Recognition of Nonconserved Bases in the P22 Operator by P22 Repressor Requires Specific Interactions between Repressor and Conserved Bases*

(Received for publication, March 19, 1997, and in revised form, June 12, 1997)

Shannon P. Hilechey, Lin Wu, and Gerald B. Koudelka‡
From the Department of Biological Sciences, State University of New York, Buffalo, New York 14260-1300

The ability of P22 repressor protein to distinguish between the six naturally occurring operator binding sites is critically important in determining whether the bacteriophage chooses to grow lytically or lysogenically. We have shown that changes in the highly conserved bases at P22 operator positions 3, 5, 6, and 7 prevent specific binding of P22 repressor. Moreover, studies of mutant proteins identified the three repressor amino acids that directly contact these conserved bases. The pattern of operator sequence conservation permits these direct amino acid-base pair interactions to occur in all except one of the 12 operator half-sites in the phage chromosome. Therefore, repressor differential affinity for these sites cannot be due to these highly conserved base-pair-amino acid interactions. Our binding studies show that the nonconserved bases at positions 2 and 4 also play an important role in determining the relative affinity of the naturally occurring P22 operators for P22 repressor. Our data indicate that the direct contacts between the three solvent-exposed amino acids and the conserved bases in the binding site lock these amino acids in place, forming a scaffold allowing the rest of the amino acids side chains to form weaker interactions with the nonconserved bases in the binding site.

The c2 gene of the lambdoid bacteriophage P22 codes for a DNA-binding protein called the P22 repressor (1). P22 repressor displays both structural and functional homology with the two other well-studied repressors of lambdoid bacteriophages, λ and 434 (2). The repressor proteins of all these lambdoid phages regulate the developmental fate of the phage by acting as transcriptional regulators (3). This regulation is mediated by the binding of repressor to two operator regions on the phage chromosome. These two regions, designated O1R and O1L, are each further divided into three closely spaced binding sites, called operators. In all three phages, the repressor displays a hierarchy of affinities for these six operators. Binding of repressor to its highest affinity binding sites activates transcription of the genes responsible for the maintenance of lysogeny and concurrently inhibits transcription of the genes responsible for lytic development. The phage's decision between lytic or lysogenic development is thus critically dependent on the ability of the repressor to discriminate between these different operators. The amino-terminal domains of the bacteriophage repressors contain the amino acids that are responsible for making all base-specific interactions with its DNA binding site (4–6). The sequences of the DNA binding domains of the repressor proteins from 434 and P22 bacteriophages are highly homologous, sharing 31% amino acid identity (2), and their three-dimensional structures are virtually identical (7, 8). Both proteins contain a helix-turn-helix structural motif and bind as homodimers to partially rotationally symmetric, naturally occurring operator sequences. Despite the high degree of both functional and structural homology between these two proteins, they bind to different recognition sequences (see Fig. 1).

To a first approximation, the basis for the different operator specificities of the P22 and 434 repressors is known. Wharton and Ptashe (9) replaced the amino acids along the solvent-exposed face of the second helix of the 434 repressor helix-turn-helix unit with the corresponding amino acids from the recognition helix of P22 repressor in the so-called "helix swap" experiment. The resulting hybrid 434 repressor protein binds specifically to the 18-base pair P22 operator instead of the 14-base pair 434 operator (9). This result shows that the amino acids in the recognition helix determine the DNA binding specificity of helix-turn-helix-containing proteins.

Although the results of the helix swap experiments suggest that the P22 and 434 repressors employ a nearly identical strategy for operator sequence recognition, several observations suggest that this is not the case. First, as mentioned above, the size of the consensus operator sequences of the two proteins differ (see Fig. 1). Second, inspection of the sequences of naturally occurring 434 and P22 operators reveals that the sequence conservation patterns of these binding sites are also different (see Fig. 1). Each 434 operator half-site contains four highly conserved positions at its outer edge, but the P22 operator half-site contains five highly conserved bases. Moreover, the completely conserved bases in each of the naturally occurring 434 operators are the first four contiguous outer base pairs in each half-site. The highly conserved bases in the P22 repressor binding sites are not contiguous and are located at positions 3, 5, 6, and 7 of the P22 operator. The observed differences between the recognition site sequences of the P22 and 434 repressor are difficult to reconcile with the implied similarity in their DNA recognition mechanisms (9), especially since the conserved bases are thought to be directly contacted by amino acids of repressors.

Although the affinities of both 434 and P22 repressors for their respective operators vary with the sequence of noncontacted bases at the center of the operator, these two proteins also employ different strategies in recognizing these bases. Whereas 434 repressor is apparently sensitive to DNA sequence-dependent differences in DNA flexibility (10–12), differences in the structure of the unbound P22 operator central sequence variants influence the affinity and specificity by modulating...
the strength of specific contacts between the P22 repressor and operator (13–15).

