An Alkaline Phosphatase Protection Assay To Investigate *trp* Repressor/Operator Interactions†

Ronen Q. Marmorstein,‡,§,‖ Mathias Sprinzl,⊥ and Paul B. Sigler*,‡

Department of Biochemistry and Molecular Biology and Department of Chemistry, The University of Chicago, Chicago, Illinois 60637, and Department of Molecular Biophysics and Biochemistry and the Howard Hughes Medical Institute, Yale University, New Haven, Connecticut 06511

Received December 21, 1989; Revised Manuscript Received May 3, 1990

ABSTRACT: We have used an alkaline phosphatase protection assay to investigate the interaction of the *trp* repressor with its operator sequence. The assay is based on the principle that the *trp* repressor will protect a terminally 5′-^{32}P-labeled operator DNA fragment from attack by alkaline phosphatase. The optimal oligonucleotide for investigating the *trp* repressor/operator interaction extends two base pairs from each end of the genetically defined target sequence predicted by in vivo studies [Bass et al. (1987) *Genes Dev. 1*, 565–572]. The assay works well over a 10000-fold range of protein/DNA affinity and is used to show that the corepressor, L-tryptophan, causes the liganded repressor to bind a 20 base pair *trp* operator duplex 6400 times more strongly than the unliganded aporepressor. The affinity of the *trp* repressor for operators containing symmetrical mutations was interpreted in terms of the *trp* repressor/operator crystal structure as follows: (1) Direct hydrogen bonds with the functional groups of G_{-9} of the *trp* operator and the side chain of Arg 69 of the *trp* repressor contribute to DNA-binding specificity. (2) G_{-6} of the *trp* operator is critical for DNA-binding specificity probably because of the two water-mediated hydrogen bonds between its functional groups and the N-terminus of the *trp* repressor's E-helix. (3) Sequence-dependent aspects of the *trp* operator's conformation help stabilize the *trp* repressor/operator complex. The operator-binding activity of mutationally altered *trp* aporepressor proteins when interpreted in terms of the relevant crystal structures showed the following: (1) The small side chain of alanine at position 77, in the turn of the helix-turn-helix, helps the unliganded aporepressor's DNA-interacting domain to collapse into an inactive conformation. (2) Unlike the terminal segments of some related repressors, the extended N-terminal arms of the *trp* repressor do not interact with the DNA in a sequence-specific way. The assay demonstrated the preferential affinity of the *gal* repressor for a 20 base pair *gal* operator, suggesting that alkaline phosphatase protection may be generally useful in measuring the affinity of DNA-binding proteins for their target DNA sequences.

The *trp* aporepressor of *Escherichia coli* exists as a homodimer of two 107 amino acid chains and is activated to the *trp* repressor by the noncooperative binding of two molecules of the corepressor, L-tryptophan (Joachimiak et al., 1983). Once activated, the *trp* repressor binds with specificity and high affinity (Joachimiak et al., 1983; Klig et al., 1987; Carey, 1988) to the operators of three operons (*trpEDCBA*, *trpR*, and *aroH*), preventing initiation of transcription (Somerville, 1983).

The *trp* repressor/operator crystal structure (Otwinowski et al., 1988) is one of several protein/DNA crystalline complexes that have been described (Frederick et al., 1984; Anderson et al., 1987; Aggarwal et al., 1988; Jordan & Pabo, 1988; Wolberger et al., 1988). In addition, there are several other protein/DNA cocrystals reported in the literature [examples given by Brennan et al. (1986) and Chandrasegaran et al. (1986)] and more cocrystallization attempts are under way. In addition, spectroscopic techniques have been exploited to investigate several protein/DNA interactions [example given by Thomas et al. (1989)]. Such biophysical studies demand that the DNA component of the complexes be represented by duplex fragments that are short but still maintain a high affinity for the protein of interest. We have used the *trp* repressor/operator system to develop a quick and convenient alkaline phosphatase protection assay by which such simulated DNA target sites can be identified. Once the appropriate oligonucleotide was identified, we show that the assay can be used conveniently and quantitatively over a wide range of protein/DNA affinities to investigate the interaction of wild-type and mutationally altered *trpR* proteins with wild-type and mutationally altered *trp* operator target sites.

Materials and Methods

Proteins

The *trpR* protein and its mutational variants were overproduced and purified as described (Joachimiak et al., 1987; Marmorstein, 1989). All preparations were determined to be homogeneous by electrophoresis in polyacrylamide gels containing sodium dodecyl sulfate and stained with silver. The α-chymotrypsin-treated *trp* repressor protein, des7, and *gal* repressor were gifts from Jannette Carey and Sankar Adhya, respectively. The concentration of *trp* repressor and its mutational variants was determined by absorbance at 280 nm (ϵ_{280}

† This work was supported by grants to P.B.S. from USPHS (GM15225) and The American Cancer Society (NP-444) and grants to P.B.S. and M.S. from NATO.

* To whom correspondence should be addressed at Department of Molecular Biophysics & Biochemistry, Yale University/Howard Hughes Medical Institute, 260 Whitney Ave./JWG 402, New Haven, CT 06511.

‡ Yale University.

§ Department of Chemistry, The University of Chicago.

‖ Present Address: Department of Biochemistry and Molecular Biology, Harvard University, 7 Divinity Ave., Cambridge, MA 02138.

⊥ Permanent address: Laboratorium für Biochemie der Universität Bayreuth, Postfach 10 12 51, D-8580 Bayreuth, Federal Republic of Germany.

0006-2960/91/0430-1141$02.50/0

Table I: Fragments That Simulate the *trp* Operator Used in the Alkaline Phosphatase Protection Assay

% protection[a]	operator sequence	
	-10 -5 -1+1 +5 +10	
E. coli. trpEDCBA operator	CGAACTAGTTAACTAGTAGC GCTTGATCAATTGATCATCG	
0	ACTAGTTAACTAGT TGATCAATTGATCA	(a)
0	GTACTAGTTAACTAGTAC CATGATCAATTGATCATG	(b)
16	TGTACTAGTTAACTAGTAC CATGATCAATTGATCATGT	(c)
35	GGTACTAGTTAACTAGTAC CATGATCAATTGATCATGG	(d)
10	TTGTACTAGTTAACTAGTAC CATGATCAATTGATCATGTT	(e)
55	CGTACTAGTTAACTAGTACG GCATGATCAATTGATCATGC	(f)
4.5	TCGTACTAGTTAACTAGTACG GCATGATCAATTGATCATGCT	(g)

[a] Refer to Materials and Methods.

