MICROBIOLOGY AND MOLECULAR BIOLOGY REVIEWS, Dec. 1997, p. 393–410 Vol. 61, No. 4
1092-2172/97/$04.00+0

AraC/XylS Family of Transcriptional Regulators

MARÍA-TRINIDAD GALLEGOS,[1]† ROBERT SCHLEIF,[2] AMOS BAIROCH,[3] KAY HOFMANN,[4] AND JUAN L. RAMOS[1]*

Department of Biochemistry and Molecular and Cellular Biology of Plants, Consejo Superior de Investigaciones Científicas, Estación Experimental del Zaidín, Granada, Spain[1]; Department of Biology, Johns Hopkins University, Baltimore, Maryland[2]; and Department of Medical Biochemistry, Centre Medicale Universitaire, University of Geneva, Geneva,[3] and Bioinformatics Group, Swiss Institute for Experimental Cancer Research Center (ISREC), Lausanne,[4] Switzerland

INTRODUCTION

Searches for homology among protein sequences can identify well-conserved motifs such as cofactor binding domains, transient peptides, helix-turn-helix and zinc finger DNA-binding motifs, and others. This approach can also identify families of related proteins in which homology extends over one or several domains of proteins that possess similar functions (5, 7, 21, 102, 103, 167, 195–197, 257).

Within the current sequence databases, the AraC/XylS family of regulators is one of the most common positive regulators (84, 209, 243). Other common regulator families are ArsR (175), AsnC (140), Crp (233), DeoR (17), GntR (29, 104, 213), IclR (214), LacI (253), LuxR/UphA (112), LysR (111), MarR (56), MerR (107), NtrC (189), TetR (198), YedF/YeeD/YhhP (14), and YhdG/YjbN/YohI (15).

The AraC/XylS family is characterized by significant amino acid sequence homology extending over a 100-residue stretch constituting the DNA binding domain of the family members. The domain is most often found in oligomeric proteins, but in a few natural cases (4, 18, 47, 86, 151, 260) and in artificial cases (37, 143, 170) the single conserved domain itself can bind to DNA and activate transcription from cognate promoters. The domain does not appear to bind effector molecules, this function being provided either by additional domains in the family members or by other proteins that regulate the synthesis of AraC/XylS family members.

AraC, the regulator of the L-arabinose operon in *Escherichia coli*, was the first member to be identified, purified, and characterized biochemically (95, 223–226). Tobin and Schleif (243) envisaged that AraC, RhaS, and RhaR defined a group of transcriptional regulators. Later, Ramos et al. (209) and Henikoff et al. (112) suggested that eight proteins (AraC, RhaR, RhaS, MelR, and Rns from *E. coli*; XylS from *Pseudomonas putida*; AraC from *Erwinia carotovora*; and VirF from *Yersinia enterocolitica*) formed an incipient family. In 1993, Gallegos et al. (84) extended the family to include 27 proteins with the addition of AdaA from *Bacillus subtilis*; AraC from *Citrobacter freundii*; AppY (also called M5), CelD, CfaD, EnvY, FapR, SoxS, TetD from *E. coli*; ExsA and MmsR from *Pseudomonas aeruginosa*; VirF from *Shigella flexneri*; AraC and RhaS from *Salmonella typhimurium*; TcpN (also called ToxT) from *Vibrio cholerae*; LcrF from *Yersinia pestis*; and several natural XylS proteins from different TOL plasmids. These proteins were aligned with the PILEUP program, which made it possible to define a 99-amino-acid stretch of homology at the C terminus of these proteins.

In this review, we have extended the family to include more than 100 proteins and polypeptides derived from open reading frames (ORFs) translated from DNA sequences. Here we summarize and discuss the general distinguishing characteris-

* Corresponding author. Mailing address: CSIC-Estación Experimental del Zaidín, Apdo. Correos 419, E-18008 Granada, Spain. Phone: 34-58-121011. Fax: 34-58-129600.

† Present address: Department of Biology, Imperial College of Science, Technology and Medicine, London, United Kingdom.

tics of the family, the structure-function organization of the AraC/XylS family of polypeptides, and the biochemical and molecular aspects of their mode(s) of action.

CURRENT MEMBERS OF THE AraC/XylS FAMILY

Successive Search for Members of the Family

To identify new members of the AraC/XylS family of transcriptional regulators, the 99 amino acids of the C-terminal end of the 27 proteins identified as members of the family (84) were aligned and analyzed by the algorithm of Lüthy et al. (160). This allowed us to define a matrix for the profile of the aligned sequences. This profile was then used to search for new members of the family within protein databases (SWISSPROT and PIR). The newly identified proteins were retrieved and aligned with the previously identified members of the family, and a new profile was again defined. This new profile was then used to search for putative polypeptides as members of the family by searching in nucleic acid databases (EMBL, GenBank, Genpept, and TREMBL), where we identified ORFs whose translated sequences gave a polypeptide that was a probable regulatory protein and a member of this family. Finally, a new alignment of all found sequences was carried out. This again defined the 99-amino-acid stretch as the most highly conserved region of this family of proteins, and a new profile was defined by analyzing the segment with the algorithm of Lüthy et al. (160). This new profile was used to search for members of the family in all available protein and nucleic acid databases (in March 1997), but no new members were identified. This profile therefore now defines the AraC/XylS family of transcriptional regulators. It can be accessed from the PROSITE database as entry PS01124.

Table 1 lists the current proteins identified from protein and DNA databases (March 1997) as members of the AraC/XylS family. All characterized proteins of the family are positive transcriptional regulators except CelD, which seems to be a repressor (199). As shown in Table 1, members of the AraC/XylS family regulate very diverse genes and functions. Some members of the family control single operons or genes; others control multiple, unlinked target genes (regulons); while others are themselves regulated by other genes, forming complex regulatory networks (184).

Analyses of a protein sequence with the matrix assigned a value to the query sequence. The value assigned by the matrix to each of the family members ranged from 30.74 to 12.52, with small variations between two consecutive proteins identified as members of the family (the complete set of values is available from K. Hofmann). However, a difference of 4.7 points was observed between the last member of the family assigned by the profile, namely, Hrp from *Xanthomonas oryzae*, and the closest value of a protein not identified as a member of the family, MutS from *Thermus thermophilus*. Therefore, we propose that a protein belongs to the AraC/XylS family if the value after analysis with the profile defined in PROSITE PS01124 is above 12.52.

Nonetheless, comparison of a query sequence with the conserved domain of any of the family members can identify the query sequence as a member of the family. The sequence can then be rapidly aligned to any of the homologs with the FASTA program.

Functions Regulated by Members of the Family

As mentioned above, all proteins in the AraC/XylS family are positive transcriptional activators except CelD, which seems to act as a repressor (199). Two members of the family, the AraC protein from *E. coli* and the YbtA protein from *Y. pestis*, can function both as a repressor and as a positive regulator (73, 224, 225) in different promoters or in the same promoter depending on the presence or the absence of appropriate effectors.

Two types of proteins are distinguished in the family: in one group, the signal receptor resides in the same polypeptide as the regulatory function (i.e., AraC, XylS, RhaR, and UreR) (59, 63, 172, 185, 210, 232, 243); in the other group, transcription of the regulatory protein is controlled by another regulator. This regulator can be an activator or a repressor, so that stimulation or derepression of transcription leads to the overexpression of the member of the AraC/XylS family, which in turn regulates transcription from cognate promoters (i.e., MarA, SoxS, and TcpN) (4, 113, 129, 190, 260). The proteins belonging to the family have three main regulatory functions in common: carbon metabolism, stress response, and pathogenesis.

Regarding carbon metabolism, members of the family control the degradation of sugars such as arabinose (AraC), cellobiose (CelD), melibiose (MelR and MsmR), raffinose (RafR and MsmR), rhamnose (RhaR and RhaS), and xylose (XylR); certain amino acids such as valine (MmsR), arginine (AdiY), and ornithine (OruR); alcohols such as 1,2-propanediol (PocR); alkylbenzoates (XylS); *p*-hydroxyphenylacetic acid (HpaA); and herbicides such as *S*-ethyl dipropylthiocarbamate (TchR). These transcriptional regulators are characterized by the fact that they stimulate transcription from cognate promoters in response to the presence of the effector. All of them are about 300 amino acids long (30, 32, 74, 105, 169, 178, 199, 204, 218, 238, 248, 251). The three best-characterized proteins in this subgroup of the family are AraC, RhaR, and XylS (84, 224, 244; see below for further details). Certain regulatory proteins are involved in the production of virulence factors in infections of plants (HrpB from *Burkholderia solanacearum*) or mammals. Among the latter, these regulatory factors have been found in microbes that colonize mainly the gastrointestinal tract but also the respiratory tract or the urinary system. These factors include AfrR, AggR, CfaD, CsvR, FapR, PerA, and Rns from *E. coli*; CafR and LcrF from *Y. pestis*; ExsA and PchR from *P. aeruginosa*; InvF from *S. typhimurium*; MxiE from *Shigella flexneri*; TcpN from *V. cholerae*; and VirF from *Shigella* and *Y. pestis* (36, 55, 77, 85, 91, 106, 113, 127, 129, 133, 183, 192, 222, 258, 263).

These proteins are involved in stimulation of the synthesis of proteins that play a role in adhesion to epithelial tissues, such as fimbriae (AfrR, AggR, CfaD, CsvR, FapR, PerA, Rns, and TcpN), components of the cell capsule (CafR), and invasins (ExsA, HrpB, InvF, MxiE, and VirF). Some members of the family control the production of other virulence factors such as siderophores (PchR) and urease (UreR). These regulators can be plasmid or chromosomally encoded.

Except for UreR, which binds urea to become active (58), it has not been demonstrated that regulators of this group bind specific effectors, although all of them respond to environmental factors such as temperature, osmolarity of the medium, and concentration of Ca^{2+} (20, 53, 55, 76, 115, 124, 141, 182, 202, 230, 242, 261).

Some regulators are involved in the response to stressors, e.g., response to alkylating agents (Ada from *E. coli*, *S. typhimurium*, and *Mycobacterium tuberculosis* and AdaA from *Bacillus subtilis*) (57, 100, 176, 177); response to oxidative stress (SoxS from *E. coli* and *S. typhimurium*) (4, 260); tolerance to antibiotics, organic solvents, and heavy metals (AarP from *Providencia stuartii*, MarA and Rob from *E. coli*, PqrA from *Proteus vulgaris*, and RamA from *Klebsiella pneumoniae*) (90, 121, 161, 229, 240); and transition from exponential growth to

TABLE 1. Members of the AraC/XylS family of transcriptional regulators

Protein	Microorganism	Accession no.[a]	Function	No. of residues	Reference(s)
AarP	*Providencia stuartii*	SP:P43463	Transcriptional activator of *acc* (2′) *Ia* gen for 2′-*N*-acetyltransferase	135	161
Ada	*Escherichia coli*	SP:P06134	Repair of alkylated guanine in DNA by stoichiometrically transferring the alkyl group at the O-6 position to a cysteine residue in the enzyme in a suicide reaction, because the enzyme is irreversibly inactivated; can also repair *O*-4-methylthymine. The methylated Ada protein is a positive regulator of its own synthesis (*ada*) and that of other genes, such as *alkA*, *alkB* and *aidB*	354	57, 138, 150, 180
Ada	*Mycobacterium tuberculosis*	SP:Q10630	Similar to *E. coli* Ada	496	177
Ada	*Salmonella typhimurium*	SP:P26189	Similar to *E. coli* Ada	352	100
AdaA	*Bacillus subtilis*	SP:P19219	One of the two proteins required for the adaptative response to alkylating agents. It accepts a methyl group from methylphosphotriesters and then acts as a transcriptional activator of the *ada* operon	211	176
AdiY	*Escherichia coli*	SP:P33234	Transcriptional activator of the *adiA* gene for biodegradative acid-induced arginine decarboxylase	253	32, 237
AfrR	*Escherichia coli*	TE:Q07681	Probable transcriptional activator of the *afrABRS* operon for expression of AF/R1 fimbria in *E. coli* RDEC-1, a rabbit pathogen	272	258
AggR	*Escherichia coli*	SP:P43464	Transcriptional activator of the *aggA* gene for aggregative adherence fimbria I (AAF/I) expression in enteroaggregative *E. coli* strains	265	183
AppY	*Escherichia coli*	SP:P05052	Transcriptional activator of the *cyxAB*, *hyaABCDEF* and *appA* operons during the deceleration phase of growth	243	12, 131
AraC	*Citrobacter freundii*	SP:P11765	Regulator of several operons involved in the transport and catabolism of L-arabinose (similar to *E. coli* AraC)	281	30
AraC	*Escherichia coli*	SP:P03021	Activator of the expression of the *araBAD*, *araFGH* and *araE* operons, which are involved in the transport and catabolism of L-arabinose. Repressor of its own synthesis	292	174, 238, 248, 266
AraC	*Erwinia chrysanthemi*	SP:P07642	Similar to *E. coli* AraC	310	149
AraC	*Salmonella typhimurium*	SP:P03022	Similar to *E. coli* AraC	281	46
AraL	*Streptomyces antibioticus*	SP:Q03320	Unknown	303	265
AraL	*Streptomyces lividans*	SP:P35319	Unknown	304	43
CafR	*Yersinia pestis*	SP:P26950	Positive regulator of *caf1MA* and *caf1* operons for the production and transport of the capsule antigen F1	301	85, 128
CelD	*Escherichia coli*	SP:P17410	Repressor of the *celABCF* operon involved in the degradation of cellobiose, arbutin, and salicin	280	199
CfaD	*Escherichia coli*	SP:P25393	Transcriptional activator of the *cfaABCE* operon for the production of CFA/I fimbriae in enterotoxigenic *E. coli* strains	265	222
CsvR	*Escherichia coli*	SP:P43460	Transcriptional activator of the operon involved in the production of CS5 fimbriae in enterotoxigenic *E. coli* strains	301	55
EnvY	*Escherichia coli*	SP:P10805	Transcriptional temperature-dependent activator of several *E. coli* envelope proteins, most notably the porins OmpF and OmpC and the λ receptor, LamB	253	159
ExsA	*Pseudomonas aeruginosa*	SP:P26993	Transcriptional activator of the *exsCBA* operon and *exsD*, *exoS*, and *ORF1* genes required for the synthesis and secretion of exoenzyme S	298	77
FapR	*Escherichia coli*	SP:P23774	Transcriptional activator of the 987P operon for fimbrial proteins in enterotoxigenic *E. coli* strains	260	133
HpaA	*Escherichia coli*	TE:Q46985	Transcriptional activator of *hpaBC* operon for catabolism of *p*-hydroxyphenylacetic acid	295	204

