Division of Nutritional Science, Washington University School of Medicine, Campus Box 8031, 660 South Euclid Avenue, St. Louis, MO 63110, USA.

Abbreviation used:
EDTA; ethylenediaminetetraacetic acid.

References

1. Rohs, R., Jin, X., West, S. M., Joshi, R., Honig, B. & Mann, R. S. (2010). Origins of specificity in protein–DNA recognition. *Annu. Rev. Biochem.* **79**, 233–269.
2. Rohs, R., West, S. M., Liu, P. & Honig, B. (2009). Nuance in the double-helix and its role in protein–DNA recognition. *Curr. Opin. Struct. Biol.* **19**, 171–177.
3. Koudelka, G. B., Harrison, S. C. & Ptashne, M. (1987). Effect of non-contacted bases on the affinity of 434 operator for 434 repressor and Cro. *Nature*, **326**, 886–888.
4. Olson, W. K., Gorin, A. A., Lu, X. J., Hock, L. M. & Zhurkin, V. B. (1998). DNA sequence-dependent deformability deduced from protein–DNA crystal complexes. *Proc. Natl Acad. Sci. USA*, **95**, 11163–11168.
5. McFail-Isom, L., Sines, C. C. & Williams, L. D. (1999). DNA structure: cations in charge? *Curr. Opin. Struct. Biol.* **9**, 298–304.
6. Rohs, R., West, S. M., Sosinsky, A., Liu, P., Mann, R. S. & Honig, B. (2009). The role of DNA shape in protein–DNA recognition. *Nature*, **461**, 1248–1253.
7. Koudelka, G. B. & Carlson, P. (1992). DNA twisting and the effects of non-contacted bases on affinity of 434 operator for 434 repressor. *Nature*, **355**, 89–91.
8. Zhang, Y., Xi, Z., Hegde, R. S., Shakked, Z. & Crothers, D. M. (2004). Predicting indirect readout effects in protein–DNA interactions. *Proc. Natl Acad. Sci. USA*, **101**, 8337–8341.
9. Wolberger, C., Dong, Y., Ptashne, M. & Harrison, S. C. (1988). Structure of a phage 434 Cro/DNA complex. *Nature*, **335**, 789–795.
10. Eliseo, T., Sanchez, I. E., Nadra, A. D., Dellarole, M., Paci, M., de Prat Gay, G. & Cicero, D. O. (2009). Indirect DNA readout on the protein side: coupling between histidine protonation, global structural cooperativity, dynamics, and DNA binding of the human papillomavirus type 16 E2C domain. *J. Mol. Biol.* **388**, 327–344.
11. Hobaika, Z., Zargarian, L., Boulard, Y., Maroun, R. G., Mauffret, O. & Fermandjian, S. (2009). Specificity of LTR DNA recognition by a peptide mimicking the HIV-1 integrase α4 helix. *Nucleic Acids Res.* **37**, 7691–7700.
12. Watkins, D., Mohan, S., Koudelka, G. B. & Williams, L. D. (2010). Sequence recognition of DNA by protein-induced conformational transitions. *J. Mol. Biol.* **396**, 1145–1164.
13. Hilchey, S. P., Wu, L. & Koudelka, G. B. (1997). Recognition of nonconserved bases in the P22 operator by P22 repressor requires specific interactions between repressor and conserved bases. *J. Biol. Chem.* **272**, 19898–19905.
14. Wu, L., Vertino, A. & Koudelka, G. B. (1992). Non-contacted bases affect the affinity of synthetic P22 operators for P22 repressor. *J. Biol. Chem.* **267**, 9134–9135.
15. Watkins, D., Hsiao, C., Woods, K. K., Koudelka, G. B. & Williams, L. D. (2008). P22 c2 repressor–operator complex: mechanisms of direct and indirect readout. *Biochemistry*, **47**, 2325–2338.
16. Hud, N. V. & Plavec, J. (2003). A unified model for the origin of DNA sequence-directed curvature. *Biopolymers*, **69**, 144–158.
17. Wu, L. & Koudelka, G. B. (1993). Sequence-dependent differences in DNA structure influence the affinity of P22 operator for P22 repressor. *J. Biol. Chem.* **268**, 18975–18981.
18. Hilchey, S. P. & Koudelka, G. B. (1997). DNA-based loss of specificity mutations. *J. Biol. Chem.* **272**, 1646–1653.
19. Burkhoff, A. M. & Tullius, T. D. (1987). The unusual conformation adopted by the adenine tracts in kinetoplast DNA. *Cell*, **48**, 935–943.
20. Burkhoff, A. M. & Tullius, T. D. (1988). Structural details of an adenine tract that does not cause DNA to bend. *Nature*, **331**, 455–457.
21. Price, M. A. & Tullius, T. D. (1993). How the structure of an adenine tract depends on sequence context: a new model for the structure of TnAn DNA sequences. *Biochemistry*, **32**, 127–136.
22. Mack, D. R., Chiu, T. K. & Dickerson, R. E. (2001). Intrinsic bending and deformability at the T-A step of CCTTTAAAGG: a comparative analysis of T-A and A-T steps within A-tracts. *J. Mol. Biol.* **312**, 1037–1049.
23. Tevis, D. S., Kumar, A., Stephens, C. E., Boykin, D. W. & Wilson, W. D. (2009). Large, sequence-

