The EMBO Journal vol.6 no.10 pp.3145–3153, 1987

The interaction of the recognition helix of *lac* repressor with *lac* operator

Norbert Lehming, Jürgen Sartorius, Monika Niemöller, Gabriele Genenger, Brigitte v.Wilcken-Bergmann and Benno Müller-Hill

Institut für Genetik der Universität zu Köln, Weyertal 121, 5000 Köln 41, FRG

Communicated by B.Müller-Hill

We have constructed a system which allows systematic testing of repressor–operator interactions. The system consists of two plasmids. One of them carries a *lac* operon in which *lac* operator has been replaced by a unique restriction site into which synthetic operators can be cloned. The other plasmid carries the gene coding for the repressor, in our case a semi-synthetic *lacI* gene of which parts can be exchanged in a cassette-like manner. A *galE* host allows us to select for mutants which express repressors with altered specificities. Here we report the change of specificity in the *lac* system by changing residues 1 and 2 of the recognition helix of *lac* repressor. The specificity changes are brought about cooperatively by the change of both residues. Exchanges of just one residue broaden the specificity. Our results hint that the recognition helix of *lac* repressor may possibly have the opposite orientation to those in Lambda *cro* protein or *434 CI* repressor.
Key words: lactose system/*lac* repressor/recognition helix/protein–DNA recognition/*Escherichia coli*

Introduction

Accurate recognition of DNA sequences by proteins is the prerequisite of accurate regulation of transcription. Several possible modes of such recognition exist. Here we are interested in alpha helices recognizing details of the structure of the major groove of B-DNA. The first evidence for such a recognition was seen in the X-ray structure of Lambda *cro* protein. Its dimer contains two alpha helices just 34 Å apart (Ohlendorf *et al.*, 1982). These recognition helices are part of a helix-turn-helix motif which has been found to be structurally similar in many repressors and analogous proteins (Matthews *et al.*, 1982; Sauer *et al.*, 1982; Gicquel-Sanzey and Cossart, 1982; Sauer and Pabo, 1984). In the *trp* repressor (Schevitz *et al.*, 1985) and the *cap* protein (Weber and Steitz, 1984) the recognition helix has been identified by X-ray analysis. Recently X-ray analysis of repressor–operator cocrystals has proven that the recognition helix of *434 CI* repressor indeed recognizes the major groove of *434* operator (Anderson *et al.*, 1987). Moreover, the specificities of the *cap* protein (Ebright *et al.*, 1984) and of *434 CI* repressor (Wharton and Ptashne, 1987) have been altered by exchanges of residues of the recognition helix.

The *lac* repressor belongs to the class of DNA recognizing proteins which have a helix-turn-helix motif (Matthews *et al.*, 1982). Although no suitable crystals exist for X-ray analysis, the *lac* system seems a good candidate for a detailed analysis of protein–DNA recognition. It has been extensively analyzed by bacterial genetics and its physiology is well known (Müller-Hill, 1975; Miller, 1978). In the wild-type a more than 1000-fold repression is observed. This strong repression is predominantly due to interaction of *lac* repressor with the first *lac* operator close to the *lac* promoter. If the second *lac* operator in the *lacZ* (β-galactosidase) gene is destroyed a 300-fold repression is still observed (Eismann *et al.*, 1987). This suggests an elaborate specificity of recognition of the first *lac* operator by *lac* repressor. Moreover a fully symmetric 'ideal' *lac* operator has been described (Sadler *et al.*, 1983; Simons *et al.*, 1984) which binds *lac* repressor *in vitro* eight times tighter than wild-type *lac* operator.

Besse *et al.* (1986) have shown that natural *lac* operator can be replaced by a unique restriction site into which synthetic ideal *lac* operator (Simons *et al.*, 1984) can be cloned. The detailed knowledge of the physiology of the *lac* system allows one to set up selection systems for exceedingly rare I^+ mutants (Miller, 1978). We have made use of these advantageous properties of the *lac* operon and have set up a system which allows detection and selection of repressor mutants with changed specificity. Similar systems have been recently described by Wharton and Ptashne (1987) and Benson *et al.* (1986). Preliminary accounts of our results can be found in Lehming *et al.* (1987a,b). Here we report several specificity changes in *lac* repressor and discuss the possibility first raised by Adler *et al.* (1972) of a code governing this type of protein DNA recognition.

Results

The system

We have constructed a system which allows systematic testing of the interactions between any repressor and operator. The system consists of two plasmid vectors which are mutually compatible (they have different origins), which can be selected for individually (they carry an *amp* or a *tet* resistance gene respectively) and which can be easily separated on gels because of their different sizes (Figure 1).

One of the plasmids contains a rudimentary *lac* operon consisting of a *lac* promoter and a *lacZ* gene. The *lac* operator is deleted and is replaced by a unique *Xba*I site (Besse *et al.*, 1986) into which synthetic operators can be cloned. Thus the level of expression of β-galactosidase depends on the free or occupied state of this operator. The absence of a *lacZ* gene and the presence of a functional *lacY* (permease) gene in the host allow selection of phenotypic repressor I^+ mutants with the help of the *galE* selection system (Miller, 1978). The *galE* mutation kills *Escherichia coli* cells in the presence of a sufficient amount of D-galactose which will arise when a constitutive *lac* system hydrolyzes lactose to D-galactose and glucose. Under suitably repressed conditions this level of galactose is too low to be toxic.

The other plasmid contains the repressor gene. Since we wanted to study the *lac* repressor–operator system we introduced the *lacI* (repressor) gene. We synthesized its first 70 codons and introduced as many unique restriction sites as possible without changing the protein sequence of *lac* repressor. This semi-synthetic

