^{31}P NMR Spectra of an Oligodeoxyribonucleotide Duplex *Lac* Operator–Repressor Headpiece Complex†

Christine Karslake, Stephen Schroeder, Pei Ling Wang, and David G. Gorenstein*
Department of Chemistry, Purdue University, West Lafayette, Indiana 47907
Received February 6, 1990; Revised Manuscript Received April 11, 1990

ABSTRACT: The interaction of a symmetric *lac* operator duplex, d(TGTGAGCGCTCACA)$_2$, with the N-terminal 56-residue headpiece fragment of the *lac* repressor protein was monitored by ^{31}P NMR spectroscopy. The changes in the ^{31}P chemical shifts upon addition of the headpiece demonstrated an end point of two headpiece fragments per symmetric 14-mer duplex with each headpiece binding to the T1pG2pT3pG4pA5 ends of the duplex. The specific phosphate ^{31}P perturbations observed are consistent with those residues implicated in protein binding by previous NMR, molecular biological, and biochemical techniques. Upon complexation, the ^{31}P signals of phosphates G2–A5 showed upfield or downfield shifts (<0.2 ppm) while most other residues were unperturbed. The interactions were dependent on ionic strength. The ^{31}P NMR data provide direct evidence for predominant recognition of the 5′ strand of the 5′-TGTGA/3′-ACACT binding site.

The *lac* repressor and operator system has been the prototype for studying protein–DNA interactions (Gilbert et al., 1975). Although extensively studied (Caruthers, 1980; Ptashne, 1986), it is still not understood at a detailed molecular level. The *lac* repressor protein is proposed to bind to its operator through a helix–turn–helix motif (Boelens et al., 1987), which shares homology with other bacterial DNA-binding proteins (Ptashne, 1986). Both molecular biology and NMR have proved to be fruitful in understanding the molecular basis for the *lac* repressor and operator system. The DNA-binding proteins that have been studied by X-ray crystallography (unfortunately not including the *lac* repressor) have each been found to bind to DNA as a dimer so that the protein–DNA complex exhibits a 2-fold axis of symmetry, with one of the α-helices in the monomer unit binding in the major groove of the operator. X-ray crystallographic structural refinements and extensive mutagenesis experiments [cf. Betz (1986, 1987), Ebright (1986), Eisenbeis et al. (1985), Koudelka et al. (1987), Wharton et al. (1984), Wharton and Ptashne (1985, 1987), and Lehming et al. (1987, 1988)], as well as model building on a number of operator DNA complexes, suggest that certain side-chain residues recognize the individual bases of nucleic acids, while others "recognize" the backbone phosphates, possibly through sequence-specific variations in the DNA conformation.

The *lac* repressor system has been extensively studied by ^{1}H NMR (Buck et al., 1980, 1983; Hogan et al., 1981; Nick et al., 1982; Scheek et al., 1983). It appears to be possible to duplicate the basic *lac* operator–*lac* repressor protein interaction by using the smaller *lac* repressor headpiece N-terminal domain fragment (Adler et al., 1972; Buck et al., 1978; Wade-Jardetzky et al., 1979; Wemmer & Kallenbach, 1983; Boelens et al., 1987) and small 14–22-bp DNA fragments that comprise the *lac* operator (two subunits of the tetrameric *lac* repressor bind to the two halves of the partially palindromic *lac* operator; Pilz et al., 1980).

These studies have attempted to gain a better understanding of the detailed binding interactions between the regulatory proteins, such as the *lac* repressor, and DNA operator sequences. Most efforts directed toward defining the binding specificity between amino acid sequences and DNA sequences have centered on hydrogen-bonding to the acceptor/donor groups on the Watson–Crick base pairs in the major groove [cf. Landschulz et al. (1988)]. At present we do not understand this "second genetic code" of protein–DNA recognition. Perhaps one reason for the inability to dissect the basis for this specificity is the emphasis on base-pair interactions alone. As described in this paper, the conformation and position of the phosphates may be an important component of protein–DNA recognition.

Both ionic and nonionic interactions are believed to be responsible for the tight binding of repressor protein to operator. Indeed, ionic interactions between proteins and their recognition sequences have been shown to be of major importance. Thus, the recent crystal structure of the *trp* repressor demonstrated that every one of the direct protein contacts was mediated through interactions with the phosphate backbone (Otwinowski et al., 1988). Ionic interactions involving the phosphate backbone have been implicated as being important factors in the recognition of the *lac* operator as well. Specifically, deHaseth et al. (1977) determined that *lac* repressor–operator interaction involves approximately 10–12 ionic interactions per operator duplex. Therefore, the repressor protein may be recognizing the geometry of the phosphate backbone as a very important feature. It is not known whether any of these ionic interactions provides a specific recognition mechanism for the repressor. As one might expect, ionic strength plays a large role in these ionic interactions.

^{31}P Chemical Shifts as a Probe of DNA Structure. It is now widely appreciated that duplex DNA can show local conformational heterogeneity in the base-pair geometry as well as the deoxyribose phosphate backbone (Calladine, 1982; Dickerson, 1983; Dickerson & Drew, 1981). While base-pairing and stacking interactions can often be adequately defined from ^{1}H NMR, NOESY-derived distances (Nilges et al., 1987), detailed sequence-specific conformational information with regard to the deoxyribose phosphate backbone

† Supported by NIH (AI27744), the Purdue University Biochemical Magnetic Resonance Laboratory, which is supported by NIH (Grant RR01077 from the Biotechnology Resources Program of the Division of Research Resources), the NSF National Biological Facilities Center on Biomolecular NMR, Structure and Design at Purdue (Grants BBS 8614177 and 8714258 from the Division of Biological Instrumentation), and the National AIDS Research Center at Purdue (AI727713).

0006-2960/90/0429-6578$02.50/0 © 1990 American Chemical Society

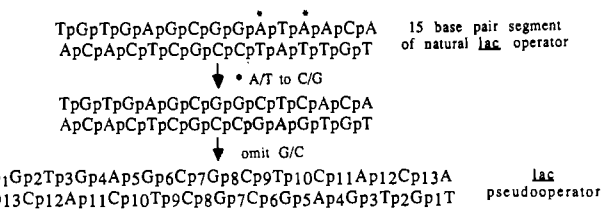

FIGURE 1: Natural operator and symmetric pseudooperator [14-mer, d(TGTGAGCGCTCACA)$_2$]. Asterisks indicate base-pair alterations.

is often not available. Recently we have attempted to utilize ^{31}P NMR to define sequence-specific variations in the *backbone* conformation of oligonucleotides (Gorenstein et al., 1988; Nikonowicz et al., 1989a,b; Schroeder et al., 1989; Roongta et al., 1990; Powers et al., 1990).

