THE JOURNAL OF BIOLOGICAL CHEMISTRY
© 1990 by The American Society for Biochemistry and Molecular Biology, Inc.

Vol. 265, No. 7, Issue of March 5, pp. 3844-3850, 1990
Printed in U.S.A.

Use of Saturation Mutagenesis to Localize Probable Functional Domains in the NahR Protein, a LysR-type Transcription Activator*

(Received for publication, July 24, 1989)

Mark A. Schell‡, Pamela H. Brown, and Satanaryana Raju

From the Department of Microbiology, University of Georgia, Athens, Georgia 30602

The NahR protein of the *Pseudomonas* naphthalene degradation plasmid NAH7 encodes a 300-residue transcription activator which is very similar to the NodD transcription activator of *Rhizobium* and other proteins in the LysR activator family. NahR binds to conserved sequences upstream (nucleotides −80 to −47) of the *nah* and *sal* promoters and activates transcription of genes for naphthalene catabolism in response to the inducer salicylate. Transformation of an *Escherichia coli gal* deletion strain (containing a *sal* promoter-*galK* fusion plasmid) with hydroxylamine-treated *nahR* DNA and selection on galactose/salicylate plates allowed isolation of 30 unique activation-deficient *nahR* alleles which fell into two classes: class I, defective in both activation and specific binding to the NahR activation site of the *sal* promoter; and class II, defective in activation, but with wild-type DNA binding activity. DNA sequence analysis showed that the amino acid substitutions eliminating DNA binding activity were mostly clustered in an NH_2-terminal helix-turn-helix motif (residues 23–45) or a COOH-terminal domain (residues 239–291). Similar analysis of class II mutants identified a domain (residues 126–206) possibly involved in inducer binding and/or transcription activation functions. The partial *trans*-dominance of many mutant alleles and the size of NahR-specific DNA binding activity measured by gel filtration suggest that the active NahR protein may be a tetramer.

The *nahR* gene encodes a 34-kDa polypeptide which activates transcription of two operons on plasmid NAH7 of *Pseudomonas putida* (1, 2) in response to the inducer salicylate. The two operons controlled by *nahR* (*nah* and *sal*) encode 13 enzymes necessary for catabolism of naphthalene or salicylate as sole carbon and energy source (3, 4). The *nah* and *sal* (P_{sal})[1] promoters have similar DNA sequences (5), and fusion of either promoter to the *Escherichia coli galK* gene produces *E. coli* strains which show a 30-fold induction of *galK* expression only in the presence of both *nahR* and salicylate (1). The NahR protein protects a homologous region of both promoters (between nucleotides −80 to −47) from DNase I digestion; mutagenesis experiments identified three nucleotides in this region (at positions −74, −73, and −61) required for NahR binding and/or activation (6).[2] Whereas NahR binding to the region at nucleotides −80 to −47 is required for transcription activation, NahR binding does not appear to require the inducer salicylate (6).

A data base search using the NahR open reading frame identified NahR as a member of the LysR family of prokaryotic transcription activators (7). Members of this family of almost identically sized transcription activators (including LysR, NodD, IlvY, AmpR, etc.) show major amino acid sequence similarities in their NH_2-terminal halves; all the activators are transcribed from promoters that are very close to and divergent from the promoters of one set of the structural genes they control (8). Since the NahR amino acid and DNA sequence is 45% similar to that of the NodD transcription activator of *Rhizobium meliloti* and appears to recognize a similar sequence, it was proposed that *nodD* and *nahR* may have evolved from a common ancestor (7).

The NH_2-terminal regions of NodD, NahR, and LysR which have the greatest amino acid similarity (residues 23–45) show evidence of a helix-turn-helix (HTH) DNA binding motif (7, 8) similar to those of other Cro-like DNA-binding proteins (9–11). In addition, portions of the DNA-binding sites of NahR, NodD, and Cro show major homologies, suggesting that NahR and NodD (and maybe other LysR-like proteins) utilize a Cro-like mechanism to bind to DNA. However, direct evidence for involvement of the putative NahR or NodD HTH motifs in DNA binding or transcription activation is lacking. Identification of any homologous domains of NodD and NahR possibly involved in transcription activation is also lacking since significant amino acid sequence homologies between these and other activators (*e.g.* cAMP receptor protein, AraC, etc.) are absent. In contrast to cAMP receptor protein, most evidence suggests that the amount of NahR or NodD bound to their promoter target sites is not greatly affected by the presence of their inducers. Thus, the apparent role of inducer may not be to stimulate binding to the promoter, but rather to convert the bound activator into a form which can promote transcription (6, 12, 13).

The available data suggest that NodD and NahR utilize a similar mechanism for inducer-responsive transcription activation. To identify and define important functional domains in these activators, we isolated and characterized 30 unique activation-negative NahR mutants. Our results identified two domains probably involved in DNA binding and one domain involved in inducer-responsive transcription activation. These domains are apparently present in many LysR family members and thus may be of widespread significance.

MATERIALS AND METHODS

Plasmids and Bacterial Strains Used—Plasmid pMS104 (2) has a 3.5-kilobase fragment of NAH7 DNA containing the entire *nahR* gene

* This work was supported by Grant GM-32255-07 from the National Institutes of Health. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

‡ To whom correspondence should be addressed: Dept. of Microbiology, University of Georgia, 816 Biological Science Bldg., Athens, GA 30602. Tel.: 404-542-0512.

[1] The abbreviations used are: P_{sal}, *sal* promoter; HTH, helix-turn-helix; Ap^r, ampicillin-resistant; Tc^r, tetracycline-resistant.