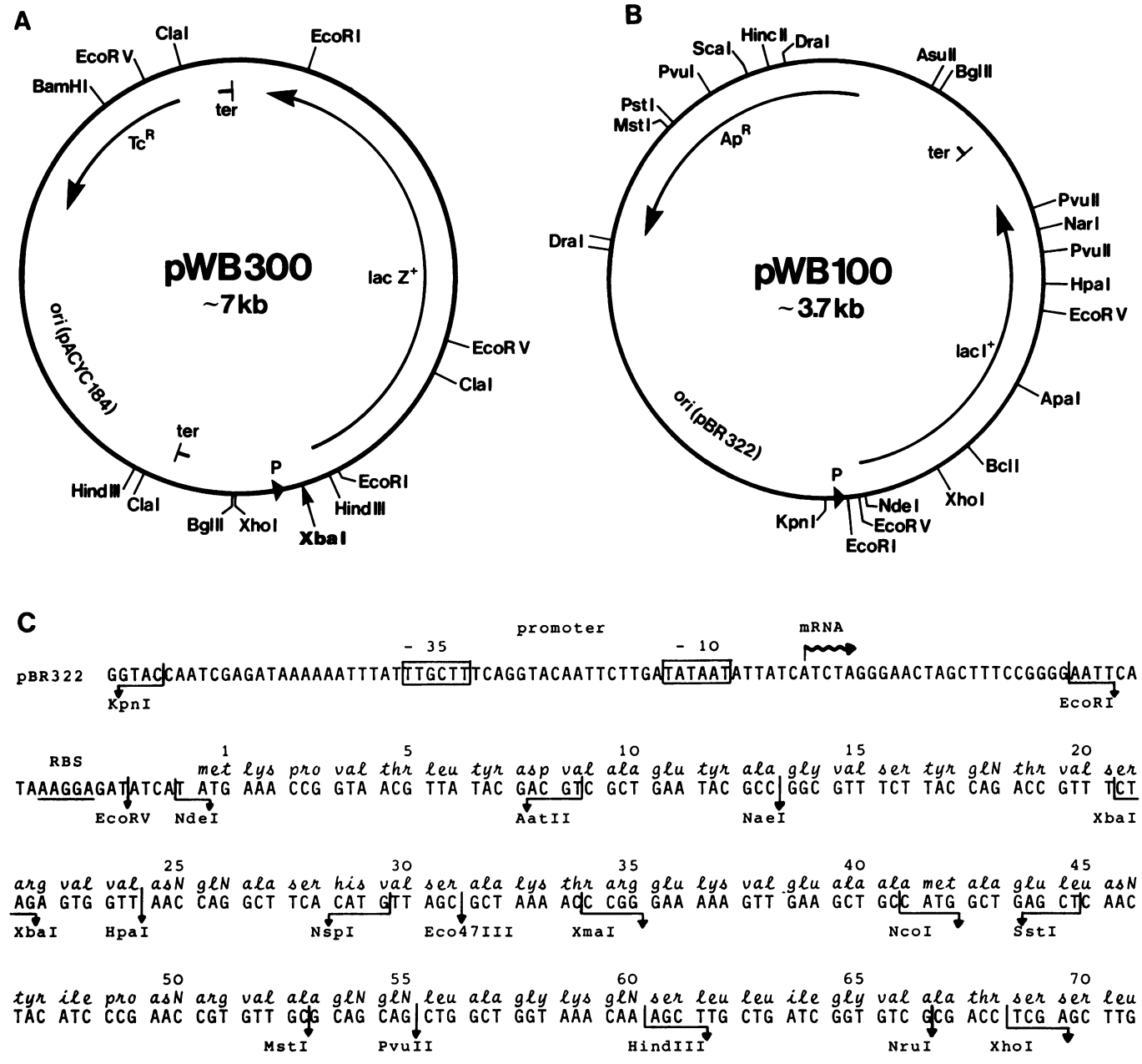

Fig. 1. The plasmid system used to detect and select *lac* repressor variants with altered specificity. **(A)** The plasmid pWB 300 which carries the various *lac* operators in its *XbaI* site. **(B)** The plasmid pWB 100 which carries the *lacI* gene. **(C)** The synthetic part of the *lacI* gene. Unique or almost unique useful restriction sites are indicated.

lacI gene allows the introduction of specific mutations by the replacement of small regions of double-stranded synthetic DNA in a cassette-like manner, or the introduction of a maximum of random mutations in discrete codons by appropriate random chemical DNA synthesis.

The 18 bp ideal lac operator is optimal in lac repressor recognition

We first addressed the question whether the symmetrical 'ideal' *lac* operator described by Sadler *et al.* (1983) and Simons *et al.* (1984) had indeed the optimal structure to bind *lac* repressor. We synthesized all possible *lac* operator variants with symmetric changes in just one base on both sides of its axis of symmetry in base pairs 1 to 9 (Figure 2). We then determined the levels of β-galactosidase of all variants under repressed and derepressed conditions. We found a 200-fold repression in the case of ideal *lac* operator. One may recall that repression in the wild-type *lac* system on the chromosome or an episome is about 1000-fold (Jacob and Monod, 1961). This discrepancy in the extent of repression is at least in part due to the occasional loss of the repressor-carrying plasmid. If we plate an aliquot of the bacteria we use to assay β-galactosidase, we observe consistently that about 0.5% of them have lost the *ampR lacI*$^+$ plasmid and express β-galactosidase constitutively. The value of 200-fold repression we observe with ideal *lac* operator in the plasmid system thus marks the limit of our system and could actually be substantially higher.

The levels of β-galactosidase differ 20-fold in the *lac* operator variants under derepressed conditions. This could be due to various reasons, such as changes of promoter strength or messenger stability (Besse *et al.*, 1986). We assume here that the ratio of

Plasmid pWB	Operator Sequence	Spec. Activity of ß-Galactosidase repressed	Spec. Activity of ß-Galactosidase not repressed	Repression in vivo	T/2 in vitro min
310	9 8 7 6 5 4 3 2 1 AATTGTGAGC * GCTCACAATT TTAACACTCG * CGAGTGTTAA 1 2 3 4 5 6 7 8 9	≤21	4,400	≥200	13
311	AATTGTGAGA TCTCACAATT	3,300	7,500	2	
313	G C	1,400	5,200	4	
314	T A	1,100	22,000	20	
321	AATTGTGAAC GTTCACAATT	210	16,000	76	
322	C G	1,000	16,000	16	
324	T A	650	7,100	11	
332	AATTGTGCGC GCGCACAATT	2,800	5,800	2	
333	G C	900	4,200	5	
334	T A	2,800	5,300	2	
341	AATTGTAAGC GCTTACAATT	740	3,800	5	
342	C G	730	4,500	6	
344	T A	1,000	4,000	4	
351	AATTGAGAGC GCTCTCAATT	370	5,900	16	
352	C G	660	7,000	11	
353	G C	400	3,600	9	
361	AATTATGAGC GCTCATAATT	2,400	6,300	3	
362	C G	2,400	4,500	2	
364	T A	3,600	10,000	3	
371	AATAGTGAGC GCTCACTATT	240	1,100	5	
372	C G	300	4,000	13	
373	G C	290	6,700	23	
381	AAATGTGAGC GCTCACATTT	30	1,500	50	6
382	C G	85	2,900	34	2,5
383	G C	88	2,100	24	2,5
392	ACTTGTGAGC GCTCACAAGT	31	2,300	74	2,5
393	G C	96	2,700	28	3,5
394	T A	35	1,800	51	3,0

Fig. 2. The 18 bp ideal *lac* operator is optimal for *in vivo* repression and *in vitro lac* repressor binding. Ideal *lac* operator (Besse *et al.*, 1986) and its variants have been cloned into the *Xba*I site of plasmid pWB 300 (see Figure 1). All operators share the consensus sequence AATTGTGAGCGCTCACAATT. The host was the *E. coli* strain DC 41-2 $(lacpro)_{\Delta}$ *RecA GalE SmR*. Plasmid pWB 100 (see Figure 1) was used as source of *lac* repressor. To test for the state of no repression, we used a variant of pWB 100 which had a deletion within the *lacI* gene (codon 14–60) making it phenotypically I^-. Bacteria were grown at 37°C in glycerol minimal medium in the presence of proline, ampicillin and tetracycline to an OD_{580} of 1.0. Repression is defined as specific activity of ß-galactosidase in the absence of *lac* repressor divided by the specific activity of ß-galactosidase in the presence of *lac* repressor. Unavoidable loss of the *lacI amp* plasmid limits repression to a factor of 200. We note that some values reported earlier (Lehming *et al.*, 1987a,b) had to be corrected. Occasionally the loss of the *lacI amp* plasmid had exceeded 0.5%. Here we checked all cultures for plasmid loss and discarded those in which plasmid loss was excessive. Furthermore earlier occasional inadvertent mixup of plasmids was found and corrected by sequencing again all *lac* operator variants.