Continued on following page

TABLE 1—*Continued*

Protein	Microorganism	Accession no.	Function	No. of residues	Reference(s)
HrpB	*Burkholderia solanacearum*	SP:P31778	Transcriptional activator of the hypersensitive response genes (*hrp*) involved in plant pathogenicity	477	89
HrpXc	*Xanthomonas campestris*	TE:Q56801	Similar to *B. solanacearum* HrpB	503	193
HrpXv	*Xanthomonas campestris* pv. vesicatoria	TE:Q56790	Similar to *B. solanacearum* HrpB	476	254
HrpXo	*Xanthomonas oryzae*	TE:Q56831	Similar to *B. solanacearum* HrpB	502	193
InvF	*Salmonella typhimurium*	SP:P39437	Transcriptional activator of the *inv* operon required for epithelial tissue invasion	216	127
LcrF	*Yersinia pestis*	SP:P28808	Transcriptional activator of the virulence regulon (similar to *Y. enterocolitica* VirF)	271	116
LumQ	*Photobacterium leiognathi*	SP:Q51872	Probable transcriptional regulator	248	153
LumQ	*Synechocystis* sp.	TE:P73364	Unknown	241	126
MaoB	*Escherichia coli*	SP:Q47129	Transcriptional activator of the *maoA* gene coding a monoamine oxidase	301	263
MarA	*Escherichia coli*	SP:P27246	Transcriptional activator of the *sodA*, *zwf*, *micF*, *slp*, *fpr*, *fumC*, and *nfo* genes, which are involved in the multiple antibiotic resistance (mar) phenotype	129	47, 86
MarA	*Salmonella typhimurium*	SP:Q56070	Similar to *E. coli* MarA	129	240
MelR	*Escherichia coli*	SP:P10411	Transcriptional activator of the *melAB* operon for transport and catabolism of melibiose	302	32, 251
MmsR	*Pseudomonas aeruginosa*	SP:P28809	Transcriptional activator of the *mmsAB* operon for valine catabolism	307	236
MsmR	*Streptococcus mutans*	SP:Q00753	Transcriptional activator of the *msm* operon (*msmEFGK*, *aga*, *dexB*, *gftA*) required for the transport of melibiose, raffinose, and isomaltotriose and for melibiose, saccharose, and isomaltosaccharide catabolism	278	219
MxiE	*Shigella flexneri*	SP:Q04642	Transcriptional activator of *mxi* and *spa* operons involved in the synthesis and secretion of the Ipa proteins required for the epithelial tissue invasion	210	3
MxiE	*Shigella sonnei*	SP:Q55292	Similar to *S. flexneri* MxiE	210	6
NitR	*Rhodococcus rhodochrous*	TE:P72312	Transcriptional activator of *nitA*, which codes for a nitrilase	319	137
OruR	*Pseudomonas aeruginosa*	TE:P72171	Probable transcriptional activator of the ornithine utilization operon	339	105
PcrR	*Synechocystis* sp.	TE:P72600	Unknown	346	126
PchR	*Synechocystis* sp.	TE:P72595	Unknown	326	126
PchR	*Synechocystis* sp.	TE:P72608	Unknown	330	126
PchR	*Pseudomonas aeruginosa*	SP:P40883	Transcriptional activator of the pyochelin and ferripyochelin receptor	296	106
PerA	*Escherichia coli*	SP:P43459	Transcriptional activator of the *eaeA* gene for intimin, a protein required for adherence to the host cell membrane, in enterohemorrhagic and enteropathogenic *E. coli* strains	205	91
PobR	*Pseudomonas aeruginosa*	TE:Q51543	Probable transcriptional activator of *pobA*, which codes the *p*-hydroxybenzoate hydroxylase	288	68
PocR	*Salmonella typhimurium*	SP:Q05587	Transcriptional activator of *cbiABCDETFGHJKLMNQOP* and *cobUST* operons, required for the adenosyl-cobalamine (vitamin B_{12}) synthesis, and *pduABC* and *pduF*, required for 1,2-propanediol catabolism. Also regulates its own synthesis	303	42, 218
PqrA	*Proteus vulgaris*	SP:Q52620	Probable transcriptional activator of genes and/or operons responsible of multidrug resistance	122	121
RafR	*Pediococcus pentosaceus*	SP:P43465	Transcriptional activator of the operon for raffinose catabolism	277	147
RamA	*Klebsiella pneumoniae*	SP:Q48413	Probable transcriptional activator that confers multidrug resistance phenotype	113	90
RhaR	*Escherichia coli*	SP:P09378	Transcriptional activator of the operon *rhaSR* involved in the regulation of rhamnose catabolism	312	201, 243
RhaR	*Salmonella typhimurium*	SP:P40865	Similar to *E. coli* RhaR	106 (partial)	241

Continued on following page

TABLE 1—*Continued*

Protein	Microorganism	Accession no.	Function	No. of residues	Reference(s)
RhaS	*Escherichia coli*	SP:P09377	Transcriptional activator of genes required for the L-rhamnose catabolism (*rhaBAD*) and the genes which codify the rhamnose transporter (*rhaT*)	278	201, 243
RhaS	*Salmonella typhimurium*	SP:P27029	Similar to *E. coli* RhaS	277	187
Rns	*Escherichia coli*	SP:P16114	Transcriptional activator of the *csoBACE* operon, which codes the protein for CS1 or CS2 fimbriae in enterotoxigenic *E. coli* strains	265	36
Rob	*Escherichia coli*	SP:P27292	Binds to the right arm of the replication origin *oriC* of the *E. coli* chromosome; also involved in resistance to antibiotics, heavy metals, and superoxide stress and in tolerance to organic solvents	289	32, 229
SoxS	*Escherichia coli*	SP:P22539	Transcriptional activator of the superoxide response regulon which includes at least 10 genes such as *acnA* (aconitase), *fpr* (NADPH-ferredoxin oxidoreductase), *fumC* (fumarase C), *inaA* (unknown), *micF* (an antisense inhibitor of *ompF*), *nfo* (endonuclease IV), *pqi-5* (unknown), *ribA* (GTP cyclohydrolase), *sodA* (Mn-superoxide dismutase), and *zwf* (glucose-6-phosphate dehydrogenase)	106	4, 18, 151, 260
SoxS	*Salmonella typhimurium*	SP:Q56143	Similar to *E. coli* SoxS	106	166
TcpN	*Vibrio cholerae*	SP:P29492	Transcriptional activator of *tcpABYCDZEFMONJacfBC*, *tcpI*, *tcpH*, *acfA*, *acfD*, *ctxAB* operons required for epithelial tissue colonization	276	113, 130, 192
TetD	Tn*10*	SP:P28816	Unknown	138	22, 227
ThcR	*Rhodococcus* sp.	SP:P43462	Transcriptional activator of the *thc* operon for the degradation of the thiocarbamate herbicide EPTC	332	178
UreR	*Enterobacteriaceae*		Transcriptional activator of the *ureDABCEFG* operon for urease production; *P. stuartii* and *Salmonella* proteins are 98% identical to the *E. coli* protein		
	Escherichia coli	SP:P32326		296	58, 185
	Proteus vulgaris	SP:Q02458		293	
	Providencia stuartii				
	Salmonella sp.				
V38K	*Mycobacterium tuberculosis*	SP:Q06861	Probable role in the regulation of proteins necessary for virulence	339	97
VirF	*Shigella dysenteriae*	SP:Q04248	Transcriptional activator of the *virB* and *virG* genes. VirB is itself an activator of the *ipaABCD* virulence regulon; *S. flexneri* and *S. sonnei* proteins are identical to the *S. dysenteriae* protein	262	129, 220, 264
	Shigella flexneri				
	Shigella sonnei				
VirF	*Yersinia enterocolitica*	SP:P13225	Transcriptional activator of the *Yersinia* virulence regulon comprising *yop*, *ysc*, *yadA* and *ylpA* genes; the *Y. pseudotuberculosis* protein is 99% identical to the *Y. enterocolitica* protein	271	53
	Yersinia pseudotuberculosis				
XylR	*Escherichia coli*	SP:P37390	Probable transcriptional activator of the *xylBAFGHR* operon, which seems to be implicated in the catabolism of xylose	392	231
XylR	*Haemophilus influenzae*	SP:P45043	Similar to *E. coli* XylR	387	74
XylS	*Pseudomonas putida*	SP:P07859	Transcriptional activator of the pWW0 plasmid *meta* operon (*xylXYZLTEGFJQKIH*), required for the degradation of benzoate and substituted derivatives	321	119, 169, 235
XylS1	*Pseudomonas putida*	SP:Q04710	Transcriptional activator of the pWW53 plasmid *meta* 1 and 2 operons for benzoate catabolism and substituted derivatives	321	10

Continued on following page

TABLE 1—*Continued*

Protein	Microorganism	Accession no.	Function	No. of residues	Reference(s)
XylS2	*Pseudomonas putida*	SP:Q05092	Pseudogen present in pDK1 and pWW53 plasmids	157	10
XylS3	*Pseudomonas putida*	SP:Q05335	Transcriptional activator of the pWW53 plasmid *meta* 1 and 2 operons for benzoate catabolism and substituted derivatives	331	10
XylS4	*Pseudomonas putida*	SP:Q04713	Transcriptional activator of the pDK1 plasmid *meta* 1 and 2 operons for benzoate catabolism and substituted derivatives	331	10
Ya52	*Haemophilus influenzae*	SP:P45008	Unknown	298	74
Ybbb	*Bacillus subtilis*	SP:P40408	Unknown	529	205
YbtA	*Yersinia pestis*	TE:Q56951	Unknown	319	73
Ycgk	*Alteromonas carragenovora*	SP:P43461	Unknown	166	16
Yfeg	*Escherichia coli*	SP:P36547	Unknown	350	245
Yfif	*Bacillus subtilis*	SP:P54722	Unknown	314	262
Yhiw	*Escherichia coli*	SP:P37638	Unknown	242	231
Yhix	*Escherichia coli*	SP:P37639	Unknown	274	231
Yidl	*Escherichia coli*	SP:P31449	Unknown	307	31
Yijo	*Escherichia coli*	SP:P32677	Unknown	283	18
YisR	*Bacillus subtilis*	SP:P40331	Unknown	195 (partial)	34
Ymcr	*Streptomyces lavendulea*	SP:P43458	Unknown	281	13
	Mycobacterium tuberculosis	TE:P71663	Unknown	360	200
	Escherichia coli	TE:P76241	Unknown	273	19
	Escherichia coli	TE:P77379	Unknown	284	61
	Escherichia coli	TE:P77396	Unknown	285	2, 19
	Escherichia coli	TE:P77402	Unknown	303	1
	Escherichia coli	TE:P77601	Unknown	239	61
	Escherichia coli	TE:P77634	Unknown	265	19, 45
	Salmonella typhimurium	TE:Q04819	Unknown	259	80
	Azorhizobium caulinodans	TE:Q43970	Unknown	227	88
	Escherichia coli	TE:Q46855	Unknown	375	19
	Lactobacillus helveticus	TE:Q48557	Unknown	87 (partial)	60
	Burkholderia cepacia	TE:Q51600	Unknown	53 (partial)	98
	Pseudomonas diminuta	TE:Q51695	Unknown	168 (partial)	148
	Rhizobium leguminosarum	TE:Q52799	Unknown	296	259
	Streptomyces aureofaciens	TE:Q53603	Unknown	137 (partial)	139
	Streptomyces hygroscopicus	TE:Q54308	Unknown	330	228
	Mycobacterium tuberculosis	GP:1781124	Unknown	263	200
	Mycobacterium tuberculosis	GP:1806231	Unknown	259	200

[a] SP, SWISSPROT; TE, TREMBL; GP, Genpept.

the stationary phase (AppY from *E. coli*) (12, 131). Some members of this group of proteins are highly homologous to each other, and some of them—SoxS, MarA, and Rob—cross-regulate certain genes (8, 48). These proteins apparently need to be overproduced to exert their regulatory role (9, 86, 181).

No specific regulatory function has yet been assigned to several members of the family (Table 1) (EnvY, Yfeg, Yhiw, Yhix, Yidl, and Yijo from *E. coli*; AraL from *Streptomyces antibioticus* and *Streptomyces lividans*; TetD from Tn*10*; Ya52 from *Haemophilus influenzae*; YcgK from *Alteromonas carragenovora*; PccR, PchR, and LumQ from *Synechocystis* sp.; AraC from *Azorhizobium caulinodans*; PobR from *Rhizobium leguminosarum*; Hpr from *Xanthomonas campestris* and *Xanthomonas oryzae*; YmcR from *Streptomyces lavendulae*; and Ybbb, Yfif, and YisR from *B. subtilis*).

Distribution and Evolution

Members of the AraC/XylS family are widely distributed in diverse prokaryote genera (Table 1). The G+C content of genes encoding AraC/XylS family members vary from 28% for *E. coli rns* (36) to at least 67% for *Streptomyces araL* (43, 265). Most of the genes encoding members of this family are in the genomes of the gamma subdivision of the proteobacteria (purple bacteria) (194). A few have been found in low G+C and high G+C gram-positive bacteria and in cyanobacteria, but none have been found in archaebacteria or eukaryotes (194). However, because many prokaryotic genera have not been subjected to extensive genetic characterization, the observed distribution of AraC/XylS proteins may be nonrepresentative. The large genetic distances between prokaryotes with AraC/XylS regulators and the vast differences in G+C content suggest that a progenitor arose early in prokaryotic evolution. Because the conserved sequences within the members of the AraC/XylS are a series of well-established domains involved in DNA binding and stimulation of transcription, this family probably evolved through the recruitment of new domains of key importance in determining which function the regulator carries out. A phylogenetic tree in which no relationship between the branches and the function regulated by each subgroup is evident can be obtained upon request from M. T. Gallegos.

DOMAIN ORGANIZATION OF AraC/XylS POLYPEPTIDES

Size and Location of the Conserved Domain in AraC/XylS Members

Most members of the AraC/XylS family of regulators are 250 to 300 residues long, although a few exceptions are found:

HprB from *Burkholderia solanacearum*, Ada from *M. tuberculosis*, Ybbb from *B. subtilis*, and Hrp from *X. campestris* and *X. oryzae* are about 500 amino acids long (Table 1). A few proteins and hypothetical polypeptides were found to be particularly short (106 to 166 residues), e.g., AarP from *Providencia stuartii*, MarA and SoxS from *E. coli*, PqrA from *Proteus vulgaris*, RamA from *Klebsiella pneumoniae*, TetD from Tn*10*, and YcgK from *Alteromonas carragenovora* (Table 1).

The region of greatest amino acid sequence homology identified in XylS/AraC members is clearly a set of 99 residues found in most of the proteins at the C-terminal end of the regulators, although in some cases it is at the N-terminal end (CafR and Rob from *E. coli*) or in the central domain (Ada from *E. coli* and *S. typhimurium* and Ybbb from *B. subtilis*).

Conserved Domain

The alignment of the 99 amino acids that are highly conserved in the proteins of the AraC/XylS family of regulators is shown in Fig. 1. By using Matrix Blosum45 (112), a histogram showing the degree of similarity at each position was obtained (Fig. 2). With a cutoff point of 0.5 for similarity, 17 residues showed a high degree of conservation and represent the consensus for the family (A-----S---L---F----G----------R---A---L--------(I/V)--(I/V)----G(F/Y)-----F---F(R/K)---G--P, where - is any amino acid).

This sequence was conserved in at least 60% of the aligned proteins. The sequence is similar but not identical to that previously proposed by Gallegos et al. (84) based on the alignment of 27 proteins. From a statistical point of view, the present sequence is more accurate, because it includes 109 proteins and extends for 75 amino acids within the stretch of 99 residues. Analyses of the structures and sequences of proteins have established that sequence homology greater than 25% between two proteins extending for 50 amino acids is sufficient to ensure their identical tertiary structure (221). Given that members of the AraC/XylS family are transcriptional regulators and that the region of similarity extends for a region of nearly 100 amino acids with an overall similarity greater than 20%, these proteins can be assumed to possess identical tertiary structures in the conserved region. However, no tertiary structure for this domain is available, mainly because of the low solubility of the proteins of this family (64, 226).

Secondary-structure predictions were made with the entire alignment of the 99-amino-acid homologous segment by using the algorithm of Rost and Sander (217). This analysis suggested the existence of two potential α-helix–turn–α-helix (HTH) DNA binding motifs (23, 195–197, 257). In the XylS regulator, the first HTH motif is located at positions 228 to 251 and the second HTH motif is located at positions 281 to 305; these correspond to positions 195 to 218 and 245 to 270, respectively, in the AraC regulator (26–28, 169).

Evidence that the first HTH motif constitutes the DNA binding motif in AraC is based on the following findings. (i) Interference binding assays suggested that residues in the second α-helix of the motif made specific contacts with target DNA sequences at the P_{araBAD} promoter (27, 28). (ii) Mutations within residues in this region in AraC (Cys204 → Tyr, Ser208 → Ala, Arg210 → Cys, and His212 → Tyr or Ala) reduced binding to and decreased transcriptional activation from the P_{araBAD} promoter (26–28, 39, 78). The presence of two mutations in the XylS protein supports a role for these helices in promoter recognition: Ser229 → Ile (the first amino acid of the first α-helix) and substitution of Cys for Phe248 (in the second α-helix) resulted in mutant regulators with increased affinity for target sequences and the ability to mediate transcription from the cognate Pm promoter constitutively (81, 83, 162, 267).

The second HTH motif has been proposed for all proteins in the family. This motif contains an extra amino acid in the turn with respect to canonical HTH DNA binding motifs. Its biochemical role is unknown. Mutations within these helices are available for some members of the family: the substitution of Ala and Asn for Ser271 and Arg272, respectively, has been achieved in MelR, and Val has been substituted for Asp288 in XylS. These mutants behaved similarly to the wild-type regulator (41, 208). In the case of AraC, the picture arising from the analysis of mutants with mutations in these helices (Gly249 → Asp, Arg250 → His, Gly253 → Ser, Asp256 → Ala, Gln257 → Ala, Ser261 → Ala, and Val264 → Ile) is more complex, since certain mutants lost contact with multiple bases or bound to DNA in a pattern not fully consistent with a canonical HTH DNA binding motif (28, 39, 78).

It was recently suggested that AraC might contact target DNA sequences through the two HTH motifs (186). Although this might be the case for AraC (see below), it may not be a general rule for members of the family. This statement is based on comparisons of each of the HTH motifs of each member in the family with the corresponding aligned HTH motif in the rest of the family. Our results showed that sequence conservation at the HTH comprising the first HTH motif was low and that with certain pairs of sequences it was highly divergent. In contrast, no such variation was found when the second HTH motifs were compared (Fig. 2). We suggest that the variation in the first HTH motif represents the recognition of different target sequences at the cognate promoters by different regulators; conservation at the second HTH motif may thus represent a common function for all members of the family, e.g., contact with the transcriptional machinery. However, this hypothesis needs to be tested in vitro.

A small region of high sequence conservation was found outside the second HTH motif and toward the C-terminal end. Its most characteristic feature was the presence of a proline in more than 90% of the proteins in the family.

