Construction of a Synthetic Gene for the Metalloregulatory Protein MerR and Analysis of Regionally Mutated Proteins for Transcriptional Regulation†

Kenneth M. Comess,*,‡ Lisa M. Shewchuk,‡ Kathryn Ivanetich,§ and Christopher T. Walsh*,‡

Department of Biological Chemistry and Molecular Pharmacology, Harvard Medical School, 240 Longwood Avenue, Boston, Massachusetts 02115, and Biomolecular Resource Center, University of California, San Francisco, California 94143

Received November 3, 1993; Revised Manuscript Received February 14, 1994®

ABSTRACT: The transcriptional control protein MerR is a metalloregulatory switch, activating transcription of a mercury resistance operon in the presence of mercuric ions and repressing transcription in their absence. We report here the construction and utilization of a synthetic *merR* gene and a single-copy *merT'–lacZ* fusion reporter for mutagenic analysis of the MerR protein's function. Site-directed mutagenesis of clustered acidic residues within the central region of the MerR protein indicated that these residues are important to the protein's ability to repress transcription. Quadruple or sextuple mutations involving residues E83 and E84 and other nearby acidic residues result in a repression-deficient (RD) phenotype. One of the mutant proteins was purified and shown by gel shift assay to retain binding to its operator DNA with an affinity similar to wild-type protein, suggesting that transcriptional repression does not correlate with MerR binding affinity. A small region of *merR* corresponding to residues 81–92 also was mutagenized in a search for other RD mutants and for mutants displaying sufficient transcriptional activation in the absence of mercuric ion to be classified as constitutive activation (CA) mutants. In this case, oligonucleotide-directed randomization of the target region and a screening/selection protocol were employed. Sixteen different mutants with an RD phenotype were identified, as well as eight different mutants with a CA phenotype. A high frequency of S87C mutations is evident in the RD set of mutants. The CA mutants have a high incidence of S86C and A89V mutations. The CA double mutant S86C/A89V was purified and found to bind to its DNA site with an affinity similar to that of the wild-type protein. Chemical nuclease activity assays indicate that the nonmercurated S86C/A89V CA mutant has a DNA distortion activity identical to that of mercurated wild-type MerR. A unique disulfide bond bridging this CA mutant's dimer interface was found and is proposed to constrain protein conformation in a manner analogous to mercuric ion binding in the wild-type protein.

A novel mechanism of transcriptional regulation is utilized in the procaryotic mercury resistance operon and appears to be conserved between Gram-positive and Gram-negative bacteria (Helmann et al., 1990b; Wright et al., 1990; Misra, 1992; Summers, 1992). The single transcript encoding the structural genes for mercury detoxification is under transcriptional control by the protein MerR. The promoter for this operon is poorly recognized by RNA polymerase (RNAP),[1] although the −35 and −10 σ^{70} promoter elements themselves are very well conserved (O'Halloran et al., 1989; Heltzel et al., 1990). An overlong spacing between the −35 and −10 elements apparently prevents transcription initiation, as nucleotide deletions in the spacer region of the *Escherichia coli* P_{merT} promoter lead to constitutive high-level *mer* gene expression (Parkhill & Brown, 1990). The regulatory protein MerR binds as a dimer between the −35 and −10 promoter elements (O'Halloran & Walsh, 1987). This binding reduces an already low level of expression severalfold. Apo MerR therefore is a repressor. Mechanistically, it acts by preventing isomerization of the closed complex to an open complex (Heltzel et al., 1990). Upon diffusion of a small amount of mercuric salts into a cell from a contaminated medium, MerR becomes a powerful transcriptional activator of the detoxification genes while remaining bound at the same site (O'Halloran et al., 1989). *E. coli* HB101 cells containing no mercury resistance determinants are killed by 50 μM $HgCl_2$, whereas cells containing a functionally intact *mer* operon grow well in media containing 250 μM $HgCl_2$ (Moore & Walsh, 1989). Nanomolar concentrations of Hg(II) are sufficient to induce high-level transcriptional activation by MerR *in vitro* (Ralston & O'Halloran, 1990). A structural model has been proposed for the homologous MerR protein from *Bacillus* species RC607 in which the protein dimer selectively binds a single mercuric ion, forming an asymmetric intersubunit bridge between the thiol groups of two cysteines on one subunit and a third cysteine on the other subunit (Helmann et al., 1990a). The Hg(II)–MerR complex elicits a localized underwinding of the DNA in the center of its binding site upon binding mercuric ion (Frantz & O'Halloran, 1990; Ansari et al., 1992). This change in helical twist realigns the flanking −35 and −10 promoter elements to a spatial orientation akin to that of a shorter, consensus length, spacer.

Random and site-directed mutagenesis of a variety of residues has been used in the absence of NMR or crystallographic structural data to identify specific regions of the MerR protein involved in its various functions. A mutation

† This work was supported in part by a grant from the National Institutes of Health (GM20011) and by an American Cancer Society Postdoctoral Fellowship to K.M.C.

* Authors to whom correspondence should be addressed.

‡ Harvard Medical School.

§ University of California.

® Abstract published in *Advance ACS Abstracts,* March 15, 1994.

[1] Abbreviations: β-gal, β-galactosidase; bp, base pair: *E. coli*, *Escherichia coli*; EDTA, ethylenediaminetetraacetic acid; IPTG, isopropyl β-D-thiogalactoside; PAGE, polyacrylamide gel electrophoresis; PCR, polymerase chain reaction; P_{merR}, promoter for regulatory transcript; P_{merT}, promoter for structural transcript; RNAP, *Escherichia coli* RNA polymerase; SDS, sodium dodecyl sulfate; TPEG, phenylethyl β-D-thiogalactoside; Xgal, 5-bromo-4-chloro-3-indolyl β-D-galactoside.

0006-2960/94/0433-4175$04.50/0

impairing DNA binding maps to a putative helix–turn–helix motif in the N-terminus (Ross et al., 1989; Shewchuk et al., 1989a). Mutations impairing transcriptional activation map throughout the distal two-thirds of the MerR protein, and include the three residues implicated in Hg(II) binding (Ross et al., 1989). These are cysteines at positions 82, 117, and 126 in the Tn*501* MerR protein. Lesions which allow for moderately Hg(II)-independent transcriptional activation, A89V[2] and S131L, have also been identified (Ross et al., 1989). When these mutations are combined to create the double mutant A89V/S131L, higher levels of transcription are observed (Parkhill et al., 1993). Interestingly, C82 has been shown to be positioned at the dimer interface (Shewchuk et al., 1989a,b,c).

The synthetic Tn*501 merR* gene reported here has greatly facilitated further molecular genetic analysis of MerR structure and function. Mutants have been constructed either by simple replacement of small restriction fragments with synthetic DNA duplexes carrying the desired mutation or by PCR, utilizing nearby upstream restriction sites. Mutations conferring a deficiency in some activity may indicate residues critical to that activity. Mutants which gain an activity normally present in the wild-type protein only when a coregulator or allosteric effector [e.g., Hg(II)] is bound may indicate the site of activity acted upon by allosteric interaction. The work described here focuses primarily on the latter type of mutants.

MATERIALS AND METHODS

Restriction endonucleases and the Klenow fragment of *E. coli* DNA polymerase I were from New England Biolabs (Beverly, MA). T4 DNA ligase was purchased from Boehringer Mannheim (Indianapolis, IN). T4 polynucleotide kinase and Sanger sequencing reagents were from United States Biochemical (Cleveland, OH). [α-^{35}S]dATP, [α-^{32}P]-dATP, and [α-^{32}P]TTP were from New England Nuclear (Boston, MA).

Oligonucleotide Synthesis. Tritylated oligonucleotides for the synthetic gene were synthesized at the UCSF Biomolecular Resource Center on an Applied Biosystems 380B DNA synthesizer (Foster City, CA) using cyanoethyl phosphoramidite chemistry. Initially, tritylated oligonucleotides were purified and detritylated on NENSORB PREP cartridges (Dupont, Boston, MA), and subsequently purified by denaturing PAGE using 8% acrylamide/8 M urea gels. DNA was eluted from gels with 0.3 M sodium acetate, extracted with phenol/chloroform/isoamyl alcohol (26:25:1), and precipitated with ethanol. Oligonucleotides for mutagenesis were synthesized at Harvard Medical School on an Applied Biosystems 380B DNA synthesizer, or purchased from Oligos Etc., Inc. (Wilsonville, OR). Degeneracy was introduced by making a mixture of all four phosphoramidite precursors and then adding a small percentage of this mixture to the correct precursor during individual synthesis steps.

Plasmid pSG Assembly and Cloning Strategy. The 432 bp synthetic *merR* gene was assembled and cloned in five sections, as shown in Figure 1. Purified oligonucleotides were phosphorylated with T4 polynucleotide kinase as described (Sambrook et al., 1989). Complementary oligonucleotides were annealed by heating to 70 °C for 5 min and then slowly cooling to room temperature. Annealed oligonucleotides for sections 1 (nucleotides 1–123), 2 (nucleotides 124–179), and 5 (nucleotides 370–463) were directly ligated into pUC18 using T4 DNA ligase to create pSG1, pSG2, and pSG5, respectively. Ligations were carried out overnight at room temperature with T4 DNA ligase as described by Sambrook et al. (1989). Section 3 (nucleotides 180–279) was ligated into pSG2 as a *Xma*I/*Sac*I fragment to give pSG23. Section 4 (nucleotides 280–369) was ligated into pSG5 as a *Sac*I/*Sph*I fragment to give pSG45. Sections 1, 2, and 3 were then combined by ligation of an *Eco*RI/*Xba*I fragment from pSG1 into pSG23, yielding pSG123. The gene was completed by the ligation of an *Eco*RI/*Sac*I fragment from pSG123 into pSG45. The complete product, containing the synthetic *merR* gene cloned into the *Eco*RI and *Hin*dIII sites of pUC18, is designated pSG. All ligation mixtures were transformed into competent *E. coli* DH5α [*endA*1, *recA*1, *hsdR*17 ($r_k^-m_k^+$), *supE*44, *thi*-1, λ-, *gyrA*96, *relA*1, Δ(*lacZYA-argF*)*U169*] cells according to standard procedures (Sambrook et al., 1989). Colonies were screened for inserts of the appropriate size and restriction fragment patterns. The sequence of the cloned gene was confirmed by Sanger sequencing (Sanger et al., 1977).

