DNA-based Loss of Specificity Mutations

EFFECTS OF DNA SEQUENCE ON THE CONTACTED AND NON-CONTACTED BASE PREFERENCES OF BACTERIOPHAGE P22 REPRESSOR

(Received for publication, September 17, 1996, and in revised form, October 31, 1996)

Shannon P. Hilchey and Gerald B. Koudelka

From the Department of Biological Sciences, State University of New York, Buffalo, New York 14260-1300

Although the two central bases of the P22 operator are not contacted by the P22 repressor, changes in these bases alter the affinity of operator for repressor. Previous studies (Wu, L., and Koudelka, G. B. (1993) J. Biol. Chem. 268, 18975–18981) show that the structure of the P22 repressor-operator complex varies with central base sequence. Here we show that central base sequence composition affects the strength of two, and likely all, specific amino acid-base pair contacts between synthetic P22 operators and P22 repressor. However, altering a specific protein-DNA contact via a loss-of-contact mutation in repressor results in a loss of specificity at only one contacted position. Thus, only changing the sequence of non-contacted bases affects repressor’s global base specificity. The observed effects of ionic concentration on the affinities of various operators for repressor and the DNase I patterns of protein complexes indicate certain central base repressor and the DNase I patterns of protein complexes concentration on the affinities of various operators for repressor and the DNase I patterns of protein complexes evidence that central bases sequence. Here we show that central base sequence composition affects the strength of two, and likely all, specific amino acid-base pair contacts between synthetic P22 operators and P22 repressor. However, altering a specific protein-DNA contact via a loss-of-contact mutation in repressor results in a loss of specificity at only one contacted position. Thus, only changing the sequence of non-contacted bases affects repressor’s global base specificity. The observed effects of ionic concentration on the affinities of various operators for repressor and the DNase I patterns of protein complexes with these binding sites indicate certain central base sequences facilitate optimal juxtaposition of repressor with its contacted bases, while others prevent it. The existence of different structural forms of the repressor-operator complexes explains how the relative energetic importance of specific amino acid-base pair edge contacts is modulated.

The c2 gene of the lambdoid bacteriophage P22 codes for a DNA-binding protein called the P22 repressor. P22 repressor displays both structural and functional homology with the repressors of other lambdoid bacteriophages (1). The repressor proteins of all lambdoid phages control the developmental fate of the phage by acting as transcriptional regulators. This regulation is mediated by the binding of a repressor dimer to two operator regions on the phage chromosome (2). These two regions, designated OX and OY, are each further divided into three closely spaced binding sites, called operators. In all these phages, the repressor displays a hierarchy of affinities for these six operators. This ability of repressor to discriminate between binding sites is critical for the phage’s decision between lytic or lysogenic development. Binding of repressor to sites for which its has the highest affinity activates transcription of genes responsible for the maintenance of lysogeny and concurrently inhibits transcription of the genes responsible for lytic development. The developmental fate of the phage is thus critically dependent on the ability of repressor to discriminate between these different operators.

Sequence-specific DNA recognition requires precise matching between the molecular surfaces of protein and DNA. At the atomic level, direct reading of the DNA sequence occurs by interaction of amino acid side chains and/or peptide backbone atoms with the functional groups in DNA. These interactions require chemical complementarity between the interacting atoms. Ensuring that these chemically complementary interactions can occur requires structural complementarity along the interacting surfaces of the protein and DNA molecules. If the functional groups in the protein are not properly juxtaposed with those in the DNA, the stability of the protein DNA complex is compromised. Non-contacted bases within or adjacent to the binding site can therefore indirectly affect the affinity of DNA for protein by modulating the proper alignment between the functional groups of the protein and the DNA. Indirect effects could arise from sequence-dependent differences in the structure of the unbound and/or bound DNA, and/or the resistance of the DNA to distortion into the configuration appropriate for stable complex formation.

The sequences of the P22 phage operators are partially rotationally symmetric (Fig. 1). The repressor binding site is 18 base pairs in length (3). Inspection of these binding site sequences reveals that the outermost bases are highly conserved while the innermost bases are highly divergent (Fig. 1). Specific contacts made between the P22 repressor and the DNA are restricted to the outermost bases of the operator, while the central four bases are not contacted (4). Although central bases of the operators are not contacted, the binding affinity of P22 repressor for operator is modulated by the sequence composition of these bases (4, 5). These modulations in affinity are due, at least in part, to sequence-dependent structural differences among operators bearing different central base sequences (5).

Although the torsional flexibility of the DNA does not appear to play a significant role in determining the affinity of P22 operator for P22 repressor (5), other differences in DNA structure do appear to play a part in governing the strength of this complex. For example, P22 repressor prefers binding sites with a narrower minor groove at the center of the binding site. We have shown that central base sequence directs the geometry of this groove (5).