[2] S. Smith and M. A. Schell, unpublished data.

inserted into the *Pst*I site of the tetracycline-resistant (Tc^r) vector pRK404 (14). Plasmid pMS160 (1) contains a 110-base pair *Hae*III-*Rsa*I fragment with the *sal* promoter sequences nucleotides −85 to +27) transcriptionally fused to the *E. coli galK* gene on an ampicillin-resistant (Ap^r) derivative of plasmid pKO1 (15). Plasmid pWS22 was derived from pMS15 (1) by deletion of the NahR-binding site sequences between the *Eco*RI and *Eag*I sites utilizing end filling with Klenow fragment of DNA polymerase and self-ligation. This resulted in fusion of the *lac* promoter of pUC8 to nucleotide +5 of the *nahR* promoter; cells containing this plasmid showed isopropylthiogalactoside-inducible synthesis of NahR. Plasmids pMG250 and pMAS510 were constructed by insertion of an *Xba*I termination linker (New England BioLabs 1062) into the *Sal*I and *Ssp*I sites of the *nahR* coding region, respectively.

Hydroxylamine Mutagenesis—Ten μg of pMS104 DNA was incubated for 15 h at 37 °C in 0.25 ml of 0.1 M hydroxylamine + 0.1 M potassium phosphate buffer (pH 6.0) freshly prepared as described by Davis *et al.* (16). The DNA was dialyzed twice against 1000 volumes of TE buffer (10 mM Tris-HCl (pH 7.0), 2 mM EDTA) for 4 h followed by phenol/$CHCl_3$ extraction and ethanol precipitation.

Selection of nahR Mutants—Plasmid pMS160 (Ap^r) containing P_{sal} transcriptionally fused to the *E. coli galK* gene was transformed into the *gal* operon deletion strain *E. coli* S165. Fifteen 0.02-ml aliquots of *E. coli* S165 (pMS160) competent cells were separately transformed with 0.2 μg of hydroxylamine-treated pMS104 DNA (Tc^r); and after a short recovery period (15 min, 37 °C), *nahR* mutants were selected by plating on LB plates (17) containing 0.1% galactose, 0.03% salicylate, 50 μg/ml ampicillin, and 25 μg/ml tetracycline. Transformants containing wild-type *nahR* cannot grow on these plates (see "Results"). Overall transformation efficiency (to Tc^r) was reduced 10–20-fold by hydroxylamine treatment.

Analysis of DNA Binding Activity—The ability of NahR proteins to specifically bind to P_{sal} was measured by gel retardation assay utilizing a ^{32}P-labeled *Eco*RI fragment of pMG1 (6) which contains P_{sal} sequences from nucleotides −175 to +26. The methods for preparation of cell-free extracts from *E. coli* strains, labeled pMG1 substrate, conditions for incubation of extracts, and analysis of binding by gel electrophoresis were described previously (6).

Assay for Transcription Activation—*E. coli* S165 (pMS160) strains containing each mutant allele of *nahR* were grown in the presence or absence of salicylate, and cell-free extracts were prepared and then assayed for galactokinase activity. Methods for growth, preparation of extracts, and galactokinase assay were described previously (1).

Assay for trans-Dominance—Each plasmid with a unique mutant *nahR* allele (designated pMASn) was transferred from *E. coli* S165 (pMS160) to *P. putida* 277 (NAH7) by triparental mating with *E. coli* HB101 (pRK2013) followed by selection for *P. putida* (NAH7) (pMASn) strains at 30 °C on BSM/glucose plates with 30 μg/ml each tetracycline and tryptophan. Methods and media for conjugal transfer were described previously (2). Cell-free extracts of *P. putida* 277 (NAH7) (pMASn) merodiploids were prepared from cells grown in LB medium at 30 °C in the presence and absence of 0.05% salicylate and assayed for the levels of catechol 2,3-dioxygenase (encoded by *nahH*) by the method of Feist and Hegeman (18). Protein in extracts was determined by the method of Kalb and Bernlohr (19).

DNA Sequencing—The entire DNA sequence of each *nahR* allele (from −50 to 30 bp downstream of the termination codon) was determined on one strand by a modification of the dideoxy chain termination method of Sanger *et al.* (20) using ^{35}S-labeled nucleotides and *nahR*-specific 17-mer oligonucleotides synthesized on an Applied Biosystems Model 380A DNA synthesizer to prime synthesis. Some mutant alleles were first cloned in two overlapping contiguous fragments (*Hin*dIII-*Hin*dIII and *Pst*I-*Bgl II*) into the phagemid vector pTZ19U (21, 22); after infection with a M13KO7 helper phage, single-stranded template was prepared from JM107 cells and sequenced with Klenow fragment of DNA polymerase I as described previously (7). Otherwise, CsCl-purified double-stranded plasmid DNA was sequenced with primers and Sequenase 2.0 as recommended by the manufacturer (United States Biochemical Corp., Cleveland, OH).

Maxicell Analysis—Cells of *E. coli* SK6501, a derivative of RYC1000 (23), containing mutant *nahR* plasmids were UV-irradiated, recovered, and labeled with [^{35}S]methionine by a modification of the method of Sancar *et al.* (24). Labeled proteins in maxicells were analyzed on sodium dodecyl sulfate-10% polyacrylamide gels prepared and run as described by Laemmli (25) followed by fluorography by the method of Laskey and Mills (26).

Molecular Biological Techniques—Procedures for restriction enzyme digestions, ligations, transformations, electrophoresis, plasmid DNA isolation, and electroelution were standard and have been described previously (1, 6, 17). DNA enzymes were used according to manufacturers' instructions. Antibiotic levels utilized to select and maintain transformants were 100 μg of ampicillin and 25 μg of tetracycline/ml.

Chemicals—All molecular biology reagents were purchased from Bethesda Research Laboratories, New England BioLabs, or Boehringer Mannheim. Media components were from Difco. T4 DNA polymerase and nucleoside triphosphates were from Pharmacia LKB Biotechnology Inc. Antibiotics, electrophoresis chemicals, and other chemicals were from Sigma or were reagent-grade.