A 747 bp *Pvu*II fragment from pSG containing the *merR* gene and the P_{lac} promoter was ligated to a 3000 bp *Bam*HI/*Pst*I fragment from pGP1–2 (Tabor & Richardson, 1985) to yield plasmid pSGK. Plasmid pSGK contains a P15A origin and a kanamycin resistance marker.

Construction of Reporter Strain. We placed a single-copy *mer* reporter into the *E. coli* chromosome by using a vector, pDHB3201, constructed and provided for that purpose by N. Kuldell and A. Hochschild (unpublished results). Plasmid pDHB3201 contains the *malF* gene from *E. coli* and a tetracycline resistance gene placed into the *Bss*HII site within the *malF* gene (Figure 4). It is a derivative of plasmid pDHB32 (Boyd et al., 1987). A unique *Sac*I site is within the *malF* gene, upstream of the tetracycline resistance gene. The tetracycline gene in pDHB3201 is oriented in the same direction as the *malF* gene. The P_{merT} operator/promoter–*lacZ* transcriptional fusion from pWR2 (Ross et al., 1989) was excised with *Eag*I and *Tth*111I, the ends were made blunt by treatment with the Klenow enzyme, and the fragment was isolated from an agarose gel after electrophoretic separation. This fragment includes a nonfunctional truncation of the *merR* gene encoding the first 16 residues only. Plasmid pDHB3201 was linearized with *Sac*I, the ends were made blunt with Klenow fragment, and the pWR2 fragment was allowed to ligate into this site. Clones were characterized in XL1 blue cells [*supE*44 *hsd*R17 *rec*A1 *end*A1 *gyr*A46 *thi rel*A1 *lac*$^-$ F′ (*proAB*$^+$ *lacI*^{q} *lacZ*ΔM15 Tn*10*)] expressing *merR* on the compatible pSGK plasmid. A clone (designated pKMC4) which demonstrated repression and activation of *lacZ* activity in the absence and presence of Hg(II), respectively, was found by restriction enzyme mapping to have the insert oriented such that *malF* transcription is opposite to *lacZ* transcription.

Plasmid pKMC4 was linearized with *Pst*I and *Sca*I and electro-transformed into an MC1000 *recD* derivative [F$^-$ *araD139* Δ(*araABC-leu*)*7679 recD galU galK* Δ(*lac*)X74 *rpsL thi*]. Colonies were selected/screened on maltose MacConkey plates (Silhavy et al., 1984) containing 2.5 μg/mL tetracycline. Maltose-negative (white) colonies were purified on maltose MacConkey plates containing 5 μg/mL tetracycline. After being screened for ampicillin sensitivity, the strain was transformed with pSG to check for reporter activity. Finally, P1 transduction (Silhavy et al., 1984) was

[2] Single-letter amino acid codes: A, alanine; C, cysteine; D, aspartate; E, glutamate; F, phenylalanine; G, glycine; H, histidine; I, isoleucine; K, lysine; L, leucine; M, methionine; N, asparagine; P, proline; Q, glutamine; R, arginine; S, serine; T, threonine; V, valine; W, tryptophan; Y, tyrosine.

employed to transfer the reporter into strain TG1 [*supE hsdΔ5 thi* Δ(*lac-proAB*) F′ (*traD36 proAB*$^+$ *lacI*^{q} *lacZ*ΔM15)]. The transductant, TG104, was recovered by screening for β-galactosidase activity on Xgal agar after an initial selection on tetracycline agar.

β-Galactosidase Activity Assays. Transcriptional regulation of the P_{merT} promoter by the mutant and wild-type proteins was monitored *in vivo* by assaying *lacZ* expression either from a *Eco*RI–*Eag*I *merR*-deleted derivative of pLS-ROpZ (Shewchuk et al., 1989b), designated pLS-OpZ, or from the single-copy reporter host TG104. Overnight cultures grown in M63 salts (Silhavy et al., 1984) supplemented with thiamin, magnesium sulfate, 0.2% glucose, and appropriate antibiotics were subcultured 1:100 into identical medium containing 1 mM IPTG. After growth to an OD_{595} of ca. 0.4, cultures were split, and one portion was treated with 1.5 μM $HgCl_2$. These induced and uninduced cultures were grown for an additional 30 min and then rapidly chilled on ice. β-Galactosidase assays using *o*-nitrophenyl β-D-galactopyranoside as substrate were performed as described by Miller (1972), by treating cells with SDS and chloroform. In either reporter system, the standard deviations derived from triplicate analyses were 5% for values obtained from uninduced (no mercuric ion added) cultures and 15% for induced samples. As reported previously (Ross et al., 1989), variation in activation levels was observed on different days for the wild-type and mutant strains, although the data for mutant proteins maintain consistent relationships to the wild-type uninduced and induced values. Miller assay data are therefore reported relative to wild-type values in Tables 1–3 and in Figure 5.

Site-Directed Mutagenesis. Paired acidic residues within proposed helical regions of Tn*501* MerR (D68E69, E77D78, and E83E84) were mutated to the corresponding glutamine and asparagine residues by cassette mutagenesis. Sets of complementary oligonucleotides were synthesized such that when annealed they would contain the 5′- or 3′-overhanging ends corresponding to the *Bss*HII and *Bsu*36I sites (D68N/E69Q mutagenesis), *Bsu*36I and *Kpn*I sites (E77Q/D78N mutagenesis), or *Kpn*I and *Sac*I sites (E83Q/E84Q mutagenesis). For example, the E77Q/D78N mutagenesis cassette was comprised of the following oligonucleotide pair (nucleotide substitutions are underlined): 5′-d(TGAGGCTGCAGAACGGTAC)·5′-d(CGTTCTGCAGCC). The complementary oligonucleotides were phosphorylated, annealed, and ligated into digested pSGK as described above. Mutations were confirmed by Sanger sequencing (Sanger et al., 1977). Quadruple and sextuple mutants were constructed by exchanging the appropriate sections of the double mutants.

Mutagenesis of the other residues, C82, S86, S87, A89, C117, and C126, was carried out by PCR, using as an upstream primer an oligonucleotide containing the codon change, a unique restriction site 5′ to this codon, and a d(GCGC) sequence 5′ to the restriction site. The technique is analogous to that described below for degenerate oligonucleotide cassette mutagenesis and shown in Figure 4B. For example, the S86C PCR mutagenesis primer was as follows (nucleotide substitutions are underlined): 5′-d(GCGCGGTACCCACTGCGAAGAAGCGTGTTCTCTGGCTGAACACAA).

Degenerate Oligonucleotide Cassette Mutagenesis and Screening. A small region of MerR comprising resiudes 81–92 was partially randomized by using a degenerate oligonucleotide as one primer for PCR amplification of codons 79–144. The 61-mer oligonucleotide for cassette mutagenesis (Figure 4B) is comprised of an upstream *Kpn*I site, several additional upstream nucleotides [d(CGCG)] to facilitate fragment *Kpn*I cleavage, an internal stretch of 36 partially degenerate nucleotides corresponding to residues H81–K92, and a downstream stretch of 15 nucleotides of exact complementarity to the *merR* gene. The internal randomized region was doped to a level of 10% per base with a mixture of the other nucleotide precursors so that the average number of incorrect nucleotides per molecule is (0.1 × 36) = 3.6. The actual number of mutations per molecule is described by a Poisson distribution. The internal region also was synthesized to contain the d(TCT) codon for S86 rather than d(AGC). This modification destroys a unique *Sac*I site and allows randomized genomes to be enriched over residual wild type by *Sac*I incubation prior to transformation.

Second-strand synthesis and amplification of the degenerate oligonucleotide were carried out by using a 17-mer M13(-20) sequencing primer (New England Biolabs) as a downstream partner for 20 cycles of PCR. Thus, the resultant fragment and pSG template both contain *Kpn*I and *Hin*dIII sites (Figures 1 and 4B). The mutagenic library was constructed in one pot: a portion of the *merR* gene was amplified by using the degenerate oligonucleotide and M13 primer, the pSG template and 236 base pair amplified fragment were digested with *Kpn*I and *Hin*dIII, the digested fragment (now 208 bp) was allowed to ligate into the pSG backbone, and wild-type pSG, generated via religation of the liberated *Kpn*I/*Hin*dIII pSG fragment, was linearized by digestion with *Sac*I. Ligation of the mutagenized fragment is expected to predominate due to its amplified level, but any wild-type pSG regenerated will have a reduced transformation efficiency due to its linearization. Aside from its ease of use, this technique avoids the technical difficulties of mutually primed synthesis (Sambrook et al., 1989), as well as the need for two adjacent and unique restriction sites for cloning. Ligations were carried out with T4 DNA ligase after extraction with phenol/chloroform/isoamyl alcohol and precipitation by sodium acetate and ethanol (Sambrook et al., 1989). The plasmid library was transformed into TG104 and plated on minimal 2% agar plates (1× M63 salts) supplemented with $MgSO_4$, thiamin, and 0.5% lactose, and containing kanamycin (25 μg/mL), IPTG (1 mM), and Xgal (75 μL of 50 mM spread on top ca. 1 h before use) to give ca. 150 colonies per plate. Colonies which were larger and bluer were transferred to Hg(II)-containing indicator plates for secondary screening. This step was carried out on plates identical to those above, except containing 0.1% rather than 0.5% lactose and also including 2 mM phenylethyl β-D-thiogalactoside (TPEG, a competitive inhibitor of β-galactosidase) and 1.5 μM $HgCl_2$. The agar conditions were determined empirically to select for growth and screen for blue color. Approximately 5000 colonies were selected and screened, and 43 colonies that expressed relatively high levels of β-galactosidase in the presence and absence of $HgCl_2$ were assayed directly for β-galactosidase activity.

Coexpression of Two Plasmids. For experiments requiring transformation of two plasmids into the same host cell, transformations were done sequentially. The relative copy number of the two plasmids pSG and pSGK was determined by restriction endonuclease digestion and gel electrophoresis, as the pSG-derived and pSGK plasmids are of different size. On the basis of ethidium bromide-stained agarose gels, the two plasmids were present at an approximate ratio of (4–6):1 for pSG-derived to pSGK plasmids.

Subcloning for Overexpression and Purification of Wild-Type and Mutant Proteins. The wild-type synthetic gene from pSG as well as consensus RD and CA mutants was

subcloned into vector pET-22b(+) (Novagen; Madison, WI) via the *Nde*I and *Hin*dIII restriction sites (Figure 1). Ampicillin-resistant colonies were screened by examining whole cell extracts on SDS–PAGE gels under reducing conditions (Sambrook et al., 1989) for IPTG-dependent expression of a 16-kDa protein, and their plasmids were analyzed by restriction mapping. Colonies were screened and expressed in host BL21(DE3)pLysS (Novagen). Overexpression constructs thus identified for the wild-type, consensus RD mutant, and consensus CA mutant proteins are designated pKMC60, pKMC61, and pKMC62, respectively.

