supported by the presence of Lower Mississippean conodonts in chert and tuff that may overlie unconformably the Nitinat tuff.
27. R. W. Yole, *Proc. Geol. Assoc. Can.* **20**, 30 (1969).
28. J. E. Muller, K. E. Northcote, D. Carlisle, *Geol. Surv. Can. Pap. 74-8* (1974), p. 1.
29. F. Barker, A. S. Brown, J. R. Budahn, G. Plafker, *Chem. Geol.* **75**, 81 (1989).
30. J. A. Jeletzky, *Geol. Surv. Can. Pap. 69-14* (1970), p. 1.
31. E. Carlisle and T. Susuki, *Can. J. Earth Sci.* **11**, 254 (1974).
32. $\varepsilon_{Nd} = 10^4[(^{143}Nd/^{144}Nd_{sample})/(^{143}Nd/^{144}Nd_{CHUR}) - 1]$ where CHUR is the chondritic uniform reservoir. See D. J. DePaolo, *Neodymium Isotope Geochemistry* (Springer-Verlag, New York, 1988). ε_{Sr} is the similar notation for the $^{87}Sr/^{86}Sr$ ratio.
33. S. D. Samson, P. J. Patchett, G. E. Gehrels, R. G. Anderson, *J. Geol.*, in press.
34. R. W. Griffiths and I. H. Campbell, *Earth Planet. Sci. Lett.* 99, 66 (1990).
35. E. M. MacKevett, Jr., and D. H. Richter, *Geol. Assoc. Can. Cordilleran Sect. Progr. Abstr.* xx, 13 (1974).
36. A. Davis and G. Plafker, *Can. J. Earth Sci.* **22**, 183 (1985).
37. J. W. Hawkins and J. T. Melchior, *J. Geophys. Res.* **90**, 11431 (1985).
38. D. Weiss, Y. Bassias, I. Gautier, J.-P. Mennessier, *Geochim. Cosmochim. Acta* **53**, 2125 (1989).
39. Y. Bassias, H. L. Davies, L. Leclaire, D. Weis, *Bull. Mus. Natl. Hist. Nat. Paris* **4/9**, 367 (1987); L. Leclaire *et al.*, *Geo-Marine Lett.* 7, 169 (1987).
40. H. L. Davies *et al.*, *Contrib. Mineral. Petrol.* **103**, 457 (1989).
41. R. A. Duncan, *J. Volcanol. Geotherm. Res.* **4**, 283 (1978); R. Duncan, unpublished data; H. Whitechurch, R. Montigny, J. Sevigny, M. Storey, V. Salters, *Proc. ODP Sci. Results 120*, in press.
42. A. K. Baksi, *Chem. Geol.* **70**, 46 (1986).
43. J.-Y. Royer and M. F. Coffin, *Proc. ODP Sci. Results 120*, in press.
44. D. Weiss and F. A. Frey, *Eos* **71**, 657 (1990).
45. C. Alibert, *Proc. ODP Sci. Results 119* (1991), p. 293; M. Storey *et al.*, *Proc. ODP Sci. Results 120*, in press.
46. M. Recq and P. Charvis, *Geophys. J. Roy. Astron. Soc.* **84**, 529 (1986); M. Recq, D. Brefort, J. Malod, J.-L. Veinante, *Tectonophysics* **182**, 227 (1990); D. M. Hussong, L. K. Wipperman, L. W. Kroenke, *J. Geophys. Res.* **84**, 6003 (1979).
47. L. W. Kroenke, *Haw. Inst. Geophys. Rep. 72-5* (Univ. of Hawaii, Honolulu, 1972); J. J. Mahoney, in *Seamounts, Islands, and Atolls, AGU Monogr. 43*, B. H. Keating, P. Fryer, R. Batiza, G. W. Boehlert, Eds. (American Geophysical Union, Washington, D.C., 1987), pp. 207–220; The preexisting oceanic lithosphere could not have been more than about 50 million years old at the time of plateau eruption [R. L. Larson, S. M. Smith, and C. G. Chase, *Earth Planet. Sci. Lett.* **15**, 315 (1972).
48. J. J. Mahoney and K. J. Spencer, *Earth Planet. Sci. Lett.*, in press.
49. It has also been postulated that the plateau was principally continental [R. L. Carlson, N. I. Christensen, R. P. Moore, *Earth Planet. Sci. Lett.* **51**, 171 (1980); A. Nur and Z. Ben-Avraham, *J. Geophys. Res.* **87**, 3644 (1982)] but recent geochemical studies do not support this hypothesis (*48*).
50. J. E. Andrews *et al.* (shipboard scientific party), *Init. Rep. Deep Sea Drilling Proj.* **30**, 27 (1975).
51. R. J. Musgrave [*Tectonics* **9**, 735 (1990)] argued on the basis of magnetostratigraphic data that Malaita was not formed as part of the Ontong-Java plateau.
52. R. Duncan, unpublished data.
53. J. A. Tarduno, W. V. Sliter, H. Mayer, paper presented at the Caltech Plume Symposium, Pasadena, CA, May 1991; J. Tarduno, unpublished manuscript.
54. J. Mahoney, personal communication.
55. C. B. Officer, J. I. Ewing, R. S. Edwards, H. R. Johnson, *Geol. Soc. Am. Bull.* **68**, 359 (1957).
56. J. B. Diebold, P. L. Stoffa, P. Buhl, M. Truchan, *J. Geophys. Res.* **86**, 7901 (1981).
57. C. L. Bowland and E. Rosencrantz, *Geol. Soc. Am. Bull.* **100**, 534 (1988).
58. T. W. Donnelly, K. Melson, R. Kay, J. J. W. Rogers, *Init. Rep. Deep Sea Drilling Proj.* **15**, 989 (1973).
59. T. W. Donnelly *et al.*, in *The Geology of North America*, vol. H, *The Caribbean Region*, (Geological Society of America, Washington, DC, 1990), pp. 339–374.
60. M. Storey *et al.*, *Geology* 19, 376 (1991); L. M. Echeverria, *Contrib. Minaral. Petrol.* **73**, 253 (1980).
61. R. A. Duncan and R. B. Hargraves, *Geol. Soc. Am. Mem.* **162**, 81 (1984).
62. E. M. Klein and C. H. Langmuir, *J. Geophys. Res.* **92**, 8089 (1987).
63. D. P. McKenzie and M. J. Bickle, *J. Petrol.* **29**, 625 (1988).
64. R. W. Griffiths and I. H. Campbell, *J. Geophys. Res.*, in press.
65. R. L. Larson, *Geology* 19, 547 (1991).
66. M. A. Richards and D. C. Engebretson, *Eos* **71**, 1575 (1990).
67. We thank N. Sleep, M. Coffin, J. Mahoney, N. Donnelly and two anonymous reviewers for helpful comments on the manuscript. Some of the ideas in this paper were developed over the past few years with R. Griffiths, I. Campbell, G. Davies, and R. Hill. We thank R. Griffiths for providing the photograph in Fig. 1.