RESULTS

Isolation of nahR Mutants—*E. coli* cells containing plasmid pMS104 (with *nahR*) and plasmid pMS160 (with *galK* transcriptionally fused to P_{sal}) produce high levels of galactokinase activity only when grown with the inducer salicylate (1, 6) (Table I). *E. coli* cells containing high levels of galactokinase but low levels of the *galT* gene product (galactose-1-phosphate uridylyltransferase) will accumulate inhibitory levels of galactose-1-PO_4 when grown in the presence of galactose (15). Accordingly, we found that *E. coli gal* deletion strain S165 containing pMS160 and pMS104 could form colonies on LB plates with 0.1% galactose, but not on LB plates with 0.1% galactose and 0.02% salicylate. However, if the S165 (pMS160) cells contained a mutant *nahR* allele defective in salicylate-induced activation of *galK* expression from P_{sal}, it would be expected that they could form colonies on galactose/salicylate plates.

When *E. coli* S165 (pMS160) was transformed with hydroxylamine-mutagenized pMS104 DNA, 1.5% of the total Ap^r/Tc^r transformants grew well on galactose/salicylate plates, whereas less than 0.1% of the Ap^r/Tc^r colonies from transformations with untreated pMS104 DNA grew on the same selection plates. Over 50 galactose/salicylate-resistant colonies, presumably containing unique mutant *nahR* alleles, were isolated from transformations with hydroxylamine-mutagenized pMS104 DNA. When the putative mutants were assayed for salicylate-induced levels of galactokinase, none showed more than a 2-fold increase in galactokinase activity when grown under inducing conditions, whereas cells with wild-type *nahR* showed a 20-fold increase in response to salicylate (Tables I and II). Uninduced levels of *galK* expression were the same in both wild type and mutants. These results demonstrate the effectiveness of this selection procedure for isolation of mutant *nahR* alleles defective in salicylate-induced transcription activation of P_{sal}.

Analysis of Specific DNA Binding Activity of Mutant NahR Proteins—Plasmid DNA was prepared from each transformant, and the mutant pMS104 derivatives (pMASn) were separated from pMS160 by transformation into *E. coli* SK6501 and selection for Ap^s/Tc^r transformants. Cell-free extracts were prepared from each, and the relative specific DNA binding activity of each mutant NahR protein was measured by gel retardation assay with a ^{32}P-labeled P_{sal} DNA fragment as substrate (Fig. 1). DNA binding assays clearly differentiated two classes of mutants (Tables I and II).

Class I mutants had dramatically lowered binding activity (<10% of wild type) for the NahR-binding site of P_{sal}. We could not detect any specific DNA binding activity for P_{sal} using up to 5 μg of protein extract from cells containing most class I mutant plasmids; extracts from cells with wild-type NahR showed significant binding activity even with 10-fold less extract protein. Maxicell analysis showed that all class I mutant *nahR* alleles tested directed synthesis of NahR proteins that were the same in size and amount as the wild-type allele (Fig. 2). Thus, the phenotype of many class I NahR mutants probably resulted from amino acid substitutions that

TABLE I
Properties of class I NahR mutants

Mutant plasmid	Amino acid change	Salicylate activation of GalK[a]		Promoter binding activity[b]	Inhibition of wild-type allele[c]
		NI	I		
pMS104	Wild type	1.3	23.0	1.0	1.70
pMG250	$Arg^{21} \rightarrow$ stop	1.2	1.4	<0.05	0.85 (0)
pMAS43	$Ala^{27} \rightarrow Thr$	1.5	1.9	<0.05	0.45 (47)
pMAS42	$Thr^{56} \rightarrow Met$	1.6	1.7	<0.05	0.48 (44)
pMAS51	$Arg^{215} \rightarrow Trp$	1.3	1.8	0.10	0.62 (30)
pMAS63	$Arg^{43} \rightarrow His$	1.8	2.0	<0.05	0.43 (50)
pMAS11	$Arg^{45} \rightarrow His$	1.7	1.4	<0.05	0.40 (53)
pMAS222	$Pro^{155} \rightarrow Leu$	1.3	1.6	<0.05	0.70 (17)
pMAS92	$Thr^{26} \rightarrow Ile$	2.5	2.5	<0.05	0.46 (46)
pMAS131	$Thr^{244} \rightarrow Ile$	1.6	1.8	<0.05	0.83 (2)
pMAS101	$Pro^{269} \rightarrow Leu$	1.4	1.9	<0.05	0.82 (2)
pMAS12	$Thr^{239} \rightarrow Ile$	1.8	2.0	<0.05	0.70 (17)
pMAS341	$Arg^{291} \rightarrow Trp$	1.5	2.0	<0.05	0.73 (14)
pMAS170	$Pro^{246} \rightarrow Ser$	1.2	2.3	<0.05	0.80 (6)
pMAS120	$Asp^{105} \rightarrow Asn$	1.1	1.7	0.1	0.72 (15)
pMAS34	$Gln^{85} \rightarrow$ stop	1.5	1.6	<0.05	0.80 (6)
pMAS31	$Gln^{292} \rightarrow$ stop	1.3	1.9	<0.05	0.51 (40)
pMAS244	$Gln^{122} \rightarrow$ stop	1.2	1.3	<0.05	0.70 (15)
pMAS33	$Gln^{168} \rightarrow$ stop	1.6	2.0	<0.05	0.79 (7)
pMAS245	$Gln^{49} \rightarrow$ stop	1.4	1.7	<0.05	0.84 (1)
pMAS510	$His^{278} \rightarrow$ stop	1.3	1.9	<0.05	0.55 (40)

[a] Values are the specific activity ((nanomoles of galactose phosphorylated $minute^{-1}$ $milligram^{-1}$ of protein) × 10^{-2}) of galactokinase (GalK) in *E. coli* S165 (pMS160) with the indicated plasmids from cells grown with (I) or without (NI) salicylate.

[b] DNA binding activity of mutant NahR proteins relative to pMS104 (=1.0) was determined by gel retardation assay with a ^{32}P-labeled P_{sal} fragment and quantified by densitometry of autoradiographs. Values are the average of three determinations with various amounts of protein similar to those in Fig. 1.

