

Recognition of DNA structure by 434 repressor

Gerald B. Koudelka*

Department of Biological Sciences, State University of New York at Buffalo, Cooke Hall, North Campus, Buffalo, NY 14260-1300, USA

Received July 3, 1997; Revised and Accepted November 9, 1997

ABSTRACT

In complexes of bacteriophage 434 binding sites with 434 repressor the central 4 bp of the 14 bp site are not contacted by the protein, although changes in these bases alter binding site affinity for the repressor. Our previous data suggested that the ability of the non-contacted central bases to be overtwisted in repressor–DNA complexes governs affinity of the binding site for 434 repressor. This idea was tested by examining the affinity of two central sequence variant 434 binding sites for 434 repressor as a function of binding site average twist. The 434 repressor preferred the relatively overwound binding site to the two more underwound forms. The greatest affinity enhancement resulting from increasing twist was observed with a binding site that is relatively underwound and more resistant to twisting deformation. Consistent with the idea that 434 repressor overtwists its binding site upon DNA binding, we show that 434 repressor is capable of binding to sites bearing a single base insertion in their center (a 15mer), but binds poorly to binding sites bearing central base deletions (12mer and 13mer). The N-terminal dimer interface plays a large role in determining 434 repressor central base preferences. Mutations in this interface eliminate central base discrimination and/or site size preferences. These mutations also lead to changes in the size of the repressor footprint on the various sized DNA sites that are consistent with their binding characteristics.

INTRODUCTION

Sequence-specific DNA recognition by proteins requires appropriate juxtaposition of amino acids and base pairs. This juxtaposition can be modulated not only by specialized protein architectures (1,2), but also by sequence-specific effects on DNA. It is therefore apparent that a complete understanding of how DNA binding proteins seek out and bind specifically and with high affinity only to their cognate binding sites will require knowledge of three areas: (i) the nature of sequence-specific effects on DNA structure; (ii) the role of protein architecture in determining possible protein–DNA contacts; (iii) the extent to which these elements interplay during formation of specific DNA–protein complexes. In this paper we utilize 434 repressor to begin to assess the role that each of these factors plays in DNA sequence-specific recognition.

434 repressor is required by bacteriophage 434 for establishment and maintenance of the lysogenic state (3). 434 repressor belongs to the class of proteins containing a helix–turn–helix structural motif (4,5). In complexes of 434 repressor with its binding site one dimer of the protein is bound to a partially two-fold rotationally symmetric sequence on B-form DNA. The two-fold related ‘recognition’ α -helices, one from each protein monomer, protrude from the surface of the protein and insert in successive major grooves on one face of the DNA (6). Each recognition helix is positioned in the major groove so that its sidechains make base-specific contacts with the *outermost* 4 or 5 bp of the 14 bp binding site (5,7).

The phage contains two operator regions, O_R and O_L , both of which consist of three closely spaced binding sites. The sequences of the outermost base pairs in the naturally occurring 434 operators are completely conserved in five of the six binding site sites and the remaining site differs only at one base. Despite the fact that the repressor makes specific contacts with the bases in the conserved region of the binding site, the affinities of these six sites for repressor vary over a 40-fold range. This differential affinity of the various operator binding sites is crucial for functioning of the genetic switch that directs the phage to develop lytically or lysogenically. Juxtapositioning of the conserved bases in the 434 binding site with the 434 repressor helix–turn–helix motif shows that the direct protein–DNA interactions that occur in this region of the complex *cannot* be responsible for governing the ability of 434 repressor to distinguish among the naturally occurring binding sites. We have shown that the non-conserved bases at the center of the binding site direct differential affinity of the operators for protein, despite the fact that these bases are not contacted by the protein (8).

Structural and biochemical studies reveal that the DNA in the center of the 434 repressor–DNA complex is overtwisted by $\sim 30^\circ$ and the minor groove in this region is narrowed relative to canonical B-DNA (7,9). These DNA deformations are required to align the halves of the binding site so that each monomer of the bound repressor dimer can make optimal contacts with one half-site (7). These deformations appear to be induced by repressor binding (9). Relief of the strain in the protein and DNA that results from repressor-induced alteration of DNA conformation by mutating the dimer interface of the repressor protein (9) or by nicking the binding site DNA at its center (10) eliminates ability of the repressor to distinguish between various central base sequences.

The twist of several uncomplexed synthetic 434 binding sites varies in a central base sequence-dependent manner that parallels their affinities for 434 repressor. Binding sites containing A+T-rich central base sequences that have a higher affinity for repressor are overtwisted relative to G+C-rich sites that have lower affinity (11). The twist of all these DNAs in repressor–DNA complexes is the

*To whom correspondence should be addressed. Tel: +1 716 645 3489; Fax: +1 716 645 2975; Email: koudelka@acsu.buffalo.edu

same. This observation indicates that in repressor–DNA complexes conformation of the DNA sugar–phosphate backbone is not affected by the central base sequence. Since the weaker 434 binding sites are undertwisted relative to the stronger sites and all repressor-ligated sites have identical twists, the repressor induces larger changes in the twist of the weaker binding sites than it does in the stronger binding sites (9). This observation suggests that the difference in affinity of the various binding sites must depend, at least in part, on the central sequence dependence of the equilibrium twist of the uncomplexed binding site. However, the variation in twist with central sequence does not completely account for variation in strength of 434 binding sites. In addition, our previous results show that different base compositions have different twisting flexibilities. Therefore, the base composition at the center of the 434 binding site influences its affinity for 434 repressor by affecting both the equilibrium twist of the unbound site and its twisting flexibility.