Variations of ^{31}P chemical shifts of individual phosphates in duplex oligonucleotides have been suggested to be attributable to torsional angle changes in the deoxyribose phosphate backbone. Various chemical shift calculations (Gorenstein, 1984, 1987; Gorenstein & Kar, 1975) and ^{31}P signal assignments in oligonucleic acids (Gorenstein et al., 1987; Lai et al., 1984; Schroeder et al., 1986; Shah et al., 1987a,b) have confirmed that the ^{31}P resonance of a phosphate diester in a g^-,g^- conformation should be 1.6 ppm upfield of the ^{31}P signal of an ester in a t,g^- conformation (t, trans or 180° dihedral angle) (Gorenstein, 1984; Roongta et al., 1990). [The notation for the P–O ester torsion angles follows the convention of Seeman et al. (1976) with the ζ, P–O-3′ angle given first followed by the α, P–O-5′ angle; both P–O ester bonds have dihedral angles of −60° or −gauche, g^-.]

Thus one of the major contributing factors that we have hypothesized determines ^{31}P chemical shifts is the main-chain torsional angles of the individual phosphodiester groups along the oligonucleotide double helix. It should be noted that environmental effects on the ^{31}P chemical shifts of nucleic acids (Costello et al., 1976; Lerner & Kearns, 1980) are generally smaller than the intrinsic conformational factors discussed above, assuming comparisons are made under similar solvation conditions. Similarly, other possible effects, such as ring-current shifts are not likely to be responsible for shifts >0.01 ppm (Giessner-Prettre et al., 1976; Gorenstein, 1984).

As described in this paper, we have assigned the ^{31}P NMR spectrum of the 14-base-pair *lac* "pseudooperator" segment d(TGTGAGCGCTCACA)$_2$ bound to the *lac* repressor headpiece. This symmetrical base sequence is about two-thirds the length of the 21-base-pair wild-type sequence and is believed to contain most of the important contact sites (Sadler et al., 1983; Simons et al., 1984). The pseudooperator segment is derived from the natural operator sequence as shown in Figure 1. The pseudooperator differs from the natural operator sequence by two A-T to G-C base-pair conversions, indicated by the two asterisks in Figure 1, and by deletion of a single G-C base pair at its center. Interestingly, these sequence alterations result in a 10-fold increase in repressor binding affinity over that for the natural operator sequence (Sadler et al., 1983). Here we present the first ^{31}P spectral changes of an operator duplex upon binding repressor protein. The data suggest that specific phosphates of the tight-binding operator do interact in an ionic fashion with residues of the headpiece.

Experimental Procedures

Synthesis. The self-complementary 14-mer d(TGTGAGCGCTCACA)$_2$ was synthesized by a manual modification of the solid-phase phosphite triester method (using 10 μmol of the starting nucleoside-derivatized support at a 28 mmol/g loading level for each separate synthesis) as previously described (Lai et al., 1984; Schroeder et al., 1987; Shah et al., 1984a,b). The resulting tritylated product was cleaved from the support and purified by C-18 reverse-phase HPLC with an acetonitrile/triethylammonium acetate buffer gradient as previously described (Gorenstein et al., 1988). The amounts of purified oligonucleotide obtained through several separate syntheses were 43.6 mg (871 OD units, 12.0 μmol).

NMR Measurements. ^{1}H spectra were obtained on a Varian VXR-500 NMR spectrometer. The ^{31}P 1D NMR spectra and the 2D ^{1}H/^{31}P correlation COLOC (PAC) and J-resolved spectra were run on Varian XL-200A spectrometer at ambient temperature (ca. 25 °C) operating at 81.1 MHz. A sweep width of 172 Hz, an acquisition time of 2.98 s, a block size of 1K zero filled to 16K, and a pulse width of 7 ms were use for the 1D spectra. Spectra were resolution-enhanced by using a combination of positive exponential and Gaussian apodization functions. Typical values were 0.1–0.2 resolution enhancement values and 0.5–0.6 apodization function values, respectively. The values were adjusted in accordance with the signal to noise ratio in the ^{31}P NMR spectrum. The number of acquisitions for each spectrum was typically between 2000 and 3000. The ^{31}P resonances were referenced to an external sample of trimethyl phosphate (TMP) at 0.0 ppm, which is 3.53 ppm downfield of 85% phosphoric acid.

A Bio-Ion Cf desorption mass spectrometer was used to determine the molecular weight of the headpiece. A 10-μL aliquot of a 10 μg/mL sample followed by 1 μL of ethanol was applied to a nitrocellulose coated mylar target and absorbed for a half hour. The sample was then washed with 50 mL of a 0.1% trifluoroacetic acid in water solution and dried. The instrument was set in the positive mode at 15000 V. Data were collected every nanosecond for 4 h. The TDC group size was set for 8192.

Lac Repressor Headpiece. Plasmid pHIQ3 and *Escherichia coli* strain MC1061 were gifts from Dr. Joan Betz. This overproducing strain provided ca. 200 mg of *lac* repressor protein/100 L of culture.

Lac repressor headpiece consisting of the N-terminal 56 residues was prepared according to Arndt et al. (1981) with the additional quenching of the α-chymotrypsin (Worthington) digest with a 4 molar equiv excess of turkey trypsin inhibitor (Sigma).

^{31}P NMR samples were prepared by dissolving 4–6 mg (3 mM duplex) of the lyophilized DNA in 0.4 mL of buffer solution in D_2O. Both the 14-mer pseudooperator and the headpiece were centrifuged in Sephadex G-10 (Pharmacia) spun columns (Maniatis et al., 1982) to change the buffer conditions. Both headpiece and DNA concentrations were measured by UV spectroscopy. Amounts of the 14-mer were determined spectrophotometrically by using the relationship of 20 absorbance units (optical density, OD units)/1.0 mg of DNA at 260 nm.

Results

Assignment of ^{31}P Signals of 14-mer Oligonucleotide Duplex. The ^{31}P signals of the tetradecamer were previously assigned by both 2D pure absorption phase constant time (PAC) heteronuclear correlation NMR (Fu et al., 1988) and ^{17}O/^{18}O labeling methodologies (Schroeder et al., 1987). The ^{31}P spectrum for the 14-mer is shown in Figure 2.