Given that AarP, MarA, PqrA, RamA, SoxS, and TetD, the shortest members of the XylS family (106 to 166 amino acids long), consist mainly of the homologous segment, the stretch of conserved residues most probably contains all the domains necessary for these regulators to interact with target DNA sequences and RNA polymerase and thus activate transcription from target promoters. Furthermore, for regulators whose recognition site has been defined, the target sequences in the cognate promoters have been located adjacent to or overlapping the −35 region of the promoter, as is the case in other positively regulated promoters (33, 49, 122). This suggests that the mechanism of transcription activation by AraC/XylS family members may involve direct interactions with RNA polymerase.

Nonconserved Domain

Data available for the nonconserved domain are scarce and basically limited to the AraC protein of the family *Enterobacteriaceae*; much less is known about the XylS and the other proteins. The nonhomologous N-terminal and central regions of the regulators recognizing chemical signals are presumed to contain binding sites for activator molecules that confer specificity (41, 172, 208, 232). Whether this information also holds for other members of the family is unknown.

The AraC protein, which regulates the L-arabinose operons in *E. coli*, consists of two domains that function in chimeric proteins. One provides the ability to form dimers (residues 1 to 170) and binds the ligand arabinose, and the other provides

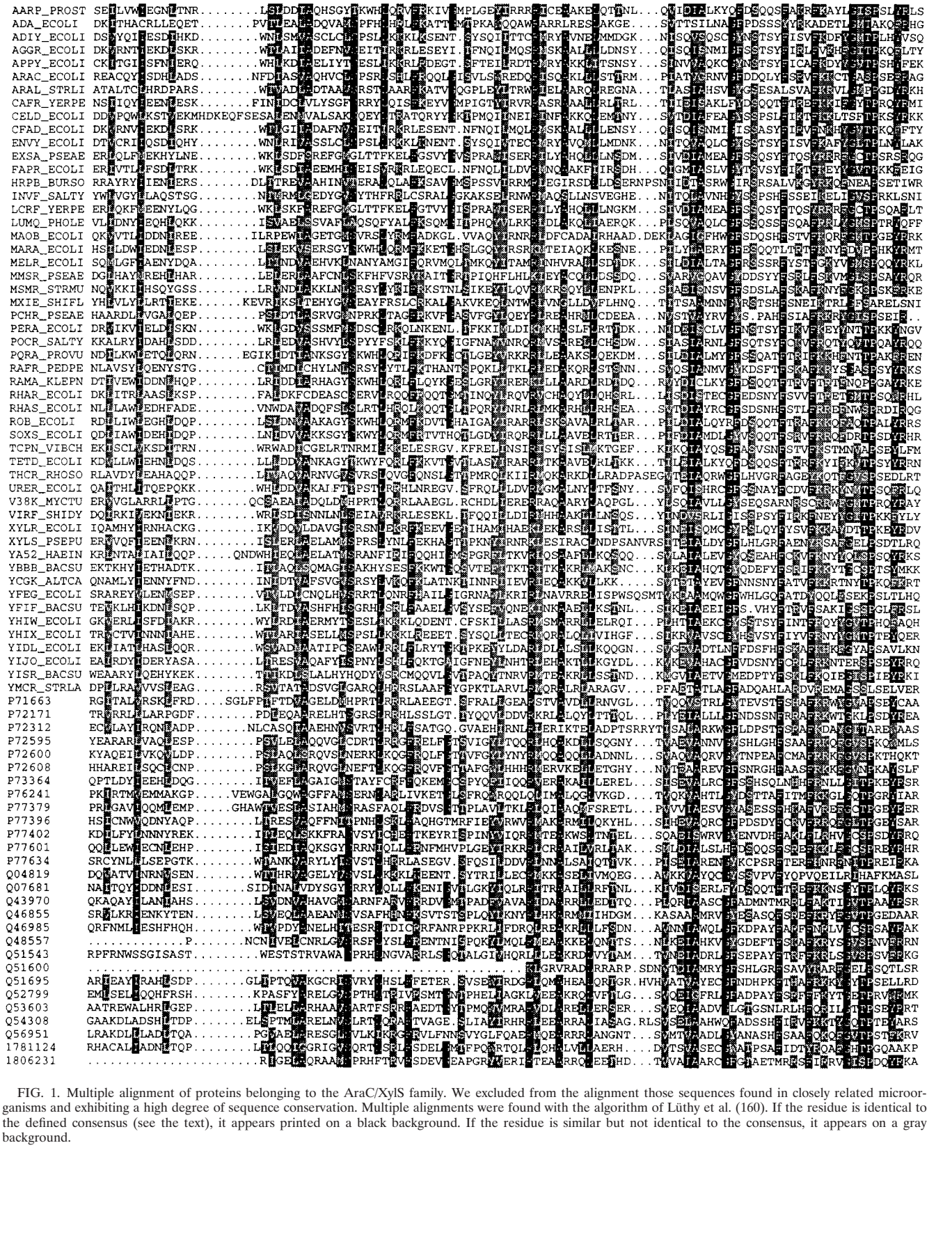

FIG. 1. Multiple alignment of proteins belonging to the AraC/XylS family. We excluded from the alignment those sequences found in closely related microorganisms and exhibiting a high degree of sequence conservation. Multiple alignments were found with the algorithm of Lüthy et al. (160). If the residue is identical to the defined consensus (see the text), it appears printed on a black background. If the residue is similar but not identical to the consensus, it appears on a gray background.

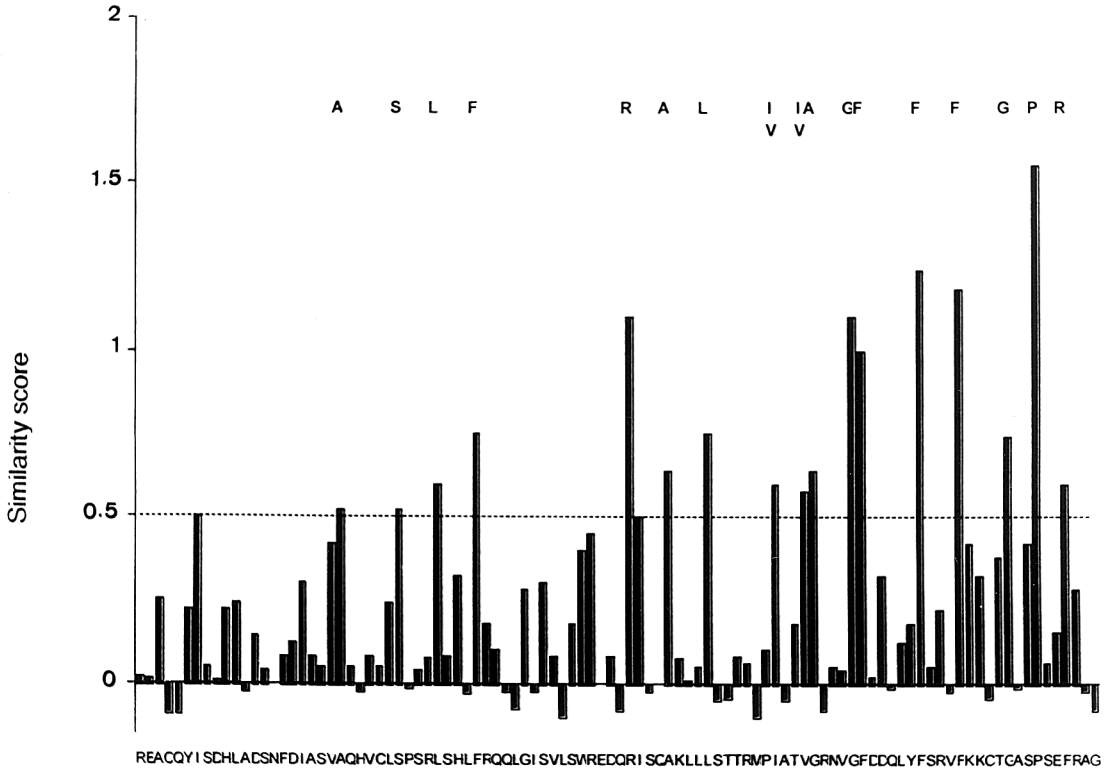

FIG. 2. Sequence conservation similarity within the conserved domain of the AraC/XylS family. The similarity score was calculated by using Matrix Blosum45 (112). A similarity score greater than 0.5 was chosen to establish the cutoff point to derive the consensus sequence of the family, which is shown from the portion above the corresponding bar in the histogram. The chosen amino acid was present at the given position in at least 60% of the aligned sequences shown in Fig. 1. The corresponding AraC sequence within this conserved domain is shown at the bottom of the figure, where the first amino acid residue corresponds to Arg 180 and the last residue corresponds to Gly 279.

site-specific DNA binding capability and activates transcription (residues 178 to 286) (35, 143, 168). These domains are connected by a flexible linker (69). In vivo and in vitro experiments showed that a chimeric protein consisting of the N-terminal half of the AraC protein and the DNA binding domain of the LexA repressor dimerizes, binds to a LexA operator, and represses the expression of a LexA operator in an arabinose-responsive manner (35). This suggests that at least in the case of AraC, the ligand domain and the DNA binding domain are independent (35).

Conclusive evidence for effector binding and dimerization of AraC in this nonconserved domain is provided by the crystal structure of this domain in the presence of arabinose (232). This domain contains an eight-stranded antiparallel β-barrel with "jelly-roll" topology, followed by two turns of 3_{10} helix, followed by a ninth β-strand that form part of one sheet of the β-barrel. The last β-strand is followed by two α-helices packed against the outer surface of the barrel. Each monomer of AraC binds one molecule of α-L-arabinose. The sugar stacks against the indole ring of Trp-95 and is stabilized by direct hydrogen bonds with the side chains of Pro-8, Thr-24, Arg-38, Tyr-82, and Hys-92, as well as hydrogen bonds with water molecules in the binding pocket. The sugar binding site is completed by the N-terminal arm of the protein (residues 7 to 18), which loops around to close off the end of the β-barrel in which arabinose is bound.

AraC is a dimer in both the presence and the absence of arabinose (255). Crystallographic data for the N-terminal domain of AraC showed that the two monomers are associated by an antiparallel coiled coil formed between the terminal α-helix of each monomer, with each end of the coiled coil anchored by a triad of leucine residues that pack together in a knobs-into-holes manner.

Schleif's group investigated whether any of the amino acids in the linker region between the nonconserved and the conserved domains play active, specific, and crucial structural roles or whether these amino acids merely serve as passive spacers between the functional domains. They found that all but one of the linker amino acids could be substituted by other amino acids individually and in small groups with no substantial effect on the ability of AraC protein to activate transcription when arabinose is present. However, when the entire linker region is replaced with linker sequences from other proteins, the functioning of AraC is impaired (69, 70).

MECHANISMS OF ACTION OF INDIVIDUAL FAMILY MEMBERS

The XylS Regulator Controls Expression from the Pm Promoter

The growth of *P. putida* (pWW0) on alkylbenzoates requires expression of the *meta* pathway operon, mediated by the XylS protein (79, 83, 118, 208). The *xylS* gene is expressed at low constitutive levels from a σ^{70}-dependent promoter called Ps2; on the addition of a *meta*-cleavage pathway substrate, expression from the Pm promoter occurs immediately, suggesting that the regulator becomes active after effector binding (82). The XylS protein is 321 amino acid residues long (119, 169, 235). The first two-thirds of the protein sequence, i.e., the

amino-terminal and central regions of the protein, seem to be involved in interactions with effectors (172, 208). Interactions between effector molecules and the regulator have been studied by analyzing XylS-dependent transcriptional activation from the Pm promoter in the presence of different benzoate analogs. These studies revealed that substituted benzoates are XylS effectors, although not all positions in the planar benzoate molecule are equivalent. For example, position 3 is highly permissible (-CH_3, -C_2H_5, and -OCH_3 groups and F, Cl, Br, and I atoms are permissible substituents), whereas positions 2 and 4 pose some restrictions to substituents (-CH_3, -F, and -Cl groups are allowed, whereas -C_2H_5 and -I are not) (210, 211). Although disubstitutions involving positions 2 and 3 and positions 3 and 4 are permissible, other combinations are usually nonpermissible, which suggests that interactions between the effector and the regulator are nonsymmetrical. Ramos et al. (208) and Michán et al. (172) isolated and sequenced a series of mutant regulators able to recognize substituted benzoate effectors that are not recognized by the wild-type regulator. Critical mutations were found to be clustered at positions 37 to 45. Arg-41 seems to be a critical residue for interaction(s) with effectors, since changes at this position result in multiple different phenotypes. For example, XylSArg41Gly is a mutant regulator that has lost the ability to recognize *o*- and *p*-methylbenzoate, although it remained activatable by *m*-methylbenzoate. Substitution of Arg41 with Leu resulted in a mutant unable to respond to benzoate effectors (172).

XylS mutants such as XylSArg41Cys, XylSPro37Gly, XylSSer229Ile, XylSAsp274Val, and XylSAsp274Glu mediated transcription from Pm in the absence of effectors (172, 267). These results support the hypothesis that XylS exists in vivo in a dynamic equilibrium between an inactive and an active form with respect to transcriptional stimulation. Therefore, transition from the inactive to the active form may be mediated by effector binding. How the interaction between benzoates and XylS leads to an active regulator is not yet understood, but regardless of the mechanism, the effector binding pocket and the DNA binding motif are not independent domains, as shown by intramolecular dominance of C-terminal mutations over N-terminal ones and by the reversal of this dominance in double mutants constructed in vitro (171).

Overproduction of XylS via a natural cascade regulatory system—involving expression from tandem Ps1 and Ps2 promoters (82; see reference 206 for a review)—or after expression from strong promoters (120, 169, 207, 234) leads to stimulation of transcription from the Pm promoter in the absence of effectors. This finding further supports the idea that XylS may exist in an equilibrium between an inactive and an active form, so that if the total amount of XylS protein is increased in the cell, some of the XylS molecules become active from a transcriptional point of view (164).

Stimulation of transcription from the Pm promoter requires a DNA sequence extending to about 80 bp upstream of the transcription initiation point (83, 132, 207). In the architecture of the Pm promoter, two regions can be distinguished on the basis of genetic data: the XylS interaction region, which extends from about bp −46 to −80, and the region between −41 and +1 for RNA polymerase recognition, which exhibits atypical −35 and −10 DNA sequences. XylS-dependent transcription from Pm can be mediated by RNA polymerase with either σ^{70} or σ^{S} (163).

Gallegos et al. (83) and González (92) have studied in detail the organization of XylS binding sites in the Pm promoter. They generated a series of 5′ sequential deletions and a large series of point mutations in the promoter and analyzed transcription from the resulting mutant promoters mediated by the wild-type XylS protein and by mutant XylS regulators that were constitutive. It was found that Pm promoter variants deleted up to −60 could be activated by constitutive XylS mutants (but not by the wild-type regulator) and that extension of the deletion to −51 prevented transcription. On the basis of sequence analyses, it was proposed that the XylS binding site was probably represented by the motif T(C/A)CAN_4TGCA, which appears twice, such that the exact location of the RNA polymerase binding site proximal motif was between −46 and −57 and the distal motif was between −67 and −78 (82). The −46 to −57 proximal site constitutes the minimum sequence required for transcription stimulation. Point mutations suggest that the TGCA submotif may be the primary recognition site, with the remaining sequences contributing to overall affinity (92, 132).

Kaldalu et al. (125) reported the immunopurification of a functionally active XylS protein bearing a hemagglutinin epitope fused at its N terminus (N-XylS). This N-XylS variant was able to specifically bind and retain a DNA fragment bearing the proposed XylS binding region in Pm. A set of footprinting experiments indicated that N-XylS binds along one side of the DNA, covering four helix turns (from −28 to −72) and making base-specific contacts in four adjacent major groove regions on the same helix face. This footprinting extended beyond the site defined by genetic means; as in other members of the family (28, 37, 59, 71, 72, 110, 152, 158, 250), this may reflect oligomerization of N-XylS after recognition of a primary binding site. Further in vitro studies with purified RNA polymerase and XylS are needed to determine whether the binding sites for each protein overlap. The observation that overproduction of the regulator is sufficient to activate Pm in vivo in the absence of effector (120, 169, 207, 234) supports the hypothesis that effectors increase the cellular concentration of XylS in its active conformation ($XylS_a$ may exhibit higher affinity for its target DNA sequence) at the DNA target site.

Arabinose Metabolism in *E. coli*

Four transcriptional units in *E. coli* are involved in the utilization of L-arabinose: *araBAD*, which encodes three enzymes responsible for L-arabinose catabolism (65); *araE* and *araFGH*, which encode proteins responsible for low-affinity and high-affinity transport of L-arabinose (25, 239); and the regulatory gene *araC*, which encodes a protein that controls the expression of these genes as well as autoregulating its own synthesis (38, 66, 67, 95, 256).