ß-galactosidase under constitutive and repressed conditions is *not* influenced by the absolute level of expression. We cite as evidence for this assumption the fact that repression in the *lac* promoter mutant L8, in which expression is reduced 12-fold, is the same as in wild-type promoter (Miller, 1978). We observe that the level of repression is drastically reduced in all but one of the variants which have exchanges in base pairs 1 to 7. The repression in the variants in base pairs 8 and 9 is low but significant.

To control the significance of these *in vivo* measurements, we measured *in vitro* the dissociation rates of these *lac* repressor complexes with mutant *lac* operators and compared them with the dissociation rate of the complex of ideal *lac* operator with *lac* repressor. We used for these measurements the filter binding assay of Riggs *et al.* (1970). The half-life of the complex between ideal *lac* operator and *lac* repressor is 13 min under our conditions. The half lives of all complexes with *lac* operator variants in base pairs 8 and 9 were at least two times lower (Figure 2). We conclude that ideal *lac* operator is optimal for recognition of *lac* repressor.

Altering the specificity of lac repressor by design

The second question we addressed was whether the specificity of the *lac* repressor–operator system could be altered. There are two ways to do that. One may try to predict specific exchanges in the recognition helix which may alter the recognition of one particular base of the operator. Or one may induce random changes in one or several residues of the recognition helix. A selection system must then be used to select for mutants with an altered specificity. We have used both approaches successfully.

The *lac* and *gal* repressors and operators are rather homologous in sequence (Wilcken-Bergmann and Müller-Hill, 1982; Adhya and Miller, 1979). The recognition helix of *gal* repressor differs from the recognition helix of *lac* repressor in three positions (Figure 3). Residues 1 and 2 of the recognition helix are Val1 Ala2 instead of Tyr1 Gln2 and residue 8 is an Ile8 instead of Val8. The *gal* operator differs from ideal *lac* operator in positions 2 and 4. The change in position 4 is a symmetric G-C to A-T exchange, whereas the change in position 2 of *gal* operator seems not to be conserved (Figure 3). The A which is occasionally found in position 2 of the *gal* operator is also well tolerated by wild-type *lac* repressor when it is introduced into *lac* operator (Figure 2). Thus it seemed possible that *lac* repressor, in which just residues 1 and 2 of the recognition helix are replaced by Val1 Ala2, would recognize *lac* operator variant 341 5′TGTAAGC–

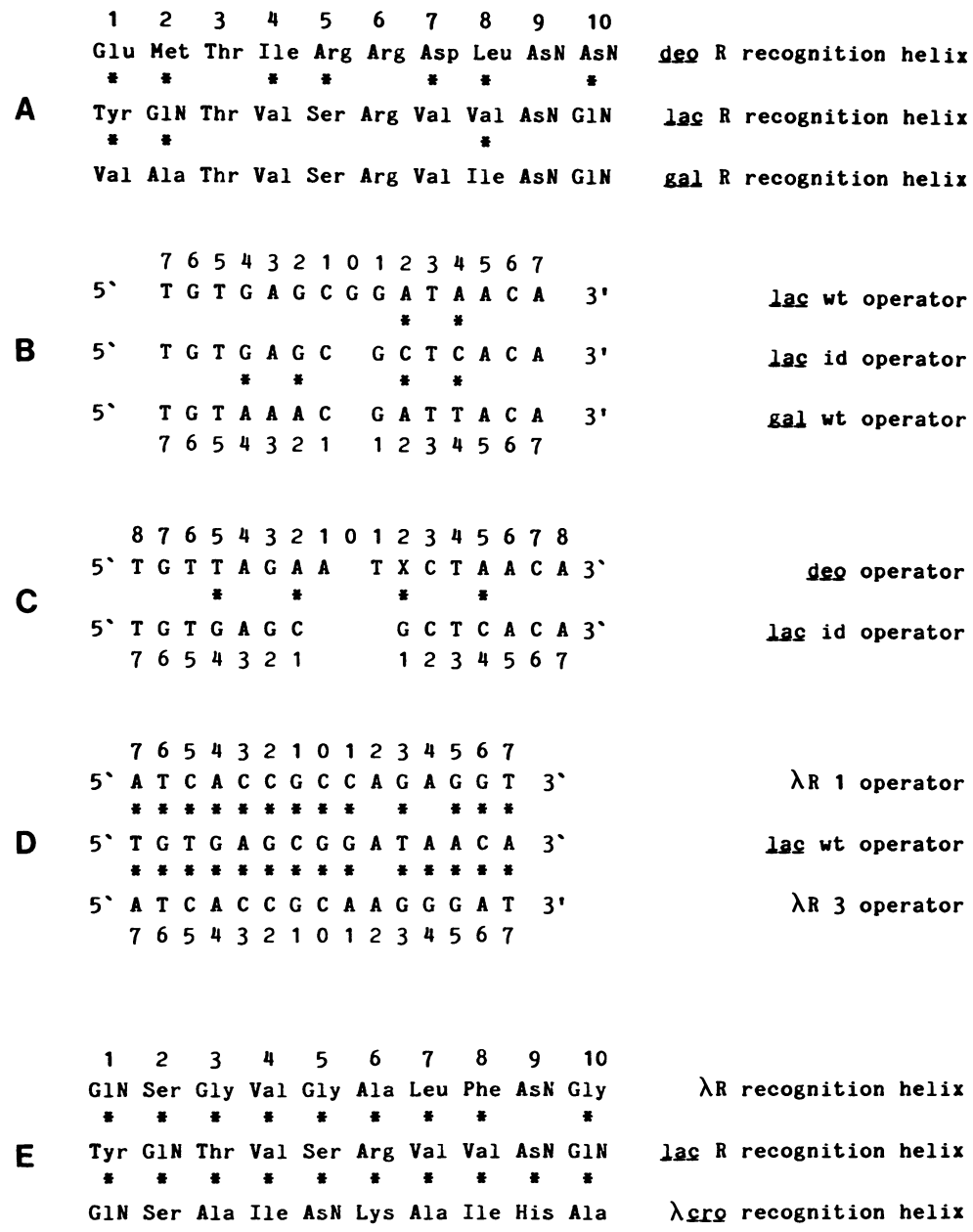

Fig. 3. The sequences of the recognition helices of various repressors and the sequences of various operators. **(A)** and **(E)**: residues of the recognition helices are numbered from 1 to 10. Residues differing from those of *lac* repressor are marked by stars. **(B), (C)** and **(D)**: the bases of the various operators are all numbered in the same systematic manner. The (sometimes lacking) central base is numbered 0. The base to the left or right of the axis of symmetry 1, and so on. Bases which differ from the *lac* operator are marked by stars. *gal* and *deo* operators are those close to the promoter. References for the various sequences are given in the text.