Headpiece. The 56-residue *lac* repressor headpiece was characterized by both proton NMR and mass spectroscopy. The proton NMR spectrum of the purified headpiece was identical with that previously published (Arndt et al., 1981). A Bio-Ion Cf desorption mass spectrometer, accurate to ±5

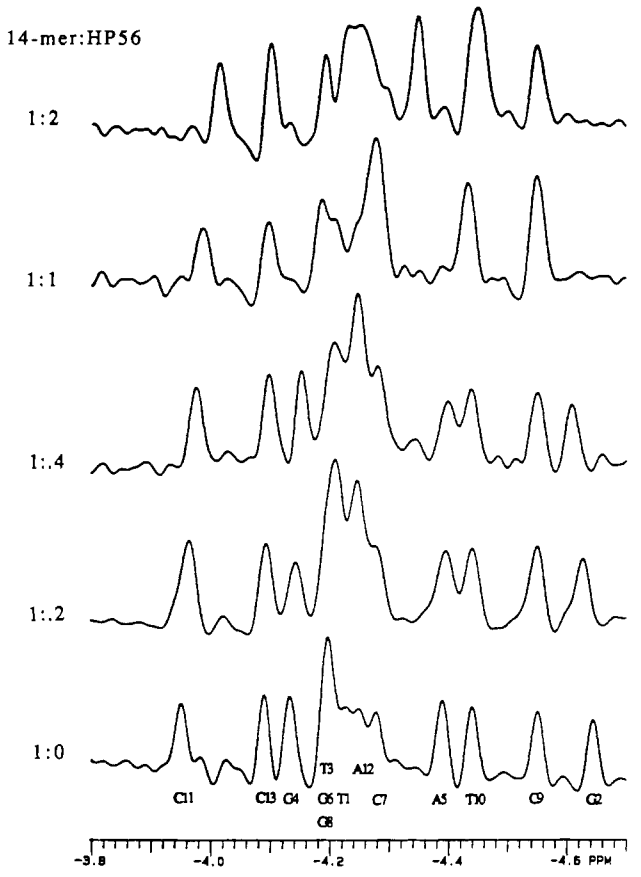

FIGURE 2: ^{31}P NMR spectra and phosphate assignments of the 14-mer, d(TGTGAGCGCTCACA)$_2$, at various *lac* headpiece:DNA ratios. Numbering corresponds to the phosphate position from the 5′-end of the duplex.

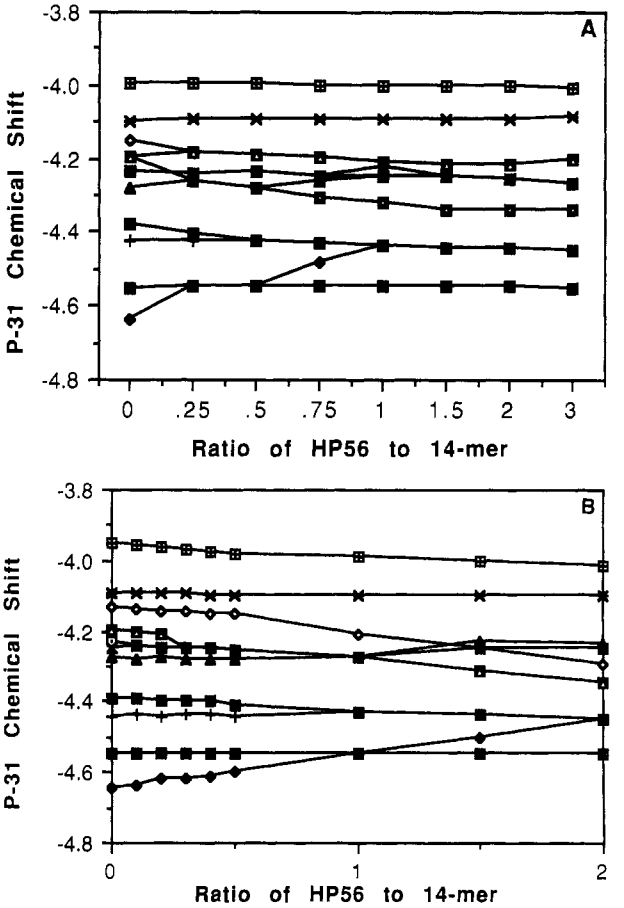

FIGURE 3: Titration of the ^{31}P chemical shifts and assignments of the 14-bp oligonucleotide duplex, d(TGTGAGCGCTCACA)$_2$, as a function of the relative ratio of [*lac* repressor headpiece]:[14-bp "super" operator] at moderate ionic strength, 0.09–0.17 M (A), and low ionic strength, 0.067 M (B).

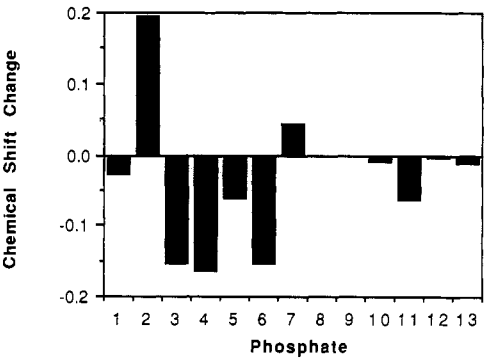

FIGURE 4: Summary bar plot of perturbation of ^{31}P chemical shift of individual phosphates of the 14-mer operator upon binding 2 equiv of the 56-residue *lac* repressor headpiece. There is some ambiguity in the signal assignment of phosphates 3, 6, and 8. Thus, only two of the ^{31}P signals of phosphates 3, 6, and 8 are shifted upfield; the third (assumed to be phosphate 8) is not perturbed.

mass units, was used to determine the molecular weight of the headpiece. The calculated molecular mass was 6227 daltons while the measured mass was 6231 daltons, which is within the experimental error for this technique.

Headpiece Titration. The ^{31}P NMR spectra at various [operator]:[*lac* repressor headpiece] ratios are shown in Figure 2. Plots of the ^{31}P chemical shift vs [operator]:[*lac* repressor headpiece] ratio are shown in Figure 3 at two different ionic strengths. Because of peak overlap during various additions of headpiece to the operator, it is necessary to clarify how the phosphate resonances were assigned during the titration. We assumed that perturbations in the ^{31}P chemical shift were a continuous function throughout the titration. Thus, peaks that did not shift with added protein remained unshifted once the peak had merged with another peak. In addition, the slope of the perturbation was assumed to stay constant during and after merging of the peaks.