A major stumbling block for obtaining information regarding how P22 repressor discriminates between the naturally occurring operator sites is the absence of a three-dimensional structure of the P22 repressor-operator complex. We have obtained detailed insights into how the noncontacted base pairs at the center of the P22 operator contribute to operator specificity (13–15). However, the role of contacted bases located at the outer edges of the binding site in this process is unclear. Wüthrich and co-workers (7) report the three-dimensional structure of the amino-terminal domain of P22 repressor in the absence of DNA. The tertiary structure of the helix-turn-helix unit of the P22 repressor revealed in their studies is similar to that of bacteriophage repressors. Unfortunately, large differences between the P22 protein and the other phage repressors elsewhere in their structures, especially in the region of the protein likely for dimerization, prohibit us from building a reliable model of the P22 repressor-operator interactions. The purpose of the work reported here is to obtain data that will provide us with the insight needed to understand the mechanism of operator recognition and discrimination by P22 repressor.

EXPERIMENTAL PROCEDURES

Binding Sites, Plasmids, DNA Fragments, and End-labeling—DNA manipulations were performed as described previously (16). The binding site oligonucleotides were purchased (Integrated DNA Technologies), annealed, and ligated into the unique StuI site in pUC18 (17). The positions of the individual bases in the rotationally symmetric synthetic binding sites and the partially symmetric naturally occurring sites are numbered consecutively from 1 to 18 along one strand (see Fig. 1A). The center of symmetry of all the sequences lies between bases 9 and 10. Our studies used operators bearing doubly symmetric base substitutions at positions 2–7 of the operator. For simplicity in the text, base changes are referred to by the identity of the base at a particular position on the upper strand; however, all changes made are alterations of the base pair at the specified position in both half-sites. The sequences of the resulting plasmids were confirmed by dideoxy sequencing (18). The apparent affinity of the repressor for nonspecific DNA was assessed using unmodified pUC18 that did not contain a P22 operator site.

The DNA molecules were cleaved at the EcoRI site and 3′-end-labeled by recovering the recessed ends with Klenow fragment and [α-32P]dATP. The resulting 3000-base pair linear EcoRI cut DNAs were used directly in some filter binding studies as noted in the legends to the figures. In other experiments, this labeled DNA was subsequently cleaved with HindIII, and the resulting 93-base pair DNA fragments were gel electrophoresed by gel electrophoreses for experiments that use this fragment. Control experiments showed that the affinity of repressor for DNA is not significantly dependent on DNA length.

Protein Preparation—P22 repressor was isolated from the Escherichia coli strain XA90 bearing either the pTP125 (6), pTP1125-Ala45I, pTP1125-Ala32, pTP7-Gln45, pTP7-Ala32, or pTP-Ala37 plasmids, which causes the overproduction of wild-type or mutant P22 repressors bearing either an Asn-32 → Ala, Val-33 → Ala, Asn-32 → Gln, Gln-37 → Ala, or Gln-37 → Asn change, respectively. pTP125 was a gift from A. Poteete (University of Massachusetts Medical Center). pTP1125-Ala32 and -Ala45I were 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 deoxynucleotides and subsequently cleaved with HindIII and the resulting 83-base pair DNA fragment was isolated by gel electrophoresis (see legends to the figures). The sequences of the resulting plasmids were confirmed by dideoxy sequencing (18). The apparent affinity of the repressor for nonspecific DNA was assessed using unmodified pUC18 that did not contain a P22 operator site.

The DNA molecules were cleaved at the EcoRI site and 3′-end-labeled by recovering the recessed ends with Klenow fragment and [α-32P]dATP. The resulting 3000-base pair linear EcoRI cut DNAs were used directly in some filter binding studies as noted in the legends to the figures. In other experiments, this labeled DNA was subsequently cleaved with HindIII, and the resulting 93-base pair DNA fragments were gel electrophoresed by gel electrophoreses for experiments that use this fragment. Control experiments showed that the affinity of repressor for DNA is not significantly dependent on DNA length.

Protein Preparation—P22 repressor was isolated from the Escherichia coli strain XA90 bearing either the pTP125 (6), pTP1125-Ala32, pTP1125-Ala45I, pTP7-Gln45, pTP7-Ala32, or pTP-Ala37 plasmids, which causes the overproduction of wild-type or mutant P22 repressors bearing either an Asn-32 → Ala, Val-33 → Ala, Asn-32 → Gln, Gln-37 → Ala, or Gln-37 → Asn change, respectively. pTP125 was a gift from A. Poteete (University of Massachusetts Medical Center). pTP1125-Ala32 and -Ala45I were 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 deoxynucleotides and subsequently cleaved with HindIII and the resulting 83-base pair DNA fragment was isolated by gel electrophoresis (see legends to the figures). The sequences of the resulting plasmids were confirmed by dideoxy sequencing (18). The apparent affinity of the repressor for nonspecific DNA was assessed using unmodified pUC18 that did not contain a P22 operator site.

The DNA molecules were cleaved at the EcoRI site and 3′-end-labeled by recovering the recessed ends with Klenow fragment and [α-32P]dATP. The resulting 3000-base pair linear EcoRI cut DNAs were used directly in some filter binding studies as noted in the legends to the figures. In other experiments, this labeled DNA was subsequently cleaved with HindIII, and the resulting 93-base pair DNA fragments were gel electrophoresed by gel electrophoreses for experiments that use this fragment. Control experiments showed that the affinity of repressor for DNA is not significantly dependent on DNA length.