The AraC protein is predominantly a dimer in solution (35, 168, 255). In the absence of arabinose, AraC protein represses expression of the *araBAD* and *araC* promoters (called P_{araBAD} and P_{araC}, respectively) (62, 99, 114, 117, 144–146, 156, 157, 165). With arabinose, AraC activates transcription from the promoters of the catabolic operons (Fig. 3). The response of the wild-type *ara* operons to arabinose was found to occur within 3 s of inducer addition (114), and mRNA was detected within 15 to 30 s (114, 123). Expression from all four of these promoters is also regulated by the cyclic AMP-catabolite activator protein (99). AraC protein interactions with the *ara* promoters were determined by chemical interference assays and by mutagenesis of the protein and the promoters (26–28, 37, 40, 101, 109, 144, 156, 168). A consensus sequence for AraC binding was obtained by comparing the sites from *E. coli* and *S. typhimurium ara* promoters (28, 108–110, 158). An $AGCN_7TCCATA$ sequence is conserved in all sites and appears as a tandem repeat (Fig. 4). The $araO_2$ site, which is needed for inhibition of transcription at P_{araBAD} (see below), is apparently only half of a site (Fig. 4).

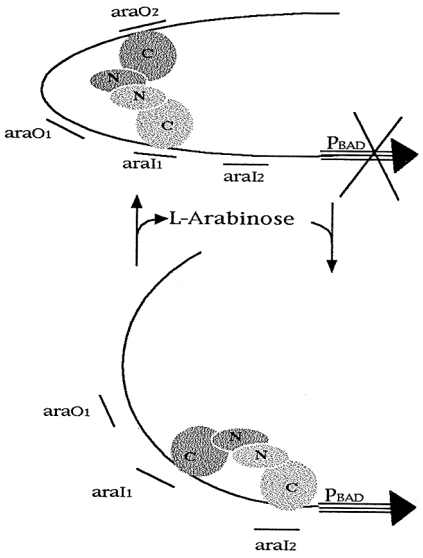

FIG. 3. Model of AraC induction by L-arabinose at the *araBAD* promoter and relevant sites at the divergent *araC* promoter. Details of functioning are explained in the text.

Regulation of the *araBAD* (P_{araBAD}) and *araC* (P_{araC}) promoters by the AraC protein has been extensively characterized (37, 156). In the absence of arabinose, one monomer of the AraC dimer occupies the $araI_1$ site while the other occupies a half-site approximately 200 bp away, known as $araO_2$ (Fig. 3). The dimer bound to target sequences in this way generates a DNA loop, which prevents transcription from P_{araBAD} and P_{araC} (37, 156, 157). When arabinose is added, the AraC protein undergoes a conformational change and shifts to occupy the adjacent half-sites, $araI_1$ and $araI_2$ (156, 157). As a result, P_{araBAD} is induced. Therefore, the main consequence of arabinose binding on AraC protein is to change the affinity of AraC for different spatial arrangements of half-sites. In the absence of arabinose, AraC favors binding to half-sites separated by more than one helical turn of DNA (Fig. 3), whereas in the presence of arabinose, AraC favors binding to half-sites separated by less than one helical turn of DNA (37). Therefore, arabinose destabilizes AraC protein binding to the I_1-O_2 looped complex but stabilizes binding to the I_1-I_2 site. Furthermore, because the loop is disorganized, free access of RNA polymerase to the P_{araC} promoter is transitorily facilitated and transcription increases. Subsequently, P_{araC} shuts down as a result of AraC protein binding to the $araO_1$ site, which blocks the access of RNA polymerase to the P_{araC} promoter (225).

It was shown that to activate transcription in P_{araBAD}, the AraC protein binding site must overlap the −35 region of the promoter by 4 bp (212). AraC protein was located on one side of four adjacent helix-turn regions of the DNA, and there is evidence that each AraC monomer requires two direct repeats in successive turns of the DNA helix for binding (37, 110, 158). In light of the strict spacing and orientation requirements for AraC activation, interactions between AraC and RNA polymerase are likely to be specific and inflexible. Providing further support for this theory is the almost identical arrangement of the protein binding sites for *araBAD* and *araE*. Surprisingly, the *araFGH* promoter (P_{araFGH}) possesses a radically different structure. In P_{araFGH}, the catabolite activator protein site, rather than the AraC site, overlaps the −35 recognition sequence of RNA polymerase. In addition, the AraC sites in *araFGH* are arranged in the opposite direct-repeat orientation (108).

Niland et al. (186) systematically substituted every base pair in a synthetic 17-bp $araI_1$ target (5-TAGCATTTTTATCCATA-3′ [the underlined bases correspond to those conserved in the consensus]) with each of the three possible alternatives and then used qualitative gel shift analysis to test the binding of AraC to these 51 DNA targets in the presence of L-arabinose. They found that every substitution of the underlined bases reduced AraC binding to 1/10 or less whereas substitutions at other bases had little or no effect. In the absence of L-arabinose, the binding of AraC to $araI_1$ was reduced to one-sixth or less.

Two possible HTH motifs were proposed in the C-terminal domain of AraC (78), but contact to DNA was demonstrated only for the first (27, 28). This first motif binds the first major groove of the DNA. These results were confirmed by Niland et al. (186). The finding of Niland et al. (186) with AraC mutant Asp256 → Ala (in the second helix of the second HTH motif) provided evidence that the second HTH contacts the second major groove.

SoxS Regulator and Sox-Box

Redox cycling compounds such as paraquat and menadione are a continuous source of superoxide in the cell as a consequence of repeated cycles of oxidation and reduction. Exposure of *E. coli* cells to these compounds induces the synthesis of about 40 proteins (93). A subset of these proteins are produced by a regulon controlled by two genes, *soxR* and *soxS*, which constitute the so-called *soxRS* regulon (4, 190, 260). The following genes are known to be members of this regulon: *acnA* (aconitase), *fpr* (NADPH:ferredoxin oxidoreductase), *fumC* (fumarase C), *inaA* (function unknown), *micF* (antisense regulator of *ompF*), *nfo* (DNA repair enzyme endonuclease IV), *pqi-5* (function unknown), *ribA* (GTP cyclohydrolase II), *sodA* (manganese superoxide dismutase), *soi-17* (function unknown), *soi-28* (function unknown), and *zwf* (glucose-6-phosphate dehydrogenase) (44, 94–96, 134–136, 154, 155, 173, 216, 246). Both SoxR (17 kDa) and SoxS (13 kDa) are DNA binding proteins. Induction of the *soxRS* regulon occurs in two steps. An intracellular signal of oxidative stress (reduction in the cellular NADPH/$NADP^+$ ratio or exposure to superoxide) converts preexisting SoxR protein into a transcriptional activator of the *soxS* gene. The overproduced SoxS protein in turn activates the transcription of target genes of the regulon (4, 87, 190, 191, 260).

In vitro studies have demonstrated that purified SoxS and

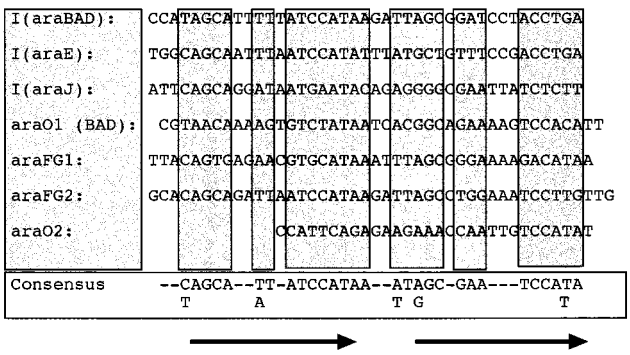

FIG. 4. Multiple alignment of AraC binding sites. A residue was chosen for the consensus sequence when it appeared in more than half of the sequences. The lines underneath show the direct repeats.

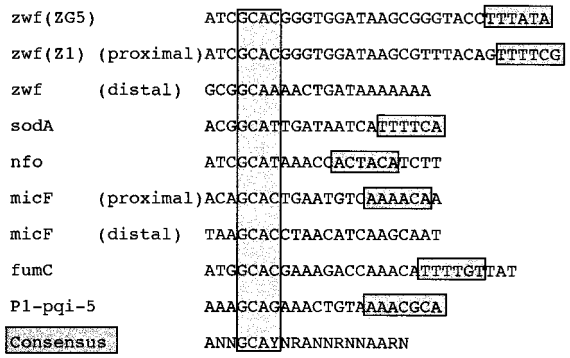

FIG. 5. Multiple alignment of primary SoxS binding sites in different promoters. The −35 region of each promoter is shown in a horizontal box. The vertical box shows the highly conserved GCAPy motif in all promoters (further details are given in references 71, 72, 151, and 152).

MalE-SoxS fusion proteins activate transcription from the promoter of target genes and can specifically bind and form multiple DNA-protein complexes thanks to the presence of multiple binding sites at cognate promoters (71, 72, 151, 152). DNase I footprinting assays have shown that promoters whose transcription is activated by SoxS seem to fall into two classes with respect to the location of the proximal site relative to the −35 hexamers of the promoters (Fig. 5). In one class, the primary protected region completely covers the −35 hexamers of the *micF*, *nfo*, P1-*pqiA*, and *sodA* promoters, whereas it is adjacent to or only partially overlaps the −35 hexamers of the *fumC* and *zwf* promoters (Fig. 5). *ribA* seems to be an exception, since the putative SoxS binding site is located from −146 to −118, far upstream from the −35 element (134). The SoxS distal sites at the *micF* and *zwf* promoters (Fig. 5) have been characterized by a combination of DNase I footprinting and methylation interference assays (71, 72, 151, 152). The alignment of the protected regions (Fig. 5) revealed a "Sox-box" consensus whose sequence is ANNGCAPyNPuANNPuNN AAPu, where N is any base, Py is a pyrimidine, and Pu is a purine (72).

A potentially important feature of the 19-bp consensus sequence is the GCAPy motif that lies near the 5′ end. This short sequence is conserved among the proximal and distal sites of *fumC*, *micF*, *nfo*, P1-*pqiA*, *sodA*, and *zwf*. Therefore, the GCAPy motif may be a primary recognition element for SoxS, with the remaining positions of the Sox-box sequence contributing to the overall affinity. The dissociation constant for chimeric MalE-SoxS binding to DNA sequences that contain this element is about 10^{-8} to 10^{-9} M. This relatively weak interaction suggests that additional free energy for binding might come from cooperative interactions with either RNA polymerase or a second SoxS molecule.

The importance of the GCAPy motif is also substantiated by the properties of several *sodA* mutants. Naik and Hassan (179) and Compan and Touati (50) described *sodA* mutants that do not respond to superoxide stress. In one mutant, the 5′-GCAT-3′ sequence, which lies within the proximal protected site of *sodA*, was changed to 5′-TACG-3′; in another mutant, the sequence was deleted. Presumably the uninducible phenotype of these mutants was derived from the destruction of this GCAT sequence. Furthermore, single base pair substitutions at any position in the GCAY motif greatly reduced SoxS binding to synthetic oligonucleotides bearing the *micF*-proximal site (152).

VirF Regulator in *Yersinia*

Pathogenic bacteria of the genus *Yersinia* (*Y. enterocolitica*, *Y. pseudotuberculosis*, and *Y. pestis*) cause diseases in rodents and humans, with symptoms ranging from enteritis to septicemia and death. All three species carry closely related plasmids of about 70 kb, generally called pYV (for *Yersinia* virulence), which are essential for virulence. At 37°C and in medium without calcium, the pYV plasmids direct the secretion of at least 10 proteins called Yops (YopB, YopD, YopE, YopH, YopM, YopN, YopO, YopP, and YopQ) and LcrV, the protective antigen (51, 75).

The synthesis of plasmid-encoded virulence proteins in yersinae is positively controlled at the transcriptional level by the *virF* gene, the key activator of the systems. VirF forms dimers in solution and stimulates transcription from the *yopE*, *yopH*, *lcrGVH-yopBD*, and *virC* operon promoters (53, 115, 141, 249).

At low temperatures, i.e., 25°C, transcription of these genes is under negative control by YmoA, a histone-like global regulator (52, 215). Transcription of the *virC* operon and *yop* genes is also repressed by Ca^{2+} (54, 76, 203).

VirF binds to multiple sites in the promoter region of *yopE*, *yopH*, *virC*, and *lcrG* (141, 249, 250). DNase I and hydroxyl radical footprinting identified the corresponding binding sites, and a 13-bp TTTaGYcTtTat (capital letters indicate bases conserved in more 60% of the sequences) was inferred. VirF bound tightly to this sequence when it appeared as an inverted repeat separated by a single base pair and weakly when the sequence appeared alone (251). The strong sites were occupied before the weak sites. The position of the binding sites with respect to the −35 region varied depending on the promoter.

Activator Sequences in Some Promoters Controlled by Members of the AraC/XylS Family

Apart from the detailed footprinting analysis of the P_{araBAD} and P_{rha} promoters and the *soxS*-regulated promoters and a thorough analysis of the Pm region, little evidence for other promoters is available. Without attempting an exhaustive review, we summarize below some findings in other promoters for which data are available.

The P_{ureD} promoter is activated by UreR. The chromosomal *Proteus mirabilis*, the plasmid-encoded *E. coli*, and other urease loci in the family *Enterobacteriaceae* comprise seven contiguous structural and accessory genes (*ureDABCEFG*) and the divergently transcribed *ureR* gene, which encodes the transcriptional regulator (58, 185). Physical mapping identified the region between −61 and −86 with respect to the transcription initiation point from P_{ureD} as sine qua non for the transcription from this promoter and also found that sequences up to −135 increased transcription from P_{ureD}. Gel shift assays with the DNA fragment up to the −135 point revealed multiple binding of UreR to this promoter. This suggested that UreR binds as a multimer or exhibits multiple binding sites (59). The exact position of UreR binding is unknown, but our inspection of the sequence revealed a direct TATTTT repeat in the −61 to −86 region, which was also found (slightly distorted) upstream from −86. Whether these repeats constitute the actual sites recognized by UreR is unknown.

The YbtA protein controls its own synthesis and expression from *psn* and *irp2*. The pesticin receptor (Psn) of *Y. pestis* confers sensitivity to bacteriocin and pesticin and is an integral component of an inorganic iron transport system that functions at 37°C. YbtA controls the synthesis of Psn and proteins encoded by the *irp2* operon and also controls its own synthesis. Sequence alignment of the promoters controlled by YbtA re-

vealed a consensus sequence, aACCCgWWWcgGG (where W is A or T), which appears twice in each promoter. No clear symmetry was found in this sequence, although the nature of the highly conserved residues suggested that the binding sites are recognized as two direct repeats (73).

RhaS is one of the regulators of rhamnose metabolism in *E. coli*. RhaS activates transcription from *rhaBAD*, and transcription from the *rhaS* gene is controlled by RhaR (63, 244). Both RhaR and RhaS bind rhamnose to stimulate transcription from the corresponding cognate promoters. Full transcription from the *rhaBAD* promoter requires CRP (catabolite repression protein). Deletion analysis at the promoter of the *rhaBAD* operon revealed the requirement for a stretch of about 80 bp upstream from the main transcription initiation point. The CRP binding site was located adjacent to this sequence and was centered at bp −92.5 (63, 243, 244). By DNase footprinting and mutational analysis, it has been shown that RhaS binds in P_{rhaBAD} to an inverted repeat of two 17-bp half-sites separated by 16 bp. These findings were made possible by the discovery that the normally insoluble RhaS protein could be renatured in active form by the slow removal of urea while in the presence of DNA. This technique will probably prove useful in the study of the AraC/XylS family members (243).

The MelR regulator controls transcription from the *melAB* operon promoter. The *melAB* operon encodes proteins essential for melibiose metabolism in *E. coli*. Transcription initiation from P_{melAB} is stimulated by the MelR regulator with melibiose. Scrutiny of the nucleotide sequence at this promoter revealed a relatively well-conserved σ^{70} −10 box and an unconserved −35 region. Upstream in the P_{melAB} promoter, two identical 18-bp elements are organized as an inverted repeat from positions −109 to −92 and from positions −54 to −71; these are the MelR binding sites (41, 252).

CONCLUDING REMARKS AND FUTURE PERSPECTIVES

Profile PROSITE PS01124 defines for AraC/XylS family members a matrix that has been established on the basis of successive searches in protein and nucleic acid databases. More than 100 proteins and polypeptides deduced from ORFs have been included in the family. The matrix assigns to these proteins and polypeptides a normalized score equal to or higher than 12.52. Once a more exhaustive analysis of prokaryote genomes is carried out, we predict that the number of proteins assigned to the AraC/XylS family on the basis of PROSITE profile PS01124 will increase significantly.

Multiple alignments of the proteins in this family revealed a stretch of 99 conserved amino acids. This conserved domain comprises all functions required for DNA binding and RNA polymerase contact and stimulation of transcription. Secondary-structure analysis predicts the presence of two potential HTH structures. The set formed by the first HTH motif seems to be the actual DNA binding domain of the members of the family; however, the possibility that the other HTH domain also functions as a DNA binding domain cannot be ruled out. Exhaustive analyses of mutations in different members of the family are needed to further define the role of these HTH. Furthermore, efforts to crystallize this stretch are needed to determine the actual tertiary structure of the members of this family.

