JOURNAL OF BACTERIOLOGY, Feb. 1993, p. 1103–1109
0021-9193/93/041103-07$02.00/0

Vol. 175, No. 4

The Oxygen Sensor FixL of *Rhizobium meliloti* Is a Membrane Protein Containing Four Possible Transmembrane Segments

AUGUSTO F. LOIS,† GARY S. DITTA, AND DONALD R. HELINSKI*

Department of Biology and Center for Molecular Genetics, University of California, San Diego, La Jolla, California 92093-0634

Received 18 September 1992/Accepted 8 December 1992

Regulation of nitrogen fixation genes in *Rhizobium meliloti* is mediated by two proteins, FixL and FixJ, in response to oxygen availability. FixL is an oxygen-binding hemoprotein with kinase and phosphatase activities that is thought to sense oxygen levels directly and to transmit this signal to FixJ via phosphorylation-dephosphorylation reactions. FixJ controls the expression of other regulatory genes, including *nifA*, that regulate the transcription of genes required for symbiotic nitrogen fixation. We have been studying the structural and functional features of FixL that are required for oxygen sensing. We constructed mutant derivatives and confirmed that FixL consists of 505 amino acids instead of 464, as originally reported. Hydropathy plots of the full-length protein, together with Tn*phoA* insertional analysis, lead us to propose that FixL is likely to be a polytopic integral membrane protein containing four membrane-spanning segments. We have also constructed an N-terminal deletion of the FixL protein whose in vivo activity indicates that the hydrophobic membrane-spanning regions are not absolutely required for oxygen sensing in vivo. We also report that FixL shares homology in its N terminus with other sensor proteins, including KinA from *Bacillus subtilis* and NtrB from *Bradyrhizobium parasponia*. The region of homology comprises a 70-amino-acid residue stretch that is also conserved in two oxygenases, P-450 and isopenicillin synthase.

Sensory transduction of extracellular signals is an important process that permits both prokaryotic and eukaryotic cells to respond to different environmental stimuli. During nodulation of alfalfa (*Medicago sativa*) by *Rhizobium meliloti*, a continuing signal exchange takes place between the plant and the bacterium that culminates in a nitrogen-fixing symbiosis within the root cortex. An important signal that regulates the expression of bacterial genes involved in nitrogen fixation appears to be the concentration of oxygen within the nodule (1, 16, 43). A heme-containing protein, the product of the *fixL* gene (FixL), is thought to be involved in sensing oxygen and in transducing this signal via kinase and phosphatase activities to a positive gene regulator, FixJ (20). The *fixJ* gene has been shown to regulate the expression of other regulatory genes, *nifA* and *fixK*, that in turn control transcription from a number of genes, including those coding for nitrogenase (*nifHDK*), the enzyme directly involved in fixing atmospheric nitrogen (3, 13, 41).

FixL and FixJ belong to a group of regulatory proteins that consists of two components (a sensor and a regulator) that serve as mediators of environmental signal transduction in bacterial cells (13, 39). FixL has sequence similarity with other sensor proteins in its C-terminal end, while FixJ is related to other regulator proteins in its N-terminal region. Signal transduction in these systems occurs via phosphorylation-dephosphorylation reactions (40). Typically, the sensor protein is autophosphorylated at a conserved histidine residue in the C-terminal domain and transfers the phosphate to a conserved aspartate residue in the regulator's N-terminal region in response to a specific stimulus.

The mechanism whereby sensor proteins perceive a signal and modify their kinase activity accordingly is largely unknown. Several of the sensor proteins, including EnvZ (40), PhoR (30), VirA (35), and perhaps KinB (8), are thought to be transmembrane proteins that perceive an extracellular signal via their N-terminal domain and, either by conformational changes or by changes in oligomeric state, alter their C-terminal kinase activity (22). Other sensor proteins of two-component systems, such as CheA (38), NtrB (39), and KinA (8), are soluble and may require additional proteins to sense the environmental signal.

The FixL protein from *R. meliloti* contains several possible transmembrane segments in its N-terminal region. The nucleotide sequence of FixL, determined by David et al. (13), shows two putative starting codons at the 5′ end of the gene. On the basis of the coding probability, David et al. assigned the second ATG as the starting codon. However, the stretch of amino acids that would be translated between the first and second ATGs provides an additional hydrophobic region that could serve as a membrane spanner in FixL. In this study, we have used site-directed mutagenesis coupled with Tn*phoA* insertional analysis to present evidence that translation of FixL begins preferentially at the first ATG codon, producing a polytopic transmembrane protein likely to contain four membrane-spanning segments. We have recently shown (20, 28, 37) that a truncated soluble version of FixL that retains the heme moiety is autophosphorylated in vitro, can transfer a phosphate group to FixJ, and responds to oxygen in vitro by modifying its kinase and phosphatase activities. Since FixL binds oxygen and can therefore sense oxygen concentration directly, we also were interested in determining whether the N-terminal hydrophobic region of FixL was necessary for oxygen sensing in vivo. A deletion analysis of the N-terminal region of FixL revealed that oxygen sensing in vivo also does not require the presence of the predicted membrane-spanning segments.

* Corresponding author.

† Present address: Regulatory Biology Laboratory, The Salk Institute for Biological Studies, San Diego, CA 92186-5800.

MATERIALS AND METHODS

Bacterial strains and phage. All strains used in this work have been described previously. The *R. meliloti* wild-type strain is 102F34 (12); T219 is a derivative of 102F34 that contains an unspecified deletion including the *fixL* and *fixJ* genes (43). *Escherichia coli* strains include C2110, a *polA* mutant (24) that was used to propagate cointegrates of pUC and pRK290 derivatives; TB1 [F' *ara*Δ(*lac-proAB*) *rpsL* (φ80dΔ(*lacZ*)M15)] (2); CC118, a *phoA* mutant used for Tn*phoA* mutagenesis (31); and XL1-BLUE [*recA1 endA1 gyrA96 thi hsdR17 supE44 relA1 lac* (F' *proAB lacI*^{q}ZΔM15) Tn*10* (Tet^r)] (Stratagene Corp.), which was used to propagate plasmids carrying Tn*phoA* insertions. Phage λ Tn*phoA* (23) was obtained from C. Manoil.