The ^{31}P signals of phosphates 2, 4, and 5 are well resolved at the beginning of the titration and clearly show significant perturbation upon addition of headpiece. As shown in the summary bar plot of Figure 4, the G2 phosphate shifts 0.20 ppm downfield while the G4 phosphate shifts 0.16 ppm upfield during the titration. The A5 phosphate shifts 0.06 ppm upfield during the titration. Unfortunately the ^{31}P signals of phosphates 3, 6, and 8 are coincident in the absence of repressor, and in the 2:1 complex two of the three shift 0.15 ppm upfield during the titration. It is most likely that phosphate 3 is one of the two that is perturbed because it is adjacent to phosphates 2 and 4, which experience the greatest perturbation. The ^{31}P signals of the remaining phosphates 1, 7, 9, 10, 12, and 13 show either no or small perturbations (<0.05 ppm) upon titration with headpiece (11 shows a modest upfield shift of 0.07 ppm).

Assuming that the magnitude of the ^{31}P shift reflects the degree of interaction of the phosphate with the headpiece, we can conclude that phosphates 2, 3, 4, and 5 (and possibly 6) represent a major binding site for the headpiece.

The interaction of the headpiece with operator is dependent on ionic strength. At an ionic strength of 0.067, two headpieces are bound per operator duplex (Figure 3A). At an ionic strength of 0.514, no interaction is observed (spectra not shown). During a titration encompassing ionic strengths ranging from 0.087 to 0.169 only one headpiece is bound per duplex operator (Figure 3B). Although the ionic strength varied during the titration, the final ^{31}P spectrum was unaltered

upon dilution to the initial ionic strength.

Discussion

The ^{31}P spectral changes upon binding the 56-residue headpiece to the 14-mer operator (Figure 2) demonstrate that all of the phosphate resonances remain in fast exchange during the entire course of the titration. This fact greatly simplified the interpretation of the results because only one signal was observed per phosphate rather than two, which would have been the case under slow exchange.

At low ionic strength the ^{31}P titration curves (Figure 3A) level off at a ratio of two headpiece molecules (HP) per one operator duplex (O). Therefore, one headpiece is bound at each of the putative dTGTGA recognition sequences that are located on either side of the operator's 2-fold axis of symmetry. It is assumed that each headpiece makes the same contacts as the other headpiece on the basis of symmetry considerations. This is supported by the titration curves, which show that the chemical shift perturbations at a 1:1 ratio are ca. $^1/_2$ those at the final 2:1 [HP]:[O] ratios. Therefore the observed chemical shift represents the population weighted average of the chemical shift of free operator and the headpiece-bound state, with little chemical shift difference between the 1:1 and 2:1 complexes. These ^{31}P titration results are consistent with earlier ^{1}H NMR titration data (Lamerichs et al., 1989).

The strength of the *lac* repressor–operator interaction is very much dependent on the ionic strength of the solution (Riggs et al., 1970; Whitson et al., 1986; Winter & von Hippel, 1981; Record et al., 1977). This is consistent with a large contribution of ionic stabilization of the complex. The perturbations in the ^{31}P spectra of the operator in the presence of headpiece are strongly influenced by the ionic strength of the solution, consistent with the ionic association mechanism. When the titration was carried out in a buffer of 0.05 M phosphate, pH 7.4, and 0.4 M KCl, no phosphate ^{31}P signals were perturbed upon addition of headpiece, indicating that no interaction occurred (data not reported). Surprisingly, we find that in moderate ionic strength solutions (0.09–0.17 M) the ^{31}P titration curve levels off at a 1:1 complex (Figure 3B). This suggests that at higher ionic strength, where the binding of headpiece to operator is weaker, a second headpiece protein is prevented from binding to the operator. This ionic-strength-related, anticooperative behavior is likely associated with the net reduction of the duplex anionic electrostatic potential upon binding of the first headpiece molecule.

The interaction is highly specific, with the only major ^{31}P changes occurring at phosphates at positions 2 (downfield) and 3, 4, 5, and 6 (upfield) (i.e., TpGp*Tp*Gp*Ap*Gp*, asterisks indicating position of interaction). Smaller perturbations are found at phosphates 1, 7, and 11.

Most importantly, the interaction is asymmetric, with little perturbation of the complementary phosphates 9–13 on the opposite 3′ strand (Figure 4). Note that of the two ^{31}P signals of complementary phosphates 2 and 12 in the G2pT3/C13pA12 base step only the phosphate 2 signal shifts downfield upon addition of headpiece. Similarly, the ^{31}P signals of complementary phosphates 3 and 11 in the T3pG4/A12pC11 base step both shift upfield upon addition of headpiece, although again the major perturbation is observed in the 5′ strand (T3 phosphate). The complementary phosphate (T10) of the most upfield-shifted ^{31}P signal (phosphate G4) is also not perturbed at all. Clearly, only the 5′ strand of the operator is influenced by binding to the repressor while the base-pairing 3′ strand is essentially unaltered.

Origin of ^{31}P Chemical Shift Perturbations. As indicated above, studies carried out on oligonucleotide duplexes and drug–duplex complexes indicate that ^{31}P chemical shifts are strongly dependent on the phosphate diester conformation. Thus, any induced structural change of the pseudooperator occurring upon repressor headpiece binding will be evident from resonance shifts in the ^{31}P spectrum. ^{31}P chemical shifts are also dependent on the electrostatic environment, which can be significantly altered through binding contact with an amino acid side chain. Thus, changes occurring in the ^{31}P spectrum upon operator–repressor complex formation should lead to better understanding of the binding interaction.

The ^{31}P chemical shifts of four of the phosphates were shifted either downfield or upfield during the titration by at most 0.1–0.2 ppm. These ^{31}P chemical shift changes upon binding the headpiece to the operator suggest trhat recognition occurs via electrostatic, groove-binding interactions rather that intercalation, since intercalation is expected to produce a 2–3 ppm downfield ^{31}P shift while electrostatic association should produce considerably smaller (and generally upfield shifts) for the phosphates directly involved in the interaction (Gorenstein, 1978, 1981, 1984; Gorenstein et al., 1981; Patel, 1979; Reinhardt & Krugh, 1977).

