Volume 8 Number 3 1980 Nucleic Acids Research

# Sequence-specific interactions of the tight-binding I12-X86 lac repressor with non-operator DNA

Albert Schmitz and David J.Galas

Departement de Biologie Moléculaire, Université de Genève, CH-1211, Genève 4, Switzerland

Received 11 December 1979

## ABSTRACT

The tight-binding I12-X86 lac repressor binds to non-operator DNA in a sequence-specific fashion. Using the DNA of the E. coli I gene we have investigated these sequence-specific interactions and compared them to the operator binding of wild-type repressor. The specific, non-operator DNA interactions are sensitive to the inducer IPTG. One strong binding site in the I gene DNA was found to be one of two expected on the basis of their homology with the lac operator. The binding of I12-X86 repressor to this site was visualized using the footprinting technique, and found to be consistent with an operator-like binding configuration. The protection pattern extends into an adjacent sequence suggesting that two repressor tetramers are bound in tandem.

## INTRODUCTION

The highly specific binding of the E. coli lac repressor to the operator has been extensively studied by both biochemical and genetic methods (1-6). The discovery of altered repressor molecules with increased affinity for DNA (8, 13) together with the development of the footprinting method (7), to define protein binding sites on DNA, makes it possible now to extend the study of repressor-DNA interactions to a very large range of DNA sequences.

Previous studies have defined missense mutations in the I gene of E. coli that result in the synthesis of a repressor with a highly increased affinity for operator and non-operator DNA, as compared to wildtype repressor. The X86 mutation results in the replacement of the amino acid serine 61 in the repressor

polypeptide chain by leucine (9, 12), and the I12 mutation changes proline 3 to tyrosine (13). Both altered repressors exhibit a 50 to 100-fold increase in affinity for the operator, and a concomitant increase in affinity for non-operator DNA (10, 11, 13). The repressor synthesized by a strain carrying the double mutation I12-X86 (I12-X86 repressor) shows an increase in operator affinity of about $10^4$-fold. Its IPTG (isopropyl-β, D-thiogalactosid) binding affinity and aggregation properties, however, are unaltered by these amino acid substitutions (13).

Two properties of the I12-X86 repressor make its interaction with non-operator DNA a useful model system. First, the DNA binding seems to be specific for certain sequences and is sensitive to the inducer IPTG. Thus, the binding is similar to the interaction of wildtype repressor with the operator. Since the I12-X86 repressor can bind to specific sites on non-operator DNA, to what extent do these sites resemble the lac operator ? Two operator-like sequences have been reported which bind repressor specifically (3, 14, 15), and do resemble the operator: one in the Z gene and one just following the I gene. In an arbitrary sample of DNA sequences we might expect then that those sequences that best "resemble" the operator itself would have the highest affinity for the repressor. The extent to which this reasonable, but not compelling, expectation is true is what we wish to determine. The degree of resemblance to the operator of non-operator binding sites will help to elucidate the basis for the sequence-recognition capacity of the lac repressor.

## MATERIALS AND METHODS

Desoxyribonuclease I (DNase I) was purchased from Worthington Biochemical Corp., the restriction enzymes HpaII and MboII from New England BioLabs and [$^{14}$C] isopropyl-β, D-thiogalactoside (IPTG, 25 mCi/mmol) from Schwarz/Mann. The restriction enzymes HindII/III were purified as described in reference 20.

### Preparation of DNA

The source of DNA for the repressor binding studies was the

plasmid pMC1 (16, 17). This plasmid is a pMB9 carrying a 2800 base-pair insertion of E. coli DNA containing the I gene. Plasmid DNA was prepared according to a published procedure (13). After digestion of pMC1 DNA with the restriction enzymes Hind II and Hind III, a 935 base-pair long fragment of I gene DNA was isolated by gel-electrophoresis on a 6% polyacrylamide gel. This fragment carries part of the I gene promoter and most of the structural I gene DNA (16, 18, see figure 1).

Labelled Fragments for Filter Binding Studies

The restriction enzyme HpaII cleaves the 935 fragment into 5 fragments which are 49, 67, 78, 108, 241 and 392 nucleotides long (counted on the upper DNA strand as shown in figure 1). The fragments were end-labelled as described in reference 21 and used for nitrocellulose filter-binding assays according to Riggs et al (22, 23). The binding buffer contained 10 mM Tris-HCl (pH 7.4), 10 mM Mg-acetate, 0.1 mM EDTA, 0.1 mM DTT and 10 mM KCl, unless otherwise indicated. The reaction volume was 0.35 ml.

Repressor

The I12-X86 repressor was purified as described in reference 13 and was greater than 85% pure as judged by SDS gel electrophoresis (24).

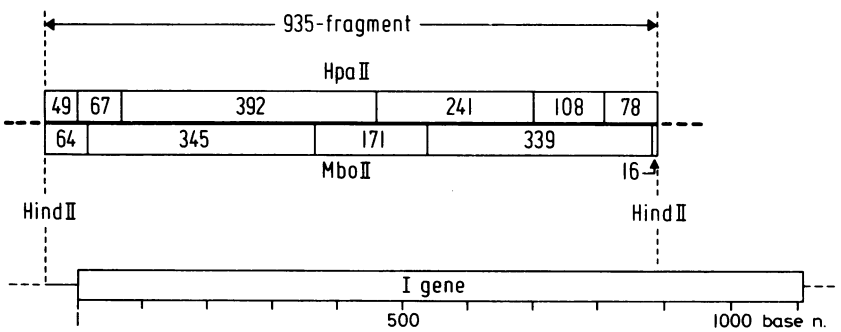


Figure 1. Characterization of the 935 fragment of I gene DNA. The 935 fragment is flanked by HindII restriction sites. The number of nucleotides in the upper DNA strand of the fragments produced by restriction nucleases HpaII and MboII are indicated in the boxes (for full sequence and restriction map see ref. 18).