Plasmid constructs and site-directed mutagenesis. Plasmids pCHK57 (16), pRK404 (15), pSelect (Promega Corp.), pBluescript KS$^+$ (Stratagene), and pMW2 (20) have been described elsewhere. pGD31 is an intermediate vector used both to construct pGD311 and to facilitate the transfer of mutated *fixL* genes to pGD311. It was generated from pMW2 in several steps. A *Hin*dIII-*Sna*BI fragment was deleted, the ends were filled in with Klenow enzyme, and the DNA was religated to eliminate one of the two *Bam*HI sites. Since cointegrates with pCHK57 were to be made at a unique *Hin*dIII site, sequences homologous with the β-galactosidase gene carried on PCHK57 were deleted and a downstream *Hin*dIII site was introduced by replacing a 168-bp *Bam*HI-*Nar*I fragment with a *Hin*dIII linker, yielding pGD28. To eliminate the *lac* promoter from pGD28, first the *Pvu*II-*Bgl*II fragment of pGD28 was replaced with the *Pvu*II-*Bgl*II fragment of pMW2, and then the *Pvu*II-*Bam*HI fragment in this construct was deleted, yielding pGD31. The cointegrate plasmid pGD311 was constructed by combining pGD31 and pCHK57, using the single *Hin*dIII site of each, with the orientation being such that the *lac* promoter reads in the direction opposite that of the *nifA* promoter (P*nifA*).

pGG101 was used to create plasmid pAL401Δ and was constructed by Marie Gilles-Gonzalez as follows. Plasmid pMW2 was cut with *Eco*RI and then ligated; this step removes the 3' end of *fixL* and all of *fixJ*. The deleted plasmid was then cut with *Stu*I and *Rsr*II and ligated; this step deletes amino acids 50 through 126 in the FixL-coding region and leaves the reading frame intact. To restore the 3' end of *fixL* and the *fixJ* coding region, the *Eco*RI fragment from pMW2 was cloned back into the plasmid, yielding pGG101.

The plasmid used to generate Tn*phoA* insertions in *fixL* (pAF20) was made by cloning a *Hin*dIII-*Bgl*II fragment from pMW2 into pBluescript KS$^+$. This fragment contains the entire *fixL* and *fixJ* genes and the 5' end of the *fixK* gene.

Site-directed mutagenesis was accomplished by means of the Altered Sites *in vitro* Mutagenesis System (Promega Corp.), following the manufacturer's specifications. Two plasmids were used to generate all of the site-specific mutations described in this work, pAL40 and pAL41. A *Bam*HI-*Eco*RI fragment from pGD31 or pGG101 which specifies the N-terminal region of the *fixL* gene was cloned into pSelect to yield pAL40 or pAL41, respectively. Plasmids pAL420 and pAL430 are derivatives of pAL40; plasmid pAL401Δ is a derivative of pAL41. These plasmids were constructed as described below.

The oligomer 5'-CGCTTTTTGACTGCAGCCAGTTATCC-3' was used to create a *Pst*I restriction site at the first ATG codon of *fixL* to yield pAL430 (methionine residue changed to leucine residue), and the oligomer 5'-TGCGATCAGATC TATACGCGCCG-3' was used to generate a *Bgl*II site at the second ATG codon, yielding pAL420 (methionine changed to serine). Deletions within *fixL* were made by using the oligomer 5'-TTCTGATCGCAAGAACAAGCATCCAGTTA TCC-3'. Use of this oligomer produces a deletion of the codons for amino acid residues 3 through 36 within the *fixL* coding region and yielded plasmid pAL401Δ.

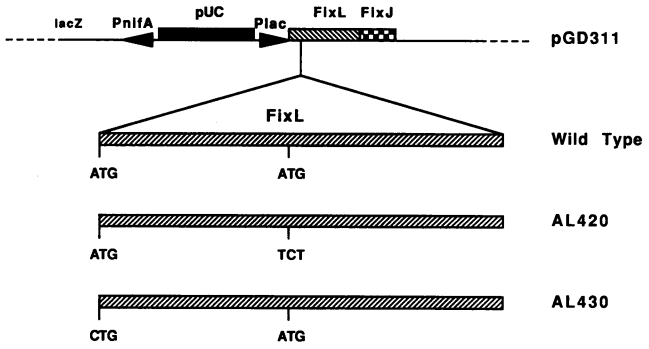

FIG. 1. Plasmid constructs expressing FixL, FixJ, and mutant derivatives. pGD311 contains wild-type *fixL* and *fixJ* genes inserted into the broad-host-range plasmid pCHK57 (2). The expanded view represents the first 100 amino acids of the N-terminal end of FixL. The positions of the two ATGs correspond to residues 1 and 41 from the amino end of the polypeptide.

After site-directed mutagenesis was carried out, the *Bam*HI-*Eco*RI fragment was cloned back into pGD31 deleted for the same restriction fragment. The pGD31 derivatives were then cut with *Hin*dIII and cloned into *Hin*dIII-digested pCHK57. pGD311 contains the wild-type *fixL* and *fixJ* genes in pCHK57 (16) (Fig. 1). pCHK57 contains a translational fusion of P*nifA* with *lacZ* to monitor the induction of P*nifA*. Plasmids pAL420, pAL430, and pAL401Δ are similar in structure to pGD311 but contain the mutations or the deletions described above and shown in Fig. 1. These plasmids were transferred to *R. meliloti* by triparental matings involving *E. coli* C2110 cells carrying the designated plasmid and *R. meliloti* T219 and TB1 cells carrying the helper plasmid pRK2073 (12, 14, 16).

