JOURNAL OF BACTERIOLOGY, Apr. 1995, p. 1742–1750 Vol. 177, No. 7
0021-9193/95/$04.00+0

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

WILLIAM P. FAWCETT AND RICHARD E. WOLF, JR.*

Department of Biological Sciences, University of Maryland Baltimore County, Baltimore, Maryland 21228

Received 9 November 1994/Accepted 17 January 1995

In *Escherichia coli* K-12, transcription of *zwf*, the gene for glucose 6-phosphate dehydrogenase, is subject to growth rate-dependent regulation and is activated by SoxS in response to superoxide stress. To define genetically the site of SoxS activation, we undertook a detailed deletion analysis of the *zwf* promoter region. Using specifically targeted 5′ and 3′ deletions of *zwf* sequences, we localized the SoxS activation site to a 21-bp region upstream of the *zwf* promoter. This minimal "soxbox" was able to confer paraquat inducibility when placed upstream of a normally unresponsive *gnd-lacZ* protein fusion. In addition, we used these findings as the basis for resecting unnecessary sequences from the region upstream of the promoters of two other SoxS-regulated genes, *sodA* and *nfo*. Like the *zwf* soxbox, the regions required for SoxS activation of *sodA* and *nfo* appear to lie just upstream of or overlap the −35 hexamers of the corresponding promoters. Importantly, the sequence boundaries established here by deletion analysis agree with the primary SoxS recognition sites of *zwf*, *sodA*, and *nfo* that we previously identified in vitro by gel mobility shift and DNase I protection assays with a purified MalE-SoxS fusion protein.

In *Escherichia coli*, the oxidative branch of the pentose phosphate pathway provides ribose for nucleoside biosynthesis and NADPH for reductive biosynthesis (9, 12). Two of the enzymes of this pathway, glucose 6-phosphate dehydrogenase (encoded by *zwf*) and 6-phosphogluconate dehydrogenase (encoded by *gnd*) are subject to growth rate-dependent regulation (40). The specific activities of these enzymes increase in proportion to increased growth rate during steady-state growth on different carbon sources. Studies on the growth rate-dependent regulation of *zwf* and *gnd* have shown that at least part of the increase in the specific activities of these enzymes is attributable to an increase in the transcription rates of the respective genes (25, 26).

Apart from its basic growth rate-dependent regulation, *zwf* is also a member of at least two other regulons. Thus, unlike *gnd*, *zwf* expression is transcriptionally activated by SoxS during episodes of oxidative stress induced by exposure of *E. coli* to superoxide-generating agents such as paraquat (17, 18, 39). Indeed, *E. coli* strains deleted for *zwf* are hypersensitive to oxidative stress, thus making *zwf* a pivotal member of the *soxRS* regulon (18, 21). In addition, treatment of cells with several antibiotics and aromatic weak acids appears to activate *zwf* transcription through the action of *marA*, thus making *zwf* a member of the multiple antibiotic resistance (*mar*) regulon (2, 6, 7, 14, 16).

The work presented here has several overlapping long-range objectives. One is to understand the physiological basis for *zwf*'s membership in the *soxRS* and *marRAB* regulons. The other is to identify the *cis*-acting regulatory sites of *zwf* that are involved in superoxide control for subsequent comparison with the sites required for growth rate control and for induction by antibiotics and aromatic weak acids.

Recently, we described the purification of a MalE-SoxS fusion protein and its use in in vitro determination of the control sites of *zwf* and other member genes of the *soxRS* regulon, such as *sodA*, *nfo*, *micF*, and *fumC* (11). Gel mobility shift assays and DNase I protection assays revealed that specific sequences near the *zwf* promoter are recognized by MalE-SoxS (11). While these in vitro studies were under way, we also attempted to define genetically the SoxS activation site on *zwf*. Taking advantage of PCR and recombinant DNA techniques, we made specific 5′ and 3′ deletions of the *zwf* promoter region and examined the effects of these deletions on the paraquat inducibility of *lacZ* protein fusions in vivo. The results of this deletion analysis of *zwf*, as well as of *sodA* and *nfo*, form the basis of this report.

MATERIALS AND METHODS

Chemicals and enzymes. Methyl viologen (paraquat), ampicillin, chloramphenicol, and ONPG (*o*-nitrophenyl-β-D-galactopyranoside) were purchased from Sigma, St. Louis, Mo. IPTG (isopropyl-β-D-thiogalactopyranoside), X-Gal (5-bromo-4-chloro-3-indolyl-β-D-galactoside), and β-mercaptoethanol were obtained from Boehringer Mannheim Biochemicals (Indianapolis, Ind.), 5 Prime 3 Prime, Inc. (Boulder, Colo.), and Bio-Rad Laboratories (Richmond, Calif.), respectively. Restriction enzymes were purchased from Bethesda Research Laboratories (Gaithersburg, Md.), Boehringer Mannheim Biochemicals, New England Biolabs (Beverly, Mass.), and Promega Corporation (Madison, Wis.). Deoxynucleoside triphosphates (dNTPs) were obtained from Pharmacia LKB Biotechnology Inc. (Piscataway, N.J.) or Promega. Phage T4 DNA ligase, phage T4 polynucleotide kinase, phage T4 DNA polymerase, and *Taq* DNA polymerase were from Promega. The *fmol* DNA sequencing system was from Promega, and Sequenase version 2.0 was from United States Biochemical (Cleveland, Ohio). [α-^{35}S]dATP, [α-^{32}P]dATP, and [γ-^{32}P]ATP were obtained from New England Nuclear (Boston, Mass.). All other chemicals and reagents not listed here were obtained either from Fisher Scientific (Fair Lawn, N.J.), Sigma, or VWR Scientific (Piscataway, N.J.).

Media, growth conditions, and β-galactosidase assays. LB or LB supplemented with glucose (2.0 mg/ml) was used for genetic and paraquat inducibility experiments. Ampicillin was added at 100 μg/ml for growth of plasmid strains and at 50 μg/ml for growth of lysogen strains.

For the paraquat inducibility experiments, cell cultures were diluted 1:200 from overnight cultures in 50 ml of fresh medium and incubated in an orbit shaker water bath at 37°C with vigorous agitation. Cultures were grown to an optical density at 600 nm of 0.300, and 25 ml was removed to a second flask. Paraquat was added to one flask at a final concentration of 0.5 mM, and incubation continued for 90 min. At the end of this growth period, 2 ml of each

* Corresponding author. Mailing address: Department of Biological Sciences, University of Maryland Baltimore County, Baltimore, MD 21228. Phone: (410) 455-2268. Fax: (410) 455-3875. Electronic mail address: wolf@umbc.edu.

culture was transferred to culture tubes prechilled on ice, and chloramphenicol was added at 0.1 mg/ml (final concentration). β-Galactosidase activity was determined by the colorimetric assay of Miller (23) on cells permeabilized with chloroform and sodium dodecyl sulfate. For each experiment, strains were routinely grown as duplicate cultures and each culture was assayed five times. All experiments were repeated at least two to three times on different days. On a given day, the coefficient of variation of the β-galactosidase activities for a bacterial fusion strain was generally less than 10% of the mean.

Bacterial strains. All experiments were performed with *E. coli* K-12 strains. Strain HB301 [W3110 Δ(*argF-lac*)*U169*] (5) was used as the transformation recipient of the *lacZ* protein fusions in plasmid pMLB1034. HB301 was also used as the host for integration of the λ phages carrying various protein fusions and as the source of genomic DNA for PCR (see below). Strain DH5α [*endA1 hsdR17 supE44 thi-1 recA1 gyrA relA1* Δ(*argF-lac*)*U169 deoR* (ϕ80 *lacZ*ΔM15)] was used for the selection of DNA inserts in pUC18 and pUC19. Strain NF3079 [*dam-3 araD139* Δ(*araABOIC-leu*)*7679 rpsL galU galK* Δ(*lac*)*X74*] was used for the preparation of plasmid DNA requiring restriction with *Bcl*I. The properties of strain HB301(λDR101), which carries a *zwf-lac* protein fusion, have been described elsewhere (27).

