Repressor of Phage 16-3 with Altered Binding Specificity Indicates Spatial Differences in Repressor-Operator Complexes

Szilamér Ferenczi,1 László Orosz,1,2 and Péter P. Papp1*

Institute of Genetics, Agricultural Biotechnology Center, Gödöllő, Szent-Györgyi A. 4., H-2100, Hungary,1 and Department of Genetics, Eötvös Loránd University and Research Group for Molecular Genetics of the Hungarian Academy of Sciences, Pázmány P. Sétány 1/C, H-1117, Budapest, Hungary2

Received 23 September 2005/Accepted 23 November 2005

The C repressor protein of phage 16-3, which is required for establishing and maintaining lysogeny, recognizes structurally different operators which differ by 2 bp in the length of the spacer between the conserved palindromic sequences. A “rotationally flexible protein homodimers” model has been proposed in order to explain the conformational adaptivity of the 16-3 repressor. In this paper, we report on the isolation of a repressor mutant with altered binding specificity which was used to identify a residue-base pair contact and to monitor the spatial relationship of the recognition helix of C repressor to the contacting major groove of DNA within the two kinds of repressor-operator complexes. Our results indicate spatial differences at the interface which may reflect different docking arrangements in recognition of the structurally different operators by the 16-3 repressor.

The regulation of gene expression occurs mostly at the level of transcription through sequence-specific DNA-protein interactions. In several cases, even a single change in the binding sequences results in loss of function due to loss or weakened binding of the protein. Alternatively, examples are also known of when the DNA binding protein has relaxed binding specificity, that is, when several changes at given positions still allow proper binding (12, 23). Only a few naturally existing systems are known where the protein has the ability to bind specifically to sequences with different lengths. The Escherichia coli cyclic AMP receptor protein (CRP) recognizes 16- and 18-bp-long binding sites, in which 6- and 8-bp central spacers, respectively, are bracketed by the recognition sequences. According to Adhya’s “geometric homeostasis” method of resolution in CRP-DNA complexes, a conformational shift from B- to A-form DNA over one helical turn covering the longer spacer allows sequence-specific binding of CRP (13). The CytR repressor recognizes binding sites consisting of two octamer repeats, in direct or inverted orientation, separated by 2 bp. However, in the presence of cyclic AMP-CRP, CytR instead recognizes inverted repeats separated by 10 to 13 bp or direct repeats separated by 1 bp. It was shown that the bases for the structurally different sites were conformational changes within the CytR protein induced by protein-protein interactions between CRP and CytR (14, 21). The CI repressor of E. coli phage 186 was found to recognize two distinct DNA sequences, termed A-type and B-type sites. The A-type binding sites were different in length since half-sites were separated by either 4-bp or 5-bp spacers (26). The binding of 16-3 C repressor to its operators is another example for recognition of structurally different DNA sites by the same protein. In this system, the possibility that either conformational shifts from B- to A-form DNA within the binding sites or interactions of the repressor with another protein is involved in formation of 16-3 repressor-operator complexes was ruled out (20).

Genes, proteins, and chromosomal sites for several functions of the temperate phage 16-3 of Rhizobium meliloti 41 have been studied in detail (1, 2, 5, 8, 9, 11, 19, 22, 24, 25). The most thoroughly analyzed region of the 16-3 genome has been the immC regulatory region which encodes a lambda type CI repressor (16-3 CI repressor), and the cognate c cistron is flanked by two operator regions, O1 and OR (3, 4, 6, 16–18, 20). An interesting feature of the immC region is its partial cross-compatibility with the phage 434 immC region: although the overall sequence homology is not significant (below 15%), the helix-turn-helix motifs of the two repressors (55% identity) and the OR operators are highly similar (4). In both of the 16-3 O1 and OR operator regions, two repressor binding sites were found: the O1-type operators were 12 bp long (5'-ACAA-4 bp-TTGT-3'), while the OR-type operators were 14 bp long (5'-ACAA-6 bp-TTGT-3'). To explain all previous in vivo and in vitro binding data, we proposed the “rotationally flexible protein homodimers” model. Our model hypothesizes alternative interactions for the dimerization domains as the basis for structural flexibility in recognizing the two types of binding sites (20). The model supposes, based on analyses of repressor binding to mutant and truncated versions of both types of operators (for a review, see Fig. 2 in reference 20), that the spatial relationships of the appropriate side chains of the residues in the recognition helix and the contacting bases of the DNA at the interface are the same in both types of repressor-operator complexes. We thought that an altered binding specificity repressor mutant can be used suitably to test this assumption.