Labelled DNA Fragments for the Footprinting Experiments

The 78 nucleotide fragment contains the I gene promoter-distal end of the 935 fragment (fig. 1). The 935 fragment was end-labelled on the 5' ends with $^{32}P$ using polynucleotide kinase (21), and then cleaved with HpaII. The radioactively labelled 78 fragment isolated from this reaction mixture carries the label on the lower strand (fig. 1). End-labelling the isolated 78 fragment, followed by cleavage with MboII yields a 62 nucleotide fragment labelled on the upper strand, as shown in figure 1. These two fragments, labelled on opposite strands, were used in the sequencing and footprinting reactions, to permit the protection pattern to be seen on both strands. The footprinting reactions were carried out as described in reference 7. The resolution of the fragments on denaturating polyacrylamide gels was done in parallel with fragments cleaved chemically in standard sequencing reactions to allow the assignment of the DNaseI cleavage sites*.

Computer Analysis

The searches for certain sequence homologies in the I gene were carried out using a set of interactive programs written in BASIC for the HP-9845 computer. The I gene sequence was taken from reference 18, entered into the computer and checked visually for accuracy three independent times. As a further check the sequence was then translated and compared with the repressor amino-acid sequence.

RESULTS

Operator Homologies in the I Gene

From previous studies of the lac repressor-operator system in E. coli it seems likely that a major fraction of the sequence-specific interactions take place within the central 13 nucleotides of

*In reference 28 we incorrectly stated that the 3'-phosphate group present on the ends of the chemically cleaved fragments and not on the DNaseI-cleaved fragments causes a decrease in the relative electrophoretic mobility of the fragment. In fact, the mobility is increased by this added, negatively charged phosphate group (35).

the operator, extending 6 nucleotides to either side of the center of symmetry (6, 15, 25). This segment includes all of the known sites of operator constitutive ($O^c$) mutations (3, 26). We scanned the I gene DNA sequence by computer, from base -50 in the I gene promoter to the end of the structural gene, for homologies with the operator sequence in both orientations. Table I shows the most extensive homologies found: 8 or 9 out of 13. They occur at about the frequency expected in random sequences. Below table I the positions of these sequences in the structural I gene are indi-

| | Operator position | 5'-GTGAGCGGATAAC-3' | | Homology | Homology w/$O^{c,s}$ |
|---|---|---|---|---|---|
| 1 | 387 | GTGGGCTGATCAT | (U) | 9 | 5 |
| 2 | 657 | TTCAGCCGATAGC | (U) | 9 | 5 |
| 3 | 820 | TGGTGCGGATATC | (U) | 9 | 6 |
| 4 | 43 | GACATCGTATAAC | (L) | 9 | 6 |
| 5 | 23 | GTGAATGTGAAAC | (U) | 8 | 5 |
| 6 | 480 | CTGACCAGACACC | (U) | 8 | 4 |
| 7 | 965 | GTGAAGGGCAATC | (U) | 8 | 5 |
| 8 | 998 | GTGAAAAGAAAAA | (U) | 8 | 3 |
| 9 | 402 | TCCAGCGGATAGT | (L) | 8 | 5 |
| 10 | 503 | GGGAGAAAATAAT | (L) | 8 | 5 |
| 11 | 929 | GCAAGCGGTCCAC | (L) | 8 | 6 |

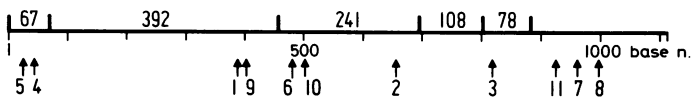


Table I. Homologous sequences with the operator in the lac I gene. The central 13 positions of the lac operator, shown at the top, were found to be partially homologous to many sequences in I gene. The sequences with 8 or more positions matching are shown here. The symmetry of the operator is indicated by the boxes. The dots show the positions at which $O^c$ mutations have been found. The arrow indicates the center of symmetry. The position of the I gene sequence is numbered by the left-most position of the sequence as numbered by Farabaugh (1978), regardless of which strand is homologous to the upper strand of the operator. The U or L following the sequence indicates whether the homology was found on the upper (U) or lower (L) strand. The extent of the homology is indicated in the next column, and the extent of match with the $O^c$ positions is indicated in the final column. Below the table the positions of the homologies within the I gene are shown. The gene is diagrammed as in figure 1.

cated (there are none in the I gene promoter). Of particular interest are the sequences numbered 3 and 4, which match the operator at 9 out of 13 positions, and include matches in 6 out of 8 $O^c$ positions in the gene. These sequences fall, quite conveniently, on two small restriction fragments of about the same size: 78 and 67 nucleotides long. We focussed special attention in this study on these two sites.

