Structural homology between *rbs* repressor and ribose binding protein implies functional similarity

C.A. MAUZY AND M.A. HERMODSON

Department of Biochemistry, Purdue University, West Lafayette, Indiana 47907

(RECEIVED November 18, 1991; REVISED MANUSCRIPT RECEIVED February 12, 1992)

Abstract

The deduced amino acid sequence of the *rbs* repressor, RbsR, of *Escherichia coli* is homologous over its C-terminal 272 residues to the entire sequence of the periplasmic ribose binding protein. RbsR is also homologous to a family of bacterial repressor proteins including LacI. This implies that the structure of the repressor consists of a two-domain binding protein portion attached to a DNA-binding domain having the four-helix structure of the LacI headpiece. The implications of these relationships to the mechanism of this class of repressors are discussed.

Keywords: binding proteins; chemotaxis; *Escherichia coli*; evolution; protein homology; repressors; ribose transport

The three-dimensional structures of only a few repressor proteins have been determined at atomic resolution. To our knowledge only in the case of the *trp* repressor is there enough structural information to define the transition from the aporepressor to the holo-repressor at a resolution that explains the change from low to high DNA-binding specificity (Zhang et al., 1987; Otwinowski et al., 1988). Working hypotheses have also been put forward (McKay et al., 1982) to describe the allosteric transition in the catabolite gene activator protein (CAP). However, little is known about the transitions in other proteins that regulate transcription from conformations that bind the appropriate target DNA sequences to those that do not. In many cases the physiological effectors are not even known.

A common theme in many DNA-binding proteins is the presence of a helix–turn–helix motif at the DNA-binding site (Weber & Steitz, 1984; Kaptein et al., 1985; Zhang et al., 1987; Aggarwal et al., 1988; Jordan & Pabo, 1988; Otwinowski et al., 1988; Wolberger et al., 1988). The location of the motif in the linear sequences of the proteins varies from the N-terminus (e.g., LacI), to the middle (e.g., TrpR), to the C-terminus (e.g., CAP). There is enough similarity in the sequences that form this

motif to raise the possibility that they all may have arisen from a common ancestor (Weber et al., 1982). However, the motif is a peptide sequence of only 22 amino acids, and the constraints required to form the helix–turn–helix structure may be the basis for the apparent similarity, at least for comparisons of widely separated species like viruses to bacteria (Weber et al., 1982, p. 5089).

We show here that RbsR (Mauzy & Hermodson, 1992) is a member of a family of homologous repressor proteins, including LacI. The amino-terminal 60 residues of each member of this protein family are predicted to form the bundle of four helices characteristic of proteins employing the helix–turn–helix DNA contact motif. Of special interest, however, is the observation that the C-terminal 270 residues of RbsR are homologous to the ribose binding protein. Homology between short sequence fragments of arabinose (ABP), galactose (GBP), and ribose (RBP) binding proteins and the *lac* and *gal* repressors had been noted in 1983 (Müller-Hill, 1983), and hypotheses relating to the sugar binding site in LacI (Sams et al., 1984) and various functional sites on LacI (Gordon et al., 1988) were proposed. We now show that the entire sequences of the mature sugar binding proteins are homologous to the repressors. This has substantial implications relative to the structures of this family of repressors and to the way in which effectors bind the repressors and exert their influence on transcription.

However, the observation (Fig. 2) that RbsR is homologous to the entire sequence of RBP makes possible the comparison of the C-terminal 270 residues of the repressor to known protein crystal structures. The ABP, GBP, and RBP structures have been solved at atomic resolution (Gilliland & Quioco, 1981; Vyas et al., 1983; Mowbray & Cole, 1992). The alignment of the primary structures of ABP and GBP shows 21% amino acid identities over 300 residues compared (Fig. 2), and the three-dimensional structures are remarkably similar, indeed almost superimposable (Vyas et al., 1983). In the RBP/GBP alignment (Fig. 2), 27% of the residues are identical. For the RBP/RbsR comparison, 63 of the 268 residues compared (23%) align with identical residues. This allows one to predict with confidence that the three-dimensional structures of RBP and RbsR will be very similar within that