Growth and purification of the wild-type and RD mutant proteins were carried out as described previously (Helmann et al., 1989; Shewchuk et al., 1989a), with several modifications. After lysis of 1 L of cells in low-salt buffer and extraction of MerR from the cell debris with high-salt buffer (O'Halloran & Walsh, 1987), the protein was selectively precipitated with ammonium sulfate. A 15% ammonium sulfate cut removed some UV-absorbing material without precipitating MerR. A subsequent 25% cut precipitated MerR while leaving much of the DNA behind. After reconstitution in a buffer containing 0.12 M NaCl, 2 mM EDTA, 50 mM Tris·HCl (pH 7.4), 5% glycerol, and 10 mM 2-mercaptoethanol, the protein was slurried with DE-52 anion-exchange resin (Whatman). The resin selectively binds DNA at this salt concentration. Heparin–Sepharose CL-6B (Pharmacia) chromatography was then carried out on the supernatant as before (Helmann et al., 1989; Shewchuk et al., 1989a). Approximately 9 mg of purified protein was obtained per liter of cells.

The construct pKMC62 was grown with shaking in 1-L Luria broth cultures at 37 °C to an A_{595} of 0.5 and then induced with 375 μM IPTG and moved to a 30 °C shaking incubator for 4 h of induction. Purification was carried out as for the other proteins.

Purification and *in vitro* characterization of the E77Q/D78N/E83Q/E84Q mutant protein were as described for the wild-type protein (Helmann et al., 1989; Shewchuk et al., 1989a).

Gel Shift Assay. Purified mutant and wild-type proteins were used in gel shift assays to compare directly affinities for the P_{merT} binding site. At a fixed concentration of 10^{-12} M DNA, well below the K_D of ca. 1×10^{-10} M, the protein concentration was varied to determine the protein concentration at which half of the DNA was bound. This concentration is the apparent K_D for protein operator interaction. Labeled DNA containing the P_{merT} site was obtained in several steps from plasmid pMerOp (Shewchuk et al., 1989a). Digestion by *Eco*RI and filling-in with Klenow fragment using [α-^{32}P]dATP and [α-^{32}P]TTP were followed by precipitation and digestion with *Hin*dIII. The 305 base pair end-labeled fragment then was isolated by agarose gel electrophoresis and eluted from the gel slice by using a Gene Clean II kit (Bio 101; LaJolla, CA). This fragment is labeled at one end only. DNA concentrations were determined by the fluorometric method of Le Pecq and Paoletti (1966). Assays were run as previously described (Shewchuk et al., 1989c), except that protein–DNA complexes were allowed to equilibrate for 30 min at 4 °C (Helmann et al., 1989) and 6.75% polyacrylamide gels were electrophoresed in Tris–acetate/EDTA buffer as described by O'Halloran et al. (1989). Quantitation of bound and free DNA was carried out by exposing dried gels to storage phosphor screens and analyzing the data on a Molecular Dynamics Phosphorimager.

Bis(5-phenyl-1,10-phenanthroline)copper(I) Hypersensitivity Assay. The 305 base pair DNA fragment prepared for gel shift assays also was used for this assay. The labeled fragment was incubated at a concentration of 3 nM with purified protein (at 30 nM) in 10 mM Tris·HCl (pH 8), 50 mM KCl, 5% glycerol, 0.05% NP40, 100 μM EDTA, 50 μg/mL BSA, 10 μg/mL sonicated salmon sperm DNA, and 1 mM 3-mercaptopropionic acid. Binding was allowed to proceed on ice for 30 min. The cleavage reagent was added by first mixing 10 μL of 32 mM 5-phenyl-1,10-phenanthroline (Sigma) (dissolved in ethanol) with 10 μL of 7.2 mM aqueous cupric sulfate (Baker), immediately adding 80 μL of water, and then adding 1 μL of this complex to the protein/DNA incubation mixture (final volume of 20 μL). The reaction was initiated by addition of 1.2 μL of 100 mM 3-mercaptopropionic acid and allowed to proceed at 37 °C for 2 min; 1.3 μL of 50 mM 2,9-dimethyl-1,10-phenanthroline was then added to quench the reaction, and the DNA was precipitated by addition of sodium acetate, yeast tRNA, and ethanol according to standard techniques. Samples were reconstituted in formamide loading dye (USB) and separated on 6% denaturing polyacrylamide sequencing gels. A DNA sequence ladder was generated by Maxam–Gilbert guanine-specific cleavage of the same DNA fragment (Maxam & Gilbert, 1980).

RESULTS

Synthetic Tn501 merR Gene. A synthetic gene based on the amino acid sequence of Tn*501* MerR was constructed by synthesis of five pairs of oligonucleotides, lengths 122, 56, 100, 91, and 93 bp, and sequential cloning as depicted in Figure 1. The full-length gene was sequenced and encodes the wild-type Tn*501* MerR protein. This synthetic *merR* gene facilitates the rapid generation of directed mutations by inclusion of 25 unique restriction sites within the *merR* coding sequence (Figure 1). The *in vivo* and *in vitro* properties of the synthetic MerR protein are indistinguishable from wild-type MerR. The synthetic protein represses and activates transcription of the Tn*501 mer* promoter in the absence and in the presence of Hg(II), respectively. *In vitro*, the protein binds its DNA operator with an apparent K_D of 9×10^{-11} and 1×10^{-10} M in the absence and presence of Hg(II), respectively (Figure 2), and ligates a single mercuric ion in the presence of a 1000-fold excess of buffer thiols (data not shown).

Mutagenesis of Paired Acidic Residues. The synthetic Tn*501 merR* gene proved useful in assessing possible roles of acidic residues in the central region of the MerR protein. Pairs of acidic residues (D68E69, E77D78, and E83E84) were mutated to the corresponding glutamine or asparagine by cassette mutagenesis as described under Materials and Methods. Transcriptional activation and repression by the mutant proteins were monitored *in vivo* at the P_{merT} promoter residing on the multicopy pLS-OpZ reporter plasmid. The three pairs of acidic residues were changed singly and in combination. Mutation of either the pair at 68 and 69 or the pair at 77 and 78, or both pairs together, had little or no effect on activation or repression (rows 2, 3, and 5, Table 1). In contrast, mutation of the pair at 83 and 84 results in a slight RD phenotype (row 4, Table 1). Additional mutation of either or both of the other two pairs of acidic residues resulted in a protein even more deficient in repression but still retaining wild-type activation function (rows 6, 7, and 8, Table 1).

The E77Q/D78N/E83Q/E84Q quadruple mutant MerR was purified and assayed *in vitro* for operator DNA binding. Essentially wild-type binding was observed (Figure 3).

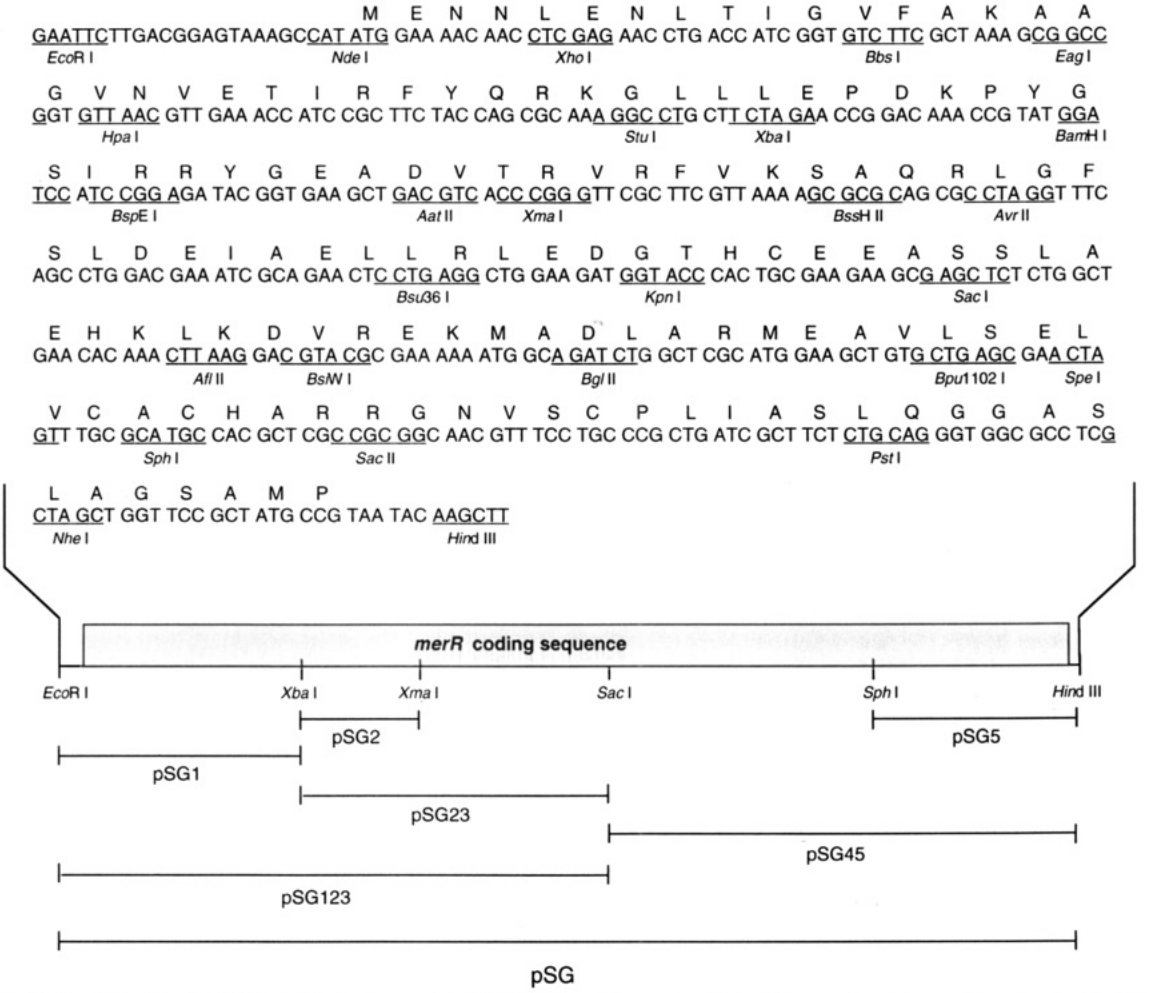


FIGURE 1: DNA sequence and cloning strategy for the synthetic Tn*501 merR* gene. Details of the construction are described under Materials and Methods. The amino acid sequence is shown above the DNA sequence. Unique restriction sites are also indicated.