Our view of how the central operator sequence affects affinity of 434 operator for 434 repressor suggests that the energy needed to overwind this DNA determines the relative affinities of different binding sites for the protein. This idea was tested by examining relative affinities of the repressor for binding sites that are comparatively overwound or underwound prior to binding the protein. In order to further explore the role of the geometry and rigidity of the dimer interface of the 434 repressor in directing its central base preferences we introduced mutations into this interface that should alter the relative orientations of the two repressor monomers and examined their DNA sequence preferences.

MATERIALS AND METHODS

Binding sites, plasmids, DNA fragments and end-labeling

DNA manipulations were performed as described previously (12). The binding site oligonucleotides were purchased (Integrated DNA Technologies), annealed and ligated into the unique *SalI* site in pUC18 (13). The sequences of the synthetic 434 binding sites are given in Table 1. The positions of the individual bases are numbered consecutively along one strand (Table 1). For studies on the effect of central sequence and length variation, these plasmids were cleaved with *EcoRI* and 3'-end-labeled by repairing the recessed ends with Klenow fragment and [α - 32 P]dATP. The resulting ~2700 bp linear *EcoRI*-cut DNAs were used directly in filter binding experiments. For DNase I footprinting the binding site-containing plasmids were cleaved with *HindIII* and *HaeIII* and the recessed *HindIII* ends were repaired in the presence of [α - 32 P]dATP. This protocol generates a labeled 260 bp binding site-containing fragment and a 14 bp labeled fragment that does not interfere with cleavage analysis.

Table 1. Sequences of synthetic 434 binding sites

Operator	Sequence
Reference	A ₁ C ₂ A ₃ A ₄ T ₅ A ₆ T ₇ A ₈ T ₉ A ₁₀ T ₁₁ T ₁₁ G ₁₃ T ₁₄
7G	A ₁ C ₂ A ₃ A ₄ T ₅ A ₆ G ₇ C ₈ T ₉ A ₁₀ T ₁₁ T ₄₄ G ₁₃ T ₁₄
12	A ₁ C ₂ A ₃ A ₄ T ₅ A ₆ T ₇ A ₈ T ₉ T ₁₀ G ₁₁ T ₁₂
13	A ₁ C ₂ A ₃ A ₄ T ₅ A ₆ A ₇ T ₈ A ₉ T ₁₀ T ₁₁ G ₁₂ T ₁₃
14	A ₁ C ₂ A ₃ A ₄ T ₅ A ₆ T ₇ A ₈ T ₉ A ₁₀ T ₁₁ T ₁₁ G ₁₃ T ₁₄
15	A ₁ C ₂ A ₃ A ₄ T ₅ A ₆ T ₇ A ₈ A ₉ T ₁₀ A ₁₁ T ₁₂ T ₁₂ G ₁₄ T ₁₅

Shown are the sequences along one strand of each binding site. The upper two rows of the table shows the sequences of the synthetic 14mer sites used in these studies. The lower part of the table depicts the sequences of the binding site sized variants.

To examine the effect of topoisomer-induced twist differences on affinity of repressor dephosphorylated 258 bp linear DNA fragments containing either the reference or 7G synthetic 434 binding site and *BanI* cohesive ends were gel isolated from the plasmids constructed as described (14). These were 5'-end-labeled by incubating DNA with polynucleotide kinase and [γ - 32 P]ATP, followed by a chase with unlabeled ATP. Fragments were purified away from nucleotides and kinase by phenol extraction followed by ethanol precipitation. A portion of the reaction was reserved for use as the linear control in the filter binding study. The remainder of the labeled DNA fragments were incubated for 10 min in buffer containing 50 mM Tris–HCl, 10 mM dithiothreitol (DTT), 50 μ g/ml bovine serum albumin (BSA) and either 10 mM MgCl₂ at 25°C, to create the 24 turn topoisomer, or 100 mM MgCl₂ at 0°C, to create the 25 turn topoisomer. To this mixture sufficient DNA ligase was added to give >50% ligated circles in an additional 10 min incubation. The reactions were quenched by phenol extraction followed by ethanol precipitation. The remaining linear DNA in each of the two product mixtures was dephosphorylated by incubation with alkaline phosphatase, followed by phenol extraction and ethanol precipitation. The purities of the topoisomer preparations were assessed on polyacrylamide gels. As determined by phosphorimager analysis, the 24 turn preparation was not contaminated with either linear DNA or 25 turn topoisomer. The 25 turn preparations contained no detectable linear DNA, but ~15% of the total radioactivity in the sample was present as 24 turn topoisomer. These DNA preparations were used directly in filter binding experiments.

Proteins

Site-directed mutagenesis of the repressor gene was performed essentially as described (15). Wild-type and mutant 434 repressors were isolated from *Escherichia coli* strain X90 bearing a plasmid that causes overproduction of 434 repressor (16).

DNA binding assays

Nitrocellulose filter binding experiments were performed as described (17). The reaction buffer for these assays contained 10 mM Tris–HCl, pH 7.8, 50 mM KCl, 1 mM DTT, 100 μ g/ml BSA and <1 nM labeled DNA. The counts retained on the filter as a function of 434 repressor concentration were converted to fractional saturation values (18). Fractional saturation values were determined in duplicate. Three to five duplicate measurements were averaged and apparent dissociation constants (K_D) determined from non-linear least squares fits to the data. The standard deviations of the reported dissociation constants reflect the propagated errors of the data and the fits. These are <5% of the value.