We have suggested that the central sequence-dependent modulation of DNA twist and geometry of the minor groove at the center of the binding site ultimately results in differences in the affinity of various operators for repressor due to an alteration in the number and/or geometry of protein-DNA contacts. In this paper we further examine the effect that central sequence-based structural differences have on the specificity of interaction between synthetic P22 operators and P22 repressor

*This work was supported by United States Public Health Service Grant GM42138 from the National Institutes of Health. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked “advertisement” in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

†To whom correspondence should be addressed: Dept. of Biological Sciences, State University of New York at Buffalo, Cooke Hall, North Campus, Buffalo, NY 14260-1300. Tel.: 716-645-3489; Fax: 716-645-2975; E-mail: CAMGBK@UBVMS.CC.BUFFALO.EDU.

1 S. Hilchey, L. Wu, and G. Koudelka, manuscript in preparation.
protein. Data presented here demonstrate that the strength of specific amino acid-base pair interactions in the P22 repressor-operator complex are influenced by central sequence context. Changes in the central sequence can make repressor blind to mutations in bases which are contacted by the protein. We also present data which gives insight into how these non-contacted base pairs are modulating these protein DNA contacts.

EXPERIMENTAL PROCEDURES

Binding Sites, Plasmids, DNA Fragments, and End-labeling—DNA manipulations were performed as described (9). The sequences of the binding sites are given in Fig. 1. The central sequence is the portion of the sequence labeled with a horizontal line. The number of base pairs are modulating these protein DNA contacts. In all cases, the protein concentration corresponding to the reaction product was visualized by exposing the gel to a Molecular Technololgy, annealed, and ligated into the unique sites in pUC18 (7). The sequences of the resulting plasmids were confirmed by dideoxy sequencing (8). These DNA molecules were cleaved at the EcoRI site and 3’-end-labeled by repairing the recessed ends with Klenow fragment and [α-32P]dATP. The resulting ~2700-base pair linear DNA were used directly in filter binding studies. For the DNase I footprinting experiments, the labeled DNA was cut with HindIII, and the resulting ~83-base pair operator-containing DNA fragment was isolated from agarose gels. Control experiments showed that within this size range the affinity of repressor for DNA and the structure of the repressor-DNA complexes are independent of DNA length.

Protein Preparation—P22 repressor was isolated from the Escherichia coli strain XA90 bearing either the pTP125 (9) or the pTP125-ALA-32 plasmids that cause the overproduction of wild-type or mutant P22 repressor bearing a 33 → Ala change, respectively. pTP125 was a gift from A. Poteete (University of Massachusetts Medical Center, Worcester, MA). pTP125-ALA-32 was constructed in two steps. First, one of the two EcoRI sites (upstream of the tac promoter in pTP125) was destroyed by partial digestion of the plasmid pTP125 with EcoRI followed by repair of the recessed ends with Klenow in the presence of deoxyribonucleotides and subsequent ligation. This created the plasmid pTP125. Second, the 220-base pair EcoRI-Csp 451 fragment of pTP125 which encodes the amino-terminal portion of P22 repressor was replaced with a DNA fragment bearing the mutation which changes A3n-32 to Ala. Both P22 repressors were purified as described by Della et al. (9).

DNA Binding Assays—Nitrocellulose filter binding experiments were performed as described (3). The reaction buffer for these assays contained 10 mM Tris HCl, pH 7.8, and 50 mM, 100 mM, or 150 mM KCl. The counts retained on the filter as a function of P22 repressor concentration were converted to fractional saturation values (10). For each operator and buffer condition, the range protein concentrations used to determine the Kp were adjusted so that the fractional saturation values between 10 and 90% were obtained over a 15-fold protein concentration range. In all cases, the protein concentration corresponding to the Kp was located in the middle of this range. Fractional saturation values were determined in duplicate. Three to five duplicate measurements were averaged and dissociation constants (Kp) were determined from non-linear least squares fits to the data. The maximum variation in the Kp determined from each duplicate measurement was less than 2-fold. The standard deviations of the reported dissociation constants were less than 5% of the value. DNase I protection assays were performed essentially as described (11). Briefly, binding reactions were performed in 10 mM Tris-HCl, pH 7.8, 1 mM MgCl2, 1 mM dithiothreitol, 2.5 µg/ml of bovine serum albumin, ~1 µM operator DNA, either 50 µM or 150 µM KCl, and an appropriate amount of P22 repressor. After a 10-min incubation at 23°C, sufficient DNase I was added to give, on average, one cleavage/DNA molecule in 5 min of further incubation. Following reaction with DNase I, the ethanol-precipitated samples were suspended in 100% formamide dye mixture and denatured by heating at 90°C. The samples were subjected to electrophoresis on 30 cm long denaturing 7.5% polyacrylamide gels containing 7 M urea, 89 mM Tris-HCl, 89 mM sodium borate, and 1 mM EDTA. Reaction products were visualized by exposing the gel to a Molecular Dynamics PhosphorImaging screen overnight. The screen was scanned on a Molecular Dynamics Model 425E PhosphorImager at 80 µm resolution. The resulting bands were then quantified using Molecular Dynamics ImageQuantNT™ software (version 4.1). Line graphs (Fig. 4) represent one-pixel-wide areas generated by drawing lines down the center of individual lanes (Fig. 3). Peak heights were normalized to peaks on either side of the DNase I protected area that do not change in intensity upon repressor binding. All intensity values reported in the text are pixel volumes calculated by the ImageQuantNT™ software using a local average background. The quantitative or qualitative appearance of the footprinting results were detected between replicate experiments. All footprinting results were obtained at saturating concentrations of binding protein. Results of experiments performed with a 3-fold higher protein concentration were not quantitatively different from the results reported here.