Construction and Screening of a Regionally Randomized MerR Library. In order to reduce the copy number of the transcriptional activity reporter, a *merT'–lacZ* transcriptional fusion was placed into the *E. coli* chromosome in single copy. By expressing the *merR* gene on a multicopy plasmid, the ratio of DNA-binding protein to binding sites is enhanced, and the background of readthrough due to unoccupied operator sites is decreased. An enhancement in the ability to discriminate between repressed and basal-level activation therefore should be elicited. The *mer* reporter from plasmid pWR2 (Ross et al., 1989) was cloned into a shuttle vector (Boyd et al., 1987; N. Kuldell and A. Hochschild, unpublished results) for integration into the *E. coli* chromosome (Figure 4A). The resultant vector, pKMC4, contains two disruptions of the nonessential *E. coli* gene *malF*. A double homologous recombination event, in which the ends of the *malF* gene recombine with the host chromosome, will disrupt the host *malF* gene and confer tetracycline resistance and MerR-regulated *lacZ* expression on the *E. coli* recipient strain. The reporter is integrated into a known site, and the rest of the plasmid, including the origin of replication and the β-lactamase gene, is degraded or otherwise lost from the cell. Transformations of linearized plasmid pKMC4 were carried out by electroporation. Putative integrants selected on tetracycline agar were screened for loss of maltose utilization on maltose MacConkey plates and for appropriate integration of the interrupted *malF* gene on ampicillin-containing plates. Finally, the ability to distinguish repression, activation, and basal-level expression phenotypes on plates was optimized in the presence of the plasmid pSG, in the presence of pSG + $HgCl_2$, and in the absence of plasmid, respectively. The reporter strain is designated TG104.

A library of MerR mutant plasmids was constructed by spiked oligonucleotide mutagenesis for screening in strain TG104 (Figure 4B). The mutagenized region, spanning the codons for residues H81 to K92, was chosen for its apparent importance in mediating activation and repression (see Discussion). A primary screen was carried out on minimal media plates without $HgCl_2$, supplemented with IPTG, Xgal, and lactose as the sole carbon source. Approximately 1% of the colonies were large and blue, indicating increased levels of β-galactosidase synthesis. A secondary screen was carried out to distinguish colonies capable of high-level transcriptional activation from those only capable of basal-level transcription. The secondary screen allowed discrimination of nonfolded or truncated mutants from functional repression-deficient (RD) or constitutive activation (CA) mutants.

Forty-three mutant candidates that expressed elevated levels of β-galactosidase in the absence of $HgCl_2$ and approximately wild-type levels in the presence of $HgCl_2$ were analyzed in liquid culture. Mutants found to have a transcription level upon induction with $HgCl_2$ less than half of wild type were rejected. The remaining mutants were found to cluster into two distinct classes. The degree of induced activity varied with different experiments, but values for the mutants relative to wild type were invariant within each experiment. The best discriminator for dividing the RD and CA classes was the relative induction ratio. Quantitatively, the RD mutant

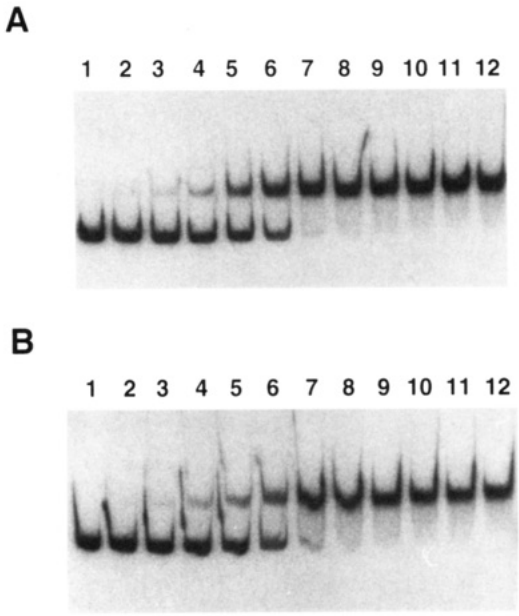


FIGURE 2: Concentration dependence of P_{merT} binding by synthetic gene-derived MerR. An autoradiogram of a typical gel binding assay in either the absence (A) or the presence (B) of Hg(II), performed as described under Materials and Methods, is shown. The lower band corresponds to free DNA while the upper band corresponds to a stoichiometric MerR–DNA complex. The protein concentration in each lane is as follows: (1) no protein, (2) 7×10^{-12} M, (3) 1×10^{-11} M, (4) 4×10^{-11} M, (5) 7×10^{-11} M, (6) 1×10^{-10} M, (7) 4×10^{-10} M, (8) 7×10^{-10} M, (9) 1×10^{-9} M, (10) 4×10^{-9} M, (11) 7×10^{-9} M, and (12) 1×10^{-8} M. The K_D determined from this experiment is (8–9) $\times 10^{-11}$ and (1–1.1) $\times 10^{-10}$ M in the absence and presence of Hg(II), respectively.

Table 1: Regulatory Abilities of Tn*501* Acidic Residue Mutants[a]

merR mutant	activity ratio, uninduced[b]	Hg(II) induction ratio[c]
wild type	1.0	166
D68N/E69Q	1.0	115
E77Q/D78N	0.9	176
E83Q/E84Q	6.7	23
D68N/E69Q/E77Q/D78N	0.9	149
D68N/E69Q/E83Q/E84Q	27	4.5
E77Q/D78N/E83Q/E84Q	21	5.7
D68N/E69Q/E77Q/D78N/ E83Q/E84Q	15	8.3

[a] Assayed by using the multicopy reporter pLS-OpZ. [b] Ratio of β-galactosidase activity of uninduced mutant (no $HgCl_2$) to uninduced wild-type MerR. [c] Ratio of β-galactosidase activity of $HgCl_2$-induced mutant to uninduced mutant MerR.

proteins were chosen as those exhibiting an induction ratio greater than 2.0 (but less than 60% of wild type) and CA mutants as those with a ratio less than 2.0. Sixteen unique members of the repression-deficient class were found and have 3–23-fold greater β-galactosidase activity than wild type in the absence of mercury and induction ratios ranging from 2.4-fold to 14-fold. Eight constitutive class mutants had 22–79-fold greater activity in the absence of mercury than wild type and induction levels ranging from 0.8-fold to 1.9-fold. The induced levels of activity were between 0.5 and 1.7 times the activity of induced wild type. Nineteen other mutants did not meet these criteria and were not examined further.

Mutants from both selected classes were sequenced, as well as several random members of the parent library (Figure 5). A high frequency of substitutions at position S87 in the RD class and positions S86 and A89 in the CA class is evident in Figure 5A, relative both to other positions in the same class and between classes. Substitutions at these positions were

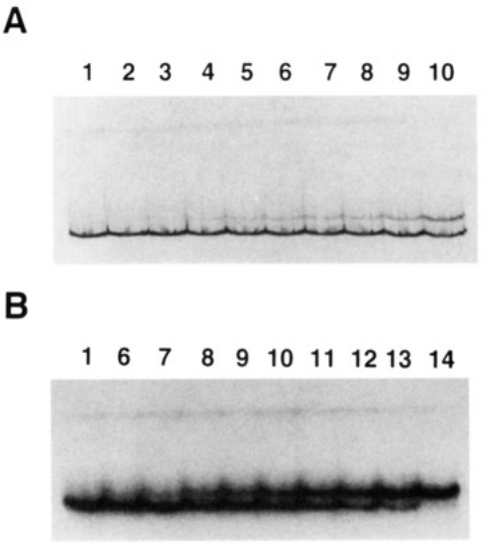


FIGURE 3: Concentration dependence of E77Q/D78N/E83Q/E84Q protein–operator DNA binding in the absence (A) and presence (B) of Hg(II). The protein concentration in each lane is as follows: (1) no protein, (2) 2×10^{-11} M, (3) 3×10^{-11} M, (4) 4×10^{-11} M, (5) 5×10^{-11} M, (6) 6×10^{-11} M, (7) 7×10^{-11} M, (8) 8×10^{-11} M, (9) 9×10^{-11} M, (10) 1×10^{-10} M, (11) 2×10^{-10} M, (12) 3×10^{-10} M, (13) 4×10^{-10} M, and (14) 5×10^{-10} M. The K_D determined from this experiment is $\sim 10^{-10}$ M in both the presence and absence of Hg(II).

nonrandom. For the RD class, 69% of the unique mutants had a mutation at position S87 (11 of 16 unique mutants), and all of these were S87C substitutions (Figure 5B). For the CA class, 75% of the unique mutants had mutations at positions S86 and A89 (6 of 8 unique CA mutants), and all of these were S86C and A89V mutations (Figure 5C). Although mutations were found at these positions in the parent library, none of them were S86C, S87C, or A89V. The sequenced mutations found in this library were S86S (silent mutation), S87A, S87stop, and A89S. Although other substitutions were present in every mutant, the frequency and identity of mutations at positions 86, 87, and 89 in the RD and CA phenotype-selected libraries suggest that these may be good solutions for generating the RD and CA phenotypes.

Consensus RD and CA Mutants. Since none of the mutants examined contained a single S87C mutation or a double S86C/A89V mutation in the absence of other mutations, these mutants were prepared by site-directed mutagenesis. A PCR technique analogous to that employed for degenerate oligonucleotide library construction (Figure 4B) was employed. This technique resulted in mutagenesis efficiencies greater than 85%. Miller assay analysis of the consensus RD mutant (S87C) and consensus CA mutant (S86C/A89V), along with "tester" mutants S86C, A89V, and S87C/A89V, demonstrated that the S87C mutation alone is sufficient for the RD phenotype and that the S86C/A89V double mutation is necessary and sufficient for constitutive activation (Table 2). The single mutants S86C and A89V do not fit the criteria described above for either RD or CA mutants. The S87C/A89V mutant also fit CA class criteria, but the uninduced and induced levels of activity were higher for the S86C/A89V mutant. Under inducing conditions, the consensus CA mutant S86C/A89V yields almost twice as much β-gal activity as the induced wild-type MerR protein. It is both a constitutive activator and a mercury-inducible superactivator.