Since there are weak interaction sites for repressor outside of the central 13 nucleotides (4, 5), whose presence in a potential binding site could compensate for the absence of extensive homology within, we also scanned for homologies to the central 21 nucleotides of the operator sequence. The largest homology to this sequence is 11 out of the 21 positions. This occurs at 7 sites in the I gene. Three of these 7 are also represented in the homologous sequences of table I : they are numbered 5, 6 and 10. It is possible, on the other hand, that fewer selected nucleotides are required to specify a binding site of a reasonable strength. With this mind we looked for homologies to the $O^c$ positions in the operator, that is, to the sequence 5'-GNGAGCGNNTNNC-3'. There are 10 sites in the I gene that match this sequence in 6 out of the 8 specified positions and no more extensive matches. Of this set only 3 are represented in table I : numbers 3, 4 and 11. A similar analysis shows that for the central 9 positions of the operator there are only two that match in 8 positions, one in 7 positions and 24 in 6 positions. The two 8 out of 9 matches are numbered 3 and 9 in table I, the 7 out of 9 is numbered 2. As shown in a later section, the I12-X86 repressor binds with high affinity to DNA sequence number 3 in table I, while under the same conditions it binds very poorly to sequence number 4.

In an attempt to define the sequence properties responsible for the affinity of sequence 3 for the repressor, we scanned for homology to sequence 3 in the rest of the I gene DNA. There is a sequence within the 108 nucleotide fragment (see fig. 1) that matches sequence number 3 in 9 of its 13 positions, but matches operator in only 4 positions. This sequence binds repressor poorly (see next section).

Filter Binding Experiments

Most of the I gene is contained within the 935 nucleotide HindII/III fragment, as shown in figure 1. To compare the repressor affinity for fragments containing the various candidate sequences, we measured the retention of the Hpa II fragments of the 935 fragment on nitrocellulose filters by the I12-X86 repressor. Figure 2 shows the results of such an experiment, carried out at high salt concentration (80 mM KCl), where the effect of increasing repressor concentration is demonstrated.

Repressor-bound fragments containing high affinity sites are identified by their gradual disappearance from the filter run-through since the repressor preferentially retains these fragments on the filter. The autoradiograph in figure 2A, and especially the scans in figure 2B, indeed demonstrate that the repressor preferentially retains the 78 fragment, and, to a lesser extent, the longer 241 and 392 fragments. This retention cannot be simply a function of the size of the fragments, since the 108 fragment, of intermediate size, is not retained. These experiments also show the retention to be IPTG-sensitive, a characteristic of the wild-type repressor-operator interactions, since the presence of IPTG in the reaction mixture results in the release of otherwise filter-bound fragments. Note that the 67 fragment is not bound by the I12-X86 repressor even though it contains sequence 4, with an operator homology similar to that contained in the 78 fragment. The fact that we have two fragments with similar operator homologies and of similar size, only one of which binds repressor well, makes the detailed characterization of the binding site on the 78 fragment essential. The differences in binding properties could thereby be attributed to a small set of positions (discussion section).

We performed a series of filter-binding experiments to study more accurately the binding of I12-X86 repressor to these fragments. The results of these experiments are shown in figure 3A, where the retention of the radioactively labelled fragments is plotted against repressor concentration in the presence and absence of IPTG. The stronger binding of the I12-X86 repressor to

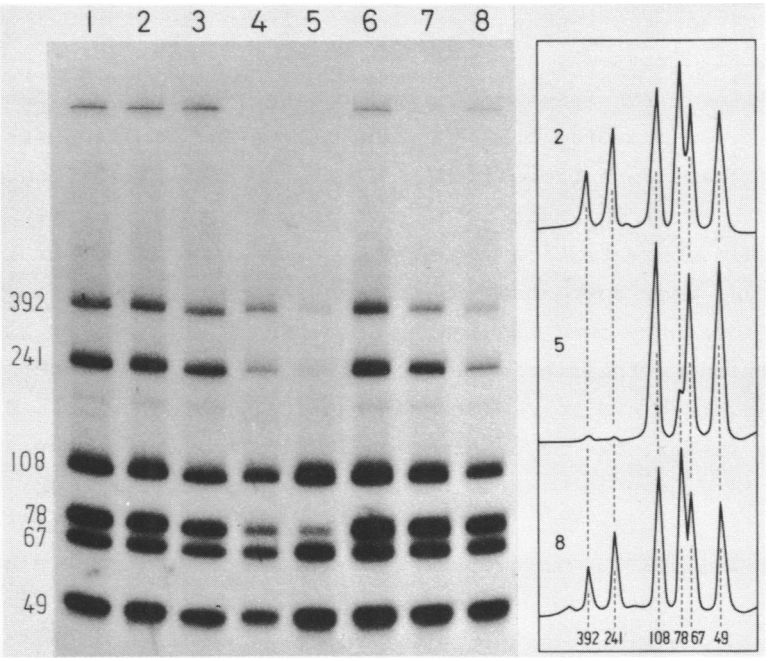


Figure 2. Size-fractionated HpaII fragments not retained by the I12-X86 repressor on filters.
A. Autoradiograph of the polyacrylamide gel. Increasing amounts of repressor were incubated with 5'-end-labelled HpaII fragments of the 935 fragment in the presence and absence of 20 mM IPTG. The DNA in this reaction mixture that passed through nitrocellulose filters was concentrated by ethanol precipitations and size fractionated on a 10% polyacrylamide gel, with subsequent autoradiograph of the gel. Lane 1 : the input DNA (no filtration), lane 2 : no repressor, but with filtration, lane 3 : 3 x $10^{-8}$M repressor, lane 4 : $10^{-7}$M repressor, lane 5 : 3 x $10^{-7}$M repressor, lane 6 : 3 x $10^{-8}$M repressor + IPTG, lane 7 : $10^{-7}$M repressor + IPTG, lane 8 : 3 x $10^{-7}$M repressor + IPTG. The binding buffer contained 80 mM KCl.
B. Densitometric scans of the autoradiograph. 1) Scan of lane 2. 2) Scan of lane 5. 3) Scan of lane 8.