RESULTS

Effect of Central Sequence on Specificity of Interaction between P22 Repressor and Synthetic Operators—Our earlier results suggested that sequence changes at the center of the P22 operator may affect the strength of specific amino acid-base pair contacts made at the outer edges of the binding site (4). To test this idea, we have determined the relative affinity of P22 repressor protein for synthetic operators bearing mutations at contacted base positions 3 and 5 in the context of two different central sequence variants (9T and 9C, see Fig. 1 for sequences) at 100 mM KCl. We have already shown that the structures of the 9T and 9C synthetic P22 operators vary in the presence and absence of P22 repressor and that their affinities for P22 repressor are different (4, 5). In our experiments, all possible changes at position 3 were examined, while at position 5 two sequences, 5A and 5C, were studied. Except for varying the sequence at either positions 3 or 5 in the context of either the 9T and 9C central sequences (Fig. 1), the sequences of the binding sites are identical.

Although the three-dimensional structure of the P22 repressor-DNA complex has not yet been determined, two lines of evidence suggest that operator positions 3 and 5 are specifically contacted by repressor amino acids. First, the base sequence at position 3 is conserved in all but 1 of the 12 naturally occurring half-sites, while the base sequence at position 5 is completely conserved in all naturally occurring P22 operators (Fig. 1). Second, P22 repressors bearing Ala substitutions at amino acids 32 or 33 are unable to distinguish between base substitutions at operator positions 3 or 5, respectively (12).

In the context of the 9T central sequence, substituting the base at position 3 in one half-site results in a loss of affinity ranging from 16- to 28-fold (Table I, lines 1–4). This observation is consistent with the suggestion that the base at position 3 is specifically contacted by P22 repressor. Single base substitutions are expected to result in a loss of affinity of operator for repressor (12). In contrast to results obtained in the context of the 9T central sequence, mutating position 3 in the context of a 9C central sequence had virtually no effect on the binding affinity of repressor to P22 repressor (Table I, lines 5–8). Thus, in the presence of the 9C central sequence, P22 repressor protein is unable to discriminate between base substitutions at operator position 3.

We performed a similar analysis at position 5 in the context of both the 9T or 9C central sequence. As with the position 3 substitutions, the preference of P22 repressor for the consensus base at position 5 is only maintained within the 9T central sequence context (Table I, compare lines 1 and 9): repressor is unable to distinguish between consensus and non-consensus bases at position 5 in the context of a 9C central sequence (Table I, lines 5 and 10). Together, these data support our...
suggestion that central sequence changes affect the strength of specific P22 repressor-operator complexes by altering the strength of specific amino acid-base pair contacts. Despite this loss of specificity, however, P22 repressor binds 40-fold more tightly to any of the 9C operator variants than it does to non-specific DNA.

Effect of the Loss of Specific Contact on the Sensitivity of P22 Repressor to Central Sequence Substitutions—The data in Table I show that central sequence changes can clearly blind repressor to substitutions in the contacted region of the operator. We wished to determine whether changes in the contacted region can similarly blind repressor to central sequence substitutions. Changing the central sequence from 9T → 9C, when all the contacted positions are consensus, decreases the affinity of the operator for repressor by about 20-fold (Table I, lines 1 and 5). In contrast to this result, the affinity of repressor for operators bearing non-consensus mutations at position 3 is relatively insensitive to central sequence substitutions; changing 9T → 9C in the context of a base mutation at position 3 has at most a 2-fold effect.

The presence of mutations in other contacted positions also affects P22 repressor’s central sequence preferences. Changing the central sequence from 9T to a 9C in the context of a non-consensus 5 G:G base pair decreases the affinity of the operator for repressor by only about 2-fold, compared to the 20-fold effect of substituting central sequence in the context of the wild-type position 5 base (Table I, lines 1, 5, 9, and 10).