Protein Preparation—P22 repressor was isolated from the Escherichia coli strain XA90 bearing either the pTP125 (6), pTP1125-Ala32, pTP1125-Ala45I, pTP7-Gln45, pTP7-Ala32, or pTP-Ala37 plasmids, which causes the overproduction of wild-type or mutant P22 repressors bearing either an Asn-32 → Ala, Val-33 → Ala, Asn-32 → Gln, Gln-37 → Ala, or Gln-37 → Asn change, respectively. pTP125 was a gift from A. Poteete (University of Massachusetts Medical Center). pTP1125-Ala32 and -Ala45I were 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 deoxynucleotides and subsequently cleaved with HindIII and the resulting 83-base pair DNA fragment was isolated by gel electrophoresis (see legends to the figures). The sequences of the resulting plasmids were confirmed by dideoxy sequencing (18). The apparent affinity of the repressor for nonspecific DNA was assessed using unmodified pUC18 that did not contain a P22 operator site.

The DNA molecules were cleaved at the EcoRI site and 3′-end-labeled by recovering the recessed ends with Klenow fragment and [α-32P]dATP. The resulting 3000-base pair linear EcoRI cut DNAs were used directly in some filter binding studies as noted in the legends to the figures. In other experiments, this labeled DNA was subsequently cleaved with HindIII, and the resulting 93-base pair DNA fragments were gel electrophoresed by gel electrophoreses for experiments that use this fragment. Control experiments showed that the affinity of repressor for DNA is not significantly dependent on DNA length.

Protein Preparation—P22 repressor was isolated from the Escherichia coli strain XA90 bearing either the pTP125 (6), pTP1125-Ala32, pTP1125-Ala45I, pTP7-Gln45, pTP7-Ala32, or pTP-Ala37 plasmids, which causes the overproduction of wild-type or mutant P22 repressors bearing either an Asn-32 → Ala, Val-33 → Ala, Asn-32 → Gln, Gln-37 → Ala, or Gln-37 → Asn change, respectively. pTP125 was a gift from A. Poteete (University of Massachusetts Medical Center). pTP1125-Ala32 and -Ala45I were 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 deoxynucleotides and subsequently cleaved with HindIII and the resulting 83-base pair DNA fragment was isolated by gel electrophoresis (see legends to the figures). The sequences of the resulting plasmids were confirmed by dideoxy sequencing (18). The apparent affinity of the repressor for nonspecific DNA was assessed using unmodified pUC18 that did not contain a P22 operator site.

The DNA molecules were cleaved at the EcoRI site and 3′-end-labeled by recovering the recessed ends with Klenow fragment and [α-32P]dATP. The resulting 3000-base pair linear EcoRI cut DNAs were used directly in some filter binding studies as noted in the legends to the figures. In other experiments, this labeled DNA was subsequently cleaved with HindIII, and the resulting 93-base pair DNA fragments were gel electrophoresed by gel electrophoreses for experiments that use this fragment. Control experiments showed that the affinity of repressor for DNA is not significantly dependent on DNA length.
for P22 repressor. The apparent affinities are expressed as the relative dissociation constants of these operators $K_d$.

The protein concentrations used were corrected for activity (eq 1).

The * indicates that the repressor binds to the operator only nonspecifically. Nonspecific binding is observed at concentrations $> 3 \times 10^{-6}$ M.

superimposed over the initial complex was 2.34 Å for all atoms. To check the reliability of the final model, the minimization procedure was repeated several times. The final structures resulting from each round were highly similar and displayed a root mean square deviation of $\pm 0.5$ Å.

**RESULTS**

Effect of Operator Mutations on Wild-type P22 Repressor Binding—As a first step toward identifying the protein-DNA contacts made by P22 repressor, we mutated the bases at position 2–7 in the P22 operators and measured the apparent affinity of these mutant operators for repressor. We did not examine the effect of position 1 on repressor-operator interaction, since the results of other experiments indicate that operator position 1 does not influence repressor binding (data not shown). The role of the four noncontacted bases at the center of the P22 operator on determining operator affinity for repressor has been the subject of our other work (13–15).

As shown in Fig. 2, repressor only binds specifically to operators that contain the consensus bases at positions 3, 5, 6, and 7. Under the conditions of this experiment, the affinities of repressor for the operators bearing mutations at these positions are indistinguishable from its affinity for nonspecific DNA (see also below). These data are consistent with the suggestion that P22 repressor does not directly contact the bases at operator positions 1, 2, 3, and 4.

By contrast with the drastic effects of changing the conserved bases at positions 3, 5, 6, and 7 on operator affinity for repressor, substituting the nonconserved bases at operator positions 2 and 4 has much less dramatic effects on operator strength (Fig. 2). At 100 mM KCl, changing position 2 from the consensus T:A to either a C:G or G:C decreases the apparent affinity of repressor by only 4.3- and 3.2-fold, respectively. The affinities of operators bearing A:T, G:C, or C:G base pairs at position 4 for repressor are decreased 3.8-, 7.6-, and 5-fold, respectively, relative to the operator bearing the consensus T:A base pair at this position. These data indicate that wild-type P22 repressor does not directly contact the bases at operator positions 2 and 4, but they do suggest that the repressor does recognize the bases at these positions in some way. The observed magnitude of changes in operator affinity for repressor upon operator position 2 and/or 4 substitution is consistent with disruptions of weak contacts between the protein and the bases at these positions. Alternatively, changes in operator position 2 and/or 4 may modulate affinity of operator for repressor indirectly by altering DNA structure.

