Proc. Natl. Acad. Sci. USA
Vol. 84, pp. 6673–6676, October 1987
Biochemistry

Missing contact probing of DNA–protein interactions

(λ repressor/depurination/depyrimidation/"footprinting")

ALAN BRUNELLE AND ROBERT F. SCHLEIF

Department of Biochemistry, Brandeis University, Waltham, MA 02254

Communicated by Mark Ptashne, May 19, 1987 (received for review March 3, 1987)

ABSTRACT **We have examined the positions of contact between λ phage repressor protein and operator O_R1 DNA by scanning populations of lightly depurinated or depyrimidated DNA for bases essential to or irrelevant to repressor binding. This global scanning technique delineates the apparent contact region between λ repressor and operator and shows bases previously demonstrated or predicted to be contacted plus some additional bases. A mutant repressor, previously shown to contact DNA as wild-type repressor does with the exception of a missing contact to guanosine G4′ [Hochschild, A. & Ptashne, M. (1986) *Cell* 44, 925–933], similarly failed to contact G4′ when assayed by this method. Coupled with altering a test residue of a DNA-contacting protein to glycine or alanine so as to eliminate a specific contact, the method appears to provide an efficient means of scanning for specific residue–base contacts.**

Although x-ray crystallography has the potential for determining the general structure of protein–nucleic acid complexes (1, 2), chemical modification experiments may be a better approach for high-resolution probing of the details of these interactions (3). Additionally, chemical modification experiments yield information in instances where x-ray crystallography cannot be applied. Two general types of chemical probing experiments have been performed: (*i*) premethylation and preethylation interference experiments in which guanines and phosphates are identified whose alkylation interferes with binding of the protein (4) and (*ii*) specific missing contact experiments in which a small amino acid residue is substituted for a larger one in the binding protein and a specific base pair in the DNA is also altered (3, 5, 6). If the amino acid at the altered position previously contacted the DNA, then the substitution of a smaller residue eliminates contact at this position, which leaves the mutant protein indifferent to the actual identity of the base at this one position.

The approaches mentioned above are valuable, but they possess important limitations. First, with premethylation interference, any guanine whose methylation interferes with the protein's binding is revealed as a contact, independent of whether the guanine participated in any specific hydrogen bonding. In addition, with the premethylation approach, only guanines may easily be examined. The specific missing contact approach requires that both mutant protein and mutant DNAs be prepared and tested for each specific amino acid–base contact being tested.

Here we test and extend an approach first developed and reported only by Majors (7) that can, in principle, overcome some of these difficulties. The basic idea is that, instead of methylation of guanines, the DNA is sparingly depurinated or depyrimidated; then binding, separation, and determination of relevant and irrelevant bases is performed as in the premethylation interference approach (4, 8). Meaningful interpretation of the data obtained from such experiments depends crucially, however, on the absence of significant structure alterations generated by the absence of a base. If there is essentially no alteration in the positions of the bases not removed or of the phosphodiester backbone of the DNA and if the protein structure, once bound to the DNA, is not altered because of the missing base, then this approach could be helpful in mapping protein–DNA interactions.

Today, with detailed structural information of several DNA–protein interactions available on specific amino acid residue–base contacts, it is possible to test the missing contact approach. The λ phage repressor–operator system is a good test case. Structure determination (9), chemical modification experiments (10), genetics (10–12), and combined genetics and chemical modification experiments (3) suggest a number of bases likely contacted or not contacted by the repressor. Most notably, premethylation interference experiments with the wild-type repressor show that a serine at position 45 closely approaches G4′ on the bottom strand of O_R1 (see Fig. 1), but a mutant with alanine at position 45 does not contact there (3). As required for the base removal approach described in this paper, we find that the wild-type repressor but not the alanine-45 mutant contacts G4′ on both strands. Additionally, other bases contacted by repressor are revealed, most of which were seen in the previously applied methods.

MATERIALS AND METHODS

Enzymes and Chemicals. Piperidine and formic acid (88%) were from Fisher. Anhydrous hydrazine and dimethyl sulfate were from Aldrich. [γ-^{32}P]ATP was from New England Nuclear.