DNase I protection assays were performed essentially as described (19). Briefly, binding reactions were performed in 10 mM Tris–HCl, pH 7.8, 1 mM MgCl₂, 1 mM DTT, 2.5 μ g/ml sonicated chick blood DNA, 100 μ g/ml BSA, \leq 1 nM binding site DNA, 50 mM KCl and an appropriate amount of 434 repressor or its variants. After a 10 min incubation at 23°C sufficient DNase I was added to give, on average, 1 cleavage/DNA molecule in 5 min further incubation. Following reaction with DNase I the ethanol-precipitated samples were suspended in 90% formamide dye mix and denatured by heating at 90°C. The samples were subjected to electrophoresis on 25 \times 30 cm denaturing 7.5% polyacrylamide gels containing 7 M urea, 89 mM Tris–HCl, 89 mM sodium borate and 1 mM EDTA. Reaction products were

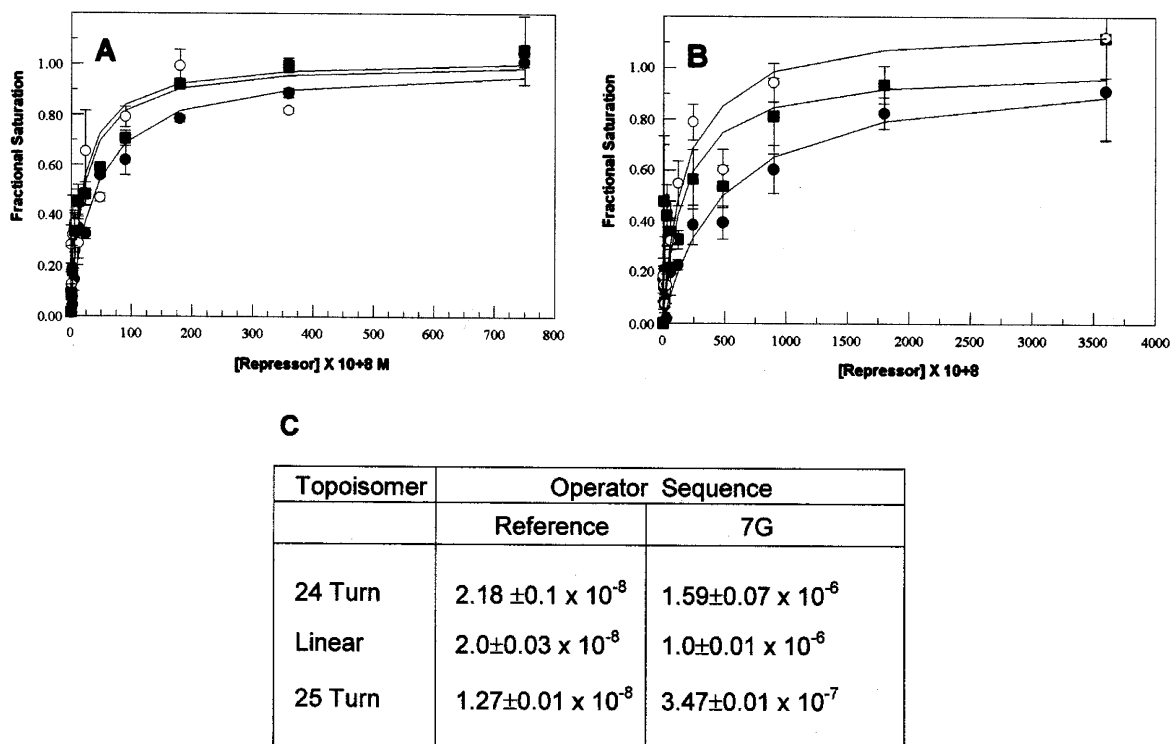


Figure 1. Effect of changing DNA structure on affinities of 434 repressor for synthetic 434 reference (A) and 7G (B) binding sites. The sequences of the binding sites used are given in Table 1 and were prepared as described in Materials and Methods. The values of apparent dissociation constants (K_D) of repressor binding site complexes, determined as described in Materials and Methods, are given in (C).

visualized by exposing the gel to a Molecular Dynamics phosphorimaging screen overnight.

RESULTS

Dependence of 434 binding site strength on DNA twist

The view of how the central sequence affects affinity of 434 binding site for 434 repressor implies that the energy needed to overwind this DNA determines the relative affinities of different binding sites for the protein. Hence, a binding site that is appropriately overwound prior to binding protein should have a higher affinity for that protein than binding sites that are relatively underwound. To test this idea the affinity of repressor for two different 434 binding sites (see Table 1 for sequences) in three different topoisomeric forms, linear and 258 bp circles with a L_k (linking number) of 24 or 25, was determined. Since circles of this size display no significant amount of writhe, L_k is directly related to twist (20,21). Thus the higher linking number topoisomer ($L_k = 25$) has an average higher twist than the lower linking number ($L_k = 24$) topoisomer. The average twist of the linear form lies between the 24 and 25 turn topoisomers. Determining the affinity of 434 repressor for 434 binding sites embedded in the three different topoisomeric forms allows examination of the effects of changing the average twist of DNA on binding site affinity for repressor.

The affinity of 434 repressor for each of the 434 binding sites varies with the topoisomeric form of the DNA (Fig. 1). With either binding site the 434 repressor has the highest affinity for the 25 turn topoisomer, followed by the linear DNA, and binds weakest to the 24 turn topoisomer. This data indicates that the

repressor can discriminate between binding sites on the basis of their average twist and that it prefers to bind to a relatively overwound DNA site.