**Elucidation of Base Contacts by Loss of Contact Mutations**—We initially changed three amino acids along the solvent-exposed face of the recognition helix to alanine residues and examined the ability of these mutant proteins to distinguish between the various mutant operators. Alanine substitutions were used for two reasons. First, mutant proteins bearing alanine substitutions at DNA contacting positions would be expected to lose specificity at the operator position that is contacted by the wild-type protein. Second, based on its relatively strong propensity to form $alpha$-helices, we reasoned that substituting alanine would not dramatically affect the structure of the recognition helix.

**Base Preferences Determined by Asn-32**—Changing Asn-32 to Ala eliminates the ability of P22 repressor to discriminate between various base substitutions at operator position 3 (compare Figs. 2 and 3). Its affinity for operators bearing the different bases at position 3 varies by no more than 1.7-fold. This is in marked contrast to the large deleterious effect of position 3 substitutions on binding of wild-type repressor (Fig. 2). These results are consistent with the suggestion that Asn-32 makes a direct contact with the base(s) at operator position 3.

The data in Fig. 3 show that the Ala-32 mutant protein is largely unable to distinguish between operators bearing position 4 changes. Its affinity for these various operators spans only a 2-fold range as compared with the 4–7-fold effect that substituting position 4 bases has on operator affinity for the wild-type protein (Fig. 2). Similarly, as opposed to the 4–7-fold effect of position 2 substitution on operator affinity for wild-type protein (Fig. 2), the alanine substitution has less than a 2-fold effect on operator binding to the Ala-32 mutant repressor (Fig. 3). These data suggest that amino acid 32 is also at least partly responsible for P22 repressor base preferences at operator positions 2 and 4.

Although the Ala-32 protein displays decreased sensitivity to position 2, 3, and 4 operator substitutions relative to the
observed at concentrations.

The ability of wild-type P22 repressor to discriminate between base substitutions at operator positions 2 and 4 appears to direct the base preferences of the P22 repressor at operator positions 32. Since the P22 Gln-32 repressor has a new position 3 specificity (see above), it nearly retains wild-type specificity at operator position 5 (Fig. 5). Substituting position 5 in the operator sequence as well as the binding affinities are expressed as described in the legend to Fig. 2. The value used for normalization is 1 = 7.0 × 10⁻² M. Nonspecific binding is observed at concentrations >2 × 10⁻² M. K₈, K₉.

Asn-32 wild-type repressor (compare Figs. 2 and 3), the mutant protein still discriminates between position 5 changes as well as does wild-type repressor. This observation shows that the Asn-32 → Ala mutation does not eliminate base discrimination at all positions in the P22 operator. Thus, the amino acid at position 32 appears to direct the base preferences of the P22 repressor at operator positions 2, 3, and 4.

The ability of wild-type P22 repressor to discriminate between base substitutions at operator positions 2 and 4 appears to require the direct interaction between Asn-32 and operator position 3. To test this idea, we constructed a mutant P22 repressor that is able to make a direct, but non-native contact at position 3 and then examined the base specificity of this mutant protein. We reasoned that a protein bearing an Asn-32 → Gln substitution would display "434 repressor-like" base preference at position 3 (Ref. 20 and Fig. 4) and bind specifically only to operators bearing an A:T base pair at this operator position. As is evident in Fig. 5, the Gln-32 repressor protein displays the predicted new specificity at position 3, binding only to binding sites bearing an A:T base pair at position 3 as opposed to the wild-type protein requirement for an T:A base pair at this position.

Since the P22 Gln-32 repressor has a new position 3 preference, we are unable to assess the mutant protein position 2 and 4 base specificities in the background of the P22 consensus T:A base at position 3. We therefore constructed a series of operators bearing an A:T base pair at position 3 and various bases at positions 2 and 4. Fig. 5 shows that the Gln-32 mutant repressor is able to discriminate between base changes at operator position 4. Moreover, this protein displays the same order of preference for base changes at operator positions 2 and 4. The value used for normalization is 1 = 7.0 × 10⁻² M. K₈, K₉.

Base Preferences Determined by Val-33—In contrast to the large deleterious effects of substituting the bases at operator position 5 on the apparent operator affinity of wild-type repressor (Fig. 2), the P22 repressor bearing an Ala-33 substitution is capable of specifically binding all operator position 5 variants (Fig. 6). Substituting the 5 A:T consensus with either a G:C, T:A, or C:G decreases the apparent affinity of operator for the Gln-32 repressor by 4.7–39.5-fold.

The amino acid at position 32 appears to direct the base preferences of the P22 repressor at operator positions 2, 3, and 4. Moreover, this protein displays the same order of preference for base changes at operator positions 2 and 4. The value used for normalization is 1 = 7.0 × 10⁻² M. K₈, K₉. Nonspecific binding is observed at concentrations >2 × 10⁻² M. K₈, K₉.
substitution is a chemically conservative change, it is possible that despite its shorter side chain, the alanine may continue to function in position 5 discrimination. Although the Ala-33 protein displays a range of affinities for the position 5 substituted operators, it is clear that the Val-33 → Ala substitution results in a mutant protein that has not completely lost specificity; rather, this substitution causes a relaxed specificity at this position. Nonetheless, these data indicate that Val-33 contacts the base pair at operator position 5.