These results also demonstrate that the conformations of the operator phosphates are not significantly perturbed upon binding of headpiece. Thus the ^{31}P resonance of a phosphate diester in a g^-,g^- conformation (both P–O ester bonds have dihedral angles of −60° or −gauche, g^-) has been shown to be 1.5–1.6 ppm upfield of the ^{31}P signal of an ester in a t,g^- conformation (t, trans or 180° dihedral angle) (Gorenstein, 1984). The ^{31}P chemical shifts are also sensitive to local helical parameters such as helix twist and base-pair roll (Schroeder et al., 1989). These geometric changes influence the conformation of the sugar phosphate backbone and are reflected in changes in the phosphate ester conformation and hence ^{31}P chemical shift (Schroeder et al., 1989).

The small upfield and downfield shifts of the phosphates upon addition of headpiece suggest that there is very little perturbation of the local helix geometry and backbone conformation in the complex relative to the free operator. As shown in Figure 2, the spread in the ^{31}P chemical shifts of the phosphates in the complex is narrower (0.5 ppm) than in the free operator (0.7 ppm). The backbone conformation is more uniformly g^-,g^- in the complex than in the free operator, although the degree of this perturbation is small.

Previously we have shown that the ^{31}P chemical shift of a phosphate in oligonucleotide duplexes and duplex–drug complexes is shifted downfield by ca. 1.0 ppm per 20° of helix unwinding (Lai & Gorenstein, 1989). The repressor protein is believed to unwind the operator by 90° (Barkley & Bourgeois, 1978; Wang et al., 1974). Even if this unwinding were uniformly distributed over all of the operator base steps, this degree of unwinding should have resulted in a major change in the conformation of the phosphates and hence should have produced large ^{31}P chemical shift changes. We do not observe any net downfield shift of the phosphates. This result is thus inconsistent with a major unwinding of the operator in the duplex complex. Indeed, our results suggest that the headpiece perturbs the sugar phosphate backbone very little. The discrepancy between the earlier unwinding data and the ^{31}P results could possibly be attributed to an inability of the headpiece molecule to unwind the duplex to the same degree as the entire *lac* repressor protein.

Comparison with Other Data. Phosphate ethylation experiments demonstrate that phosphates 2, 3, 8, and 9 (numbering based upon 14-mer sequence) are ethylated to either a greater or lesser degree upon addition of repressor (Barkley

& Bourgeois, 1978). Our results definitely show phosphate 2 being involved and phosphate 3 probably being involved in the recognition process. Phosphate 9, which remains well resolved during the entire titration, was definitely not perturbed upon addition of the headpiece. It is possible that residues not included in the headpiece are responsible for the observed ethylation increase by ethylnitrosourea observed for phosphate 9.

Noble et al. (1984), through the use of various methyl phosphonate substitutions along the 21-base-pair natural operator sequence, have shown that at least seven phosphate positions may be recognized through electrostatic interaction with *lac* repressor. In addition, the interaction at any particular phosphorus is stereospecific, which one might expect owing to the specificity of the binding interaction as a whole.

^{1}H NOEs are observed in the headpiece–operator complex to Tp1, p7Gp8, p8Cp9, and p9Tp10 bases and to two bases (an adenine and a thymine) that are not included in our 14-mer operator sequence (Lamerichs et al., 1989). These NOEs and model-building studies of Kaptein and co-workers are not entirely consistent with our ^{31}P NMR data. However, ^{1}H NMR studies in the imino region have shown perturbation in the AT* base pair in the TpGpT*pG portion of a symmetrical 18-bp operator–51-residue headpiece complex (Buck et al., 1986). Our results indicate very little (if any) interaction of the headpiece with phosphates 7–10. Of course, it is entirely possible (and even quite likely) that interactions observed in the ^{1}H NMR NOESY experiment are not reflected in the interactions observed in our ^{31}P NMR spectra. Thus, if the protein makes contact to bases 7–10, it may not be simultaneously binding to the phosphates in this region of the operator as well. Similarly, if the contacts to the operator in the G2–A5 region are to the phosphate ester backbone, then contacts to the bases may not be present. We should emphasize again that our ^{31}P data provide the only direct evidence for predominant recognition of the 5′-strand of the 5′-TGTGA/3′-ACACT binding site.

A number of site-specific mutants of the *lac* repressor (Lehming et al., 1987, 1988; Betz, 1986, 1987; Ebright, 1986) and the relative binding affinities to various wild-type and mutant operators provide support for the model proposed by Kaptein and co-workers (Lamerichs et al., 1989). Betz has characterized several tight-binding mutants ($Val_{52} \rightarrow Asn_{52}$ and $Val_4 \rightarrow Ala_4$), and Lehming et al. have constructed a number of *lac* repressor mutants that have different affinities for various operators. The data of Lehming et al. confirm that the major recognition site of the operator is the 5′-TGTGA sequence starting at the first G. Mutagenesis studies by Muller-Hill and co-workers (Lehming et al., 1987, 1988) on the repressor protein have implicated p1Gp2, p2Tp3, and p3Gp4 as sites of contact. Our results confirm that p2–p6 are likely all involved in the recognition of the operator DNA.

The data of Lehming et al. suggest that residue 1 of the recognition helix (Tyr_{17}) interacts with base pairs 4 and 5 (see above numbering convention), which is consistent with our ^{31}P NMR data. Note, however, that Kaptein's results (Lamerichs et al., 1989) are not entirely consistent with the mutagenesis experiments. Thus Ebright (1986) and Lehming et al. (1987) suggest that residues 17/18 recognize base pairs 4 and/or 5, while Kaptein indicates that Tyr_{17} interacts with base pairs 2/3. Also, Arg_{22} (residue 6 of the recognition helix) contacts base pair 6 (Lehming et al., 1988).

This study defines the points of phosphate perturbation by *lac* repressor headpiece upon interaction with the tight-binding 14-mer operator. These sites agree in general with the recognition sites that have been observed previously by using NMR, molecular biological and phosphate ethylation techniques. The perturbed phosphates account for the 12 ionic interactions that have been defined for the *lac* repressor–operator complex (deHaseth et al., 1977).