In Vivo Characterization of RD and CA Consensus Mutant Proteins. Competition in transcriptional regulation between the mutant and wild-type proteins was assessed *in vivo* by

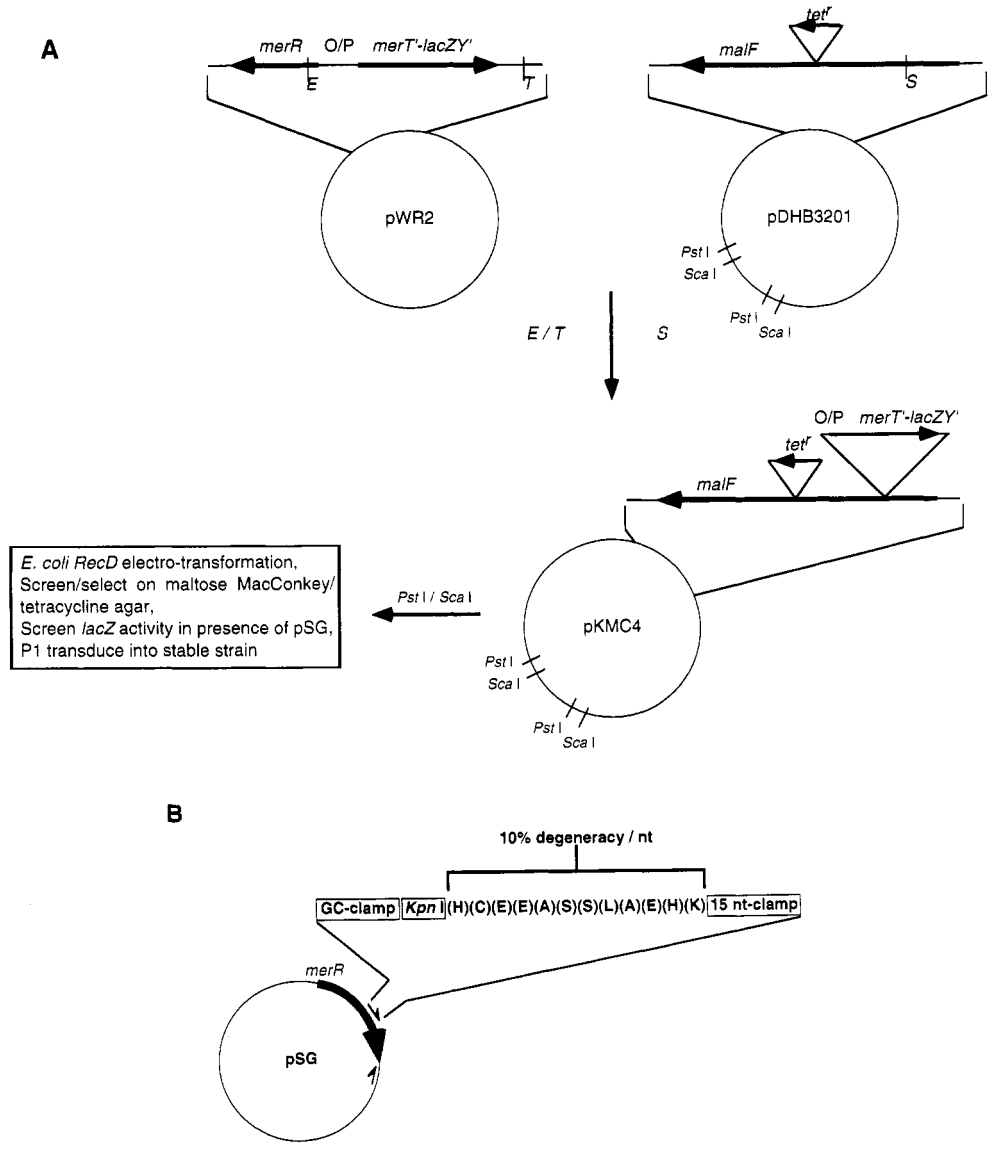


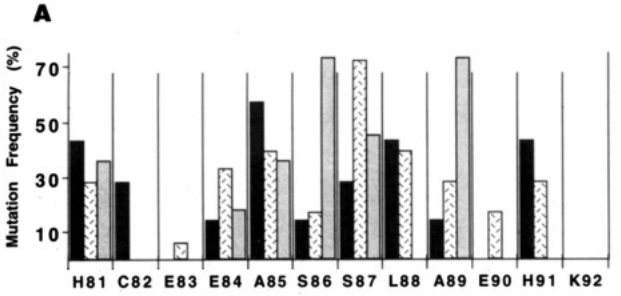
FIGURE 4: Construction of host-integrated reporter strain (A) and strategy for regional randomization of residues H81 to K92 of MerR (B). In (B), the mutagenic oligonucleotide primer is shown schematically, with randomized nucleotide codons represented by the wild-type amino acid sequence. The other primer is a standard 17-mer (–20) M13 sequencing primer. O/P, P_{merT} operator/promoter region; nt, nucleotide; *E*, *Eag*I; *T*, *Tth*111I; *S*, *Sac*I.

introducing plasmids expressing both proteins into individual TG104 cells containing the chromosomally encoded reporter. Wild-type and mutant alleles were expressed from identical P_{lac} promoters. The mutant *merR* genes were placed on high-copy plasmids, however, and the wild-type gene was expressed on a compatible low-copy plasmid. In Miller assays, the level of constitutive expression elicited by the consensus CA mutant is reduced considerably (ca. 5-fold) in the presence of wild-type MerR (data not shown).

We also wished to ascertain the effects *in vivo* of removing each of the three conserved cysteine residues normally involved in coordinating Hg(II). The triple mutants S86C/A89V/C82A, S86C/A89V/C117A, and S86C/A89V/C126A were constructed by oligonucleotide-directed PCR cassette mutagenesis as above, and examined by using the Miller assay (Table 3). Some activity was retained in the C82A and C117A triple mutants, in contrast to the C126A triple mutant and the single mutants C82A, C117A, or C126S described previously (Shewchuk et al., 1989b). Interestingly, Hg(II) induction also occurs for the former triple mutants (S86C/A89V/C82A and S86C/A89V/C117A), indicating altered metal ligation from the wild-type coordination environment.

The stability of these mutants *in vivo* was confirmed by assaying repression of the P_{merR} promoter. β-Galactosidase activity of a single-copy P_{merR}–*lacZ* fusion reporter is repressed strongly in the presence of wild-type MerR, the consensus CA mutant, or any of the three constructed triple mutants (B. T. Ballard and C. T. Walsh, unpublished results).

In Vitro Characterization of the Consensus RD and CA Mutants. The unique restriction sites of the synthetic *merR* gene were exploited for rapid subcloning of the wild-type and consensus mutant genes into overexpression vectors. An *Nde*I site at the methionine start codon (Figure 1) allowed cloning in one step into an optimally aligned and regulated translational start site (see Materials and Methods). Figure 6 demonstrates that expression of all three proteins is tightly repressed in the absence of IPTG induction and induced to similarly high levels in the presence of IPTG. Induction with 1.2 mM IPTG at early to mid-log growth phase and subsequent growth at 37 °C work well for the wild-type and RD (S87C) mutants, but not for the CA (S86C/A89V) mutant. Under these conditions, the CA mutant broth was found as >90% insoluble aggregated material. Shifting the growth cultures from 37 to 30 °C at the point of IPTG induction and reducing the amount of IPTG



B

81 ... 92	β-Gal Activity, uninduced	Induction Ratio
----H-C-E-E-A-S-S-L-A-E-H-K----		
----------V-S---C----------	18	2.4
----------C-G----------	23	2.6
----------V-S---C-------Q------	15	3.0
----------Y-D-Q------	16	3.0
----------C-P-----Q------	18	3.2
----D-----V-----C-R----------	11	3.8
----R-----G---------V-K--------	15	3.9
----Q-------V-------V----------	11	4.0
----------G---C-----G--------	3.0	4.1
----------C-R----------	14	4.0
----------C-C-A----------	8.0	4.2
----------C---C-V-----Y------	11	4.3
----R----------C---T----------	16	4.4
----------G---C-------N------	10	5.7
--------D-A-V-C----------	6.0	7.6
----Q-----G---C-----G----------	3.0	14

C

----------G---C---V----------	22	1.3
----------E---C----------	25	1.3
----------G-C-----V----------	29	1.1
----------C-A---V----------	36	1.4
----------D---C-----V----------	50	1.9
----R---------C-----V----------	51	1.4
----R-------Q-C-----V----------	59	0.8
----------G---C-----V----------	79	1.1

FIGURE 5: Results of mutant screening protocol. (A) Positional mutation frequencies for the parent library (black) and for the RD (mottled) and CA (light gray) phenotype subclasses. The high frequency of substitutions at position S87 in the RD class and at positions S86 and A89 in the CA class is clearly visible. (B) Transcriptional activity and sequence identity of RD class mutants. (C) Transcriptional activity and sequence identity of CA class mutants. All mutants display activities under inducing conditions that are at least half that of induced wild type. Assays were conducted on the single-copy reporter in TG104. An induction ratio of 51 was found for the wild-type MerR protein. β-Gal activity and induction ratio, see Table 1 for definitions.

Table 2: β-Galactosidase Activity of Selected Single and Double MerR Mutants[a]

merR mutant	activity ratio, uninduced[b]	Hg(II) induction ratio
wild type	1.0	14
S86C	1.2	12
S87C	1.7	7.9
A89V	1.4	4.5
S86C/A89V	17	1.5
S87C/A89V	8.8	1.7

[a] Assayed by using the single-copy reporter strain TG104. [b] See Table 1 footnotes for definitions.

by a factor of 3 obviated this problem, presumably by slowing the production of protein sufficiently to allow proper folding within the cell. All three proteins were purified by identical protocols, employing heparin–Sepharose chromatography in the final step (Helmann et al., 1989; Shewchuk et al., 1989a,c). Each of the expression vectors pKMC60, pKMC61, and pKMC62 produced 8–9 mg of pure protein from 1 L of culture.