22 February 1991; accepted 23 July 1991

The DNA Binding Arm of λ Repressor: Critical Contacts from a Flexible Region

NEIL D. CLARKE,* LESA J. BEAMER,† HARRY R. GOLDBERG, CAROL BERKOWER, CARL O. PABO‡

Segments of protein that do not adopt a well-ordered conformation in the absence of DNA can still contribute to site-specific recognition of DNA. The first six residues (NH_2-Ser^1-Thr^2-Lys^3-Lys^4-Lys^5-Pro^6-) of phage λ repressor are flexible but are important for site-specific binding. Low-temperature x-ray crystallography and codon-directed saturation mutagenesis were used to study the role of this segment. All of the functional sequences have the form [X]^1-[X]^2-[Lys or Arg]^3-[Lys]^4-[Lys or Arg]^5-[X]^6. A high-resolution (1.8 angstrom) crystal structure shows that Lys^3 and Lys^4 each make multiple hydrogen bonds with guanines and that Lys^5 interacts with the phosphate backbone. The symmetry of the complex breaks down near the center of the site, and these results suggest a revision in the traditional alignment of the six λ operator sites.

ALTHOUGH THE WELL-CHARACTERIZED DNA binding motifs (helix-turn-helix, zinc finger, and so forth) use stably folded units of secondary structure for recognition, flexible regions can also contribute to site-specific recognition. Unfortunately, crystallographic studies of such segments are difficult because residual disorder in these segments may result in a poor electron density map. In order to address these problems in studying a flexible DNA-binding segment of λ repressor, we (i) determined the crystal structure of the repressor-operator complex at low temperature to reduce thermal motion and (ii) performed an exhaustive genetic and biochemical analysis to test the functional relevance of the observed crystal structure.