Table II: Mutant *trp* Operators Used in the Alkaline Phosphatase Protection Assay

operator[a]	sequence
	-10 -5 -1+1 +5 +10
wild type	CGTACTAGTTAACTAGTACG GCATGATCAATTGATCATGC
CT-6	CGTA**T**TAGTTAACTA**A**TACG GCAT**A**ATCAATTGAT**T**ATGC
GT-9	C**T**TACTAGTTAACTAGTA**A**G G**A**ATGATCAATTGATCAT**T**C
TA-2	CGTACTAG**A**TA**T**CTAGTACG GCATGATC**T**AT**A**GATCATGC
gal	CG**GTG**TA**AACGTT**TA**CAC**CG GC**CAC**AT**TTGCAA**AT**GTG**GC

[a] Mutant operators are named as two letters followed by a number. The first letter is the one-letter abbreviation for the wild-type base; the second letter is the one-letter abbreviation for the replacement; the number is the base position. Base positions that differ from the idealized *trp* operator are indicated in bold face type.

= 1.48×10^4 M^{-1} cm^{-1}) (Joachimiak et al., 1983). The concentration of *gal* repressor was approximated to be 1.0 mg/mL at an absorbance of 1.0 at 280 nm (Scopes, 1987). Before use, all proteins were dialyzed against 250 mM NaCl and 10 mM Tris·HCl, pH 7.4, and diluted to 2.5–70 μM dimer in 250 mM NaCl, 10 mM Tris·HCl, pH 7.4, and 100 μg/mL BSA.

DNA Fragments

The simulated *trp* operators are shown in Table I. Crude oligonucleotides a, b, c, and f in Table I were gifts from M. Haden, W. Mandecki, and J. L. Fox of Abbott laboratories (Abbott Park, IL), and oligonucleotides d, e, and g in Table I were purchased from the University of Pennsylvania (Philadelphia, PA). These oligonucleotides were purified by reverse-phase HPLC as described (Joachimiak et al., 1987).

Table II shows the mutant *trp* operators whose binding to the *trp* repressor was investigated with the alkaline phosphatase protection assay. These oligonucleotides were purified in the same way as the simulated *trp* operators except that a NaCl gradient (0.54–0.62 M), in 10 mM NaOH, pH 12.0, and a MonoQ FPLC column (Pharmacia; Piscataway, NJ) was used in place of an acetonitrile gradient (5–12%) on a reverse-phase HPLC column in the last chromatographic step. The peak containing pure oligonucleotide was collected and adjusted to pH 7.4 with 100 mM cacodylic acid and desalted through a reverse-phase C-18 Sep-Pack cartridge (Waters; Milford, MA) (Lo et al., 1984). The applied oligonucleotide was washed with 20 mM triethylammonium bicarbonate (TEAB), pH 7.4, and eluted with 30% acetonitrile in 100 mM TEAB, pH 7.4. The oligonucleotides were freeze-dried, redissolved in water, and freeze-dried again to remove all traces of volatile salt.

Each oligonucleotide, 40 ± 5 μg, was 5′-^{32}P labeled with T4 polynucleotide kinase (Boehringer Mannheim; Indianapolis, IN) and [^{32}P]dATP (Amersham; Arlington Heights, IL) by scaling up the procedure of Maxam and Gilbert (1977). Initially, the 5′-phosphorylated oligonucleotides were separated from the unlabeled fragments by electrophoresis on 40-cm 25% polyacrylamide gels in 7 M urea. The largest molecular weight radioactive bands were cut out of the gels and the oligonucleotides extracted with 0.5 M ammonium acetate and 0.1 mM EDTA, pH 7.0. Since the background of unphosphorylated DNA fragments does not materially affect the reaction, this separation proved to be unnecessary and was omitted in the later experiments. The oligonucleotides, either extracted from the polynucleotide gel or heat treated for 2 min at 90 °C (to inactivate the T4 polynucleotide kinase) following the kinase reaction, were passed through a reverse-phase C-18 Sep-Pack cartridge and desalted and dried as described earlier. The dried oligonucleotides were dissolved in 250 mM KCl and 10 mM Tris·HCl, pH 7.4, and renatured at a duplex concentration of about 0.5 mM by heating to 80 °C for 5 min and slowly cooling to 5 °C. The oligonucleotides were finally diluted to about 500 nM duplex with 250 mM NaCl and 10 mM Tris·HCl, pH 7.4, and stored at 4 °C. Duplex oligonucleotide concentrations were determined at 20 °C from their absorbance spectra. Extinction coefficients of oligonucleotide duplexes were calculated from their constituent bases (Pabst Laboratories Circular, 1956) and corrected for their hypochromic effect. The molar extinction coefficients for native oligonucleotide duplexes a–g are as follows: $\epsilon(a)_{260} = 2.62 \times 10^5$, $\epsilon(b)_{260} = 3.04 \times 10^5$, $\epsilon(c)_{260} = 3.11 \times 10^5$, $\epsilon(d)_{260} = 3.12 \times 10^5$, $\epsilon(e,f)_{260} = 3.32 \times 10^5$, and $\epsilon(g)_{260} = 3.46 \times 10^5$ M^{-1} cm^{-1} (oligonucleotides a–g are identified in Table I).

Alkaline Phosphatase Protection Assay

Alkaline phosphatase from calf intestine (Boehringer Mannheim) will liberate ^{32}P from a 5′-^{32}P end-labeled oligonucleotide (Charonas & Sande, 1980; Maxam & Gilbert, 1980). The binding of the *trp* repressor or a varient thereof to a ^{32}P-labeled oligonucleotide duplex of the appropriate length presumably protects against enzymatic removal of ^{32}P. Trichloroacetic acid (TCA) precipitation of the oligonucleotide followed by filtration separates the liberated ^{32}P from the ^{32}P-labeled oligonucleotide. Therefore, measuring the residual activity in the precipitate provides a convenient method to track the rate of hydrolysis and hence the degree of protection. A quantitative treatment that relates this degree of protection from hydrolysis to the K_D of the ^{32}P-labeled oligo-

nucleotide/protein complex is developed under K_D determination.