Closer inspection of this data reveals that the greatest enhancement in affinity in response to increasing average twist is seen with the 7G binding site (Fig. 1). This result would be expected if the repressor can distinguish between these two DNAs based solely on their twist. We have shown elsewhere that the affinities of uncomplexed linear reference and 7G binding sites for 434 repressor increase in parallel with the resting twists of the uncomplexed linear forms (11). Since these sites have the same twist when complexed with repressor (9), the repressor must induce larger changes in twist of the weaker 7G binding site than it does in the stronger reference binding site. The free energy of twisting is proportional to the square of the change in twist (22). Hence, decreasing the amount of additional twist that must be induced by repressor binding would have a proportionately larger effect on binding sites that are undertwisted, relative to those binding sites whose twist more closely resembles that found in the repressor–DNA complex.

Similar to the effect of overtwisting, decreasing twist of the DNA by forming the 24 turn topoisomer has a greater effect on the relative affinity of the 7G binding site for repressor than for the reference operator. Since the free energy of twisting deformations also depends on the torsional spring constant (22), this observation suggests that in addition to being affected by the twist of the DNA, the affinity of repressor for DNA also depends on the flexibility of the DNA. Our measurements indicate that the 7G binding site is more resistant to twisting deformation than is the reference binding site.

Table 2. Relative affinities of 434 repressor and dimer interface mutants of various length synthetic 434 binding sites

Operator	Relative K_D					
	Wild-type	Ala47	Asp47	Ala46	Trp44	Trp44Asp47
12	>100	20	2	ND	2	0.3
13	23	6.7	2.5	6.3	2	1
14	1	1	1	1	1	1
15	7	22	6	40	2	4

Values are normalized vertically to the apparent K_D of each protein for the reference 14mer binding site. The values (unity) used for normalization with each protein are: wild-type 2×10^{-8} M; Ala47 1.6×10^{-7} M; Asp47 1.7×10^{-7} M; Ala46 8×10^{-8} M; Trp44 3×10^{-7} M; Trp44Asp47 8×10^{-6} M. ND, not determined.

If 434 repressor distinguishes between various central base sequences on the basis of different twists then the repressor should be sensitive to various fixed orientations of the two half-sites. The total amount of overtwisting induced by 434 repressor binding is $\sim 30^\circ$ (7). In mixed sequence DNA each base is rotated around the helix axis relative to its neighbor by an average of 34.5° . Hence, adding one base to the center of the 434 binding site, creating a 15 bp binding site, rotates the ends of the DNA to nearly the same degree as does repressor binding. On this basis it is reasonable to suggest that 434 repressor should be capable of binding specifically to a 15 bp binding site with good affinity. Similarly, deleting either one or two central base pairs creates sites that bear the contacted region of the binding sites in positions that are underwound by $\sim 34.5^\circ$ (13mer) and 69° (12mer) respectively. Since the repressor must severely overwind these DNAs in order to appropriately rotationally position the two half-sites, our hypothesis predicts that binding sites bearing these deletions should be poor substrates for repressor binding. Consistent with our assertions, Table 2 shows that wild-type 434 repressor binds to a 15 base binding site better than it does to 12 or 13 base long binding sites. The weaker binding of repressor for the 15mer as opposed to the 14mer site may reflect strain in the dimer interface when the protein attempts to bind the longer binding site (see below).

N-Terminal dimer interface and central sequence recognition

To examine the role of inter-subunit interactions in determining the specificity of 434 repressor for the base composition at the center of the binding site mutations were directed to this region of the protein. The relative affinities of the mutant proteins for 434 central sequence variant binding sites were determined. The dimer contacts in the repressor N-terminal domain are formed in part by a symmetrical pair of salt bridges between Arg41 in one subunit and Glu47 in the other (Fig. 2). The polypeptide lying between these salt links forms an extended hydrophobic interaction surface. Changing Glu47 to Ala should prevent formation of the inter-subunit salt link with Arg41. The data in Table 2 show that the Glu47 \rightarrow Ala mutation reduces the ability of 434 repressor to discriminate between 12, 13 and 14 base binding sites, but increases ability of the repressor to distinguish between sites larger than 14 bases long. This result suggests that the inter-subunit salt link has a role in determining the site size preference of 434 repressor. The overall effect of the Ala47 substitution is to alter the site size order of preference of the repressor from $14 > 15 > 13 > 12$ to $14 > 13 > 12 > 15$. These data

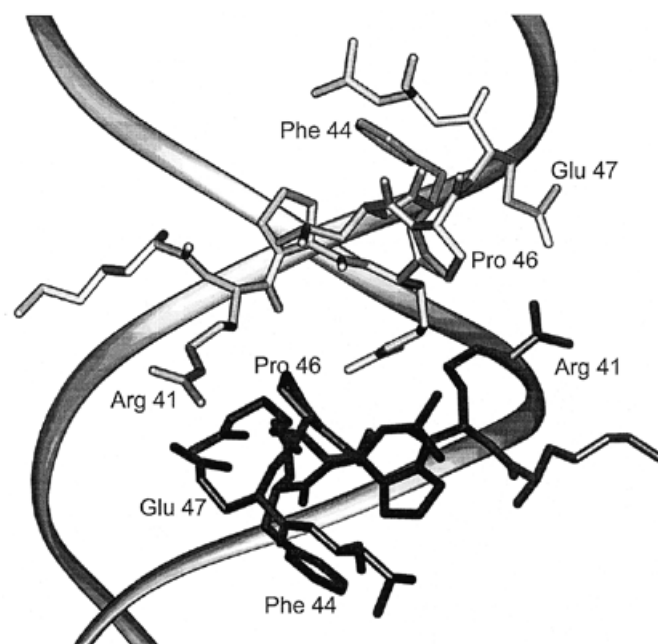


Figure 2. N-Terminal dimer interface of the 434 repressor. Shown are amino acids 41–47 for each of two subunits of repressor bound to O_g1. The view is into the minor groove at the center of the binding site. The DNA backbone is represented as the gray ribbon. The two subunits are distinguished by different shadings.

indicate that the salt bridges between Arg41 in one subunit and Glu47 in the other facilitate ability of the repressor to bind tightly to larger binding sites (see also below). However, the change in site size specificity conferred by this mutation is incomplete and is consistent with involvement of other amino acids in the dimer interface in determining this specificity.