DNA Preparation. Plasmid pAH5h3 was a gift from A. Hochschild (3). O_R1 is contained on a 90-base-pair *Hin*dIII–*Bst*NI fragment of plasmid pAH5h3. The top strand was end-labeled at the *Hin*dIII site, and the bottom strand was end-labeled at the *Bst*NI site. The labeled 90-base-pair fragment was isolated by cutting with the opposite enzyme, electrophoresing, and then electroeluting the end-labeled fragment.

Repressors. Preparations of the wild-type (serine-45) and mutant (alanine-45) λ repressors were a gift from J. Douhan III (prepared as described in ref. 3).

DNA Premodification. *G reaction.* In a 1.5-ml Microfuge tube, end-labeled O_R1 fragment was resuspended in a final volume of 200 μl of 0.05 M sodium cacodylate buffer (pH 8.0). Two minutes after the addition of 1 μl of dimethyl sulfate, the reaction was quenched with the addition of 50 μl of 1 M 2-mercaptoethanol/1.5 M sodium acetate containing 5 μg of tRNA as carrier (13). It was ethanol-precipitated twice, rinsed, dried, and resuspended in 25 μl of 10 mM Tris·HCl, pH 8.0/1 mM EDTA (TE). Five microliters of 2 M NaCl was added, and the sample was incubated at 90°C for 10 min to depurinate the methylated bases. The sample was then cooled slowly to room temperature to allow denatured DNA

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

to renature, ethanol-precipitated, washed with 95% ethanol, dried, and resuspended in 20 μl of TE.

G+A reaction. In a 1.5-ml Microfuge tube, end-labeled DNA was resuspended in 15 μl of TE that contained 5 μg of tRNA as carrier. One and one-half microliters of 1 M formate (pH 2.0) (4% vol/vol formic acid) was added, and the reaction was incubated in a sealed tube at 37°C for 30 min (13). The reaction was ethanol-precipitated twice, rinsed with 95% ethanol, dried, and resuspended in 20 μl of TE.

C+T reaction. End-labeled DNA was resuspended in 25 μl of glass distilled water in a Microfuge tube. Hydrazine (15 μl) was added, and the reaction was allowed to proceed for 30 min at room temperature in the sealed tube (13). The DNA was precipitated by the addition of 1 ml of 1-butanol, vigorous mixing, and centrifugation at room temperature. The supernatant was discarded, and the pellet was resuspended in 100 μl of 0.3 M sodium acetate. It was ethanol-precipitated, rinsed, dried, and resuspended in TE. All premodified DNA could be stored at −20°C for up to 2 weeks without significant degradation.

Binding Reaction. Approximately 1–2 × 10^5 cpm of end-labeled O_R1 fragment was used per reaction. The binding buffer was 50 mM KCl/0.5 mM EDTA/10 mM Tris acetate, pH 7.4/1 mM dithioerythritol/5 μM bovine serum albumin/5% glycerol. Repressor sufficient to bind all the premodified O_R1 DNA, as determined by a titration carried out before the actual experiment, was added to the 20-μl reaction mix. For wild-type repressor, binding was allowed to proceed for 10 min at 25°C after which a 50- to 100-fold molar excess of unlabeled pAH5h3 was added as competitor in a volume of 1 μl with gentle mixing. Dissociation of the wild-type repressor was allowed to proceed for 5 min, after which the whole reaction was loaded onto a native electrophoresis gel (8). Because the mutant repressor had a significantly faster dissociation rate, it was allowed to bind for 10 min at 25°C, after which the reaction was immediately chilled on ice. This slowed the dissociation sufficiently, so that upon addition of 50- to 100-fold molar excess of competitor, the reaction could be loaded immediately onto the gel. These conditions for the dissociation of wild-type and mutant repressors from the premodified O_R1 fragment yielded ≈15% dissociation. The bound and free O_R1 DNA were well resolved on the gels. They were cut out, electroeluted, and ethanol-precipitated. The isolated premodified DNA was then subjected to piperidine cleavage (13), separated on a sequencing gel, and autoradiographed to display the cleavage patterns, and densitometry of relevant bands was performed with a Joyce–Lobel densitometer.