Fig. 1. The six repressor sequences (Farabaugh, 1978; vonWilcken-Bergmann & Müller-Hill, 1982; Valentin-Hansen et al., 1986; Rolfes & Zalkin, 1988; Reidl et al., 1989; this work) were aligned employing minimal numbers of gaps in the sequences. Residues were highlighted when three or more amino acids in a column were identical. This understates the relationships somewhat in that there are many columns where two pairs are identical. The standard one-letter code for the amino acids is used, and gaps are indicated by dashes. * In order to save space, the N-terminal four residues of Mall and six residues of CytR are omitted as are the C-terminal seven residues of LacI. R, RbsR; G, GalR; L, LacI; M, Mall; C, CytR; P, PurR. The numbers refer to the residue numbers of RbsR with the zero above the indicated residue.

Fig. 2. GBP, RBP, and ABD sequences were aligned optimally (M.A. Hermodson, unpubl.), as were the six homologous repressors (Fig. 1). Without changing the GalR to RbsR or the binding protein alignments, the two sets were aligned. Residues were highlighted where two residues in a column were identical. Thus, some columns are highlighted with different shadings. The code is the same as in Figure 1. The entire sequences of the mature binding proteins are shown. The N-terminal 57 residues of RbsR and 58 residues of GalR, homologous to the LacI DNA-binding domain, are not shown. The numbers above the alignments refer to the ABP sequence and those below to the RbsR sequence with the last numeral above or below the indicated residue.

portion of RbsR that is homologous to RBP (Fig. 2). By extension, the three-dimensional structures of the five repressor proteins, where each of the pairwise comparisons gives a level of sequence identity of 30% or more (Fig. 1), must also be very similar to each other.

Thus, the repressors must be composed of a DNA binding domain of roughly the first 60 residues, having the bundle of helices found for the *lac* repressor head-piece (Kaptein et al., 1985) attached to the characteristic two-domains-on-a-hinge (Fig. 3) structure of the periplasmic sugar binding proteins (Gilliland & Quiocho, 1981; Mao et al., 1982; Vyas et al., 1983). The N-terminus of ABP (or GBP or RBP) is on the surface at one end of the molecule with the cleft in which the sugar binds on the opposite side of the domain (Gilliland & Quiocho, 1981; Vyas et al., 1983). There are only three residues at the N-terminus prior to the beginning of the first β -strand in the

binding proteins. The β -strand formed by residues 4–10 of the binding proteins is the second strand of a six-strand β -sheet forming the core of the lower domain shown in Figure 3 (Gilliland & Quiocho, 1981). Because the observed sequence homology implies that such key structural elements will be found in equivalent positions in the homologs (Matthews & Rossmann, 1985), the DNA-binding domain must be attached to the N-terminal end of the binding protein portion of the repressor molecule (Fig. 3) on the opposite side of the N-terminal domain from the effector-binding cleft. The 60 residues of the DNA-binding domain form a compact four-helix structure (Kaptein et al., 1985), which could not contact the effector-binding cleft. It is conceivable that the binding and release of effector could be "sensed" by the DNA-binding domain through contacts made by that domain with the "hinge" of the binding protein portion of the

Fig. 3. Stereo ribbon drawing of ABP (Gilliland & Quijcho, 1981) in an orientation that shows the binding cleft in the closed formation (arabinose bound) on the right side of the illustration. The N-terminus is at the bottom left and leads into the second strand of the six-strand β -sheet core of the N-terminal domain. The graphics program of J.P. Priestle (1988) was employed.

molecule (left side, Fig. 3), since the C-terminal domain (top, Fig. 3) rotates some 20° relative to the N-terminal domain when a binding protein releases the ligand, bringing the C-terminal domain closer to the point where the DNA-binding domain is attached (Mao et al., 1982; Sack et al., 1989; Luck & Falke, 1991). However, there is good reason to expect that this family of repressor proteins forms dimers or higher oligomers upon binding DNA, and in the *lac* repressor the sites that have been identified with dimer or tetramer contacts map to the C-terminal domain (see discussion below). If the ternary complex with DNA in this family of repressors involves at least a dimer of the repressor protein in each case, and if the protein-protein contacts are mainly or exclusively through the C-terminal domain of the binding protein-like part of the molecules, clearly the affinity of the protein for the

DNA would be drastically altered by the closing of the hinge upon inducer binding. Obviously, this model needs to be tested by further studies employing purified repressor protein.