Changing Glu47 to Asp conserves the potential for forming an inter-subunit salt link with Arg41 while shortening the sidechain by one $-\text{CH}_2-$ group. Like the Ala47 mutation, the Asp47 substitution alters the order of preference of repressor for binding sites of different size, again resulting in a repressor that prefers smaller binding sites. The site size preference of Asp47 repressor is $14 > 12 \approx 13 > 15$ (Table 2), with the protein essentially unable to distinguish between 12, 13 and 14 base long sites. We suggest that this change in site size preference results from the ability of the Asp47 to form a salt link with Arg41 in the 12, 13 and 14 base binding sites, while in the 15 base binding site, because of the short sidechain, such a link cannot be formed. Such a link would stabilize the smaller binding site–repressor complexes relative to the 15mer–repressor complex.

The relative orientation of the dimer interface segment comprised of residues 41–47 is fixed by the presence of a proline at position 46 (Fig. 2). Changing this residue should lead to an increased ability of the residues that form the dimer interface to re-orient themselves with respect to each other and alter the repressor's DNA recognition properties. The data in Table 2 shows that a Pro46 \rightarrow Ala mutation alters the repressor site size preference. The mutant protein binds tightest to the 14 bp binding site followed by the 13 bp binding site and binds the 15 bp site with lowest affinity. We were initially surprised by this result, since substituting an alanine for Pro46 could be expected to increase flexibility of the dimer interface, facilitating extension of this interface to allow reasonably tight binding to the

15 bp binding site. However, inspection of energy minimized molecular models of the wild-type and Ala46 mutant repressor (not shown) reveals that the Pro46→Ala substitution allows collapse of the polypeptide backbone in the region between amino acids 41 and 47 against the remainder of the protein structure. This ~1 Å shift in the backbone in this region in each DNA-bound monomer allows the two subunits to more closely approach each other. This observation is consistent with increased preference of the mutant repressor for the shorter 13 bp binding site over the longer 15 bp site.

Structural analysis indicates that packing of Phe44 into a cleft in the protein restricts conformational flexibility of the polypeptide segment between amino acids 41 and 47 (Fig. 2). Model building shows that, because of its larger sidechain, a Trp44 could not fit into this cleft. As a result, the amino acids that form the repressor dimer interface will become disengaged from the surface of the protein, thereby freeing the two monomers to re-orient themselves with respect to the DNA. Consistent with this idea, the data in Table 2 show that the Trp44 protein has essentially lost its ability to distinguish between the different length binding sites.

Since the Trp44 mutation appears to eliminate site size preference of the repressor, the idea that the Asp47 mutation affects site size selection by forming a shortened salt link with Arg41 was tested by combining the two mutations. Table 2 shows that the Trp44Asp47 doubly mutant protein has a completely altered site size preference order, as compared with the wild-type repressor. The doubly mutant protein prefers a 12 bp site over a 14 bp site by ~3-fold. This protein also binds weakest to the 15 bp long 434 binding site. Hence, the added flexibility conferred by the Trp44 substitution further accentuates the preference of a repressor bearing an Asp47 change for small binding sites (Table 2).

Table 3 shows that except for mutant proteins containing a Phe44→Trp substitution the N-terminal dimer interface mutants do not display a central sequence preference that is significantly different from that of the wild-type protein. The ability of Phe44→Trp repressor to discriminate between the reference and 7G site is reduced ~15-fold. The Trp44Asp47 double mutant is also ~15-fold reduced in ability to distinguish between the reference and 7G binding sites. These phenotypes are consistent with the hypothesis that these two proteins have a more flexible dimer interface and therefore can readily adjust to the structure of the DNA site, rather than imposing a particular structure on the DNA.

Table 3. Relative affinities of 434 repressor and dimer interface mutants for synthetic 434 binding sites

Operator	Relative K_D					
	Wild-type	Ala47	Asp47	Ala46	Trp44	Trp44Asp47
Reference	1	1	1	1	1	1
7G	60	70	55	ND	4	4

Values are normalized vertically to the apparent K_D of each protein for the reference binding site. The values (unity) for normalization in each protein are given in the legend to Table 2.

ND, not determined.

Structural basis for site size selection by the repressor

In an effort to understand how the structures of 434 repressor–DNA complexes are affected by changes both in binding site size and dimer interface mutation DNase I footprinting was used to characterize the complexes between wild-type repressor or dimer

interface mutants with the 13, 14 and 15 bp sites. Inspection of Figure 3 reveals that the ability of repressor to efficiently protect binding site DNA from cleavage varies with both DNA site size and dimer interface structure. For example, at saturating concentrations the footprint of wild-type repressor on the 13 base binding site extends only over the 13 bases that comprise the consensus region of the binding site. In contrast, the footprint of this protein on the 14 base site is 24 bases long, extending four and six bases downstream and upstream of the consensus sequence respectively. On the 15 base site wild-type repressor protects the entire binding site plus two bases upstream and six bases downstream of the consensus sequence. This 23 base footprint is similar to that seen on the 14 base site and indicates that repressor is able to form a more ‘wild-type-like’ complex with the 15 base site than it can on the 13 base site. The observed differences in the size of the repressor footprint on these three DNAs suggests that the wild-type repressor is more intimately associated with the DNA in the 14 and 15 base sites than it is in the 13 base site and provides a rationale for its extremely low affinity for the 13 base site.