Changing Val-33 → Ala reduces but does not completely abolish the ability of P22 repressor to discriminate between base changes at position 4 of the operator (Fig. 6). We suggest that the residual position 4 base preferences displayed by the Ala-33 mutant protein are due to contacts made by Asn-32, as presented above (see also below). Hence, P22 repressor position 4 base preferences appear to be determined by two amino acids even though the base sequence at this operator position is not well conserved among the naturally occurring binding sites. Thus, whereas position 4 does not appear to be important in defining the overall consensus sequence of the P22 operator, it may be important in modulating the ability of the repressor to discriminate between the six P22 operators in the phage chromosome.

The Ala-33 mutant repressor protein is capable of binding specifically to operators bearing substitutions at position 6 (compare Figs. 2 and 6). The affinities vary from 17- and 18- to 24-fold as the position 6 A:T base pair is substituted with a T:A, G:C, or C:G, respectively (Fig. 6). This observation suggests that the Val-33 residue contacts the bases at operator positions 6 as well as 4 and 5; however, the loss of specificity at position 6 conferred by this alanine substitution is not as complete as that seen at operator position 5 (Fig. 6). The Ala-33 repressor does retain overall sequence specificity as evidenced by the observation that any position 3 substitution prevents specific operator binding of the Ala-33 repressor (Fig. 6 and data not shown). Taken together, the results of binding studies performed with wild-type and Ala-33 repressors indicate that Val-33 is largely responsible for P22 repressor position 5 base specificity and, to a lesser extent, its specificity at operator positions 4 and 6.

Base Preferences Determined by Gln-37—Changing Gln-37 to alanine resulted in a mutant repressor protein that no longer specifically bound to any one of our sets of synthetic P22 operators, although this protein does bind tightly to DNA derived from the dicAB operon (21–23).1 This observation suggested that position 37 plays a unique role in determining the P22 repressor operator sequence preferences. As an alternative strategy to probe the potential DNA contacts made by Gln-37, we changed Gln-37 → Asn. The Asn-37 protein is able to bind specifically to P22 operators; however, the poor filter retention of the Asn-37 repressor-operator complexes at 100 mM KCl prevented us from investigating the full range of the alteration in specificity of this protein under these conditions (data not shown). We were able to perform this study at 50 mM KCl. Thus, the results of the binding studies performed with the Asn-37 mutant and the results from a parallel set of experiments performed with the wild-type repressor under these conditions are shown in Fig. 7.

As compared with the results obtained at 100 mM KCl (Fig. 2), the apparent operator affinity of wild-type repressor at 50 mM KCl is less sensitive to changes at operator positions 3 and 5 (Fig. 7). Under these conditions, changing the bases at these positions does not eliminate specific binding (Fig. 7). Examination of Fig. 7 shows that the position 3 and 5 specificities of the Asn-37 mutant repressor are nearly identical to that of wild-type P22 repressor, consistent with our data showing that other amino acids in the repressor are involved in determining the base preferences of the protein at these positions. These data also show that the Gln-37 → Asn substitution does not cause P22 repressor to have a generalized loss of base specificity over the entire operator.

The wild-type repressor protein strongly prefers the consensus base at operator positions 6 and 7 at lower salt (Fig. 7). The affinity of operator for repressor decreases from 10- to >20-fold as either one of these positions is altered (Fig. 7). In contrast, changes at these positions have at most a 10-fold effect on the operator affinity of the Asn-37 mutant protein. Position 6 substitutions result in a decrease in the apparent affinity of mutant operator for Asn-37 ranging from only 3.3–10-fold as compared with the 10- to >20-fold decrease observed with the Gln-37 wild-type protein (Fig. 7). A loss of specificity by the Asn-37-substituted protein is also observed at position 7. When position 7 is changed to C:G, A:T, or T:A, the apparent affinity of operator for Asn-37 repressor decreases by 4.3, 5.6-, and 9.3-fold (Fig. 7). Not only is this range of affinities lower in magnitude as compared with those for the Gln-37 wild-type repressor, but the order of binding to the Asn-37 mutant protein also differs from the wild-type protein. The Gln-37 wild-type repressor prefers 7C > 7A = 7T (Fig. 7), whereas the Asn-37 mutant prefers 7T > 7A > 7C. Together, these results suggest that the glutamine residue at amino acid position 37

---

1 S. Hilchey, D. Mydlarz, J. Xu, and G. Koudelka, unpublished observations.
specifically contacts the bases both at position 6 and 7 in the P22 operator.

Model Building Studies of the P22 Repressor-Operator Complex—The contacts described (see Figs. 8B and 9 for a summary) do not agree with the pattern of protein-DNA contacts we had initially predicted (compare Fig. 5B with Fig. 9). We turned to molecular modeling to both help us understand the basis of this difference and to gain an insight into how the amino acids Asn-32, Val-33, and Gln-37 specify all the contacted bases in the P22 operator. As described under “Experimental Procedures,” our model building efforts utilized the 434 repressor-operator complex as a starting template (8). The root mean square difference between the starting and final structures is 2.24 Å for all atoms, indicating that most traces of the starting structure have been effaced during the minimization. Although we do not suggest that the P22 repressor-operator model structure is flawless in every detail, we assert that it does provide an accurate framework for understanding the molecular basis of P22 repressor operator specificity determinants.