Defining the Length of the trp Operator. Forty micromolar (1.12 μg/mL) *trp* repressor and 1 mM L-tryptophan (the corepressor) were incubated in a 100-μL solution (250 mM KCl, 10 mM Tris·HCl, pH 7.4) that was 2 μM in one of the duplexes in Table I that had been labeled at both 5'-termini with ^{32}P. After 30 min at 22 °C, digestion of the oligonucleotide was initiated by adding 3 units of calf intestine alkaline phosphatase (Boehringer Mannheim; Indianapolis, IN) (3 μL from a 1 unit/μL stock stored in 3 M NaCl, 3 mM $MgCl_2$, 0.1 mM $ZnCl_2$, and 30 mM triethylamine, pH 7.6). After 2 min, alkaline phosphatase activity was stopped with 0.15 M KH_2PO_4/$KHPO_4$, pH 7.0, and the ^{32}P-labeled oligonucleotide was precipitated by adding 1 volume of 20 mg/mL bulk tRNA followed by 2 mL of 5% TCA and incubated for 5 min at room temperature. The TCA-treated incubation mixture was applied to a nitrocellulose membrane (Millipore Corporation, 45 μm) and filtered. The reaction tube was washed with an additional 2 mL of 5% TCA and filtered through the same nitrocellulose membrane. The nitrocellulose membrane was then dried and the precipitated ^{32}P-labeled oligonucleotide was counted in toluene-based scintillation fluid. Table I shows that the blunt-ended 20 base pair fragment (f in Table I) was the oligonucleotide most effectively protected from alkaline phosphatase digestion by the binding of *trp* repressor. All further binding studies were done with 20 base pair fragments.

Alkaline Phosphatase Mediated Digestion of 20 Base Pair 5'-^{32}P-Labeled trp Operators. Twenty nanomolar oligonucleotide duplex was digested at 22 °C with 2.5–10 units of calf intestine alkaline phosphatase (previously dialyzed against 250 mM NaCl and 10 mM Tris·HCl, pH 7.4) in a 200-μL reaction volume comprised of 250 mM NaCl, 100 μg/mL BSA, and 10 mM Tris·HCl, pH 7.4. Aliquots were removed at prescribed times and the reaction was quenched, DNA precipitated, and the filter-bound DNA counted as described above.

To determine the dissociation constant of the *trp* repressor/operator complex, the procedure described above was carried out in the presence of 500 nM *trp* repressor (25-fold excess over operator) and 0.5 mM L-tryptophan. In addition, dissociation constants were also determined for the unliganded *trp* aporepressor/operator complex and the *trp* repressor/operator complex formed in the presence of 0.5 mM indole-3-propionic acid in place of L-tryptophan.

Dissociation constants of *trp* repressor/operator complexes with symmetrical operator mutations were determined by carrying out digestions of each of the altered operators in the presence or absence of *trp* repressor as described above. Dissociation constants of complexes containing mutationally altered *trp* repressor and a wild-type 20 base pair *trp* operator were also determined by the procedure described in the paragraph above.

K_D Determination

The derivation that follows is based on Segel's (1975) general equation for the case of "inhibition by substrate depletion". In our specific case the substrate is a ^{32}P-labeled oligonucleotide and the inhibitor is the *trp* repressor. ^{32}P-Labeled oligonucleotide (Np) in the presence of alkaline phosphatase (AP) and *trp* repressor (R) can undergo either of two reactions: (a) hydrolysis by alkaline phosphatase yielding the dephosphorylated oligonucleotide (N) and inorganic phosphate ([^{32}P]P_i) or (b) complex formation with the *trp* repressor (R·Np) in which the phosphorylated oligonucleotide becomes protected from enzymatic hydrolysis. These reactions are schematized below:

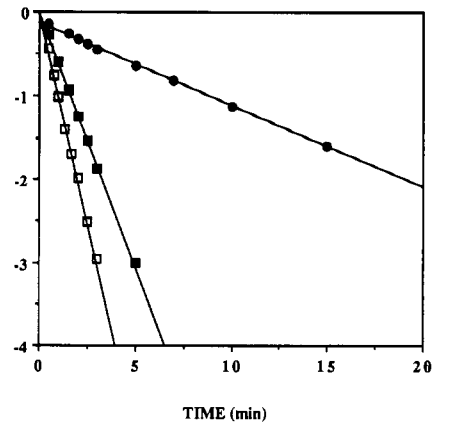
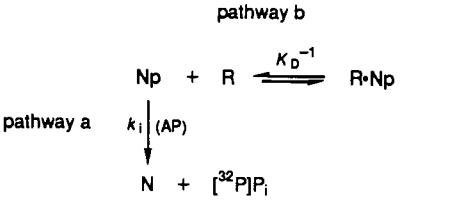


FIGURE 1: Effect of operator sequence on the rate of alkaline phosphatase digestion in the presence of *trp* repressor. The natural logarithm of the ratio of precipitated, (CPM), to total radioactivity, $(CPM)_0$, is plotted as a function of time for the *trp* operator mutants CT-6 (■) and GT-9 (●). The digestion of the *trp* operator mutant CT-6 (□), in the absence of protein, is shown for reference (the rate of digestion of the GT-9 mutant is similar). The experimental conditions are described under Materials and Methods.



where k_i is the rate constant for dephosphorylation catalyzed by alkaline phosphatase and K_D is the dissociation constant of the repressor/^{32}P-labeled oligonucleotide complex. Since first-order kinetics applies (Figure 1), the concentration of ^{32}P-labeled oligonucleotide at a given time, t, can be described by eq 1 where [Np] is the concentration of free ^{32}P-labeled

$$\ln [Np] = -k_i[AP]t + \ln [Np]_{T,0} \quad (1)$$

oligonucleotide at time t and $[Np]_{T,0}$ is the concentration of free ^{32}P-labeled oligonucleotide at zero time when no repressor is present. This represents the concentration of total ^{32}P-labeled oligonucleotide used in the incubation.

In the presence of *trp* repressor (R) the rate of dephosphorylation will be lower due to R·Np complex formation, which will decrease the concentration of free Np available for enzymatic hydrolysis. Combining the expression for the dissociation of [R·Np] and the mass conservation equation for the total amount of ^{32}P-labeled oligonucleotide normalized to a given volume leads to eq 2 where $[Np]_T$ is the total con-

$$[Np] = [Np]_T/(1 + [R]/K_D) \quad (2)$$

centration of free and repressor-bound ^{32}P-labeled oligonucleotide at a given time, t. Assuming that the equilibrium value of $[N_p]$ is continually reestablished, i.e., that the system is at "steady state", then substituting [Np] from eq 2 into eq 1 followed by integration gives

$$\ln [Np]_T = -k_i[AP]t(1/(1 + [R]/K_D)) + \ln [N_p]_{T,0} \quad (3)$$

Since the fraction of ^{32}P-labeled oligonucleotide remaining after alkaline phosphatase digestion is equal to the fraction of filter-retained radioactivity at time t