We have focused on the first six residues of λ repressor. This segment has been referred to as an "arm" because of the way it wraps around the operator DNA (see Fig. 1). Nuclear magnetic resonance (NMR) experiments clearly show that this segment is flexible in solution (*1*). Previous crystallographic studies of the protein by itself or of the protein bound to DNA showed only weak electron density for the arm (*2, 3*). However, despite its inherent flexibility, the arm plays a critical role in DNA binding and deletion of the arm results in a greater than 8000-fold reduction in DNA binding affinity (*4*).

The NH_2-terminal arm of λ repressor has the sequence (NH_2-Ser^1-Thr^2-Lys^3-Lys^4-Lys^5-Pro^6-). In order to understand how this segment contributes to DNA binding, we first tested the functional significance of each residue. In separate experiments, each codon in the arm was randomly mutated by oligonucleotide cassette mutagenesis of a plasmid-borne gene encoding the amino-terminal domain of λ repressor (*5*). Subsequent transformation into *Escherichia coli* and selection for resistance to phage infection allowed us to identify a large number of functional sequence variants in the

N. D. Clarke, Howard Hughes Medical Institute and Department of Molecular Biology and Genetics, Johns Hopkins University School of Medicine, Baltimore, MD 21205.
L. J. Beamer and C. Berkower, Department of Molecular Biology and Genetics, Johns Hopkins University School of Medicine, Baltimore, MD 21205.
H. R. Goldberg, Department of Biophysics and Biophysical Chemistry, Johns Hopkins University School of Medicine, Baltimore, MD 21205.
C. O. Pabo, Howard Hughes Medical Institute and Departments of Molecular Biology and Genetics and of Biophysics and Biophysical Chemistry, Johns Hopkins University School of Medicine, Baltimore, MD 21205.

*Present address: Center for Advanced Research in Biotechnology, National Institute of Standards and Technology, Rockville, MD 20850.
†Present address: Molecular Biology Institute, University of California, Los Angeles, CA 90024.
‡Present address: Department of Biology, Massachusetts Institute of Technology, Cambridge, MA 02139.

NH_2-terminal arm. A list of the functional substitutions we identified is given in the top part of Table 1.

It is clear from Table 1 that residues 3, 4, and 5 are very important for repressor function: only Arg and Lys are acceptable at positions 3 and 5 and only Lys is functional at position 4. Residues 1, 2, and 6 are less important and there are many functional substitutions at these positions (*6*). Variants with a deletion of the Ser^1 codon (Δ1) or a deletion of both the Ser^1 and Thr^2 codons (Δ1-2) also were functional, and thus it does not appear that Ser^1 and Thr^2 make any critical contacts with the DNA (*7*).

We verified the importance of residues 3, 4, and 5 by introducing specific substitutions at these positions by site-directed cassette mutagenesis and testing the ability of these variants to confer resistance to phage infection. In particular, we introduced Ala and Met at residues 3, 4, and 5 and introduced Arg at residue 4. None of these variants conferred phage resistance.

This test for phage resistance gives only a crude indication of the affinity with which a repressor variant binds DNA. In order to further assess the DNA binding activity of the variant proteins, we measured their ability to repress expression of β-D-galactosidase from a lacZ fusion under the control of the λP_R promoter (*8*). Results of the β-D-galactosidase repression assay for substitutions at positions 3, 4, and 5 are shown in Fig. 2. Each of the Ala and Met substitutions at these critical positions results in similarly poor repression of β-D-galactosidase expression. It is possible that the DNA-induced structure adopted by these residues is highly cooperative and that mutation of any of the residues disrupts the contacts made by all. These results also show that Arg is not functional at position 4; in fact, it is no better at this position than Ala or Met. Only the Lys side chain is capable of making the appropriate contacts with the operator DNA.

Gel-shift assays were used to determine the in vitro binding activities for eight mutants (*9*). Most of these had been selected as functional variants (Ala^2, Lys^2, Δ1-2, Arg^3, Arg^5, Ala^6, and Gly^6). One nonfunctional, site-directed mutant (Ala^3) was also tested. The functionally selected variants all have binding affinities for the O_L1 operator site within threefold that of wild type, while the nonfunctional Ala^3 mutant binds about 300-fold worse than wild type. These results confirm that the repression of lytic growth provides a tight selection for functional repressor activity and that the in vivo repression of β-D-galactosidase from λP_R-lacZ fusions (which was measured for all of the mutants in this study) is a useful measure of relative binding affinity to the O_R region of λ.