Changing the amino acids that form the dimer interface affect ability of the repressor to protect the various size binding sites in a manner that is consistent with their effects on site size preference. Both the Trp44 repressor and the Trp44Asp47 doubly mutant repressor, which display markedly altered site size preferences, give the largest size footprint on the 13 base site (Fig. 3). Their ≥16 base long footprint compares with the 14 base long footprint of Asp47 and Ala47 proteins and the 15 bp long footprint of the Ala46 protein. The large footprint sizes of the Trp44 and Trp44Asp47 proteins suggests that they form a more intimate complex at the 13 bp site than do the other proteins.

All the dimer interface mutants form protein–DNA complexes that give smaller footprints on the 14 base site than does the wild-type repressor (Fig. 3). Interestingly, three of the four proteins that have an altered site size preference, Asp47, Trp44 and Trp44Asp47, have the smallest footprints on the 14 base site. The Asp47 repressor does not protect the base at +6, although the Ala47 and Pro46 proteins do at least partially protect this base. More dramatically, the Trp44 and Trp44Asp47 mutant repressors not only fail to protect the base at +6 but also do not completely protect the bases at –2 to –4. These bases are protected by all the other proteins examined.

Examination of Figure 3 shows that all mutant proteins form a very similar footprint on the 15 base long binding site and that this footprint is one base smaller than that made by the wild-type protein. Wild-type repressor protects the DNA site from two bases upstream to four bases downstream of the 434 binding site consensus sequence, but each mutant protein protects only up to the base at –2 in the upstream half of the consensus binding site.

DISCUSSION

The sequence of bases at the center of the 434 binding site have a significant role in determining its affinity for 434 repressor, despite X-ray crystallographic and biochemical evidence showing that these bases are not contacted by the protein. In complexes with 434 repressor these bases are overwound with respect to B-form DNA. The data presented here show that changing the relative rotational juxtaposition of the two DNA half-sites from a more underwound to a more overwound configuration increases the affinity of the binding site for the repressor. These observations confirm that the