GCTTACA3′ in which G4 is replaced symmetrically by A, better than ideal *lac* operator 5′TGTGAGC−GCTCACA3′.

The data shown in Table I confirm this expectation. Val1 Ala2 *lac* repressor represses ideal *lac* operator at least eight times less than wild-type *lac* repressor. It represses however variant 341 of *lac* operator at least 40 times better than wild-type *lac* repressor. We synthesized the Ile8 variant to check whether the Val8 to Ile8 exchange in the recognition helix of *lac* repressor influences the specificity of binding. It does not; it does however increase binding of this repressor to all *lac* operator variants 2-fold in a nonspecific manner (data not shown).

To find out whether Val1 *lac* repressor and/or Ala2 *lac* repressor are able to recognize *lac* operator variant 341 we synthesized *lac* repressor variants carrying these single exchanges. The data in Table I show that Val1 *lac* repressor represses *lac* operator variant 344 to a greater extent than does wild-type repressor. It represses variants 351, 352 and 353 as well as does wild-type repressor. But it does not repress variant 341. In contrast Ala2 *lac* repressor strongly represses *lac* operator variant 341 but no other one. These results suggest that residue 1 of the recognition helix interacts with base pairs 4 and 5 and that Ala2 of the recognition helix interacts with base pair 4. In addition we find weak effects in the case of the Ala1 variant on base pair 2 of *lac* operator (see Discussion).

Filter binding assays confirm and extend this conclusion (Table II). They show that the complex between Val1 Ala2 *lac* repressor and *lac* operator variant 341 is exceedingly stable, whereas the corresponding complexes with Val1 or Ala2 *lac* repressor are

Table I. Repression of *lac* operator variants by *lac* repressor variants

lac operator variant	Repression: *lac* repressor variant											
	wt	V1	A1	Q1	H1	A2	S2	M2	V1A2	A1A2	Q1S2	Q1M2
310	≥200	64	120	35	≥200	20	20	≥150	24	16	24	27
321	76	34	31	12	76	2	1	70	3	1	1	5
322	16	11	32	6	42	1	1	21	2	2	3	5
324	11	4	12	1	3	1	1	7	1	1	1	1
332	2	2	2	1	2	1	1	3	1	1	1	1
333	4	2	5	2	3	1	1	3	1	1	2	1
334	2	1	2	1	1	1	1	3	1	1	1	1
341	5	2	3	1	2	50	14	10	≥200	120	10	1
342	6	2	10	1	1	1	2	5	2	1	1	1
344	4	18	50	30	≥200	1	1	12	18	10	6	100
351	16	15	45	7	11	2	12	4	25	60	30	7
352	10	8	18	3	1	2	2	2	3	3	4	3
353	9	12	20	8	3	1	12	12	20	28	50	12
361	3	2	2	1	2	1	1	1	1	1	1	1
362	2	1	2	1	1	1	1	1	1	1	1	1
364	3	2	2	1	3	1	1	2	1	1	1	1

The *lac* operator carrying plasmids are the same as in Figure 2. The *lac* repressor variants are named after the exchange they carry in their recognition helix. The usual one-letter code is used. The number indicates the number in the recognition helix. Thus V1 indicates that tyrosine 17 (residue number one of the recognition helix) is exchanged for a valine. Repression is defined as in Figure 1.

at least 100 times less stable. Val1 and Ala2 of the recognition helix of *lac* repressor apparently cooperate in recognizing A-T in position 4 of *lac* operator.

Wild-type *lac* repressor protects the guanosines in positions 4 and 6 of ideal *lac* operator against methylation (Lehming *et al.*, 1987a). We wondered whether protection of G4 by Val1 Ala2 *lac* repressor might be reduced. We found some reduction in methylation protection of G4 (data not shown), but the effect was rather weak. We reasoned that the bulky side chain of Val1 might be the cause of this and synthesized Ala1 Ala2 *lac* repressor. It repressed the operator variant 341 *in vivo* strongly but less well than Val1 Ala2 (Table I). In a methylation protection experiment with ideal operator Ala1 Ala2 *lac* repressor proved to be deficient in methylation protection of G in position 4 (Figure 4). This confirms that residues 1 and/or 2 of the recognition helix of *lac* repressor interact with the base pair in position 4 of ideal *lac* operator.

The second system we used for designing a specificity change was the *deo* system (Valentin-Hansen *et al.*, 1985). At first sight the recognition helix of the *deo* repressor and the *deo* operator look rather different from the corresponding *lac* structures (Figure 3). A closer look at the *deo* operator however reveals that it contains a variant of the 5′ TGTGA 3′ motif found in *lac* operator. The *deo* motif 5′ TGTTA 3′ differs from the *lac* and *gal* motifs in one base and in its location one additional base pair upstream from the centre of symmetry. If this motif were recognized by the *deo* recognition helix, exchanging residues 1 and 2 of the *lac* recognition helix for those of the *deo* recognition helix should suffice to repress plasmid 344, i.e. ideal *lac* operator variant 5′ TGTTAGC−GCTAACA 3′. We knew that a replacement of GlN2 with Glu abolishes repression in ideal *lac* operator and all variants completely (data not shown). We knew in addition that a Tyr1 to GlN1 exchange broadens the specificity (Table I and below). We therefore decided not to synthesize the Glu1 Met2 but the GlN1 Met2 variant. It can be seen in Table I that our expectations were borne out. The GlN1 Met2 *lac* repressor variant represses specifically plasmid 344 and not at all plasmid 341. This result confirms that residues 1 and 2 of the *lac* recognition helix interact specifically with base pair 4 of the *lac* operator. The single exchanges of GlN1 and Met2 again broaden the specificity of repressor−operator recognition (Table I).

Knowing that the specific recognition of base pair 4 of *lac* operator could be altered in two directions by exchanges in residues 1 and 2 of the recognition helix of *lac* repressor, we reasoned that it might be revealing to replace residues 1 and 2 of the recognition helix of *lac* repressor with the residues found both in Lambda repressor and Lambda *cro* protein (i.e. GlN1 Ser2). The Lambda operators and recognition helices of Lambda repressor and Lambda *cro* show no apparent homologies with the corresponding structures of the *lac* system (Figure 3). It is known that the serines in position 2 of the recognition helices of Lambda repressor and Lambda *cro* protein interact with the G of the opposite strand of base pair 5 (our numbering system!) of Lambda operator (Hochschild and Ptashne, 1986). It seemed worthwhile to study the effect of a GlN1 Ser2 replacement in the recognition helix of *lac* repressor. We synthesized the particular *lac* repressor variant and tested its capability to repress the *lac* operator variants. We found a broadening of the specificity. GlN1 Ser2 *lac* repressor represses ideal *lac* operator at least seven times less well than wild-type *lac* repressor, but it represses *lac* operator variants 341, 351 and 353 about as well as ideal *lac* operator (Table I). We will come back to the significance of these results in the Discussion.