To measure alkaline phosphatase activity of the different *fixL*-Tn*phoA* fusions in *R. meliloti*, the DNA region of the pAL20 derivative containing the insertion was subcloned into pRK404 as follows. A *Bam*HI fragment from pAL20 which contains the N-terminal region of FixL and the entire *phoA* gene of the transposon was inserted into pRK404 such that the *lac* promoter reads into the *fixL* gene. These plasmids were then mated into wild-type *R. meliloti* as described below.

Media, growth conditions, and enzyme assays. *E. coli* strains were grown in LB broth (GIBCO-BRL); rhizobia were grown in either TY (16) or YMB (42). Antibiotics were used as follows for *E. coli*: penicillin G, 250 μg/ml; kanamycin, 30 or 300 μg/ml; and tetracycline, 15 μg/ml. For *R. meliloti* strains, we used tetracycline at 5 μg/ml, kanamycin at 30 μg/ml, and nalidixic acid at 15 μg/ml.

To detect alkaline phosphatase activity on plates, cells were streaked on TY containing 40 mg of 5-bromo-4-chloro-3-indolyl phosphate (BCIP) per ml. Alkaline phosphatase activity of liquid cell cultures was determined as described previously (29), following the modification of an earlier assay (7). β-Galactosidase activity was measured in liquid cultures (36) and in plant nodules (4). The stoppered tube assay (16)

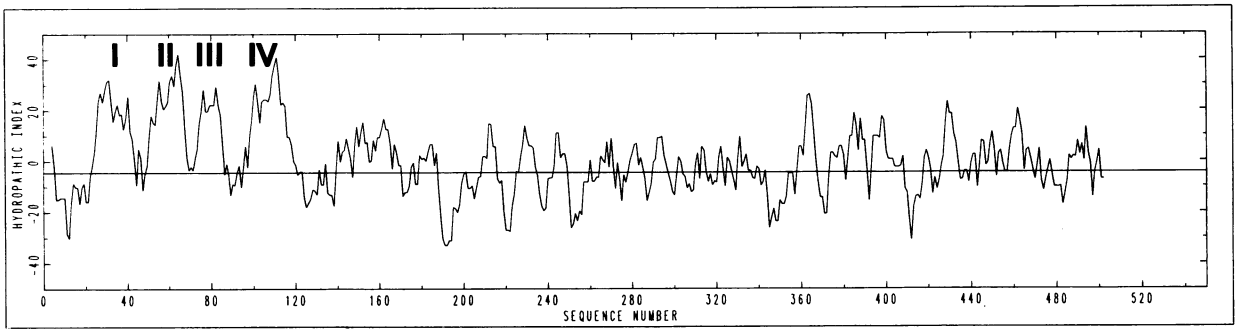

FIG. 3. Hydropathy profile of FixL. TS I to IV are indicated.

A hydropathy plot constructed with the amino acid sequence of the complete FixL protein revealed four hydrophobic stretches that could serve to anchor the protein to the membrane (Fig. 3). Additional evidence that translation starts at the first ATG was obtained by Tn*phoA* insertional analysis as described below. A *fixL-phoA* fusion (pAL37) with high alkaline phosphatase activity was present as an insert between the first and second ATGs at codon 37 (Fig. 4 and Table 2), suggesting that this region of the protein is being translated in vivo.

Tn*phoA* insertional analysis. *E. coli* CC118 carrying the *fixLJ* genes on plasmid pAL20 was used as the bacterial host for the transpositions. The colonies that appeared after infection with λ Tn*phoA* had at least three distinct phenotypes: white, light blue, and dark blue. White colonies were the most abundant and contained cells harboring plasmids with insertions in many different regions. Cells of one white colony were analyzed, and the site of insertion was mapped in *fixL* in a position corresponding to the presumed cytoplasmic region of the FixL protein (pAL386; Fig. 4). Restriction enzyme analysis of the plasmids isolated from light blue colonies revealed in every case a rearrangement of the Tn*phoA* transposon. Each of the dark blue colonies analyzed contained Tn*phoA* insertions that were mapped either to the β-lactamase gene or to a location in the *fixL* gene corresponding to the predicted periplasmic regions shown in Fig. 4. The exact location of each insertion in the periplasmic region was determined by DNA sequencing. As shown in Fig. 4, one insertion was in the first periplasmic loop, while four others were located in the second loop.

To determine whether the properties of the Tn*phoA* inserts in *R. meliloti* are similar to those observed in *E. coli*, the *fixL-phoA* fusions were transferred by in vitro recombinant DNA procedures to the broad-host-range plasmid pRK404 (15) as described in Materials and Methods. The pRK404 plasmid derivatives were then established in wild-type *R. meliloti* by conjugal mating. As shown in Table 2, the phenotype of each of the translational fusions was the same in *R. meliloti* as in *E. coli* CC118. The enzymatic activities of the translational protein fusions in *Rhizobium* liquid cultures revealed a 13- to 18-fold difference between putative cytoplasmic and periplasmic locations. These results are consistent with differences in the level of alkaline phosphatase activity of fusions obtained for several other proteins (32, 33). We have constructed a topological map of FixL (Fig. 4) that is based on the hydropathy plots and on the results of the Tn*phoA* mutagenesis. According to this model, it is likely that there are two distinct periplasmic loops. The first one extends from amino acids 39 through 44, and the second one extends from residues 86 through 97 (Fig. 5). *phoA* fusions with high alkaline phosphatase activity were isolated at residue 37 and between amino acids 83 through 89, consistent with this interpretation.