In order to understand the structural basis for the importance of Lys^3, Lys^4, and Lys^5 in recognition, we collected high-resolution crystallographic data at low temperature (about −15°C). We expected that this would reduce radiation damage to the crystals and also hoped it would reduce the thermal motion of the arm. This data set was used to refine the structure of the λ repressor-operator complex to a resolution of 1.8 Å (*10*). Electron density for the arm that binds to the consensus half-site improved dramatically with this new data (*11*). Our model for this arm and the electron density in which it was built are shown in Fig. 3A. The electron density is well defined for residues 3 through 6. The density remains poor for Ser^1 and Thr^2, but the genetic analysis indicated that these two residues do not make any important contacts to the operator, and it does not appear that they are constrained to any particular conformation in the complex. On the other half-site, the arm is not well ordered; this may reflect a fundamental difference (discussed below) in the rec-

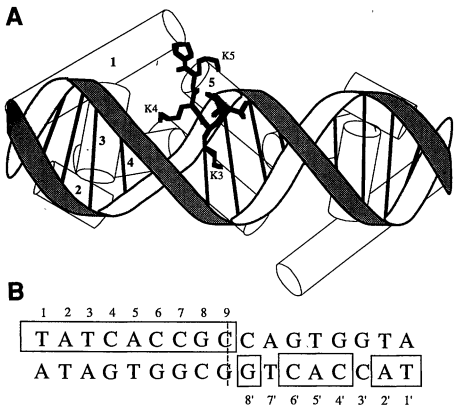

Fig. 1. (**A**) Schematic drawing of the repressor O_L1 complex, with an emphasis on the conformation of the NH_2-terminal arm. The five α-helices of the consensus half-site monomer are labeled 1 through 5. The arm is drawn in thick black lines and represents the conformation of the consensus half arm as determined in this work. The side chains for Lys^3, Lys^4, and Lys^5 are labeled K3, K4, and K5. The nonconsensus half arm is poorly defined in the electron density maps and is not shown in this schematic. (**B**) The O_L1 operator sequence. The consensus base pairs are enclosed in boxes. One half-site (that includes the central base pair) matches the consensus sequence [TATCACCGC] determined from the six λ operators and is called the consensus half-site. The other half differs from the consensus sequence at two positions, in addition to the central base pair, and is referred to as the nonconsensus half-site.

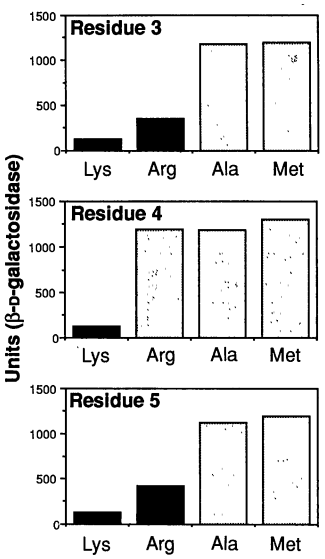

Fig. 2. In vivo repression of the O_R operator region by sequence variants at positions 3, 4, and 5. The bar graphs show β-D-galactosidase activities expressed from an O_R::lacZ fusion in cells containing repressor with the indicated arm sequence (*8*). The residue number is indicated at the top of each graph; the bottom of each graph gives the amino acid at that position. A black bar indicates that the amino acid was identified by the functional selection. A gray bar indicates an amino acid substitution that was introduced by site-directed cassette mutagenesis and that was verified to be inactive by the phage-resistance criterion. β-D-Galactosidase activity was assayed in US3/λ200 (*8*) cells transformed with the relevant repressor gene and with a compatible plasmid containing the $LacI^Q$ gene under non-inducing conditions.

Table 1. Listing by residue number of the functional sequence variants and of some nonfunctional variants; the first entry in each column is the wild-type residue. At each position, the wild-type residue was recovered in the functional selection. Nonfunctional variants at positions 3, 4, and 5 were introduced by site-directed mutagenesis. Nonfunctional variants at residue 6 were obtained by sequence screening of a restricted random-oligonucleotide cassette mutagenesis experiment; Δ in the residue 2 column indicates a deletion of codons 1 and 2. The order in which the variants are listed has no particular significance. For details of the mutagenesis experiments, see (*5*).