[c] Values are the salicylate-induced specific activity (nanomoles of 2-hydroxymuconic semialdehyde formed $minute^{-1}$ $milligram^{-1}$ of protein) of catechol 2,3-dioxygenase in *P. putida* (NAH7) merodiploids containing the indicated plasmids. Values in parentheses are percent inhibition caused by the mutant allele and were calculated as the difference between merodiploid levels and levels in *P. putida* (NAH7) (pMG250) divided by the levels in *P. putida* (NAH7) (pMG250).

TABLE II
Properties of class II NahR mutants

Mutant plasmid	Amino acid change	Salicylate activation of GalK[a]		Promoter binding activity[b]	Inhibition of wild-type allele[c]
		NI	I		
pMS104	Wild type	1.3	23.0	1.0	1.70
pMG250	$Arg^{21} \rightarrow$ stop	1.2	1.6	<0.05	0.85 (0)
pMAS30	$Thr^{130} \rightarrow Met$	2.4	2.9	0.4	0.40 (53)
pMAS40	$Cys^{173} \rightarrow Tyr$	1.3	1.8	0.5	0.48 (44)
pMAS50	$His^{206} \rightarrow Tyr$	1.2	1.2	0.8	0.25 (70)
pMAS60	$Gly^{203} \rightarrow Asp$	1.3	1.8	1.0	0.61 (25)
pMAS61	$Pro^{35} \rightarrow Ser$	1.5	2.2	0.3	0.65 (24)
pMAS94	$Pro^{227} \rightarrow Ser$	1.4	2.4	1.0	0.49 (44)
pMAS151	$Thr^{204} \rightarrow Ile$	1.4	2.0	0.3	0.22 (76)
pMAS256	$Cys^{253} \rightarrow Tyr$	1.3	2.3	1.0	0.20 (78)
pMAS242	$Gly^{152} \rightarrow Ser$	1.3	1.7	0.3	0.47 (46)
pMAS258	$Gly^{107} \rightarrow Asp$	2.3	2.7	1.0	0.10 (88)
pMAS241	$Cys^{126} \rightarrow Tyr$	1.6	1.9	0.7	0.12 (86)

[a] Values are the specific activity ((nanomoles of galactose phosphorylated $minute^{-1}$ $milligram^{-1}$ of protein) × 10^{-2}) of galactokinase (GalK) in *E. coli* S165 (pMS160) with the indicated plasmids from cells grown with (I) or without (NI) salicylate.

[b] DNA binding activity of mutant NahR proteins relative to pMS104 (=1.0) was determined by gel retardation assay with a ^{32}P-labeled P_{sal} fragment and quantified by densitometry of autoradiographs. Values are the average of three determinations with various amounts of protein similar to those in Fig. 1.

[c] Values are the salicylate-induced specific activity (nanomoles of 2-hydroxymuconic semialdehyde formed $minute^{-1}$ $milligram^{-1}$ of protein) of catechol 2,3-dioxygenase in *P. putida* (NAH7) merodiploids containing the indicated plasmids. Values in parentheses are percent inhibition caused by the mutant allele and were calculated as the difference between merodiploid levels and levels in *P. putida* (NAH7) (pMG250) divided by the levels in *P. putida* (NAH7) (pMG250).

altered residues in the DNA-binding domain(s) of NahR, although some of the substitutions could affect overall protein structure and folding, indirectly leading to loss of DNA binding activity.

All class II mutants, however, produced NahR proteins which still retained a significant portion (>30% of wild type) of their DNA binding activity (Fig. 1 and Table II); this binding activity was still specific for P_{sal} (data not shown).

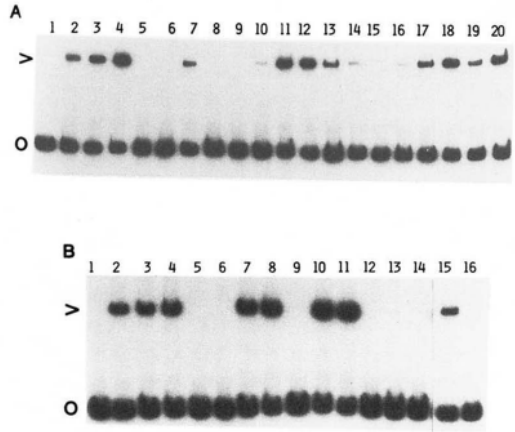

FIG. 1. **Analysis of mutant NahR proteins for specific DNA binding.** *A*, ^{32}P-labeled 250-base pair *Eco*RI fragment containing the *sal* operon promoter sequences between nucleotides −175 and +27 was incubated with 5 μg (unless otherwise indicated) of cell-free extract protein from *E. coli* SK6501 containing the following plasmids. *A, lane 1,* pRK404 vector; *lane 2,* pMS104 (0.5 μg); *lane 3,* pMS104 (1.5 μg); *lane 4,* pMS104; *lane 5,* pMAS92; *lane 6,* pMAS43; *lane 7,* pMAS120; *lane 8,* pMAS11; *lane 9,* pMAS63; *lane 10,* pMAS222; *lane 11,* pMAS258; *lane 12,* pMAS241; *lane 13,* pMAS120; *lane 14,* pMAS51; *lane 15,* pMAS170; *lane 16,* pMAS131; *lane 17,* pMAS61; *lane 18,* pMAS50; *lane 19,* pMAS151; *lane 20,* pMAS60. *B, lane 1,* pRK404 vector; *lane 2,* pMS104 (0.5 μg); *lane 3,* pMS104 (1.5 μg); *lane 4,* pMS104; *lane 5,* pMAS12; *lane 6,* pMAS101; *lane 7,* pMAS30; *lane 8,* pMAS40; *lane 9,* pMAS341; *lane 10,* pMAS94; *lane 11,* pMAS256; *lane 12,* pMAS31; *lane 13,* pMAS42; *lane 14,* pMAS245; *lane 15,* pMAS242; *lane 16,* pMAS33. After incubation at 37 °C for 10-min, samples were analyzed by electrophoresis and autoradiography. Migration positions of the bound retarded promoter fragment (>) and free unbound promoter fragment (O) are indicated.