One of the striking features of the AraC/XylS family is the paucity of biochemical data. This reflects the difficulty of handling these proteins. Most of them are highly insoluble and are thus difficult to purify. Because several members of the family possess this property and because the dimerization domain of AraC is soluble (232), it seems that it is the DNA binding domain which makes these proteins poorly soluble (64). Efforts to improve the solubility of this domain are essential to facilitate purification and crystallization.

The conserved domain is usually connected to a nonconserved domain via a linker. The nonconserved domain is critical for signal recognition in members of the family activated by effector binding. However, it is not known how the linker transfers a signal from the signal reception site to the DNA binding site or how the active regulator interacts with RNA polymerase to drive transcription from cognate promoters. The role of the nonconserved domain in proteins involved in pathogenesis is an area that deserves particular attention, since practically no data are available.

No general conclusions can be drawn regarding the promoters regulated by members of the family. However, it has been found that these promoters usually contain more than one binding site for the regulator. Many sites for which the regulator has high and low affinity have been identified. The site proximal to the RNA polymerase binding site has been found in most cases to overlap or abut the −35 region of the promoter, but cases exist in which sites are located at about 100 bp from the main transcription initiation point. Whether this reflects the possibility that different members of the family contact RNA polymerase in different ways is unknown (11, 24, 142, 163, 188, 247). Another feature is the organization of the binding sites. It has been suggested that for some promoters the binding sites are organized as inverted repeats whereas for others they are organized as direct repeats. However, few symmetry studies are available, and this deserves attention.

It should be noted that in spite of the high homology among AraC/XylS members, transcription stimulation mediated by these proteins from the corresponding promoters shows interesting diversity. In addition to the specific regulator, transcription from certain promoters regulated by members of this family requires other proteins for maximal activity (e.g., the CRP in the P_{rhaBAD} promoter), or histone-like proteins that act as negative regulators (e.g., YomA in the VirF-regulated P_{yop} promoters) (52, 124, 242). This is clear evidence that the expression of genes controlled by members of this family is integrated in overall cellular control.

ACKNOWLEDGMENTS

Work in Granada was supported by a grant from the CICYT (AMB 1038-C02-1) and a grant from the Biotechnology Program of the EC (BIO4-CT97-2270). Work in Baltimore was supported by a grant from NIH (GM18277).

We thank Karen Shashok for linguistic improvement of the manuscript, Maxi Manzanera for assistance with artwork, and Javier González for help with the computers. We are indebted to M. M. Fandila for preparing the manuscript.

REFERENCES

1. **Aiba, H., T. Baba, K. Fujita, K. Hayashi, A. Honjo, T. Horiuchi, K. Ikemote, T. Inada, K. Isono, S. Isono, T. Itoh, K. Kanai, H. Kasai, K. Kashimoto, S. Kim, S. Kimura, M. Kitagawa, M. Kitakawa, K. Makino, S. Masuda, T. Miki, K. Mizobuchi, H. Mori, K. Motomura, Y. Nakamura, H. Nashimoto, Y. Nishio, T. Oshima, N. Saito, G. Sampei, Y. Seki, H. Tagami, K. Takemoto, C. Wada, Y. Yamamoto, and M. Yano.** 1996. SPTREMBL. Accession no. P77402.
2. **Aiba, H., T. Baba, K. Fujita, K. Hayashi, A. Honjo, T. Horiuchi, K. Ikemote, T. Inada, K. Isono, S. Isono, T. Itoh, K. Kanai, H. Kasai, K. Kashimoto, S. Kim, S. Kimura, M. Kitagawa, M. Kitakawa, K. Makino, S. Masuda, T. Miki, K. Mizobuchi, H. Mori, K. Motomura, Y. Nakamura, H. Nashimoto, Y. Nishio, T. Oshima, N. Saito, G. Sampei, Y. Seki, H. Tagami, K. Takemoto, C. Wada, Y. Yamamoto, and M. Yano.** 1997. SPTREMBL. Accession no. P77396.
3. **Allaoui, A., P. J. Sansonetti, and C. Parsot.** 1993. MxiD, an outer membrane protein necessary for the secretion of the *Shigella flexneri* Ipa invasins. Mol. Microbiol. **7:**59–68.

4. **Amábile-Cuevas, C. F., and B. Demple.** 1991. Molecular characterization of the *soxRS* genes of *Escherichia coli*: two genes control a superoxide stress regulon. Nucleic Acids Res. **19:**4479–4484.
5. **Anderson, W. F., D. H. Ohlendorf, Y. Takeda, and B. W. Matthews.** 1981. Structure of the *cro* repressor from bacteriophage lambda and its interaction with DNA. Nature **290:**754–758.
6. **Arakawa, E., J. I. Kato, K. I. Ito, and H. Watanabe.** 1995. SWISSPROT. Accession no. Q55292.
7. **Argos, P., M. Hanei, J. M. Wilson, and W. N. Kelley.** 1983. A possible nucleotide-binding domain in the tertiary fold of phosphoribosyltransferases. J. Biol. Chem. **258:**6450–6457.
8. **Ariza, R. R., S. P. Cohen, N. Bachhawat, S. B. Levy, and B. Demple.** 1994. Repressor mutations in the *marRAB* operon that activate oxidative stress genes and multiple antibiotic resistance in *Escherichia coli*. J. Bacteriol. **176:**143–148.
9. **Ariza, R. R., Z. Li, N. Ringstad, and B. Demple.** 1995. Activation of multiple antibiotic resistance and binding of stress-inducible promoters by *Escherichia coli rob* protein. J. Bacteriol. **177:**1655–1661.
10. **Assinder, S. J., P. de Marco, D. J. Osborne, C. L. Poh, L. E. Shaw, M. K. Winson, and P. A. Williams.** 1993. A comparison of multiple alleles of *xylS* carried by TOL plasmid pWW53 and pDK1 and its implications for their evolutionary relationship. J. Gen. Microbiol. **139:**557–568.
11. **Atlung, T., K. Knudsen, L. Heerfordt, and L. Brøndsted.** 1997. Effects of σ^S and the transcriptional activator AppY on induction of the *Escherichia coli* and *cbdAB-appA* operons in response to carbon and phosphate starvation. J. Bacteriol. **179:**2141–2146.
12. **Atlung, T., A. Nielsen, and F. G. Hansen.** 1989. Isolation, characterization and nucleotide sequence of *appY*, a regulatory gene for growth-phase-dependent gene expression in *Escherichia coli*. J. Bacteriol. **171:**1683–1691.
13. **August, P. R., M. C. Flickinger, and D. H. Sherman.** 1994. Cloning and analysis of a locus (*mcr*) involved in mitomycin C resistance in *Streptomyces lavendulae*. J. Bacteriol. **176:**4448–4454.
14. **Bairoch, A., and K. E. Rudd.** 1995. PROSITE. Accession no. PDOC00884.
15. **Bairoch, A., and K. E. Rudd.** 1995. PROSITE. Accession no. PDOC00874.
16. **Barbeyron, T., B. Henrissat, and B. Kloareg.** 1994. The gene encoding the kappacarragenase of *Alteromonas carragenovora* is related to β-1,3-1,4-glucanases. Gene **139:**105–109.
17. **Beck von Bodman, S., G. T. Hayman, and S. K. Farrand.** 1992. Opine catabolism and conjugal transfer of the nopaline Ti plasmid pTiC58 are coordinately regulated by a single repressor. Proc. Natl. Acad. Sci. USA **89:**643–647.
18. **Blattner, F. R., V. D. Burland, G. Plunkett III, H. J. Sofia, and D. L. Daniels.** 1993. Analysis of the *Escherichia coli* genome. IV. DNA sequence of the region from 89.2 to 92.8 minutes. Nucleic Acids Res. **21:**5408–5417.
19. **Blattner, F. R., G. Plunkett III, G. F. Mayhew, N. T. Perna, and F. D. Glasner.** 1997. SPTREMBL. Accession no. P76241, P77396, P77634, and Q46855.
20. **Bolin, I., and H. Wolf-Watz.** 1988. The plasmid-encoded Yop2b protein of *Yersinia pseudotuberculosis* is a virulence determinant regulated by calcium and temperature at the level of transcription. Mol. Microbiol. **2:**237–245.
21. **Branden, C., and J. Tooze.** 1991. Introduction to protein structure. Garland Publishing Co., New York, N.Y.
22. **Braus, G., M. Argast, and C. F. Beck.** 1984. Identification of additional genes on transposon Tn*10*: *tetC* and *tetD*. J. Bacteriol. **160:**504–509.
23. **Brennan, R. G., and B. W. Matthews.** 1989. The helix-turn-helix DNA binding motif. J. Biol. Chem. **264:**1903–1906.
24. **Brøndsted, L., and T. Atlung.** 1996. Effect of growth conditions on expression of the acid phosphatase (*cyx-appA*) operon and the *appY* gene, which encodes a transcriptional activator of *Escherichia coli*. J. Bacteriol. **178:** 1556–1562.
25. **Brown, C. E., and R. W. Hogg.** 1972. A second transport system of L-arabinose in *Escherichia coli* controlled by the *araC* gene. J. Bacteriol. **111:** 606–613.
26. **Brunelle, A., W. Hendrickson, and R. Scheif.** 1985. Altered DNA contacts made by a mutant AraC protein. Nucleic Acids Res. **13:**5019–5026.
27. **Brunelle, A., and R. F. Schleif.** 1987. Missing contact probing of DNA-protein interactions. Proc. Natl. Acad. Sci. USA **84:**673–676.
28. **Brunelle, A., and R. F. Schleif.** 1989. Determining residue-base interactions between AraC protein and *ara*I DNA. J. Mol. Biol. **209:**607–622.
29. **Buck, D., and J. R. Guest.** 1989. Overexpression and site-directed mutagenesis of the succinyl-CoA synthetase of *Escherichia coli* and nucleotide sequence of a gene (*g*30) that is adjacent to the *suc* operon. Biochem. J. **260:**737–747.
30. **Burke, K. A., and G. Wilcox.** 1987. The *araC* gene of *Citrobacter freundii*. Gene **61:**243–252.
31. **Burland, V., G. Plunkett III, D. L. Daniels, and F. R. Blattner.** 1993. DNA sequence and analysis of 136 kilobases of the *Escherichia coli* genome: organizational symmetry around the origin of replication. Genomics **16:** 551–561.
32. **Burland, V. D., G. Plunkett III, H. J. Sofia, D. L. Daniels, and F. R. Blattner.** 1995. Analysis of the *Escherichia coli* genome VI: DNA sequence of the region from 92.8 through 100 minutes. Nucleic Acids Res. **23:**2105–2119.
33. **Busby, S., and R. H. Ebright.** 1994. Promoter structure, promoter recognition and transcription activation in prokaryotes. Cell **79:**743–746.
34. **Bussey, L. B., and R. L. Switzer.** 1993. The *degA* gene product accelerates degradation of *Bacillus subtilis* phosphoribosylpyrophosphate amidotransferase in *Escherichia coli*. J. Bacteriol. **175:**6348–6353.
35. **Bustos, S. A., and R. F. Schleif.** 1993. Functional domains of the AraC protein. Proc. Natl. Acad. Sci. USA **90:**5638–5642.
36. **Caron, J., L. M. Coffield, and J. R. Scott.** 1989. A plasmid-encoded regulatory gene, *rns*, required for expression of the CS1 and CS2 adhesins of enterotoxigenic *Escherichia coli*. Proc. Natl. Acad. Sci. USA **86:**963–967.
37. **Carra, J. H., and R. F. Schleif.** 1993. Variation of half-site organization and DNA looping by AraC protein. EMBO J. **12:**35–44.
38. **Casadaban, M. J.** 1976. Regulation of the regulatory gene for the arabinose pathway, *araC*. J. Mol. Biol. **104:**557–566.
39. **Cass, L. G., and G. Wilcox.** 1986. Mutations in the *araC* regulatory gene of *Escherichia coli* B/r that affect repressor and activator functions of AraC protein. J. Bacteriol. **166:**892–900.
40. **Cass, L. G., and G. Wilcox.** 1988. Novel activation of *araC* expression and DNA site required for *araC* autoregulation in *Escherichia coli* B/r. J. Bacteriol. **170:**4174–4180.
41. **Caswell, R., J. Williams, A. Lyddiatt, and S. Busby.** 1992. Overexpression, purification and characterization of the *Escherichia coli* MelR transcription activator protein. Biochem. J. **287:**493–499.
42. **Chen, P., D. L. Andersson, and J. R. Roth.** 1994. The control region of the *pdu*/*cob* regulon in *Salmonella typhimurium*. J. Bacteriol. **176:**5474–5482.
43. **Chen, C. W., T. W. Yu, H. M. Chung, and C. F. Chou.** 1992. Discovery and characterization of a new transposable element, Tn*4811*, in *Streptomyces lividans* 66. J. Bacteriol. **174:**7762–7769.
44. **Chou, J. H., J. T. Greenberg, and B. Demple.** 1993. Posttranscriptional repression of *Escherichia coli* OmpF protein in response to redox stress: positive control of the *micF* antisense RNA by the *soxRS* locus. J. Bacteriol. **175:**1026–1031.
45. **Chung, E., E. Allen, R. Araujo, A. Aparicio, K. Davis, M. Duncan, N. Federspiel, R. Hyman, S. Kalman, C. Komp, O. Kurdi, H. Lew, D. Lin, A. Namath, P. Oefner, D. Roberts, S. Schramm, and R. W. Davis.** 1997. SPTREMBL. Accession no. P77634.
46. **Clarke, P., J. H. Lee, K. Burke, and G. Wilcox.** 1992. Mutations in the *araC* gene of *Salmonella typhimurium* LT2 which affect both activator and autoregulatory functions of the AraC protein. Gene **117:**31–37.
47. **Cohen, S. P., H. Hachler, and S. B. Levy.** 1993. Genetic and functional analysis of the multiple antibiotic resistance (*mar*) locus in *Escherichia coli*. J. Bacteriol. **175:**1484–1492.
48. **Cohen, S. P., L. M. McMurry, and S. B. Levy.** 1988. *marA* locus causes decreased expression of OmpF porin in multiple-antibiotic-resistant (Mar) mutants of *Escherichia coli*. J. Bacteriol. **170:**5416–5422.
49. **Collado-Vides, J., B. Magasanik, and J. D. Gralla.** 1991. Control site location and transcriptional regulation in *Escherichia coli*. Microbiol. Rev. **55:** 371–394.
50. **Compan, I., and D. Touati.** 1993. Interaction of six global transcription regulators in expression of manganese superoxide dismutase in *Escherichia coli* K-12. J. Bacteriol. **175:**1686–1696.
51. **Cornelis, G. R.** 1994. *Yersinia* pathogenicity factors. Curr. Top. Microbiol. Immunol. **192:**243–263.
52. **Cornelis, G. R., C. Sluiters, I. Delor, D. Geib, K. Kaniga, C. Lambert de Rouvroit, M. P. Sory, J. C. Vanooteghem, and T. Michiels.** 1991. *ymoA*, a *Yersinia enterocolitica* chromosomal gene modulating the expression of virulence functions. Mol. Microbiol. **5:**1023–1034.
53. **Cornelis, G. R., C. Sluiters, C. Lambert de Rouvroit, and T. Michiels.** 1989. Homology between VirF, the transcriptional activator of the *Yersinia* virulence regulon, and AraC, the *Escherichia coli* arabinose operon regulator. J. Bacteriol. **171:**254–262.
54. **Cornelis, G. R., M. P. Sory, Y. U. Laroche, and I. Derclaye.** 1986. Genetic analysis of the plasmid region controlling virulence in *Yersinia enterocolitica* O:9 by mini-Mu insertions and *lac* fusions. Microb. Pathog. **1:**349–359.
55. **de Haan, L. A. M., G. A. Willshaw, B. A. M. van der Zeijst, and W. Gaastra.** 1991. The nucleotide sequence of a regulatory gene present on a plasmid in an enterotoxigenic *Escherichia coli* strain of serotype O167:H5. FEMS Microbiol. Lett. **67:**341–346.
56. **Dehoux, P., and P. Cossart.** 1995. Homologies between salmolysin and some bacterial regulatory proteins. Mol. Microbiol. **15:**591.
57. **Demple, B., B. Sedgwick, P. Robins, N. Totty, M. D. Waterfield, and T. Lindahl.** 1985. Active sites and complete sequence of the suicidal methyltransferase that counters alkylation mutagenesis. Proc. Natl. Acad. Sci. USA **82:**2688–2692.
58. **D'Orazio, S. E. F., and C. M. Collins.** 1993. The plasmid-encoded urease gene cluster of the family *Enterobacteriaceae* is positively regulated by UreR, a member of the AraC family of transcriptional activators. J. Bacteriol. **175:**3459–3467.
59. **D'Orazio, S. E. F., V. Thomas, and C. M. Collins.** 1996. Activation of transcription at divergent urea-dependent promoters by the urease gene