Residue number					
1	2	3	4	5	6
Functional variants					
Ser	Thr	Lys	Lys	Lys	Pro
Ala	Ala	Arg		Arg	Lys
Δ	Δ				Arg
Lys	Lys				Ala
Asn	Ser				Val
Val	Asn				Cys
Thr	Met				Ser
Gly	Tyr				Leu
Arg	His				Gly
Pro	Leu				
Ile	Ile				
His	Val				
Leu	Phe				
	Gln				
	Pro				
	Asp				
	Cys				
Nonfunctional variants					
		Ala	Arg	Ala	Asp
		Met	Ala	Met	Asn
			Met		Glu
					Gln

ognition of the two half-sites (*12, 13*).

As was apparent in the initial structure determination (done at 2.5 Å resolution and room temperature), the arm extends away from the globular part of repressor's NH_2-terminal domain and wraps around the DNA, roughly following the major groove toward the center of the operator site (*3*). Lys^5 and Pro^6 appear to help direct the arm into the major groove: Lys^5 makes hydrogen bond or salt-bridge interactions with the phosphate backbone, and Pro^6 restricts the number of accessible conformational states. Although Pro^6 is not essential, both the β-D-galactosidase repression assay and the gel shift assay show that Pro is better than Ala at this position and that Ala is better than Gly.

The improved electron density obtained in this study has given us a better understanding of the contacts made by Lys^4 and has led us to completely reinterpret the role of Lys^3. The side chain of Lys^4 makes specific contacts to guanines at positions 6 and 7 of the operator (Fig. 3B). The side-chain nitrogen (which is protonated and has three hydrogens at physiological pH) participates in three hydrogen bonds: to the O-6 oxygen of guanine 6 (G6), to the O-6 oxygen of G7, and to the oxygen of the amide side chain of Asn^{55}. The nitrogen of the Asn amide hydrogen bonds to the N-7 nitrogen of G6. This extensive hydrogen-bond network involving Lys^4, two bases in the DNA, and a side chain from the globular part of the protein clearly explains the strict requirement for Lys at position 4 (*14*).

Lys^3 also hydrogen bonds to two guanines in the major groove: to G8 and to G9, which is on the opposite strand at the center of the operator site (Fig. 3B). Since earlier models had either the main chain or side chain of Thr^2 at this position, we performed two experiments to substantiate this new model and to provide evidence for this structure in solution. First, we found that protein expressed from a gene with a deletion of codons 1 and 2 fully protects G8 from methylation by dimethysulfate, whereas there is no protection if the protein lacks residues 1, 2, and 3 (*4, 15*). This result is consistent with a direct contact between Lys^3 and G8. Second, competition experiments in which a gel-shift assay was used indicate that DNA binding by wild-type protein (which has Lys at position 3) is much more sensitive to a GC → AT change at base pair 8 than is the Ala^3 mutant protein (*16*). Again, this result is consistent with the direct Lys^3-G8 contact we observe crystallographically (*17*).

The high-resolution structure also reveals a direct contact between the backbone carbonyl for Lys^3 and the extracyclic amine of C8 (Fig. 3B). Thus, the side chain of Lys^3 interacts with the guanine of base pair 8 while the backbone of Lys^3 interacts with the cytosine of the same base pair. To our knowledge, a contact between the polypeptide backbone and a base in the major groove of DNA has not been reported previously. In addition, there appears to be a water-mediated hydrogen bond between the backbone amide of Lys^5 and a phosphate.

Because the arm makes specific contacts to base pairs 6, 7, 8, and 9 on one half of the O_L1 site, it may be appropriate to revise the traditional sequence alignment of λ operator sites (*18*). If the alignment of O_L2 is reversed, as it has been in Fig. 4, we find that the consensus half-site of each of the six λ operators perfectly matches these four base pairs (with the exception of a mismatch at base pair 9 in one operator) (*19*). In this new alignment, the consensus half-sites are also identical at base pairs 4 and 5, which are known to be energetically important residues on the consensus half-site of O_R1 (*12*). The sequence conservation of the bases contacted by the arm suggests that the asymmetry of arm binding evident in the O_L1 complex structure also occurs with the other operators, and that the structure of the consensus half-site arm is likely to be similar for each repressor-operator complex. For similar reasons, we expect the nonconsensus half-site arm for each repressor-operator complex to be less well ordered and less important energetically than the consensus half-site arm.