Because ribose induces the *rbs* operon, RbsR with ribose (or a ribose metabolite) bound, i.e., the closed conformer in the binding protein portion, must be the conformation that has the lower affinity for the operator. The open conformation (best observed in Mao et al. [1982] and Sack et al. [1989]) thus becomes the form that binds DNA, acting as a repressor. The operons regulated by LacI and GalR are also subject to induction by sugars, a situation entirely analogous to that for the *rbs* operon. However, PurR regulates biosynthetic pathways, and end-products of the pathway act as corepressors (Rolfes & Zalkin, 1990a). It seems very unlikely that the corepressor would bind to the open form of the binding protein or that the aporepressor would assume the closed form that opens in the presence of the corepressor. Thus, it is most likely that the RbsR and PurR proteins have opposite binding characteristics to their respective operators with highest affinity being the open form for RbsR and the closed form for PurR. It is interesting to note that PurR and RbsR are the most similar proteins in this family at the sequence level (47% identical), something one may not expect for proteins with significantly different mechanisms.

Whereas earlier models relating LacI to the periplasmic binding proteins (Müller-Hill, 1983; Sams et al., 1984; Gordon et al., 1988) were based on very fragmentary sequence information (Müller-Hill, 1983) and were thus highly speculative, the basic concepts have now been shown to be correct. The models advanced here are far more precise and detailed and can be used to correct and refine the earlier ones. For instance, the models of Müller-Hill (1983) must be revised to reflect the two-domain structure of the binding proteins, with both domains being involved in sugar binding (Gilliland & Quijcho, 1981; Vyas et al., 1983) and also in transport (A. Binnie, H. Zhang, & M.A. Hermodson, unpubl.). The DNA binding domain of the repressors is an added third domain to that general structure. The predicted sugar binding site for LacI advanced by Sams et al. (1984) and the extensive use of that model to explain the results of a mutagenesis analysis of LacI by Gordon et al. (1988) must be largely revised to account for the more accurate alignment of sequences (Fig. 2) made possible by the complete data set. Sams et al. (1984) correctly aligned the segments in the vicinity of residues 150 and 205 of ABP with those of LacI, two regions where the binding protein/repressor similarity is particularly strong. But over half the proposed binding-site residues for LacI (Sams et al., 1984; Gordon et al., 1988) are far from the binding site based on the alignments in Figures 1 and 2. Equivalent residues in the sequences of the three binding proteins and four repressors are shown in Table 1. It is obvious from the lists of hy-

Table 1. Equivalent residues in the three binding proteins and four repressors^a

ABP	GBP	RBP	RbsR	PurR	GalR	LacI
K10	(Y10)	(S9)	T66	T67	G67	S69
(Q11)	(K11)	(T10)	A67	S68	D68	S70
E14	D14	N13	N70	A71	D71	L73
D89	N91	D89	D145	D146	[N145]	D149
D90	(K92)	R90	W146	W147	[R146]	V150
(L145)	H152	(T135)	K188	R190	I188	S191
(T147)	D154	(A137)	P190	T192	D190	S193
R151	R158	R141	R194	R196	R194	R197
N205	N211	N190	N245	G247	N245	N246
N232	(F233)	(V212)	I267	V269	I267	I268
(D235)	D236	D215	I270	I272	I270	V271
(P254)	N256	Q235	Q290	Q292	Y290	Q291

^a For ABP, GBP, and RBP, the residues that form hydrogen bonds to galactose (Vyas et al., 1991), glucose (Vyas et al., 1988), and ribose (Mowbray & Cole, 1992), respectively, are given with no parentheses. Residues shown in parentheses for the sugar binding proteins are in equivalent positions in the alignment in Figure 2, but do not make a hydrogen bond to the sugar. Based on the alignments in Figures 1 and 2, the equivalent residues in four of the repressors are shown (counting the initiator methionyl residue as number 1 even though it is not usually present in the mature protein). The brackets indicate a position of uncertainty due to placement of the gaps between residues 135 and 155 of the repressors.