regulator UreR. Mol. Microbiol. **21:**643–655.
60. **Dudley, E. G., A. C. Husgen, W. He, and J. L. Sttele.** 1996. Sequencing, distribution, and inactivation of the dipeptidase A gene (*pepDA*) from *Lactobacillus helveticus* CNRZ32. J. Bacteriol. **178:**701–704.
61. **Duncan, M., E. Allen, R. Araujo, A. M. Aparicio, E. Chung, K. Davis, N. Federspiel, R. Hyman, S. Kalman, C. Komp, O. Kurdi, H. Lew, D. Lin, A. Namath, P. Oefner, D. Roberts, S. Schramm, and R. W. Davis.** 1996. SPTREMBL. Accession no. P77379 and P77601.
62. **Dunn, T. M., S. Hahn, S. Ogden, and R. F. Schleif.** 1984. An operator at −280 base pairs that is required for repression of *araBAD* operon promoter: addition of DNA helical turns between the operator and promoter cyclically hinders repression. Proc. Natl. Acad. Sci. USA **81:**5017–5020.
63. **Egan, S. M., and R. F. Schleif.** 1993. A regulator cascade in the induction of *rhaBAD*. J. Mol. Biol. **234:**87–98.
64. **Egan, S. M., and R. F. Schleif.** 1994. DNA-dependent renaturation of an insoluble DNA binding protein. J. Mol. Biol. **243:**821–829.
65. **Englesberg, E., R. L. Anderson, R. Weinberg, N. Lee, P. Hoffee, G. Huttenhauer, and H. Boyer.** 1962. L-Arabinose-sensitive, L-ribulose 5-phosphate 4-epimerase-deficient mutants of *Escherichia coli*. J. Bacteriol. **84:**137–146.
66. **Englesberg, E., J. Irr, J. Power, and N. Lee.** 1965. Positive control of enzyme synthesis by gene C in the L-arabinose system. J. Bacteriol. **90:**946–957.
67. **Englesberg, E., C. Squires, and F. Meronk.** 1969. The L-arabinose operon in *Escherichia coli* B/r: a genetic demonstration of two functional states of the product of a regulator gene. Proc. Natl. Acad. Sci. USA **62:**1100–1107.
68. **Entsch, B., L. Squire, and R. E. Wicks.** 1994. SPTREMBL. Accession no. Q51543.
69. **Eustance, R. J., S. A. Bustos, and R. F. Schleif.** 1994. Locating and lengthening the interdomain linker in AraC protein. J. Mol. Biol. **242:**330–338.
70. **Eustance, R. J., and R. F. Schleif.** 1996. The linker region of AraC protein. J. Bacteriol. **178:**7025–7030.
71. **Fawcett, W. P., and R. E. Wolf, Jr.** 1994. Purification of a MalE-SoxS fusion protein and identification of the control sites of *Escherichia coli* superoxide-inducible genes. Mol. Microbiol. **14:**669–679.
72. **Fawcett, W. P., and R. E. Wolf, Jr.** 1995. Genetic definition of the *Escherichia coli zwf* "Soxbox," the DNA binding site for SoxS-mediated induction of glucose 6-phosphate dehydrogenase in response to superoxide. J. Bacteriol. **177:**1742–1750.
73. **Fetherston, J. D., S. W. Bearden, and R. D. Perry.** 1996. YbtA, an AraC-type regulator of the *Yersinia pestis* pesticin/yersiniabactin receptor. Mol. Microbiol. **22:**315–325.
74. **Fleischmann, R. D., M. D. Adams, O. White, R. A. Clayton, E. F. Kirkness, A. R. Kerlavage, C. J. Bult, J. F. Tomb, B. A. Dougherty, J. M. Merrick, K. McKenney, G. Sutton, W. FitzHugh, C. Fields, J. D. Gocayne, J. Scott, R. Shirley, L. I. Liu, A. Glodek, J. M. Kelley, J. F. Weidman, C. A. Phillips, T. Spriggs, E. Hedblom, M. D. Cotton, T. R. Utterback, M. C. Hanna, D. T. Nguyen, D. M. Saudek, R. C. Brandon, L. D. Fine, J. L. Fritchman, J. L. Fuhrmann, N. S. M. Geoghangen, C. L. Gnehm, L. A. McDonald, K. V. Small, C. M. Fraser, H. O. Smith, and J. C. Venter.** 1995. Whole-genome random sequencing and assembly of *Haemophilus influenzae*. Science **269:**496–512.
75. **Forsberg, A., R. Rosqvit, and H. Wolf-Watz.** 1994. Regulation and polarized transfer of the *Yersinia* outer proteins (Yops) involved in antiphagocytosis. Trends Microbiol. **2:**14–19.
76. **Forsberg, A., and H. Wolf-Watz.** 1988. The virulence protein Yop5 of *Yersinia pseudotuberculosis* is regulated at transcriptional level by plasmid-pIB1-encoded *trans*-acting elements controlled by temperature and calcium. Mol. Microbiol. **2:**121–133.
77. **Frank, D. W., and B. H. Iglewski.** 1991. Cloning and sequence analysis of a *trans*-regulatory locus required for exoenzyme S synthesis in *Pseudomonas aeruginosa*. J. Bacteriol. **173:**6460–6468.
78. **Franklyn, C. S., and N. Lee.** 1988. AraC proteins with altered DNA sequence specificity which activate a mutant promoter in *Escherichia coli*. J. Biol. Chem. **263:**4400–4407.
79. **Franklyn, F. C. H., P. R. Lehrbach, R. Lurz, B. Rueckert, M. Bagdasarian, and K. N. Timmis.** 1983. Localization and functional analysis of transposon mutations in regulatory genes of the TOL catabolic pathway. J. Bacteriol. **154:**676–685.
80. **Friedrich, M. J., N. E. Kinsey, J. Vila, and R. J. Kadner.** 1993. Nucleotide sequence of a 13.9 kb segment of the 90 kb virulence plasmid of *Salmonella typhimurium*: the presence of fimbrial biosynthetic genes. Mol. Microbiol. **8:**543–558.
81. **Gallegos, M. T.** 1996. Caracterización del regulador transcriptional XylS del plásmido TOL de *Pseudomonas putida*. Ph.D. dissertation. University of Granada, Granada, Spain.
82. **Gallegos, M. T., S. Marqués, and J. L. Ramos.** 1996. Expression of the TOL plasmid *xyl*S gene in *Pseudomonas putida* occurs from a σ^{70}-dependent promoter or from σ^{70}- and σ^{54}-dependent tandem promoters according to the aromatic compound used for growth. J. Bacteriol. **178:**2356–2361.
83. **Gallegos, M. T., S. Marqués, and J. L. Ramos.** 1996. The $TACAN_4TGCA$ motif upstream from the −35 region on the σ^{70}/σ^{S}-dependent Pm promoter of the TOL plasmid is the minimum DNA segment required for transcription stimulation by XylS regulators. J. Bacteriol. **178:**6427–6434.
84. **Gallegos, M. T., C. Michán, and J. L. Ramos.** 1993. The XylS/AraC family of regulators. Nucleic Acids Res. **21:**807–810.
85. **Galyov, E. E., A. V. Karlishev, T. V. Chernovskaya, D. A. Dolgikh, O. Y. Smirnov, K. L. Volkovoy, V. M. Abramov, and V. P. Zav'yalov.** 1991. Expression of the envelope antigen F1 of *Yersinia pestis* is mediated by the product of *caf1M* gene having homology with the chaperone protein PapD of *Escherichia coli*. FEBS Lett. **286:**79–82.
86. **Gambino, L., S. J. Gracheck, and P. F. Miller.** 1993. Overexpression of the MarA positive regulator is sufficient to confer multiple antibiotic resistance in *Escherichia coli*. J. Bacteriol. **175:**2888–2894.
87. **Gardner, P. R., and I. Fridovich.** 1993. NADPH inhibits transcription of the *Escherichia coli* mangenese superoxide dismutase gene (*sodA*) *in vitro*. J. Biol. Chem. **268:**12958–12963.
88. **Geelen, D., K. Goethals, M. Van Montagu, and M. Holsters.** 1995. The *nodD* locus from *Azorhizobium caulinodans* is flanked by two repetitive elements. Gene **164:**107–111.
89. **Genin, S., C. L. Gough, C. Zischeck, and C. A. Boucher.** 1992. Evidence that the *hrpB* gene encodes a positive regulator of pathogenicity genes from *Pseudomonas solanacearum*. Mol. Microbiol. **6:**3065–3076.
90. **George, A. M., R. M. Hall, and H. W. Stokes.** 1995. Multidrug resistance in *Klebsiella pneumoniae*: a novel gene, *ramA*, confers a multidrug resistance phenotype in *Escherichia coli*. Microbiology **141:**1909–1920.
91. **Gómez-Duarte, O. G., and J. B. Kaper.** 1995. A plasmid encoded regulatory region activates chromosomal *eaeA* expression in enteropathogenic *Escherichia coli*. Infect. Immun. **63:**1767–1776.
92. **González, M. M.** 1997. Personal communication.
93. **Greenberg, J. T., and B. Demple.** 1989. A global response induced in *Escherichia coli* by redox-cycling agents overlaps with that induced by peroxide stress. J. Bacteriol. **171:**3933–3939.
94. **Greenberg, J. T., P. Monach, J. H. Chou, P. D. Josephy, and B. Demple.** 1990. Positive control of a global antioxidant defense regulon activated by superoxide-generating agents in *Escherichia coli*. Proc. Natl. Acad. Sci. USA **87:**6181–6185.
95. **Greenblatt, J., and R. Schleif.** 1971. Arabinose C protein: regulation of the arabinose operon in vitro. Nature (London) New Biol. **233:**166–170.
96. **Gruer, M. J., and J. R. Guest.** 1994. Two genetically-distinct and differentially-regulated aconitases (AcnA and AcnB) in *Escherichia coli*. Microbiology **140:**2531–2541.
97. **Gupta, S., and A. K. Tyagi.** 1993. Sequence of a newly identified *Mycobacterium tuberculosis* gene encoding a protein with sequence homology to virulence-regulating proteins. Gene **126:**157–158.
98. **Haak, B., S. Fetzner, and F. Lingens.** 1995. Cloning, nucleotide sequence, and expression of the plasmid-encoded genes for the two-component 2-halobenzoate 1,2-dioxygenase from *Pseudomonas cepacia* 2CBS. J. Bacteriol. **177:**667–675.
99. **Hahn, S., T. Dunn, and R. Schleif.** 1984. Upstream repression and CRP stimulation of the *Escherichia coli* L-arabinose operon. J. Mol. Biol. **180:**61–72.
100. **Hakura, A., K. Morimoto, T. Sofuni, and T. Nohmi.** 1991. Cloning and characterization of the *Salmonella typhimurium ada* gene, which encodes *O*-6-methylguanine-DNA methyltransferase. J. Bacteriol. **173:**3663–3672.
101. **Hamilton, E. P., and N. Lee.** 1988. Three binding sites for AraC protein are required for autoregulation of *araC* in *Escherichia coli*. Proc. Natl. Acad. Sci. USA **85:**1749–1753.
102. **Harrison, S. C.** 1991. A structural taxonomy of DNA-binding domains. Nature **353:**715–719.
103. **Harrison, S. C., and A. K. Aggarwal.** 1990. DNA recognition by proteins with the helix-turn-helix motif. Annu. Rev. Biochem. **59:**933–969.
104. **Haydon, D. J., and J. R. Guest.** 1991. A new family of bacterial regulatory proteins. FEMS Microbiol. Lett. **63:**291–295.
105. **Hebert, M. D., and J. E. Houghton.** 1996. SPTREMBL. Accession no. P72171.
106. **Heinrichs, D. E., and K. Poole.** 1993. Cloning and sequence analysis of a gene (*pchR*) encoding an AraC family activator of pyochelin and ferripyochelin receptor synthesis in *Pseudomonas aeruginosa*. J. Bacteriol. **175:**5882–5889.
107. **Helmann, J. D., B. T. Ballard, and C. T. Walsh.** 1990. The MerR metalloregulatory protein binds mercuric ion as a tricoordinate, metal-bridged dimer. Science **247:**946–948.
108. **Hendrickson, W., C. Flaherty, and L. Molz.** 1992. Sequence elements in the *Escherichia coli araFGH* promoter. J. Bacteriol. **174:**6862–6871.
109. **Hendrickson, W., and R. Schleif.** 1985. A dimer of AraC protein contacts three adjacent major groove regions of the *araI* DNA site. Proc. Natl. Acad. Sci. USA **82:**3129–3133.
110. **Hendrickson, W., C. Stoner, and R. Schleif.** 1990. Characterization of the *araFGH* and *araJ* promoters. J. Mol. Biol. **215:**497–510.
111. **Henikoff, S., G. W. Haughn, J. M. Calvo, and J. C. Wallace.** 1988. A large family of bacterial activator proteins. Proc. Natl. Acad. Sci. USA **85:**6602–6606.
112. **Henikoff, S., J. C. Wallace, and J. P. Brown.** 1990. Finding protein similar-