Altering the specificity of lac repressor by random exchanges of codons of the recognition helix

The experiments so far involved residues 1 and 2 of the recognition helix of *lac* repressor and base pairs 4 and 5 of *lac* operator. We could not find one case which allowed us to predict the effect of possible exchanges further downstream in the recognition helix.

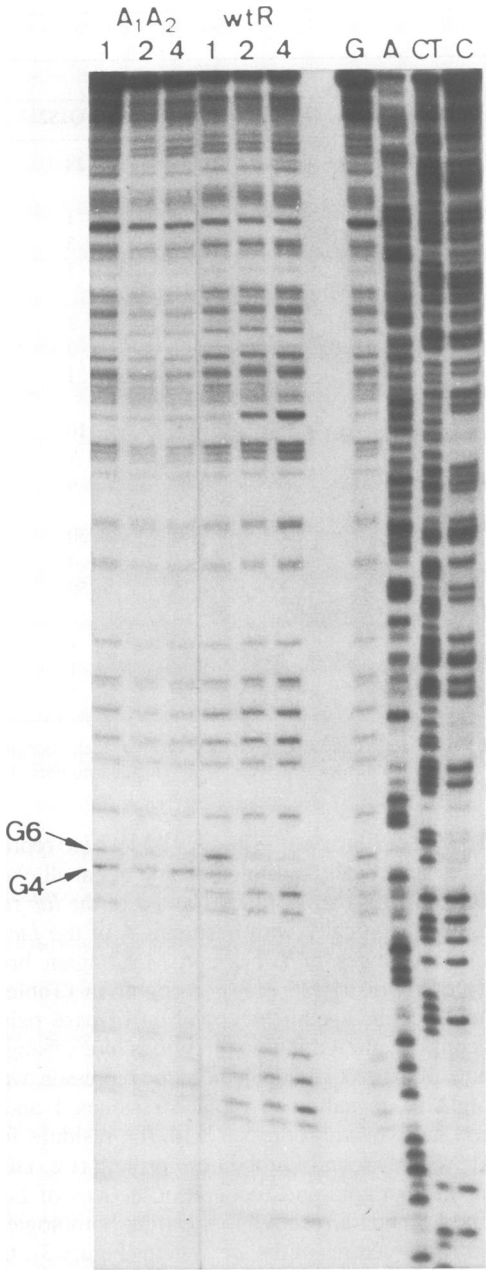

Fig. 4. Methylation protection experiments. Ala1 Ala2 *lac* repressor protects only G6 whereas wild-type *lac* repressor protects G4 and G6 of ideal *lac* operator against methylation. 1 pmol of labeled DNA fragment containing ideal *lac* operator was mixed with 1, 2 or 4 pmol of wild-type *lac* repressor or *lac* repressor having residues 1 and 2 of its recognition helix exchanged for alanines. Then the methylation was performed according to Ogata and Gilbert (1978).

We therefore set up a selection system which allowed the selection of rare I^+ mutants with altered specificity out of a mass of irrelevant I^- mutants. The selection makes use of the property of *Gal*E mutants to be killed by D-galactose. The I^+O^+ *lac* system is sufficiently repressed under our conditions (see Materials and methods) to reduce the level of β-galactosidase so that toxic amounts of D-galactose are not produced. If *lac* repressor does not recognize *lac* operator properly β-galactosidase produces sufficient amounts of D-galactose in the presence of β-galactosides to kill the cells.

We first investigated whether a single amino acid exchange in residues 1 to 9 of the recognition helix would change the specificity of recognition of one base pair of the ideal *lac* operator. We introduced 10% of wrong bases by chemical synthesis in all positions from codon 1 to codon 9 of the recognition helix. We produced banks of about 40 000 clones in our host carrying plasmids 332 to 383 (see Figure 2). In the host carrying plasmid 344 we found one positive clone. Tyr1 was exchanged for His1. In Tables I and II we present the data of this variant. It represses *in vivo* ideal *lac* operator as well as plasmid 344. *In vitro* it binds to 310 and 344 DNA at least 100 times stronger than wild-type *lac* repressor binds to ideal *lac* operator, i.e. 310 DNA. Thus, it is not surprising that it produces a lac^- phenotype when introduced into a chromosomal $I^-Z^+Y^+$ background.

Discussion

Do rules exist which determine the recognition of different surfaces of the major groove of B-DNA by different alpha helices of proteins? The question of the existence of such a code was raised by Adler *et al.* (1972). Most recent observers have been rather pessimistic about its possible existence. The arguments used against it are the following.

(i) X-ray analysis has indicated that the recognition helices of Lambda *cro* protein (Ohlendorf *et al.*, 1982), *434 CI* repressor head piece (Anderson *et al.*, 1987) and *cap* protein (Weber and Steitz, 1984) enter the major groove of their respective operators at different angles. This does not exclude the possibility that an optimal angle exists which is used for optimal recognition.

(ii) Recent data suggest that the fine structure of B-DNA may be sequence-dependent and merely by that fact influence the recognition of nearby bases (Koudelka *et al.*, 1987). This fine structure may favour some and exclude other sequences from being recognized properly.

(iii) Finally, the amino acid sequences of recognition helices which recognize the same or rather similar operator sequences are vastly different (see for example the review of Pabo and Sauer, 1984). Most illuminating in this respect are the cases of Lambda and *434* repressor and *cro* proteins. The repressors and *cro* proteins of both phages have almost completely different recognition helices but recognize rather similar operator sequences (Ptashne, 1978, 1986; Hochschild *et al.*, 1986; see also Figure 3).

We note that the Lambda and *434* repressors and *cro* proteins must bind cooperatively to two adjacent operators to perform significant repression. Any system of adjacent operators invites degeneracy of recognition. Wharton and Ptashne (1987) found very little repression by *434* repressor in a system similar to ours where one *434* operator replaced *lac* operator. When we replaced the first two residues of the recognition helix of *lac* repressor with those of Lambda repressor or Lambda *cro* protein, which are the same, we found a 10-fold decrease of *lac* repression and a broadening of specificity (Table I). The Lambda or *434* data do not exclude the possibility that a code may exist, but they indicate that, if it exists, natural selection has made parsimonious use of it. To phrase it differently, some sloppiness in protein–DNA recognition has been advantageous, and extreme accuracy was not always asked for. We may also cite here the case of the His1 *lac* repressor variant. In nature His1 is not found in any recognition helix known so far. This may be due to the drastic increase of the binding constant (Table II) which may be simply too high to be useful.