Plasmids. The plasmids constructed for this study are listed in Table 1. Cloning vectors, pUC18 and pUC19, were used for the isolation of specific DNA fragments. Plasmid pMLB1034, a derivative of pBR322, has an *Eco*RI-*Sma*I-*Bam*HI cloning site that can be used for construction of protein fusions to *lacZ* (34). Plasmids pDR17 (*zwf*), pCMP78(+) (*gnd*), and pSOD-1 (*sodA*) have been described elsewhere (5, 27, 38). Plasmid pIT3 (*nfo*) carries the *Bam*HI-*Eco*RI fragment from pRPC124 (30) in pRS551 (36). The remaining plasmids were constructed as detailed in Table 1 by using the standard recombinant methods outlined below. Plasmids pSOD-1 and pIT3 were kindly provided by H. M. Steinman (Albert Einstein College of Medicine, Yeshiva University) and B. Weiss (University of Michigan Medical School), respectively.

Oligonucleotide primers. The synthetic oligonucleotide primers used for PCR and DNA sequencing are listed in Table 2. Except for Q4(LacZ) (Oligos Etc., Wilsonville, Ore.) and −40SP (−40 sequencing primer; New England Biolabs), the oligonucleotide primers were synthesized in the Biopolymer Laboratory, University of Maryland, Baltimore, or the Department of Biological Sciences, University of Maryland Baltimore County, Baltimore. For each oligonucleotide primer, the 5′ and 3′ positions indicated in Table 2 refer to sequence positions relative to the proposed transcriptional start sites of the genes, with the following exceptions. The positions of pMLB1034-5′, pMLB1034-3′, pBR322CW, and Q4(LacZ) are numbered with respect to the first base of the *Eco*RI site of pMLB1034. Similarly, the position of −40SP is numbered with respect to the first base of the *Eco*RI site of pUC19.

To facilitate subcloning of PCR DNA fragments and construction of the protein fusions to *lacZ*, specific restriction sites were incorporated into the oligonucleotide primers with the exception of the *Bcl*I site in Zwf5′-1. These restriction sites are identified in Table 2.

PCR. PCR was performed on either purified plasmid DNA or genomic DNA. Typical reaction volumes were 0.100 ml and contained 50 mM KCl, 10 mM Tris-HCl (pH 9.0), 0.10% Triton X-100, 200 μM each dNTP, 1.0 μM each oligonucleotide primer, and 2.5 U of *Taq* DNA polymerase. Plasmid DNA was added at 1 to 10 ng per reaction mixture. For reactions using genomic DNA, a small amount of cells from −80°C freezer stocks was added directly to the reaction mixture. The primers and DNA templates used for each PCR are indicated in Table 1.

Amplification was performed as follows in a Techne PHC-2 dry block thermocycler (Techne Incorporated, Princeton, N.J.). Reaction samples were initially heated to 95°C for 5 min. Afterwards, the reaction samples were subjected to 35 cycles of 95°C for 75 s, 55°C for 75 s, and 72°C for 75 s. A final cycle of 95°C for 75 s, 55°C for 75 s, and 72°C for 240 s was used to provide maximum extension by *Taq* DNA polymerase.

Following amplification, the PCR products were purified by using the Magic PCR Preps DNA purification system (Promega) as specified by the manufacturer. Yield and quality of the purified PCR products were examined by visualization of sample aliquots on agarose or polyacrylamide gels after staining with ethidium bromide. PCR products were either flush ended with T4 DNA polymerase (3) or digested with the appropriate restriction enzymes prior to subcloning.

Recombinant methods and DNA sequencing. Standard recombinant DNA methods were used (3, 23, 28, 34). Purified plasmid DNA was prepared by the alkaline lysis method, using the Magic Minipreps DNA purification system (Promega) as specified by the manufacturer. Bacteriophage λ DNA was purified by using the Magic Lambda Preps DNA purification system (Promega) as specified by the manufacturer.

DNA inserts in pUC18 and pUC19 were selected from *E. coli* DH5α transformants exhibiting the loss of *lacZ*α complementation on LB plates supplemented with ampicillin (0.1 mg/ml), X-Gal (40 μg/ml), and IPTG (0.2 mg/ml). Protein fusions in pMLB1034 were selected from HB301 transformants exhibiting β-galactosidase expression on LB plates supplemented with ampicillin (0.1 mg/ml) and X-Gal (40 μg/ml). PCR and restriction analysis were used to confirm the identities of selected transformants. Plasmid DNA (e.g., pDR17 and pZ1) to be restricted with *Bcl*I was prepared from *E. coli* NF3079.

In vivo recombination between plasmids carrying the various protein fusions and λRZ5 was used to transfer the fusions to bacteriophage λ (27). Recombinant bacteriophage were selected as Lac^+ plaques on strain HB301 plated with top agar on H plates (23) with X-Gal and purified to homogeneity. Lac^+ Amp^r lysogens of the recombinant phages were selected by streaking HB301 transductants on either lactose MacConkey or H X-Gal plates supplemented with ampicillin (25 μg/ml). The lysogens were screened with λ *c*I and λ *c*I *c*17 to isolate monolysogens (33). To confirm that the lysogens retained the fusion junctions of the respective plasmid fusions, the lysogens were screened by PCR using flanking oligonucleotide primers.

In addition to the selection criteria mentioned above, the identities of the protein fusions were confirmed by DNA sequencing. DNA sequencing was performed by the chain termination method (29), using either Sequenase version 2.0 or the *fmol* DNA sequencing system. DNA sequencing templates included purified plasmid DNA, purified λ phage DNA, and purified PCR DNA fragments from the HB301 monolysogens.

Nomenclature. As a matter of convention, the protein fusions in this report will be referred to by their common fusion names without the plasmid (p) or bacteriophage (λ) prefix unless necessary to avoid confusion. The K designation (Z1K, Z3K, Z4K, etc.) indicates that the construct has an engineered internal *Kpn*I site immediately upstream of the *zwf* promoter's −35 hexamer. In addition, fusion names were assigned for identification purposes and do not necessarily reflect the order in which the protein fusions were constructed.

DNA sequences. The DNA sequences of the genes in this report are available in the following references: *zwf* (27), *gnd* (24), *sodA* (38), and *nfo* (30). The GenBank accession numbers for these sequences are M55005 (*zwf*), K02072 (*gnd*), M94879 (*sodA*), and M22591 (*nfo*). R. Cunningham (State University of New York at Albany) was kind enough to provide us with unpublished *nfo* sequences.

RESULTS

Preparation of *zwf-lacZ* protein fusion Z1. Our initial approach to delineating the *cis*-acting site in *zwf* that is involved in the *soxRS*-mediated response to oxidative stress was based on the observation that a *zwf-lacZ* operon fusion, λDR52 (26), was regulated by *soxRS* at the transcriptional level (18). This operon fusion carries a 686-bp *Bcl*I promoter-containing *zwf* fragment (positions −140 to +546) fused to *′trpA-′lacOZYA′* (26). Fortuitously, the *Bcl*I site within the *zwf* coding sequence is in frame, and taking advantage of this fact, we isolated the 686-bp *Bcl*I fragment from NF3079(pDR17) and subcloned it into the protein fusion expression vector pMLB1034 (34). The resultant *zwf-lacZ* protein fusion, pZwf, was subsequently transferred to λRZ5 by in vivo recombination, and the λ *zwf-lacZ* protein fusion phage was integrated into the *att*λ site of strain HB301.

Following treatment with paraquat, the β-galactosidase activity of strain HB301(λZwf) was induced 5.7-fold (Table 3). Similarly, it was determined that HB301(λDR101) (26), which carries *zwf* sequences from −646 to +119, was induced 6.4-fold. The data obtained with these two λ *zwf-lacZ* protein fusions indicate that *zwf* sequences between −140 and +119 are sufficient for paraquat inducibility. On the basis of these results, we constructed the *zwf-lacZ* fusion Z1, which carries *zwf* sequences from −140 to +118. The 1-bp difference at the 3′ end of fusion Z1 (+118), compared with fusion DR101 (+119), is due to the incorporation of a *Bam*HI site in the oligonucleotide primer, Zwf3′-2 (Table 2). The C-to-G transition at position +119 shortens the *zwf* sequence by 1 bp but does not alter the amino acid (Gly) at codon 19 (27). The β-galactosidase activity of strain HB301(λZ1) was induced by treatment with paraquat to the same extent (6.8-fold) as that observed for strains HB301(λDR101) and HB301(λZwf) (Table 3). We do not know why the β-galactosidase activity of strain HB301(λZwf) is twofold higher than that of strains HB301(λDR101) and HB301(λZ1). Regardless, the factor of induction is the same for the three strains.