Determination of repression levels in vivo. To evaluate the strength of interaction between the wild-type and mutant repressors and the different operators, we used a single-copy reporter system (9). All measurements were done in R. meliloti at 28°C. Derivatives of pGSB1, containing the different oper-
ties (expressed in Miller units) were carried out as described for the no-repressor control) or one of its derivatives containing an OR2 operator. Another plasmid, either pSEM91 (25) (used as a helper) or pPM238 is the source of the repressor.

β-Galactosidase assays and calculations of promoter activities (expressed in Miller units) were carried out as described previously (15). Repression (R) values were calculated using the following equation (9, 20): \( R = 1 - \frac{(\text{promoter activity in the cell when repressor was added from plasmid/promoter activity in the cell without repressor})}{\text{promoter activity in the cell without repressor}} \).

Role of residues of the recognition helix in repressor binding. Plasmids pGSB42 and pGSB62, which contain the sequences of an OR2 operator and an OR3 operator, respectively, between the −35 and −10 regions of a promoter, were constructed as previously described (9). Plasmid pSEM91 was used to express the I6-3 C repressor protein and its mutant derivatives. pPM232 (20) expressed the wild-type repressor (CME). Plasmid pPM238 was created. pPM238 is identical to pPM232, except that MscI and EagI restriction sites, in front of the repressor gene, were introduced by on May 19, 2009

TABLE 1. R values determined by binding of wild-type and mutant repressors to OR2 and OR3 operators

<table>
<thead>
<tr>
<th>Operator</th>
<th>CME</th>
<th>CME/B</th>
<th>CQ37A</th>
<th>CQ38A</th>
<th>CQ40A</th>
<th>CQ41A</th>
<th>CQ42A</th>
<th>CQ43A</th>
</tr>
</thead>
<tbody>
<tr>
<td>OR2</td>
<td>0.80</td>
<td>0.76</td>
<td>0.12</td>
<td>0.30</td>
<td>0.31</td>
<td>0.21</td>
<td>0.34</td>
<td>0.37</td>
</tr>
<tr>
<td>OR3</td>
<td>0.81</td>
<td>0.79</td>
<td>0.13</td>
<td>0.24</td>
<td>0.25</td>
<td>0.23</td>
<td>0.43</td>
<td>0.39</td>
</tr>
</tbody>
</table>

* Residues of the wild-type recognition helix were Gln-Gln-Ile-Asn-Asn-Leu-Phe-Ala. Each value represents the average of three parallel experiments with deviations less than or equal to ±0.02.

b CME is a wild-type repressor at the protein level, produced when the C gene with silent mutations in plasmid pPM238 is the source of the repressor.

CME and mutant repressors to OR2-type and OR3-type operators was measured in vivo, and the calculated repression values are summarized in Table 1. Repression values of CME and CME are basically the same for each operator, indicating that the changes (silent at the protein level) which were introduced to create restriction sites in the C gene had no effect on the repression level. Replacement of the residues that were most likely to make contact with the operators (positions 37, 38, 41, and 42) by alanine resulted in significant loss in repressor binding activity irrespective of the type of operator used. Alaine at positions 40 and 43 also resulted in reduced binding, indicating the importance of the side chains of the original residues, which may be involved in maintaining the overall structure of the repressor in the DNA-protein complex.