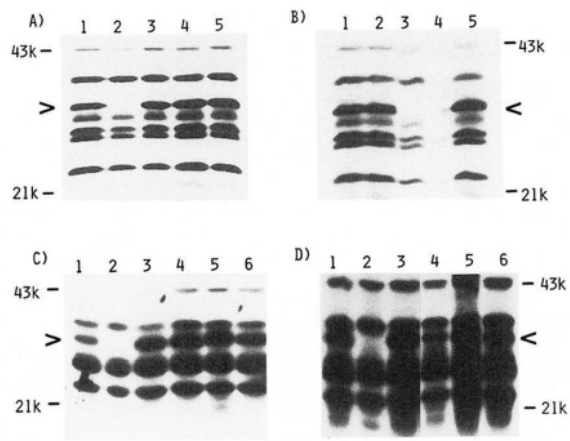

FIG. 2. **Maxicell analysis of mutant NahR proteins.** *E. coli* SK6501 cells containing mutant NahR plasmids were UV-irradiated and labeled with [^{35}S]methionine, and total protein samples were electrophoresed on a sodium dodecyl sulfate-10% polyacrylamide gel followed by fluorography. Samples are as follows. *A, lane 1,* pMS104; *lane 2,* pRK404 vector; *lane 3,* pMAS43; *lane 4,* pMAS63; *lane 5,* pMAS11. *B, lane 1,* pMAS131; *lane 2,* pMAS12; *lane 3,* pMAS245; *lane 4,* SK6501 only; *lane 5,* pMAS17. *C, lane 1,* pMAS42; *lane 2,* pMAS34; *lane 3,* pMAS31; *lane 4,* pMAS61; *lane 5,* pMAS241; *lane 6,* pMAS92. *D, lane 1,* pMA341; *lane 2,* pRK404; *lane 3,* pMAS101; *lane 4,* pMAS30; *lane 5,* pMAS40; *lane 6,* pMAS151. The migration positions of 34-kDa NahR protein (>) and the molecular size standards ovalbumiun (43 kDa) and chymotrypsin (21 kDa) are indicated.

Thus, whereas most class II mutants had a relatively high specific DNA binding activity, they were nonetheless totally defective in activation of transcription from P_{sal} in response to the inducer. Maxicell analysis showed that class II mutants produced the same amount of NahR polypeptide as the wild-type allele (Fig. 2 and data not shown). These mutants probably contain amino acid substitutions which render them defective either in inducer binding or in the ability to activate transcription in response to binding of inducer.

Analysis of Mutant Alleles for trans-Dominance—Each mutant *nahR* plasmid was transferred by conjugative mobilization into a *P. putida* strain containing the wild-type NAH7 plasmid. When the resultant merodiploids, *P. putida* (NAH7) (pMASn), were grown on salicylate as sole carbon and energy source, most of the class II and some of the class I merodiploids were significantly reduced in apparent growth rate. Since a functional NahR protein is required for rapid growth on salicylate, it was possible that the reduced growth rate was caused by interference of some mutant NahR proteins with the normal functioning of wild-type NahR.

To quantify the interference, transcription activation of the *sal* operon on NAH7 in *P. putida* merodiploids was measured by determination of the levels of catechol 2,3-dioxygenase (encoded by *nahH*) in cells grown with salicylate. Since *nahH* is the second gene in the *sal* operon transcript, the levels of catechol 2,3-dioxygenase activity should indicate the magnitude of *sal* operon transcription activation.

The presence of several of the class II mutant alleles (pMAS50, pMAS151, pMAS256, pMAS258, and pMAS241) reduced the induction of the *sal* operon in the merodiploids by 70–88%, whereas a few had a smaller inhibitory effect (Table II). This is not unexpected since the most inhibitory alleles produce mutant proteins which retain DNA binding activity and thus could interfere with binding and transcription activation by wild-type NahR. On the other hand, most class I alleles did not have a major effect on the expression of *nahH* in the merodiploids (Table I). However, a few (pMAS11, pMAS92, pMAS63, pMAS42, and pMAS43) apparently inhibited transcription activation by NahR by nearly 50%. Since these mutants apparently cannot bind to P_{sal}, their inhibitory effect is more difficult to explain. One possibility is that *in vivo*, the active form of NahR is a multimer. The formation of inactive heteromultimers in the merodiploids would result in a lower amount of active NahR, leading to lowered transcription activation.

Biochemical Evidence That NahR Is a Multimer—Further evidence that the native active form of NahR is a multimer was obtained from gel filtration analysis of the molecular size of P_{sal}-specific DNA binding activity. It was previously reported that in *E. coli* maxicells, most NahR protein is tightly associated with DNA, probably resulting from tight binding of NahR to its activation site on P_{sal}, which partially overlaps the divergent promoter for *nahR* (5). Thus, its true molecular size could not be accurately determined due to its migration as a heterogeneous population of DNA-protein complexes (1). Therefore, a plasmid (pWS22) was constructed which had the *nahR* coding region (lacking P_{sal}, the *nahR* promoter, and the NahR-binding site) fused to the *lac* promoter.

Gel filtration analysis utilizing Bio-Gel A-1.5m was employed to determine the molecular size of P_{sal}-specific DNA binding activity produced in crude extracts of *E. coli* (pWS22) cells. The molecular size of P_{sal}-specific DNA binding activity was found to be 125 kDa in reference to several size standards (Fig. 3). No evidence for 70- or 35-kDa species was obtained. In contrast to previous results (1), DNase I treatment did not alter the elution position of NahR (data not shown). These results suggest that NahR *in vivo* may bind as a tetramer.