dependent effects on DNA conformation by minor groove binding compounds. *Nucleic Acids Res.* **37**, 5550–5558.
24. Watkins, D., Koudelka, G. B. & Williams, L. D. (2009). Minor groove solvation as a determinant of indirect readout.
25. Martin, S. R. & Schilstra, M. J. (2008). Circular dichroism and its application to the study of biomolecules. *Methods Cell Biol.* **84**, 263–293.
26. Cary, P. D. & Kneale, G. G. (2009). Circular dichroism for the analysis of protein–DNA interactions. *Methods Mol. Biol.* **543**, 613–624.
27. Bloomfield, V. A., Crothers, D. M. & Tinoco, I. (2000). *Nucleic Acids: Structures, Properties, and Functions.* University Science Press, Mill Valley, CA.
28. Johnson, B. B., Dahl, K. S., Tinoco, I., Ivanov, V. I. & Zhurkin, V. B. (1981). Correlations between DNA structural parameters and calculated circular dichroism spectra. *Biochemistry*, **20**, 73–78.
29. Baase, W. A. & Johnson, W. C. (1979). Circular dichroism and DNA secondary structure. *Nucleic Acids Res.* **6**, 797–814.
30. Whipple, F., Kuldell, N. H., Cheatham, L. A. & Hochschild, A. (1994). Specificity determinants for the interaction of lambda and P22 repressor dimers. *Genes Dev.* **8**, 1212–1223.
31. Bell, C. E., Frescura, P., Hochschild, A. & Lewis, M. (2000). Crystal structure of the lambda repressor C-terminal domain provides a model for cooperative operator binding. *Cell*, **101**, 801–811.
32. Bell, C. E. & Lewis, M. (2001). Crystal structure of the lambda repressor C-terminal domain octamer. *J. Mol. Biol.* **314**, 1127–1136.
33. Watkins, D., Harris, L., Koudelka, G. B. & Williams, L. D. (2011). Protein: feeling the groove of DNA. In *Frontiers in Nucleic Acids* (Sheardy, R. D. & Winkle, S. A., eds), pp. 147–166, American Chemical Society, Washington, DC.
34. Ivanov, V. I., Minchenkova, L. E., Schyolkina, A. K. & Poletayev, A. I. (1973). Different conformations of double-stranded nucleic acid in solution as revealed by circular dichroism. *Biopolymers*, **12**, 89–110.
35. Sarai, A., Kumar, S., Kitajima, K., Wang, Y. F. & An, J. (2007). BAInt: Base–Amino Acids Interactions Database, Vol. 2011.
36. García-Castellanos, R., Mallorquí-Fernández, G., Marrero, A., Potempa, J., Coll, M. & Gomis-Rüth, F. X. (2004). On the transcriptional regulation of methicillin resistance. *J. Biol. Chem.* **279**, 17888–17896.
37. Safo, M. K., Zhao, Q., Ko, T.-P., Musayev, F. N., Robinson, H., Scarsdale, N. *et al.* (2005). Crystal structures of the Blai repressor from *Staphylococcus aureus* and its complex with DNA: insights into transcriptional regulation of the *bla* and *mec* operons. *J. Bacteriol.* **187**, 1833–1844.
38. Joshi, R., Passner, J. M., Rohs, R., Jain, R., Sosinsky, A., Crickmore, M. A. *et al.* (2007). Functional specificity of a Hox protein mediated by the recognition of minor groove structure. *Cell*, **131**, 530–543.
39. Bishop, E. P., Rohs, R., Parker, S. C., West, S. M., Liu, P., Mann, R. S. *et al.* (2011). A map of minor groove shape and electrostatic potential from hydroxyl radical cleavage patterns of DNA. *ACS Chem. Biol.* **6**, 1314–1320.
40. Locasale, J. W., Napoli, A. A., Chen, S., Berman, H. M. & Lawson, C. L. (2009). Signatures of protein–DNA recognition in free DNA binding sites. *J. Mol. Biol.* **386**, 1054–1065.
41. Drew, H. R. & Travers, A. A. (1985). DNA bending and its relation to nucleosome positioning. *J. Mol. Biol.* **186**, 773–790.
42. Satchwell, S. C., Drew, H. R. & Travers, A. A. (1986). Sequence periodicities in chicken nucleosome core DNA. *J. Mol. Biol.* **191**, 659–675.
43. Greenbaum, J. A., Pang, B. & Tullius, T. D. (2007). Construction of a genome-scale structural map at single-nucleotide resolution. *Genome Res.* **17**, 947–953.
44. Slattery, M., Riley, T., Liu, P., Abe, N., Gomez-Alcala, P., Dror, I. *et al.* (2011). Cofactor binding evokes latent differences in DNA binding specificity between Hox proteins. *Cell*, **147**, 1270–1282.
45. DeAnda, J., Poteete, A. R. & Sauer, R. T. (1983). P22 c2 repressor. Domain structure and function. *J. Biol. Chem.* **258**, 10536–10542.
46. Coulandre, C. & Miller, J. H. (1977). Genetic studies of the *lac* repressor. III. Additional correlation of mutational sites with specific amino acid residues. *J. Mol. Biol.* **117**, 525–567.
47. Riggs, A. D., Suzuki, H. & Bourgeoss, S. (1970). *lac* repressor–operator interaction. I. Equilibrium studies. *J. Mol. Biol.* **48**, 67–83.
48. Mauro, S. A. & Koudelka, G. B. (2004). Monovalent cations regulate DNA sequence recognition by 434 repressor. *J. Mol. Biol.* **340**, 445–457.