ities with nucleotide sequence databases. Methods Enzymol. **183:**111–132.
113. **Higgins, D. E., E. Nazareno, and V. J. Dirita.** 1992. The virulence gene activator ToxT from *Vibrio cholerae* is a member of the AraC family of transcriptional activators. J. Bacteriol. **174:**6974–6980.
114. **Hirsh, J., and R. Schleif.** 1973. On the mechanism of action of L-arabinose C gene activator and lactose repressor. J. Mol. Biol. **80:**433–444.
115. **Hoe, N. P., and J. D. Goguen.** 1993. Temperature sensing in *Yersinia pestis*: translation of the LcrF activator protein is thermally regulated. J. Bacteriol. **174:**7901–7909.
116. **Hoe, N. P., F. C. Minion, and J. D. Goguen.** 1992. Temperature sensing in *Yersinia pestis*: regulation of *yopE* transcription by *lcrF*. J. Bacteriol. **174:** 4275–4286.
117. **Huo, L., K. J. Martin, and R. Schleif.** 1988. Alternative DNA loops regulate the arabinose operon in *Escherichia coli*. Proc. Natl. Acad. Sci. USA **85:** 5444–5448.
118. **Inouye, S., A. Nakazawa, and T. Nakazawa.** 1981. Molecular cloning of gene *xyl*S of the TOL plasmid: evidence for positive regulation of the *xyl*DEGF operon by *xyl*S. J. Bacteriol. **148:**413–418.
119. **Inouye, S., A. Nakazawa, and T. Nakazawa.** 1986. Nucleotide sequence of the regulatory gene *xyl*S on the *Pseudomonas putida* TOL plasmid and identification of the protein product. Gene **44:**235–242.
120. **Inouye, S., A. Nakazawa, and T. Nakazawa.** 1987. Overproduction of the *xyl*S gene product and activation of the *xyl*DLEGF operon on the TOL plasmid. J. Bacteriol. **169:**3587–3592.
121. **Ishida, H., H. Fujiwara, T. Horiuchi, K. Sato, Y. Osada, and M. Furusawa.** 1994. SWISSPROT. Accession no. W52620.
122. **Ishihama, A.** 1993. Protein-protein communication within the transcription apparatus. J. Bacteriol. **175:**2483–2489.
123. **Johnson, C., and R. F. Schleif.** 1995. In vivo kinetics of the arabinose promoters in *Escherichia coli*. J. Bacteriol. **177:**3438–3442.
124. **Jordi, B. J. A. M., B. Dagberg, L. A. M. de Haan, A. M. Hamers, B. A. M. van der Zeijst, W. Gaastra, and B. E. Uhlin.** 1992. The positive regulator CfaD overcomes the repression mediated by histone-like protein H-NS (H1) in the CFA/I fimbrial operon of *Escherichia coli*. EMBO J. **11:**2627–2632.
125. **Kaldalu, N., T. Mandel, and M. Ustav.** 1996. TOL plasmid transcription factor XylS binds specifically to the Pm operator sequence. Mol. Microbiol. **20:**569–579.
126. **Kaneko, T., S. Sato, H. Kotani, A. Tanaka, E. Asamizu, Y. Nakamura, N. Miyajima, M. Hirosawa, M. Sugiura, S. Sasamoto, T. Kimura, T. Hosouchi, A. Matsuno, A. Muraki, N. Nakazaki, K. Naruo, S. Okumura, S. Shimpo, C. Takeuchi, T. Wada, A. Watanabe, M. Yamada, M. Yasuda, and S. Tabata.** 1996. Sequence analysis of the genome of the unicellular cyanobacterium *Synechocystis* sp. strain PCC 6803. Sequence determination of the entire genome and assignment of potential coding regions. DNA Res. **3:**109–136.
127. **Kaniga, K., J. C. Bossio, and J. E. Galán.** 1994. The *Salmonella typhimurium* invasion genes *invF* and *invG* encode homologues of the AraC and PulD family of proteins. Mol. Microbiol. **13:**555–568.
128. **Karlyshev, A. V., E. E. Galyov, V. M. Abramov, and V. P. Zav'yalov.** 1992. *caf1R* gene and its role in the regulation of capsule formation of *Y. pestis*. FEBS Lett. **305:**37–40.
129. **Kato, J. I., K. I. Ito, A. Nakamura, and H. Watanabe.** 1989. Cloning of regions required for contact hemolysis and entry into LLC-MK2 cells from *Shigella sonnei* form I plasmid: *virF* is a positive regulator gene for these phenotypes. Infect. Immun. **57:**1391–1398.
130. **Kaufman, M. R., C. E. Shaw, I. D. Jones, and R. K. Taylor.** 1993. Biogenesis and regulation of the *Vibrio cholerae* toxin-coregulated pilus: analogies to other virulence factor secretory systems. Gene **126:**43–49.
131. **Kemp, E. H., N. P. Minton, and N. H. Mann.** 1987. Complete nucleotide sequence and deduced amino acid sequence of the M5 polypeptide gene of *Escherichia coli*. Nucleic Acids Res. **15:**3924.
132. **Kessler, B., V. de Lorenzo, and K. N. Timmis.** 1993. Identification of a *cis*-acting sequence within the Pm promoter of the TOL plasmid which confers XylS-mediated responsiveness to substituted benzoates. J. Mol. Biol. **230:**699–703.
133. **Klaasen, P., and F. K. de Graaf.** 1990. Characterization of FapR, a positive regulator of expression of the 987P operon in enterotoxigenic *Escherichia coli*. Mol. Microbiol. **4:**1779–1783.
134. **Koh, Y.-S., J. Choih, J. H. Lee, and J. H. Roe.** 1996. Regulation of the *ribA* gene encoding GTP cyclohydrolase II by the *soxRS* locus in *Escherichia coli*. Mol. Gen. Genet. **251:**591–598.
135. **Koh, Y.-S., and J. H. Roe.** 1995. Isolation of a novel paraquat-inducible (*pqi*) gene regulated by the *soxRS* locus in *Escherichia coli*. J. Bacteriol. **177:**2673–2678.
136. **Koh, Y.-S., and J.-H. Roe.** 1996. Dual regulation of the paraquat-inducible gene *pqi*-5 by SoxS and RpoS in *Escherichia coli*. Mol. Microbiol. **22:**53–61.
137. **Komeda, H., Y. Hori, M. Kobayashi, and S. Shimizu.** 1996. Transcriptional regulation of the *Rhodococcus rhodochrous* J1 *nitA* gene encoding nitrilase. Proc. Natl. Acad. Sci. USA **93:**10572–10577.
138. **Kondo, H., Y. Nakabeppu, H. Kataoka, S. Kuhara, S. Kawabata, and M. Sekiguchi.** 1986. Structure and expression of the *alkB* gene of *Escherichia coli* related to the repair of alkylated DNA. J. Biol. Chem. **261:**15772–15777.
139. **Kormanec, J., A. Lempelova, M. Farkasovsky, and D. Homerova.** 1995. Cloning, sequencing and expression in *Escherichia coli* of a *Streptomyces aureofaciens* gene encoding glyceraldehyde-3-phosphate dehydrogenase. Gene **165:**77–80.
140. **Kyrpides, N. C., and C. A. Ouzounis.** 1995. Nucleic acid-binding metabolic enzymes: living fossils of stereochemical interactions? J. Mol. Evol. **40:**564–569.
141. **Lambert de Rouvroit, C., C. Sluiters, and G. R. Cornelis.** 1992. Role of the transcriptional activator, VirF, and temperature in the expression of the pYV plasmid genes of *Yersinia enterocolitica*. Mol. Microbiol. **6:**395–409.
142. **Landini, P., and M. R. Volkert.** 1995. RNA polymerase alfa subunit binding site in positively controlled promoters: a new model for RNA polymerase-promoter interaction and transcriptional activation in the *Escherichia coli ada* and *aidB* genes. EMBO J. **14:**4329–4335.
143. **Lauble, H., Y. Georgalis, and U. Heinemann.** 1989. Studies on the domain structure of the *Salmonella typhimurium* AraC protein. Eur. J. Biochem. **185:**319–325.
144. **Lee, D. H., L. Huo, and R. Schleif.** 1992. Repression of the *araBAD* promoter from $araO_1$. J. Mol. Biol. **224:**335–341.
145. **Lee, D. H., and R. Schleif.** 1989. *In vivo* DNA loops in *araCBAD*: size limit and helical repeat. Proc. Natl. Acad. Sci. USA **86:**476–480.
146. **Lee, N., C. Franklyn, and E. P. Hamilton.** 1987. Arabinose-induced binding of AraC protein to $araI_2$ activates the *araBAD* operon promoter. Proc. Natl. Acad. Sci. USA **84:**8814–8818.
147. **Leenhouts, K. K. J., A. A. Bolhuis, J. J. Kok, and G. G. Venema.** 1994. SWISSPROT. Accession no. P43465.
148. **Lehmann, M., B. Tshisuaka, S. Fetzner, and F. Lingens.** 1995. Molecular cloning of the isoquinoline 1-oxidoreductase genes from *Pseudomonas diminuta*, structural analysis of *iorA* and *iorB*, and sequence comparisons with other molybdenum-containing hydroxylases. J. Biol. Chem. **270:**14420–14429.
149. **Lei, S. P., H. C. Lin, L. Heffernan, and G. Wilcox.** 1985. *araB* gene and nucleotide sequence of *araC* gene of *Erwinia caratovora*. J. Bacteriol. **164:** 717–722.
150. **Lemotte, P. K., and G. C. Walker.** 1985. Induction and autoregulation of *ada*, a positively acting element regulating the response of *Escherichia coli* K-12 to methylating agents. J. Bacteriol. **161:**888–895.
151. **Li, Z., and B. Demple.** 1994. SoxS, an activator of superoxide stress genes in *Escherichia coli*: purification and interaction with DNA. J. Biol. Chem. **269:**18371–18377.
152. **Li, Z., and B. Demple.** 1996. Sequence specificity for DNA binding by *Escherichia coli* SoxS and Rob proteins. Mol. Microbiol. **20:**937–945.
153. **Lin, J. W., K. Y. Yu, Y. F. Chao, and S. F. Weng.** 1995. The *lumQ* gene is linked to the *lumP* gene and the *lux* operon in *Photobacterium leiognathi*. Biochem. Biophys. Res. Commun. **217:**684–695.
154. **Liochev, S. I., and I. Fridovich.** 1992. Fumarase C, the stable fumarase of *Escherichia coli*, is controlled by the *soxRS* regulon. Proc. Natl. Acad. Sci. USA **89:**5892–5896.
155. **Liochev, S. I., A. Hausladen, W. F. Beyer, and I. Fridovich.** 1994. NADPH: ferredoxin oxidoreductase acts as a paraquat diaphorase and is a member of the *soxRS* regulon. Proc. Natl. Acad. Sci. USA **91:**1328–1331.
156. **Lobell, R., and R. Schleif.** 1990. DNA looping and unlooping by AraC protein. Science **250:**528–532.
157. **Lobell, R. B., and R. Schleif.** 1991. AraC-DNA looping orientation and distance-dependent loop breaking by the cyclic AMP receptor protein. J. Mol. Biol. **218:**45–54.
158. **Lu, Y., C. Flaherty, and W. Hendrickson.** 1992. AraC protein contacts asymmetric sites in the *Escherichia coli araFGH* promoter. J. Biol. Chem. **267:**24848–24857.
159. **Lundrigan, M. D., M. J. Friedrich, and R. J. Kadner.** 1989. Nucleotide sequence of the *Escherichia coli* porin thermoregulatory gene *envY*. Nucleic Acids Res. **17:**800.
160. **Lüthy, R., I. Xenarios, and D. Eisenberg.** 1994. Profile analysis: detection of distantly related proteins. Proc. Natl. Acad. Sci. USA **84:**4355–4358.
161. **Macinga, D. R., M. M. Parojcic, and P. N. Rather.** 1995. Identification and analysis of *aarP*, a transcriptional activator of the 2′-*N*-acetyltransferase in *Providencia stuartii*. J. Bacteriol. **177:**3407–3414.
162. **Manzanera, M.** 1997. Personal communication.
163. **Marqués, S., M. T. Gallegos, and J. L. Ramos.** 1995. Role of σ^S in transcription from the positively controlled Pm promoter of the TOL plasmid of *Pseudomonas putida*. Mol. Microbiol. **18:**851–857.
164. **Marqués, S., and J. L. Ramos.** 1993. Transcriptional control of the *Pseudomonas putida* TOL plasmid catabolic pathways. Mol. Microbiol. **9:**923–929.
165. **Martin, K., L. Huo, and R. Schleif.** 1986. The DNA loop model for *ara* repression: AraC occupies the proposed loop sites *in vivo* and repression negative mutations lie in these same sites. Proc. Natl. Acad. Sci. USA **83:**3654–3658.
166. **Martins, E. A. L., and B. Demple.** 1996. SWISSPROT. Accession no. Q56143.
167. **McKay, D. B., and T. A. Steitz.** 1981. Structure of catabolite gene activator

protein at 2.9 Å resolution suggests binding to left-handed *B*-DNA. Nature **290:**744–749.
168. **Menon, K. P., and N. L. Lee.** 1990. Activation of *ara* operons by a truncated AraC protein does not require inducer. Proc. Natl. Acad. Sci. USA **87:** 3708–3712.
169. **Mermod, N., J. L. Ramos, A. Bairoch, and K. N. Timmis.** 1987. The XylS gene positive regulator of TOL plasmid pWWO: identification, sequence analysis and overproducing leading to constitutive expression of *meta* cleavage operon. Mol. Gen. Genet. **207:**349–354.
170. **Michán, C. M., S. J. W. Busby, and E. I. Hyde.** 1995. The *Escherichia coli* MelR transcriptional activator: production of a stable fragment containing the DNA-binding domain. Nucleic Acids Res. **23:**1518–1523.
171. **Michán, C., B. Kessler, V. de Lorenzo, K. N. Timmis, and J. L. Ramos.** 1992. XylS domain interactions can be deduced from intraallelic dominance in double mutants. Mol. Gen. Genet. **235:**406–412.
172. **Michán, C., L. Zhou, M. T. Gallegos, K. N. Timmis, and J. L. Ramos.** 1992. Identification of critical amino-terminal regions of XylS. The positive regulator encoded by the TOL plasmid. J. Biol. Chem. **267:**22897–22901.
173. **Mito, S., Q. M. Zhang, and S. Yonei.** 1993. Isolation and characterization of *Escherichia coli* strains containing new gene fusions (*soi*::*lac*Z) inducible by superoxide radicals. J. Bacteriol. **175:**2645–2651.
174. **Miyada, C. G., A. H. Horwitz, L. G. Cass, J. Timko, and G. Wilcox.** 1980. DNA sequence of the *araC* regulatory gene from *Escherichia coli* B/n. Nucleic Acids Res. **8:**5267–5274.
175. **Morby, A. P., J. S. Turner, J. W. Huckle, and N. J. Robinson.** 1993. SmtB is a metal-dependent repressor of the cyanobacterial metallothionein gene *smt*A: identification of a Zn inhibited DNA-protein complex. Nucleic Acids Res. **21:**921–925.
176. **Moroshi, F., K. Hayashi, and N. Munakata.** 1990. *Bacillus subtilis ada* operon encodes two DNA alkyltransferases. Nucleic Acids Res. **18:**5473–5480.
177. **Murphy, L., D. Harris, B. G. Barrell, M. A. Rajandream, and S. V. Walsh.** 1996. SWISSPROT. Accession no. Q10630.
178. **Nagy, Y., G. Schoofs, F. Compernolle, P. Proost, J. Vanderleyden, and R. De Mot.** 1995. Degradation of thiocarbamate herbicide EPTC (*S*-ethyl dipropylcarbamothioate) and biosafening by *Rhodococcus* sp. strain NI86/21 involve an inducible cytochrome P-450 system and aldehyde dehydrogenase. J. Bacteriol. **177:**676–687.
179. **Naik, S. M., and H. M. Hassan.** 1990. Use of site-directed mutagenesis to identify an upstream regulatory sequence of *sodA* gene of *Escherichia coli* K-12. Proc. Natl. Acad. Sci. USA **87:**2618–2622.
180. **Nakabeppu, Y., H. Kondo, S. Kawabata, S. Iwanaga, and M. Sekiguchi.** 1985. Purification and structure of the intact Ada regulatory protein of *Escherichia coli* K12, *O*-6-methylguanine-DNA methyltransferase. J. Biol. Chem. **260:**7281–7288.
181. **Nakajima, H., K. Kobayashi, M. Kobayashi, H. Asako, and R. Aono.** 1995. Overexpression of the *robA* gene increases organic solvent tolerance and multiple antibiotic and heavy metal ion resistance in *Escherichia coli*. Appl. Environ. Microbiol. **61:**2302–2307.
182. **Nakayama, S., and H. Watanabe.** 1995. Involvement of *cpxA*, a sensor of a two-component regulatory system in the pH-dependent regulation of expression of the *Shigella sonnei virF* gene. J. Bacteriol. **177:**5062–5069.
183. **Nataro, J. P., D. Yikang, D. Yingkang, and K. Walker.** 1994. AggR, a transcriptional activator of aggregative adherence fimbria I expression in enteroaggregative *Escherichia coli*. J. Bacteriol. **176:**4691–4699.
184. **Neidhardt, F. C.** 1987. Multigene systems and regulons, p. 1313–1317. *In* F. C. Neidhardt, J. L. Ingraham, K. B. Low, B. Magasanik, M. Schaechter, and H. E. Umbarger (ed.), *Escherichia coli* and *Salmonella typhimurium*: cellular and molecular biology, vol. 2. American Society for Microbiology, Washington, D.C.
185. **Nicholson, E. B., E. A. Concaugh, P. A. Foxall, M. D. Island, and H. L. T. Mobley.** 1993. *Proteus mirabilis* urease: transcriptional regulation by UreR. J. Bacteriol. **175:**465–473.
186. **Niland, P., R. Hühne, and B. Müller-Hill.** 1996. How AraC interacts specifically with its target DNAs. J. Mol. Biol. **264:**667–674.
187. **Nishitani, J., and G. Wilcox.** 1991. Cloning and characterization of the L-rhamnose regulon in *Salmonella typhimurium* LT2. Gene **105:**37–42.
188. **Niu, W., Y. Kim, G. Tan, T. Heyduk, and R. H. Ebright.** 1996. Transcription activation at class II CAP-dependent promoters: two interactions between CAP and RNA polymerase. Cell **87:**1123–1134.
189. **North, A. K., K. E. Klose, K. M. Stedman, and S. Kustu.** 1993. Prokaryotic enhancer-binding proteins reflect eukaryote-like modularity: the puzzle of nitrogen regulatory protein C. J. Bacteriol. **175:**4267–4273.
190. **Nunoshiba, T., E. Hidalgo, C. F. Amábile-Cuevas, and B. Demple.** 1992. Two-stage control of an oxidative stress regulon: the *Escherichia coli* SoxR protein triggers redox-inducible expression of the *soxS* regulatory gene. J. Bacteriol. **174:**6054–6060.
191. **Nunoshiba, T., E. Hidalgo, Z. Li, and B. Demple.** 1993. Negative autoregulation by the *Escherichia coli* SoxS protein: a dampening mechanism for the *soxRS* redox stress response. J. Bacteriol. **175:**7492–7494.
192. **Ogierman, M. A., and P. A. Manning.** 1992. Homology of TcpN, a putative regulatory protein of *Vibrio cholerae*, to the AraC family of transcriptional activators. Gene **116:**93–97.
193. **Oku, T., A. M. Alvarez, and C. I. Kado.** 1995. Conservation of the hypersensitivity-pathogenicity regulatory gene *hrpX* of *Xanthomonas campestris* and *oryzae*. DNA Seq. **5:**245–249.
194. **Olsen, G. J., C. R. Woese, and R. Overbeek.** 1994. The winds of (evolutionary) change: breathing new life into microbiology. J. Bacteriol. **176:**1–6.
195. **Pabo, C. O., and M. Lewis.** 1992. The operator-binding domain of lambda repressor: structure and DNA recognition. Nature **298:**443–447.
196. **Pabo, C. O., and R. T. Sauer.** 1984. Protein-DNA recognition. Annu. Rev. Biochem. **53:**293–321.
197. **Pabo, C. O., and R. T. Sauer.** 1992. Transcription factors: structural families and principles of DNA recognition. Annu. Rev. Biochem. **61:**1053–1095.
198. **Pan, W., and B. G. Spratt.** 1994. Regulation of the permeability of the gonococcal cell envelope by the *mtr* system. Mol. Microbiol. **11:**769–775.
199. **Parker, L. L., and B. G. Hall.** 1990. Characterization and nucleotide sequence of the cryptic *cel* operon of *Escherichia coli* K12. Genetics **124:**455–471.
200. **Philipp, W. J., S. Poulet, K. Eiglmeier, L. Pascopella, V. Balasubramanian, B. Heym, S. Bergh, B. R. Bloom, W. R. Jacobs Jr., and S. T. Cole.** 1996. An integrated map of the genome of the tubercle bacillus, *Mycobacterium tuberculosis* H37Rv, and comparison with *Mycobacterium leprae*. Proc. Natl. Acad. Sci. USA **93:**3132–3137.
201. **Plunkett, G., III, V. D. Burland, D. L. Daniels, and F. R. Blattner.** 1993. Analysis of the *Escherichia coli* genome. III. DNA sequence of the region from 87.2 to 89.2 minutes. Nucleic Acids Res. **21:**3391–3398.
202. **Porter, M. E., and C. J. Dorman.** 1994. A role of H-NS in the thermo-osmotic regulation of virulence gene expression in *Shigella flexneri*. J. Bacteriol. **176:**4187–4191.
203. **Price, S. T., K. Y. Leung, S. S. Barve, and S. C. Straley.** 1989. Molecular analysis of LcrGVH, the V antigen operon of *Yersinia pestis*. J. Bacteriol. **171:**5646–5653.
204. **Prieto, M. A., and J. L. García.** 1994. Molecular characterization of 4-hydroxyphenylacetate 3-hydroxylase of *Escherichia coli*. J. Biol. Chem. **269:** 22823–22829.
205. **Quirk, P. G., A. A. Guffanti, S. Clejan, J. Cheng, and T. A. Krulwich.** 1994. Isolation of Tn*917* insertion mutants of *Bacillus subtilis* that are resistant to the protonophore carbonyl cyanide *m*-chlorophenylhydrazone. Biochim. Biophys. Acta **1186:**27–34.
206. **Ramos, J. L., S. Marqués, and K. N. Timmis.** Transcriptional control of the *Pseudomonas* TOL plasmid catabolic operons is achieved through an interplay of host factors and plasmid encoded regulators. Annu. Rev. Microbiol., in press.
207. **Ramos, J. L., N. Mermod, and K. N. Timmis.** 1987. Regulatory circuits controlling transcription of TOL plasmid encoding *meta*-cleavage pathway for degradation of alkylbenzoates by *Pseudomonas*. Mol. Microbiol. **1:**293–300.
208. **Ramos, J. L., C. Michán, F. Rojo, D. Dwyer, and K. N. Timmis.** 1990. Signal-regulator interactions: genetic analysis of the effector binding site of *xyl*S, the benzoate-activated positive regulator of *Pseudomonas* TOL plasmid *meta*-cleavage pathway operon. J. Mol. Biol. **211:**373–382.
209. **Ramos, J. L., F. Rojo, L. Zhou, and K. N. Timmis.** 1990. A family of positive regulators related to the *Pseudomonas putida* plasmid XylS and the *Escherichia coli* AraC activators. Nucleic Acids Res. **18:**2149–2152.
210. **Ramos, J. L., A. Stolz, W. Reineke, and K. N. Timmis.** 1986. Altered effector specificities in regulators of gene expression: TOL plasmid *xyl*S mutants and their use to engineer expansion of the range of aromatics degraded by bacteria. Proc. Natl. Acad. Sci. USA **83:**8467–8471.
211. **Ramos, J. L., A. Wasserfallen, K. Rose, and K. N. Timmis.** 1987. Redesigning metabolic routes: manipulation of TOL plasmid pathway for catabolism of alkylbenzoates. Science **235:**593–596.
212. **Reeder, T., and R. Schleif.** 1993. AraC protein can activate transcription from only one position and when pointed in only one direction. J. Mol. Biol. **231:**205–218.
213. **Reizer, A., J. Deutscher, M. H. Saier, and J. Reizer.** 1991. Analysis of the gluconate (*gnt*) operon of *Bacillus subtilis*. Mol. Microbiol. **5:**1081–1089.
214. **Reverchon, S., W. Nasser, and J. Robert-Baudouy.** 1991. Characterization of *kdg*R, a gene of *Erwinia chrysanthemi* that regulates pectin degradation. Mol. Microbiol. **5:**2203–2216.
215. **Rohde, J. R., J. M. Fox, and S. A. Minnich.** 1994. Thermoregulation in *Yersinia enterocolitica* is coincident with changes in DNA supercoiling. Mol. Microbiol. **12:**187–199.
216. **Rosner, J. L., and J. L. Slonczewski.** 1994. Dual regulation of *inaA* by the multiple antibiotic resistance (Mar) and superoxide (SoxRS) stress response systems in *Escherichia coli*. J. Bacteriol. **176:**6262–6269.
217. **Rost, B., and C. Sander.** 1993. Improved prediction of protein secondary structure by use of sequence profiles and neural networks. Proc. Natl. Acad. Sci. USA **90:**7558–7562.
218. **Roth, J. R., J. G. Lawrence, M. Rubenfield, S. Kieffer-Higgins, and G. M. Church.** 1993. Characterization of the cobalamin (vitamin B_{12}) biosynthetic genes of *Salmonella typhimurium*. J. Bacteriol. **175:**3303–3316.
219. **Russell, R. R. B., J. Aduse-Opoku, I. C. Sutcliffe, L. Tao, and J. J. Ferretti.**