Comparing the structures of the model-built complex with that of the 434 repressor-operator complex reveals that the overall positioning of the recognition helix with respect to the major groove of the DNA is different in the two complexes (Fig. 8A). In general, the C-terminal end of the recognition helix of P22 repressor is shifted closer to the bases at positions 6 and 7 than this helix is in the 434 complex (Fig. 8A). The difference in the angle of helix insertion into the major groove between the 434 and P22 repressor-DNA complexes causes a marked difference in the relative juxtapositioning of their respective amino acids and base pairs. In addition, the model of the P22 repressor-DNA complex suggests that the operator DNA is distorted from the B-form helix between bases 3 and 5 (Fig. 8A). This distortion is an overall increase in the propeller twist of the thymine bases at positions 3, 4, and 5, forming a van der Waals pocket into which the Val-33 residue is inserted (Fig. 8B). Together, the DNA distortions and positioning of the recognition helix in the major groove are critical in accounting for the results of our biochemical studies.

The model-built structure of the P22 repressor-operator complex shows that the first amino acid of the recognition helix, in this case δ-NH$_2$ of Asn-32, is positioned to make a hydrogen-bonded interaction with the O-4 of the thymine base at P22 operator position 3. Consistent with biochemical data indicating that Asn-32 is the sole determinant of P22 repressor position 3 specificity, the model shows that no other amino acid is in close proximity to the bases at this position (Fig. 8B).

As opposed to the direct contact between Val-33 and the operator position 4 that we initially predicted (Fig. 4), both the model building and binding studies show that this amino acid directly interacts with the bases at operator position 5. The close approach of Val-33 and the base pair at position 5 suggests a van der Waals contact between the C-5 CH$_3$ group of the thymine on the bottom strand at this position and the γ-CH$_3$ of Val-33. Our current model does not explain the residual discrimination of the Ala-33 mutant repressor against operators bearing G:C base pairs at position 5 (Fig. 6, compare position 5 T:A and C:G versus G:C substitutions). It appears unlikely that this residual discrimination is due to contacts from other amino acids in the recognition helix, since no other amino acids appear to closely approach these bases.

The model building suggests that the P22 repressor preference for A:T base pairs at position 4 is determined solely by hydrophobic interaction with the C-5 CH$_3$ of the thymine base at this position. No other amino acids appear to closely approach any other functional groups of the bases on either strand at this operator position. The β- and δ-CH$_3$ of both Val-33 and Asn-32 apparently contribute to position 4 specificity by forming a hydrophobic pocket for the C-5 CH$_3$ group of the thymine at this position. From careful inspection of our model, the positioning of the side chains of Val-33 and Asn-32 could not contribute to recognition of position 4 without the primary contacts made by these amino acids to positions 5 and 3, respectively. A mechanism for recognition of the base in the homologous position in the 434 operator by 434 repressor has already been proposed (8).

Biochemical studies show that changing Gln-37 → Ala renders P22 repressor unable to bind any wild-type or single-
position mutant operator (data not shown). This observation implies that Gln-37 contributes a larger fraction of the binding energy of the complex than does either Asn-32 or Val-33. Consistent with this, the model suggests that Gln-37 donates a hydrogen bond to the O-4 of thymine of the A:T base at position 6 with the $\gamma$-NH$_2$, and the $\gamma$-O accepts a hydrogen bond from the N-4 of the cytosine on the bottom strand at position 7 (Fig. 8B). The formation of this bridging configuration placed the torsion angles of Gln-37 in the fully extended configuration, indicating that these hydrogen-bonded contacts may mutually support each other in this unfavorable glutamine conformation. Another possibility is that either the recognition helix is moved even closer to bases 6 and 7 in this region than our modeling has demonstrated, and/or there exists a distortion in the structure of the DNA within this region, allowing for contact between the Gln-37 residue and positions 6 and 7, simultaneously, without fully extending the Gln side chain.

**DISCUSSION**

The ability of P22 repressor protein to distinguish between the six naturally occurring operator binding sites is critically important in determining whether the bacteriophage grows lytically or lysogenically. Examination of the naturally occurring operator sequences reveals highly conserved bases at positions 2 and 4, except one of the 12 operator half-sites in the phage chromosome. Therefore, differential affinity of the repressor for the consensus base sequence at operator positions 3, 5, 6, and 7 is conferred by direct contacts from three amino acids on the solvent-exposed surface of the recognition helix of P22 repressor. The pattern of operator sequence conservation permits these amino acid-base pair contacts to occur in all except one of the 12 operator half-sites in the phage chromosome. Therefore, differential affinity of the repressor for these sites cannot be due to these conserved amino acid-base pair interactions. We have shown in other work that the noncontacted bases at the center of the P22 operator affect the relative affinity of P22 operator for repressor by modulating the strength of specific protein-DNA contacts via DNA sequence-specific effects on the structure of the repressor-operator complex. Our recent studies presented here show that the nonconserved bases at positions 2 and 4 also play an important role in determining the relative affinity of repressor for its naturally occurring operators. The results of the mutagenesis studies indicate that direct interactions between the three solvent-exposed amino acids and the conserved bases in the binding site lock these amino acids in place, forming a scaffold allowing the rest of the amino acid side chains to form weaker interactions with the nonconserved bases in the binding site.