Conclusions and Some Implication toward DNA–Protein Recognition and Specificity. Most attention on understanding the binding specificity between amino acid sequences and DNA sequences has centered on hydrogen-bonding to the acceptor/donor groups on the Watson–Crick base pairs in the major groove [cf. Landschulz et al. (1988)]. At present we do not understand this second genetic code of protein–DNA recognition. Perhaps one reason for the inability to dissect the basis for this specificity is the emphasis on base-pair interactions alone. On the basis of analysis of ^{31}P chemical shifts and coupling constants, it has been possible to demonstrate that the conformations of the phosphates in duplex oligonucleotides show sequence specificity (Ott & Eckstein, 1985a,b; Schroeder et al., 1986; Gorenstein et al., 1988, 1990). Do DNA-binding proteins recognize these sequence-specific backbone structural features of DNA? It is not unrealistic to assume that at least a portion of protein–DNA recognition derives from recognition of the sequence-specific variation in the geometry and (in the case of the phosphates) the electrostatic potential of the DNA (an indirect readout mechanism). DNA is not a uniform cylinder of equivalently placed negative charges but rather one in which the phosphate charges can significantly migrate across the surface of the cylinder in response to changes in local helical parameters such as helix twist (in turn responsive to sequence specificity). With helical twist variations of 12–14° expected in normal B-DNA, the phosphate groups can vary from their uniform B-DNA position by ±7°, which translates into linear displacements of 1 Å (indeed, variations of phosphate position by 6–7 Å are observed in DNA–protein complexes; Otwinowski et al., 1988). It would appear that the location of these charges could represent a convenient signpost for designating a specific DNA sequence.

Certainly electrostatic interactions can be a major factor in stabilizing these complexes. In phage 434 repressor, the protein makes three or four hydrogen bonds to the phosphates of the backbone (Anderson et al., 1987). *Extensive* networks of hydrogen bonds are made with the oxygens of the phosphates in the direct operator recognition site of various operators (Jordan & Pabo, 1988; Wolberger et al., 1988; Otwinowski et al., 1988). Most significantly, in *trp* repressor it appears that *all* of the 24 specific, direct protein contacts to the operator are to the phosphate ester phosphoryl oxygens (Otwinowski et al., 1988). *No* direct hydrogen bonds exist between the base pairs of the operator and the repressor.

Dickerson (Dickerson, 1983; Dickerson & Drew, 1981) has noted that significant variation in the sugar phosphate ϵ and ζ torsional angles is observed in the crystal state. Most phosphates are in the B_I conformational states ($\epsilon = t$, $\zeta = g^-$) while several are in the B_{II} conformational state ($\epsilon = g^-$, $\zeta = t$). A crankshaft motion interconverts these extreme phosphate ester conformations. It is important to point out that changes in the torsional angles ϵ and ζ similar to the B_I to B_{II} conformational transitions not only shift the position of the phosphate but also rotate the orientation of the phosphoryl oxygens relative to the helix axis. Figure 5 shows a cylindrical projection of the phosphate ester backbone of a portion of one strand of a 12-mer (Dickerson, 1983) between residues 9 and 12. The phosphate of the G10pC11 base step is in a B_{II} conformation while the other phosphates shown are in the B_I conformation. Note the relative orientation of the two

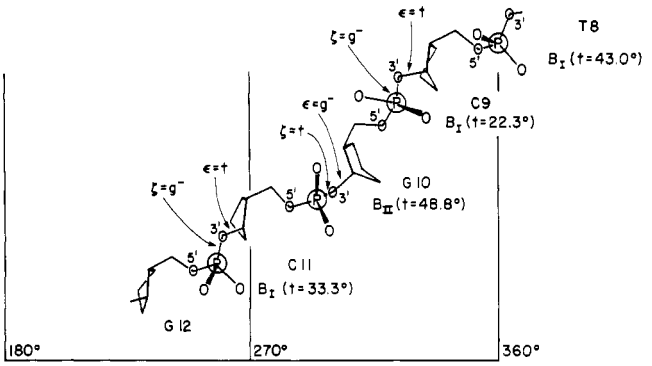

FIGURE 5: Cylindrical projection of the phosphate ester backbone of the Dickerson 12-mer between residues 9 and 12 [derived from Dickerson (1983)]. The phosphate of the G10pC11 base step is in a B_{II} conformation while the other phosphates shown are in the B_I conformation.

phosphoryl oxygens in the figure. The plane formed by the phosphorus atom and the phosphoryl oxygens of the phosphate in the B_I conformation is nearly perpendicular to the plane formed by the phosphorus atom and the phosphoryl oxygens of the phosphate in the B_{II} conformation. Any set of H-bond donors or positively charged groups could readily discriminate between these two conformations that differ significantly in their orientation of the electrostatic potential.

We have recently shown that the population of the B_I/B_{II} conformations is controlled by sequence-specific variation in the local helical structure which can alter the length of the sugar phosphate backbone. Thus if the helix unwinds (for example, in response to purine–purine clash), the sugar phosphate "tether" between a base step must contract. In turn this alters the conformation of the phosphate (Schroeder et al., 1989; Roongta et al., 1990). The response of the phosphate conformation to the sequence-specific variation in local helical parameters can thus explain the "indirect readout" of the operator sequence through the protein–phosphate contacts.

We have shown that the repressor interacts with the phosphates in the 5′-TGTGA-3′ region, which may present a unique set of phosphate conformations and electrostatic potentials to be recognized by the repressor protein. Of note, the 5′-TGTGA-3′ sequence may be conserved in other protein recognition processes (Nussinow, 1986a), and the 5′-TGTG-3′ sequence has been implicated as a recognition signal required for transcription termination in eukaryotes (Nussinov, 1986b). Perhaps a portion of the second genetic code will be found in the sequence-specific variation in the phosphate ester backbone, an often overlooked component of DNA structure.

ACKNOWLEDGMENTS

We greatly appreciate the contributions of Paul Wisniowiski, Dr. Claude Jones, and Dr. Robert Santini, and we thank Dr. Joan Betz for the gift of the plasmid and *E. coli* cell line.

REFERENCES

Adler, K., Beyreuther, K., Fanning, E., Geisler, N., Gronenborn, B., Klemm, A., Muller-Hill, B., & Pfahl, M. (1972) *Nature 72*, 322–327.

Anderson, J. E., Ptashne, M., & Harrison, S. C. (1987) *Nature 326*, 846–852.

Arndt, K. T., Boschelli, F., Lu, P., & Miller, J. H. (1981) *Biochemistry 20*, 6109–6118.

Barkley, M. D., & Bourgeois, S. (1978) in *The Operon* (Miller, J. H., & Reznikoff, W. S., Eds.) pp 177–220, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.

Berg, O. G., Winter, R. B., & von Hippel, P. H. (1982) *Trends Biochem. Sci. 7*, 52.