Isolation of mutant repressors with altered binding specificities. Plasmid pGSB1 was used to construct different promoter/operator units in front of a promoterless lacZ gene. To isolate altered binding specificity mutant repressors, synthetic oligonucleotides were used to introduce all possible symmetrical changes into an OR2 operator at positions 1 and −1 (OR2−1), 2 and −2 (OR2−2), and 3 and −3 (OR2−3), as well as 4 and −4 (OR2−4) (Table 2). Due to our experimental setup, the R values directly correlate with the in vivo binding efficiencies of the repressors to the operators. Repression values for the binding of the CME repressor to mutant OR2-type operators are listed in Table 3. To correspond the results to the complex functional level of lysogenization, we can consider the pheno-

TABLE 2. Mutant OR2-type operators with symmetrical changes used in binding studies

<table>
<thead>
<tr>
<th>Mutations in the indicated operator*</th>
</tr>
</thead>
<tbody>
<tr>
<td>OR2−1</td>
</tr>
<tr>
<td>GCAATTGTAGTTGC</td>
</tr>
<tr>
<td>CCAATTGTAGTTGC</td>
</tr>
<tr>
<td>TCAATTGTAGTTGC</td>
</tr>
<tr>
<td>OR2−2</td>
</tr>
<tr>
<td>AGAATTGTAGTTCT</td>
</tr>
<tr>
<td>AAAATTGTAGTTTC</td>
</tr>
<tr>
<td>ATAATTGTAGTTAC</td>
</tr>
<tr>
<td>OR2−3</td>
</tr>
<tr>
<td>ACAATTGTAGTTGG</td>
</tr>
<tr>
<td>ACCAATTGTAGTTG</td>
</tr>
<tr>
<td>ACTAATTGTAGTTA</td>
</tr>
<tr>
<td>OR2−4</td>
</tr>
<tr>
<td>AGCAATTGTAGCTGG</td>
</tr>
<tr>
<td>ACCAATTGTAGCTG</td>
</tr>
<tr>
<td>ACACAATTGTAGCTGT</td>
</tr>
</tbody>
</table>

* Underlining indicates bases in the palindromic positions; boldface letters mark symmetrical changes. Numbering of operator positions (OR2): 1/cg ............................................................................................ 0.66
  2/ta ............................................................................................ 0.60
  2/gc ............................................................................................ 0.58
  3/gc ............................................................................................ 0.42
  3/ta ............................................................................................ 0.37
  3/cg ............................................................................................ 0.34
  3/ta ............................................................................................ 0.31
  3/cg ............................................................................................ 0.30
  2/ta ............................................................................................ 0.25
  2/gc ............................................................................................ 0.23
  2/ta ............................................................................................ 0.21
  2/cg ............................................................................................ 0.18
  1/ta ............................................................................................ 0.43
  1/cg ............................................................................................ 0.40
  1/ta ............................................................................................ 0.37
  1/cg ............................................................................................ 0.34
  1/ta ............................................................................................ 0.31
  1/cg ............................................................................................ 0.28
  1/ta ............................................................................................ 0.25
  1/cg ............................................................................................ 0.22
  1/ta ............................................................................................ 0.19
  1/cg ............................................................................................ 0.16
  4/ta ............................................................................................ 0.43
  4/cg ............................................................................................ 0.40
  4/ta ............................................................................................ 0.37
  4/cg ............................................................................................ 0.34
  4/ta ............................................................................................ 0.31
  4/cg ............................................................................................ 0.28
  4/ta ............................................................................................ 0.25
  4/cg ............................................................................................ 0.22
  4/ta ............................................................................................ 0.19
  4/cg ............................................................................................ 0.16

TABLE 3. Binding of CME repressor to mutant OR2-type operators

<table>
<thead>
<tr>
<th>Operator</th>
<th>CME binding</th>
</tr>
</thead>
<tbody>
<tr>
<td>OR2−1</td>
<td>0.32</td>
</tr>
<tr>
<td>OR2−2</td>
<td>0.50</td>
</tr>
<tr>
<td>OR2−3</td>
<td>0.60</td>
</tr>
<tr>
<td>OR2−4</td>
<td>0.39</td>
</tr>
<tr>
<td>OR2−5</td>
<td>0.37</td>
</tr>
<tr>
<td>OR2−6</td>
<td>0.31</td>
</tr>
<tr>
<td>OR2−7</td>
<td>0.73</td>
</tr>
<tr>
<td>OR2−8</td>
<td>0.42</td>
</tr>
<tr>
<td>OR2−9</td>
<td>0.58</td>
</tr>
<tr>
<td>OR2−10</td>
<td>0.45</td>
</tr>
<tr>
<td>OR2−11</td>
<td>0.66</td>
</tr>
<tr>
<td>OR2−12</td>
<td>0.53</td>
</tr>
<tr>
<td>OR2−13</td>
<td>0.81</td>
</tr>
</tbody>
</table>