The *zwf-lacZ* protein fusion Z1 became the starting point for the preparation of the additional *zwf* deletion mutants described in this report. The nucleotide sequence of the *zwf* promoter region of construct Z1 is shown in Fig. 1. Relevant

TABLE 1. Plasmids used in this study

Plasmid	Genotype or description[a]	Source or reference
pUC18, pUC19	Cloning vectors with *bla lacI OPZ′*α	Laboratory stock
pMLB1034	Translation fusion cloning vector carrying *′lacZ-lacY′* in pBR322, Amp^r	34
pDR17	pBR322 *zwf*$^+$ (−308 to +1682) Amp^r	27
pCMP78(+)	pMLB1034::Φ(*gnd′-′lacZY′*)(−135 to +290)(Hyb)(PQ$^+$)	5
pIT3	pRS551::Φ(*nfo′-lacZ*) [operon fusion] Amp^r Kan^r	B. Weiss
pSOD-1	pBR322 *sodA*$^+$ Amp^r	H. M. Steinman
pUC18Z1	274-bp PCR product (Zwf5′-1 and Zwf3′-2 on pDR17) in *Hin*cII site of pUC18 (−140 to +118)	This study
pUC19Z1	275-bp *Eco*RI-*Bam*HI fragment from pZ1 in pUC19 (−140 to +118)	This study
pUC18Z2	220-bp PCR product (Zwf5′-2 and Zwf3′-2 on pDR17) as a 214-bp *Kpn*I-*Bam*HI fragment in pUC18 (−86 to +118)	This study
pUC18Z9	172-bp PCR product (Zwf5′-8 and Zwf3′-2 on pDR17) as a 165-bp *Kpn*I-*Bam*HI fragment in pUC18 (−35 to +118)	This study
pUC18Z1K-3′	159-bp PCR product (−40SP and Zwf3′-6 on pUC19Z1) as a 116-bp *Eco*RI-*Kpn*I fragment in pUC18 (−140 to −42)	This study
pUC18Z1K	165-bp *Kpn*I-*Bam*HI fragment from pUC18Z9 in pUC18Z1K-3′ (−140 to −42::*Kpn*I::−35 to +118)	This study
pUC18Z3K	pUC18Z1K digested with *Bss*HII, filled in with Klenow fragment and dNTPs, digested with *Sma*I, ligated (−76 to −42::*Kpn*I::−35 to +118)	This study
5′ and internal deletions of *zwf-lacZ* protein fusions		
pZwf	686-bp *Bcl*I fragment from pDR17 in *Bam*HI site of pMLB1034; pMLB1034::Φ(*zwf′-′lacZY′*)(−140 to +546)(Hyb)(PQ$^+$)	This study
pZ1	265-bp *Bcl*I-*Bam*HI fragment from pUC18Z1 in *Bam*HI site of pMLB1034; pMLB1034::Φ(*zwf′-′lacZY′*)(−140 to +118)(Hyb)(PQ$^+$)	This study
pZ2	226-bp *Eco*RI-*Bam*HI fragment from pUC18Z2 in pMLB1034; pMLB1034::Φ(*zwf′-′lacZY′*)(−86 to +118)(Hyb)(PQ$^+$)	This study
pZ3K	208-bp *Eco*RI-*Bam*HI fragment from pUC18Z3K in pMLB1034; pMLB1034::Φ(*zwf′-′lacZY′*)(−76 to −42::*Kpn*I::−35 to +118)(Hyb)(PQ$^+$)	This study
pZ4	204-bp PCR product (Zwf5′-3 and Zwf3′-2 on pDR17) as a 197-bp *Eco*RI-*Bam*HI fragment in pMLB1034; pMLB1034::Φ(*zwf′-′lacZY′*)(−67 to +118)(Hyb)(PQ$^+$)	This study
pZ4K	204-bp PCR product (Zwf5′-3 and Zwf3′-2 on pUC18Z1K) as a 197-bp *Eco*RI-*Bam*HI fragment in pMLB1034; pMLB1034::Φ(*zwf′-′lacZY′*)(−67 to −42::*Kpn*I::−35 to +118)(Hyb)(PQ$^+$)	This study
pZ5	199-bp PCR product (Zwf5′-4 and Zwf3′-2 on pDR17) as a 192-bp *Eco*RI-*Bam*HI fragment in pMLB1034; pMLB1034::Φ(*zwf′-′lacZY′*)(−62 to +118)(Hyb)(PQ$^+$)	This study
pZ6	193-bp PCR product (Zwf5′-5 and Zwf3′-2 on pDR17) as a 186-bp *Eco*RI-*Bam*HI fragment in pMLB1034; pMLB1034::Φ(*zwf′-′lacZY′*)(−57 to +118)(Hyb)(PQ$^-$)	This study
pZ7	191-bp PCR product (Zwf5′-6 and Zwf3′-2 on pDR17) as a 184-bp *Eco*RI-*Bam*HI fragment in pMLB1034; pMLB1034::Φ(*zwf′-′lacZY′*)(−55 to +118)(Hyb)(PQ$^-$)	This study
pZ8	188-bp PCR product (Zwf5′-7 and Zwf3′-2 on pDR17) as a 181-bp *Eco*RI-*Bam*HI fragment in pMLB1034; pMLB1034::Φ(*zwf′-′lacZY′*)(−51 to +118)(Hyb)(PQ$^-$)	This study
pZ9	177-bp *Eco*RI-*Bam*HI fragment from pUC18Z9 in pMLB1034; pMLB1034::Φ(*zwf′-′lacZY′*)(−35 to +118)(Hyb)(PQ$^-$)	This study
pZ1K	275-bp *Eco*RI-*Bam*HI fragment from pUC18Z1K in pMLB1034; pMLB1034::Φ(*zwf′-′lacZY′*)(−140 to −42::*Kpn*I::−35 to +118)(Hyb)(PQ$^+$)	This study
pZ1KΔ1	140-bp PCR product (pBR322CW and Zwf3′-7 on pDR17) as a 111-bp *Eco*RI-*Kpn*I fragment in pZ9; pMLB1034::Φ(*zwf′-′lacZY′*)(−140 to −47::*Kpn*I::−35 to +118)(Hyb)(PQ$^-$)	This study
pZ1KΔ2	135-bp PCR product (pBR322CW and Zwf3′-8 on pDR17) as a 106-bp *Eco*RI-*Kpn*I fragment in pZ9; pMLB1034::Φ(*zwf′-′lacZY′*)(−140 to −52::*Kpn*I::−35 to +118)(Hyb)(PQ$^-$)	This study
3′ deletions of *zwf-gnd-lacZ* protein fusions		
pUC19*gnd*	446-bp PCR product [Gnd5′-3 and Q4 on pCMP78(+)] as a 337-bp *Kpn*I-*Bam*HI fragment in pUC19 (−36 to +290)	This study
p1034*gnd*	349-bp *Eco*RI-*Bam*HI fragment from pUC19*gnd* in pMLB1034; pMLB1034::Φ(*gnd′-′lacZY′*)(−36 to +290)(Hyb)(PQ$^-$)	This study
pZ1G1	243-bp PCR product (pBR322CW and Zwf3′-3 on pZ1)::*Sty*I::346-bp PCR product [Gnd5′-1 and Q4 on pCMP78(+)] as a 452-bp *Eco*RI-*Bam*HI fragment in pMLB1034; pMLB1034::Φ(*zwf′-′gnd′-′lacZY′*)(−140 to +60::*Sty*I::+62 to +290)(Hyb)(PQ$^+$)	This study
pZ1G2	215-bp PCR product (pBR322CW and Zwf3′-4 on pZ1)::*Nsp*V::403-bp PCR product [Gnd5′-2 and Q4 on pCMP78(+)] as a 481-bp *Eco*RI-*Bam*HI fragment in pMLB1034; pMLB1034::Φ(*zwf′-′gnd′-′lacZY′*)(−140 to +32::*Nsp*V::+4 to +290)(Hyb)(PQ$^+$)	This study
pZ1G3	184-bp PCR product (pBR322CW and Zwf3′-5 on pZ1)::*Nsp*V::403-bp PCR product [Gnd5′-2 and Q4 on pCMP78(+)] as a 450-bp *Eco*RI-*Bam*HI fragment in pMLB1034; pMLB1034::Φ(*zwf′-′gnd′-′lacZY′*)(−140 to +1::*Nsp*V::+4 to +290)(Hyb)(PQ$^+$)	This study
pZ4G1	482-bp PCR product (Zwf5′-3 and Q4 on pZ1G1) as a 374-bp *Eco*RI-*Bam*HI fragment in pMLB1034; pMLB1034::Φ(*zwf′-′gnd′-′lacZY′*)(−67 to +60::*Sty*I::+62 to +290)(Hyb)(PQ$^+$)	This study
pZ4G2	511-bp PCR product (Zwf5′-3 and Q4 on pZ1G2) as a 403-bp *Eco*RI-*Bam*HI fragment in pMLB1034; pMLB1034::Φ(*zwf′-′gnd′-′lacZY′*)(−67 to +32::*Nsp*V::+4 to +290)(Hyb)(PQ$^+$)	This study