The idea that the direct contacts between repressor and conserved bases in the operator facilitate recognition of the nonconserved bases in the operator is best illustrated by examining recognition of the bases at position 2, 3, and 4. Recognition of the bases at these positions is eliminated by mutating the residue Asn-32, which only directly contacts position 3. Restoring a non-native contact at position 3 by substituting a Gln at this position creates a P22 repressor with an alternate position 3 preference and also restores the ability of repressor to discriminate between base changes at operator positions 2 and 4. This result strongly suggests that recognition of positions 2 and 4 by the amino acid at position 32 in the protein requires a direct interaction of amino acid 32 with the base at operator position 3.

Although the ability of the repressor to express base preferences at operator positions 2 and 4 requires contacts between Asn-32 and/or Val-33 at positions 3 and 5, respectively, it is not evident from the biochemical data how the amino acid-conserved base pair contacts help in recognizing the nonconserved bases. Two nonexclusive hypotheses can explain these results. First, direct amino acid-base pair contacts at the conserved positions may serve to properly position the protein side chains so that atoms not involved in making the primary contact can form weak interactions between the amino acids and functional groups on the bases. Alteration of these weak interactions would have measurable but not catastrophic effects on repressor affinity. Alternatively, the nonconserved bases may influence the range of DNA structures that the operator can form or alter the ease of forming the DNA structure that is most compatible with complex formation. According to this idea, in order for the repressor to form the appropriate contacts with the conserved bases, the P22 operator must be deformed. The changes in the sequences of nonconserved bases would affect operator affinity for protein by changing the energy needed to deform the DNA or alter the strength of conserved base pair amino acid contacts.

Despite making base-specific contacts, changing either amino acids 32 or 33 to alanine has at most an 8-fold effect on the apparent affinity of the repressor for operator. This observation suggests that the contacts made by these residues to the DNA do not make a large contribution to the overall stability of the repressor-operator complex. Instead these amino acids function in a relatively energetically neutral fashion to provide sequence specificity to the repressor. This suggestion is consist-
ent with our previous observation that changing the composition of the noncontacted central sequence can eliminate repressor base preferences at positions 3 or 5, with only a ~20-fold decrease in affinity of repressor for operator (15).

Binding data show that P22 repressor position 4 base preference is modulated by Asn-32 and Val-33. An examination of the model of the P22 repressor-operator complex suggests that the P22 repressor preference for a T:A base pair at this position may result from both weak interactions of protein groups with this base pair and sequence-dependent DNA structural effects. The Asn-32 β-CH₂ and δ-CH₃ of Val-33 form a hydrophobic pocket into which nestles the thymine C-5 CH₃ at operator position 4. These amino acid side chain groups are held in this position via their interactions with positions 3 and 5, respectively. Changing the base at position 4 would remove the CH₃ group from this pocket, resulting in the observed decrease in affinity. Structurally, the model of the P22 repressor-DNA complex reveals that the DNA deformation observed in the complex is essential in allowing the contact between Val-33 and the bases at position 5. The T:A at position 4 may help facilitate the observed DNA deformation within this region.

Although the detailed analysis of our data provides a plausible explanation for P22 repressor position 4 base preferences, we are unable to gain a similar insight into how P22 repressor specifies the position 2 T:A base pair. Either or both of two circumstances may prevent us from determining the structural basis for repressor operator position 2 base preferences. First, our modeling may not take into account DNA distortions that could be occurring within this region of the operator in the complex. If the DNA substantially bends toward the protein in this region, a direct interaction between either the Asn-32 residue or some other portion of the protein may be possible. An examination of the structure of homeodomain protein anti-DNA complex shows that a DNA distortion within a region of that complex that is analogous to the region around P22 operator position 2 allows for contacts to be made between the protein and the DNA that could not be made with B-form DNA (24). There is currently no biochemical evidence indicating whether or not the P22 repressor protein significantly bends the DNA in the region lying near amino acid 32. Alternatively, the observed position 2 specificity may be determined indirectly through DNA sequence-dependent structural variation. For example, bases not directly contacted by the CAP protein influence the stability of the CAP-DNA complexes by changing the ease with which this can be bent (25–27). The position 2 T:A base pair may play a similar role in the P22 repressor-operator complex. This T:A base pair forms part of a stretch of poly(dA·dT) sequence at the end of the P22 operator. Structural studies suggest that these sequences are rigid and straight (28–31) and that they introduce kinks or bends into the DNA at the junction between the A-tract and the surrounding sequence. Any base pair substitution may result in a deviation from the preferred DNA conformation.