Chou and Fasman (9, 10) and Doolittle (17) have studied the probability of bend occurrence in polypeptides. We calculated the probability of a β turn occurring at each of the three predicted bends (Fig. 4) by the method of Chou and Fasman (10) or by means of the computer program COMFOR (17). It was found that there is a high β-turn probability between transmembrane segments II and III (TS II and III) and also between TS III and IV (Fig. 5), as calculated by these two different methods (Table 3). Chou and Fasman

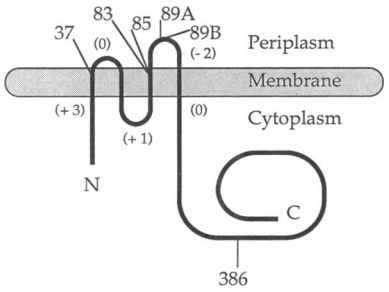

FIG. 4. Predicted topology of FixL in the membrane. Numbers indicate amino acid positions from the N-terminal end (using the first ATG start codon) of the FixL-Tn*phoA* fusion junctions. Numbers in parentheses show the total charge of a 10-amino-acid stretch from the end of the transmembrane segment.

TABLE 2. Alkaline phosphatase activities of Tn*phoA*-generated fusions

Fusion plasmid[a]	*R. meliloti*		*E. coli* phenotype
	Alkaline phosphatase activity (U/OD_{600}[b])	Phenotype	
pAL37	27	Blue	Blue
pAL85	35	Blue	Blue
pAL88	25	Blue	Blue
pAL89A	29	Blue	Blue
pAL89B	31	Blue	Blue
pAL386	2	White	White
pRK404	0.8	White	White

[a] Plasmid name denotes the C-terminal-most amino acid position within FixL before the junction with Tn*phoA*.

[b] OD_{600}, optical density at 600 nm.

was used to determine the amount of transcription from P*nifA* under microaerobic conditions.

TnphoA mutagenesis. To isolate TnphoA insertions in *fixL*, *E. coli* CC118 cells carrying pAL20 were infected with λ TnphoA at a multiplicity of infection of 1 according to the procedure of Long et al. (29). DNA isolated from kanamycin-resistant colonies that had a stable phenotype was used to transform XL1-BLUE cells. This strain was used to propagate plasmids carrying the transposon because insertions appear to be unstable in *recA*$^+$ cells and the yield of plasmid DNA from CC118 cells was poor.

To determine the location of the *fixL*-TnphoA fusion junction, the plasmids were initially analyzed by cutting with *Kpn*I and *Sph*I. All plasmids except pAL386 were then sequenced by using a kit purchased from U.S. Biochemical to determine the exact locations of the insertions.

FixL antiserum. Antibodies were raised in rabbits by using a highly purified C-terminal-end fragment (37) of the FixL protein.

Gel electrophoresis and immunoblots. Sodium dodecyl sulfate-polyacrylamide gel electrophoresis was carried out according to Laemmeli (27), using 4% stacking and 8% separating gels. For immunoblots, the proteins were transferred to either Immobilon-P (Millipore Corp.) or Hybond-N (Amersham Corp.) membranes. After treatment with antibodies, the membranes were exposed either to BCIP (Immobilon-P) for permanent staining or to Lumiphos from Boehringer Mannheim (Hybond-N) for autoradiography. For quantitation, autoradiograms were scanned on a laser densitomer, and the data were processed by using software written by Ambis (San Diego, Calif.).

Secondary structure analysis of FixL. Hydropathy plots of FixL were carried out as described by Kyte and Doolittle (26). β-Turn predictions were made by using the program COMFOR, which is based on algorithms of Chou and Fasman (9) and Garnier et al. (19), as described previously (10).

A procedure designed to determine the significance of a sequence alignment between two sequences was kindly provided by R. Doolittle. We used the computer program DAYJUM to calculate the number of standard deviations from the mean alignment score of randomized sequences relative to the authentic alignment. The relatedness of two sequences is considered significant if the standard deviation is >3. When two sequences score above 5 in the jumbling test, they are considered to be genuinely related.

RESULTS

Translational start of FixL. The amino acid sequence of FixL, deduced from the nucleotide sequence, shows two possible ATGs as start codons (Fig. 1), with the second ATG start corresponding to codon 41. David et al. (13) assigned the second ATG as the translational start of the *fixL* gene. To determine which start codon is preferentially used in vivo, each of these two ATGs was changed separately by site-directed mutagenesis and the amount of FixL protein present in crude lysates prepared from cells carrying the mutant *fixL* genes was determined by Western immunoblots. These experiments were carried out with *R. meliloti* T219 that carries a deletion of the *fixLJ* operon in the chromosome and that contained plasmids producing either wild-type FixL and FixJ (pGD311) or mutant derivatives of FixL together with FixJ (pAL420 and pAL430, respectively; Fig. 1). As shown in Fig. 2, lanes 1 and 3, similar levels of FixL are produced from plasmid pAL420, which carries a mutation in the second ATG, and the pGD311 (wild-type) control. When

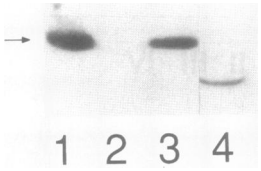

FIG. 2. Immunoblot analysis of FixL mutants in *R. meliloti*. Crude lysates of T219 cells harboring pGD311 (lane 1), pAL430 (lane 2), pAL420 (lane 3), and pAL401Δ (lane 4) were electrophoresed, and the proteins were blotted onto Hybond-N membranes. The structures of pAL420 and pAL430 are shown in Fig. 1, and the structure of pAL401Δ is shown in Fig. 5. Proteins were detected with FixL antiserum, using a chemiluminescent substrate, followed by autoradiography.

the first codon is mutated to the leucine codon CTG (pAL430), however, the amount of FixL present in the cell is dramatically reduced (Fig. 2, lane 2). These data suggest that the first ATG is used preferentially in vivo as the starting codon. A small amount of protein identical in size to wild-type FixL can be detected with pAL430. We assume that this protein arises from low-level initiation at the non-canonical start codon CUG by *R. meliloti* ribosomes. Initiation at this codon does not occur in *E. coli* (25) but has been observed frequently in eukaryotes (34).