Continued on following page

TABLE 1—*Continued*

Plasmid	Genotype or description[a]	Source or reference
pZ4G3	480-bp PCR product (Zwf5′-3 and Q4 on pZ1G3) as a 372-bp *Eco*RI-*Bam*HI fragment in pMLB1034; pMLB1034::Φ(*zwf′-′gnd′-′lacZY′*)(−67 to +1::*Nsp*V::+4 to +290)(Hyb)(PQ^+)	This study
pZ3G4	49-bp *Eco*RI-*Kpn*I fragment from pUC18Z3K::*Kpn*I::337 bp *Kpn*I-*Bam*HI fragment from pUC19*gnd* as a 380-bp *Eco*RI-*Bam*HI fragment in pMLB1034; pMLB1034::Φ(*zwf′-gnd′-′lacZY′*)(−76 to −42::*Kpn*I::−36 to +290)(Hyb)(PQ^+)	This study
pZ4G4	477-bp PCR product (Zwf5′-3 and Q4 on pZ3G4) as a 369-bp *Eco*RI-*Bam*HI fragment in pMLB1034; pMLB1034::Φ(*zwf′-gnd′-′lacZY′*)(−67 to −42::*Kpn*I::−36 to +290)(Hyb)(PQ^+)	This study
pZ5G4[b]	472-bp PCR product (Zwf5′-4 and Q4 on pZ3G4) as a 364-bp *Eco*RI-*Bam*HI fragment in pMLB1034; pMLB1034::Φ(*zwf′-gnd′-′lacZY′*)(−62 to −42::*Kpn*I::−36 to +290)(Hyb)(PQ^+)	This study
5′ deletions of *sodA-lacZ* and *nfo-lacZ* protein fusions		
pUC18*sodA*-1	204-bp PCR product (SodA5′-1 and SodA3′-1 on pSOD-1) as a 196-bp *Kpn*I-*Bam*HI fragment in pUC18 (−89 to +98)	This study
p*sodA*-1	208-bp *Eco*RI-*Bam*HI fragment from pUC18*sodA*-1 in pMLB1034; pMLB1034::Φ(*sodA′-′lacZY′*)(−89 to +98)(Hyb)(PQ^+)	This study
p*sodA*-2	169-bp PCR product (SodA5′-2 and SodA3′-1 on pSOD-1) as a 162-bp *Eco*RI-*Bam*HI fragment in pMLB1034; pMLB1034::Φ(*sodA′-′lacZY′*)(−55 to +98)(Hyb)(PQ^+)	This study
p*nfo*-1	197-bp PCR product (Nfo5′-1 and Nfo3′-1 on pIT3) as a 192-bp *Eco*RI-*Bam*HI fragment in pMLB1034; pMLB1034::Φ(*nfo′-′lacZY′*)(−103 to +78)(Hyb)(PQ^+)	This study
p*nfo*-2	149-bp PCR product (Nfo5′-2 and Nfo3′-1 on pIT3) as a 144-bp *Eco*RI-*Bam*HI fragment in pMLB1034; pMLB1034::Φ(*nfo′-′lacZY′*)(−58 to +78)(Hyb)(PQ^+)	This study

[a] Numbers in parentheses refer to sequence positions with respect to the proposed transcriptional start sites. PQ^+, paraquat inducible. PQ^-, paraquat uninducible.
[b] Named ZG5 in reference 11.

sequence features (the −35 and −10 hexamers of the *zwf* promoter, the transcriptional and translational start sites, the Shine-Dalgarno sequences, and selected restriction sites) are indicated.

Basis for the deletion analysis of *zwf*. Prior to initiating the deletion analysis described below, we attempted to define in vitro the binding site on *zwf* for the native SoxS protein (10). Using a partially purified extract from an *E. coli* strain overproducing SoxS, we performed gel mobility shift assays on the isolated *Eco*RI-*Bam*HI fragment of Z1 (10). These initial experiments indicated that the extract contained a protein, presumably SoxS, that recognized *zwf* sequences between the *Bss*HII and *Alu*I sites (Fig. 1) (10). In addition, preliminary DNase I protection assays revealed a hypersensitive cleavage site near the *Bss*HII site and an area of protection extending from there toward the promoter (10). Unfortunately, as described elsewhere (11), we were unable to continue these in vitro experiments with the native SoxS protein because of its insolubility. Nevertheless, these preliminary studies did suggest possible boundaries for SoxS binding that we subsequently investigated by deletion analysis.

Delineation of the 5′ boundary of the SoxS activation site. Since our initial in vitro experiments had indicated that the *cis*-acting site for SoxS binding of *zwf* was probably upstream of the *zwf* promoter, we first determined the 5′ boundary of the activation site. The endpoints of the 5′ deletions are depicted in Fig. 1, and the effects of these deletions on paraquat-inducible β-galactosidase expression are shown in Table 3.

Our first effort was to determine whether a *zwf-lacZ* fusion deleted for all *zwf* sequences upstream of the *zwf* promoter retained paraquat inducibility. As expected from the preliminary in vitro experiments, this minimal *zwf-lacZ* fusion, Z9, was uninducible when challenged with paraquat. However, when the upstream *zwf* sequences to position −140 were restored (Z1K), the β-galactosidase activity was induced 7.1-fold. Although the extent of paraquat inducibility of Z1K was the same as the 6.8-fold induction observed for Z1, we did note that its overall β-galactosidase activity level was 50% lower than that of Z1 (Table 3). We attribute this lower-level activity to the presence of the *Kpn*I site that we had constructed upstream of the promoter.

We then prepared a set of deletions with 5′ endpoints at −86 (Z2), −76 (Z3K), −67 (Z4 and Z4K), −62 (Z5), −57 (Z6), −55 (Z7), and −51 (Z8) and determined the effects of the deletions on paraquat-inducible β-galactosidase expression (Fig. 1 and Table 3). The data show that the 5′ boundary of the site for SoxS-mediated activation lies between positions −62 and −57, as defined by the *zwf-lacZ* fusions Z5 and Z6 (Table 3).

Delineation of the 3′ boundary of the SoxS activation site. The experiments described above left open the possibility that sequences within the *zwf* promoter or downstream of it were required for SoxS binding and transcriptional activation. Therefore, starting with construct Z1, we constructed a series of deletions with 3′ endpoints at +60 (Z1G1), +32 (Z1G2), and +2 (Z1G3) (Fig. 1). Then, we placed the resulting *zwf* segments upstream of appropriate 5′ deletions of the *gnd-lacZ* protein fusion CMP78(+) (Fig. 2). The data in Table 4 show that *zwf* sequences downstream of +2 are not required for paraquat-inducible gene expression. We further tested these same 3′ *zwf* deletions by deleting *zwf* sequences upstream of position −67 (constructs Z4G1, Z4G2, and Z4G3; Fig. 2). As expected, these three constructs remained paraquat inducible (Table 4).