The six naturally occurring operator sites bind P22 repressor with different affinities, and these differences are crucial to the developmental decisions of the bacteriophage (32). We have previously shown that changes in the sequence of bases at the center of the P22 operator can markedly affect operator affinity for P22 repressor. These effects alone cannot account for the observed differential affinities of the naturally occurring binding sites for P22 repressor. We suggest that variation at positions 2 and 4 together with central base sequence differences act in concert to determine relative operator affinity for repressor. For example, based on their similar central base sequence, we would predict that O₂ and O₃ would have the identical affinity for repressor, but instead O₂ binds repressor 2-fold more tightly than does O₃. This small but biologically significant difference in relative affinity for repressor can be readily accounted for by the observation that O₃ bears a A:T base at position 4, which should reduce its affinity for repressor relative to O₂ by about this amount. Unfortunately, taking into account the effect of variations in both the central base sequence (13–15) and positions 2 and 4 on operator affinity for repressor does not allow us to completely account for the relative affinities of all the naturally occurring binding sites. We hypothesize that this inability to predict the relative affinity of an operator for repressor may be due to the long distance effects of central base sequence on recognition of contacted bases. Further operator mutagenesis and binding studies will be required to test this idea.

Acknowledgments—The authors thank Dr. Paul Gollnick for critical reading of the manuscript. We also acknowledge members of our laboratory for helpful discussions.

REFERENCES

1. Poteete, A., Ptashe, M., Ballivet, M., and Eisien, H. (1980) J. Mol. Biol. 137, 81–91
2. Sauer, R. T., Yucum, R. R., Doolittle, R. F., Lewis, M., and Pabo, C. O. (1982)
3. Nature 296, 447–451
4. Ptashe, M. (1986) A Genetic Switch, Blackwell Publishing Co., Los Angeles
5. Sauer, R. T., Pabo, C. O., Meyer, B. J., Ptashe, M., and Backman, K. (1979)
6. Nature 279, 396–400
7. Wharton, R. F., Brown, E. L., and Ptashe, M. (1985) Cell 38, 361–369
8. DeAnda, J., Poteete, A. R., and Sauer, R. T. (1983) J. Biol. Chem. 258, 10536–10542
9. Sevilla-Sierra, P., Otting, G., and Wuthrich, K. (1994) J. Mol. Biol. 235, 1003–1020
10. Aggarwal, A., Rodgers, D. W., Drott, M., Ptashe, M., and Harrison, S. C. (1988) Science 242, 889–897
11. Wharton, R. F., and Ptashe, M. (1985) Nature 316, 601–605
12. Koudelka, G. B., Harrison, S. C., and Ptashe, M. (1987) Nature 326, 886–888
13. Koudelka, G. B., and Carlson, P. (1992) Nature 355, 89–91
14. Koudelka, G. B., Harbury, P. B., Harrison, S. C., and Ptashe, M. (1988) Proc. Natl. Acad. Sci. U. S. A. 85, 4633–4637
15. Wu, L., Vertino A., and Koudelka, G. B. (1992) J. Biol. Chem. 267, 3134–9135
16. Wu, L., and Koudelka, G. B. (1993) J. Biol. Chem. 268, 18975–18981
17. Ilchey, S. P., and Koudelka, G. B. (1997) J. Biol. Chem. 272, 1646–1653
18. Maniatis, T., Fritsch, E. F., and Sambrook, J. (1982) Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory, NY
19. Messing, J. (1983) Methods Enzymol. 101, 29–79
20. Sanger, F., Donelson, J. E., Couplan, C. L., and Fischer, D. (1973) Proc. Natl. Acad. Sci. U. S. A. 70, 1209–1213
21. Tabor, S. (1990) in Current Protocols in Molecular Biology (Ausubel, F., Brent, R., Kingston, R. E., Moore, D. D., Seidman, J. G., Smith, J. A., and Struhl, K., eds) pp. 16.2.1–16.2.11, Wiley-Liss, Inc., New York
22. Wharton, R. P., Harrison, S. C., and Ptashe, M. (1987) Nature 326, 888–891
23. Bejar, S., Bouche, F., and Bouche, J. P. (1988) Mol. Gen. Genet. 212, 11–19
24. Bejar, S., Cam, K., and Bouche, J. P. (1989) Nucleic Acids Res. 17, 6821–6833
25. Bejar, S., and Bouche, J. P. (1985) Mol. Gen. Genet. 201, 146–150
26. Otting, G., Qian, Y. Q., Helled, T., Muller, M., Affolter, M., Gehring, W. J., and Wuthrich, K. (1990) EMBO J. 9, 3085–3092
27. Schulte, S. C., Shiel, C., and Steitz, T. A. (1991) Science 253, 1001–1007
28. Gartenberg, M. R., and Crothers, D. M. (1986) Nature 323, 824–829
29. Dalma-Weiszhausz, D., Gartenberg, M. R., and Crothers, D. M. (1991) Nucleic Acids Res. 19, 611–616
30. Nelson, H. C. M., Finch, J. T., Luisi, B. F., and Klug, A. (1987) Nature 320, 221–226
31. Lomonosoff, G. P., Butler, P. J. G., and Klug, A. (1981) J. Mol. Biol. 149, 745–760
32. DiGabriele, A. D., Sandersen, M. R., and Steitz, T. A. (1989) Proc. Natl. Acad. Sci. U. S. A. 86, 1816–1820
33. DiGabriele, A. D., and Steitz, T. A. (1993) J. Mol. Biol. 231, 1024–1039
34. Poteete, A., and Ptashe, M. (1982) J. Mol. Biol. 157, 21–48