$$[Np]_T/[Np]_{T,0} = (CPM)/(CPM)_0$$

where (CPM) is the radioactivity per unit volume retained on

the filter after filtration and $(CPM)_0$ is the radioactivity per unit volume retained at zero time. Therefore in the absence of repressor eq 1 can be rewritten to give

$$\ln (CPM) = -k_i[AP]t + \ln (CPM)_0 \quad (4)$$

and in the presence of repressor eq 3 can be rewritten as

$$\ln (CPM) = -k_i'[AP]t + \ln (CPM)_0 \quad (5)$$

where k_i' is the apparent rate constant for dephosphorylation in the presence of repressor:

$$k_i' = k_i/(1 + [R]/K_D)$$

K_D can be solved for explicitly from eq 5 to give

$$K_D = [R]/(k_i/k_i' - 1) \quad (6)$$

Filter-retained radioactivity, (CPM), is measured as a function of time in the absence and presence of protein, R, and constants $k_i[AP]$, $k_i'[AP]$, and $(CPM)_0$ are adjusted to optimize the fit of the experimental points to eq 4 and 5 by employing a linear least-squares fitting program (Cricket Graph, Apple Computer, Inc.; Cupertino, CA). Equation 6 is then used to calculate the K_D of the respective protein/DNA complex. $(CPM)_0$ is left as a parameter to correct for the small fraction of DNA, which may not be in the duplex conformation and, therefore, could not be bound by repressor. Adjusted values for $(CPM)_0$ were all within 5% of the values determined analytically.

The *trp* repressor can form complexes with dephosphorylated oligonucleotide (N), as well as with phosphorylated oligonucleotide (Np). Therefore, the mass balance equation for repressor becomes

$$[R]_0 = [R] + [R{\cdot}Np] + [R{\cdot}N]$$

However, since the total concentration of repressor, $[R]_0$, is in large excess with respect to the total concentration of ^{32}P-labeled oligonucleotide, $[Np]_{T,0}$ (200 nM to 10 μM versus 20 nM, respectively), formation of R·N and R·Np complexes will not significantly deplete the concentration of total repressor, $[R]_0$, that is

$$[R] = [R]_0 - ([R{\cdot}Np] + [R{\cdot}N]) \simeq [R]_0$$

Results and Discussion

trp Repressor Binding to Operators of Various Length

Table I shows that a 2 μM solution of either the 14 or 18 base pair fragments is not protected from 5′-^{32}P dephosphorylation, even at a *trp* repressor concentration of 40 μM. These operator fragments were also found to bind weakly to *trp* repressor ($K_D \geq 10^{-6}$ M) by a filter-binding assay (Wagner et al., manuscript in preparation). A minimum of 19 nucleotides and 18 base pairs is required to produce significant protection, while a blunt 20 base pair duplex is best. Extending the duplex by just one more nucleotide apparently places the phosphate outside the range of the repressor's full protection. Bass et al. (1987) showed that the principal sequence determinants of specific binding in vivo reside in the central 14 base pairs, suggesting that the three symmetrically flanking base pairs impart stability to the complex without specificity. The 14 base pair operator is not sufficiently long to support repressor binding, probably because such a complex would make only 14 of the 24 amino acid/phosphate oxygen contacts found in the contact surface of the crystalline complex (Otwinowski et al., 1988). The 18 base pair operator was also a poor substrate for the *trp* repressor even though both in vivo studies (Bass et al., 1987) and the crystal structure of the complex (Otwinowski et al., 1988) showed that all the *trp* repressor/operator contacts occur within nine base pairs from the central dyad. It is important, however, to stress that the sequence elements responsible for discriminating the operator are not the only contributions to binding. The phosphates of the extra flanking nucleotides may enhance affinity by augmenting the electrostatic component of the interaction energy. Table I shows that the 19, 20, and 21 base pair fragments all bind more tightly to the repressor than the 18 base pair fragment.

Table III: Effect of Ligation on the Dissociation Constant of the *trp* Repressor/20 Base Pair Operator Complex

ligand	$k_i[AP]^a$ (min^{-1})	$k_i'[AP]^a$ (min^{-1})	$K_D{}^a$ (M)
L-tryptophan	1.43	0.0139	2.00×10^{-9}
indole-3-propionate	0.528	0.470	2.06×10^{-5}
no ligand	0.528	0.441	1.28×10^{-5}

[a] The calculation of $k_i[AP]$, $k_i'[AP]$, and K_D is described under Materials and Methods. All values are accurate to within 5%.

Dissociation Constants of Aporepressor/Operator in the Presence of Various Ligands

Table III shows the dissociation constants for the 20 base pair fragment (Table I, f) complexed to the *trp* repressor, aporepressor, and the "pseudorepressor" (inactive adduct formed by binding indole-3-propionate in place of L-tryptophan). We found that converting aporepressor to repressor by binding L-tryptophan increases the affinity of the protein for operator DNA by a factor of 6400. In contrast, converting aporepressor to pseudorepressor by binding the "inducer" indole-3-propionate decreases the affinity of the protein for operator DNA by about 2-fold. The K_D of 2 nM calculated for the repressor/operator complex is in fair agreement with the values determined by filter-binding techniques. The value reported by Klig and co-workers (1987) is 3 nM and that reported by Marmorstein and Sigler (1989) is 5.9 nM. The K_D for the repressor/operator complex is, however, four times higher than the value of 0.5 nM reported by Carey (1988) using a gel shift assay at a pH of 6.0. It is possible that at the lower pH used in the gel shift assay the protein may gain sufficient positive charge to produce a 4-fold increase in electrostatic attraction.

trp Repressor Binding to trp Operators with Symmetrical Mutations

Bass et al. (1987) have used an in vivo "challenge assay" (Benson et al., 1986) to study *trp* repressor binding to operators with symmetrical base pair changes. Their experiments show which base pairs have the most influence on the in vivo function of the operator as a negative regulatory target for repressor in a P22 *ant* promoter construct. Presumably this assay depends on the equilibrium affinity of the protein for the operator sequence. They found that the five nucleotides, 5′-ACTAG-3′, at positions −7 through −3 comprise the most important region for stabilizing the *trp* repressor/operator complex in vivo. In order to quantitatively assess under controlled in vitro conditions the functional significance of some of the interactions observed in the crystalline *trp* repressor/operator complex, we developed the alkaline phosphatase protection assay to study *trp* repressor binding to operators with symmetrical mutations.