Betz, J. L. (1986) *Gene 42*, 283–292.

Betz, J. L. (1987) *J. Mol. Biol. 195*, 495–504.

Boelens, R., Scheek, R. M., van Boom, J. H., & Kaptein, R. (1987) *J. Mol. Biol. 193*, 213–234.

Buck, F., Ruterjans, H., & Beyreuther, K. (1978) *FEBS Lett. 78*, 335–338.

Buck, F., Ruterjans, H., Kaptein, R., & Beyreuther, K. (1980) *Proc. Natl. Acad. Sci. U.S.A. 77*, 5145–5148.

Buck, F., Hahn, K. D., Zemann, W., Ruterjans, H., Sadler, J. R., Beyreuther, K., Kaptein, R., & Scheek, R. (1983) *Eur. J. Biochem. 132*, 321–327.

Buck, F., Hahn, K. D., Brill, W., Ruterjans, H., Chernov, B. K., Skryabin, K. G., Kirpichnikov, M. P., & Bayev, A. A. (1986) *J. Biomol. Struct. Dyn. 3*, 899–911.

Calladine, C. R. (1982) *J. Mol. Biol. 161*, 343–352.

Caruthers, M. H. (1980) *Acc. Chem. Res. 13*, 155–160.

Connolly, B. A., & Eckstein, F. (1984) *Biochemistry 23*, 5523–5527.

Costello, A. J. R., Glonek, T., & Van Wazer, J. R. (1976) *Inorg. Chem. 15*, 972–974.

deHaseth, P. L., Lohman, T. M., & Record, M. T., Jr. (1977) *Biochemistry 16*, 4783–4790.

Dickerson, R. E. (1983) *J. Mol. Biol. 166*, 419–441.

Dickerson, R. E., & Drew, H. R. (1981) *J. Mol. Biol. 149*, 761–786.

Ebright, R. H. (1986) in *Protein Structure Folding and Design* (Oxender, D., Ed.) pp 207–219, Alan R. Liss, New York.

Eisenbeis, S. J., Nasoff, M. S., Noble, S. A., Bracco, L. P., Dodds, D. R., & Caruthers, M. H. (1985) *Proc. Natl. Acad. Sci. U.S.A. 82*, 1084–1088.

Frechet, D., Cheng, D. M., Kan, L.-S. & Ts'o, P. O. P. (1983) *Biochemistry 22*, 5194–5200.

Fu, J. M., Schroeder, S. A., Jones, C. R., Santini, R., & Gorenstein, D. G. (1988) *J. Magn. Reson. 77*, 577–582.

Giessner-Prettre, C., Pullman, C., Borer, B., Kan, L. S., & Ts'o, P. O. P. (1976) *Biopolymers 15*, 2277.

Gilbert, W., Gralla, J., Major, J., & Maxam, A. (1975) *Protein–Ligand Interactions*, pp 270–288, de Gruyter, Berlin, FRG.

Gorenstein, D. G. (1978) in *Jerusalem Symposium, NMR in Molecular Biology* (Pullman, B., Ed.) pp 1–15, D. Reidel Publishing Co., Dordrecht, The Netherlands.

Gorenstein, D. G. (1981) *Annu. Rev. Biophys. Bioeng. 10*, 355.

Gorenstein, D. G. (1983) *Prog. Nucl. Magn. Reson. Spectrosc. 16*, 1–98.

Gorenstein, D. G. (1984) in *Phosphorus-31 NMR: Principles and Applications* (Gorenstein, D. G., Ed.) Academic Press, New York.

Gorenstein, D. G. (1987) *Chem. Rev. 87*, 1047–1077.

Gorenstein, D. G., & Kar, D. (1975) *Biochem. Biophys. Res. Commun. 65*, 1073–1080.

Gorenstein, D. G., & Goldfield, E. M. (1984) in *Phosphorus-31 NMR: Principles and Applications* (Gorenstein, D. G., Ed.) Academic Press, New York.

Gorenstein, D. G., Findlay, J. B., Momii, R. K., Luxon, B. A., & Kar, D. (1976) *Biochemistry 15*, 3796–3803.

Gorenstein, D. G., Goldfield, E. M., Chen, R., Kovar, K., & Luxon, B. A. (1981) *Biochemistry 20*, 2141–2150.

Gorenstein, D. G., Schroeder, S., Miyasaki, M., Fu, J., Jones, C., Roongta, V., & Abuaf, P. (1987) *Proc. Int. Conf. Phosphorus Chem., 10th 30*, 567–570.

Gorenstein, D. G., Schroeder, S. A., Fu, J. M., Metz, J. T., Roongta, V. A., & Jones, C. R. (1988) *Biochemistry 27*, 7223–7237.

Gorenstein, D. G., Meadows, R. P., Metz, J. T., Nikonowicz, E., & Post, C. P. (1990) in *Advances in Biophysical Chemistry* (Bush, C. A., Ed.) JAI Press, Greenwich, CT (in press).

Groeddel, D. V., Yanasura, D. G., & Caruthers, M. H. (1978) *Proc. Natl. Acad. Sci. U.S.A. 75*, 3578–3582.

Helene, C., & Lancelot, G. (1982) *Prog. Biophys. Mol. Biol. 39*, 1–68.

Hogan, M., Wemmer, D., Bray, R. P., Wade-Jardetzky, N., & Jardetzky, O. (1981) *FEBS Lett. 124*, 202–203.

Jordan, S. R., & Pabo, C. O. (1988) *Science 242*, 893–899.

Koudelka, G. B., Harrison, S. C., & Ptashne, M. (1987) *Nature 326*, 886.

Lai, K., & Gorenstein, D. G. (1989) *Biochemistry 28*, 2804–2812.

Lai, K., Shah, D. O., Derose, E., & Gorenstein, D. G. (1984) *Biochem. Biophys. Res. Commun. 121*, 1021.

Lamerichs, R. M. J. N., Boelens, R., van der Marel, G. A., van Boom, J. H., Kaptein, R., Buck, F., Fera, B., & Ruterjans, H. (1989) *Biochemistry 28*, 2985–2991.

Landschulz, W. H., Johnson, P. E., & McKnight, S. L. (1988) *Science 240*, 1759–1764.

Lehming, N., Sartorius, J., Niemoller, M., Genenger, G., von Wilcken-Bergmann, B., & Muller-Hill, B. (1987) *EMBO J. 6*, 3145–3153.