drogen-bonding residues for the three binding proteins that the contacts between the various sugars and the side chains in the binding pocket vary a great deal. The reasons for this are clearly shown in the crystallographic studies of Vyas et al. (1988, 1991). Although the protein structures are very similar, the position of the sugar molecule varies drastically with respect to its orientation in the binding pocket. Thus, most predictions of liganding residues based on homology have been wrong for these proteins in the past (Argos et al., 1981; Sams et al., 1984; Gordon et al., 1988), at least in some details, and that is likely to be the case in any attempts to predict liganding residues for the repressors. One can say with confidence that the residues in Table 1 for the various proteins are in the vicinity of the binding cleft, and examination of the walls of the binding cleft from the ABP, GBP, and RBP crystal structures would add to the list. For other than predicting the residues lining the cleft, more detailed models are probably not warranted.

The extensive mutagenesis studies of Kleina and Miller (1990) identified a number of mutations that led to the I^s phenotype of the *lac* repressor. I^s mutants bind to the DNA even in the presence of inducer and thus can arise from mutations that reduce the affinity of the repressor for the inducer, those that result in increased affinity for the operator, or those that fail to undergo the proper allosteric transition upon inducer binding (Kleina & Miller, 1990). Most of the mutations that led to the I^s pheno-

type (Fig. 4 of Kleina & Miller, 1990) map to the inducer binding pocket by our analysis (Table 1). In particular, the clusters about residues 191–197, 246–249, 273, and 292 are all in the vicinity of the binding cleft. This would suggest that lowered inducer binding affinity accounts for the I^s phenotype at these sites. The residues at positions 84 and 88 of the LacI sequence correspond to residues near the hinge of the binding protein structures. Thus, failure of the proper allosteric transition is a probable explanation for the I^s phenotype at those sites. The residues from positions 95 to 99 and the one at 220 are surface residues near the inducer binding cleft; what effect they may have on inducer binding or the allosteric transition is hard to predict.

The alignment in Figure 2 of the binding protein sequences employs primary structure comparisons with minimal insertions of gaps. This makes the alignments somewhat different from the one based on the tertiary structures of ABP and GBP by Vyas et al. (1991), which employs large gaps. Nevertheless, only around residue 230 (ABP numbering) is there a disagreement regarding alignment of residues in the binding sites. The high degree of sequence similarity in this region makes the alignment in Figure 2 quite convincing.

Very recent work in Eisenberg's laboratory (Bowie et al., 1991) employing a three-dimensional (3D) profiling method has also detected the relationship between the sugar binding proteins and this family of repressors. This method employs searches for 3D structures into which a given sequence could fold. Thus, sequence similarity per se is not the criterion on which this approach is based. The results are entirely compatible with our sequence-based comparisons.

Although patterns of secondary and tertiary structures are highly conserved in homologous proteins and allow us to propose that the family of repressors (Fig. 1) is composed of a two-domain binding protein attached to a DNA reading head, much less is known regarding quaternary structures and their preservation in homologous systems. That issue is relevant for this family of repressor proteins, since LacI is known to form a tetramer when binding productively to the target DNA sequences (Platt et al., 1973). In addition, the symmetry of the operator sequences (Rolfes & Zalkin, 1990b; Mauzy & Hermodson, 1992) and almost universal precedent in other repressor systems makes it very likely that binding to the DNA requires at least a dimer of the repressor. The sites that bind inducer and those that are required to form dimers and tetramers in LacI all map to the region C-terminal to the 59-amino acid N-terminal headpiece (Platt et al., 1973), the region homologous to the binding proteins. There is no strong evidence for the sugar binding proteins acting as oligomers. One report (Richarme, 1983) suggests that GBP and maltose binding protein can form dimers under certain conditions, but binding of the respective sugars induces return to the monomeric state.