the 78 fragment, as compared with each of the others, manifests itself in the steepness of the slope of the curve for total DNA binding (no IPTG present). The higher binding specificity of the repressor for the 78 fragment can be seen in the increased ratio

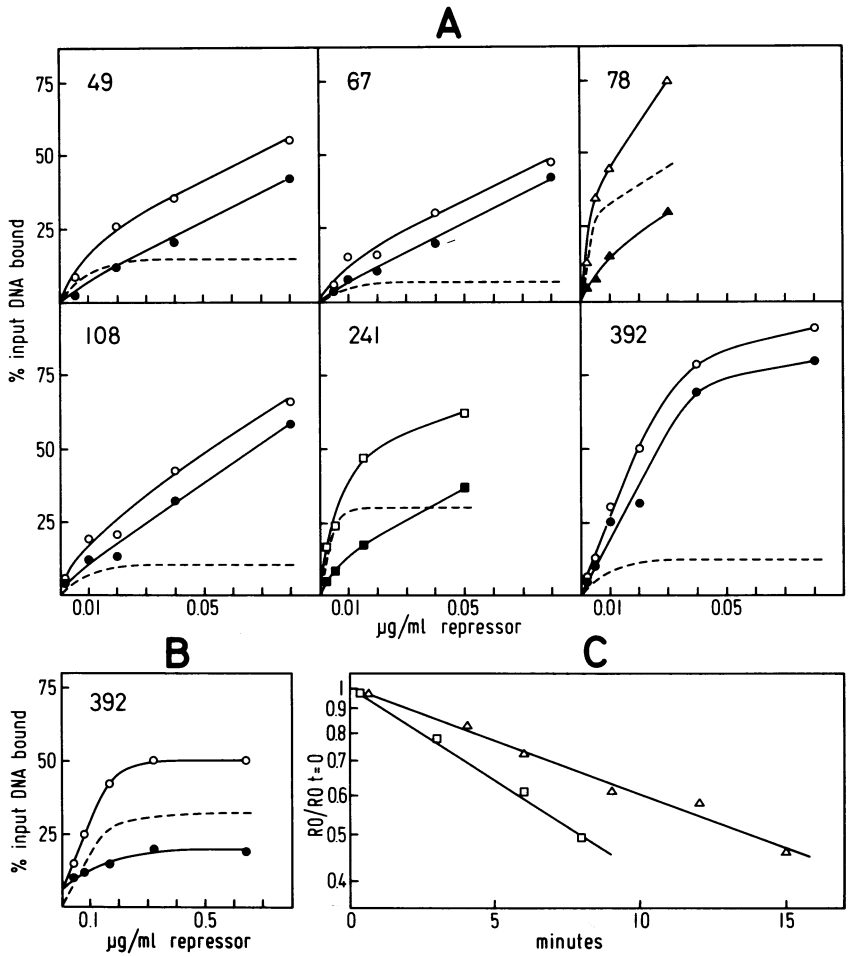


Figure 3. A. DNA binding curves using I12-X86 repressor and the HpaII fragments. The filter binding assays measure the amount of DNA retained on the filter by repressor. The data are plotted here for each of the HpaII fragments as indicated in the upper left of the panals as a function of the repressor concentration, in the presence of IPTG (filled in symbols) and in the absence of IPTG (open symbols). The broken line is the difference between the two. B. The identical experiments as for 3A were performed with the 392 fragment, except that 160 mM KCl was used in the binding buffer. C. The dissociation kinetics of the I12-X86 repressor complex with the HpaII 78 and 241 fragments. The dissociation was measured as described in reference 22. The decay of the IPTG-sensitive DNA complex formed at a repressor concentration of $7 \times 10^{-11}$M and a DNA concentration of $10^{-10}$M was monitored following the addition of lac operator containing λplac5 DNA to a concentration of $2 \times 10^{-10}$M.

of IPTG-sensitive to IPTG-insensitive binding. In fact, these fragments exhibit almost identical binding curves in the presence of IPTG, with the exception of the large fragment 392. The fraction of IPTG-sensitive repressor binding to the 392 fragment can be magnified by increasing the ionic strength of the binding buffer, which is in agreement with the finding that non-operator DNA interactions of wildtype repressor are more salt-sensitive than the operator binding (23, 27). Figure 3B shows the more pronounced IPTG-sensitive binding of the repressor to the 392 fragment in binding buffer containing 160 mM KCL instead of 10 mM KCl as in figure 3A. The small effect of IPTG on the binding of repressor to the 392 fragment at low ionic strength (fig. 3A) is thus seen to be a real effect.

The equilibirum binding experiments can only give a rough estimate of the dissociation constant of these repressor-DNA interactions, since neither the fraction of repressor active in IPTG-sensitive binding, nor the number of binding sites on these DNA fragments are known. To circumvent these complications in estimating affinity constants we measured the rate of dissociation of these repressor-DNA complexes. Figure 3C shows the dissociation rates of the IPTG sensitive complexes formed between the I12-X86 repressor and the 78 and 241 fragments. Both dissociations give straight lines in the semilogarithmic plot, implying that the dissociation is from a single binding site. The half-lives of the complexes are 8 minutes for the 241 fragment and 14 minutes for the 78 fragment. Since the wildtype repressor-operator complex has a similar half-life of about 15 minutes ($K_{diss} = 10^{-13}M$), one can estimate the dissociation constant of the I12-X86 repressor interactions with the 241 and 78 fragment to be in the range of $K_{diss} = 10^{-13}M$ (assuming, that the association rates considering the short fragments used are similar to the one for the wildtype-repressor-operator interactions under these conditions (13, 22, 23).