Oxygen-regulated activity of wild-type FixL and of the mutant proteins was determined in *R. meliloti* by measuring the amount of transcription from the FixL-dependent P*nifA* under microaerobic conditions. For this purpose, a P*nifA-lacZ* fusion present in plasmid pGD311 (Fig. 1) was used, and the level of β-galactosidase activity was determined in the *fixLJ* deletion strain T219. Plasmid pGD311 carries both the FixL and FixJ genes (Fig. 1) and produces approximately five times more FixL than does the single genomic copy of *fixL*, as judged from Western blot analysis of FixJ levels. Previous work has shown that oxygen regulation and FixL-dependent P*nifA* induction by pGD311 are equivalent to those displayed by wild-type *R. meliloti* 102F34 (16). Table 1 shows that, as expected, the mutation in the second ATG had no effect on microaerobic induction of *nifA* transcription. A mutation in the first ATG reduced *nifA* transcription by about 30% (Table 1). The relatively modest effect on P*nifA* expression is probably due to the fact that even with a 90% reduction in the amount of FixL produced from pAL430, cells would still be expected to contain approximately one-half the amount of FixL that is present in wild-type *R. meliloti* 102F34, since the plasmid-borne *fixL* gene overproduces FixL protein by about fivefold, as discussed earlier. It is also likely that the amount of FixL normally present in *R. meliloti* 102F34 is in excess of that needed for proper oxygen regulation (27a).

TABLE 1. Characterization of FixL mutants in strain T219

Plasmid (phenotype)	β-Galactosidase activity (U)[a]		
	Aerobic	Microaerobic	Symbiotic
pCHK57 (no FixL)	43	81	19
pGD311 (wild type)	45	1,495	100
pAL420 (ATG-2 mutation)	49	1,440	
pAL430 (ATG-1 mutation)	34	957	
pAL401Δ (deletion in TS I–IV[b])	23	223	18

[a] Average of triplicate experiments.
[b] As shown in Fig. 5.

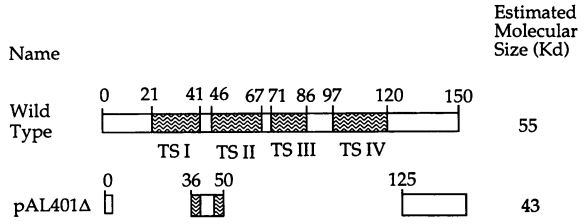

FIG. 5. Map of the N-terminal domain of wild-type FixL and mutant derivative pAL401Δ. Transmembrane segments are represented as wavy regions. Numbers indicate the amino acid positions from the N-terminal end.

rules (10) also predict a β-turn occurrence between TS I and TS II.

Deletion analysis of FixL. To determine the functional requirements of the membrane-spanning segments in oxygen regulation, we constructed a deletion that removes most of the N-terminal end of FixL (Fig. 5). The deletion (pAL401Δ) was then tested for the activation of P*nifA* in vegetative cells of the *fixLJ* deletion strain *R. meliloti* T219 in response to oxygen deprivation. A similar analysis was carried out on bacteroids prepared from alfalfa nodules infected with T219. As shown in Table 1, plasmid pAL401Δ produces a protein with less than 15% of the wild-type activity when cells are induced microaerobically and with no activity above background in nodules. The residual β-galactosidase activity produced with pAL401Δ in vegetative cells increases approximately 10-fold when the cells are incubated under microaerobic conditions compared with aerobic conditions. These data suggest that the truncated protein produced by this plasmid is able to sense oxygen in vivo, although with reduced efficiency.

Figure 2 shows a Western blot analysis of crude lysates of *R. meliloti* T219 carrying the different plasmid-borne derivatives of FixL. Each mutant produces a protein of the approximate size expected. As shown, a lower amount of FixL per cell is produced by plasmid pAL401Δ compared with the amount of protein produced with plasmid pGD311, which expresses wild-type FixL. However, the protein level expressed by pAL401Δ is higher than the level of FixL produced by wild-type *R. meliloti* 102F34, which cannot be detected under these conditions. The reduction in overall *nifA* expression seen for pAL401Δ is therefore unlikely to be due to less FixL being present in the cell. The membrane-spanning region of FixL may be necessary for optimal activity in vivo.

Sequence similarity studies of the N-terminal region of FixL. The C-terminal 270 amino acids of FixL starting near residue 230 show significant sequence identity with other two-component sensor proteins (13). The heme-binding region of FixL is located in the amino-terminal region, between the hydrophobic segments and the conserved C-terminal half of the protein (37). This N-terminal region of the protein is not expected to be related to other two-component sensors since it contains the regulator-specific sensing domain. We nevertheless carried out a computer search with the N-terminal segment between amino acids 86 to 220 to determine whether conserved sequences that might represent a common structural feature within the sensing domain could be identified. Surprisingly, a significant level of identity was found with the N-terminal end of KinA, the sensor protein that is involved in regulating endospore formation in *Bacillus subtilis*. A conserved region of about 70 amino acids has 25% identity and scored well above random in the jumbling test described by Doolittle (17) (Fig. 6). The same region also showed similarity to the N-terminal region of *Bradyrizobium parasponiae* NtrB, the sensor that regulates nitrogen metabolism in this bacterium.

Two other proteins, both of which are oxidases, also have sequences with similarity to the sequence of the sensing domain of FixL (Fig. 6). Isopenicillin synthase (IPS) is a nonheme iron-binding oxidase that catalyzes the cyclization of δ-(L-α-aminoadipoyl)-L-cysteinyl-D-valine to isopenicillin (11). The region of this oxidase related to FixL is well conserved among all of the bacterial and fungal oxidase proteins and may be involved in coordinating metal binding (11). Another oxygenase, P-450 IID1 from humans, a B-type hemoprotein that functions in the oxidation of hydrocarbon substrates (21), also shows significant identity with FixL. The conserved region in P-450 has no known structural or functional properties, however, and it should be noted that this was the only P-450 identified in the sequence search.