Another indicates that GBP forms a very stable dimer (Rasched et al., 1976). Other investigations show no tendency for binding proteins to form dimers or higher ordered states (Anraku, 1968). Because the LacI tetramer is formed in the absence of the inducer, it is conceivable that the weak tendency of the apo-binding proteins to form dimers as observed by Richarme (1983) is a vestige of the common ancestor to the binding protein and the repressors, but detailed structural analyses of the repressors in both the holo- and apo-repressor forms will be needed to address such points. The region from Ser 279 to Tyr 282 of LacI (analogous to residues 278–281 of RbsR, Fig. 1) has been suggested to be important for tetramer formation (Gordon et al., 1988); this is not a region of striking similarity between the binding proteins and the repressors (Fig. 2). Recent identification of a leucine zipper motif at the C-terminus of LacI (Alberti et al., 1991; Chakerian et al., 1991), which appears to be essential for the dimer-to-tetramer transition, does not appear to be relevant to the rest of this family, since it occurs in LacI C-terminal to the termination of the other members in the alignment in Figure 1.

The observation that a set of repressor proteins is related structurally, and by extension functionally, to a set of proteins that function in the initial events leading to membrane transport and chemotaxis is an excellent example of the parsimony of the evolutionary process. In both cases the presence or absence of a small effector molecule has to be "sensed" by a macromolecular complex in order for the proper biological response to occur. The simple solution of using a two-domain, hinged effector binding protein that has drastically different conformations depending on whether or not the effector is present and bound solves the first requirement, namely that the presence of the effector must cause a detectable change in the macromolecule that is its primary receptor. Once that requirement is met, gene duplication and domain shuffling (i.e., attachment of function-specific domains [Doolittle, 1985]) is all that is required to obtain molecules that perform completely different biological functions in response to the same (or similar) effectors. In this case, a signal sequence must be attached to the binding protein to effect its transport to the periplasm, and the DNA-binding domain is needed for the repressor function.

Just as engineers and scientists are advised not to reinvent the wheel, the process of evolution also employs successful inventions in multiple ways.

Methods

Protein alignments

The predicted amino sequence of RbsR was searched against the Protein Data Bank using the FastP protein alignment program from Dr. William Pearson, University

of Virginia (Lipman & Pearson, 1985). The data obtained from the RbsR-RBP alignment, along with previous binding protein alignments (Hermodson, unpubl.), were used to align the three sugar binding protein sequences with RbsR and GalR, inserting a minimum number of gaps in the sequences and maintaining consistency within the repressor protein alignments and the binding protein alignments. There is one region of uncertainty regarding placement of gaps between residues 135 and 155 of RbsR where Figures 1 and 2 differ slightly.

The alignments of the family of repressors were performed on the VAX system at the AIDS Center Laboratory for Computational Biochemistry (ACLCB) at Purdue University using the Wisconsin software package.

Note added in proof

Dr. Milton H. Saier, Jr., finds homology between the FruR protein, a new sequence he has determined, and the sugar binding proteins. The FruR protein is a member of this family of repressors (M. Saier, pers. comm.).

Acknowledgments

This work was supported by PHS grant GM24602 and NSF grant DMB 9003012. The AIDS Center Laboratory for Computational Biochemistry assisted with alignments using published software (Devereux et al., 1984). The Computational Biochemistry Laboratory is supported by PHS grant AI 27713. The assistance of Rick Westerman with computer analyses is gratefully acknowledged. The authors thank Howard Zalkin for helpful discussion of the concepts. The artistry of Dr. Jean-Pierre Wery with the graphics system is deeply appreciated, as is Ms. Jo Crain's expert assistance with the assembly of this manuscript. This is journal paper number 13233 of the Purdue University Agricultural Experiment Station.

References

- Aggarwal, A.K., Rodgers, D.W., Drott, M., Ptashne, M., & Harrison, S.C. (1988). Recognition of a DNA operator by the repressor of phage 434: A view at high resolution. *Science* 242, 899–907.
- Alberti, S., Oehler, S., vonWilcken-Bergmann, B., Krämer, H., & Müller-Hill, B. (1991). Dimer-to-tetramer assembly of *lac* repressor involves a leucine heptad repeat. *New Biol.* 3, 57–62.
- Anraku, Y. (1968). Transport of sugars and amino acids in bacteria. *J. Biol. Chem.* 243, 3123–3127.
- Argos, P., Mahoney, W.C., Hermodson, M.A., & Hane, M. (1981). Structural prediction of sugar-binding proteins functional in chemotaxis and transport. *J. Biol. Chem.* 256, 4357–4361.
- Bowie, J.U., Lüthy, R., & Eisenberg, D. (1991). A method to identify protein sequences that fold into a known three-dimensional structure. *Science* 253, 164–170.
- Chakerian, A.E., Tesmer, V.M., Manly, S.P., Brackett, J.K., Lynch, M.J., Hoh, J.T., & Matthews, K.S. (1991). Evidence for leucine zipper motif in lactose repressor protein. *J. Biol. Chem.* 266, 1371–1374.
- Devereux, J., Haeberli, P., & Smithies, O. (1984). A comprehensive set of sequence analysis programs for the VAX. *Nucleic Acids Res.* 12, 387–395.
- Doolittle, R.F. (1985). The genealogy of some recently evolved vertebrate proteins. *Trends Biochem. Sci.* 10, 233–237.
- Farabaugh, P. (1978). Sequence of the *lacI* gene. *Nature* 274, 765–769.