Table IV shows that, of the mutant *trp* operators investigated, the symmetrical C to T change at position −6 is most detrimental. The *trp* repressor's affinity for this oligonucleotide is roughly the same as its affinity for the unrelated *gal* operator (Tables II and IV). The importance of a C·G base pair at position −6 is corroborated by the in vivo study of Bass et al. (1987), which showed that any symmetrical substitution for $(C{\cdot}G)_{-6}$ severely disrupts operator binding. Figure 2A, which shows a base pair step from the crystalline repressor/operator

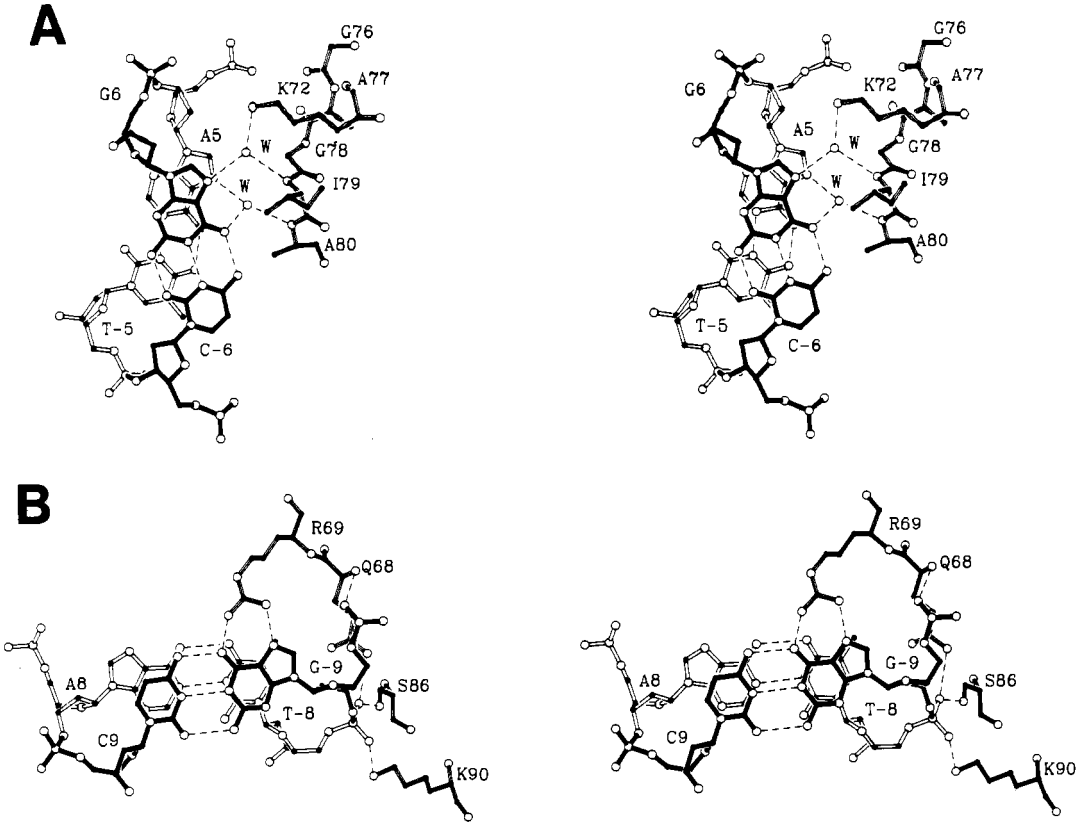


FIGURE 2: Structural details of the *trp* repressor/operator interface. Two base pair steps are shown. Only protein residues (on the right) in the proximity of the respective DNA step are shown for clarity. The bonds of the nucleotide closest to the viewer are filled; the protein's bonds are stippled. Carbon atoms are small filled spheres; oxygen, nitrogen, and phosphorus atoms are large open spheres; hydrogen bonds (2.6–3.1 Å) are dashed. A signifies both adenosine and alanine, G is for both guanosine and glycine, and W labels water molecules. (A) Protein/DNA interactions with A5 and G6; (B) protein/DNA interactions with G-9. Reprinted with permission from *Nature*. Copyright 1988 Macmillan Magazines Ltd.

Table IV: Effect of Operator Sequence on the Dissociation Constant of the *trp* Repressor/20 Base Pair Operator Complex

operator[a]	k_i[AP][b] (min^{-1})	k_i'[AP][b] (min^{-1})	K_D[b] (M)
wild type	1.43	0.0139	2.00×10^{-9}
CT-6	1.00	0.609	1.55×10^{-5}
GT-9	0.801	0.0987	1.41×10^{-6}
TA-2	0.757	0.0359	1.24×10^{-7}
gal	0.394	0.232	1.43×10^{-5}

[a] Mutant operators are named as two letters followed by a number. The first letter is the one-letter abbreviation for the wild-type base; the second letter is the one-letter abbreviation for the replacement; the number is the base position (see Table II). [b] The calculation of k_i[AP], k_i'[AP], and K_D is described under Materials and Methods. All values are accurate to within 5%.

complex (Otwinowski et al., 1988), explains why a C·G base pair at position −6 plays a critical role in stabilizing the complex. Two water molecules bridge the peptide nitrogens of Ile 79 and Ala 80 at the amino terminus of the E-helix to the major groove functional groups of A5 and G6. In addition, the water molecule that bridges N79 or Ile to N7 of G6 is firmly positioned by the Nζ of Lys 72. This arrangement of hydrogen bonds requires a purine at position 6 that accepts two hydrogen bonds and channels the positive charge potential of the E-helix dipole to the most electronegative portion of the bases, that is N7 and N6 or O6 of the purines. Changing C·G to T·A at position −6 would clearly disrupt at least one of these water-mediated hydrogen-bonding networks since an N6 of an A would not serve as a hydrogen bond acceptor, a requirement fulfilled by the O6 or G. A secondary disruptive effect of a mutation at this position might be the change in context for the 11° roll angle at the adjacent $(T\cdot A)_{-5}$ base pair, which, by allowing the helix to bend, facilitates many of the amino acid–phosphate oxygen contacts present in the complex.

Table IV also shows that a symmetrical G to T change at position −9 reduces repressor affinity by about 700-fold. Surprisingly, Bass et al. (1987) identified this position to be relatively insensitive to mutational change in their in vivo assay. The importance of having a G at position −9 is demonstrated on two grounds. First, a G at position −9 is conserved in all three *E. coli trp* operators. Second, the crystalline repressor/operator complex shows that the guanidino group of Arg 69 makes direct hydrogen bonds to O6 and N7 at position −9 (Figure 2B). Replacing a G with a T at position −9 would disrupt at least one of the two hydrogen bonds and possibly both since the O4 of T would be in a different position than the hydrogen bond acceptors, O6 and N7 of G. The substantial difference between the in vivo results and those reported here for the G_{-9} substitution may be more apparent than real. A stabilizing interaction at or near the very end of the operator fragment may lash down an otherwise pliable part of the complex and thereby diminish the phosphatase's access to the labeled termini. This local effect would exaggerate the importance of the G_{-9} to the overall stability of the complex and may represent a weakness of this type of "protection" assay. A possibly related observation is noted in the case of phage 434 repressor/operator cocrystals, where the orientation of some amino acid side chains relative to the ends of the oligonucleotide in the cocrystals changes as the length of the oligonucleotide changes from 14 to 20 base pairs (Anderson et al., 1987; Aggarwal et al., 1988).