Although all purification steps were carried out in the presence of 5–10 mM 2-mercaptoethanol, analysis on non-reducing SDS–polyacrylamide gels revealed the freshly purified CA protein to be present as a disulfide-linked dimer

Table 3: β-Galactosidase Activity of Triple Mutants

merR mutant	activity ratio, uninduced[a]	Hg(II) induction ratio
wild type[b]	1.0	36.8
S86C/A89V[b]	31.4	1.90
C82A[c]	0.71	0.91
S86C/A89V/C82A[b]	3.17	3.12
C117A[c]	0.84	0.85
S86C/A89V/C117A[b]	2.11	10.6
C126S[c]	1.36	0.86
S86C/A89V/C126A[b]	1.53	0.93

[a] See Table 1 footnotes for definitions. [b] Assayed by using the single-copy reporter strain TG104. [c] Assayed by using the multicopy reporter pLS-OpZ.

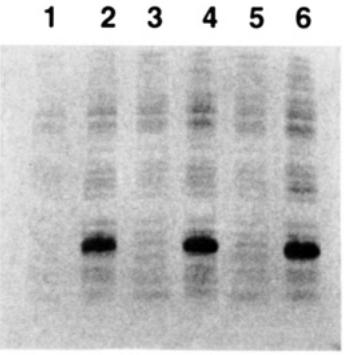


FIGURE 6: Amounts of wild-type and RD and CA consensus mutant proteins in crude extracts. Lanes 1–6 show proteins resolved by denaturing 20% SDS–PAGE and stained with Coomassie brilliant blue. (Lanes 1, 3, 5) Uninduced extracts; (lanes 2, 4, 6) extracts induced with 1 mM IPTG for 2 h. Lanes 1 and 2, pKMC60; lanes 3 and 4, pKMC61; lanes 5 and 6, pKMC62.

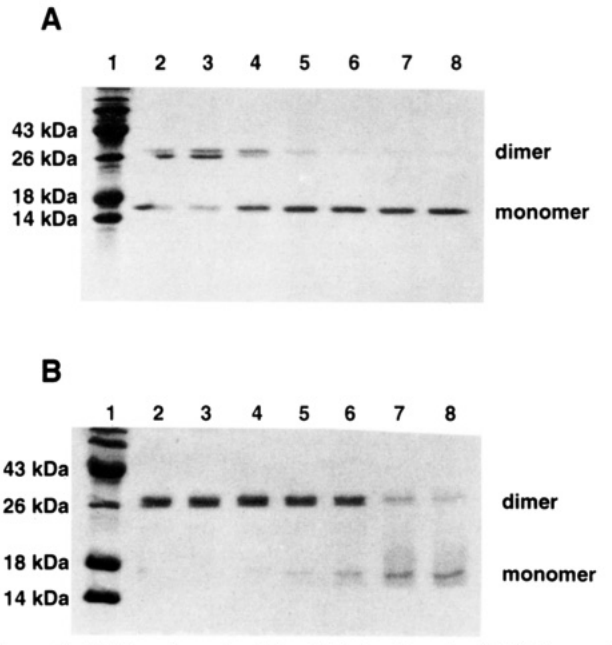


FIGURE 7: SDS–polyacrylamide gel electrophoresis of wild-type and CA mutant MerR proteins incubated with varying amounts of 2-mercaptoethanol. Protein samples were buffer-exchanged into thiol-free solutions containing 2% SDS prior to incubation. Incubations were carried out for 15 min at room temperature. (A) Wild-type protein. (B) CA mutant protein. Lane 1, molecular mass standards. (Lanes 2–8) Protein incubated in the following concentrations of 2-mercaptoethanol: (2) 10^{-5} M, (3) 10^{-4} M, (4) 10^{-3} M, (5) 5×10^{-3} M, (6) 10^{-2} M, (7) 5×10^{-2} M, and (8) 10^{-1} M.

(Figure 7). The wild-type MerR protein can be induced to form an intersubunit disulfide bond only after removal of buffer thiols or prolonged storage (Shewchuk et al., 1989c). The

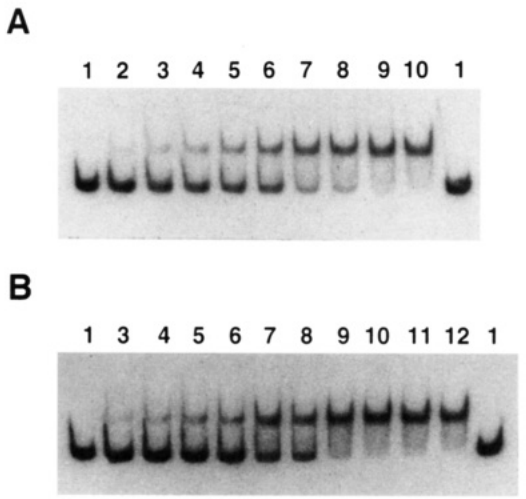


FIGURE 8: Concentration dependence of consensus CA protein–operator DNA binding in the absence (A) and presence (B) of Hg(II). The protein concentration in each lane is as follows: (1) no protein, (2) 1×10^{-11} M, (3) 2.5×10^{-11} M, (4) 4×10^{-11} M, (5) 5.5×10^{-11} M, (6) 7×10^{-11} M, (7) 8.5×10^{-11} M, (8) 1×10^{-10} M, (9) 2.5×10^{-10} M, (10) 4×10^{-10} M, (11) 5.5×10^{-10} M, (12) 7×10^{-10} M. The K_D determined from this experiment is (7–8) $\times 10^{-11}$ and (9–10) $\times 10^{-11}$ M in the absence and presence of Hg(II), respectively.

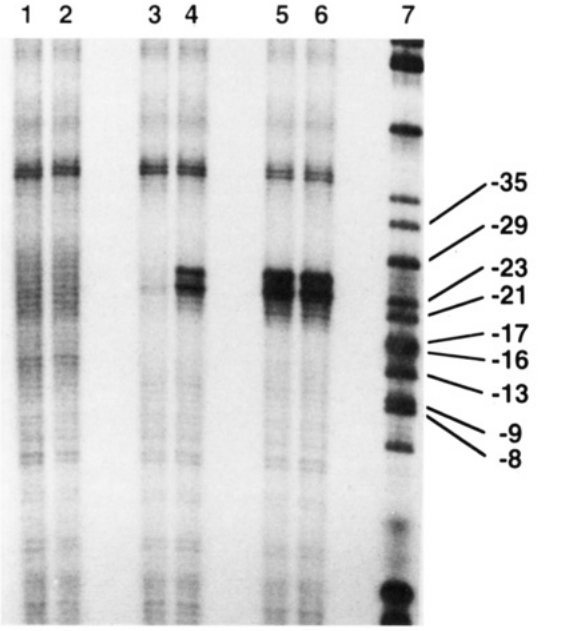


FIGURE 9: Chemical nuclease footprinting assay. An autoradiogram of a denaturing polyacrylamide gel analysis of the results of bis(5-phenyl-1,10-phenanthroline)copper(I) hypersensitivity footprinting is shown. Promoter DNA numbering system is according to Shewchuk et al. (1989c) and Frantz and O'Halloran (1990). Lanes 1 and 2, no protein; lanes 3 and 4, wild-type MerR; lanes 5 and 6, consensus CA mutant MerR; lane 7, Maxam–Gilbert G-specific reaction lane. Lanes 1, 3, and 5, no Hg(II) incubation; lanes 2, 4, and 6, samples incubated with 1 μM $HgCl_2$.

wild-type MerR protein, allowed to dimerize by autoxidation in the absence of 2-mercaptoethanol, can be reduced to 50% monomer by incubation with 1 mM 2-mercaptoethanol for 15 min (Shewchuk et al., 1989c) (Figure 7A). The CA mutant requires 100 mM 2-mercaptoethanol for 50% reduction (Figure 7B). The freshly purified RD mutant S87C does not contain an intersubunit disulfide cross-link and behaves similarly to wild-type MerR upon oxidation (data not shown).

The DNA binding affinity of purified CA protein was measured by gel shift assay (Figure 8). In contrast to the Hg(II)-independent transcriptional activation mutant A89V/S131L recently studied by Parkhill et al. (1993), the CA mutant binds its DNA site with similar affinity to wild-type MerR both in the absence and in the presence of $HgCl_2$.

Bis(5-phenyl-1,10-phenanthroline)copper(I) Hypersensitivity Footprinting. The chemical nuclease activity of 1,10-phenanthroline–cuprous ion complexes depends on *in situ* generation of a reactive tetrahedral cuprous chelate and accessibility of the preformed complex to C1′ hydrogens in the minor groove of DNA (Spassky & Sigman, 1985; Spassky et al., 1988; Thederahn et al., 1989). Frantz and O'Halloran (1990) have demonstrated that a 5-phenyl-substituted 1,10-phenanthroline complex is able to distinguish wild-type repressed and activated MerR states, by exploiting localized underwinding produced in the DNA binding site by the activation-proficient (mercuric ion-bound) form of MerR. The underwound DNA structure apparently allows increased access and cleavage by the reagent. The CA mutant is demonstrated to confer cleavage hypersensitivity both in the presence and in the absence of Hg(II) ions (Figure 9). Cleavage occurs at the same position as for the activated wild-type MerR protein. The cleavage positions are assigned by the guanine-specific fragment cleavage pattern shown in lane 7. As before (Frantz & O'Halloran, 1990; Parkhill et al., 1993), cleavage occurs between the –35 and –10 promoter elements, within the center of an interrupted 14-nucleotide palindrome that forms the MerR binding site (O'Halloran & Walsh, 1987; Shewchuk et al., 1989c).

DISCUSSION

Transcriptional control of mercury resistance in bacteria takes place with a remarkably efficient utilization of protein and DNA sequence. The 144 amino acid MerR protein from Tn*501* functions as both a transcriptional repressor and an activator while bound at the same ca. 30-nucleotide DNA site (O'Halloran et al., 1989; Shewchuk et al., 1989c). Although MerR to date has been intractable to physical characterization by NMR or X-ray crystallography, molecular genetics and biochemistry have been useful in defining amino acid residues important to specific MerR functions. Residues with critical functions in repression, activation, dimerization, DNA binding, and mercuric ion binding can in principle be identified by mutagenesis and screening for a loss of function. Such analyses, together with subsequent biochemical studies, have implicated residues 10–29 as a helix–turn–helix DNA binding motif (Ross et al., 1989; Shewchuk et al., 1989a), residues C82, C117, and C126 as metal binding ligands (Ross et al., 1989; Shewchuk et al., 1989a,b; Helmann et al., 1990a), and residues near the latter two cysteines as important to activation (Ross et al., 1989). Other mutations have been identified which lead either to a loss of repressor function or to a gain of ability to activate transcription in the absence of Hg(II). Mutations A89V or S131L in the Tn*21* MerR protein (94% identical in sequence to Tn*501* MerR) confer a repression-deficient phenotype (Ross et al., 1989). When placed into the Tn*501* protein, the double mutant A89V/S131L has a fully constitutive phenotype and confers bis(5-phenyl-1,10-phenanthroline)copper(I) chemical nuclease hypersensitivity on its DNA binding site in the absence of Hg(II). For wild-type Tn*501* MerR, hypersensitivity is only observed in the presence of Hg(II) (Parkhill et al., 1993). Gain of function mutations such as this one demonstrate alternate structural solutions to the wild-type mechanism of activation and may indicate residues critical to allosteric modulation in the wild-

type protein. We have constructed a synthetic *merR* gene containing 25 unique restriction sites, as well as a single-copy reporter for monitoring transcriptional activation by MerR in *E. coli*. Specific regions indicated by previous studies to be important to function were examined by mutagenesis, and a number of new mutants have been characterized.