TABLE 3. Calculation of β-turn probability in the membrane-anchoring domain

Tripeptide	Position	P_t[a]	Chou and Fasman[b]	Garnier et al.[b]
EPIG	TS I-TS II	0.82	+	−
GRNP	TS II-TS III	3.5	+	+
DGPS	TS III-TS IV	4.95	+	+

[a] Calculated from Chou and Fasman (10) by multiplying the probability (p_t) of each amino acid in the tripeptide occurring at the β turn. $P_t > 1$ is considered to represent a probable bend.

[b] The program COMFOR was used to calculate the probability of bend occurrence according to Chou and Fasman (10) or Garnier et al. (19). +, probable bend occurrence.

DISCUSSION

The results presented in this study show that FixL consists of 505 amino acids rather than 464 as initially proposed (13). This interpretation is based on the results of site-specific mutagenesis of the two possible starting codons and from a Tn*phoA* insertional analysis.

A hydropathy plot carried out with full-length FixL shows that there are four possible membrane-spanning segments in the N-terminal region. If these hydrophobic domains are used by FixL to span the cytoplasmic membrane, the only likely topological arrangement based on structural considerations would position the C-terminal end facing the cytoplasmic side for the following reasons. Regulatory proteins belonging to the sensor class of the two-component system family contain a region of conserved sequences usually found in the C-terminal end (39). In this region of homology, most sensor proteins contain a conserved histidine residue which is thought to be the site of phosphorylation (40). In FixL, this histidine is located at amino acid 285. Since this C-terminal domain of FixL must physically interact with the cytoplasmic regulator FixJ in order to transduce the oxygen signal, its positioning in the cytoplasm is expected. A Tn*phoA* insertional fusion at position 386 had no significant alkaline phosphatase activity, consistent with a cytoplasmic localization for the C terminus of FixL. The topological model presented in Fig. 4 proposes that FixL is likely to contain two small regions in the N terminus that are exposed to the periplasmic side of the membrane. The Tn*phoA* insertional analysis is consistent with this model, since *fixL-phoA* fusions with a high alkaline phosphatase activity

Name	S.D.		Location
FixL		RARDAHLRSILDTVPDATVVSATDGTIVSFNAAAVRQFGYAEEEVIGQNLRILMPEPYRHEHDGYLQR	200
KinA	8.1	MEQDTQHVKPLQTKTDIHAVLASNGRIIYISANSKLHLGYLQGEMIGSFLKTFLHEEDQFLVESYFYN	68
p450	6.2	KEESGFLREVLNAVPVLLHIPALAGKVLRFQKAFLTQL DELLTEHRMTWDPAQPPRDLTEAFLAE	278
NtrB	5.9	RPVPTDGEAILNALPNPVLLVAPDGRIVDANIAAESFFEISTQFLRRQSLKELVPFGSPLLALIEQVR	76
I.S.	4.2	ASVVLIRYPYLDPYPEAAIKTAADGTKLSFEWHEDVSLITVLYQSNVQNLQVETAAGYQDIEADDTGY	250

FIG. 6. Alignment of sequences having sequence identities with FixL. The number of standard deviations (S.D.) from the mean was calculated according to Doolittle (17). Homologies with standard deviations of >3 are considered significant. Identities to FixL are indicated (*). The location represents the last C-terminal amino acid of the polypeptide fragment shown.

were isolated only in these two regions. The majority of these insertions were found between amino acids 83 and 89, the second and larger periplasmic loop (Fig. 4).

At least two features of membrane proteins serve as determinants of topological structure (for a review, see reference 6). A major determinant is the hydrophobic stretches that can span the cytoplasmic membrane. A second important determinant is the charge distribution on either side of the transmembrane regions, which is important in determining the orientation of polypeptides in the membrane. Charges that are localized immediately adjacent to the hydrophobic segment are especially crucial (5) and can serve as stop transfer regions to control the insertion of these proteins into the membrane. The positive inside rule described by von Heijne (44) suggests that a net positive charge is more predominant in the region of the protein facing the cytoplasmic side compared with the periplasmic side. As shown in Fig. 4, the net charge of residues near the transmembrane segments in the cytoplasmic side of FixL is either positive or neutral. The periplasmic regions on the other hand are overall negatively charged or neutral. On the basis of the observations discussed above, we propose that FixL is a polytopic, integral membrane protein containing four membrane-spanning segments arranged as shown in Fig. 4.

FixL is thought to contain at least two domains: an oxygen-sensing domain containing heme located in the N-terminal region of the protein and a kinase domain located in the C-terminal region (37). Data presented here suggest that the hydrophobic, membrane-spanning region in the N-terminal-most segment of the protein can also be considered a separate domain and that this domain may not be necessary for regulation by oxygen. We performed a deletion analysis of this membrane-anchoring domain that removed all four putative hydrophobic membrane-spanning regions. The mutant protein was tested in a *Rhizobium* strain lacking endogenous FixL and containing a plasmid-borne P*nifA-lacZ* fusion in *cis*. Deleting all four transmembrane segments produced a protein that was localized mostly in the soluble fraction (27a) and which lost about 83% of its P*nifA*-activating activity when tested under microaerobic conditions. In nodules, this protein appeared completely inactive. The residual *nifA* activity observed in vegetative cells as a result of the expression from plasmid pAL401Δ was oxygen regulated. The residual oxygen regulation observed for pAL401Δ is consistent with the fact that we have already shown that a truncated soluble version of FixL starting at amino acid 127 carries heme and responds to oxygen by modifying both kinase and phosphatase activities (28). These results suggest that the membrane-anchoring domain may not be absolutely necessary for oxygen sensing in vivo or in vitro but that it may affect overall activity of the protein in vivo.