Determination of Amino Acid Changes Responsible for Class I Mutant Phenotypes—The entire nucleotide sequence (from nucleotides −70 to +1000; COOH terminus = nucleotide 954) of each mutant *nahR* allele was determined on one strand

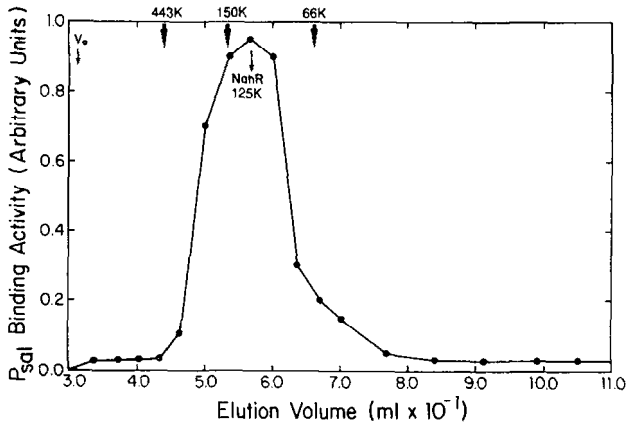

FIG. 3. **Gel filtration analysis of molecular size of active NahR protein.** A 1.5 × 65-cm column of Bio-Gel A-1.5m was loaded with 10 mg of crude extract protein from *E. coli* SK6501 (pWS22). Following elution with TE buffer + 0.1 M NaCl, the relative amount of P_{sal}-specific binding activity in fractions (2.5 ml) was quantified by densitometry of autoradiographs of gel retardation assays. Elution positions of the molecular size standards bovine serum albumin (66 kDa), alcohol dehydrogenase (150 kDa), and apoferritin (443 kDa) are indicated.

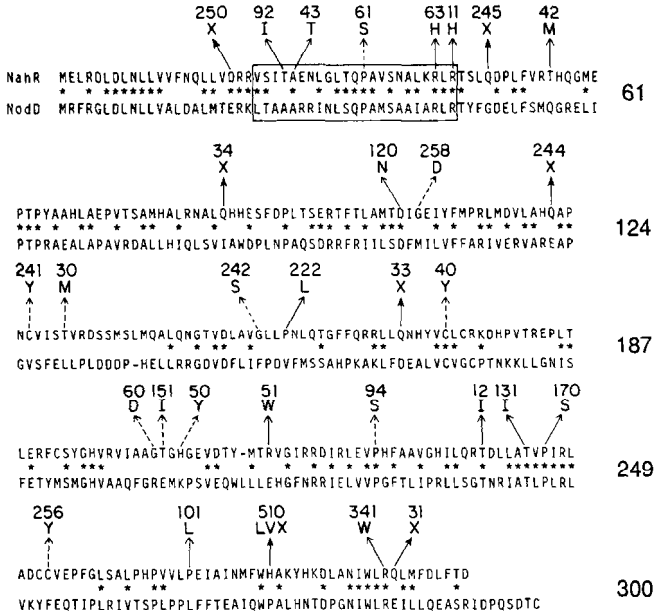

FIG. 4. **Location and nature of amino acid substitutions in NahR mutants.** The amino acid sequence (single-letter code) of the NahR protein (*upper lines*) is shown aligned with the amino acid sequence of the homologous NodD transcription activator (*lower lines*) from *R. meliloti*. Similar residues are marked with *asterisks*. *Arrows* show amino acid substitutions responsible for the altered phenotype of each numbered pMASn NahR mutant; *solid arrows* designate class I mutations, and *broken arrows* designate class II mutations. *X* designates terminations mutants. The helix-turn-helix motif is *boxed*. Residue numbers of NahR for the last residue on each line are given to the right.

using *nahR*-specific oligonucleotide primers. Among the 50 independent activation-defective mutants selected, 30 unique single-nucleotide substitution mutants were found. Consistent with the mechanism of action of hydroxylamine, all nucleotide substitutions were either A for G or T for C; no double-nucleotide change mutants or deletions were found. Twenty-six of the 50 mutations produced unique amino acid substitutions in the NahR reading frame; only five were chain termination mutants (Tables I and II and Fig. 4). Most substitutions occurred only once or twice. However, a few appeared at higher frequency in the population: $Thr^{130} \rightarrow Met$ (pMAS30) was found eight times; $His^{206} \rightarrow Tyr$ (pMAS50) and $Gly^{107} \rightarrow Asp$ (pMAS258) were each found four times.

Ten of the 13 class I amino acid substitutions which resulted in a major (>95%) loss of DNA binding activity were clustered in only two small regions at opposite ends of the NahR protein: five between residues 26 and 56 and five between residues 239 and 291 (Table I and Fig. 4). Most (70%) of these substitutions are at residues which are conserved in both NahR and NodD (7), strongly implicating these two regions in the DNA binding function of the two transcription activators. In fact, five of these substitutions are located in or near the putative HTH motif of NahR (residues 23–45). One of the mutations (pMAS43) alters a residue in the NahR HTH motif (Ala^{27}) which is conserved in nearly all HTH motifs (9–11). Changing either Arg^{43} (pMAS63) or Arg^{45} (pMAS11) to His also results in complete loss of DNA binding activity. An Arg residue corresponding to the NahR Arg^{45} is found at an analogous position near the Cro and Tn10 *tet* repressor HTH motifs (27); for Cro, this Arg residue is important for its protein/DNA interactions (10). An Arg residue corresponding to Arg^{43} is found in the proposed HTH motifs of many of the LysR family of activators (8). Substitution of Pro^{35} with Ser in the NahR HTH motif, while largely blocking salicylate-induced transcription activation, reduced the apparent binding affinity of the mutant NahR protein for P_{sal} by only 70%. An analogous Pro residue is also found at the same relative position in the HTH motifs of NodD, LysR, and TetR. In TetR, changing this Pro residue to Leu completely eliminated its ability to bind to DNA (27), further suggesting an important role for this residue in the binding and transcription activation function of NahR.