To facilitate the construction of deletions defining the 3′ boundary of the SoxS regulatory site, we prepared construct 1034*gnd*. This *gnd-lacZ* fusion is similar to the *zwf-lacZ* fusion Z9 in that all *gnd* sequences upstream of the *gnd* promoter have been deleted and a *Kpn*I site has been placed adjacent to the −35 hexamer (Fig. 2). As expected, β-galactosidase expression from this control protein fusion is not induced by paraquat (Table 4). We then placed *zwf* sequences −76 to −42 (Z3G4), −67 to −42 (Z4G4), and −62 to −42 (Z5G4) upstream of the minimal *gnd* promoter of construct 1034*gnd* (Fig. 2). Importantly, the presence of the upstream *zwf* sequences rendered

TABLE 2. Oligonucleotide primers for PCR and sequencing

Oligonucleotide primer	Sequence[a]	5′ position	3′ position
For PCR			
Zwf5′-1	5′-TAACA**TGATCA**GTGTCAG-3′	−146	−129
Zwf5′-2	5′-GG**GGTACC**GCACTTTGCGCGCTTT-3′	−86	−69
Zwf5′-3	5′-CCG**GAATTC**CCCGTAATCGCACGGGTG-3′	−67	−50
Zwf5′-4	5′-CCC**GAATTC**AATCGCACGGGTGGATAA-3′	−62	−45
Zwf5′-5	5′-CCC**GAATTC**ACGGGTGGATAAGCGTTT-3′	−57	−39
Zwf5′-6	5′-CCC**GAATTC**GGGTGGATAAGCGTT-3′	−55	−40
Zwf5′-7	5′-CCG**GAATTC**TGGATAAGCGTTTACAGT-3′	−51	−34
Zwf5′-8	5′-GCG**GGTACC**GTTTTCGCAAGCTCG-3′	−35	−21
Zwf3′-2	5′-AGCT**GGATCC**CCTTTCGCGCCGAAAATG-3′	+118	+101
Zwf3′-3	5′-CCG**CCATGG**CATTCTCCTTAAGTT-3′	+60	+46
Zwf3′-4	5′-CGC**TTCGAA**CTTAAGCCAGGGTATAC-3′	+33	+16
Zwf3′-5	5′-CGC**TTCGAA**CGGTGCACTGTACTGCTT-3′	+2	−17
Zwf3′-6	5′-GAAAAC**GGTACC**CGCTTATCCACCCGTGCG-3′	−42	−59
Zwf3′-7	5′-CCC**GGTACC**ATCCACCCGTGCGATTAC-3′	−47	−64
Zwf3′-8	5′-CCC**GGTACC**CCCGTGCGATTACGGGAA-3′	−52	−69
Gnd5′-1	5′-CCG**CCATGG**CCAAGCAACAGATCGGCG-3′	+62	+78
Gnd5′-2	5′-CGC**TTCGAA**CATTCCAGGCCGCGA-3′	−1	+18
Gnd5′-3	5′-GGAT**GGTACC**TTTATACTTTAATAAGGA-3′	−36	−19
Nfo5′-1	5′-CCC**GAATTC**TGTTTGCCTGGCTGATTG-3′	−103	−85
Nfo5′-2	5′-CCC**GAATTC**AAAGCGTCATCGCAT-3′	−58	−40
Nfo3′-1	5′-GG**GGATCC**GCCAGACCGCCAGCACAGC-3′	+78	+60
SodA5′-1	5′-CCCG**GGTACC**TCGGGCATTTTCCTGCAAAA-3′	−89	−70
SodA5′-2	5′-CCC**GAATTC**AAAGTACGGCATTGATAA-3′	−55	−38
SodA3′-1	5′-GTGC**GGATCC**AGGGCATCGTAAGCATACGG-3′	+98	+76
For sequencing			
pMLB1034-5′	5′-CCTGACGTCTAAGAAACC-3′	−78	−61
pMLB1034-3′	5′-ACGTTGTAAAACGACGGG-3′	+32	+15
pBR322CW	5′-GTATCACGAGGCCCT-3′	−26	−12
Q4(LacZ)	5′-CGGGCCTCTTCGCTA-3′	+121	+107
−40SP	5′-GTTTTCCCAGTCACGAC-3′	−37	−21

[a] *Bam*HI (GGATCC), *Bcl*I (TGATCA), *Eco*RI (GAATTC), *Kpn*I (GGTACC), *Nsp*V (TTCGAA), and *Sty*I (CCATGG) restriction sites are in boldface type.

the *gnd* promoter paraquat inducible, although the extent of induction was only about half of that conferred upon the *zwf* promoter by these sequences (Tables 3 and 4). As discussed below, the lower level of induction conferred by the regulatory sequences on the *gnd* promoter compared with the *zwf* promoter may be due to intrinsic differences between the two promoters.

Up to this point, the deletion analysis had reduced the site for SoxS activation to a 21-bp region upstream of the *zwf* promoter at positions −62 to −42, as defined by construct Z5G4. Although deletions had placed the 5′ boundary between positions −62 and −57, we had not prepared any 3′ *zwf* deletions that affected paraquat inducibility. Thus, returning to fusion Z1K, we constructed two internal deletions of 5 bp (Z1KΔ1) and 10 bp (Z1KΔ2) (Fig. 1). Both of these internal *zwf* deletions prevented induction of β-galactosidase expression by paraquat (Table 4). From these results, we infer that the 3′ boundary of the SoxS activation site lies between positions −47 and −42, as defined by Z1K and Z1KΔ1, or possibly between Z1K and Z1KΔ2 if the helical phasing of the regulatory element is critical to its function.

In summary, the deletion analysis of *zwf* has localized the *cis*-acting site for SoxS activation to a 21-bp region (positions −62 to −42) upstream of the *zwf* promoter. This 21-bp "soxbox" is sufficient for paraquat inducibility when placed upstream of a normally unresponsive *gnd-lacZ* fusion.

5′ deletions of *sodA* and *nfo*. Having defined a tentative soxbox for *zwf*, we wanted to investigate the activation sites of other member genes of the *soxRS* regulon. Accordingly, we constructed *lacZ* protein fusions of *sodA* and *nfo*. The β-galactosidase activities of protein fusions *sodA*-1 and *nfo*-1 were induced 3.6- and 6.8-fold, respectively, by treatment with paraquat (Table 5).

To define better the sequences required for SoxS activation

TABLE 3. Effects of 5′ deletions on the paraquat inducibility of *zwf-lacZ* fusions

Lysogen	Junction endpoint		β-Galactosidase activity (U)		Induction ratio, +PQ/−PQ
	5′	3′	−PQ[a]	+PQ	
HB301(λDR101)	−646	+119	72	460	6.4
HB301(λZwf)	−140	+546	153	873	5.7
HB301(λZ1)	−140	+118	79	537	6.8
HB301(λZ1K)[b]	−140	−42	37	260	7.1
	−35	+118			
HB301(λZ2)	−86	+118	61	443	7.3
HB301(λZ3K)	−76	−42	38	238	6.2
	−35	+118			
HB301(λZ4)	−67	+118	69	475	6.9
HB301(λZ4K)	−67	−42	41	246	6.0
	−35	+118			
HB301(λZ5)	−62	+118	61	517	8.5
HB301(λZ6)	−57	+118	74	72	1.0
HB301(λZ7)	−55	+118	89	90	1.0
HB301(λZ8)	−51	+118	67	65	1.0
HB301(λZ9)	−35	+118	33	33	1.0

[a] PQ, paraquat.

[b] The K designation indicates that the fusion has an engineered internal *Kpn*I site at positions −41 to −36 upstream of the *zwf* promoter (Fig. 1).