A search of gene banks with the protein sequence of FixL revealed that in addition to extensive C-terminal region homology with other two-component sensors, the N-terminal domain of FixL (excluding the transmembrane regions) contains sequences with significant homology to sensor proteins of *B. subtilis* (KinA) and *Bradyrhizobium parasponiae* (NtrB). This is unusual, since no homology has previously been reported in the sensing domains of this protein family (39). This homologous region of the FixL protein contains the heme moiety (37). On the basis of the homology, it is tempting to speculate that these two proteins may sense similar signals. Although the signal mediated by KinA during sporulation is still unknown, it is conceivable that oxygen plays a role in the initiation of sporulation. It is also possible that FixL senses other, still unknown, signals in addition to oxygen.

Interestingly, the same region in FixL was found to have low but significant homology to two proteins involved in oxidation reactions: cytochrome P-450 IID1 from humans and IPS. The homology with P-450 raises the possibility that this portion of P-450 is involved in heme binding or stabilization. Unfortunately, the crystal structure of eukaryotic P-450 enzyme has not yet been determined as it has for P-450 CAM of *Pseudomonas* spp. P-450 CAM differs in several aspects from the eukaryotic P-450 family and has no significant homology with FixL. In general, heme-binding regions of proteins are poorly conserved and have as their primary feature a number of nonpolar amino acids that serve to create a hydrophobic heme pocket. On the other hand, IPS does not contain heme but does bind exogenously added iron. It is conceivable that the homologous region of the FixL protein is important for providing the right environment for accommodating the iron in heme. It is difficult at the present time to reconcile the KinA-NtrB homology with that of P-450 and IPS. Both KinA and NtrB should be examined for the presence of iron and iron-binding capability. An alternative possibility is that this region of FixL and its homologs is important for interactions with other proteins with some similarity in structural features.

ACKNOWLEDGMENTS

We are indebted to R. Doolittle for helpful discussions and suggestions. We thank C. Manoil and J. Beckwith for providing λ Tn*phoA*, Sharon Stanfield for reading the manuscript, R. Roberts for technical assistance, and M. Levitt for preparation of the manuscript.

A. F. Lois was supported by a UCSD Chancellor Fellowship.

This work was supported by NIH grant RO1 GM44400 and NSF grant DMB-8716673.

REFERENCES

1. **Appleby, C. A.** 1984. Leghemoglobin and *Rhizobium* respiration. Annu. Rev. Plant Physiol. **35**:443–478.
2. **Baldwin, T. O., T. Berends, T. A. Bunch, T. F. Holzman, S. K. Rauch, L. Schmansky, M. L. Treat, and M. M. Ziegler.** 1984. Cloning of luciferase structural gene from *Vibrio harveyi* and expression of bioluminescence in *Escherichia coli*. Biochemistry **23**:663–3667.
3. **Batut, J., M.-L. Daveran-Mingot, M. David, J. Jacobs, A. M. Garnerone, and D. Kahn.** 1989. *fix*K, a gene homologous with *fnr* and *crp* from *Escherichia coli*, regulates nitrogen fixation genes both positively and negatively in *Rhizobium meliloti*. EMBO J. **8**:1279–1286.
4. **Better, M., G. Ditta, and D. R. Helinski.** 1985. Deletion analysis of *Rhizobium meliloti* symbiotic promoters. EMBO J. **4**:2419–2424.
5. **Boyd, D., and J. Beckwith.** 1989. Positively charged amino acid residues can act as topogenic determinants in membrane proteins. Proc. Natl. Acad. Sci. USA **86**:9446–9450.
6. **Boyd, D., and J. Beckwith.** 1990. The role of charged amino acids in the localization of secreted and membrane proteins. Cell **62**:1031–1033.
7. **Brickman, E., and J. Beckwith.** 1975. Analysis of the regulation of *Escherichia coli* alkaline phosphatase synthesis using deletions and ϕ80 transducing phages. J. Mol. Biol. **96**:307–316.
8. **Burbulys, D., T. A. Trach, and J. A. Hoch.** 1991. Initiation of sporulation in *B. subtilis* is controlled by a multicomponent phosphorelay. Cell **64**:545–552.
9. **Chou, P. Y., and G. D. Fasman.** 1974. Prediction of protein conformation. Biochemistry **13**:222–245.
10. **Chou, P. Y., and G. D. Fasman.** 1978. Empirical predictions of protein conformation. Annu. Rev. Biochem. **47**:251–276.
11. **Cohen, G., D. Shiffman, M. Mevarech, and Y. Aharonowitz.** 1990. Microbial isopenicillin N synthase genes: structure, function, diversity and evolution. Trends Biotechnol. **8**:105–111.
12. **Corbin, D., G. Ditta, and D. R. Helinski.** 1982. Clustering of nitrogen fixation (*nif*) genes in *Rhizobium meliloti*. J. Bacteriol. **149**:221–228.
13. **David, M., M. Daveran, J. Batut, A. Dedieu, O. Domergue, J. Ghai, C. Hertig, P. Boistard, and D. Kahn.** 1988. Cascade regulation of *nif* gene expression in *Rhizobium meliloti*. Cell **54**:671–683.
14. **Ditta, G.** 1986. Tn5 mapping of *Rhizobium* nitrogen fixation genes. Methods Enzymol. **118**:519–528.
15. **Ditta, G., T. Schmidhauser, E. Yakobson, P. Lu, X.-W. Liang, D. R. Finlay, D. Guiney, and D. R. Helinski.** 1985. Plasmids related to the broad host range vector, pRK290, useful for gene cloning and for monitoring gene expression. Plasmid **13**:149–153.
16. **Ditta, G., E. Virts, A. Palomares, and C. H. Kim.** 1987. The *nifA* gene of *Rhizobium meliloti* is oxygen regulated. J. Bacteriol. **169**:3217–3223.
17. **Doolittle, R. F.** 1987. Of Urfs and Orfs: a primer on how to analyze derived amino acid sequences. University Science Books, Mill Valley, Calif.
18. **Forst, S., D. Comeau, S. Norioka, and M. Inouye.** 1987. Localization and membrane topology of EnvZ, a protein involved in osmoregulation of OmpF and OmpC in *Escherichia coli*. J. Biol. Chem. **262**:16433–16438.
19. **Garnier, J., D. J. Osguthorpe, and B. Robson.** 1978. Analysis of the accuracy and implications of simple methods for predicting the secondary structure of globular proteins. J. Mol. Biol. **120**:97–120.
20. **Gilles-Gonzalez, M. A., G. S. Ditta, and D. R. Helinski.** 1991. A hemoprotein with kinase activity encoded by the oxygen sensor of *Rhizobium meliloti*. Nature (London) **350**:170–172.
21. **Gonzalez, F. J., F. Vilbois, J. P. Hardwick, O. W. McBride, D. W. Nebert, H. V. Gelboin, and U. A. Meyer.** 1988. Human debrisoquine 4-hydroxylase (P450IId1): cDNA and deduced amino acid sequence and assignment of the CYP2D locus to chromosome 22. Genomics **2**:174–179.
22. **Gross, R., B. Arico, and R. Rappuoli.** 1989. Families of bacterial signal-transducing proteins. Mol. Microbiol. **3**:1661–667.
23. **Gutierrez, C., J. Barondess, C. Manoil, and J. Beckwith.** 1987. The use of transposon Tn*phoA* to detect genes for cell envelope proteins subject to a common regulatory stimulus. J. Mol. Biol. **195**:289–297.
24. **Kahn, M., D. Ow, B. Sauer, A. Rabinowitz, and R. Calendar.** 1980. Genetic analysis of bacteriophage P4 using P4-plasmid ColE1 hybrids. Mol. Gen. Genet. **177**:399–412.
25. **Kozak, M.** 1983. Comparison of initiation of protein synthesis in procaryotes, eucaryotes, and organelles. Microbiol. Rev. **47**:1–45.
26. **Kyte, J., and R. F. Doolittle.** 1982. A simple method for displaying the hydropathy character of a protein. J. Mol. Biol. **157**:105–132.
27. **Laemmli, U. K.** 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature (London) **227**:680–685.