From the binding data on the fragments we are unable to determine the significance of sequence specificity, since the binding cannot be attributed to a particular site. Is the high-

affinity site on the 78 fragment identical to the site containing the operator homology ? The method of choice to answer this question is the footprinting technique which permits the precise determination of the site on the DNA to which a protein binds (7). A radioactively end-labelled DNA fragment is partially degraded by pancreatic DNaseI in the presence and absence of a DNA binding protein, size fractionated by gel electrophoresis and autoradiographed. The film then reveals a nested set of bands similar to the ones produced by the DNA sequencing reactions (21). A protein bound to the DNA supresses DNase cutting between the bases it shields and thereby delineates its binding site. Figure 4 shows autoradiographs of gels which reveal the footprint of the I12-X86 repressor on both strands of the 78 fragment. The ionic strength of the binding buffer was raised by increasing the KCl concentration to 80 mM,in contrast to the experiments shown in figure 3A, to reduce non-specific repressor-DNA interactions which might otherwise obscure the strong repressor binding sites. High concentrations of wildtype or I12-X86 repressor can protect DNA completely against DNaseI cleavage (data not shown).

In figure 5B a diagrammatic presentation of these results points out that the repressor does interact strongly with the DNA in the predicted region. The homologous sequence itself is protected on both strands by repressor. The protection, however, extends further into the fragment. A comparison with the footprint of wildtype repressor on the operator (fig. 5A), which is identical with the footprint of the I12-X86 repressor on operator (28) suggests that the extensive protection pattern is probably not caused by a single repressor tetramer alone. Although we were unable to define precisely the end of the protection on one strand because of a lack of DNaseI cleavage in this region, it is clear that the protection pattern is very similar to the footprint on operator DNA, both with respect to the length of the double-strand protection and the position of the repressor-induced DNase cleavage enhancements, indicated with arrows. The residual protection on the lower strand could be caused by another repressor molecule binding more weakly to this neigh-

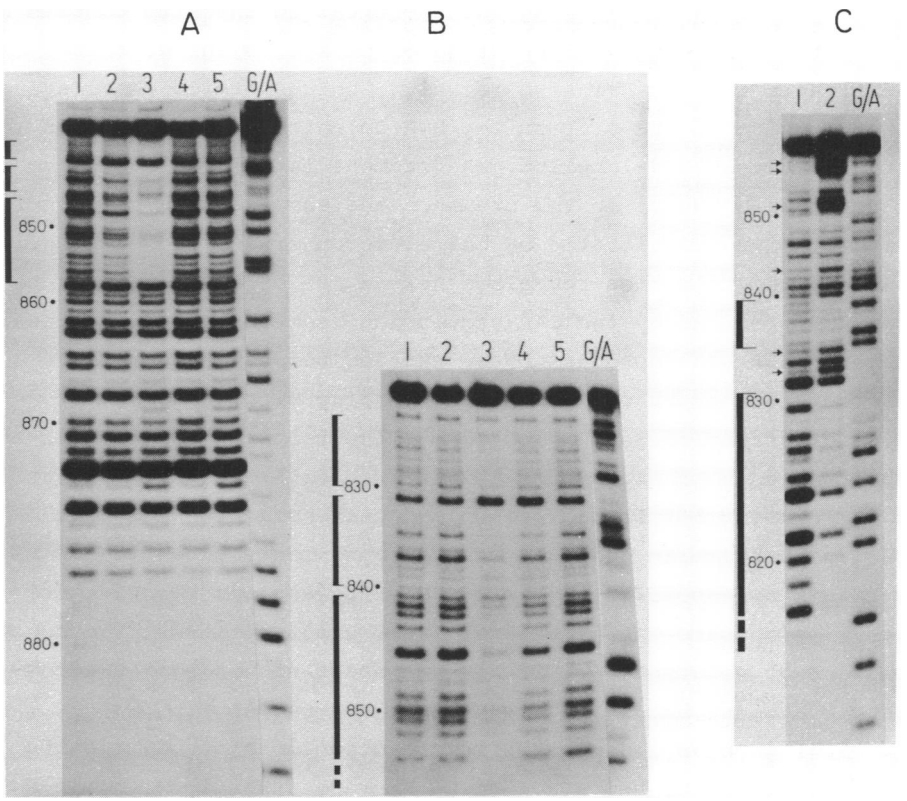


Figure 4. Footprints of the I12-X86 repressor on non-operator DNA. Footprinting reactions were carried out as described in reference 7. Gel electrophoresis of the partial DNaseI digests was on 12% denaturing polyacrylamide gel (21). Sequencing reactions were done as in reference 21. The brackets on the side of the gels indicate the regions of DNA protected by repressor against DNaseI attack. The numbering of the base-pairs is as in reference 18 : Fragments resulting from the removal of a base are given the number of that base. The binding buffer contained 80 mM KCl.
A. Footprint of I12-X86 repressor on the HpaII 78 fragment end-labelled on the lower DNA strand (see fig. 1). Lane 1 : partial DNaseI digest of the 78 fragment alone, lane 2 : $2.7 \times 10^{-8}$ M repressor, lane 3 : $8 \times 10^{-8}$ M repressor, lane 4 : $2.7 \times 10^{-8}$M repressor + IPTG, lane 5 : $8 \times 10^{-8}$M repressor + IPTG, lane 6 : sequencing reactions G > A.
B. Same reactions as in A., but electrophoresed longer for better resolution of the long DNA fragments.
C. Footprint of the I12-X86 repressor on the 62 fragment (78 fragment cut with Mbo II) 5'-end-labelled on the upper DNA strand. Lane 1 : partial DNaseI digest alone, lane 2 : $8 \times 10^{-8}$ M repressor, lane 3 : sequencing reaction G > A.