Effect of Changing Protein Contacts to DNA on Base Preferences of P22 Repressor—The above results demonstrate that the strength of the specific contacts made between the P22 repressor protein and its operator are modulated by the sequence at the center of the operator. It is also evident that alterations in the contacted bases in the binding site prevent repressor from recognizing the central base sequence. One question remains at issue; do any and all alterations of specific contacts result in global changes in base specificity at all positions, whether contacted or not contacted? The answer to this question will provide insight into the mechanism by which specificity in the protein-DNA complex is regulated.

To approach this question, we determined the effect of substituting an amino acid which specifically contacts only one base pair on the specificity of repressor for other contacted bases in the P22 operator. In particular, we present data obtained with the Asn-32 → Ala mutant repressor, although similar results were obtained when other amino acids that contact specific bases were changed (data not shown). Other studies have shown that changing Asn-32 → Ala eliminates...
P22 repressor’s ability to distinguish between base changes at operator position 3, exclusively. Here we examine the ability of this mutant protein and wild type P22 repressor to distinguish between base changes at operator position 5 in both 9T and 9C central sequence contexts. Comparing Tables I and II shows that in the context of a 9T central sequence, the affinity of both wild type and Ala-32 P22 repressor proteins for operator decreases as position 5 is changed (Table I, lines 1 and 10, with Table II, lines 1 and 2). Hence both proteins maintain specificity for the base at position 5 in the 9T context (we attribute the small decrease in position 5 specificity of the Ala-32 mutant protein to a slight mutation-induced alteration in protein structure). Therefore, our data indicate that protein-based loss of specificity mutations do not confer a global loss of specificity to the protein for the contacted region of the operator.

The presence of a 9C central sequence eliminates the ability of the mutant Ala-32 protein to discriminate between base substitutions at position 5 (Table II, lines 3 and 4). This is similar to the effect of 9C central sequence on the contacted base preferences of wild-type repressor. Therefore, changing an amino acid that specifically contacts a DNA base does not uncouple repressor’s specificity from the effects of central base composition.

The above results show that eliminating a specific protein-DNA contact by mutating the protein does not eliminate repressor’s other contacted base preferences. Elimination of contacted base specificity appears to be caused only by changing the non-contacted central sequence. Moreover, these data show that removing a specific contact does not render the repressor insensitive to the central sequence-directed loss of specificity. The latter observation implies to us that the mutant repressor would have normal ability to distinguish between operators bearing different central sequences. Surprisingly, the loss of a specific position 3 contact via an amino acid substitution also renders repressor insensitive to central sequence changes. Specifically, when 9T is substituted with 9C in the context of a consensus contacted region, the affinity of the operators for repressor bearing the Ala-32 change decreases by only 1.5-fold (Table II, lines 1 and 3). This is in marked contrast to the 20-fold decrease in affinity of operator for wild type repressor when 9T is changed to 9C (Table I, lines 1 and 5). Thus, any alteration of a specific contact, either by changing the protein or DNA sequence, desensitizes the protein to changes in non-contacted central bases, but the contacted base specificity of these proteins is still influenced by the identity of the central bases. These observations suggest that contacted base recognition by P22 repressor is intrinsically linked only to the identity of the non-contacted central base sequence.

Dependence of P22 Repressor-Operator Complex Stability on Salt Concentration—We wished to gain further insight into the mechanism by which the central bases of the P22 operator affect the strength of specific amino acid-base pair contacts. To do this, the effect of salt on the stability of several P22 repressor-operator complexes was determined. This method allowed separation of the overall binding affinities of the various operators for repressor into their ionic and non-ionic components.

We compared the effect of substituting the contacted T:A base pair at position 3 with an A:T pair in both the 9T and 9C central sequence contexts at ionic concentrations ranging from 50 to 150 mM KCl. In the operator bearing a 9T central sequence, changing position 3 from T:A to A:T results in a protein-DNA complex that is more sensitive to changes in ionic concentration. The affinity of the 3A-9T operator for P22 repressor decreases more than 4-fold between 50 and 150 mM KCl. This compares to the less than 1.7-fold effect of changing salt concentration on the stability of the 3T-9T operator-repressor complex (Fig. 2). The increased dependence of the 3A-9T mutant operator’s affinity for repressor on ionic concentration indicates that ionic interactions contribute more to the overall affinity of this repressor-operator complex than they do in the 3T-9T repressor-operator complex. This result suggests that the overall structures of these two complexes differ.

In contrast to the results obtained with the 9T-containing operators, changing the identity of the position 3 base from a T:A to A:T base pair in the context of the 9C central sequence has little, if any, additional effect on the sensitivity of the 9C operator-repressor complex to changes in ionic concentration. The affinity of the 3A-9C mutant operator decreases 14-fold between 50 and 150 mM KCl as compared to the 12-fold decrease in the affinity of 3T-9C for repressor in this range (Fig. 2).