RESULTS

Confirmation of Contacted Bases Made by the Wild-Type Repressor. The basic principle of the missing contact experiment is similar to that of premodification experiments (4, 8). A DNA fragment with a binding site for a protein is depurinated or depyrimidated at a level slightly less than one base removed per DNA fragment, after which it is rehybridized to restore the double-stranded form. Protein is added to the DNA, and the species of DNA capable of binding are separated from the species incapable of binding or from which the protein rapidly dissociates because they are missing bases important for protein binding. Subsequently, the two DNA classes are isolated, cleaved at the positions of missing bases, and separated on sequencing gels to reveal the positions irrelevant or crucial to binding. DNA with missing bases that do not interfere with protein binding enrich the bands in the DNA sample capable of binding the protein, and the DNA missing bases that interfere with protein binding enrich the bands from the DNA unable to bind the protein.

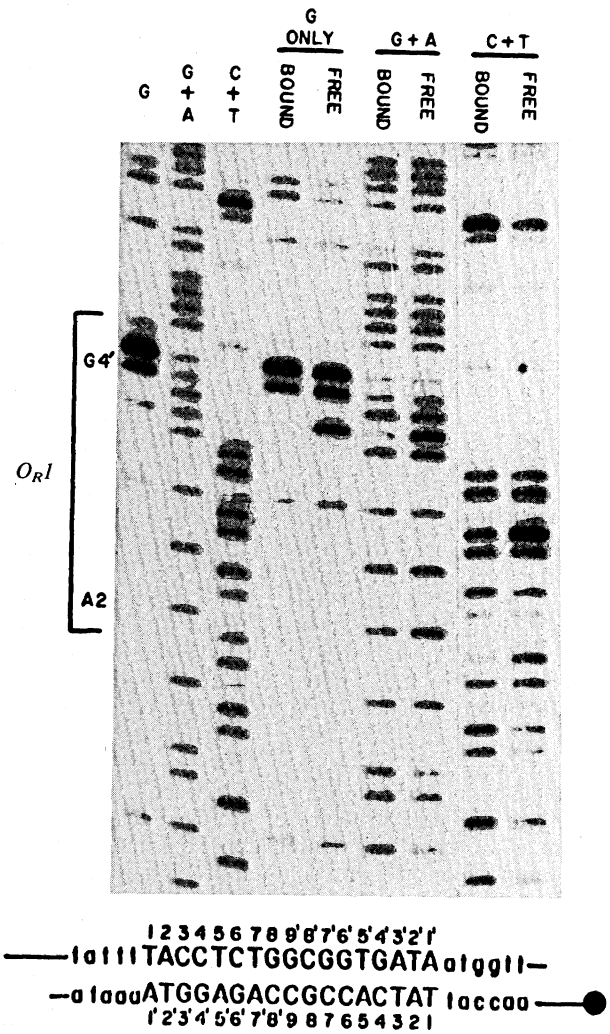

FIG. 1. Methylation depurination (G only), acid depurination (G+A), and depyrimidation (C+T) contacts of wild-type λ repressor for the bottom strand of O_R1. For reference, the diagram at the bottom is that of O_R1 with the numbering system of Hochschild and Ptashne (3). —●, position of the label; bound, gel lanes with modified DNA isolated from repressor–DNA complexes; free, gel lanes with the DNA that had dissociated and was free of bound complexes. G, G+A, and C+T are size and comparison standards, which were prepared from the same sources of DNA as used in the contact determinations but were not treated with repressor protein.

Fig. 1 shows the data obtained with wild-type repressor binding to the bottom strand of O_R1 for methylation depurination (G only), acid depurination (G+A) and depyrimidation (C+T). By comparing band intensities in the bound and free lanes to the standard tracks in Fig. 1, we see significant effects from bases missing at G6′, A7′, C7, A2, and T1 and weaker effects at G4′, C8′, C9, and C6. All other positions show either no or very weak effects. Similar measurements were performed on the top strand, and Fig. 2*A* contains a summary of the data obtained on both strands.

Fig. 2*B* is a summary of contact data from previous experiments (3, 9, 10, 14–16). It shows those contacts for which there is strong evidence, premethylation contacts, and contacts predicted from model building based on the crystal structure of repressor's amino-terminal domain (3, 9, 10, 14–16). There are nine bases within O_R1 that are predicted to be contacted by repressor amino acid side groups. These are underlined in Fig. 2*B*. Eight of these are revealed by the missing contact approach as bases that when missing affect repressor binding. Of these, the G4′ contact on the bottom