The other region of NahR containing a high percentage of the amino acid substitutions giving class I phenotype is near the COOH terminus (residues 239–290), especially between residues 239 and 246 (Fig. 4). This region does not appear to contain any HTH motifs when analyzed by the method of Dodd and Egan (9). The clustering of the mutations and the fact that they are mostly in residues conserved in NahR and NodD (and in some LysR family members) suggest this region is specifically involved in DNA binding. However, it is possible that these substitutions prevent proper folding of *NahR* and thus indirectly destroy DNA binding activity. This is even a more likely possibility for the remaining three class I mutations which are scattered at positions 105, 156, and 215 outside the two major class I domains.

The phenotypes of six chain termination mutants of NahR (pMAS31 pMAS33, pMAS34, pMAS244, pMAS245, and pMAS510) also suggest that the NH_2-terminal HTH motif alone is not sufficient for DNA binding activity. None of the putative truncated NahR proteins encoded by these mutants showed any evidence of DNA binding activity. In fact, deletion of as little as the last 9 COOH-terminal amino acids (pMAS31) results in a complete loss of DNA binding and transcription activation activity. This truncated NahR protein was produced in amounts similar to wild-type NahR (Fig. 2*C*, *lane 3*). The majority of the evidence suggests that the COOH-terminus plays an important role in DNA binding function either directly or, if the NahR protein is a tetramer as suggested by other experiments, indirectly by a possible involvement in formation of an active tetrameric species.

Analysis of Amino Acid Substitutions Resulting in Class II Phenotype—DNA sequence analysis revealed that substitutions producing mutants which have apparent wild-type binding affinity for P_{sal} yet cannot activate its transcription (class II) are also clustered in two regions: three of nine are located

between residues 203 and 207 and three between residues 107 and 130. The remaining three substitutions, while somewhat scattered, still are located within 50 residues of these other two clusters. Over 90% of the class II mutations lie in the central third of NahR (residues 107–227), whereas less than 20% of the class I substitutions were found in this region. Since all these mutants bind to, but do not activate, P_{sal}, it is possible that the central third of NahR (and residues 203–206 in particular) contains domains involved with either inducer binding and/or transcription activation. However, this hypothesis requires *in vitro* assays for salicylate binding to wild-type and mutant NahR proteins for confirmation. In contrast to class I mutants, 85% of the substitutions resulting in the class II phenotype are in residues that are not conserved in NodD and NahR. Three of the class II substitutions located outside the major cluster between residues 203 and 206 change Cys residues that might be involved in disulfide bond formation required for maintenance of protein folding.

DISCUSSION

Hydroxylamine mutagenesis and galactose/salicylate selection were used to isolate 30 unique mutant alleles of *nahR* encoding proteins defective in transcription activation of the *sal* operon promoter. Analysis of their ability to specifically bind to the activation site of P_{sal} divided the mutants into two types: class I, which do not bind to the activation site; and class II, which largely retain specific binding for P_{sal}. Since many class I substitutions were in or near a previously hypothesized HTH DNA binding motif (residues 23–45), their phenotype is probably due to amino acid substitutions for residues which are involved in interactions with the nucleotides of the NahR-binding site. In fact, several substitutions were located at or near HTH positions which, in other proteins with HTH motifs, are intimately involved in protein/DNA interactions at their target sites (10, 28, 29). Many of the altered residues in the NH_2-terminal HTH domain are highly conserved in NodD and other members of the LysR family of activators, suggesting that this HTH DNA binding domain and Ala^{27}, Pro^{35}, and Arg^{43}, in particular, are a highly conserved functional part of many transcription activators in the LysR family. Some residues (Thr^{26}, Arg^{45}, and Thr^{56}) near the HTH motif were apparently required for DNA binding by NahR, but are not conserved among the LysR family; these residues may be located in HTH positions involved in sequence recognition specificity.

The other region of NahR containing many of the substitutions eliminating DNA binding activity is that containing the last 60 COOH-terminal residues, especially residues 239–246. This is the only region in the COOH-terminal half of NahR showing concentrated amino acid sequence similarity to the NodD activator protein and some LysR family members. The clustering of the mutations in a region with a conserved amino acid sequence suggests that this region could be a second DNA-binding domain directly involved in the binding of NahR (and possibly NodD) to DNA. Analysis of *nahR* termination mutants showed that the NH_2-terminal HTH motif alone is not sufficient for DNA binding since putative truncated NahR proteins with only the first 85 (pMAS34) or 168 (pMAS33) residues could not bind DNA. Most convincing was the observation that pMAS31, producing an NahR protein lacking only the last 9 amino acids (Fig. 2*C*, *lane 3*), was completely inactive, indicating an absolute requirement for the complete COOH terminus for binding to P_{sal}. The lack of a typical DNA-binding motif structure in the COOH terminus could indicate that the COOH terminus is involved in maintenance of structure necessary for positioning of the NH_2-terminal HTH motif or, in some other manner, is indirectly involved in DNA binding. However, Burn *et al.* (30) have reported that substitution of Asn for the Asp residue located 20 residues from the COOH terminus of NodD produced mutant NodD proteins which activated expression of *nod* genes to high levels independently of the presence of flavanone inducer. Clarification of the role of the COOH terminus of NahR and other LysR family activators in their function awaits further experimentation.

Class II mutants retaining apparent wild-type binding activity for P_{sal} but which are deficient in activation function probably result from substitutions for residues involved in binding of the inducer salicylate or in the transcription activation function caused by inducer binding. Many substitutions producing the class II phenotype are in residues not conserved in NahR and NodD; since the two proteins respond to chemically distinct inducers, a difference in some of the residues involved in binding of the inducers is expected. Since 90% of the class II substitutions are located between residues 107 and 227, our hypothesis that class II substitutions affect inducer recognition is supported by the experiments of Spaink *et al.* (31) which implicate the involvement of the analogous region of NodD (residues 122–270) in flavanone inducer recognition.