Analysis of these data by the method of Record et al. (13) indicates 0.5 (±0.25) cations are displaced by each monomer of repressor binding to the 9T operator whereas 2.4 (±0.2) cations are displaced by each monomer when repressor binds to the 9C operator. When a similar analysis is performed on the 9T operators bearing consensus or non-consensus position 3 bases, the number of cations released increases from 0.5 to 1.5 (±0.5) as the consensus 3T is substituted with a 3A base. Interestingly, in the 9C central sequence context, the number of cations displaced when a 3A base is substituted for a 3T base does not significantly increase, 2.4 versus 2.6 (±0.2) cations. Hence, changing a contacted base in the context of a 9C central sequence does not appear to influence the relative importance of ionic interactions to the binding of the 9C central sequence containing operators.

DNase I Digestion Studies of P22 Operators and Repressor-Operator Complexes—The results of mutational and biochemical analyses of the various P22 repressor-DNA complexes suggest that the structures of these complexes vary with central sequence. This suggestion was tested by probing the structure of unbound and P22 repressor-bound operators by DNase I. The rate of DNase I cleavage of DNA is proportional to the
accessibility of the scissile bond and is reflected in the observed band intensity.

The conformation of the DNA phosphate backbone of four synthetic P22 operators was partially digested with DNase I in the presence or absence of protein at lower (50 mM) or higher (150 mM) KCl concentrations. Identical results were obtained at both salt concentrations (Fig. 3 and data not shown). As we observed previously, the DNase I cleavage patterns of the uncomplexed 9T and 9C operators differ from each other. Inspection of these data show that the structure of the unbound operator varies not only at the site of substitution, but also several bases away (Figs. 3 and 4A). Specifically, a comparison of the digestion patterns of these operators shows that when the central base pairs are changed from T:A-A:T to C:G-G:C, the rate of cleavage at position 9 of the operator increases slightly, while at position 11 decreases nearly 3-fold. Further away from the site of substitution, in the upstream portion of the operator, cleavage rates are also different. The intensities of the bands resulting from cleavage at positions 3, 4, 5, and 6 of the 9C operator increase between 20 and 200% as compared to cleavages at these positions in the 9T operator (Fig. 4A). Significantly, these affected bases lie in the contacted region of the operator. These data suggest that changing the central sequence in the P22 operator causes a global change in the structure of the unbound operator.

Upon binding repressor, the size and position of the DNA protected from DNase I is similar in both the 9T and 9C operators (Fig. 3). This similarity in protection pattern indicates that repressor binds specifically to both of these binding sites, despite the lack of base specificity displayed by repressor when binding to 9C-containing operators (Table I). Closer inspection of the DNase I data shows that the cleavages at positions 3 and 5 of the 9C operator are hypersensitive relative to the 9T operator, by 1.3- and 2.1-fold, respectively (Fig. 4B). Differences in the intensities are observed in the downstream portion of the operator as well. For example, position 12 is 2-fold more hypersensitive to DNase I cleavage in the 9C as compared with the 9T operator (Fig. 4B). As observed with the uncomplexed operators, these cleavage differences are in the region of the operator which is specifically contacted by the repressor, even though the only nucleotide changes are in the non-contacted central sequence. Thus, central sequence-dependent cleavage variability is observed both in the uncomplexed DNA and the repressor-operator complexes. The cleavage differences observed in the uncomplexed DNAs are exacerbated upon repressor binding such that the structures of the resulting 9T and 9C operator-repressor complexes are distinct from each other.

To further study the effect of central sequence on repressor’s interaction with the contacted region of the operator, we examined the effect of a position 3 half-site mutation on the DNase I cleavage pattern of both 9T and 9C operators. Examination of Figs. 3, 4C, and 4E clearly reveals that in the uncomplexed operators, substitution at position 3 in the context of either a 9T or 9C central sequence alters the cleavage patterns of these two DNAs. In both central sequence contexts, there is an increase in DNase I sensitivity around the site of substitution in the 3A-containing operators relative to the 3T-containing operators. These increases range from 2- to nearly 5-fold between operator positions 2 and 6. Interestingly, cleavage differences are observed several bases away from the site of substitution, some as far away as 13 bases. For example, the cleavage intensities observed at positions 7 and 14 of 3A-9T are both decreased by about 1.2-fold, relative to 3T-9T, and a 1.3-fold hypersensitivity is observed at positions 16 in the 3A-9T site. This is remarkable in that this position is 13 bases from the site of substitution (Fig. 4C). Highly reproducible long range differences are also observed when the two position 3-substituted 9C operators are compared, although these differences are more subtle than those in the 9T context. These data show that in both 9T and 9C central sequence contexts, base pair substitutions can have observable effects on the DNase I cleavage rates of bases distant from the actual site of substitution. This observation indicates that the conformation of the DNA phosphate backbone can be altered throughout the entire operator by a single base pair substitution.