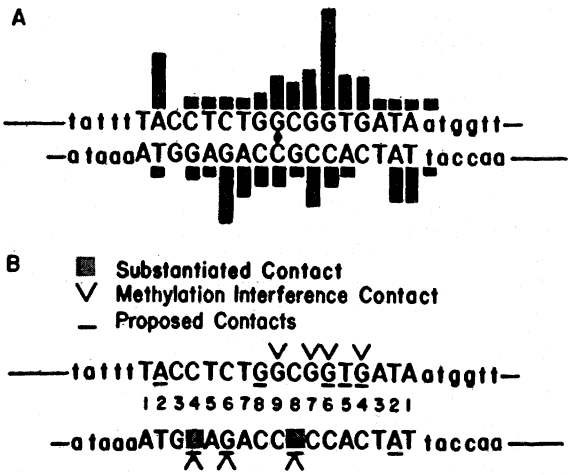

FIG. 2. Summary of data for the λ repressor at O_R1. (*A*) The complete data for both strands of O_R1 obtained by the missing contact method. The magnitude of the effect is shown by the size of the bar over or under that position of the operator. For example, the bar under position T2′ of the bottom strand indicates a very weak or insignificant effect was observed upon removal of this base. Compare this assessment with the data in Fig. 1. (*B*) Compilation of other results obtained by means of methylation interference, structure determination/model building, and mutagenesis with biochemical measurements (3, 9–12). Shaded letters are positions that are proposed specific contacts and have been substantiated to a high degree of certainty with multiple biochemical and genetic experiments. Carets indicate those positions that when methylated interfere with repressor binding. Underlined bases are those that are proposed to be specifically contacted by the repressor.

strand is weak and is discussed in the next section. The ninth, G8 of the bottom strand, generates a moderate effect in the G-only reaction but only a very weak effect in the G+A reaction. We cannot explain the discrepancy at this position, although strong evidence would argue for the presence of the bottom strand contact at G8 (6, 14, 16).

Eliason *et al.* (16) have predicted that the amino-terminal arms of the repressor dimer contact G8 of both strands. However, evidence suggests that the repressor fails to make contact to the top strand G8. More specifically, repressor bound at O_R1 does not protect the top strand G8 from methylation by dimethyl sulfate but enhances it slightly (10, 14), premethylation of the top strand guanine does not interfere with repressor binding (10), and mutation of the top strand G8 to T8 causes only a 2-fold reduction in binding (6). Our data reveal only a relatively weak effect at this position; therefore, we conclude that the top strand G8 is a relatively weak contact. Consequently, we suggest that the flexible amino-terminal arm of the repressor does not make a consistent specific contact with this base on all the related operators.

In addition to the contacts previously seen or predicted, we also see moderate to strong effects at C8′ and G7′ on the top strand as well as A7′ and C7 on the bottom strand. Somewhat weaker effects are seen at C8′, C9, and C6 on the bottom strand. Whether these derive from the flexible amino-terminal arm of repressor or represent artifacts inherent in the missing contact method cannot at present be ascertained.

Mutant Alanine-45 Repressor Fails to Contact G4′. Fig. 3*A* is data showing the effects of guanine removal from the top strand upon the binding of wild-type and mutant alanine-45 repressor. The mutant repressor fails to contact G4′ on the top strand, whereas wild-type repressor makes a significant contact there. Fig. 3*B* shows a summary of our normalized data for top and bottom strands. These data extend the results obtained by Hochschild and Ptashne (3), who obtained similar results with the bottom strand of O_R1. These data