1992. A binding protein-dependent transport system in *Streptococcus mutans* responsible for multiple sugar metabolism. J. Biol. Chem. **267:**4631–4637.
220. **Sakai, T., C. Sasakawa, S. Makino, and M. Yoshikawa.** 1986. DNA sequence and product analysis of the *virF* locus responsible for Congo red binding and cell invasion in *Shigella flexneri* 2a. Infect. Immun. **54:**395–402.
221. **Sander, C., and R. Schneider.** 1991. Database homology-derived protein structures and the structural meaning of sequence alignment. Proteins **9:**56–68.
222. **Savelkoul, P. H. M., G. A. Willshaw, M. M. McConnell, H. R. Smith, A. M. Hamers, B. A. M. van der Zeijst, and W. Gaastra.** 1990. Expression of CFA/I fimbriae is positively regulated. Microb. Pathog. **8:**91–99.
223. **Schleif, R.** 1969. An L-arabinose binding protein and arabinose permeation in *Escherichia coli*. J. Mol. Biol. **46:**185–196.
224. **Schleif, R.** 1987. The L-arabinose operon, p. 1473–1481. *In* F. C. Neidhardt, J. L. Ingraham, K. B. Low, B. Magasanik, M. Schaechter, and H. E. Umbarger (ed.), *Escherichia coli* and *Salmonella typhimurium*: cellular and molecular biology. American Society for Microbiology, Washington, D.C.
225. **Schleif, R.** 1992. Regulation of the L-arabinose catabolic operon *araBAD*, p. 643–665. *In* S. L. McKnight and K. R. Yamamoto (ed.), Transcriptional regulation, vol. 2. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
226. **Schleif, R., and M. Favreau.** 1982. Hyperproduction of AraC protein from *Escherichia coli*. Biochemistry **21:**778–782.
227. **Schollmeier, K., and W. Hillen.** 1984. Transposon Tn*10* contains two structural genes with opposite polarity between *tetA* and IS*10R*. J. Bacteriol. **160:**499–503.
228. **Schwecke, T., J. F. Aparicio, I. Molnar, A. Koening, L. E. Khaw, S. F. Haydock, M. Oliynyk, P. Caffrey, J. Cortes, J. B. Lester, G. A. Boehm, J. Staunton, and P. F. Leadlay.** 1995. The biosynthetic gene cluster for the polyketide immunosuppressant rapamycin. Proc. Natl. Acad. Sci. USA **92:**7839–7843.
229. **Skarstad, K., B. Thöny, D. S. Hwang, and A. Kornberg.** 1993. A novel binding protein of the origin of the *Escherichia coli* chromosome. J. Biol. Chem. **268:**5365–5370.
230. **Skurnik, M., and P. Toivanen.** 1992. LcrF is the temperature-regulated activator of the *yadA* gene of *Yersinia enterocolitica* and *Yersinia pseudotuberculosis*. J. Bacteriol. **174:**2047–2051.
231. **Sofia, H. J., V. Burland, D. L. Daniels, G. Plunkett III, and F. R. Blattner.** 1994. Analysis of the *Escherichia coli* genome. V. DNA sequence of the region from 76.0 to 81.5 minutes. Nucleic Acids Res. **22:**2576–2586.
232. **Soisson, S. M., B. MacDougall-Shackleton, R. Schleif, and C. Wolberger.** 1997. Structural basis for ligand-regulated oligomerization of AraC. Science **276:**421–425.
233. **Spiro, S., and J. R. Guest.** 1990. FNR and its role in oxygen-regulated gene expression in *Escherichia coli*. FEMS Microbiol. Rev. **6:**399–428.
234. **Spooner, R. A., M. Bagdasarian, and F. C. H. Franklin.** 1987. Activation of the *xylDLEGF* promoter of the TOL toluene-xylene degradation pathway by overproduction of the *xylS* regulatory gene product. J. Bacteriol. **169:**3581–3586.
235. **Spooner, R. A., K. Lindsay, and F. C. H. Franklyn.** 1986. Genetic, functional and sequence analysis of the *xyl*R and *xyl*S regulatory genes of the TOL plasmid pWW0. J. Gen. Microbiol. **132:**1347–1358.
236. **Steele, M. I., D. Lorenz, K. Hatter, A. Park, and J. R. Sokatch.** 1992. Characterization of the *mmsAB* operon of *Pseudomonas aeruginosa* PAO encoding methylmalonate-semialdehyde dehydrogenase and 3-hydroxybutyrate dehydrogenase. J. Biol. Chem. **267:**13585–13592.
237. **Stim-Herndon, K. P., T. M. Flores, and G. N. Bennett.** 1996. Molecular characterization of *adiY*, a regulatory gene which affects expression of the biodegradative acid-induced arginine decarboxylase gene (*adiA*) of *Escherichia coli*. Microbiology **142:**1311–1320.
238. **Stoner, C. M., and R. Schleif.** 1982. Is the amino acid but not the nucleotide sequence of the *Escherichia coli araC* gene conserved? J. Mol. Biol. **154:**649–652.
239. **Stoner, C. M., and R. Schleif.** 1983. The *araE* low affinity L-arabinose transport promoter. J. Mol. Biol. **170:**1049–1053.
240. **Sulavik, M. C., M. Dazer, and P. F. Miller.** 1996. SWISSPROT. Accession no. Q56070.
241. **Tate, C. G., J. A. R. Muiry, and P. J. F. Henderson.** 1992. Mapping, cloning, expression, and sequencing of the *rhaT* gene, which encodes a novel L-rhamnose-H^+ transport protein in *Salmonella typhimurium* and *Escherichia coli*. J. Biol. Chem. **267:**6923–6932.
242. **Tobe, T., M. Yoshikawa, T. Mizuno, and C. Sasakawa.** 1993. Transcriptional control of the invasion regulatory gene of *Shigella flexneri*: activation by VirF and repression by H-NS. J. Bacteriol. **175:**6142–6149.
243. **Tobin, J. F., and R. F. Schleif.** 1987. Positive regulation of the *Escherichia coli* L-rhamnose operon is mediated by products of tandemly repeated regulatory genes. J. Mol. Biol. **196:**789–799.
244. **Tobin, J. F., and R. F. Schleif.** 1990. Purification and properties of RhaR, the positive regulator of the L-rhamnose operons of *Escherichia coli*. J. Mol. Biol. **211:**75–89.
245. **Troup, B., M. Jahn, C. Hungerer, and D. Jahn.** 1994. Isolation of the *hemF* operon containing the gene for the *Escherichia coli* aerobic coproporphyrinogen III oxidase by in vivo complementation of a yeast HEM13 mutant. J. Bacteriol. **176:**673–680.
246. **Tsaneva, I. R., and B. Weiss.** 1990. *soxR*, a locus governing a superoxide response regulon in *Escherichia coli* K-12. J. Bacteriol. **172:**4197–4205.
247. **Volkert, M. R., L. I. Hajec, Z. Matijasevic, F. C. Fang, and R. Prince.** 1994. Induction of the *Escherichia coli aidB* gene under oxygen-limiting conditions requires a functional *rpoS* (*katF*) gene. J. Bacteriol. **176:**7638–7645.
248. **Wallace, R. G., N. Lee, and A. V. Fowler.** 1980. The *araC* gene of *Escherichia coli*: transcriptional and translational start-points and complete nucleotide sequence. Gene **12:**179–190.
249. **Wattiau, P., B. Bernier, P. Deslee, T. Michiels, and G. R. Cornelis.** 1994. Individual chaperones required for Yop secretion by *Yersinia*. Proc. Natl. Acad. Sci. USA **91:**1074–1078.
250. **Wattiau, P., and G. R. Cornelis.** 1994. Identification of DNA sequences recognized by VirF, the transcriptional activator of the *Yersinia yop* regulon. J. Bacteriol. **176:**3878–3884.
251. **Webster, C., K. Kempsell, I. Booth, and S. Busby.** 1987. Organisation of the regulatory region of the *Escherichia coli* melibiose operon. Gene **59:**253–263.
252. **Webster, C., L. Gardner, and S. Busby.** 1989. The *Escherichia coli melR* gene encodes a DNA-binding protein with affinity for specific sequences located in the melibiose-operon regulatory region. Gene **83:**207–213.
253. **Weickert, M. J., and S. Adhya.** 1992. A family of bacterial regulators homologous to Gal and Lac repressors. J. Biol. Chem. **267:**15869–15874.
254. **Wengelnik, K., and U. Bonas.** 1996. HrpXv, and AraC-type regulator, activates expression of five of the six loci in the *hrp* cluster of *Xanthomonas campestris* pv. *vesicatoria*. J. Bacteriol. **178:**3462–3469.
255. **Wilcox, G., and P. Meuris.** 1976. Stabilization and size of AraC protein. Mol. Gen. Genet. **145:**97–100.
256. **Wilcox, G., P. Meuris, R. Bass, and E. Englesberg.** 1974. Regulation of the L-arabinose operon BAD in vitro. J. Biol. Chem. **249:**2946–2952.
257. **Wintjens, R., and M. Rooman.** 1996. Structural classification of HTH DNA-binding domains and protein-DNA interaction modes. J. Mol. Biol. **262:**294–313.
258. **Wolf, M. K., and E. C. Boedeker.** 1990. Cloning of the genes for AF/R1 pili from rabbit enteroadherence *Escherichia coli* RDEC-1 and DNA sequence of the major structural subunit. Infect. Immun. **58:**1124–1128.
259. **Wong, C. M., M. J. Dilworth, and A. R. Glenn.** 1995. SPTREMBL. Accession no. Q52799.
260. **Wu, J., and B. Weiss.** 1991. Two divergently transcribed genes, *soxR* and *soxS*, control a superoxide response regulon of *Escherichia coli*. J. Bacteriol. **173:**2864–2871.
261. **Yahr, T. L., and D. W. Frank.** 1994. Transcriptional organization of the *trans*-regulatory locus which controls exoenzyme S synthesis in *Pseudomonas aeruginosa*. J. Bacteriol. **176:**3832–3838.
262. **Yamamoto, H., S. Uchiyama, A. N. Fajar, N. Ogasawara, and J. Sekiguchi.** 1996. Determination of a 12 kb nucleotide sequence around the 76 degrees region of the *Bacillus subtilis* chromosome. Microbiology **142:**1417–1421.
263. **Yamashita, M., H. Azakami, N. Yokoro, J. H. Roh, H. Suzuki, H. Kumagai, and Y. Murooka.** 1996. *maoB*, a gene that encodes a positive regulator of the monoamine oxidase gene (*maoA*) in *Escherichia coli*. J. Bacteriol. **178:**2941–2947.
264. **Yao, R., L. V. Reddy, and S. Palchaudhuri.** 1991. SWISSPROT. Accession no. Q04248.
265. **Yu, T. W., and C. W. Chen.** 1993. The unstable *melC* operon of *Streptomyces antibioticus* is codeleted with a Tn*4811*-homologous locus. J. Bacteriol. **175:**1847–1852.
266. **Yura, T., H. Mori, H. Nagai, T. Nagata, A. Ashihama, N. Fujita, K. Isono, K. Mizobuchi, and A. Nakata.** 1992. Systematic sequencing of the *Escherichia coli* genome: analysis of the 0–2.4 min region. Nucleic Acids Res. **20:**3305–3308.
267. **Zhou, L., K. N. Timmis, and J. L. Ramos.** 1990. Mutations leading to constitutive expression from the TOL plasmid *meta*-cleavage pathway operon are located at the C-terminal end of the positive regulator XylS protein. J. Bacteriol. **172:**3707–3710.