The above results establish that base pair substitutions have similar effects on the DNase I patterns of uncomplexed 9C and 9T operators. These results do not fully explain the dissimilar effects of base substitution on repressor affinity for operator in the two central sequence contexts. To address this question, we examined the DNase I cleavage patterns of these DNAs in complex with repressor. Inspection of Figs. 3 and 4E shows that the DNase I cleavage pattern of the 3A-9C operator-repressor complex differs significantly from that of the 3T-9C operator-repressor complex. The observed differences are found within 6 bases of the site of substitution in the upstream contacted region. Most notably, the cleavage intensities of the bands at positions 3 and 5 in the 3A-9C complex are decreased by 1.5-fold and nearly 4-fold, respectively, relative to those of the 3T-9T operator-repressor complex. Conversely, cleavage at positions 2, 4, 6, and 9 increase by 1.5-, 2.5-, 1.5-, and 1.2-fold, respectively, in the 3A-9C complex as compared with the 3T-9C complex. Many of these changes, specifically those at positions 2 and 4, appear to mimic differences present in the unbound state (compare the unbound 3A-9C versus unbound 3T-9C in Fig. 4E). By comparison with the high degree of position 3-dependent structural variability observed in response to position 3 substitution in the 9C-containing operators, the cleavage patterns of the 3T-9T and 3A-9C operator-repressor complexes are very similar (Fig. 4D). Some differences between these two repressor operator complexes are evident in Fig. 3. However, with the exception of position 4, these differences are much smaller than those seen in the complexes of repressor with the two position 3 variant 9C operators. These observations sug-
suggest that the structural variation between the various 9C-containing repressor-operator complexes is greater than that of the complexes formed between repressor and the two 9T-containing binding sites.

**DISCUSSION**

Varying the sequence of non-contacted bases at the center of the P22 operator changes the affinity of this DNA for P22 repressor by altering the relative energetic contributions made by base sequence-specific (non-ionic) versus non-specific (ionic) interactions to the overall binding affinity. This central sequence-directed alteration in the relative importance of non-specific and specific contacts is visualized as a change in the ability of repressor to discriminate between base substitutions in the contacted region of the operator. Our data suggest that the changes in specificity are either due to differences in the static structure of the complexes formed between repressor and the central sequence variant operators or in the structural flexibility of these complexes. This assertion derives from two observations. First, the stability of a 9C operator-repressor complex decreases with an increase in salt concentration, while the strength of a 9T-containing repressor operator complex is relatively insensitive to changes in ionic concentration. This shows that ionic interactions have a larger role in stabilizing the 9C operator-repressor complex than they do in the 9T complex with repressor. Second, the affinity of 9C operator for repressor is virtually unaffected by changes in the contacted base sequence, while the affinity of the 9T operator is dramatically reduced. This indicates that in the context of a 9C central sequence, repressor forms weaker interactions with the contacted bases than it does in operators containing a 9T central sequence. Thus, sequences at the center of the P22 operator that decrease the affinity of operator for repressor increase the contributions of ionic interactions to overall binding energy at the expense of the specific protein-DNA interactions. The decrease in the importance of specific interactions appears to occur through an alteration in the geometry of the specific interactions, not through a loss in the number of specific contacts (see Table I).

**FIG. 4. Quantitation of DNase I footprinting analysis of 3N-9T and 3N-9C operators at 50 mM KCl.** Refer to "Experimental Procedures" for details of graph generation. Panels A, C, and E correspond to the (−) lanes (unbound) in Fig. 3, while the B, D, and F correspond to the (+) lanes (bound). The left portion of each line graph represents the top of the gel in Fig. 3 with the right representing the bottom. The numbered scale on the x-axis denotes the operator base positions (for operator sequences, refer to Fig. 1). The y-axis is expressed in arbitrary units of intensity.
Insight into a mechanism for central base effects on contacted base specificity can be derived from analysis of the DNase I digestion patterns of various P22 repressor-operator complexes. An examination of the two 9T complexes by DNase I (Figs. 3 and 4D) shows that their protection patterns are generally similar, despite differences in the identity of the base at position 3. Both of the 9T patterns are different from either of the two 9C-containing operator-repressor complexes. Moreover, the protection patterns of the two position 3 variants of the 9C complexes are distinct from each other. The cleavage differences between the 9C and 9T operator-repressor complexes represent a lessening in the degree of protection of 9C operator by repressor, relative to the 9T operator. Together, these observations suggest that the 9T central sequence facilitates a more intimate association between protein and DNA. This is reflected in the decreased structural variation in response to contacted base substitution in the 9T operator-repressor complexes when compared with the 9C-containing protein-DNA complexes. This interpretation is consistent with the results of the salt concentration dependence studies showing that the 9T complexes have stronger specific amino acid-base pair contacts than do the 9C operator-containing protein-DNA complexes (Fig. 2). The weaker contacts in the 9C operator-repressor complexes permit more structural freedom, allowing the repressor-operator complex to rearrange in response to loss of a specific contact, which is reflected in weaker protection from DNase I cleavage.