In the *lac* system the main operator close to the promoter alone is sufficient to allow several hundred-fold repression. The second operator has only a weak auxiliary function (Eismann *et al.*, 1987). The data presented here indicate that ideal symmetric *lac*

Table II. Comparison of *in vivo* and *in vitro* binding of variant *lac* repressors to variant *lac* operators

lac operator variant	*lac* repressor	Half life of R-O complex (min)	Repression
310	wt	13	≥200
310	Val1	2.5	64
310	Ala2	≤ 1	20
310	Val1 Ala2	≤ 1	24
310	His1	≥1500	≥200
341	wt	≤ 1	5
341	Val1	≤ 1	2
341	Ala2	2.5	50
341	Val1 Ala2	320	≥200
344	His1	≥3000	≥200

lac operator variants are described and repression is defined in Figure 2. The half lives of the R-O complexes were determined with filter binding tests according to Riggs *et al.* (1970).

operator (Sadler *et al.*, 1983; Simons *et al.*, 1984) is optimal in its entire structure to recognize *lac* repressor. On the other hand, this does not imply that *lac* repressor has an optimal structure for DNA binding or *lac* operator recognition. Indeed, a Val to Ile exchange in residue 8 of the recognition helix increases binding to *lac* operator nonspecifically 2-fold.

If a code or simple rules govern protein DNA recognition, the *lac* repressor–operator system seems to be a good candidate to unravel them. We found that the exchange of one residue of the recognition helix was not sufficient to obtain a marked change to a new specificity. Exchanges of single amino acids lowered or increased repression with ideal *lac* operator but allowed at the same time comparable repression of one or more other *lac* operator variants. Thus specificity of *lac* repressor can be easily broadened through changes of one or the other amino acid in the recognition helix. However, when residues 1 and 2 of the recognition helix are changed appropriately, the specificity can be changed. We report two such cases (Table I). Val1 Ala2 binds specifically to A in position 4 whereas GlN1 Met2 binds specifically to T in position 4 of ideal *lac* operator. The *in vivo* and *in vitro* data with corresponding single amino acid exchanges suggest that the effect of the change of two amino acids is cooperative (Tables I and II).

This confirms the observation of Mossing and Record (1985) that the free-energy changes associated with single base pair exchanges in the *lac* operator are not additive. The search for specificity changes in the recognition helix has to take this cooperativity into account. Thus we will have to look for other changes in residues 1 and 2 to find the missing *lac* repressor variant, which may recognize *lac* operator variant 342. We will have to exchange other pairs of residues of the recognition helix, possibly residues 5 and 6 or 9 and 10 (Lehming *et al.*, 1987b) to obtain *lac* repressor variants which recognize *lac* operator variants with exchanges in other base pairs. These are the types of experiments which will show the existence or nonexistence of a code of protein–DNA recognition.

Finally we note that Ebright (1986) predicted that residue 2 of the recognition helix of *lac* repressor would interact with base pair 4 of *lac* operator. The results shown in Tables I and II suggest that residues 1 and 2 of the recognition helix interact predominantly with base pair 4 and to some extent with base pair 5 of ideal *lac* operator. This conclusion is confirmed by the fact that Ala1 Ala2 *lac* repressor does not protect G4 of ideal *lac* operator

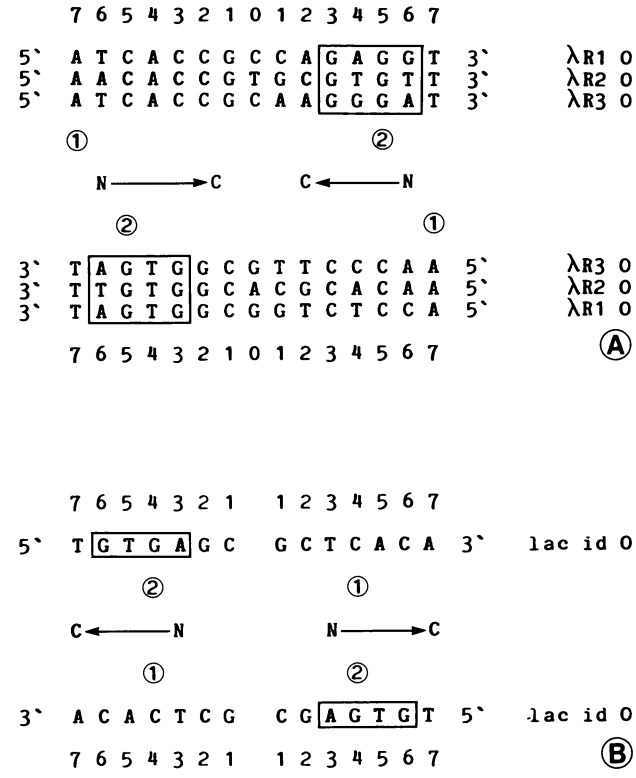

Fig. 5. Comparison of the established operator recognition of Lambda repressor and Lambda *cro* protein (Ptashne, 1986) with the proposed hypothetical operator recognition of *lac* repressor. The space between the upper and lower strands of the Lambda R operators **(A)** and ideal *lac* operator **(B)** symbolizes the deep groove. The encircled numbers within the deep groove symbolize residues 1 and 2 of the recognition helices of Lambda repressor and Lambda *cro* protein in (A) and of *lac* repressor in (B). The residues are placed just below or above the base with which they interact most strongly. The known direction of the recognition helix of Lambda repressor and Lambda *cro* protein is indicated by an arrow pointing from the N terminus of the recognition helix towards its C terminus. The indicated orientation of the recognition helix of *lac* repressor is hypothetical. The 5′ GTGA 3′ motifs in Lambda and *lac* operators are encased.

against methylation in contrast to wild-type *lac* repressor. The fact that Val1 Ala2 binds best to A4 (341) *lac* operator and that GlN1 Met2 *lac* repressor binds best to T4 (344) *lac* operator underlines the importance of van der Waals forces in these specific interactions.

This raises a paradox: if Val1 binds to the methyl group of T4 on the opposite strand of *lac* operator, and if Met2 binds to T4 of the coding strand, the recognition helix has to enter the major groove in the opposite direction to the *434* repressor (Anderson *et al.*, 1987) and the Lambda *cro* protein (Ohlendorf *et al.*, 1982). Boelens *et al.* (1987) reported n.m.r. measurements which indicate that *lac* repressor head piece interacts with a *lac* operator fragment in opposite orientation. The decisive *in vivo* or *in vitro* experiments with entire *lac* repressor and *lac* operator are lacking. We note that our data indicate weak interaction of residue 1 of the recognition helix with base pair 2 of the operator: A1 *lac* repressor represses *lac* operator variant 322 2-fold better than wild-type *lac* repressor, whereas wild-type *lac* repressor represses ideal *lac* operator better than the A1 variant (Table I). We are at the moment investigating whether exchanges downstream of codons 1 and 2 of the recognition helix of *lac* repressor influence the recognition of bases upstream or downstream of base pairs 4 and 5 of ideal *lac* operator.