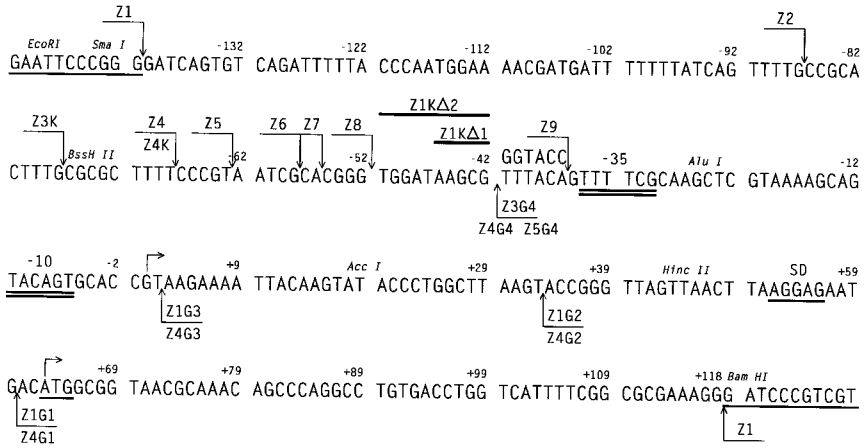

FIG. 1. Nucleotide sequence of *zwf-lacZ* protein fusion Z1. The 286-bp nucleotide sequence of *zwf-lacZ* protein fusion Z1 is presented. The following relevant features are indicated: the −35 and −10 hexamers of the *zwf* promoter (double underline), the transcription and translation start sites (horizontal arrows above the sequence), the Shine-Dalgarno sequence and first codon (underline), and selected restriction sites. The deletion endpoints are identified by the common fusion name. The 5′ deletions are marked by vertical arrows above the sequence. The 3′ deletions are marked by vertical arrows below the sequence. Internal deletions are indicated by the heavy lines above the sequence. In addition, the *Kpn*I site incorporated near the −35 hexamer of the *zwf* promoter of some of the fusions is shown.

of *sodA* and *nfo*, we reasoned, by analogy to *zwf*, that the soxbox positions for these genes would probably lie just upstream of their promoters. For *sodA*, data from two recent reports (8, 15) also enabled us to precisely locate its soxbox. In a deletion study, Compan and Touati (8) reported that a *sodA-lacZ* operon fusion (pIC4-1) which carries an internal deletion of *sodA* sequences from positions −69 to −47 was not paraquat inducible, indicating that part of the SoxS activation site for *sodA* had been removed by this deletion. In addition, Gardner and Fridovich (15) had observed that a 19-bp palindromic sequence near the *sodA* promoter was homologous to promoter sequences of *zwf* and *nfo*. Upon examining the sequence homologies observed by Gardner and Fridovich, we found that the homologous sequence in *zwf* overlapped the 3′ boundary of the soxbox that we had defined. Furthermore, in the alignment of Gardner and Fridovich (15), we noted that the *sodA* internal deletion of Compan and Touati (8) would overlap the 5′ boundary of a similarly placed soxbox for *sodA*. Accordingly,

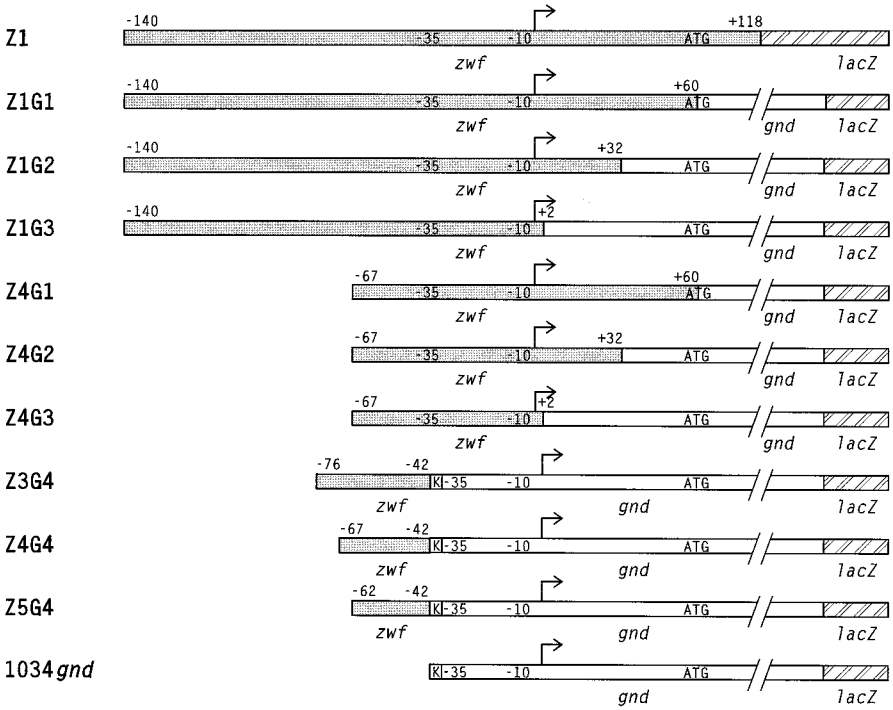

FIG. 2. Schematic representations of the *zwf-gnd-lacZ* protein fusions used to define the minimal *zwf* soxbox. The dark shaded boxes represent *zwf* sequences, while the open and hatched boxes represent *gnd* and *lacZ* sequences, respectively. The positions of the ATG start codons and of the −35 and −10 promoter hexamers are shown within the boxes, and the arrows denote the transcription start points. The boundaries of the *zwf* sequences relative to the start point of *zwf* transcription are given above each construct. The letter K represents the presence of an engineered *Kpn*I site.

TABLE 4. Effects of 3′ deletions on the paraquat inducibility of *zwf-gnd-lacZ* fusions

Lysogen	Junction endpoint		β-Galactosidase activity (U)		Induction ratio, +PQ/−PQ
	5′	3′	−PQ[a]	+PQ	
HB301(λ1034*gnd*)	NA[b]	NA	63	36	0.6
HB301(λZ1G1)	−140	+60	25	192	7.7
HB301(λZ1G2)	−140	+33	48	340	7.1
HB301(λZ1G3)	−140	+2	165	875	5.3
HB301(λZ4G1)	−67	+60	21	152	7.3
HB301(λZ4G2)	−67	+33	28	156	5.6
HB301(λZ4G3)	−67	+2	134	661	4.9
HB301(λZ3G4)	−76	−42	642	1,313	2.0
HB301(λZ4G4)	−67	−42	650	1,477	2.3
HB301(λZ5G4)[c]	−62	−42	459	1,324	2.9
HB301(λZ1K)[d]	−140	−42	37	260	7.1
	−35	+118			
HB301(λZ1KΔ1)	−140	−47	45	28	0.6
	−35	+118			
HB301(λZ1KΔ2)	−140	−52	60	57	1.0
	−35	+118			

[a] PQ, paraquat.
[b] NA, not applicable.
[c] Z5G4 was named ZG5 in reference 11.
[d] The K designation indicates that the fusion has an engineered internal *Kpn*I site at positions −41 to −36 upstream of the *zwf* promoter (Fig. 1).

we constructed a 5′ deletion of the *sodA* fusion with its endpoint at −55 (*sodA*-2). Similarly, we prepared a 5′ deletion of the *nfo* fusion with an endpoint at −58 (*nfo*-2). Figure 3 shows the 5′ ends of constructs *sodA*-2 and *nfo*-2, along with those of constructs Z5 and Z5G4, aligned according to the sequence homologies observed by Gardner and Fridovich (15).

The effects of the 5′ deletions *sodA*-2 and *nfo*-2 on paraquat inducibility were determined (Table 5). The β-galactosidase activity of the *nfo*-2 fusion strain was induced 5.5-fold, about the same extent of induction as that conferred by the shorter *nfo*-1 deletion. Thus, the sequences required for SoxS activation of *nfo* lie downstream of position −58. For the *sodA*-2 strain, the extent of induction was 1.5-fold. Although this level of induction is about 50% of that conferred by the shorter *sodA*-1 deletion, the fact that the *sodA*-2 construct is still inducible by paraquat implies that most or all of the sequences required for SoxS transcriptional activation lie downstream of −55. As discussed below, the 5′ boundaries for SoxS activation of *nfo* and *sodA* expression in vivo are consistent with the location of the primary MalE-SoxS binding sites in vitro (11).