Our current picture of how DNA sequence influences specific and non-specific base preferences of P22 repressor is that these base changes affect the structure of the P22 repressor-operator complex. An alternative view suggests that the sequence of the operator alters repressor’s base preferences by influencing the structure of the unbound operator (Figs. 3 and 4, A, C, and E). In this case, loss of specificity could be induced solely by changes in the structure of unbound DNA provided these structural changes are maintained in the repressor-operator complex. This view of the DNA sequence-induced loss of specificity implies that the operator is a static partner in the protein-DNA interaction. However, two lines of evidence suggest that the structure of DNA in the P22 repressor-operator interaction is dynamic, changing with both sequence and the presence of bound protein. First, DNase I analysis of the various repressor-operator complexes indicates that the structure of the repressor-bound operator differs from that of the unbound operator. The assertion that the solution configuration of the DNA changes upon repressor binding coincides with the previously reported repressor-induced alterations in operator twist (5). Second, the observed loss of specificity cannot be due purely to changes in unbound DNA structure because changes in the protein’s primary sequence that should only affect contacted base specificity can also cause repressor to lose its ability to discriminate between base changes at the center of the operator (Table II). This result can only be explained by a model in which the DNA sequence-induced loss of specificity is attributed to sequence-dependent changes in the structure and/or dynamics of the P22 repressor-operator complex.

The observation that the non-contacted base sequence at the center of the P22 operator globally influences specific protein-DNA contacts by altering the structure and/or dynamics of the complex represents a unique mechanism by which DNA sequence influences protein-DNA recognition. Our previous studies of the 434 repressor-operator complex showed that non-contacted bases affect affinity of DNA for protein by altering the structure of the unbound DNA, the deformability of the DNA, and the structure of the bound DNA, thus influencing the strength of non-specific contacts between protein and DNA (14–18). Hence, we are interested in the mechanism by which certain central base sequences are able to blind P22 repressor to alterations in the sequences of specifically contacted bases. As mentioned above, the strength of the 9T operator-repressor complexes are less affected by changes in ionic concentration than are the 9C complexes. A more extensive analysis of the salt dependence data supports this view. By extrapolating the salt dependence of $K_D$ data to 1 M salt concentration, we can calculate that in the context of a consensus contacted region, the non-charged interactions appear to contribute about 3 ($\pm$ 0.2) kcal/mol less to the overall binding energy of the 9C operator-repressor complex than they do to the 9T operator-repressor complex. This is in agreement with previous data from our laboratory (4). This analysis again supports the model that the non-charged interactions in 9C complexes are relatively less important than in those formed between repressor and operators bearing a 9T central sequence. More revealing is the alteration in the contribution of the non-charged interactions to the overall binding energy when a specifically contacted position is lost, for example at position 3. In the context of the 9T central sequence, non-charged interactions contribute nearly 2 ($\pm$ 0.2) kcal/mol less to overall binding energy when the position 3 contact is lost. In contrast to this result, when the same contact is lost in the 9C central sequence context, only a modest 0.4 ($\pm$ 0.05) kcal/mol loss is observed. Clearly the non-charged interactions, which include specific amino acid-base pair contacts, contribute much less to the overall energetics of binding in the 9C context leading to the apparent loss of specificity. From this finding we suggest that central sequence modulates P22 repressor’s contacted base preferences by shifting the protein between two binding modes, a relatively “more specific” mode in which the protein is in more intimate contact with the bases and a “less specific” binding mode characterized by a closer association of the protein with the phosphate backbone rather than the bases.

How can we assemble the observations presented here into a structural mechanism which explains the effects of DNA sequence on sequence-specific binding of P22 repressor? The lack of a high resolution structure of the complex prohibits us from precisely describing what kinds of three-dimensional alterations in the structure or dynamics of the DNA give rise to the difference in protection patterns seen in the footprints. Despite this limitation, a consideration of the data presented here combined with our previously published results provides some insights into this question. We have previously shown that the intrinsic twists of the 9T and 9C operators differ, with the unbound 9C operator being overtwisted as compared with the unbound 9T operator (5). Upon binding, P22 repressor protein unwinds both operators to a similar extent, such that the complexed 9C operator remains overtwisted relative to the 9T operator by a similar extent as seen in the unbound operators. We suggest that as a result of these structural differences, an optimal alignment between functional groups on repressor amino acids and DNA bases does not take place in the 9C operator-repressor complexes. This is observed as both a loss in affinity and a loss of specificity. The inability of repressor to effectively engage the base pairs in the 9C operator leads to their de-emphasis in the overall energetics of binding, causing the phosphate contacts to become relatively more important. Thus, our data are consistent with the hypothesis that the specific interactions are “looser” in the 9C complexes due to the lack of a tight fit between the protein and the DNA.