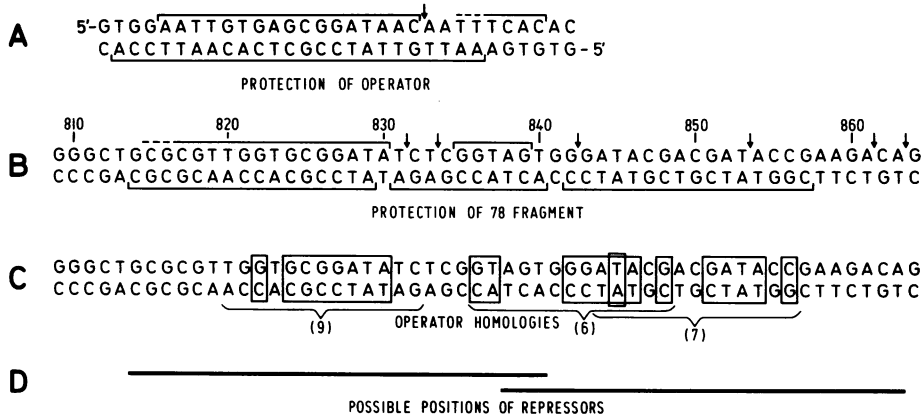


Figure 5. Diagrammatic representation of the footprinting data. In this figure the protection patterns of the footprint experiments are represented along with the footprint of wildtype repressor on the lac operator. The operator pattern is shown in A. The convention adopted here is that a nucleotide is considered protected, if the phosphodiester linkage to the 5' side of a base is shielded by repressor against DNaseI attack. The arrows indicate enhanced cutting by DNaseI, as compared to the cutting in the absence of repressor.
B. The protection pattern shown in figure 4 is diagrammed here. The numbers above the sequence are the position in the I gene sequence according to reference 18.
C. The same sequence is shown here with the homologies to the operator indicated. The brackets indicate the homologies and the number the extent of the match. The operator sequence in 5A is positioned so that the homology in the I gene is aligned with it.
D. Taking the size of the protected operator region as a standard, the possible binding position of two repressor tetramers is indicated here.

bouring sequence and/or being held in place by direct interactions with the more tightly bound repressor. In any case, if the repressor binds in register with the 7 base homology indicated in figure 5C, it would be consistent with the physical extent of the operator protection, as indicated in figure 5A. The two possible positions, indicated in figure 5D, are also consistent with there being direct interactions between the two bound

repressor molecules*. The protection data are displayed on a drawing of B-form DNA in figure 6, together with the protection pattern of wildtype repressor on operator at the same position and orientation as the homologous sequence on the 78 fragment (28), to illustrate the 3-dimensional character on the protection.

The repressor-induced enhancement of DNaseI cutting at certain sites has been observed in several other DNA-protein complexes : RNA polymerase with UV5 lac promotor and repressor with operator (28) and CAP protein with the galactose promotor (29). Two explanations can be offered to account for this phenomenon. Protein-protein interactions between the DNA-bound protein and DNase might result in an increase of local DNase concentration, or the interactions of the DNA-binding protein with the DNA might induce slight changes in DNA conformation which make it a better substrate for DNaseI (31, 32, 33). In light of the results of Wang et al., (30), who showed that lac repressor slightly unwinds the DNA on binding the operator, we prefer the latter explanation, but the question remains unresolved.

## DISCUSSION

In this study we have investigated the interactions of the tight binding I12-X86 repressor with lac I gene DNA of E. coli, in an attempt to elucidate the nature of its binding to non-operator DNA. Filter binding studies using this repressor and six fragments of I gene DNA show that the IPTG-insensitive binding to these fragments is similar (with exception of the 392 fragment, which was bound somewhat tighter, see fig. 3A). The IPTG-sensitive interactions, on the other hand, display a wider spectrum of different affinities, which is independent of the size

*Close inspection of figure 4A reveals bands (at positions 869 and 874) outside the protected region that are enhanced in the presence of repressor, with or without IPTG. This is a repeatable observation which is, at the moment, unexplained. It could be caused by an even more weakly binding repressor molecule, or by a long-distance effect of the strongly-binding repressor on the DNA conformation.