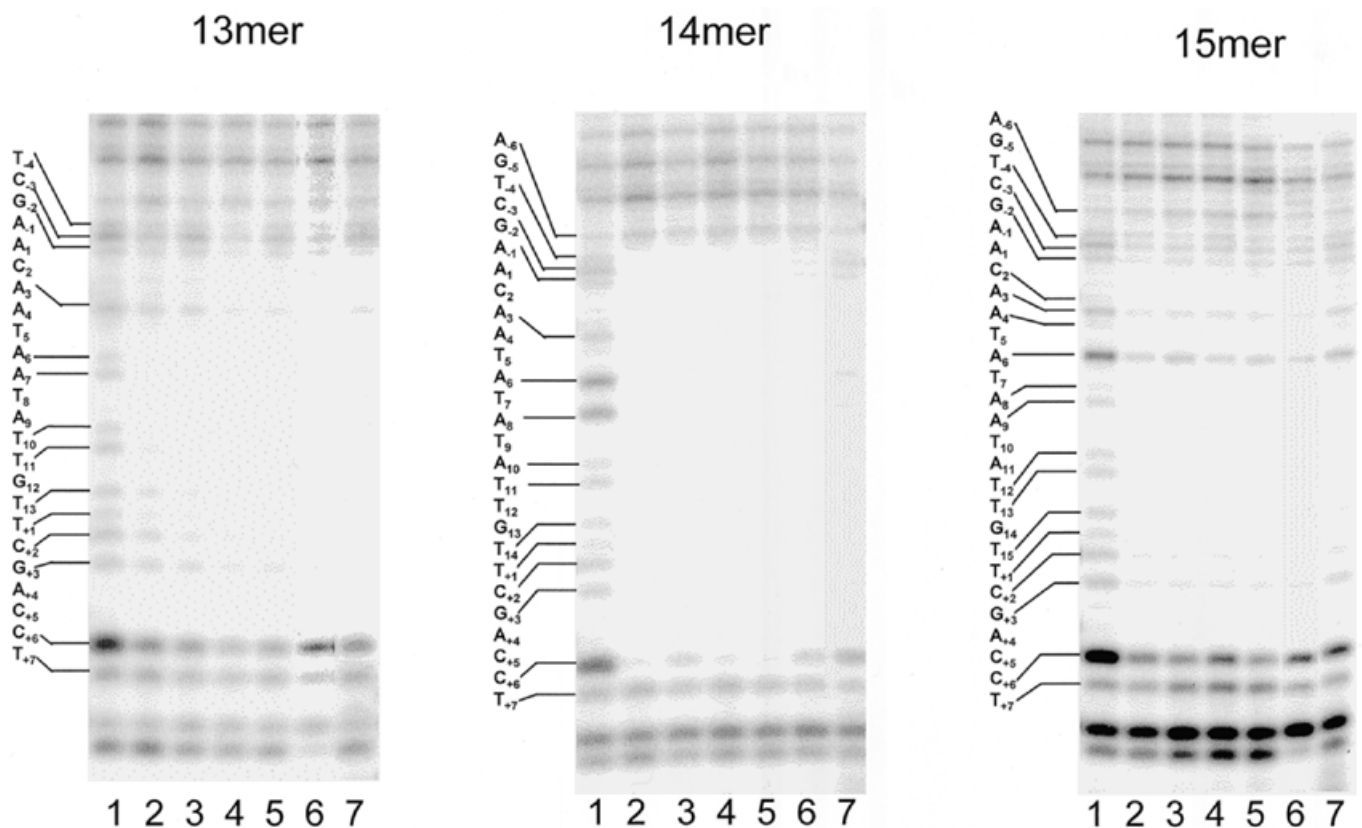


Figure 3. DNase I footprinting analysis of 434 repressor complexes with synthetic 434 binding sites of different size. DNase I experiments were performed as described in Materials and Methods. Shown is a composite of PhosphorImager scans of the gel. Each of the three panels display reactions with one binding site, as indicated. Lanes 1 in each panel are the DNase I cleavage pattern of the DNA in the absence of repressor. Lanes 2–7 are the DNase I cleavage patterns of the DNA in the presence of saturating concentrations of wild-type or dimer mutant repressor. Lanes 2, wild-type repressor; lanes 3, Ala47; lanes 4, Asp47; lanes 5, Ala46; lanes 6, Trp44; lanes 7, Trp44Asp47.

sequence of bases at the center of the 434 binding site affect affinity by influencing the ability of the binding site to be overtisted.

To confirm that the observed effects of altering the topoisomeric form of DNA on binding site affinity for repressor are related to topoisomer-induced twist changes we compared the observed difference in free energy of binding of repressor to 24 and 25 turn topoisomers with values of free energy differences calculated assuming the effect of changing topoisomer distribution affects only the resting twist of the DNA. The calculations were performed using the equation (22) $\Delta G = C \phi^2 / L$, which relates the free energy of twisting DNA to its flexibility and extent of twisting deformation, using known values for repressor-induced twist (ϕ) (7) and DNA torsional rigidity (C) (9). We chose to compare the difference in binding energy of repressor to the two circular topoisomers rather than between the linear and either of the two circular topoisomers. This is because the helical twist (H_{TW}) of the circular topoisomers can be precisely calculated from the relationship between the linking number (L_k) of a circle and the number of base pairs ($H_{TW} = \text{DNA length} / L_k$). Assuming that the twists of the binding sites embedded within the circular DNAs in complex with repressor are the same as that observed with linear DNA fragments, the calculations predict that the repressor should bind 1.3-fold better to the reference binding site in the 25 turn topoisomer than it does to this binding site in the 24 turn topoisomer. This prediction is in good agreement with the experimentally determined value of 1.7-fold. Similarly, the

calculations indicate that the repressor should prefer the 25 turn 7G binding site by 2.7-fold over the 24 turn variant. The data in Table 2 show that the repressor prefers the 25 turn site over the 24 turn site by 4.6-fold. Hence, the predicted values for the 24 versus 25 turn topoisomer accurately reflect the order of the experimentally determined effect of changing twist on the affinities of the two DNA sites for the repressor. This agreement suggests that the effect of changing the topoisomeric form of the binding site affects its affinity for repressor by influencing the twist of the unbound binding site.

With both binding sites the observed preference of the repressor for the more overwound 25 turn topoisomer is greater than that predicted. This disparity may be an artifact of a slightly underestimated torsional spring constant for both DNAs. These values were measured in the presence of 10 mM MgCl_2 (9), a condition that would be expected to slightly increase apparent DNA flexibility (23). The binding measurements were made in the absence of added MgCl_2 . Alternatively, the disparity may reflect a differential effect of circularization-induced distortion in the underwound versus overwound topoisomers. Others have proposed that linking number-dependent differences in the structure of much smaller DNAs may exist (24).

We have previously proposed that the structure of the dimer interface is rigid. We hypothesized that this rigidity fixes the relative orientation of the two repressor monomers. Thus, in order for repressor to bind specifically, the DNA must be deformed to put the

amino acid residues that specifically contact DNA in the proper juxtaposition with the conserved bases in the 434 binding site. Hence, the mutations in the dimer interface could affect the site size preference of the repressor by altering either the structure of the repressor dimer interface or the ease with which it can deform to assume the conformation needed by the repressor to permit specific complex formation. Since dimerization is a prerequisite for specific binding by the repressor, mutations that affect the structure of the dimer interface would be expected to dramatically decrease affinity of the repressor for the wild-type sized operator site. Mutations that affect dimer interface flexibility would be expected to have smaller effects on affinity of the repressor for 14 bp operators. The data in Table 2 show that the affinities of proteins bearing the F44W mutation for the 14 base site are markedly decreased, relative to wild-type repressor. The observed large decrease is consistent with the predicted structural change in the dimer interface resulting from the Trp44 substitution. Consistent with the suggestion that this change engenders a static structural change, the Trp44 and Trp44Asp47 mutant proteins only form intimate complexes with the 13mer binding site, as judged by DNase I footprinting. The affinities of the E47A, E47D and P46A dimer interface mutant proteins for the 14 bp operator differ from that of the wild-type protein by <10-fold (Table 2). These data suggest that shortening (E47D), eliminating (E47A) or changing the orientation of the dimer interface contacts (P46A) does not markedly affect the strength of the dimer interface, but only affects its flexibility. In contrast to the Trp44-containing repressors, proteins bearing the more flexible dimer interface (E47A, E47D and P46A) display DNase I footprints that expand or contract, depending upon site size.

A more extensive analysis of the footprinting data (Fig. 3) reveals that in addition to the specific differences between the footprints of the individual proteins on the different sized binding sites, an overall perspective on how the 434 repressor dimer interface determines site size preference can be deduced. This analysis shows that any mutation in the dimer interface has a two general effects: (i) a smaller footprint than wild-type protein; (ii) a relaxed discrimination against smaller binding sites. These observations suggest that the configuration of the dimer interface in the wild-type protein is constrained, thus facilitating recognition of the longer binding sites. Mutations that relax this constraint result in proteins that either have lost size selectivity or switch it completely. In other words, the default binding site for 434 repressor is shorter than 14 bp. The particular arrangement of amino acids present in the wild-type protein, however, directs repressor to bind to the larger binding sites.