The results presented here indicate that the strength of the interactions between P22 repressor and the contacted bases in the P22 operator are strongly influenced by the sequence of bases not directly in contact with the protein. One question
which remains at issue is why this mechanism for modulating DNA sequence-specific binding has evolved in this system. One possible answer to this question emerges upon consideration of the sequences of the naturally occurring P22 operators, their relative affinities for P22 repressor, and repressor’s function in the lysis-lysogeny decision. In the phage, repressor helps establish and maintain the lysogenic state by binding to OR1 and via cooperativity to OR2 at the identical concentration, despite OR2’s 30-fold lower intrinsic affinity. In a lysogen, binding to OR3 at still higher concentrations is required for proper regulation of repressor levels. Thus, the precise matching of the relative affinities of OR1, OR2, and OR3 is critical to the phage’s lifecycle decisions. We speculate that the non-contacted central sequence regulation of contacted base recognition serves to protect the phage from mutations in the contacted bases. This idea is supported by the observation that four of the six operators on the phage chromosome contain a C:G base pair at one of their two central positions (Fig. 1). It is these same four operators that repressor must always occupy in fulfilling its role in establishing and maintaining lysogeny.

The data reported here constitute the second, and more dramatic, discovery of non-contacted bases influencing the contacted base preferences of a DNA-binding protein. This phenomenon was first reported by our laboratory in our studies of the 434 repressor-operator interaction (19). We speculate that these 434 and P22 proteins may not be unique in having their base-specific contacts influenced by non-contacted bases. If non-contacted bases do commonly influence contacted base preferences, this would help to explain the poorly defined consensus sequences of several proteins which nonetheless bind DNA with high specificity (20). The effect of non-contacted bases on contacted base preference is especially problematic when the non-contacted bases exert their effects on non-adjacent bases as is true in the P22 repressor binding sites. These "long distance" effects mask the relevance of non-contacted bases in consensus protein binding sites. Thus, the implications of these consensus sequences should be interpreted with caution.

Acknowledgments—We thank Drs. Paul Gollnick and Margaret Hollingsworth and Ed Kiegle for critical reading of the manuscript. We also acknowledge members of our laboratory for helpful discussions.

REFERENCES
1. Sauer, R. T., Yocum, R. R., Doolittle, R. F., Lewis, M., and Pabo, C. O. (1982) *Nature* **298**, 447–451
2. Ptashne, M. (1986) *A Genetic Switch*, Blackwell Press, Palo Alto, CA
3. Poteete, A., Ptashne, M., Ballivet, M., and Eisen, H. (1990) *J. Mol. Biol.* **137**, 81–91
4. Wu, L., Vertino, A., and Koudelka, G. B. (1992) *J. Biol. Chem.* **267**, 9134–9139
5. Wu, L., and Koudelka, G. B. (1993) *J. Biol. Chem.* **268**, 18975–18981
6. Maniatis, T., Frisch, E. F., and Sambrook, J. (1982) *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY
7. Messing, J. (1983) *Methods Enzymol.* **101**, 20–79
8. Sanger, F., Donelson, J. E., Coulson, A. R., Kossel, H., and Fischer, D. (1973) *Proc. Natl. Acad. Sci. U.S.A.* **70**, 1209–1213
9. De Anda, J., Poteete, A. R., and Sauer, R. T. (1983) *J. Biol. Chem.* **258**, 10536–10542
10. Lin, S., and Riggs, A. D. (1972) *J. Mol. Biol.* **72**, 671–690
11. Johnson, A. D., Meyer, B. J., and Ptashne, M. (1979) *Proc. Natl. Acad. Sci. U.S.A.* **76**, 5061–5065
12. Ebright, R. H. (1996) *Methods Enzymol.* **208**, 620–640
13. Record, M. T., Jr., Lehman, T. M., and DeHaseth, P. L. (1977) *J. Mol. Biol.* **107**, 145–158
14. Koudelka, G. B., Harrison, S. C., and Ptashne, M. (1987) *Nature* **326**, 886–888
15. Koudelka, G. B., Harbury, P. H., Harrison, S. C., and Ptashne, M. (1988) *Proc. Natl. Acad. Sci. U.S.A.* **85**, 4633–4637
16. Koudelka, G. B., and Carlson, P. (1992) *Nature* **355**, 89–91
17. Bell, A. C., and Koudelka, G. B. (1993) *J. Mol. Biol.* **234**, 542–553
18. Bell, A. C., and Koudelka, G. B. (1993) *J. Mol. Biol.* **234**, 542–553
19. Bell, A. C., and Koudelka, G. B. (1995) *J. Biol. Chem.* **270**, 1205–1212
20. Segall, A. M., Goodman, S. D., and Nash, H. A. (1994) *EMBO J.* **13**, 4536–4548