In Figure 5 we compare the interaction between the recognition

helix of Lambda repressor and Lambda operator with the hypothetical interactions between the recognition helix of *lac* repressor and *lac* operator in opposite orientation. It can be seen that the respective regions of *lac* and Lambda operator look rather alike if the recognition helix of *lac* repressor enters the major groove of *lac* operator in the opposite orientation as the recognition helix of Lambda repressor enters its operator. If this were so, one may expect that the recognition of the left half of Lambda $O_R I$ operator by Lambda repressor and Lambda *cro* protein could be significantly enhanced if residues 1 and 2 and possibly residues 5, 6, 9 and 10 of their respective recognition helices were replaced by those of *lac* repressor recognition helix. Experiments are under way to test these predictions with Lambda *cro* protein.

Materials and methods

Chemicals and enzymes

Most restriction enzymes were obtained from Boehringer (Mannheim, FRG), Bethesda Research Laboratories (Eggenstein, FRG) and New England Biolabs (Beverly, USA); DNA polymerase large fragment and polynucleotide kinase from Boehringer (Mannheim, FRG); restriction endonuclease *Nsp* 7524 (*Nsp*I) and [^{32}P]deoxyribonucleotides from Amersham Buchler (Braunschweig, FRG); 5-bromo-4-chloro-3-indolyl-β-*D*-galactoside (X-gal) and isopropyl-β-*D*-thio-galactoside (IPTG) from Bachem Fine Chemicals (Torrance, USA); o-nitro-phenyl-β-*D*-thio-galactoside from Biomol (Ilvesheim, FRG); agarose and urea from Bethesda Research Laboratories (Eggenstein, FRG); the chemicals used for automated DNA synthesis from Applied Biosystems (Pfungstadt, FRG); all other chemicals were obtained from Sigma Chemie (München, FRG) or Merck (Darmstadt, FRG).

Oligonucleotides, 7–54 bases in length, were synthesized on an Applied Biosystems 380 A DNA synthesizer. They were purified on denaturing polyacrylamide gels prior to cloning and sequence analysis (Maxam and Gilbert, 1977; Chen and Seeburg, 1985). Wild-type *lac* repressor was a gift of K.Beyreuther. The *lac* repressor mutants were purified according to Müller-Hill *et al.* (1971). SM113 nitrocellulose filters were obtained from Sartorius (Göttingen, FRG).

Methods

The methods for constructing the plasmids are standard techniques (Maniatis *et al.*, 1982). Strategies for the synthesis of oligonucleotides and conditions for their cloning have been described previously (Wilcken-Bergmann *et al.*, 1986).

Strain DC 41-2 was used for the repression tests, strain DC 41-2 F′ for the selection experiments. The latter has the genotype: $(lacpro)_{\Delta}$*galE smR RecA F′ lacproI^{Q-}Z^{-}Y^{+}*. The episome carries an *I–Z* deletion leaving the 5′ end of the I^Q gene intact (Gho and Miller, 1974). β-Galactosidase was determined as in Miller (1972). To determine the half lives of *lac* repressor operator complexes we used the filter binding assay of Riggs *et al.* (1970).

For the methylation protection experiments we followed the procedure of Ogata and Gilbert (1978). The pWB 310 DNA was cleaved with *Eco*RI and *Hae*III and end-labeled with [^{32}P]dATP. The reaction mixture consisted of 1 pmol of labeled fragment in 200 μl of 50 mM cacodylate, 10 mM $MgCl_2$, 0.1 mM EDTA at pH 8 containing sonicated calf thymus DNA (10 μg/ml). About 1, 2 or 4 pmol *lac* repressor or *lac* repressor variant and 1 μl DMS were added and incubated at room temperature for 5 min. The G-reaction was carried out as described by Maxam and Gilbert (1977).

For the banks with the mutations in codon 1 to 9 of the recognition helix of *lac* repressor (Derbyshire *et al.*, 1986) two DNA strands were synthesized: 5′ GGC GTT TCT 412 213 122 344 424 131 343 344 112 CAG GCT TCA CAT GTG CGG 3′ (1 = 90% A, 4% G, 3% C, 3% T; 2 = 90% C, 4% G, 3% A, 3% T; 3 = 91% G, 3% A, 3% C, 3% T; 4 = 90% T, 4% G, 3% A, 3% T) and 5′ CCG CAC ATG TGA AGC CTG 3′ as primer. The polyacrylamide gel purified oligonucleotides were incubated at 70°C for 10 min and then cooled slowly to room temperature. They were treated twice with Klenow fragment of polymerase I and twice extracted with phenol and precipitated with ethanol. Then they were cut with *Nsp*I, purified on a 10% polyacrylamide gel and ligated with *Nsp*I–*Nae*I treated pWB 100 using T4 DNA ligase. The DNA was then transformed into strain JMX 8634 (Miller, 1978) *(lacpro)*$_{\Delta}$ *trp*$^{-}$ *ton*B *80dlac I*$^{-}$ *lac*$^{+}$ *smR*. The majority of the transformants were *lac* constitutive. Large amounts of plasmid DNA were prepared on CsCl gradients and used for the various banks.

The DNA was then transformed into bacteria of strain DC 41-2 F′ mentioned above, which carried one of the pWB 300 derivatives. We used the conditions of Hanahan (1985): 10 ng of pWB 100 DNA was transformed into 200 μl competent cells containing one of the *lac* operator–variant plasmids. After the usual 1:4 dilution the bacteria were further diluted 1:10 into nutrient broth containing 10 μg/ml tetracycline and 0.8 mg/ml ampicillin. When the bacteria had reached an OD_{600} of about 2, 10^6 cells were spread on selection plates containing 2g/l glucose, 0.1 g/l phenyl-β-*D*-galactoside, 1 g/l o-nitrophenyl-thio-β-*D*-galatoside (TONPG), 10 mg/l tetracycline and 0.4 g/l ampicillin in minimal agar. The selection plates are essentially those introduced by Miller (1978) modified by the addition of TONPG, which inhibits β-galactosidase and *lac* permease (Müller-Hill *et al.*, 1964). The use of TONPG was necessary since the repressed level of β-galactosidase in our plasmid system was too high to allow *galE* selection under normal conditions. Work is in progress to lower the expression of *lac* promoter of the pWB 300 plasmid.

After two days growth at 37°C all colonies were picked onto nutrient broth agar plates containing tetracycline and ampicillin and on identical ones containing 60 mg/l X-gal. Those which became blue on the X-gal plates were candidates to be picked for further analysis from plates containing no X-gal. For further detection of candidates we used the selection plates described above to which we added 60 mg/l X-gal or 60 mg X-gal and 0.3 g/l IPTG. Inducible colonies become more blue on the plates containing IPTG than on the others. From such colonies plasmid was extracted and treated with *Cla*I. *Cla*I cuts the pWB 300 derivatives three times, but does not cleave pWB 100. Then this DNA was retransformed into DC 41-2 F′ cells containing the same operator-variant plasmid as during selection. If the pWB 100 variant encodes a repressor variant which recognizes the operator variant, all transformants grow on the selection plate. The DNA sequence of the *Eco*RI–*Xho*I fragment of the pWB 100 mutant was determined according to Maxam and Gilbert (1977).