A symmetrical T to A change at position −2 reduces repressor affinity by about 50-fold (Table IV). This result is

Table V: Effect of Protein Species on the Dissociation Constant of the *trp* Aporepressor/20 Base Pair Operator Complex

protein species	k_i[AP][c] (min^{-1})	k_i'[AP][c] (min^{-1})	K_D[c] (M)
	+ 1 mM L-Tryptophan		
wild type	1.44	0.0139	1.95×10^{-9}
AV77[a]	1.44	0.0320	4.56×10^{-9}
EK49[a]	1.44	0.0122	1.71×10^{-9}
des7[b]	1.44	0.0297	4.22×10^{-9}
	− L-Tryptophan		
wild type	0.528	0.441	1.28×10^{-5}
AV77[a]	0.528	0.209	1.64×10^{-6}
EK49[a]	0.528	0.129	1.62×10^{-7}

[a] Mutant repressors are named as two letters followed by a number. The first letter is the one-letter abbreviation for the wild-type amino acid; the second letter is the one-letter abbreviation for the replacing amino acid; the number is the amino acid position. [b] des7 is an altered *trp* aporepressor protein missing seven N-terminal amino acids. [c] The calculation of k_i[AP], k_i'[AP], and K_D is described under Materials and Methods. All values are accurate to within 5%.

consistent with the finding of Bass et al. (1987) that this mutation results in a mild decrement in repressor affinity in vivo. The crystal structure of the repressor/operator complex does not show interactions (either direct or water mediated) involving the bases at this position, suggesting that a T to A change at position −2 does not favor the DNA conformation in the complex that facilitates the intimate protein/DNA contact surface required to form many of the contacts that stabilize the complex. Bass et al. (1987) suggested a similar role for the T·A base pair at position −2 from their in vivo repressor affinity studies.

Binding of Mutationally Altered trp Aporepressor Proteins to a 20 Base Pair Wild-Type trp Operator

trpR Mutation AV77. The DNA binding properties of the altered *trp* aporepressor protein, AV77, were of interest for two reasons. First, the sequence A77 and G78 in the last two positions of the turn is unusual for this bihelical motif, which usually requires one or both of the positions to have large nonpolar side chains. A mutational change to valine at position 77 produces a protein that represses at levels of L-tryptophan too low to support repression in a strain producing the wild-type protein (Kelly & Yanofsky, 1985). Second, several repressor proteins that bind operator DNA in the absence of ligand (such as λ*cro*, *434cro*, and λ*cI* repressors) or dissociate from their respective operators upon the binding of a ligand (such as *gal*, *lac*, and *tet* repressors) have a valine in the corresponding position.

The alkaline phosphatase protection assay was employed to quantitate the binding of AV77 to operator DNA in the presence and absence of the natural corepressor, L-tryptophan. Equilibrium dialysis experiments were employed to show that AV77 binds L-tryptophan with the same affinity as the wild-type protein (Marmorstein, 1989a). Table V demonstrates that the L-tryptophan-liganded AV77 protein binds operator DNA 2.3 times less strongly than does the comparably liganded wild-type protein. However, the opposite effect is observed when binding is assayed in the *absence* of ligand, in which case the unliganded AV77 protein binds *8 times more strongly*. The enhanced affinity of the unliganded protein for the operator is, therefore, consistent with the mutant phenotype.

The crystal structure of the aporepressor (Zhang et al., 1987) shows that the γ-methyl groups of valine at position 77 would collide with other groups of the hydrophobic brace that stabilizes the conformation of the helix-turn-helix; therefore, a valine substitution interferes with the hingelike movements associated with the repressor-to-aporepressor transition. Thus, substituting valine for alanine at position 77 freezes the helix-turn-helix motif into a conformation that is more like the active repressor than the inactive aporepressor, even in the absence of bound corepressor.

trpR Mutation EK49. Kelly and Yanofsky (1982) isolated four "super-repressor" mutants that cause the protein to repress at lower L-tryptophan concentrations than that required by the wild type. Three were found to be caused by glutamic acid to lysine changes at positions 13, 18, and 49, forming the proteins EK13, EK18, and EK49, respectively. Klig and Yanofsky (1988) used a filter-binding assay at an unusually high salt concentration (0.24 M $(NH_4)_2SO_4$) to demonstrate that EK49 binds L-tryptophan with the same affinity as the wild-type protein but binds operator DNA 10 times more strongly in the presence of excess L-tryptophan.

The alkaline phosphatase protection assay was employed to investigate the DNA-binding properties of the EK49 protein. As observed by Klig and Yanofsky (1988), at high ionic strength (0.4 M NaCl), the L-tryptophan-liganded EK49 protein binds about 10 times more tightly to operator DNA than the wild-type protein (data not shown). At more modest ionic strength (0.2 M NaCl), however, the L-tryptophan-liganded EK49 protein binds only marginally better than the wild-type protein. This is consistent with the observation of Klig and Yanofsky (1988) that in relation to the wild-type protein the enhanced binding of the EK49 protein was more noticeable at high salt concentrations. This suggests that the stabilizing influence of the lysine substitution cannot be screened, i.e., that counterion does not intervene and mitigate the augmented attractive electrostatic interaction caused by replacing glutamic acid with lysine. A more surprising result was that the unliganded EK49 protein binds about 100 times more tightly to operator DNA than the unliganded wild-type protein (Table V). Although one might have guessed that the DNA binding behavior of the EK49 protein for operator DNA was due simply to an increased electrostatic attraction, our result suggests a more subtle effect. The appropriate crystal structure with EK49 will have to be determined to establish the mechanism by which this lysine substitution affects operator binding.

trpR Chymotryptic Modification des7. In each of the crystal structures of the *trp* repressor (Schevitz et al., 1985; Lawson et al., 1988), aporepressor (Zhang et al., 1987), and repressor/operator complex (Otwinowski et al., 1988), the 14 N-terminal residues have a poorly defined conformation. In all cases, the first eight residues are crystallographically disordered; to the extent that the remaining six residues can be seen in the crystal structure, they tend to adopt different conformations and have higher thermal parameters. Since the first 14 N-terminal residues are apparently not part of the protein's globular fold, it is likely that their different conformation in each crystal structure is dictated by fortuitous crystal-packing contacts. The apparently poorly restrained conformation of the amino-terminal residues is consistent with the NMR results of Arrowsmith et al. (1989).