- Gilliland, G.L. & Quijcho, F.A. (1981). Structure of the L-arabinose-binding protein from *Escherichia coli* at 2.4 Å resolution. *J. Mol. Biol.* 146, 341–362.
- Gordon, A.J.E., Burns, P.A., Fix, D.F., Yatagai, F., Allen, F.L., Horsfall, M.J., Halliday, J.A., Gray, J., Bernelot-Moens, C., & Glickman, B.W. (1988). Missense mutation in the *lacI* gene of *Escherichia coli*. Inferences on the structure of the repressor protein. *J. Mol. Biol.* 200, 239–251.
- Jordan, S.R. & Pabo, C.O. (1988). Structure of the lambda complex at 2.5 Å resolution: Details of the repressor-operator interactions. *Science* 242, 893–899.
- Kaptein, R., Zuiderweg, E.R.P., Scheek, R.M., Boelens, R., & van Gunsteren, W.F. (1985). Structures of the lambda complex at 2.5 Å resolution: Details of the repressor-operator interactions. *J. Mol. Biol.* 182, 179–182.
- Kleina, L.G. & Miller, J.H. (1990). Genetic studies of the *lac* repressor. XIII: Extensive amino acid replacements generated by the use of natural and synthetic nonsense suppressors. *J. Mol. Biol.* 212, 295–318.
- Lipman, D.J. & Pearson, W.R. (1985). Rapid and sensitive protein similarity searches. *Science* 227, 1435–1441.
- Luck, L.A. & Falke, J.J. (1991). Open conformation of a substrate-binding cleft: ¹⁹F NMR studies of cleft angle in the D-galactose chemosensory receptor. *Biochemistry* 30, 6484–6490.
- Mao, B., Pear, M.R., McCammon, J.A., & Quijcho, F.A. (1982). Hinge-bending in L-arabinose-binding protein. *J. Biol. Chem.* 257, 1131–1133.
- Matthews, B.W. & Rossmann, M.G. (1985). Comparison of protein structures. *Methods Enzymol.* 115, 397–420.
- Mauzy, C.A. & Hermodson, M.A. (1992). Structural and functional analyses of the repressor, RbsR, of the ribose operon of *Escherichia coli*. *Protein Sci.* 1, 831–842.
- McKay, D.B., Weber, I.T., & Steitz, T.A. (1982). Structure of catabolite gene activator protein at 2.9-Å resolution. *J. Biol. Chem.* 257, 9518–9524.
- Mowbray, S.L. & Cole, L.B. (1992). 1.7 Å X-ray structure of the periplasmic ribose receptor from *Escherichia coli*. *J. Mol. Biol.* 225, 155–175.
- Müller-Hill, B. (1983). Sequence homology between Lac and Gal repressors and three sugar-binding periplasmic proteins. *Nature* 302, 163–164.
- Otwinowski, Z., Schevitz, R.W., Zhang, R.-G., Lawson, C.L., Joachimiak, A., Marmorstein, R.Q., Luisi, B.F., & Sigler, P.B. (1988). Crystal structure of *trp* repressor/operator complex at atomic resolution. *Nature* 335, 321–329.
- Platt, T., Files, J.G., & Weber, K. (1973). *Lac* repressor. Specific proteolytic destruction of the NH₂-terminal region and loss of the deoxyribonucleic acid-binding activity. *J. Biol. Chem.* 248, 110–121.
- Priestle, J.P. (1988). RIBBON: A stereo cartoon drawing program for proteins. *J. Appl. Crystallogr.* 21, 572–576.
- Rasched, I., Shuman, H., & Boos, W. (1976). The dimer of the *Escherichia coli* galactose-binding protein. *Eur. J. Biochem.* 69, 545–550.