Functional Role of Paired Acidic Residues in Repression. Although DNA unwinding appears to be critically important to the mechanism of activation, there may be an additional component of protein–protein interaction between MerR and RNA polymerase. By analogy to other transcriptional activators, both procaryotic and eucaryotic, one might expect mutation of certain clusters of acidic residues to lead to an activation-deficient phenotype (Hope et al., 1988; Sigler, 1988; Bushman et al., 1989). These acidic patch domains have proven to be a component of activation and presumably act to increase the occupancy of RNA polymerase by direct contact (Hochschild et al., 1983; Irwin & Ptashne, 1987; Gaston et al., 1990). The regions from residues R55 to R75 and from A71 to H91 in MerR have been noted previously to contain homology to known helix–turn–helix motifs (Lund et al., 1986; O'Halloran & Walsh, 1987; Ross et al., 1989). Additionally, mutation of an alanine residue at position 60, highly conserved among helix–turn–helix structures, results in a strongly activation-deficient phenotype (Ross et al., 1989). Acidic residues could be aligned on these helices to contact RNA polymerase by forming one or more acidic patch domains (Bushman & Ptashne, 1988; Ross et al., 1989).

The most likely candidate residues for transcription-stimulating RNA polymerase contacts in MerR are a cluster of 6 acidic side chains within a 17-residue region from D68 to E84. When these were substituted with neutral residues, repression- rather than activation-deficient phenotypes were obtained, however (Table 1). Thus, if there is some protein-protein contact involving these residues, then the results indicate it to be a component of repression rather than activation. In fact, apo MerR enhances the affinity of RNAP for the promoter *in vivo*, but concomitantly blocks transcription initiation (Frantz & O'Halloran, 1990; Heltzel et al., 1990). Contacts between MerR and RNAP may actually repress transcription in the wild-type apo MerR complex by holding the RNAP tightly at the promoter site and preventing open complex formation. Such contacts might be weakened in acidic to neutral mutations. An analogous mechanism has been proposed for OmpR-mediated repression at the Cd box binding site upstream of the *ompF* gene in *E. coli* (Tsung et al., 1989), as well as for OmpR repression of a synthetic chimeric promoter (Tsung et al., 1990). It should also be noted that previous studies employing chemical mutagenesis of the entire gene and genetic selection failed to identify activation-deficient mutants in any acidic residues (Ross et al., 1989).

Since the *in vitro* DNA operator binding affinity is similar for the wild-type and E77Q/D78N/E83Q/E84Q quadruple mutant proteins (Figures 2 and 3), strength as a transcriptional repressor is not simply a function of DNA binding strength. This conclusion is also supported by the fact that the Tn*501* MerR mutant C82A binds operator DNA less tightly than wild-type MerR, yet is a better transcriptional repressor (Shewchuk et al., 1989b). Additionally, several cis-acting mutants in the MerR binding site confer moderate defects in MerR–DNA binding (up to 9-fold more MerR is required to gel-shift labeled operator fragments), but are unimpaired for repressor function (Park et al., 1992).

MerR Regional Randomization and Screening. The preceding mutation analysis, as well as a variety of other studies, suggested to us that the region defined in the synthetic *merR* gene by an upstream *Kpn*I site and downstream *Afl*II site (Figure 1) may be a hot spot for other repression-deficient or constitutive activation mutations. The region includes the E83 and E84 residues, as well as A89, previously implicated in repression (Ross et al., 1989). Residues E83 and E84 were more important to repression than the upstream acidic residues studied (Table 1). Furthermore, random chemical mutagenesis coupled with a selection for activation-deficient mutants failed to identify any mutants near C82, but identified several near C117 and C126 (Ross et al., 1989). Residues near cysteines-117 and -126 but not cysteine-82 therefore appear to be important in maintaining the protein's ability to activate transcription in response to $HgCl_2$ (Ross et al., 1989), despite evidence that all three cysteines provide metal ligands. The region surrounding C82 may be of particular importance for dimerization. The C82 residue itself has been shown to be positioned at the dimer interface (Shewchuk et al., 1989a,b,c). Since activation is likely induced in the wild-type protein by Hg(II) binding across the dimer interface (Helmann et al., 1990a), a reasonable hypothesis for the location of an initial protein conformational change associated with activation is the dimer interface. The mercuric ion may act to alter contact in this region by causing otherwise unfavorable interfacial contacts. We reasoned that mutations in this region either may disrupt all MerR function by preventing dimerization or may encourage the subunit contacts characteristic of the activation-proficient conformation.

The region from H81 to K92 was examined by screening a mutant library constructed via a degenerate oligonucleotide PCR mutagenesis protocol. Since a small region was chosen for mutagenic analysis, a high density of mutations per molecule could be examined effectively. The strategy of inducing multiple mutations was useful in obtaining phenotypes that might require more than one amino acid residue change as well as to allow amino acid residue substitutions that require multiple nucleotide substitutions. Screening took place in a host containing a single-copy reporter (Figure 4). Initial screening took place on $HgCl_2$-free plates optimized empirically to distinguish basal- from repression-level *lacZ* transcriptional activity. A secondary screening step, in which positives from the primary screen were tested in the presence of Hg(II) for greater β-galactosidase activity than the unoccupied operator can produce, allowed mutations that yield nonfunctional (misfolded or truncated) MerR to be identified and discarded. Thus, mutants which bind to their operator site but do not repress could be identified. As predicted from the anticipated value of 3.6 nucleotide changes per oligonucleotide, almost all of the mutants sequenced have 3 or 4 mutations (Figure 5). Most of these codon changes were arrived at from single nucleotide substitutions. Furthermore, the strategy to find phenotypes that require multiple mutations was validated by subsequent analysis that showed the CA phenotype requires a double mutation in this region of the MerR protein (Table 2). The high frequency of mutants displaying increased transcriptional activation in the absence of Hg(II) (Figure 5B,C) indicates that our choice of targets was successful. A recent example of regional mutagenesis and selection where desired phenotypes are much less accessible is the direct selection of magnetite binding antibodies from combinatorial libraries (Barbas et al., 1993). Only 4 unique solutions were found from a pool of 10^8 library members. Although our library size was also very large, we found 16 and 8 unique solutions to the RD and CA phenotypes, respectively, by screening only ca. 5000 library members.

Identification of Consensus RD and CA Mutants. Figure 5 represents the results of sequencing a number of mutants and classifying them by their *in vivo* β-galactosidase activities. The point of division between these mutant classes is somewhat arbitrary but essentially is made according to the level of transcription activity in the absence of inducer (mercuric ion). The variation in positional mutation frequencies observed in the parent library (Figure 5A) probably reflects irregularities in oligonucleotide precursor concentrations and stabilities, synthetic coupling efficiencies, and relative codon degeneracies (Oliphant et al., 1986; Hill et al., 1987). The high frequency of S87C substitutions within the RD class and S86C and A89V substitutions within the CA class suggested that these mutations may confer the appropriate phenotypes even in the absence of other coresident mutations. Subsequent site-directed mutagenesis of the wild-type synthetic gene confirmed these predictions (Table 2). The S87C mutation is thus considered a consensus RD mutant for this region, and the S86C/A89V double mutation is considered a consensus CA mutant for this region. The S87C/A89V mutant can also be classified as constitutive, although its phenotype is weaker than that of S86C/A89V. Further studies employed the consensus RD (S87C) and CA (S86C/A89V) mutants constructed by site-directed mutagenesis.

Functional Characterization of Consensus Mutants. Since the activation-proficient conformation of the wild-type protein dimer is likely to represent an energetically demanding structure stabilized by the coordination of Hg(II), one would not predict that highly conservative substitutions, such as serine to cysteine and alanine to valine, would alter the tertiary fold of the protein significantly enough to allow Hg(II)-independent activation. It is possible that the consensus CA mutant is a true bypass mutant and activates transcription by a mechanism other than DNA distortion. The data presented in Figure 9 argue against this possibility, however. The CA mutant distorts DNA and does so at the same positions as Hg(II)-bound wild-type MerR, as probed by a bis(5-phenyl-1,10-phenanthroline)copper(I) reagent (Figure 9). This reagent shows secondary rather than primary structure specificity, although its exact recognition and cleavage specificities are not well understood (Thederahn et al., 1989, 1990). The reagent has been shown to be hyperreactive in the single-stranded regions of open and abortive initiation complexes, as well as in the double-stranded region at the leading edge of a transcription bubble (Thederahn et al., 1990). Frantz and O'Halloran (1990) demonstrated that wild-type MerR protein allows cleavage in the center of the MerR binding site, but only in the presence of mercuric ions. Chemical nuclease recognition and cleavage in this case probably correspond to an induced local untwisting of the site by the activation-proficient Hg(II)–MerR complex (Frantz & O'Halloran, 1990; Ansari et al., 1992). Lane 5 of Figure 9 indicates that constitutive untwisting by the apo form of the CA mutant occurs.

Another possible effect of the consensus CA mutations is that the new cysteine introduced at position 86 allows formation of a stable disulfide bond that locks the dimer into the activation-proficient conformation. Such a disulfide bond could contribute as much or more enthalpic and entropic stabilization as Hg(II) binding. The purified CA mutant does in fact contain an intersubunit disulfide bond, present in fresh protein preparations, that is 2 orders of magnitude more resistant to reduction than a C82–C82 intersubunit disulfide linkage artificially induced by oxidation in the wild-type protein (Figure 7). This level of resistance is probably sufficient for stability *in vivo*. It is noteworthy that the RD phenotype of the E83Q/E84Q mutant is enhanced by mutation of additional aspartate and glutamate residues (Table 1). Apparently the effect of removing several negative charges at the critical positions 83 and 84 can be potentiated by further localized charge reduction. The consensus CA mutant introduces a hydrophobic valine side chain and a hydrophobic disulfide bond, neither of which is sufficient for activity in the absence of the other (Table 2). The question of whether hydrophobic and disulfide bond modifications can be additive (or multiplicative) with reduction of negative charge is partially addressed by our isolation of an E84G/S86C/A89V triple mutant from the regionally randomized library (Figure 5C). This mutant displays a higher level of uninduced β-galactosidase reactivity than any other mutant examined.