DISCUSSION

Taking advantage of PCR and recombinant DNA techniques, we have genetically defined the soxbox of the superoxide-regulated *zwf* gene. By making preselected 5′ and 3′ deletions of *zwf* sequences and testing these deletions as protein fusions to *lacZ*, we localized the site for SoxS activation in vivo to a 21-bp region upstream of the *zwf* promoter. This 21-bp element was able to confer paraquat inducibility when placed upstream of the normally uninducible *gnd* promoter (constructs Z3G4, Z4G4, Z5G4, and 1034*gnd* in Fig. 2 and Table 4).

TABLE 5. Effects of 5′ deletions on the paraquat inducibility of *sodA-lacZ* and *nfo-lacZ* fusions

Lysogen	Junction endpoint		β-Galactosidase activity (U)		Induction ratio, +PQ/−PQ
	5′	3′	−PQ[a]	+PQ	
HB301(λ*sodA*-1)	−89	+98	904	3,245	3.6
HB301(λ*sodA*-2)	−55	+98	367	562	1.5
HB301(λ*nfo*-1)	−103	+78	24	164	6.8
HB301(λ*nfo*-2)[b]	−58	+78	63	346	5.5

[a] PQ, paraquat.
[b] This strain is a dilysogen, whereas the other strains are monolysogens.

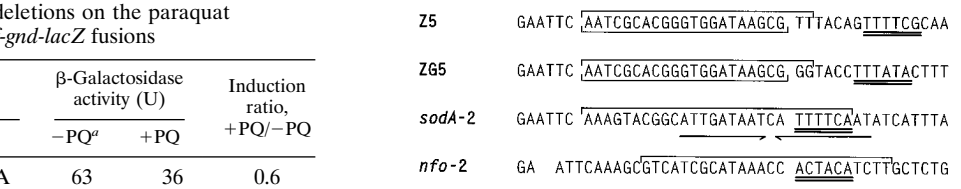

FIG. 3. Coincidence between the genetically defined boundaries of *zwf*, *sodA*, and *nfo* for SoxS transcriptional activation in vivo and the primary binding sites of MalE-SoxS in vitro. The 5′ deletion endpoints of constructs Z5, Z5G4, *sodA*-2, and *nfo*-2 are shown. In each case, an *Eco*RI restriction site is at the junction of plasmid and target gene sequences. The sequences are aligned on the basis of the 19-bp palindromic sequence homology observed by Gardner and Fridovich (15), which is indicated by the inverted horizontal arrows below the sequence of *sodA*-2. The locations of the −35 hexamers of the *zwf* (27), *gnd* (24), *sodA* (38), and *nfo* (30) promoters were previously determined by transcript mapping and are indicated by the double underlines below the sequences. Brackets above the sequences mark the primary sites for the fusions that were protected by MalE-SoxS from DNase I attack (11). Brackets below the sequences of Z5 and Z5G4 mark the minimal soxbox identified by the paraquat inducibility experiments.

The extent of induction of the *zwf-gnd-lacZ* fusions Z3G4, Z4G4, and Z5G4 was only two- to threefold (Table 4), compared with the five- to eightfold induction obtained with the related constructs that contain the *zwf* promoter, e.g., Z4 (Table 3) and Z4G3 (Table 4). This difference in the extent of induction may be attributed to either the specific location of the soxbox with respect to the *gnd* promoter or to intrinsic differences between the *zwf* and *gnd* promoters. Thus, for example, the rate-limiting step of transcription initiation may differ between the two promoters, and SoxS may primarily activate the step rate limiting in *zwf* transcription initiation. Alternatively, although the orientation of the soxbox relative to the −35 hexamer of the *gnd* promoter in constructs Z3G4, Z4G4, and Z5G4 is virtually identical to that of the native soxbox relative to the *zwf* promoter, the spacing requirement for maximum induction may differ between different promoters.

The minimal *zwf* soxbox and the 5′ boundaries of the SoxS activation sites of *sodA* and *nfo* defined by the genetic analyses presented here are consistent with the respective primary SoxS binding sites identified in vitro (11, 20). Recently, we reported the purification of a MalE-SoxS fusion protein and its use in the in vitro identification of the primary SoxS binding sites of *zwf* and four other member genes of the *soxRS* regulon (11). In particular, we showed that MalE-SoxS protects specific regions at or near the promoters of *zwf* (positions −62 to −40), *sodA* (positions −55 to −29), and *nfo* (positions −49 to −25) from DNase I attack (11). These sites protected by MalE-SoxS in vitro are illustrated in Fig. 3. For *zwf*, the genetically defined minimal soxbox lies wholly within the region protected by our purified MalE-SoxS fusion protein (11); it also coincides with the site identified by the footprinting experiments of Li and Demple (20), which used purified native SoxS protein. Similarly, the primary protected regions of *sodA* and *nfo* are downstream of the 5′ boundaries established by the 5′ deletions of constructs *sodA*-2 and *nfo*-2 (Fig. 3). Thus, the in vivo boundaries for SoxS-mediated transcriptional activation reported here agree very well with the MalE-SoxS binding sites identi-

fied in vitro. For this reason, it is not clear why the extent of induction produced by the *sodA*-1 deletion is twofold higher than that of the *sodA*-2 deletion (Table 5); as discussed below, this may be because of the multiple transcription factors that act on *sodA* (8).

The previous DNase I protection assays also revealed secondary SoxS binding sites upstream of the *zwf* (11, 20) and *sodA* promoters that could be required for transcriptional activation in vivo (11). No secondary binding site was observed in the *nfo* promoter region (11). For *zwf*, the upstream secondary site is between positions −87 and −70 (11, 20), while the *sodA* secondary site is between positions −87 and −65 (11). The genetic work presented here clearly suggests that the *zwf* secondary site is not required for SoxS activation, since construct Z5 lacks the entire secondary site but is fully paraquat inducible (Table 3).

Although the *sodA* secondary site is completely deleted in construct *sodA*-2, it is not possible at this time to decide unequivocally whether the site plays a role in SoxS activation, because the phenotype conferred by the deletion is complex. On the one hand, the *sodA*-2 construct is still partially induced by paraquat and a deletion mutant of Compan and Touati (8) remained fully inducible despite being deleted of sequences upstream of −75. However, on the other hand, the basal expression level of *sodA*-2 is lower than that of *sodA*-1, as is the extent of induction. Thus, the intermediate phenotype could be because the *sodA* promoter region appears to be capable of binding at least six transcription factors. Accordingly, in the case of *sodA*, it may be inappropriate to base expectations for in vivo effects on the binding of a purified protein to linear DNA in vitro (8). A useful approach for addressing some of these issues would be to determine the effects of the various deletions on transcriptional activation by MalE-SoxS in a minimal in vitro transcription system.

The set of constructs that genetically define the site of action of SoxS described here may prove useful in studies aimed at characterizing the DNA binding sites of the proteins most closely related to SoxS. SoxS is a member of the AraC subfamily of helix-turn-helix proteins (1, 13, 41). Within this subfamily, the proteins encoded by *soxS* (1, 41), *tetD* (4, 31), *marA* (6, 7), and *rob* (37) are highly homologous to one another over the entire 107-amino-acid length of SoxS, with the extent of homology ranging from 42% identity for the SoxS-MarA pair to 59% identity for the TetD-Rob pair. Even more striking is the fact that among the four proteins, at least three have the identical amino acid at 17 of the 21 positions of the respective helix-turn-helix motifs and at 10 of the 11 positions of the turn and the second, recognition helix. Thus, it is highly likely that the putative DNA binding sites recognized by the respective proteins will be found to have nearly identical sequences. Moreover, if all four proteins prove to be transcriptional activators (the function of TetD is unknown [35], Rob was isolated on the basis of its binding to the right border of the *E. coli* replication origin [37], and MarA is the positive regulator of multiple antibiotic resistance [14]), there should be a considerable amount of cross talk among them. Indeed, 2,4-dinitrophenol, which is an effector of the *mar* operon (7), has been known for many years to induce glucose 6-phosphate dehydrogenase expression (32), and the *soxQ1* mutation, which was initially isolated on the basis of conferring resistance to menadione (16), a redox-cycling compound, is a regulatory mutation of the *marAB* operon (2) that leads to transcriptional activation of some (e.g., *zwf*) but not all (e.g., *nfo*) member genes of the SoxRS regulon (16). Thus, according to these arguments, it should be possible to determine the degree of coincidence between the *zwf* soxbox and the DNA binding/transcriptional activation sites of MarA, TetD, and Rob by testing whether the overproduction of the respective proteins elevates β-galactosidase expression in vivo from our set of *zwf-lacZ* and *zwf-gnd-lacZ* fusions. Moreover, it would be interesting to purify the MarA, TetD, and Rob proteins and determine their respective abilities to activate in vitro transcription of the fusion constructs, as we have already done with our purified MalE-SoxS fusion protein (19).