Based on their reduced site size selectivity we were surprised that both the Asp47 and Ala47 mutant repressors retain a nearly unimpaired ability to distinguish between central sequence variant binding sites (Table 3). This observation suggests that the alterations in dimer interface flexibility imparted by these mutations is insufficient to overcome the remaining integrity of the unperturbed parts of this interface. Moreover, this observation also suggests that the reference and 7G DNA sites are more structurally malleable than are the dimer interfaces of proteins bearing either the Ala47 or Asp47 mutations. It is only by changing the size of the DNA binding site, a much more dramatic alteration in DNA structure, that the phenotypes of these mutant repressors can be detected.

For binding sites of identical sequence 434 repressor affinity varies nearly 5-fold, forcing a difference in twist of a total of ~36°

spread over 258 bp. The magnitude of the increase in repressor affinity resulting from a greater degree of pre-overtwisting of the 434 binding site may be much larger. This assertion is supported by the observation that the repressor binds reasonably well to a longer binding site in which the repressor binding surfaces on the DNA have been effectively overwound by ~34° with respect to each other, but are separated by an additional 3.4 Å. Nonetheless, even if this is the largest extent to which DNA supercoiling affects binding site affinity for repressor, the magnitude of the observed effect of changing DNA twist places it well within the range wherein it may significantly impact the biology of the phage. For example, repressor affinity for the three sites in O_R differs by at most 15-fold. Moreover, it is clear from the data presented here and elsewhere (25,26) that affinity of the repressor for naturally occurring binding sites may be differentially affected by alterations in DNA structure. A modification of DNA twist by alterations of *in vivo* levels of supercoiling may therefore profoundly affect the developmental decisions of the bacteriophage by altering the repressor order of binding to sites in the phage chromosome.

ACKNOWLEDGEMENTS

The author thanks Dr Margaret Hollingsworth for critical reading of the manuscript and acknowledges members of the laboratory for helpful discussions. This work was supported by PHS grant GM42138 from the National Institutes of Health.

REFERENCES

- Harrison,S.C. (1991) *Nature*, **353**, 715–719.
- Luisi,B.F., Xu,W.X., Otwinowski,Z., Freedman,L.P., Yamamoto,K.R. and Sigler,P.B. (1991) *Nature*, **352**, 497–505.
- Ptashne,M. (1986) *A Genetic Switch*. Blackwell Press, Palo Alto, CA.
- Sauer,R.T., Yocum,R.R., Doolittle,R.F., Lewis,M. and Pabo,C.O. (1982) *Nature*, **298**, 447–451.
- Harrison,S.C. and Aggarwal,A.K. (1990) *Annu. Rev. Biochem.*, **59**, 933–969.
- Anderson,J.E., Harrison,S.C. and Ptashne,M. (1987) *Nature*, **326**, 888–891.
- Aggarwal,A., Rodgers,D.W., Drottar,M., Ptashne,M. and Harrison,S.C. (1988) *Science*, **242**, 899–907.
- Koudelka,G.B., Harrison,S.C. and Ptashne,M. (1987) *Nature*, **326**, 886–888.
- Koudelka,G.B. and Carlson,P. (1992) *Nature*, **355**, 89–91.
- Koudelka,G.B., Harbury,P.H., Harrison,S.C. and Ptashne,M. (1988) *Proc. Natl. Acad. Sci. USA*, **85**, 4633–4637.
- Koudelka,G.B., Bell,A.C. and Hilchey,S.P. (1996) *Biol. Struct. Dyn.*, **1**, 135–153.
- Maniatis,T., Fritsch,E.F. and Sambrook,J. (1982) *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
- Messing,J. (1983) *Methods Enzymol.*, **101**, 20–79.
- Koudelka,G.B. (1991) *Nucleic Acids Res.*, **19**, 4115–4119.
- Kunkel,T.A., Roberts,J.D. and Zakour,R.A. (1987) *Methods Enzymol.*, **154**, 367.
- Wharton,R.P., Brown,E.L. and Ptashne,M. (1985) *Cell*, **38**, 361–369.
- Wu,L., Vertino,A. and Koudelka,G.B. (1992) *J. Biol. Chem.*, **267**, 9134–9135.
- Lin,S. and Riggs,A.D. (1972) *J. Mol. Biol.*, **72**, 671–690.
- Johnson,A.D., Meyer,B.J. and Ptashne,M. (1979) *Proc. Natl. Acad. Sci. USA*, **76**, 5061–5065.
- Horowitz,D.S. and Wang,J.C. (1992) *J. Mol. Biol.*, **173**, 75–91.
- Shore,D. and Baldwin,R.L. (1983) *J. Mol. Biol.*, **170**, 983–1007.
- Barkley,M.D. and Zimm,B.H. (1979) *J. Chem. Phys.*, **70**, 2991–3007.
- Hagerman,P.J. (1981) *Biopolymers*, **20**, 1503–1535.
- Heath,P.J., Clendenning,J.B., Fujimoto,B.S. and Schurr,J.M. (1996) *J. Mol. Biol.*, **260**, 718–730.
- Bell,A.C. and Koudelka,G.B. (1993) *J. Mol. Biol.*, **234**, 542–553.
- Bell,A.C. and Koudelka,G.B. (1995) *J. Biol. Chem.*, **270**, 1205–1212.