The structural and genetic studies reported here provide a clear picture of how a

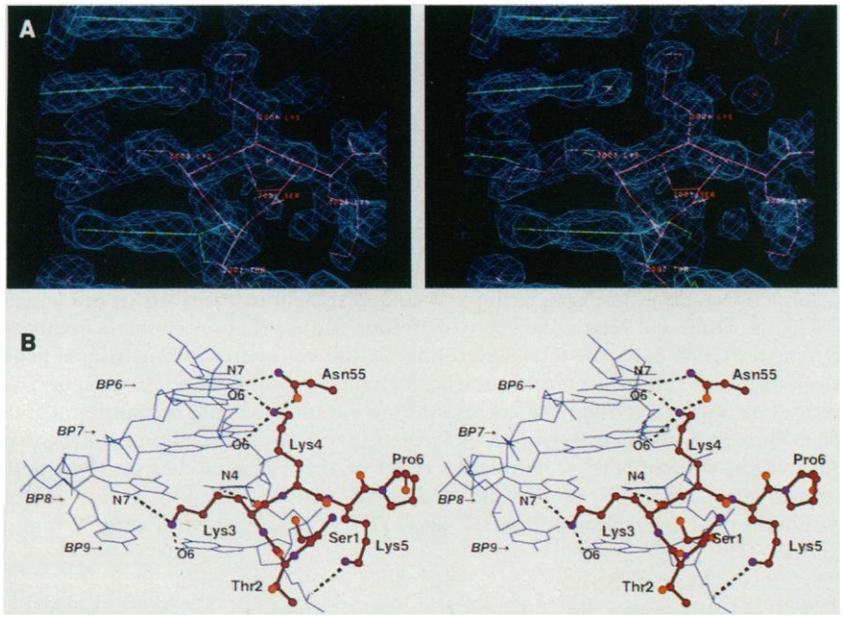

Fig. 3. (**A**) A stereo diagram of the calculated electron density for a $2F_o - F_c$ "omit" map in the region of the NH_2-terminal arm (contoured at 1σ). In order to ensure a rigorous test of the model, residues 1 to 6 were omitted from the structure and the rest of the atoms were subjected to molecular dynamics simulated annealing in XPLOR (*20*) before calculating the map. The residues of the arm are shown in red, and the DNA is in green. The orientation is similar to that shown in (B). Because of the poor electron density for Ser^1 and Thr^2, the conformations of these residues should be considered tentative. (**B**) Stereo diagram showing the NH_2-terminal arm that contacts the consensus half-site. Base pairs 6 to 9 on the consensus half-site of O_L1 are shown in blue. Residues 1 to 6 are labeled; C atoms are in red, N atoms are in purple, and O atoms are in orange. Dashed black lines indicate hydrogen bonds between the protein and the DNA or between side chains. Lys^3 and Lys^4 contact guanines in the major groove of the operator site; Lys^5 contacts a phosphate group. The hydrogen bond between the peptide backbone and C8 is also shown.

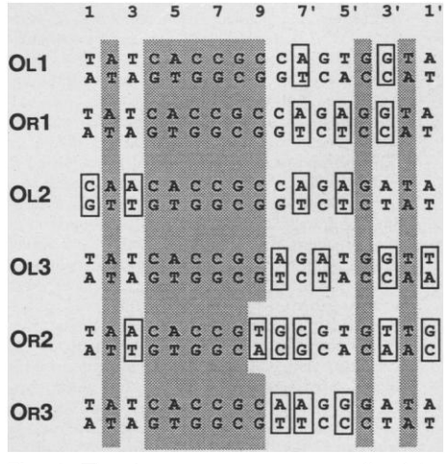

Fig. 4. The six λ operator sequences listed in order of affinity of repressor binding. Nonconsensus base pairs are shown in boxes. Highly conserved base pairs are highlighted in gray. The O_L2 sequence has been inverted relative to its conventional orientation to emphasize the highly conserved region at positions 4 to 9 on the consensus half of the operator.

particular flexible segment can contribute to site-specific DNA binding. The arm clearly has an important role in the λ repressor-operator interactions. The extent to which such segments are used in DNA binding may be underestimated, since some flexible segments may lack the sequence similarities that facilitate the study of well-structured DNA binding motifs.