ACKNOWLEDGMENTS

We thank H. M. Steinman and B. Weiss for their kind gifts of plasmids pSOD-1 and pIT3, respectively, and R. Cunningham for providing us with unpublished *nfo* sequence. We also thank K. W. Jair for help with the preparation of some of the *lacZ* fusions.

This research was supported by Public Health Service grant GM27113 from the National Institutes of Health.

REFERENCES

1. **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.
2. **Ariza, R. R., S. P. Cohen, N. Bachhawat, S. B. Levy, and B. Demple.** 1994. Repressor mutations in the *marAB* operon that activate oxidative stress genes and multiple antibiotic resistance in *Escherichia coli*. J. Bacteriol. **176:**143–148.
3. **Ausubel, F. M., R. Brent, R. E. Kingston, D. D. Moore, J. G. Seidman, J. A. Smith, and K. Struhl (ed.).** 1987. Current protocols in molecular biology. John Wiley & Sons, Inc., New York.
4. **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.
5. **Carter-Muenchau, P., and R. E. Wolf, Jr.** 1989. Growth-rate-dependent regulation of 6-phosphogluconate dehydrogenase level mediated by an anti-Shine-Dalgarno sequence located within the *Escherichia coli gnd* structural gene. Proc. Natl. Acad. Sci. USA **86:**1138–1142.
6. **Cohen, S. P., H. Hächler, and S. B. Levy.** 1993. Genetic and functional analysis of the multiple antibiotic resistance (*mar*) locus in *Escherichia coli*. J. Bacteriol. **175:**1484–1492.
7. **Cohen, S. P., S. B. Levy, J. Foulds, and J. L. Rosner.** 1993. Salicylate induction of antibiotic resistance in *Escherichia coli*: activation of the *mar* operon and a *mar*-independent pathway. J. Bacteriol. **175:**7856–7862.
8. **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:**1687–1696.
9. **Csonka, L., and D. G. Fraenkel.** 1977. Pathways of NADPH formation in *Escherichia coli*. J. Biol. Chem. **152:**3382–3391.
10. **Fawcett, W. P., and R. E. Wolf, Jr.** 1991. Characterization of the regulatory sites for oxidative stress control of *Escherichia coli* glucose 6-phosphate dehydrogenase expression, p. 28. *In* Molecular and cellular responses to oxygen, The Albany Conferences, Albany, N.Y.
11. **Fawcett, W. P., and R. E. Wolf, Jr.** 1994. Purification of a MalESoxS fusion protein and identification of the control sites of *Escherichia coli* superoxide inducible genes. Mol. Microbiol. **14:**669–679.
12. **Fraenkel, D. G.** 1987. Glycolysis, pentose phosphate pathway, and Enter-Doudoroff pathway, p. 142–150. *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.
13. **Gallegos, M. T., C. Michán, and J. L. Ramos.** 1993. The XylS/AraC family of regulators. Nucleic Acids Res. **21:**807–810.
14. **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.
15. **Gardner, P. R., and I. Fridovich.** 1993. NADPH inhibits transcription of the *Escherichia coli* manganese superoxide dismutase gene (*sodA*) *in vitro*. J. Biol. Chem. **268:**12958–12963.
16. **Greenberg, J. T., J. H. Chou, P. A. Monach, and B. Demple.** 1991. Activation of oxidative stress genes by mutations at the *soxQ*/*cfxB*/*marA* locus of *Escherichia coli*. J. Bacteriol. **173:**4433–4439.
17. **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.
18. **Greenberg, J. T., P. Monach, J. H. Chou, 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.
19. **Jair, K. W., and R. E. Wolf, Jr.** Unpublished data.
20. **Li, Z., and B. Demple.** 1994. SoxS, an activator of superoxide stress genes in *Escherichia coli*. J. Biol. Chem. **269:**18371–18377.

21. **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.
22. **Liochev, S. I., and I. Fridovich.** 1992. Effects of overproduction of superoxide dismutases in *Escherichia coli* on inhibition of growth and on induction of glucose-6-phosphate dehydrogenase by paraquat. Arch. Biochem. Biophys. **294:**138–143.
23. **Miller, J. H.** 1972. Experiments in molecular genetics. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
24. **Nasoff, M. S., H. V. Baker II, R. E. Wolf, Jr.** 1988. DNA sequence of the *Escherichia coli* gene, *gnd*, for 6-phosphogluconate dehydrogenase. Gene **27:**253–264.
25. **Pease, A. J., and R. E. Wolf, Jr.** 1994. Determination of the growth rate-regulated steps in expression of the *Escherichia coli* K-12 *gnd* gene. J. Bacteriol. **176:**115–122.
26. **Rowley, D. L., A. J. Pease, and R. E. Wolf, Jr.** 1991. Genetic and physical analyses of the growth rate-dependent regulation of *Escherichia coli zwf* expression. J. Bacteriol. **173:**4660–4667.
27. **Rowley, D. L., and R. E. Wolf, Jr.** 1991. Molecular characterization of the *Escherichia coli* K-12 *zwf* gene encoding glucose 6-phosphate dehydrogenase. J. Bacteriol. **173:**968–977.
28. **Sambrook, J., E. F. Fritsch, and T. Maniatis.** 1989. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
29. **Sanger, F., S. Nicklen, and A. R. Coulson.** 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA **74:**5463–5467.
30. **Saporito, S. M., and R. P. Cunningham.** 1988. Nucleotide sequence of the *nfo* gene of *Escherichia coli* K-12. J. Bacteriol. **170:**5141–5145.
31. **Schollmeier, K., and W. Hillen.** 1984. Transposon Tn*10* contains two structural genes with opposite polarity between *tetA* and IS*10*$_R$. J. Bacteriol. **160:**499–503.
32. **Scott, D. B. M.** 1956. The oxidative pathway of carbohydrate metabolism in *Escherichia coli*. 4. Formation of enzymes induced by 2:4-dinitrophenol. Biochem. J. **63:**593–600.
33. **Shimada, K., R. A. Weisberg, and M. E. Gottesman.** 1972. Prophage lambda at unusual chromosomal locations. J. Mol. Biol. **63:**483–503.
34. **Silhavy, T. J., M. L. Berman, and L. W. Enquist.** 1984. Experiments with gene fusions. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
35. **Simons, R. W. (University of California, Los Angeles).** 1994. Personal communication.
36. **Simons, R. W., F. Houman, and N. Kleckner.** 1987. Improved single and multicopy *lac*-based cloning vectors for protein and operon fusions. Gene **53:**85–96.
37. **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.
38. **Takeda, Y., and H. Avila.** 1986. Structure and gene expression of the *E. coli* Mn-superoxide dismutase gene. Nucleic Acids Res. **14:**4577–4589.
39. **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.
40. **Wolf, R. E., Jr., D. M. Prather, and F. M. Shea.** 1979. Growth-rate-dependent alteration of 6-phosphogluconate dehydrogenase and glucose 6-phosphate dehydrogenase levels in *Escherichia coli* K-12. J. Bacteriol. **139:**1093–1096.
41. **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.