The 14 N-terminal residues could be involved in DNA binding, as suggested for the 11 N-terminal residues of the *cI* repressor of phage λ (Pabo et al., 1982) or the C-terminal arms of *cro* (Caruthers et al., 1986). The N-terminal arms of the *trp* repressor have been implicated in DNA binding by Carey (1988). She has produced a truncated protein by removing residues 1–7 with mild α-chymotrypsin digestion. This truncated protein binds L-tryptophan as well as the full-length protein but binds operator and nonoperator DNA about 50

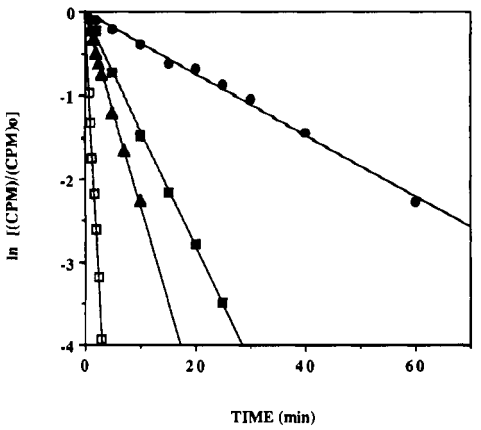


FIGURE 3: Alkaline phosphatase digestion of a 5'-terminally labeled 20 base pair idealized *gal* operator in the absence and presence of *gal* and *trp* repressors. The natural logarithm of the ratio of precipitated, (CPM), to total radioactivity, $(CPM)_0$, is plotted as a function of time in the presence of 1 mM L-tryptophan and ~1 μM *gal* repressor (●), 1 μM *trp* repressor (▲), and 10 μM *trp* repressor (■). Also plotted is the digestion in the absence of protein (□). The experimental conditions are described under Materials and Methods.

times more weakly under the conditions of the gel shift assay, which included a pH of 6.0. These results suggested that the N-terminal arms of the *trp* repressor are not important for DNA-binding *specificity* but contribute to DNA binding in general.

The alkaline phosphatase protection assay demonstrates that in the presence of L-tryptophan, des7 binds operator DNA only half as well as the wild-type protein (Table V). This decrement is not as severe as the 50-fold decrement observed by Carey (1989). The discrepancy between the two binding experiments may be due to the fact that, at pH 6.0, the N-terminal amino group of the native protein is more protonated and therefore may contribute more to the attractive electrostatic interaction with the DNA than the less protonated N-terminal amino group used in the alkaline phosphate protection assay at the higher pH of 7.4. In either case, both observations suggest that the N-terminal regions of the *trpR* protein do not play a major role in operator-specific binding.

Modeling of the N-terminal regions of the repressor onto the crystalline repressor/operator complex shows that it would be difficult for the 14 N-terminal amino acids of the repressor to wrap around the DNA in the manner seen in the λ-repressor/operator complex (Pabo et al., 1982). The uniquely complicated subunit interface of the dimeric *trp* repressor (Schevitz et al., 1985) suggests the possibility that the N-terminal segment of the repressor might be used to facilitate folding of the dimer. This speculation is consistent with the finding of Hurlburt and Yanofsky that a *trpR* mutation that begins the polypeptide at residue 8 prevents normal maturation of the truncated *trpR* protein (B, Hurlburt, personal communication).

Oligonucleotide Binding Specificity

The results shown in Figure 3 suggest that this assay might be of general use. Specifically, it shows that the presence of *gal* repressor inhibits the alkaline phosphatase mediated digestion of a ^{32}P-labeled 20 base pair *gal* operator at least 100-fold under the assay conditions used, while the *trp* repressor does a relatively poor job of protecting the *gal* operator from alkaline phosphatase digestion. Although we used 20 base pair oligonucleotides to investigate the binding of *trp* and *gal* repressors to their respective operators, the optimal fragment length required for studying other protein/DNA systems by this method may be different. Furthermore, some protein/oligonucleotide interactions could be such that the terminal phosphate may not be protected from digestion by alkaline phosphatase because the local geometry precludes protection by a firmly bound protein.

Increased interest in studying specific protein/DNA interactions by crystallographic and spectroscopic techniques has increased our need for a convenient way to identify and study "target" DNA segments (14–30 base pairs) that are compact enough to both meet the demands of the physical technique and exhibit the appropriate sequence-specific affinity. Once inferences regarding the chemistry of the interface are drawn from the structural studies, it is important to correlate the results of carefully controlled in vitro affinity measurements with structural and environmental parameters. In combination with mutational studies, we hope to ultimately relate the physical–chemical mechanism of binding to the physiological function. The most commonly used methods to quantitatively investigate protein/DNA interactions in vitro have been the nitrocellulose filter binding (Riggs et al., 1970) and gel shift (Garner & Revzin, 1981; Fried & Crothers, 1981) assays. Although of immense value, neither of these assays has been demonstrated to routinely give quantitative affinity measurements with the relatively short oligonucleotides required for crystallographic and spectroscopic studies. In addition, filter binding and gel shift studies often require a narrow range of salt and pH conditions, which are often far from physiological [examples for the *trp* repressor/operator system are given by Klig and Yanofsky (1987) and Carey (1988)]. This assay can be applied over a wide range of solvent conditions, consistent with reliable function of alkaline phosphatase. For this reason the development of this assay is of importance in our efforts to understand the physical–chemical basis for the function of *trp* repressor/operator system and possibly other specific protein/DNA regulatory complexes.

ADDED IN PROOF

Since the paper was accepted, Staake et al. (1990) have asserted that *E. coli* extracts, enriched with the *trpR* protein, react more strongly with the duplex having a top strand of TAGCGTACTAGTACGCTA (duplex β) than they do with the duplex having a top strand of TGTACTAGTTAACTAGTACA (duplex α). Duplex β in an idealized symmetrical sequence derived from the *trpEDCBA* promoter and is centered on a pseudodyad in the *trp* promoter that is 4 base pairs from the dyad of duplex α. The operator sequences chosen for both the assay presented here and for the target of the *trp* repressor in the crystallographic work (Otwinowski et al., 1988) are effectively the same as duplex α. In contradiction to the results of Staake et al. (1990), duplex α has been shown to be the target site for *trp* repressor by the band-shift and footprinting studies of Carey (1988, 1989) and to be consistent with the genetic analysis of Youderian and his colleagues (Bass et al., 1988).