REFERENCES AND NOTES

1. M. A. Weiss, J. L. Eliason, D. J. States, *Proc. Natl. Acad. Sci. U.S.A.* **81**, 6019 (1984).
2. C. O. Pabo *et al.*, *Nature* **298**, 441 (1982).
3. S. R. Jordan and C. O. Pabo, *Science* **242**, 893 (1988).
4. J. L. Eliason, M. A. Weiss, M. Ptashne, *Proc. Natl. Acad. Sci. U.S.A.* **82**, 2339 (1985).
5. Plasmid pWL103 expresses a synthetic gene for the NH_2-terminal 102 amino acids of repressor and contains a number of unique restriction sites (W. Lim and R. T. Sauer, personal communication). The plasmid was digested with Nco I (which overlaps the initiation codon), and the overhangs were filled in enzymatically. This procedure generates a blunt-ended ATG initiation codon. The DNA was then cut with BssH II (which overlaps codons 16 and 17), and the nearly full-length linear plasmid DNA was gel purified. Oligonucleotides corresponding to the coding strand of repressor from codon 1 through the bases complementary to the BssH II overhang were synthesized. For each of the six codons subjected to mutagenesis, a separate synthesis was done. At the codon of interest the oligonucleotide was synthesized with an equimolar mixture of all four nucleotides at the first two positions and C or G at the third position. A complementary primer was synthesized and annealed such that a BssH II overhang was generated at the 3′ end of the mutant strand. Enzymatic extension with the Klenow fragment of DNA Pol1 generated a pool of fragments that were then ligated into the vector, and the ligation mixture was used to transform *Escherichia coli* strain X90. Selection of functional repressor variants was performed by plating transformation mixes on plates containing ampicillin (100 mg/ml) and approximately 10^9 phage KH54, a mutant of λ that has a deletion covering the gene for repressor. In order to ensure that sampling errors were not responsible for the observed preferences at codons 3, 4, and 5, sequences were obtained from each of at least seven independent transformations at each codon. The construction of the specifically introduced nonfunctional mutants was done in a similar fashion, except that both strands were chemically synthesized with the desired mutation.
6. Although we find a number of functional substitutions at residue 6, substitutions to Asp, Asn, Glu, or Gln at this position are inactive and are expressed at undetectably low levels. These variants were identified by sequence screening of a restricted codon-directed mutagenesis experiment. This second round of mutagenesis was done because we were concerned that some problem in the oligonucleotide synthesis might have led to an underrepresentation of these codons. For this experiment, the mutagenic oligonucleotide was synthesized with the following mixture of nucleotides at codon 6: (CAG)-(A)-(GC). The Asn, Asp, Glu, and Gln substitutions were identified by sequence analysis. In pulse-chase labeling experiments, we did not observe any incorporation of radiolabeled methionine into these mutant proteins. Under the same conditions, the wild-type protein is the major radiolabeled species and is stable for at least 2 hours.
7. The deletion of codon 1 and the deletion of codons 1 and 2 probably arose as a result of oligonucleotides which were prematurely terminated during synthesis (by three and six nucleotides, respectively).
8. B. J. Meyer, R. Maurer, M. Ptashne, *J. Mol. Biol.* **139**, 163 (1980).
9. The ^{32}P end-labeled binding site consisted of the 17-bp O_L1 site in the center of a 50-bp double-stranded oligonucleotide. The assay mix was 100 mM tris (pH 7.8), 50 mM KCl, 2 mM $CaCl_2$, 5% glycerol, sonicated calf thymus DNA (40 mg/ml), $<10^{-11}$ M radiolabeled DNA, and varying amounts of repressor in a final volume of 20 ml. Samples were allowed to equilibrate for 30 min before loading onto the gel (5% polyacrylamide, 7 mM tris, 3 mM sodium acetate, pH 8.0). Prior to loading, the gel was pre-run with circulation of the gel buffer for ~1 hour. The samples were run for ~1 hour at 100 V.
10. Final crystallographic *R* factor of 0.189 for 25,252 reflections between 8.0 and 1.8 Å; root-mean-square deviation from ideal bond lengths of 0.020 Å and for bond angles of 2.1°. Details to be published elsewhere (L. J. Beamer and C. O. Pabo, manuscript in preparation).
11. Only the arm for the monomer bound to the consensus half-site has clear electron density. The arm appears to be disordered on the nonconsensus half-site, and we note that the contacts that the arm makes with G7 and G9 (the central base pair) are not available for the arm on the nonconsensus half-site. These differences in the consensus and nonconsensus half-sites are fully consistent with the differences in chemical protection patterns [R. T. Sauer, C. O. Pabo, B. J. Meyer, K. C. Backman, M. Ptashne, *Nature* **279**, 396 (1979)] and the differences in the contributions these regions make to the binding energy (*12*).
12. A. Sarai and Y. Takeda, *Proc. Natl. Acad. Sci. U.S.A.* **86**, 6513 (1989).
13. A. Hochschild, J. Douhan III, M. Ptashne, *Cell* **47**, 807 (1986).
14. In the partially refined 2.5 Å model, the Lys^4 side chain appeared to only contact base pair 6. The higher resolution structure described here shows unambiguously that the amine of Lys^4 makes contacts to both base pair 6 and to base pair 7.
15. C. O. Pabo and M. Lewis, *Nature* **298**, 443 (1982).
16. Gel shift experiments were done essentially as described in (*9*) except that various amounts of unlabeled competitor oligonucleotides were used in addition to the large amount of calf thymus DNA. The poor binding by the Ala^3 protein complicates the interpretation of this experiment somewhat, since much higher quantities of protein are required. For this protein, the dissociation constant for the cooperative nonspecific binding adjacent to a specific site appears to be close to that for binding to the specific site. The additional higher level bands seen in the gel shift experiment due to this binding are competed away equally well by O_L1 and by the operator site with the GC to AT change at base pair 8. The amount of operator DNA required to compete away these bands is much smaller on a mass basis than the amount of calf thymus DNA present, and addition of the equivalent mass of extra calf thymus DNA does not affect the gel shift pattern. In the case of wild-type protein, O_L1 is approximately a 15-fold better competitor than is the GC8 to AT8 operator variant.
17. Some of the features of this high-resolution structure that were not apparent at lower resolution were correctly predicted by Sarai and Takeda (*12*). By attempting to reconcile the 2.5 Å resolution structure and the results of binding studies to variant operator sequences, these authors correctly anticipated that Lys^4 contacts G7 (in addition to G6) and suggested that Lys^3 contacts G8.
18. M. Ptashne, *A Genetic Switch* (Cell/Blackwell, Cambridge, MA, 1986).
19. L. J. Beamer, thesis, Johns Hopkins University, Baltimore, MD (1991).
20. A. T. Brünger, J. Kuriyan, M. Karplus, *Science* **235**, 458 (1987).