- Reidl, J., Römisch, K., Ehrmann, M., & Boos, W. (1989). Mall, a novel protein involved in regulation of the maltose system of *Escherichia coli* is highly homologous to the repressor proteins GalR, CytR, and LacI. *J. Bacteriol.* 171, 4888–4899.
- Richarme, G. (1983). Associative properties of the *Escherichia coli* galactose-binding protein and maltose-binding protein. *Biochim. Biophys. Acta* 748, 99–108.
- Rolfes, R.J. & Zalkin, H. (1988). *Escherichia coli* gene *purR* encoding a repressor protein for purine nucleotide synthesis. Cloning, nucleotide sequence, and interaction with the *purF* operator. *J. Biol. Chem.* 263, 19653–19661.
- Rolfes, R.J. & Zalkin, H. (1990a). Purification of the *Escherichia coli* purine regulon repressor and identification of corepressors. *J. Bacteriol.* 172, 5637–5642.
- Rolfes, R.J. & Zalkin, H. (1990b). Autoregulation of *Escherichia coli purR* requires two control sites downstream of the promoter. *J. Bacteriol.* 172, 5758–5766.
- Sack, J.S., Saper, M.A., & Quijcho, F.A. (1989). Periplasmic binding protein structure and function. Refined x-ray structures of the leucine/isoleucine/valine-binding protein and its complex with leucine. *J. Mol. Biol.* 206, 171–191.
- Sams, C.F., Vyas, N.K., Quijcho, F.A., & Matthews, K.S. (1984). Predicted structure of the sugar-binding site of the *lac* repressor. *Nature* 310, 429–430.
- Valentin-Hansen, P., Larsen, J.E.L., Højrup, P., Short, S.A., & Barbier, C.S. (1986). Nucleotide sequence of the CytR regulatory gene of *E. coli* K-12. *Nucleic Acids Res.* 14, 2215–2228.
- von Wilcken-Bergmann, B. & Müller-Hill, B. (1982). Sequence of *galR* gene indicates a common evolutionary origin of *lac* and *gal* repressor in *Escherichia coli*. *Proc. Natl. Acad. Sci. USA* 79, 2427–2431.
- Vyas, N.K., Vyas, M.N., & Quijcho, F.A. (1983). The 3 Å resolution structure of a D-galactose-binding protein for transport and chemotaxis in *Escherichia coli*. *Proc. Natl. Acad. Sci. USA* 80, 1792–1796.
- Vyas, N.K., Vyas, M.N., & Quijcho, F.A. (1988). Sugar and signal transducer binding sites of the *Escherichia coli* galactose chemoreceptor protein. *Science* 242, 1290–1295.
- Vyas, N.K., Vyas, M.N., & Quijcho, F.A. (1991). Comparison of the periplasmic receptors for L-arabinose, D-glucose/D-galactose, and D-ribose. Structural and functional similarity. *J. Biol. Chem.* 266, 5226–5237.
- Weber, I.T., McKay, D.B., & Steitz, T.A. (1982). Two helix DNA binding motif of CAP found in *lac* repressor and *gal* repressor. *Nucleic Acids Res.* 10, 5085–5102.
- Weber, I.T. & Steitz, T.A. (1984). Model of specific complex between catabolite gene activator protein and B-DNA suggested by electrostatic complementarity. *Proc. Natl. Acad. Sci. USA* 81, 3973–3977.
- Wolberger, C., Dong, Y., Ptashne, M., & Harrison, S.C. (1988). Structure of a phage 434 cro/DNA complex. *Nature* 335, 789–795.
- Zhang, R.-G., Joachimiak, A., Lawson, C.L., Schevitz, R.W., Otwinowski, Z., & Sigler, P.B. (1987). The crystal structure of *trp* aporepressor at 1.8 Å shows how binding tryptophan enhances DNA affinity. *Nature* 327, 591–597.