ACKNOWLEDGMENTS

We thank R. L. Fox, W. Mandecki, M. Hayden, R. Metzger, and B. Powell of Abbott Laboratories for synthesizing some of the oligonucleotides and J. Carey of Stanford University and S. Adhya of the National Institutes of Health for supplying samples of des7 and *gal* repressor, respectively. We also thank J. C. Vidal and Z. Otwinowski at the University of Chicago for their help with the mathematical treatment of the binding assay.

Registry No. ACTAGTTAACTAGT, 108568-68-9; GTACTAGTTAACTAGTAC, 108568-95-2; TGTACTAGTTAACTAGTAC, 108569-01-3; GGTACTAGTTAACTAGTAC, 128495-08-9;

TTGTACTAGTTAACTAGTAC, 128495-11-4; CGTACTAGTTAACTAGTACG, 99637-95-3; TCGTACTAGTTAACTAGTACG, 128495-12-5; CGTATTAGTTAACTAATACG, 128495-09-0; CTTACTAGTTAACTAGTAAG, 128495-10-3; CGTACTAGATATCTAGTACG, 128495-13-6; CGGTGTAAACGTTTACACCG, 128526-03-4; L-Trp, 73-22-3; indole-3-propionic acid, 830-96-6.

References

Aggarwal, A. K., Rodgers, D., Drottar, M., Ptashne, M., & Harrison, S. C. (1988) *Science 242*, 899–907.

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

Arrowsmith, C. H., Carey, J., Treat-Clemons, L., & Jardetzky, O. (1989) *Biochemistry 28*, 3875–3879.

Bass, S., Sugiono, P., Arvidson, D. N., Gunsalus, R. P., & Youderian, P. (1987) *Genes Dev. 1*, 565–572.

Bass, S., Corrells, V., & Youderian, P. (1988) *Science 242*, 240–245.

Benson, N., Sugiono, P., Bass, S., Mendelman, L. V., & Youderian, P. (1986) *Genetics 114*, 1–14.

Brennan, R. G., Takeda, Y., Kim, J., Anderson, W. F., & Matthews, B. W. (1986) *J. Mol. Biol. 188*, 115–118.

Carey, J. (1988) *Proc. Natl. Acad. Sci. U.S.A. 85*, 975–979.

Carey, J. (1989) *J. Biol. Chem. 264*, 1941–1945.

Chaconas, G., & v. d. Sande, J. H. (1980) *Methods Enzymol. 65*, 75.

Chandrasegaran, S., Smith, H. O., Amzel, M. L., & Ysern, X. (1986) *Proteins 1*, 263–266.

Frederick, C. A., Grable, J., Melia, M., Samudzi, C., Jen-Jacobson, L., Wang, B.-C., Greene, P., Boyer, H. W., & Rosenberg, J. M. (1984) *Nature 309*, 327–331.

Fried, M. G., & Crothers, D. M. (1981) *Nucleic Acids Res. 9*, 6500–6525.

Garner, M. M., & Revzin A. (1981) *Nucleic Acids Res. 9*, 3047–3060.

Joachimiak, A., Kelly, R. L., Gunsalus, R. P., Yanofsky, C., & Sigler, P. B. (1983) *Proc. Natl. Acad. Sci. U.S.A. 80*, 668–672.

Joachimiak, A., Marmorstein, R. Q., Schevitz, R. W., Mandecki, W., Fox, J. L., & Sigler, P. B. (1987) *J. Biol. Chem. 262*, 4917–4921.

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

Klig, L. S., Crawford, I. P., & Yanofsky, C. (1987) *Nucleic Acids Res. 15*, 5339–5351.

Lo, K.-M., Jones, S. S., Hackett, N. R., & Khorana, G. (1984) *Proc. Natl. Acad. Sci. U.S.A. 81*, 2285–2289.

Marmorstein, R. Q. (1989) Thesis, University of Chicago, Chicago, IL.

Marmorstein, R. Q., & Sigler, P. B. (1989) *J. Biol. Chem. 264*, 9149–9154.

Maxam, A. M., & Gilbert, W. (1977) *Proc. Natl. Acad. Sci. U.S.A. 74*, 560–564.

Maxam, A. M., & Gilbert, W. (1980) *Methods Enzymol. 65*, 499.

Nelder, J. A., & Mead, R. (1963) *Computer J. 7*, 308–313.

Otwinowski, Z., Schevitz, R. W., Zhang, R.-G., Lawson, C. L., Joachimiak, A., Marmorstein, R. Q., Luisi, B., & Sigler, P. B. (1988) *Nature 335*, 321–329.

Pabo, C. O., Krouatin, W., Jeffrey, A., & Sauer, R. T. (1982) *Nature 298*, 441–443.

Pabst Laboratories Circular OR-10 (1956) *Ultraviolet Absorption Spectra of 5′ Ribonucleotides*, Milwaukee, WI.

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

Schevitz, R. W., Otwinowski, Z., Joachimiak, A., Lawson, C. L., & Sigler, P. B. (1985) *Nature 317*, 782–786.

Scopes, R. K. (1987) in *Protein Purification: Principles and Practice*, 2nd ed. (Canter, C. R., Ed.) p 280, Springer-Verlag Inc., New York.

Segel, I. H. (1975) in *Enzyme Kinetics: Behavior and Analysis of Rapid Equilibrium and Steady-State Enzyme Systems* (Segel, I. H., Ed.) pp 203–206, John Wiley & Sons, New York.

Somerville, R. L. (1983) in *Amino Acid Biosynthesis and Regulation* (Hermann, R. M., & Somerville, R. L., Eds.) pp 351–378, Addison-Wesley Publishing Co., Reading, MA.

Staake, D., Walter, B., Kisters-Woike, B., V. Wilcken-Bergmann, B., & Muller-Hill, B. (1990) *EMBO J. 9*, 1963–1967.

Thomas, G. A., Kubasek, W. L., Peticolas, W. L., Greene, P., Grable, J., & Rosenberg, J. M. (1989) *Biochemistry 28*, 2001–2009.

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

Yamaoka, K., Tanigawara, Y., Terumichi, N., & Toyozo, U. (1981) *J. Pharm. Dyn. 4*, 879–885.

Zhang, R.-G., Joachimiak, A., Lawson, C. L., Schevitz, R. W., Otwinowski, Z. O., & Sigler, P. B. (1987) *Nature 327*, 591–597.