21. We thank T. Steitz for use of the data collection facilities at Yale University and members of his laboratory, J. Friedman and M. Rould for help with data collection and processing, the National Cancer Institute for use of the CRAY supercomputer, and C. Wendling and A. Collector for oligonucleotide synthesis. Supported by NIH grant no. GM31471 and the Howard Hughes Medical Institute.

8 May 1991; accepted 23 July 1991

Structure and Stability of X·G·C Mismatches in the Third Strand of Intramolecular Triplexes

ROMÁN F. MACAYA, DARA E. GILBERT, SHIVA MALEK, JANET S. SINSHEIMER, JULI FEIGON

Intramolecular DNA triplexes that contain eight base triplets formed from the folding of a single DNA strand tolerate a single X·G·C mismatch in the third strand at acidic pH. The structure and relative stability of all four triplets that are possible involving a G·C Watson-Crick base pair were determined with one- and two-dimensional proton nuclear magnetic resonance techniques. Triplexes containing A·G·C, G·G·C, or T·G·C triplets were less stable than the corresponding parent molecule containing a C·G·C triplet. However, all mismatched bases formed specific hydrogen bonds in the major groove of the double helix. The relative effect of these mismatches on the stability of the triplex differs from the effect assayed (under different conditions) by two-dimensional gel electrophoresis and DNA cleavage with oligonucleotide EDTA·Fe(II).

TRIPLE HELICAL NUCLEIC ACID structures formed from one homopurine and two homopyrimidine RNA sequences were proposed more than 30 years ago (*1*). Recent evidence indicates that such DNA sequences may also fold back on themselves to form intramolecular triplexes, termed H-DNA (*2*), when contained in supercoiled plasmids; moreover, these structures may be relevant in vivo (*3*). Triplex formation is currently being widely investigated because of potential therapeutic applications in the specific inhibition of transcription (*4*) and for use in chromosome mapping (*5*). Proposed models for pyrimidine·purine·pyrimidine DNA triplexes have the second pyrimidine strand Hoogsteen base-paired to purines in the

Department of Chemistry and Biochemistry and Molecular Biology Institute, University of California, Los Angeles, CA 90024.