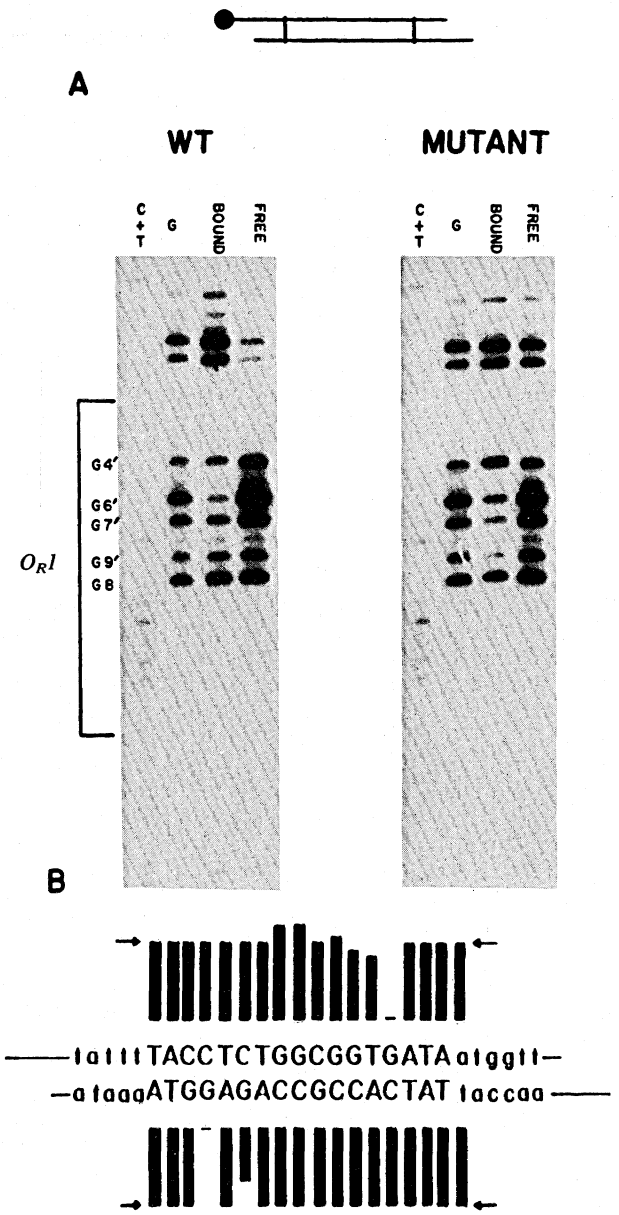

FIG. 3. (*A*) Methylation depurination (G only) contacts for wild-type and mutant repressor of the top strand of O_R1. Bound, free, and standards are as described in Fig. 1. ●—, position of the label. (*B*) A summary of the differences in the strength of the contacts derived from the wild-type and mutant repressor at O_R1. The data are normalized so that no change in the relative strength of a contact is represented by a bar reaching the height of the arrow.

demonstrate that the base removal technique is capable of identifying specific base contacts.

Depurination at position G6′ also increased the dissociation rate of the mutant relative to wild type (Fig. 3 *A* and *B*), but the difference was much less pronounced than at position G4′. Unexpectedly, removal of bases at either G7′, G9′, or G8 reduced the dissociation rate of the mutant as compared to the wild-type repressor. It is apparent that even though the primary effect caused by the change of amino acid serine-45 to alanine was to eliminate recognition of G4′, the mutation generates other subtle effects.

As visualized with the missing contact technique, mutant repressor also fails to contact the symmetrically located base G4′ on the bottom strand (Table 1) as demonstrated previously with methylation interference (3). The weak contact to this base displayed by the wild-type repressor may be related

Table 1. Relative band enrichment for the bottom strand G4′

Repressor	G4′ Bound	G4′ Free
Wild-type	0.79	1.71
Mutant alanine-45	1.00	0.96

Values are densitometric readings of position G4′ for each lane (bound or free) normalized to a band away from O_R1. The values have also been divided by the value obtained from the standard G lane so that a value close to 1.00 indicates no effect.

to the weaker contacts in general that the repressor shows to this region of operator and probably results from the fact that this half of O_R1 deviates by 3 base pairs from the consensus operator sequence. Perhaps the strong effects observed by Hochschild and Ptashne (3) at this position were the result of measuring steric interference between the N^7 methyl group at this G4′ and amino acid residue 45.

DISCUSSION

We have applied a missing contact probing technique to identify the bases of λ operator O_R1 that are apparently contacted by λ repressor. Unlike premethylation interference, this method depends on the removal of important chemical interactions from the DNA rather than the steric prevention of binding by the addition of a methyl group. When tested with the wild-type λ phage repressor and the alanine-45 mutant repressor, known not to contact base G4′ of the operator (3), we found that the wild-type repressor but not the mutant repressor contacted this base. Many other significant effects observed by this method correspond to known or predicted contacts by repressor to operator. Though a number of effects seen by this method are not explained by previous work with λ repressor and O_R1, the result with the alanine-45 mutation suggests that this method may be usable for probing contacts between other proteins and their DNA binding sites.