* Numbers are R values for indicating the strength of the repression and represent the average of the three parallel experiments with deviations less than or equal to ±0.02.
typic expression of two phase 16-3 mutants: (i) the immunity insensitivity of virulent mutant 16-3 O$_R^1-1$, which carries the O$_R^2$ operator, and (ii) the heat inducibility of the 16-3 cti3 mutant, which carries a A allele of the repressor. The phenotypes of the two mutants provide options to relate the R values to functional thresholds. Binding of C$^{wt}$ to the O$_R^2$ operator (contains O$_R^1-1$ mutation in O$_R^2$) resulted in about 70% of the wild-type repression level (i.e., C$^{wt}$ binding to the wild-type O$_R^2$ operator) (20). This result means that when R is $\approx 0.56$, the level of repression would not be enough to complete its function toward lysogenization. The repression values of C$^{153M}$ (encoded by the cti3 allele of the c gene) were 0.66 and 0.69 at 28°C for the O$_L^2$ and O$_R^2$ operators, respectively, and 0.31 at 37°C for both. This result means that when R is $\approx 0.66$, the level of repression is enough to reach the stable lysogenic stage. In the light of these data, C$^{wt}$ recognizes O$_R^2$-1cg and O$_R^2$-4cg mutant operators with near wild-type affinity, while in the rest of the cases, the strength of binding should be below the functional threshold needed for lysogen formation.

Repressor residues at positions 37 and 38 (the first and second positions in the recognition helix) were replaced with different amino acids in an attempt to identify mutations that restore the lost binding of the repressor to any of the mutant O$_R^2$-type operators. Repressors with different substitutions at position 37 were tested for binding to O$_R^2$-1 operators, and mutant repressors with changes at position 38 were tested for binding to O$_R^2$-2 and O$_R^2$-3 operators. Although we have not done a full-scale analysis (data not shown), one mutant repressor has been found to suppress an operator mutation. Repressor C$^{37A}$ binds well to O$_R^2$-1cg (Table 4). Since binding of the C$^{37A}$ repressor to the O$_R^2$ operator is very poor, like binding of C$^{wt}$ to the O$_R^2$-1cg operator (Table 4), we concluded that the C$^{37A}$ repressor had altered binding specificity for the O$_R^2$-type operators.

Identification of a residue-base pair contact in a repressor-O$_R^2$ operator complex. Sequence-specific binding of proteins to DNA depends on proper docking arrangements at the interface. It requires complementarities of the surfaces of the protein and DNA concerning size, shape, and charge. Altered binding specificity mutants have been isolated for several DNA binding proteins. They were used to identify contacts between a residue of the protein and a base pair in DNA even in the absence of the corresponding crystal structures (7, 29).

Cross-reactions between the 16-3 repressor and the 434 operator as well as between the 434 repressor and the O$_L^2$-16-3 operator (4) suggested that the glutamine residue at position 37 of the 16-3 repressor (the first residue of the recognition helix) interacts with A at the first position of the operator half-site. Isolation of the C$^{37A}$ repressor as an altered specificity mutant confirms this prediction, since its binding to the O$_R^2$-1cg mutant operator is as strong as C$^{wt}$ binding to the O$_R^2$ wild-type operator, while binding levels of C$^{wt}$ to O$_R^2$-1cg and C$^{37A}$ to O$_R^2$ were reduced significantly. Substitution of the glutamine by an alanine at the first position in the recognition helix of the 434 repressor (Gln 28 $\rightarrow$ Ala 28) resulted in altered binding specificity. The corresponding mutant 434 operator had a A · T $\rightarrow$ T · A base pair change at the first position of the 16-3 operator half-sites, since the hydrophobic and short side chain of alanine would be unlikely to find a suitable atomic group to make specific contact with the C · G base pair. One possibility is that the new contact is made with an atomic group of the nucleotide other than in the base itself. This could explain a strong binding of C$^{37A}$ repressor to the O$_R^2$-1cg mutant operator. However, C$^{37A}$ repressor has the ability to distinguish specifically between potential binding sites (R values of binding to O$_R^2$-1cg and O$_R^2$-1tet are 0.34 and 0.38, respectively).