In vitro, the affinity of the CA mutant for operator DNA is very similar to that of wild-type MerR (Figures 2 and 8). This result contrasts with that obtained recently for another constitutive activation mutant, A89V/S131L, where a 20-fold lower affinity was noted (Parkhill et al., 1993). The CA mutant studied here may be stabilized by the novel disulfide bond in its ability to bind and induce the distorted DNA conformation required in the promoter for transcriptional activation.

Coexpression of the wild-type and CA mutant proteins *in vivo* led to a surprisingly strong dampening of the mercury-independent transcriptional activation phenotype. Both proteins were expressed from P_{lac} promoters, but approximately 5-fold more copies of the CA mutant-containing plasmid were present in the cells. Surprisingly, the level of β-galactosidase activity in the absence of $HgCl_2$ was reduced 5-fold relative to the CA mutant alone. This result suggests that the CA mutant is only partially dominant and confirms that it competes for binding with wild-type MerR. It would also suggest that the CA mutant binds with significantly lower affinity *in vivo* than wild-type protein to this site, in contrast to direct measurements *in vitro* (Figure 8). Heterodimers of the two proteins perhaps are formed *in vivo* in this experiment and bind tightly to repress transcription. Alternatively, the CA mutant may not be very stable *in vivo*, and the amount of active protein present is actually much lower than expected. This would be consistent with our difficulties in expressing the protein at high levels for purification.

Figure 10 depicts a schematic for the interaction of the MerR dimer with its operator site. The proposed helix–turn–helix elements at residues 10–29 provide binding specificity at each half-site. The tridentate ligation of activator Hg(II) across the dimer interface involves C82 on one subunit. The introduction of new cysteine residues near this one can have profound effects on activity in the absence of mercury, and can substitute for the C82 or C117 positions in the presence of mercury (Table 3). A precise understanding of disulfide bond formation in the absence of inducer and altered metal ligation in its presence awaits further structural analysis.

In summary, the synthetic gene and host-integrated reporter described here have been demonstrated to be versatile and efficient tools for molecular genetic analysis of MerR structure and function relationships. We have tested the functional significance of a cluster of acidic residues proposed to be important to activation, identified a number of new repression-deficient and constitutive activation mutants, and studied one CA mutant protein in detail. The results suggest that the acidic residues function in repression rather than activation and that they may be responsible for sequestering RNAP at an inactive promoter. The presence of a stable disulfide bond

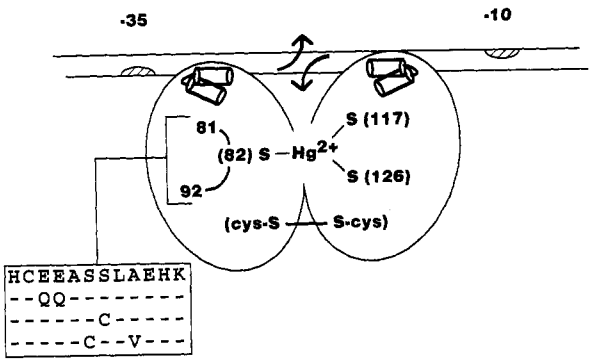


FIGURE 10: Schematic indicating structural and functional models for MerR. The MerR homodimer is shown in its transcriptional activation mode, indicated by arrows portraying localized underwinding at the center of the DNA binding site. The illustrated helix–turn–helix DNA binding motifs are comprised of residues 10–29 of each subunit. Activation is initiated either by tricoordinate ligation of a mercuric ion to cysteines-82, -117, and -126 in the wild-type protein or by the apo form of CA mutant proteins. The region studied by degenerate oligonucleotide mutagenesis is shown in the box, as well as the E83Q/E84Q, S87C, and S86C/A89V mutations. An intersubunit disulfide bond that may be responsible for activation by the S86C/A89V CA mutant is also shown, below the wild-type mercury binding site.

in an S86C/A89V CA mutant may constrain the protein to adopt the conformation and DNA unwinding activities of its activated mercurated wild-type form.

ACKNOWLEDGMENT

We are grateful to Barry T. Ballard for helpful discussions and assistance during the course of this work and to Dr. Barry L. Wanner for critical reading of the manuscript. Dr. Daniel V. Santi was instrumental in the design of the synthetic *merR* gene.

REFERENCES

Ansari, A. Z., Chael, M. L., & O'Halloran, T. V. (1992) *Nature 355*, 87–89.
Barbas, C. F., III, Rosenblum, J. S., & Lerner, R. A. (1993) *Proc. Natl. Acad. Sci. U.S.A. 90*, 6385–6389.
Boyd, D., Manoil, C. & Beckwith, J. (1987) *Proc. Natl. Acad. Sci. U.S.A. 84*, 8525–8529.
Bushman, F. D., & Ptashne, M. (1988) *Cell 54*, 191–197.
Bushman, F. D., Shang, C., & Ptashne, M. (1989) *Cell* 58, 1163–1171.
Frantz, B., & O'Halloran, T. V. (1990) *Biochemistry 29*, 4747–4751.
Gaston, K., Bell, A., Kolb, A., Buc, H., & Busby, S. (1990) *Cell 62*, 733–743.
Helmann, J. D., Wang, Y., Mahler, I., & Walsh, C. T. (1989) *J. Bacteriol. 171*, 222–229.
Helmann, J. D., Ballard, B. T., & Walsh, C. T. (1990a) *Science 247*, 946–948.
Helmann, J. D., Shewchuk, L. M., & Walsh, C. T. (1990b) in *Metal-Ion Induced Regulation of Gene Expression* (Eichorn, G. L., & Marzilli, L. G., Eds.) pp 33–61, Elsevier, New York.
Heltzel, A., Lee, I. W., Totis, P. A., & Summers, A. O. (1990) *Biochemistry 29*, 9572–9584.
Hill, D. E., Oliphant, A. R., & Struhl, K. (1987) *Methods Enzymol. 155*, 558–568.
Hochschild, A., Irwin, N., & Ptashne, M. (1983) *Cell 32*, 319–325.
Hope, I. A., Mahadevan, S., & Struhl, K. (1988) *Nature 333*, 635–640.
Irwin, N., & Ptashne, M. (1987) *Proc. Natl. Acad. Sci. U.S.A. 84*, 8315–8319.
Le Pecq, J.-B., & Paoletti, C. (1966) *Anal. Biochem. 17*, 100–107.
Lund, P. A., Ford, S. J., & Brown, N. L. (1986) *J. Gen. Microbiol. 132*, 465–480.
Maxam, A. H., & Gilbert, W. (1980) *Methods Enzymol. 65*, 499–560.
Miller, J. H. (1972) *Experiments in Molecular Genetics*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
Misra, T. K. (1992) *Plasmid 27*, 4–16.
Moore, M. J., & Walsh, C. T. (1989) *Biochemistry 28*, 1183–1194.
O'Halloran, T., & Walsh, C. (1987) *Science 235*, 211–214.
O'Halloran, T. V., Frantz, B., Shin, M. K., Ralston, D. M., & Wright, J. G. (1989) *Cell 56*, 119–129.
Oliphant, A. R., Nussbaum, A. L., & Struhl, K. (1986) *Gene 44*, 177–183.
Park, S.-J., Wireman, J., & Summers, A. O. (1992) *J. Bacteriol. 174*, 2160–2171.
Parkhill, J., & Brown, N. L. (1990) *Nucleic Acids Res. 18*, 5157–5162.
Parkhill, J., Ansari, A. Z., Wright, J. G., Brown, N. L., & O'Halloran, T. V. (1993) *EMBO J. 12*, 413–421.
Ralston, D. M., & O'Halloran, T. V. (1990) *Proc. Natl. Acad. Sci. U.S.A. 87*, 3846–3850.
Ross, W., Park, S.-J., & Summers, A. O. (1989) *J. Bacteriol. 171*, 4009–4018.
Sambrook, J., Fritsch, E. F., & Maniatis, T. (1989) *Molecular Cloning—A Laboratory Approach*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
Sanger, F., Nicklen, S., & Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. U.S.A. 74*, 5463–5468.
Shewchuk, L. M., Helmann, J. D., Ross, W., Park, S. J., Summers, A. O., & Walsh, C. T. (1989a) *Biochemistry 28*, 2340–2344.
Shewchuk, L. M., Verdine, G. L., Nash, H., & Walsh, C. T. (1989b) *Biochemistry 28*, 6140–6145.
Shewchuk, L. M., Verdine, G. L., & Walsh, C. T. (1989c) *Biochemistry 28*, 2331–2339.
Sigler, P. (1988) *Nature 333*, 210–212.
Silhavy, T. J., Berman, M. L., & Enquist, L. W. (1984) *Experiments with Gene Fusions*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
Spassky, A., & Sigman, D. S. (1985) *Biochemistry 24*, 8050–8056.
Spassky, A., Rimsky, S., Buc, H., & Busby, S. (1988) *EMBO J. 7*, 1871–1879.
Summers, A. O. (1992) *J. Bacteriol. 174*, 3097–3101.
Tabor, S., & Richardson, C. C. (1985) *Proc. Natl. Acad. Sci. U.S.A. 82*, 1074–1078.
Thederahn, T. B., Kuwabara, M. D., Larsen, T. A., & Sigman, D. S. (1989) *J. Am. Chem. Soc. 111*, 4941–4946.
Thederahn, T., Spassky, A., Kuwabara, M. D., & Sigman, D. S. (1990) *Biochem. Biophys. Res. Commun. 168*, 756–762.
Tsung, K., Brissette, R., & Inouye, M. (1989) *J. Biol. Chem. 264*, 10104–10109.
Tsung, K., Brissette, R. E., & Inouye, M. (1990) *Proc. Natl. Acad. Sci. U.S.A. 87*, 5940–5944.
Wright, J. G., Natan, M. J., MacDonnell, F. M., Ralston, D. M., & O'Halloran, T. V. (1990) *Prog. Inorg. Chem. 38*, 328–412.