The missing contact approach tested here could streamline scanning for specific amino acid residue–base contacts. By the methods used up to now, it was necessary to mutate both the residue in the protein and the contacted base in the DNA binding site. Clearly such an approach is highly inefficient in situations where accurate predictions cannot be made. Now, if other proteins behave in the same manner as λ repressor, it will suffice to predict any of the residues that contact bases. Once the contacting residue has been altered to a smaller residue, the missing contact approach should reveal what base is not contacted.

For those positions shown or predicted to be contacted by repressor, the magnitude of the effects we observed for major contacts generally corresponds to the details known about λ repressor binding. Three predicted base-specific interactions made by λ repressor to operator arise from helix 3, the DNA "recognition helix." According to model building (9, 15), glutamine-44 makes two hydrogen-bonding contacts to A2, serine-45 makes a hydrogen bond to G4′ (see also ref. 3), and alanine-49 has van der Waals contacts to T5′. As discussed above, we observe the serine-45 to G4′ contact. Additionally, we find that removing A2 from either strand significantly affects binding of repressor. T5′ on the upper strand appears to contribute significantly to binding; however, T5′ is replaced by A5′ on the bottom strand and as expected, only minor effects are observed upon removal of this base.

The proposed contacts made by residues outside the DNA recognition helix also coincide with the pattern observed with the base removal experiments. The removal of G6′ from either strand causes the most drastic effect on the dissociation rate as compared to all other positions. This position is predicted to be a bidentate hydrogen-bond acceptor of asparagine-55 (15). Not surprisingly, this base is highly conserved and is present in 11 of the 12 half-sites of the early left and right λ operators. Mutations at this position in O_R1 and O_R2 drastically affect binding of the repressor (10, 16).

The pattern observed with the base removal technique reveals an asymmetry, a reduction in the number and strength of the contacts observed in the left or nonconsensus half-site of O_R1. This asymmetry has been alluded to before when mutations in the nonconsensus half-site revealed "context dependent anomalies" (6). Two questions arise: (*i*) is the asymmetry in the structure of O_R1 or in repressor bound at O_R1 and (*ii*) would repressor make fully productive contacts to all 10 predicted contact positions if bound to the consensus operator?

We thank Ann Hochschild for helpful discussions. This work was supported by National Institutes of Health Grant GM 18277 and National Science Foundation Grant DMB-8205977 to R.F.S.

1. McClarin, J. A., Frederick, C. A., Wang, B. C., Greene, P., Boyer, H. W., Grable, J. & Rosenberg, J. M. (1986) *Science* **234**, 1526–1541.
2. Anderson, J. E., Ptashne, M. & Harrison, S. C. (1985) *Nature (London)* **316**, 596–601.
3. Hochschild, A. & Ptashne, M. (1986) *Cell* **44**, 925–933.
4. Seibenlist, U. & Gilbert, W. (1980) *Proc. Natl. Acad. Sci. USA* **77**, 122–126.
5. Ebright, R. H. (1986) *Proc. Natl. Acad. Sci. USA* **83**, 303–307.
6. Hochschild, A., Douhan, J., III, & Ptashne, M. (1986) *Cell* **47**, 807–816.
7. Majors, J. (1976) Dissertation (Harvard University, Cambridge, MA).
8. Hendrickson, W. & Schleif, R. F. (1985) *Proc. Natl. Acad. Sci. USA* **82**, 3129–3133.
9. Pabo, C. O. & Lewis, M. (1982) *Nature (London)* **298**, 443–447.
10. Johnson, A. D. (1980) Dissertation (Harvard University, Cambridge, MA).
11. Nelson, H. C. M., Hecht, M. H. & Sauer, R. T. (1983) *Cold Spring Harbor Symp. Quant. Biol.* **47**, 441–449.
12. Nelson, H. C. M. & Sauer, R. T. (1985) *Cell* **42**, 549–558.
13. Maxam, A. M. & Gilbert, W. (1980) *Methods Enzymol.* **65**, 499–560.
14. Pabo, C. O., Krovatin, W., Jeffrey, A. & Sauer, R. T. (1982) *Nature (London)* **298**, 441–443.
15. Lewis, M., Jeffrey, A., Wang, J., Ladner, R., Ptashne, M. & Pabo, C. O. (1983) *Cold Spring Harbor Symp. Quant. Biol.* **47**, 435–440.
16. Eliason, J. L., Weiss, M. A. & Ptashne, M. (1985) *Proc. Natl. Acad. Sci. USA* **82**, 2339–2343.