Lehming, N., Sartorius, J., Oehler, S., von Wilcken-Bergmann, B., & Muller-Hill, B. (1988) *Proc. Natl. Acad. Sci. U.S.A. 85*, 7947–7951.

Lerner, D. B., & Kearns, D. R. (1980) *J. Am. Chem. Soc. 102*, 7612–7613.

Maniatis, T., Fritsch, E. F., & Sambrook, J. (1982) *Molecular Cloning: A Laboratory Manual*, pp 466–467, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.

Matthews, K. S. (1979) *J. Biol. Chem. 254*, 3348–3353.

Miller, J. H., & Reznikoff, W. S. (1978) *The Operon*, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.

Nick, H., Arndt, K., Boschelli, F., Jarema, M.-A. C., Lillis, M., Sadler, J. R., Caruthers, M., & Lu, P. (1982) *Proc. Natl. Acad. Sci. U.S.A. 79*, 218–222.

Nikonowicz, E., Meadows, R., Post, C., Jones, C., & Gorenstein, D. G. (1989a) *Bull. Magn. Reson. 11*, 226–229.

Nikonowicz, E., Roongta, V., Jones, C. R., & Gorenstein, D. G. (1989b) *Biochemistry 28*, 8714–8725.

Nilges, M., Clore, G. M., Gronenborn, A. M., Brunger, A. T., Karplus, M., & Nilsson, L. (1987) *Biochemistry 26*, 3718–3733.

Noble, S. A., Fisher, E. F., & Caruthers, M. H. (1984) *Nucleic Acids Res. 12*, 3387–3404.

Nussinov, R. (1986a) *Biochim. Biophys. Acta 866*, 93.

Nussinov, R. (1986b) *J. Biomol. Struct. Dyn. 3*, 1145.

Ott, J., & Eckstein, F. (1985a) *Biochemistry 24*, 253.

Ott, J., & Eckstein, F. (1985b) *Nucleic Acids Res. 13*, 6317–6330.

Otwinowski, Z., Schevitz, R. W., Zhang, R. G., Lawson, C. L., Joachimiak, A., Marmorstein, R. Q., Luisi, B. F., & Sigler, P. B. (1988) *Nature 335*, 321.

Patel, D. J. (1974) *Biochemistry 13*, 2396–2402.

Patel, D. J. (1979) *Acc. Chem. Res. 12*, 118–125.

Petersheim, M., Mehdi, S., & Gerlt, J. A. (1984) *J. Am. Chem. Soc. 106*, 439–440.

Pilz, I., Goral, K., Kratky, O., Bary, R. P., Wade-Jarketzky, N., & Jarketzky, O. (1980) *Biochemistry 19*, 4087–4090.

Powers, R., Jones, C. R., & Gorenstein, D. G. (1990) *J. Biomol. Struct. Dyn.* (submitted for publication).

Ptashne, M. (1986) *A Genetic Switch*, Cell Press, Cambridge, MA.

Record, M. T., Jr., deHaseth, P. L., & Lohman, T. M. (1977) *Biochemistry 16*, 4791–4796.

Reinhardt, C. G., & Krugh, T. R. (1977) *Biochemistry 16*, 2890.

Riggs, A. D., Suzuki, H., & Bourgeois, S. (1970) *J. Mol. Biol. 48*, 67–83.

Roongta, V. A., Jones, C. R., & Gorenstein, D. G. (1990) *Biochemistry 29*, 5245–5258.

Sadler, J. R., Sasmor, H., & Betz, J. L. (1983) *Proc. Natl. Acad. Sci. U.S.A. 80*, 6785–6789.

Saenger, W. (1984) *Principles of Nucleic Acid Structure*, Springer-Verlag, New York.

Scheek, R. M., Zuiderweg, E. R. P., Klappe, K. J. M, van Boom, J. H., Kaptein, R., Ruterjans, H., & Beyreuther, K. (1983) *Biochemistry 22*, 228–235.

Schroeder, S., Jones, C., Fu, J., & Gorenstein, D. G. (1986) *Bull. Magn. Reson. 8*, 137–146.

Schroeder, S. A., Fu, J. M., Jones, C. R., & Gorenstein, D. G. (1987) *Biochemistry 26*, 3812–3821.

Schroeder, S. A., Roongta, V., Fu, J. M., Jones, C. R., & Gorenstein, D. G. (1989) *Biochemistry 28*, 8292–8303.

Seeman, N. C., Rosenberg, J. M., Suddath, F. L., Park Kim, J. J., & Rich, A. (1976) *J. Mol. Biol. 104*, 142–143.

Shah, D. O., Lai, K., & Gorenstein, D. G. (1984a) *Biochemistry 23*, 6717–6723.

Shah, D. O., Lai, K., & Gorenstein, D. G. (1984b) *J. Am. Chem. Soc. 106*, 4302.

Simons, A., Tils, D., von Wilcken-Bergmann, B., & Muller-Hill, B. (1984) *Proc. Natl. Acad. Sci. U.S.A. 81*, 1624–1628.

Sklenář, V., Miyashiro, H., Zon, G., Miles, H. T., & Bax, A. (1986) *FEBS Lett. 208*, 94–98.

Takeda, Y., Ohlendorf, D. H., Anderson, W. F., & Matthews, B. W. (1983) *Science 221*, 1020–1026.

Wade-Jardetzky, N., Bray, R. P., Conover, W. W., Jardetzky, O., Geisler, N., & Weber, K. (1979) *J. Mol. Biol. 128*, 259–264.

Wang, J. C., Barkley, M. D., & Bourgeois, S. (1974) *Nature 251*, 247.

Wemmer, D., & Kallenbach, N. R. (1983) *Biochemistry 22*, 1901.

Wharton, R. P., & Ptashne, M. (1985) *Nature 316*, 601–605.

Wharton, R. P., & Ptashne, M. (1987) *Nature 326*, 888.

Wharton, R. P., Brown, E. L., & Ptashne, M. (1984) *Cell 38*, 361–369.

Whitson, P. A., Olson, J. S., & Matthews, K. S. (1986) *Biochemistry 25*, 3852–3858.

Winter, R. B., & von Hippel, P. H. (1981) *Biochemistry 20*, 6948–6960.

Wolberger, C., Dong, Y., Ptashne, M., & Harrison, S. C. (1988) *Nature 335*, 789–795.

Zuiderweg, E., Scheek, R., Boelens, R., Gunsteren, W., & Kaptein, R. (1985) *Biochimie 67*, 707–715.