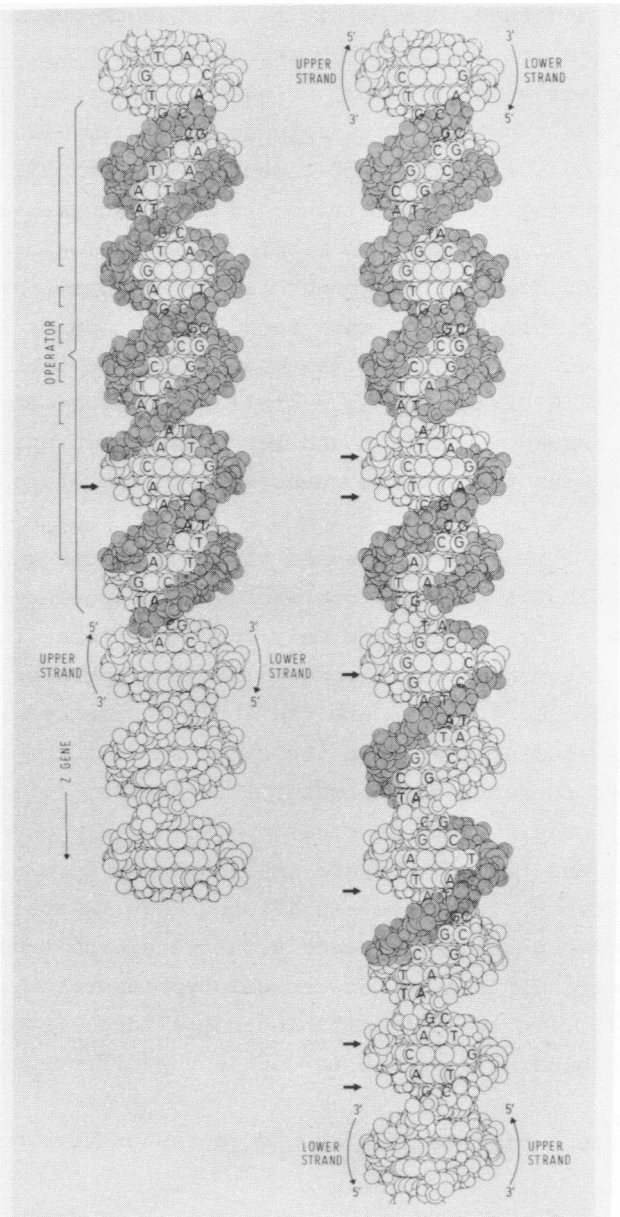


Figure 6. The footprinting data represented on B-form DNA. The patterns of protection on the operator (left) and the I gene site (right) are aligned and oriented so that the homologous bases are in precisely the same positions. The arrows indicate enhanced cutting sites for DNaseI induced by the presence of repressor.

of a fragment and therefore points to a sequence-dependent interaction, analogous to the wildtype repressor-operator binding (fig. 3A). Indeed we could show, using the footprinting technique, that the I12-X86 repressor interacts specifically with a sequence on the 78 nucleotide fragment of I gene DNA in an IPTG-sensitive fashion, protecting both DNA strands in this sequence against DNaseI attack. The preferential binding of the I12-X86 repressor to this DNA sequence had been predicted on the basis of the homologies with the operator sequence found by a computer analysis (table I, sequence 3)+. It is interesting to note that the sequence of the 67 fragment has a similar degree of homology with the operator sequence but is bound very poorly by repressor. The difference between these two sequences could therefore be revealing.

In figure 7 we have diagrammed these sequences and underlined the homologies with the operator. These sequences differ only in a few places. There are only three positions that sequence 3 (in the 78 fragment) has in common with the operator that are missing in sequence 4. These are the three G's marked in figure 7. From previous studies that have indicated some of the specific bases involved in repressor-operator interaction, it is known that all three of these G's are in close contact with the repressor (15, 25), and two of them are implicated in forming bonds with the repressor. It may be significant that in position 14 the difference between sequences 3 and 4, a transition from G:C to A:T, is the very difference between wildtype operator and a known $O^C$ operator (3). It is just possible that a large fraction of the difference in binding constant is due to the difference at position 14.

The apparent binding of a second repressor tetramer next to

+ In fact, the relevant sequence to be used to determine the homologies in a binding site should probably not be the operator itself but some near-operator sequence that reflects the complete set of possible specific interactions with the repressor in effect, a "super operator" sequence.

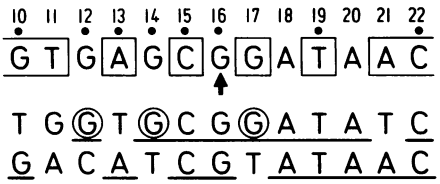


Figure 7. The operator compared with the sequences 3 and 4 from table I. The numbering system for the operator sequence is as in reference 6. The conventions here are the same as for table I. The three G's circled in the sequence 3 are the only three bases that match the operator and are present in sequence 3 while absent in 4.

the one just discussed could be due to the recognition of the operator homologous sequence indicated in figure 5C concomitant with protein-protein interactions between the DNA bound repressors. We were unable to find a repressor concentration that gave protection at only one of the two adjacent sites, up to a salt concentration of 80 mM KCl in the binding buffer. Note, however, that this situation is distinct from the operator-binding protection where we can detect no adjacent protection on either side of the operator (7, 28). It is possible, however, that at high enough repressor concentration a tandem binding mode could be seen even for operator containing DNA.

Although we consider them unlikely, it is worth taking note of several other possible explanations for our observation of the protection pattern. First of all, it is possible that there are more than two repressor tetramers involved in the protection, binding in any of several configurations. Secondly, it cannot be ruled out that the binding of the repressor to the principal site induces a peculiar conformation change in the DNA, that is effective in reducing the DNase-susceptability of the DNA even at a distance of more than 20 base-pairs (see ref. 32).

Our finding that at high repressor concentration (for both wildtype and I12-X86) non operator DNA is protected completely from DNase attack, is in agreement with the results of Zingsheim et al., (34), who showed, by electron-microscopy, complete covering of non-operator DNA by wildtype repressor.