Assays of transcription activation in merodiploids containing both the mutant and wild-type alleles showed that nearly all the class II alleles were partially *trans*-dominant, whereas most class I alleles were not. However, a few class I mutants (with substitutions located in or near the NH_2-terminal HTH motif) did show significant interference with the functioning of the wild-type protein in merodiploids. The inhibitory effect of class II alleles could result from their ability to occupy the NahR activation site of P_{sal} and restrict access by wild-type NahR. This explanation is insufficient for the inhibitory effect of the class I NahR proteins which cannot bind to DNA. One possibility is that they interfere with formation of a multimeric species. Many prokaryotic transcriptional regulators that use the HTH motif bind to DNA as dimers; the CysB protein, a transcription activator in the LysR family, has been shown to probably exist as a tetramer (32). If, as suggested by the gel filtration experiments, NahR also functions *in vivo* as a tetramer, then the formation of inactive heteromultimers in merodiploids containing mutant *nahR* alleles would significantly lower the amount of wild-type multimeric NahR protein and result in inhibition of transcription activation of P_{sal}. In this manner, class I mutant NahR proteins could partially block the functioning of wild-type NahR without having DNA binding capacity. If NahR functions as a multimer, then any mutation which affects formation of contacts between subunits or generally alters overall protein structure could affect formation of multimers and alter transcription activation. Since the class I mutations in the conserved COOH-terminal domain (residues 239–246) showed no *trans*-dominance, it is tempting to speculate that this region is directly involved in formation of multimers, but further direct evidence is necessary to demonstrate this.

Acknowledgments—We thank S. Hulsey, L. Su, and M. Williams for technical assistance and P. Bates for typing the manuscript.

REFERENCES

1. Schell, M. A., and Wender, P. E. (1986) *J. Bacteriol.* **166,** 9–14
2. Schell, M. A. (1985) *Gene (Amst.)* **36,** 301–309
3. Yen, K.-M., and Gunsalus, I. C. (1982) *Proc. Natl. Acad. Sci. U. S. A.* **79,** 874–878
4. Dunn, N. W., and Gunsalus, I. C. (1973) *J. Bacteriol.* **114,** 974–979
5. Schell, M. A. (1986) *Proc. Natl. Acad. Sci. U. S. A.* **83,** 369–373

6. Schell, M. A., and Faris, E. (1989) *J. Bacteriol.* **171,** 837–846
7. Schell, M. A., and Sukordhaman, M. (1989) *J. Bacteriol.* **171,** 1952–1959
8. Henikoff, S., Haughn, G., Calvo, J., and Wallace, J. C. (1988) *Proc. Natl. Acad. Sci. U. S. A.* **85,** 6602–6606
9. Dodd, I. B., and Egan, J. B. (1987) *J. Mol. Biol.* **194,** 557–564
10. Pabo, C. O., and Sauer, R. J. (1984) *Annu. Rev. Biochem.* **53,** 293–321
11. Brennan, R. G., and Matthews, B. W. (1989) *J. Biol. Chem.* **264,** 1903–1906
12. Hong, G.-F., Burn, J. E., and Johnston, A. W. B. (1988) *Nucleic Acids Res.* **15,** 9677–9689
13. Fisher, R. F., Egelhoff, T. T., Mulligan, J. T., and Long, S. R. (1988) *Genes & Dev.* **12,** 282–293
14. Ditta, G., Schmidhauser, T., Yakobson, E., Lu, P., Liang, Y.-W., Finlay, D. R., Guiney, D., and Helinski, D. R. (1985) *Plasmid* **13,** 149–153
15. McKenney, K., Shimtake, H., Court, D., Schmeissner, V., Brady, C., and Rosenberg, M. (1981) in *Gene Amplification and Analysis* (Chirikjan, J. C., and Papas, T. S., eds) Vol. 2, pp. 383–415, Elsevier/North-Holland Publishing Co., New York
16. Davis, R. W., Botstein, D., and Roth, J. R. (1980) *A Manual for Genetic Engineering: Advanced Bacterial Genetics*, pp. 94, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY
17. Maniatis, T., Fritsch, E. F., and Sambrook, J. (1982) *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY
18. Feist, C. F., and Hegeman, G. D. (1969) *J. Bacteriol.* **100,** 869–877
19. Kalb, V. F., and Bernlohr, R. (1977) *Anal. Biochem.* **82,** 362–371
20. Sanger, F., Miklen, S., and Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. U. S. A.* **74,** 5463–5467
21. Dente, L., Cesareni, G., and Cortese, R. (1983) *Nucleic Acids Res.* **11,** 1645–1655
22. Mead, D. A., Skorupa, E. S., and Kemper, B. (1985) *Nucleic Acids Res.* **13,** 1103–1118
23. Genilloud, O., Garrido, M. C., and Moreno, F. (1984) *Gene (Amst.)* **32,** 225–234
24. Sancar, A., Hack, A. M., and Rupp, W. D. (1979) *J. Bacteriol.* **137,** 692–693
25. Laemmli, U. K. (1970) *Nature* **227,** 680–685
26. Laskey, R. A., and Mills, A. D. (1975) *Eur. J. Biochem.* **56,** 335–341
27. Jackson, P. J., and Bertrand, K. P. (1985) *Proc. Natl. Acad. Sci. U. S. A.* **82,** 6226–6230
28. Wohlberger, C., Dong, Y., Ptashne, M., and Harrison, S. C. (1988) *Nature* **335,** 789–795
29. Jordon, S. R., and Pabo, C. (1988) *Science* **242,** 893–899
30. Burn, J., Rossen, L., and Johnston, A. W. B. (1981) *Genes & Dev.* **1,** 456–464
31. Spaink, H. P., Wijffelman, C. A., Okker, R. J., and Lugtenberg, B. E. (1989) *Plant Mol. Biol.* **12,** 59–73
32. Miller, B. E., and Kedrich, N. M. (1987) *J. Biol. Chem.* **262,** 6006–6009

Use of saturation mutagenesis to localize probable functional domains in the NahR protein, a LysR-type transcription activator.

M A Schell, P H Brown and S Raju

J. Biol. Chem. 1990, 265:3844-3850.