Binding of the altered specificity mutant repressor to O$_L^2$-type operators. In addition to identifying a residue-base pair contact, C$^{37A}$ was used to test the spatial arrangements when the 16-3 repressor binds to the structurally different O$_L^2$-type and O$_R^2$-type operators. The basic idea is that if the spatial arrangements are the same in both types of repressor-operator complexes, an altered binding specificity mutant isolated using O$_R^2$-type operators should affect binding to O$_L^2$-type operators in the same way, since the atomic groups required for the new contact are at the same place in both repressor-operator complexes. To assay binding of C$^{37A}$ repressor to O$_L^2$-type operators, its binding to O$_L^2$-1cg and O$_L^2$-1cg (5’-CCAAATTGATTGG-3’) underlined type indicates changes relative to O$_L^2$-1 was tested. Unexpectedly, the C$^{37A}$ repressor, contrary to its binding to the mutant O$_R^2$-1cg repressor, did not bind well to its parallel O$_L^2$ version, the O$_L^2$-1cg mutant (Table 4).

Although it is not clear what the exact molecular basis of the suppressor phenotype is when C$^{37A}$ repressor binds to the O$_R^2$-1cg mutant operator, the atomic groups ensuring contacts between the alanine residue and the C · G base pair in the O$_R^2$-type operator-repressor complex could be out of reach to make contact in the O$_L^2$-type complex. This means that the spatial relationship between the major grooves and the recognition helices at the interface are different in the two kinds of complexes, implying that the 16-3 repressor may use different docking arrangements to recognize O$_L^2$-type and O$_R^2$-type operators.

In the light of our new results, we can extend our previously suggested “rotationally flexible protein homodimers” model by drawing conclusions concerning the binding domains. Now we can say that the spatial arrangements and the conformation of the binding domains of the repressor dimer are somewhat different when a 16-3 repressor forms a complex with an O$_L^2$-type or an O$_R^2$-type operator. However, this finding does not influence the basic characteristic of the model, which is the “rotationally flexibility” of the dimers at the interface of the dimerization.

### Table 4. Binding of C$^{wt}$ and C$^{37A}$ repressors to wild-type and mutant (O$_R^1$-1cg) O$_L^2$- and O$_R^2$-type operators

<table>
<thead>
<tr>
<th>Operator</th>
<th>Binding of indicated repressor$^a$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>C$^{wt}$</td>
</tr>
<tr>
<td>O$_R^2$</td>
<td>0.81</td>
</tr>
<tr>
<td>O$_R^2$-1cg</td>
<td>0.50</td>
</tr>
<tr>
<td>O$_R^2$</td>
<td>0.80</td>
</tr>
<tr>
<td>O$_R^2$-1cg</td>
<td>0.24</td>
</tr>
</tbody>
</table>

$^a$ Numbers are R values for indicating the strength of the repression and represent the average of three parallel experiments with deviations less than or equal to $\pm 0.02$. 

$^n$
We thank Kornélia Szőrás Gál, Magdolna Tóth Pél, and Csilla Sánta Török for excellent technical assistance and Dhruva Chattoraj for discussion and helpful comments on the manuscript.

This work was supported by grants OTKA T 032205 to L.O., T023695 and T032255 to P.P. from the Hungarian Scientific Research Fund; by grants OM 0228/2001, OM 0278/2001 to L.O. and P.P., and OM 0320/2004 to L.O. from the National Research and Development Program NKFP; and by grants MTA/TKI/AKT-F 1999-2001 and MTA/TKI/AKT-F 2003-2006 to L.O. from the Hungarian Academy of Sciences.

REFERENCES