Acknowledgements

We thank Karin Otto and Daniela Tils for excellent technical help. This work was supported by Deutsche Forschungsgemeinschaft (SFB74-A2) and Bundesministerium für Forschung und Technologie (BCT 0 365/2).

References

Adler,K., Beyreuther,K., Fanning,E., Geisler,N., Gronenborn,B., Klemm,A., Müller-Hill,B., Pfahl,M. and Schmitz,A. (1972) *Nature*, **237**, 322–327.
Adhya,S. and Miller,W. (1979) *Nature*, **279**, 492–494.
Anderson,J.E., Ptashne,M. and Harrison,S.C. (1987) *Nature*, **326**, 846–852.
Benson,N., Sugiono,P., Bass,S., Mendelman,L.V. and Youderian,P. (1986) *Genetics*, **114**, 1–14.
Besse,M., Wilcken-Bergmann,B.v. and Müller-Hill,B. (1986) *EMBO J.*, **5**, 1377–1381.
Boelens,R., Scheek,R.M., van Boom,J.H. and Kaptein,R. (1987) *J. Mol. Biol.*, **193**, 213–216.
Chen,E.J. and Seeburg,P.H. (1985) *DNA*, **4**, 165–170.
Derbyshire,K.M., Salvo,J.J. and Grindley,N.D.F. (1986) *Gene*, **46**, 145–152.
Ebright,R.H. (1986) *Proc. Natl. Acad. Sci. USA*, **83**, 303–307.
Ebright,R.H., Cossart,P., Gicquel-Sanzey,B. and Beckwith,J. (1984) *Nature*, **311**, 232–235.
Eismann,E., Wilcken-Bergmann,B.v. and Müller-Hill,B. (1987) *J. Mol. Biol.*, **195**, 949–952.
Gicquel-Sanzey,B. and Cossart,P. (1982) *EMBO J.*, **1**, 591–595.
Gho,D. and Miller,J. (1974) *Mol. Gen. Genet.*, **131**, 137–146.
Hanahan,D. (1985) In Glover,D.M. (ed.), *DNA Cloning. A practical approach.* IRL Press, Oxford, Vol. I, pp. 109–135.
Hochschild,A. and Ptashne,M. (1986) *Cell*, **44**, 925–933.
Hochschild,A., Douhan,J. and Ptashne,M. (1986) *Cell*, **44**, 807–816.
Jacob,F. and Monod,J. (1961) *J. Mol. Biol.*, **3**, 318–356.
Koudelka,G.B., Harrison,S.C. and Ptashne,M. (1987) *Nature*, **326**, 886–888.
Lehming,N., Sartorius,J., Wilcken-Bergmann,B.v. and Müller-Hill,B. (1987a) In Guschlbauer,W. and Saenger,W. (eds), *DNA Ligand Interactions: From Drugs to Proteins.* Plenum Press, London, in press.
Lehming,N., Sartorius,J., Niemöller,M., Wilcken-Bergmann,B.v. and Müller-Hill,B. (1987b) In Quagliariello,E., Bernardi,G. and Ullmann,A. (eds), *From Enzyme Adaptation to Natural Philosophy: Heritage From Jacques Monod.* Elsevier, Amsterdam, pp. 53–68.
Maniatis,T., Fritsch,E.F. and Sambrook,J. (1982) *Molecular Cloning. A Laboratory Manual.* Cold Spring Harbor Laboratory Press, New York.
Matthews,B.W., Ohlendorf,D.H., Anderson,W.F. and Takeda,Y. (1982) *Proc. Natl. Acad. Sci. USA*, **79**, 1428–1432.
Maxam,A.M. and Gilbert,W. (1977) *Proc. Natl. Acad. Sci. USA*, **74**, 560–564.
Miller,J.H. (1972) *Experiments in Molecular Genetics.* Cold Spring Harbor Laboratory Press, New York.
Miller,J.H. (1978) In Miller,J.H. and Reznikoff,W.S. (eds), *The Operon.* Cold Spring Harbor Laboratory Press, New York, pp. 31–88.
Mossing,M.C. and Record,M.T. (1985) *J. Mol. Biol.*, **186**, 295–305.
Müller-Hill,B. (1975) *Prog. Molec. Biol.*, **30**, 227–252.
Müller-Hill,B., Beyreuther,K. and Gilbert,W. (1971) *Methods Enzymol.*, **21**, part D, 483–487.

Müller-Hill,B., Rickenberg,H.V. and Wallenfels,K. (1964) *J. Mol. Biol.*, **10**, 303–318.
Ogata,R.T. and Gilbert,W. (1978) *Proc. Natl. Acad. Sci. USA*, **75**, 5851–5854.
Ohlendorf,D.H., Anderson,W.F., Fisher,R.G., Takeda,Y. and Matthews,B.W. (1982) *Nature*, **298**, 718–723.
Ptashne,M. (1978) In Miller,J.H. and Reznikoff,W.S. (eds), *The Operon*. Cold Spring Harbor Laboratory Press, New York, pp. 325–344.
Ptashne,M. (1986) *A Genetic Switch, Gene Control and Phage* λ. Cell Press and Blackwell Scientific Publications, Oxford.
Riggs,A.D., Bourgeois,S. and Cohn,M. (1970) *J. Mol. Biol.*, **53**, 401–417.
Sadler,J.R., Sasmor,H. and Betz,J.L. (1983) *Proc. Natl. Acad. Sci. USA*, **80**, 6785–6789.
Sauer,R.T. and Pabo,C.O. (1984) *Annu. Rev. Biochem.*, **53**, 293–321.
Sauer,R.T., Yocum,R.R., Doolittle,R.F., Lewis,M. and Pabo,C.O. (1982) *Nature*, **298**, 447–451.
Schevitz,R.W., Otwinowski,Z., Joachimiak,A., Lawson,C.L. and Sigler,P.B. (1985) *Nature*, **317**, 782–786.
Simons,A., Tils,D., Wilcken-Bergmann,B.v. and Müller-Hill,B. (1984) *Proc. Natl. Acad. Sci. USA*, **81**, 1624–1628.
Valentin-Hansen,P., Hojrup,P. and Short,S. (1985) *Nucleic Acids Res.*, **13**, 5927–5936.
Weber,I.T. and Steitz,T.A. (1984) *Proc. Natl. Acad. Sci. USA*, **81**, 3973–3977.
Wharton,R.P. and Ptashne,M. (1987) *Nature*, **326**, 888–891.
Wilcken-Bergmann,B.v. and Müller-Hill,B. (1982) *Proc. Natl. Acad. Sci. USA*, **79**, 2427–2431.
Wilcken-Bergmann,B.v., Tils,D., Sartorius,J., Auerswald,E.A., Schröder,W. and Müller-Hill,B. (1986) *EMBO J.*, **5**, 3219–3225.

Received on July 3, 1987