27a. **Lois, A.** Unpublished data.

28. **Lois, A. L., M. Weinstein, G. S. Ditta, and D. R. Helinski.** J. Biol. Chem., in press.
29. **Long, S., S. McCune, and G. C. Walker.** 1988. Symbiotic loci of *Rhizobium meliloti* identified by random Tn*phoA* mutagenesis. J. Bacteriol. **170**:4257–4265.
30. **Makino, K. H., M. Shinagawa, M. Amemura, and A. Nataka.** 1986. Nucleotide sequence of the phoR gene, a regulatory gene of the phosphate regulon of Escherichia coli. J. Mol. Biol. **192**:549–556.
31. **Manoil, C., and J. Beckwith.** 1985. Tn*phoA*: a transposon probe for protein export signals. Proc. Natl. Acad. Sci. USA **82**:8129–8133.
32. **Manoil, C., and J. Beckwith.** 1986. A genetic approach to analysing membrane protein topology. Science **233**:1403–1408.
33. **Manoil, C., D. Boyd, and J. Beckwith.** 1988. Molecular genetic analysis of membrane protein topology. Trends. Genet. **4**:223–226.
34. **Mehdi, H., E. Ono, and K. C. Gupta.** 1990. Initiation of translation at CUG, GUG, and ACG codons in mammalian cells. Gene **91**:173–178.
35. **Melchers, L. S., J. G. Regensburg-Tuink, R. B. Bourret, J. A. Sedee, R. A. Schilperoort, and P. J. J. Hooykaas.** 1989. Membrane topology and functional analysis of the sensory protein VirA of *Agrobacterium tumefaciens*. EMBO J. **8**:1919–1925.
36. **Miller, J. H.** 1972. Experiments in molecular genetics. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
37. **Monson, E. K., G. S. Ditta, M. Weinstein, and D. R. Helinski.** 1992. The FixL protein of *Rhizobium meliloti* can be separated into a heme-binding oxygen-sensing domain and a functional C-terminal kinase domain. Proc. Natl. Acad. Sci. USA **89**:4280–4284.
38. **Mutoh, N., and M. I. Simon.** 1986. Nucleotide sequence corresponding to five chemotaxis genes in *Escherichia coli*. J. Bacteriol. **165**:161–166.
39. **Nixon, B. T., C. W. Ronson, and F. M. Ausubel.** 1986. Two component regulatory systems responsive to environmental stimuli share strongly conserved domains with the nitrogen assimilation regulatory genes *ntr*B and *ntr*C. Proc. Natl. Acad. Sci. USA **83**:7850–7854.
40. **Stock, J. B., A. J. Ninfa, and A. M. Stock.** 1989. Protein phosphorylation and regulation of adaptive responses in bacteria. Microbiol. Rev. **53**:450–490.
41. **Szeto, W. W., J. L. Zimmerman, V. Sundaresan, and F. M. Ausubel.** 1984. A *Rhizobium meliloti* symbiotic regulatory gene. Cell **36**:1035–1043.
42. **Vincent, J. M.** 1970. A manual for the practical study of the root nodule bacteria. International Biological Programs handbook no. 15. Blackwell Scientific Publications Ltd., Oxford.
43. **Virts, E. L., S. W. Stanfield, D. R. Helinski, and G. S. Ditta.** 1988. Common regulatory elements control symbiotic and microaerobic induction of nifA in *Rhizobium meliloti*. Proc. Natl. Acad. Sci. USA **85**:3062–3065.
44. **von Heijne, G.** 1986. The distribution of positively charged residues in bacterial inner membrane proteins correlates with the trans-membrane topology. EMBO J. **5**:3021–3027.