When the ionic strength of the environment is increased, the weaker binding of the I12-X86 repressor to non-operator DNA is preferentially inhibited thereby enhancing the relative effect of the stronger, more specific interactions. These interactions, like the one visualized by footprinting on the 78 fragment, may imitate the specific recognition of the operator by repressor. The binding of repressor to non-operator DNA then is never completely non-specific, though these terms are sometimes used interchangeably. The binding probably consists of a wide spectrum of interactions, including non-specific electrostatic interactions between positively charged groups on the repressor and negatively charged phosphate groups of the DNA, interactions with the bases in the DNA and perhaps specific binding to regions in the DNA with particular conformations.

The experiments reported here, in addition to investigating non-operator binding of the I12-X86 repressor, can be viewed as establishing the validity of a general strategy for studying the sequence-specificity of protein-DNA interactions. With the recent proliferation of DNA sequence information a very large statistical sample of known DNA sequences can, in principle, be presented to the protein of interest. Localization of the binding sites of various strenghts by footprinting could then establish a spectrum of sequences that interact with the protein. Thus the key features of the DNA recognition properties of the protein could be revealed.

ACKNOWLEDGEMENTS

We would like to thank Dr. Jeffrey Miller for his encouragement and for useful discussions throughout the course of this work. This work was supported by a grant from the Swiss National Science Foundation No 3.179.77.

REFERENCES

1. Miller, J.H. (1978) in The Operon, Miller, J.H. and Reznikoff, W.S., Eds., pp. 31-88, Cold Spring Harbor Laboratories, Cold Spring Harbor, New York
2. Bourgeois, S. and Pfahl, M. (1976) in Advances in Protein Chemistry, Anfinsen, C.B., Edsau, J.T. and Richards, F.M.,

Eds., Vol. 30, pp. 1-99, Academic Press, New York
3. Gilbert, W., Gralla, F., Majors, J. and Maxam, A. (1975) in Symposium on Protein-Ligand Interactions, Sund, H. and Blauer, G., Eds., pp. 193-206, Walter de Gruyter, Berlin
4. Ogata, R. and Gilbert, W. (1977) Proc. Nat. Acad. Sci. USA 74, 4973-4976
5. Ogata, R. and Gilbert, W. (1978) Proc. Nat. Acad. Sci. USA 75, 5851-5855
6. Goeddel, D.V., Yansura, D.G. and Caruthers, M.H. (1978) Proc. Nat. Acad. Sci. USA 75, 3578-3582
7. Galas, D. and Schmitz, A. (1978) Nucleic Acids Research 5, 3157-3170
8. Betz, J.L. and Sadler, J.R. (1976) J. Mol. Biol. 105, 293-319
9. Chamness, G.C. and Wilson, C.D. (1970) J. Mol. Biol. 53, 561-565.
10. Jobe, A. and Bourgeois, S. (1972) J. Mol. Biol. 72, 139-152
11. Pfahl, M. (1976) J. Mol. Biol. 106, 857-869
12. Files, J.G. and Weber, K. (1978) J. Mol. Biol. 123, 454-456
13. Schmitz, A., Coulondre, C. and Miller, J. (1978) J. Mol. Biol. 431-456
14. Reznikoff, W.S., Winter, R.B. and Hurley, C.K. (1974) Proc. Nat. Acad. Sci. USA 71, 2314-2318
15. Gilbert, W., Majors, J. and Maxam, A. (1976) in Dahlem Workshop on Chromosomes, pp. 167-178, Abakon Verlagsgesellschaft, Berlin
16. Calos, M.P. (1978) Nature 274, 762-765
17. Calos, M.P., Johnsrud, L. and Miller, J.H. (1978) Cell 13, 411-418
18. Farabaugh, P.J. (1978) Nature 274, 765-770
19. Tanaka, T. and Weisblum, B. (1975). J. Bacteriol. 121, 354-362
20. Bickle, T.A., Pirotta, V. and Imber, R. (1977) Nucleic Acids Research 4, 2561-2572
21. Maxam, A. and Gilbert, W. (1977) Proc. Nat. Acad. Sci. USA 74, 560-564
22. Riggs, A.D., Bourgeois, S. and Cohn, M. (1970a) J. Mol. Biol. 53, 401-417
23. Riggs, A.D., Suzuki, H. and Bourgeois, S. (1970b) J. Mol. Biol. 48, 67-83
24. Laemmli, U.K. (1970) Nature 227, 680-685
25. Yansura, D.G., Goeddel, D.V., Kundu, A. and Caruthers, M.H. (1979) J. Mol. Biol. 133, 117-135
26. Smith, T.F. and Sadler, J.R. (1971) J. Mol. Biol. 59, 273-305
27. Record, M.T., deHaseth, P.L. and Lohman, T.M. (1977) Biochemistry 16, 4791-4796
28. Schmitz, A. and Galas, D. (1979) Nucleic Acids Research 6, 111-137
29. Taniguchi, T., O'Neill, M. and Crombrugghe, B. (1979) Biochemistry 76, 5090-5094
30. Wang, J.C., Barkley, M.D. and Bourgeois, S. (1974) Nature 251, 247-248
31. Klug, A.J., Vismamitra, M.H., Kennard, O., Shakked, Z. and Steitz, T. (1979) J. Mol. Biol. 131, 669-680

32. Sobell, H.M. (1979) in Biological Regulation and Development, Goldberger, R.F. Ed., Vol. 1, pp. 171-199
33. Butler, A.P., Revzin, A. and von Hippel, P.H. (1977) Biochemistry 16, 4757-4768
34. Zingsheim, H.P., Geisler, N., Weber, K. and Mayer, F. (1977) J. Mol. Biol. 115, 565-570
35. Ross, W. and Landy, A. (1979) Cell 18, 297-307